doi:10.2533/chimia.2025.600 Chimia 79 (2025) 600–605

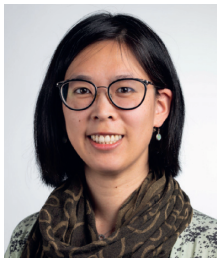
Bridging Innovation and Efficiency: The Promises and Challenges of Self-Driving Labs as Sustainable Drivers for Chemistry

Florian A. Formica[§,a], Edlyn Wu[§,a], Lucien Brey[§,a], Daniel Pacheco Gutiérrez[§,a], Riccardo Barbano[§,a], Hermann Tribukait[a], José Miguel Hernández-Lobato[b], Paco Laveille[c], and Loïc M. Roch[*,a]

Abstract: Self-driving laboratories (SDLs) are reshaping scientific discovery by combining robotics, artificial intelligence (AI), and data science to automate the full Design-Make-Test-Analyze (DMTA) cycle. This review highlights how SDLs address the inefficiencies of traditional trial-and-error methods through intelligent, autonomous experimentation. We explore key advances in AI, automation, and data infrastructure, as well as the remaining technical challenges. Applications across organic synthesis, materials science, and biotechnology (*e.g.* such as catalytic reaction optimization, solid-state synthesis, and protein engineering) demonstrate their transformative potential. A recurring theme is the role of SDLs in promoting sustainability by miniaturizing reactions and maximizing sample efficiency through AI and machine learning. Finally, we discuss the requirements for broader adoption, including robust hardware, interoperable software, and high-quality datasets, positioning SDLs as essential tools for next-generation sustainable research.

Keywords: Artificial intelligence · DMTA cycles · Self-driving labs · Sustainability

Dr. Florian A. Formica is an Application Engineer at Atinary Technologies, where he bridges science, product, and engineering to align the client needs with the evolution of the company's AI-driven platform. He is a versatile biochemist who grew up in Switzerland and holds an MSc in Biological Chemistry from EPFL and a PhD in Biomedical Sciences from ETH Zürich. As a scientist and musician, he strives to combine creativity, precision, and innovation to advance research, empower collaboration, and inspire growth.



Dr. Edlyn Wu is in Business Development at Atinary Technologies, where she works with scientists and R&D innovators to identify opportunities for accelerating research through AI-driven solutions. She earned her PhD in Biochemistry from McGill University, followed by postdoctoral research at the Max Planck Institute and Université de Lausanne. With extensive bench experience, Edlyn understands the challenges researchers face and helps them navigate complex experimental spaces, connecting their expertise with cutting-edge digital tools for transformative applications across the life sciences and beyond.

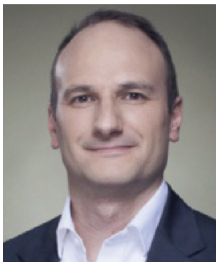
Lucien Brey is a ML Researcher at Atinary Technologies. He received a MSc in Molecular and Biological Chemistry as well as a BSc in Chemistry and Chemical Engineering from EPFL. He wrote his master thesis at MIT with Prof. Rafa Gomez-Bombarelli.

Daniel Pacheco Gutierrez is the lead of the AI & Data Engineering team at Atinary Technologies, responsible for driving the production ML systems at the SDLabs platform. He was born in Colombia, received a MSc degree in Computational Science and Engineering at EPFL and a BSc degree in Chemical Engineering from McGill University in Canada. Daniel is passionate about designing, building, and testing ML solutions applied to process optimization in the domains of chemistry and biology.

Dr. Riccardo Barbano leads the AI Research & Innovation team at Atinary Technologies, focusing on algorithmic advancement. He holds a PhD in Probabilistic Machine Learning applied to Inverse Problem in Imaging from University College London. He worked with the Medical Imaging Computing Unit, where he was also part of the Alan Turing Institute. During this time, he undertook research

*Correspondence: Dr. L. M. Roch, E-mail: loic@atinary.com
§Authors contributed equally to this work. [a]Atinary Technologies, Route de la Corniche 4, CH-1066 Epalinges, Switzerland; [b]Department of Engineering, University of Cambridge, Cambridge, United Kingdom; [c]ETH Zurich, Swiss Cat+ East, Vladimir-Prelog-Weg 1-5, CH-8093 Zurich, Switzerland.

visits to the Chinese University of Hong Kong and the Korea Advanced Institute of Science and Technology (KAIST). Before his PhD, he earned an MRes in Imaging at UCL, an MPhil in Machine Learning and Machine Intelligence at the University of Cambridge, and an MEng in Biomedical Engineering at Imperial College London.



Dr. Hermann Tribukait is an economist, entrepreneur, and leader in innovation and technology development. He is a Co-Founder and CEO of Atinary Technologies based in Silicon Valley, with over 20 years of experience in banking, consulting, and technology development. Prior to launching Atinary, Dr. Tribukait co-authored and led the Clean Energy Materials Innovation Challenge of the global initiative Mission Innovation for 3 years. He holds a BA in Economics from ITAM with highest honors, and a Masters, and PhD in Economics from Harvard. Dr. Tribukait is fluent in English, German, and Spanish.

Prof. Dr. José Miguel Hernández-Lobato is a Professor of Machine Learning at the University of Cambridge, where he leads research in probabilistic modeling, scalable Bayesian inference, and Bayesian optimization. He earned his PhD in Computer Science from Universidad Autónoma de Madrid and completed postdoctoral research at Harvard and Cambridge. His work advances data-efficient AI for scientific discovery, including applications in chemistry and drug development. Prof. Hernández-Lobato's research integrates uncertainty quantification, generative models, and automated decision-making to accelerate innovation in science and engineering.

Dr. Paco Laveille is a materials chemist specializing in catalysis and self-driving laboratories. He holds a master's in Biochemistry from Université de La Rochelle and a PhD in Materials Chemistry from Université de Montpellier. After a postdoc at IFPEN, he joined KAUST's Catalysis Center, leading high-throughput experimentation in Prof. Basset's group. He later worked at ADNOC, driving catalytic technology research from lab to commercial scale. Since 2021, he has served as Managing Director of SwissCAT+ at ETH Zurich, leading a national platform that integrates automation, data digitalization, and machine learning to accelerate catalyst discovery for sustainable fuel and chemical production.

Dr. Loïc M. Roch is the co-founder and CTO of Atinary Technologies, where he leads a multi-disciplinary team advancing machine learning for efficient experiment planning. He holds a BSc/MSc in Chemistry and Chemical Engineering from EPFL, a PhD in Quantum Chemistry from UZH and Tianjin University and completed two postdocs in AI and ML at Harvard and at the Vector Institute. Passionate about accelerating R&D, he focuses on scaling self-driving labs. As an entrepreneur, scientist, and nature-lover, Loïc envisions a world where science and technology drive sustainability and the transition to a circular economy.

1. Introduction

The chemical industry is undergoing a major shift, driven by the need for more sustainable processes and safer products. Stricter regulations are accelerating this transition, pushing companies to phase out fossil feedstocks and hazardous substances. At the heart of this change is R&D, which must evolve to support faster, smarter discovery.

Traditional chemical R&D relies heavily on slow, trial-and-error design-make-test-analyze (DMTA) cycles, often evaluating one factor at a time (OFAT). These workflows are time- and labor-intensive, typically evaluating only a handful of compounds per week. Yet chemical systems often involve thousands or millions of possible combinations across parameters like catalysts, solvents, temperature, or pressure. Given this complexity, it is no surprise that developing new chemical processes can take years – and often relies on luck.

In the context of urgent needs for sustainable chemical processes, accelerating the pace of discovery has become essential. In response, automated closed-loop experimentation, frequently named 'Self-driving labs' (SDLs), which integrate robotics, some form of artificial intelligence (AI) and data science, has emerged as a promising solution. These platforms automate not only the execution of experiments, but also the design and analysis steps, forming iterative, rational feedback loops with minimal human intervention.[1–3] Today, despite remaining challenges in robotics and application of AI in chemistry, there are several examples of R&D loops performing multiple design-make-test loops fully autonomously, or with minimal human intervention.

By enabling faster and more rational experimentation, SDLs accelerate the optimization and discovery of new molecules, materials and processes that support the development of greener technologies (Fig. 1). Before presenting some achievements of SDLs in organic chemistry, materials science, and biotechnology, we present a short overview highlighting the current development and challenges of laboratory automation and AI in this field. For a more complete understanding of the topic and broader examples, the reader can refer to recently published, more extensive reviews.[3–5]

2. Development and Challenges of Self-driving Laboratories

2.1 *Lab Automation: Status Quo*

The automation of laboratory processes has evolved steadily over the past century, beginning with the introduction of electrically-powered analytical instruments in the early 20th century.[6] In the 1960s and 70s, the integration of computers enabled automated data acquisition and motorized equipment has accelerated the adoption of more automated analytical tools such as mass spectrometry and high-performance liquid chromatography (HPLC),[7] now cornerstones of modern analytical chemistry. The 1980s and 90s witnessed the emergence of the first robotic arms and liquid handling robots and the widespread adoption of microplates.[8,9] Companies such as Hamilton, Tecan and Gilson became pioneers in this space, developing programmable robotic systems tailored for standardized microplate-based workflows, enabling parallelized assays. These early robotic platforms dramatically improved experimental throughput, particularly in combinatorial drug discovery and clinical diagnostics.[10,11]

In the following decades, increasingly complex and more advanced robotic platforms (*e.g.* ABB, Unchained labs, Chemspeed, Labman Automation) emerged, capable not only of dispensing liquids but also handling solids, performing heating, cooling, stirring, and even some forms of in-line monitoring. These latest developments in robotics form the physical backbone of the emerging SDLs. The combination of robotics, software, and AI enable high-throughput, reproducible, scalable, and digital experimenta-

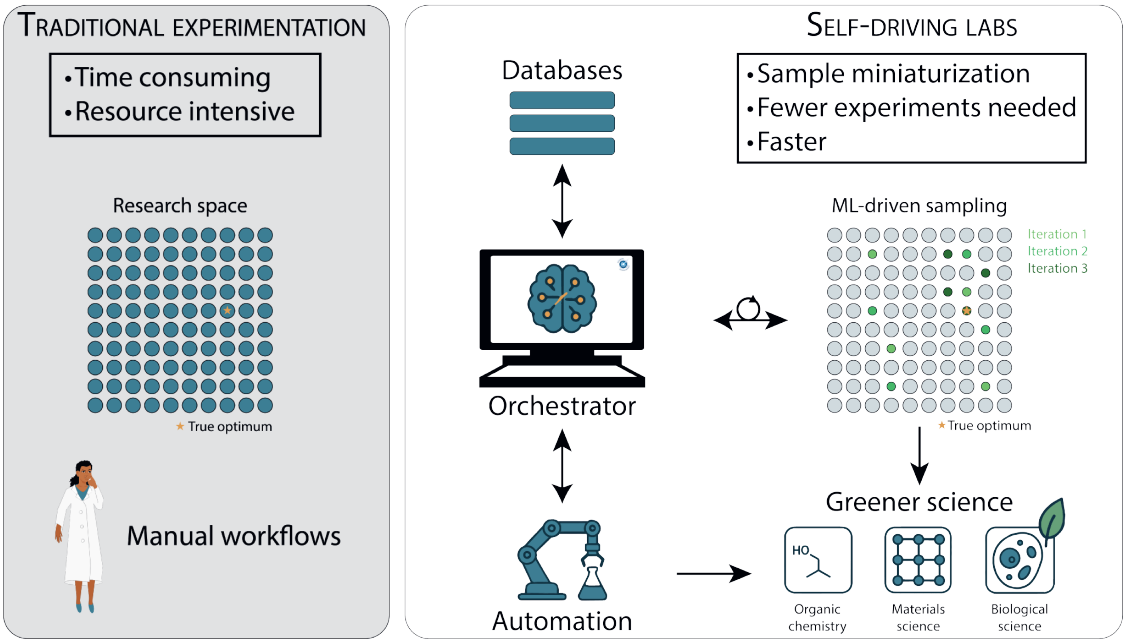


Fig. 1. Traditional experimentation typically involves a large number of experiments and relies on manual workflows, rendering the processes ineffective (left). Self-driving labs enable the simultaneous orchestration of several components such as automated equipment and external databases. In addition, the research space is sampled smartly, considerably reducing the number of experiments and quantities needed. This leads to more sustainable processes in fields such as organic chemistry, materials science and biological science (right).

tion. However, despite few historical examples of total automation of clinical and analytical labs,[12] most current systems remain limited to specific chemical workflow scopes and face a range of practical and technological challenges.[13,14]

On the data front, reaction outcomes must be digitized and autonomously analyzed for a true closed loop workflow. Techniques like UV-Vis, NMR, HPLC, and GC have already been integrated into automated workflows, enabling automated or in-line analysis. However, more complex analytical methods, in particular solid-state characterization, such as XRD (X-ray diffraction), XRF (X-ray fluorescence), TGA (thermogravimetric analysis) and ICP (inductively coupled plasma spectroscopy), remain difficult to automate due to complex sample preparation and/or manual (un) loading requirements.[15] In those examples where automation was solved, analytical throughput often becomes the next challenge or bottleneck. For example, HPLC or GC analysis of hundreds of samples may take longer than the synthesis itself, leaving robotic platforms idle and unusable for other activities. As a result, analytics are frequently decoupled from synthesis, with sample transfer performed manually.

Reliability also remains a major challenge. Software crashes, open-source code-bases that are not stable and poorly maintained, other software communication errors, and uncompleted tasks often interrupt workflows, disrupt the processes, and require human intervention and substantial time to troubleshoot. Other factors having a lack of reliability and precision are handling of small quantities (especially solids) or performing chemical reactions under pressure.

Finally, seamless integration across all software layers such as electronic lab notebooks (ELNs), machine learning (ML) models, equipment control and scheduling, as well as data management, analysis and visualization, is still rare. This results in fragmented digital ecosystems that frequently require manual workarounds to close the overall data strategy loops.

In summary, robotics and lab automation software have matured to the point where small-scale autonomous experimentation is now feasible, particularly for specific, limited experimental–analytical loops.[16] Yet the vision of a broad-purpose, fully autonomous self-driving R&D lab remains at an early stage. Benchtop wet-lab experimentation is still largely manual, relying on the expertise and craftsmanship of domain specialists. Human dexterity, however, limits precision at small scales and introduces variability and uncertainty in measurements, while requiring relatively large quantities of materials.[17] Modern lab machinery is addressing these limitations, with robots now capable of handling micro- to nano-scale volumes and executing thousands of experiments in parallel using only a fraction of the materials traditionally required.[18]

2.2 *AI in R&D*

Traditional experimental design methodologies such as OFAT or DoE are highly inefficient, scale poorly with increased dimensionality, and rely heavily on heuristics[19] – making them more of an art than a science. Leveraging increasingly powerful AI algorithms is key to unlocking new advances in chemistry. Generative AI (GenAI), Bayesian Optimization (BO), and Active Learning (AL) are distinct paradigms that serve specific purposes and can be highly valuable for SDLs.

In GenAI, the underlying training data distribution is learned in order to sample new candidates from the same distribution. For instance, candidate generation can be conditioned on contextual information or on desired chemical properties. In chemistry, GenAI has demonstrated many successes[20] – using variational autoencoders[19] for small molecule design,[21] diffusion models[23,24] for molecular docking and protein structures[25,26] or flow matching models,[27] incorporating equivariance to better capture physical and chemical properties for molecular generation.[26] Once a molecule of interest is identified, BO becomes the method of choice for optimizing the reaction conditions. This probabilistic approach to black-box optimization is particularly well-suited for expensive or time-consuming processes where building a physical model from first principles is not feasible.[29,30]

BO relies on two key components: a probabilistic surrogate model that approximates the objective function, and a utility function – often referred to as the acquisition function – that guides the selection of new candidates. The surrogate model not only predicts the objectives of interest but, crucially, also provides

uncertainty estimates associated with those predictions. The acquisition function then balances the trade-off between exploiting regions with high predicted performance and exploring areas with high uncertainty. This methodology enables efficient navigation of large, complex search spaces under data-scarce conditions, requiring only a minimal number of experimental runs. In SDLs, BO has been successfully applied to optimize synthesis parameters and discover new materials or molecules with improved target properties.[31] In cases where abundant but lower-quality data (*e.g.* from simulations) is available, multi-fidelity approaches can be leveraged to combine data sources with varying levels of accuracy. Multi-Fidelity Bayesian Optimization (MFBO) frameworks have shown significant promise in materials and molecular research,[32–34] helping to reduce the environmental footprint associated with extensive high-fidelity experimentation.

AL, also referred to as Bayesian experimental design, shares BO's iterative, model-driven loop but pursues a different goal: reducing model uncertainty rather than directly optimizing an objective. At each step, a query strategy selects the most informative candidates from the pool (or continuous design space) for labelling *via* experiment or simulation. These data points tighten the surrogate's predictive uncertainty, enabling more accurate predictions with fewer experiments. AL thus excels at building compact, high-quality datasets, accelerating tasks such as property prediction, reaction-outcome classification, and pre-training models for downstream BO or GenAI workflows.[35,36]

3. Application of Self-Driving Labs in Research and Industry

3.1 *Applications in Organic Chemistry*

SDLs are transforming (green) organic chemistry by enabling closed-loop optimization across a range of sustainable reaction modalities, most notably in flow and solid-state chemistry.[37,38] Flow systems have become increasingly popular in SDLs thanks to reaction miniaturization and precise control over temperature, mixing, and residence time, which together improve efficiency and reduce waste.[2,3,39] These systems support high throughput experimentation with minimal reagent consumption – as exemplified by a microfluidic Suzuki–Miyaura platform that delivered over 1,500 LC/MS datapoints per day using just ~50 μg of substrate per reaction.[40] Flow chemistry further enables environmentally-friendly strategies like photochemistry and flash chemistry, where highly selective and fast transformations benefit from superior heat and mass transfer and in-line sensing.[2,39,41–43] Platforms such as RoboChem exemplify this potential, integrating a microreactor and real-time NMR to autonomously optimize diverse photoredox and metallophotocatalysis reactions in less than half a milliliter.[44] Flow SDLs have also scaled to execute entire synthetic routes, from active pharmaceutical ingredients like Sonidegib[45] to complex molecules such as carpanone.[46]

Alternatively, solid-state chemistry offers a solvent-free route to sustainable synthesis, and SDLs have begun to integrate modular robotic systems to automate these workflows. One example involved a multi-robot setup that autonomously handled benzimidazole crystal growth, grinding, and XRD analysis in a closed-loop setting, achieving high-quality measurements and a throughput of 168 samples per week.[47] Together, these advances highlight how SDLs can drive green chemistry forward through automation, miniaturization, and resource-efficient experimentation.

3.2 *Application in Materials Science*

Another landscape being revolutionized by SDLs is materials science, specifically by automating the discovery and optimization of new functional materials such as thin films, battery materials, or catalysis.

In 2020, MacLeod *et al.* introduced *Ada*, a modular self-driving platform for the autonomous optimization of thin-film materials, with a focus on organic hole transport layers commonly used in perovskite solar cells.[48] The *Ada* platform combines robotic synthesis, processing, and characterization with Bayesian optimization (Phoenics), coordinated by ChemOS,[49] to autonomously explore complex experimental spaces. As a proof of concept, *Ada* optimized the pseudomobility of organic small-molecule semiconductor (spiro-OMeTAD) films by adjusting dopant concentration and annealing time. It autonomously executed all steps – from solution preparation to four-point probe measurements – using pseudomobility as a surrogate model for device performance. Two 35-experiment campaigns (each under 30 hours) converged on similar optima, validating the workflow's robustness. *Ada* also uncovered non-obvious conditions linked to improved stability, demonstrating how targeted, autonomous platforms can accelerate materials discovery.

In 2023, Szymanski *et al*. demonstrated a fully autonomous platform for the solid-state synthesis of novel inorganic materials, known as the *A-Lab*.[50] This platform combines robotics, *ab initio* thermodynamics, and ML to enable autonomous, closed-loop synthesis at the multigram scale. Target structures were selected from high-throughput predictions (Materials Project, DeepMind), with initial recipes generated by ML models trained on text-mined literature. The *ARROWS* algorithm iteratively optimized precursors and conditions using feedback from XRD measurements. Robotic arms handled all operations, from dispensing to analysis. Over 17 days, the platform attempted 58 oxide and phosphate syntheses, successfully producing 41 (71%). This work showed how AI-driven planning and robotic execution can accelerate material discovery, though application-specific performance testing remained a key next step.

In 2024, the *SwissCAT+* East hub at ETH Zurich published a practical implementation of a closed-loop, AI-guided workflow for the discovery and optimization of heterogeneous catalysts for CO_2 hydrogenation to methanol.[51,52] This study used automated high-throughput experimentation and ML-based Bayesian optimization (Atinary™ SDLabs) to explore a catalyst space of over 20 million formulations. Across six closed-loop cycles (144 catalysts), up to four-metal combinations were synthesized on a Chemspeed robot and tested in a 16-channel fixed-bed reactor (Avantium). All steps were automated, aside from manual sample transfer between stations. A Bayesian model guided catalyst selection to optimize methanol selectivity and CO_2 conversion, while minimizing methane and cost — framing it as a multi-objective, multi-constraint problem. Each loop took around five days, with the full campaign completed in six weeks, showcasing the power of closed-loop experimentation to accelerate catalyst discovery.

In 2025, Zaki *et al*. introduced *MINERVA*, a fully integrated self-driving platform for the autonomous synthesis, purification, and characterization of a wide range of nanomaterials.[53] The system features a robotic arm that automates sample transfer between several interconnected stations, including a liquid handler, a heated stirring plate, a centrifuge, and a sonicator, along with multiple spectroscopic characterization tools such as dynamic light scattering, zeta potential, UV-Vis absorbance, fluorescence, and chemiluminescence. Controlled by an orchestration system with real-time data analysis, the platform autonomously executes closed-loop synthesis–characterization cycles. It successfully reproduced diverse synthesis protocols, including gold nanoparticles, silica, metal-organic frameworks (MOFs), CuO, and core–shell nanostructures, with high reproducibility. Its modular, open design enables easy expansion to new methods. However, it should be noted that the system lacks ML-driven variable optimization, limiting its ability to explore synthesis conditions efficiently.

3.3 *Application in Biological Sciences and Biotechnology*

SDLs are central to the biological landscape by integrating robotics and AI into closed-loop experimental design. The complexity of biomolecules makes them ideal candidates to fully exploit this power. A leading example is the *SAMPLE* platform (Self-driving Autonomous Machines for Protein Landscape Exploration), which autonomously engineers proteins by learning sequence–function relationships, designing new candidates, and experimentally validating them through a fully automated robotic pipeline, from gene synthesis to cell-free protein expression to biochemical characterization.[54,55] This approach was demonstrated for glycoside hydrolase enzymes, where four SAMPLE agents independently converged on the most thermally stable variants despite noisy experimental data, highlighting both the robustness and adaptability of autonomous exploration in protein engineering and synthetic biology.[54]

Increasing accessibility and affordability of computational resources accelerates a shift from manual, linear workflows to intelligent systems capable of continuous optimization and automated operations. In synthetic biology, SDLs can exploit the central role of the genome in regulating biological function, making them ideal for high-impact tasks such as pathway engineering or directed evolution.[56] However, translating this potential into practice also requires overcoming significant technical and integration challenges.

One major hurdle is the complexity of multi-instrument workflows in biotechnology R&D. Fushimi *et al.* recently demonstrated a modular autonomous lab for medium optimization in *E. coli*, integrating hardware such as a transfer robot, liquid handler, incubator, LC-MS/MS, with closed-loop Bayesian optimization to suggest experimental conditions for optimizing both cell growth and metabolite production. The system autonomously handled culturing, measurement, and analysis, successfully scaling and identifying growth media from 1 to 100 mL. In contrast, optimization for metabolite production remained limited, highlighting the constraints of the AI algorithm in addressing multi-objective optimization within the system. This modular, reconfigurable, and cost-efficient setup improves reproducibility and operational efficiency while reducing laborious manual efforts.[57]

Another emerging application of SDLs in biology is iterative modeling in systems biology, where complex, heterogeneous data pose major challenges, thus opening opportunities for automation and closed-loop workflows. Coutant *et al.* demonstrated this by studying the diauxic shift in yeast *S. cerevisiae*, using a SDL to automate planning, execution, and model refinement. This enabled three full modeling cycles, significantly improving predictive accuracy, and uncovering new genetic interactions.[58]

In biopharmaceutical development, platforms like *NanoMAP* integrate AI-driven design with high-throughput screening to streamline the formulation of complex nanomedicines, such as lipid nanoparticles for mRNA delivery. By automating key experimental workflows, including miniaturized scale-up, characterization, and validation assays, the *NanoMAP* platform supports a more systematic and autonomous approach to developing and evaluating high-performing biologics and delivery systems.[59]

Overall, while biology-focused SDLs face similar obstacles, the convergence of maturing AI tools and scalable robotics is beginning to accelerate research across biotechnology, protein engineering, and drug delivery. These technological advances are transforming how experiments are conceived, executed, and interpreted, in an increasingly autonomous manner.[60]

4. Challenges and Future Outlook

SDLs present significant advantages yet faces dual challenges: technical and operational.

On the technical side, multidisciplinary synergy is essential, requiring the integration of robotics, software engineering, machine learning, control systems, and more. This involves smooth hardware–software alignment, where robotic planners must integrate seamlessly with the algorithmic intelligence driving experimental design through easy-to-use interfaces. A recurring theme throughout is that sustainability is achieved not just through automation, but by miniaturizing reactions and being extremely sample-efficient: a capability facilitated by AI/ML.

On the operational side, while automation increases system autonomy, human expertise remains critical. Scientists are not removed from the loop as they are repositioned as supervisors and strategic guides. Here we argue that experimental design should not rely solely on collected data but should also incorporate scientists' decision-making and domain intuition. This interaction poses challenges. In addition, data must be efficiently preprocessed, organized, and stored in cloud infrastructure, and retrieved quickly. The system must capture all relevant data, without overflowing with irrelevant information, including often-neglected failure cases, and adhere to FAIR (Findable, Accessible, Interoperable, Reusable) principles for long-term utility. The high-dimensional nature of problem spaces demands either more computational resources or well-chosen approximation schemes to maintain computational tractability within algorithmic design.

Unlocking the full potential of SDLs demands sustained collaboration across scientists, engineers, and policy makers. Building SDLs is not just about algorithms and robots, but also about people, culture, change management, and policies.

Acknowledgements

The author acknowledges the full team at Atinary Technologies for the insightful discussions that contributed to the development of this manuscript.

Author Contributions

F. F., E. W., L. B., R. B. and D. P. contributed equally to the writing of the manuscript. In addition, L. M. R., H. T., J. M. H. L. and P. L actively contributed to the writing and reviewing. All authors have given approval to the final version of the manuscript. F. F., E. W., L. B., R. B., D. P., H. T. and L. M. R. are employees of Atinary Technologies; J. M. H. L. and P. L. are scientific advisors of Atinary Technologies. The authors declare no competing financial interest.

Received: August 12, 2025

[1] M. Christensen, L. P. E. Yunker, F. Adedeji, F. Häse, L. M. Roch, T. Gensch, G. Dos Passos Gomes, T. Zepel, M. S. Sigman, A. Aspuru-Guzik, J. E. Hein, *Commun. Chem.* **2021**, *4*, 112, https://doi.org/10.1038/s42004-021-00550-x.
[2] N. Mukhin, P. Jha, M. Abolhasani, *Matter* **2025**, *8*, 102205, https://doi.org/10.1016/j.matt.2025.102205.
[3] G. Tom, S. P. Schmid, S. G. Baird, Y. Cao, K. Darvish, H. Hao, S. Lo, S. Pablo-García, E. M. Rajaonson, M. Skreta, N. Yoshikawa, S. Corapi, G. D. Akkoc, F. Strieth-Kalthoff, M. Seifrid, A. Aspuru-Guzik, *Chem. Rev.* **2024**, *124*, 9633, https://doi.org/10.1021/acs.chemrev.4c00055.
[4] F. Häse, L. M. Roch, A. Aspuru-Guzik, *Trends Chem.* **2019**, *1*, 282, https://doi.org/10.1016/j.trechm.2019.02.007.
[5] D. Pacheco Gutierrez, L. M. Folkmann, H. Tribukait, L. M. Roch, *CHIMIA* **2023**, *77*, 7, https://doi.org/10.2533/chimia.2023.7.
[6] K. Olsen, *J. Lab. Autom.* **2012**, *17*, 469, https://doi.org/10.1177/2211068212455631.
[7] Eds., C. W. Gehrke, R. L. Wixom, E. Bayer, 'Chromatography-A Century of Discovery 1900-2000.The Bridge to The Sciences/Technology', Elsevier Science, Oxford, England, **2001**.
[8] J. Boyd, *Science* **2002**, *295*, 517, https://doi.org/10.1126/science.295.5554.517.
[9] S. Borman, *Chem. Eng. News Archive* **1998**, *76*, 47, https://doi.org/10.1021/cen-v076n014.p047.
[10] D. A. Armbruster, D. R. Overcash, J. Reyes, *Clin. Biochem. Rev.* **2014**, *35*, 143.
[11] J. P. Devlin, 'High Throughput Screening: The Discovery of Bioactive Substances', CRC Press, **1997**.

[12] M. Sasaki, T. Kageoka, K. Ogura, H. Kataoka, T. Ueta, S. Sugihara, *Clin. Chim. Acta* **1998**, *278*, 217, https://doi.org/10.1016/s0009-8981(98)00148-x.
[13] M. Christensen, L. P. E. Yunker, P. Shiri, T. Zepel, P. L. Prieto, S. Grunert, F. Bork, J. E. Hein, *Chem. Sci*. **2021**, *12*, 15473, https://doi.org/10.1039/d1sc04588a.
[14] M. Seifrid, R. Pollice, A. Aguilar-Granda, Z. Morgan Chan, K. Hotta, C. T. Ser, J. Vestfrid, T. C. Wu, A. Aspuru-Guzik, *Acc. Chem. Res*. **2022**, *55*, 2454, https://doi.org/10.1021/acs.accounts.2c00220.
[15] J. A. Selekman, J. Qiu, K. Tran, J. Stevens, V. Rosso, E. Simmons, Y. Xiao, J. Janey, *Annu. Rev. Chem. Biomol. Eng*. **2017**, *8*, 525, https://doi.org/10.1146/annurev-chembioeng-060816-101411.
[16] S. Lo, S. G. Baird, J. Schrier, B. Blaiszik, N. Carson, I. Foster, A. Aguilar-Granda, S. V. Kalinin, B. Maruyama, M. Politi, H. Tran, T. D. Sparks, A. Aspuru-Guzik, *Digital Discov*ery **2024**, *3*, 842, https://doi.org/10.1039/d3dd00223c.
[17] I. Kuselman, F. Pennecchi, W. Bich, D. Brynn Hibbert, *Accredit. Qual. Assur*. **2016**, *21*, 421, https://doi.org/10.1007/s00769-016-1239-3.
[18] S. A. Biyani, Y. W. Moriuchi, D. H. Thompson, *Chem. Methods* **2021**, *1*, 323, https://doi.org/10.1002/cmtd.202100023.
[19] J. Antony, 'Design of experiments for engineers and scientists', Elsevier - Health Sciences Division, Philadelphia, PA, **2023**.
[20] P. Tiwary, L. Herron, R. John, S. Lee, D. Sanwal, R. Wang, 'Generative artificial intelligence for computational chemistry: a roadmap to predicting emergent phenomena', **2024**.
[21] D. P. Kingma, M. Welling, 'Auto-Encoding Variational Bayes', **2013**, https://doi.org/10.48550/arXiv.2409.03118.
[22] R. Gómez-Bombarelli, J. N. Wei, D. Duvenaud, J. M. Hernández-Lobato, B. Sánchez-Lengeling, D. Sheberla, J. Aguilera-Iparraguirre, T. D. Hirzel, R. P. Adams, A. Aspuru-Guzik, *ACS Cent. Sci*. **2018**, *4*, 268, https://doi.org/10.1021/acscentsci.7b00572.
[23] J. Ho, A. Jain, P. Abbeel, 'Denoising Diffusion Probabilistic Models', **2020**, https://doi.org/10.48550/arXiv.1312.6114.
[24] R. E. Turner, C.-D. Diaconu, S. Markou, A. Shysheya, A. Y. K. Foong, B. Mlodozeniec, 'Denoising Diffusion Probabilistic Models in six simple steps', **2024**, https://doi.org/10.48550/arXiv.2402.04384.
[25] G. Corso, H. Stärk, B. Jing, R. Barzilay, T. Jaakkola, 'DiffDock: Diffusion steps, twists, and turns for molecular docking', **2022**, https://doi.org/10.48550/arXiv.2210.01776.
[26] J. Yim, H. Stärk, G. Corso, B. Jing, R. Barzilay, T. S. Jaakkola, *Wiley Interdiscip. Rev. Comput. Mol. Sci.* **2024**, *14*, https://doi.org/10.1002/wcms.1711.
[27] Y. Lipman, R. T. Q. Chen, H. Ben-Hamu, M. Nickel, M. Le, 'Flow Matching for generative modeling', **2022**, https://doi.org/10.48550/arXiv.2210.02747.
[28] Y. Song, J. Gong, M. Xu, Z. Cao, Y. Lan, S. Ermon, H. Zhou, W.-Y. Ma, 'Equivariant flow matching with hybrid probability transport', **2023**, https://doi.org/10.48550/arXiv.2312.07168.
[29] J. Snoek, H. Larochelle, R. P. Adams, 'Practical Bayesian optimization of machine learning algorithms', **2012**, https://doi.org/10.48550/arXiv.1206.2944.
[30] B. Shahriari, K. Swersky, Z. Wang, R. P. Adams, N. de Freitas, *Proc. IEEE* **2016**, *104*, 148, https://doi.org/10.1109/jproc.2015.2494218.
[31] O. Schilter, D. P. Gutierrez, L. M. Folkmann, A. Castrogiovanni, A. García-Durán, F. Zipoli, L. M. Roch, T. Laino, *Chem. Sci*. **2024**, *15*, 7732, https://doi.org/ 10.1039/d3sc05607d.
[32] M. Kennedy, *Biometrika* **2000**, *87*, 1, https://doi.org/10.1093/biomet/87.1.1.
[33] D. Huang, T. T. Allen, W. I. Notz, R. A. Miller, *Struct. Multidiscip. Optim*. **2006**, *32*, 369, https://doi.org/10.1007/s00158-005-0587-0.
[34] V. Sabanza-Gil, R. Barbano, D. Pacheco Gutiérrez, J. S. Luterbacher, J. M. Hernández-Lobato, P. Schwaller, L. Roch, *Nat. Comput. Sci.* **2025**, *5*, 572, https://doi.org/10.1038/s43588-025-00822-9.
[35] T. Rainforth, A. Foster, D. R. Ivanova, F. Bickford Smith, *Stat. Sci*. **2024**, *39*, 100, https://doi.org/10.1214/23-sts915.
[36] M. Hedman, D. R. Ivanova, C. Guan, T. Rainforth, 'Step-DAD: Semi-amortized policy-based Bayesian experimental design', **2025**, https://doi.org/10.48550/arXiv.2507.14057.
[37] D. Dallinger, C. O. Kappe, *Curr. Opin. Green Sustainable Chem.* **2017**, *7*, 6, https://doi.org/10.1016/j.cogsc.2017.06.003.
[38] F. Fanelli, G. Parisi, L. Degennaro, R. Luisi, *Beilstein J. Org. Chem*. **2017**, *13*, 520, https://doi.org/10.3762/bjoc.13.51.
[39] U. K. Sharma, *CHIMIA* **2025**, *79*, 424, https://doi.org/10.2533/chimia.2025.424.
[40] D. Perera, J. W. Tucker, S. Brahmbhatt, C. J. Helal, A. Chong, W. Farrell, P. Richardson, N. W. Sach, *Science* **2018**, *359*, 429, https://doi.org/10.1126/science.aap9112.
[41] J.-I. Yoshida, A. Nagaki, T. Yamada, *Chemistry* **2008**, *14*, 7450, https://doi.org/10.1002/chem.200800582.
[42] L. M. Baumgartner, J. M. Dennis, N. A. White, S. L. Buchwald, K. F. Jensen, *Org. Process Res. Dev*. **2019**, *23*, 1594, https://doi.org/10.1021/acs.oprd.9b00236.
[43] A. D. Clayton, A. M. Schweidtmann, G. Clemens, J. A. Manson, C. J. Taylor, C. G. Niño, T. W. Chamberlain, N. Kapur, A. J. Blacker, A. A. Lapkin, R. A. Bourne, *Chem. Eng. J.* **2020**, *384*, 123340, https://doi.org/ 10.1016/j.cej.2019.123340.
[44] A. Slattery, Z. Wen, P. Tenblad, J. Sanjosé-Orduna, D. Pintossi, T. den Hartog, T. Noël, *Science* **2024**, *383*, eadj1817, https://doi.org/10.1126/science.adj1817.
[45] A. M. K. Nambiar, C. P. Breen, T. Hart, T. Kulesza, T. F. Jamison, K. F. Jensen, *ACS Cent. Sci*. **2022**, *8*, 825, https://doi.org/10.1021/acscentsci.2c00207.
[46] D. Cortés-Borda, E. Wimmer, B. Gouilleux, E. Barré, N. Oger, L. Goulamaly, L. Peault, B. Charrier, C. Truchet, P. Giraudeau, M. Rodriguez-Zubiri, E. Le Grognec, F.-X. Felpin, *J. Org. Chem*. **2018**, *83*, 14286, https://doi.org/10.1021/acs.joc.8b01821.
[47] A. M. Lunt, H. Fakhruldeen, G. Pizzuto, L. Longley, A. White, N. Rankin, R. Clowes, B. Alston, L. Gigli, G. M. Day, A. I. Cooper, S. Y. Chong, *Chem. Sci*. **2024**, *15*, 2456, https://doi.org/10.1039/d3sc06206f.
[48] B. P. MacLeod, F. G. L. Parlane, T. D. Morrissey, F. Häse, L. M. Roch, K. E. Dettelbach, R. Moreira, L. P. E. Yunker, M. B. Rooney, J. R. Deeth, V. Lai, G. J. Ng, H. Situ, R. H. Zhang, M. S. Elliott, T. H. Haley, D. J. Dvorak, A. Aspuru-Guzik, J. E. Hein, C. P. Berlinguette, *Sci. Adv*. **2020**, *6*, eaaz8867, https://doi.org/10.1126/sciadv.aaz8867.
[49] L. M. Roch, F. Häse, C. Kreisbeck, T. Tamayo-Mendoza, L. P. E. Yunker, J. E. Hein, A. Aspuru-Guzik, *Sci. Robot*. **2018**, *3*, eaat5559, https://doi.org/10.1126/scirobotics.aat5559.
[50] N. J. Szymanski, B. Rendy, Y. Fei, R. E. Kumar, T. He, D. Milsted, M. J. McDermott, M. Gallant, E. D. Cubuk, A. Merchant, H. Kim, A. Jain, C. J. Bartel, K. Persson, Y. Zeng, G. Ceder, *Nature* **2023**, *624*, 86, https://doi.org/10.1038/s41586-023-06734-w.
[51] A. Ramirez, E. Lam, D. P. Gutierrez, Y. Hou, H. Tribukait, L. M. Roch, C. Copéret, P. Laveille, *Chem. Catal*. **2024**, *4*, 100888, https://doi.org/10.1016/j.checat.2023.100888.
[52] P. Laveille, P. Miéville, S. Chatterjee, E. Clerc, J.-C. Cousty, F. De Nanteuil, E. Lam, E. Mariano, A. Ramirez, U. Randrianarisoa, K. Villat, C. Copéret, N. Cramer, *CHIMIA* **2023**, *77*, 154, https://doi.org/10.2533/chimia.2023.154.
[53] M. Zaki, C. Prinz, B. Ruehle, *ACS Nano* **2025**, *19*, 9029, https://doi.org/10.1021/acsnano.4c17504.
[54] J. T. Rapp, B. J. Bremer, P. A. Romero, *Nat. Chem. Eng*. **2024**, *1*, 97, https://doi.org/10.1038/s44286-023-00002-4.
[55] B. Gopal, *J. Biosci*. **2024**, *49*, https://doi.org/10.1007/s12038-024-00443-w.
[56] H. G. Martin, T. Radivojevic, J. Zucker, K. Bouchard, J. Sustarich, S. Peisert, D. Arnold, N. Hillson, G. Babnigg, J. M. Marti, C. J. Mungall, G. T. Beckham, L. Waldburger, J. Carothers, S. Sundaram, D. Agarwal, B. A. Simmons, T. Backman, D. Banerjee, D. Tanjore, L. Ramakrishnan, A. Singh, *Curr. Opin. Biotechnol*. **2023**, *79*, 102881, https://doi.org/10.1016/j.copbio.2022.102881.
[57] K. Fushimi, Y. Nakai, A. Nishi, R. Suzuki, M. Ikegami, R. Nimura, T. Tomono, R. Hidese, H. Yasueda, Y. Tagawa, T. Hasunuma, *Sci. Rep*. **2025**, *15*, 6648, https://doi.org/10.1038/s41598-025-89069-y.
[58] A. Coutant, K. Roper, D. Trejo-Banos, D. Bouthinon, M. Carpenter, J. Grzebyta, G. Santini, H. Soldano, M. Elati, J. Ramon, C. Rouveirol, L. N. Soldatova, R. D. King, *Proc. Natl. Acad. Sci. U. S. A*. **2019**, *116*, 18142, https://doi.org/10.1073/pnas.1900548116.
[59] R. J. Hickman, P. Bannigan, Z. Bao, A. Aspuru-Guzik, C. Allen, *Matter* **2023**, *6*, 1071, https://doi.org/10.1016/j.matt.2023.02.007.
[60] C. Cizauskas, E. DeBenedictis, P. Kelly, *New Biotechnol*. **2025**, *88*, 1, https://doi.org/10.1016/j.nbt.2025.03.004.