

Perspective

Autonomous chemistry: Navigating self-driving labs in chemical and material sciences

Oliver Bayley,^{1,2} Elia Savino,^{1,2} Aidan Slattery,^{1,2} and Timothy Noël^{1,*}

SUMMARY

Self-driving labs (SDLs) have recently emerged as one of the most significant technological developments in the chemical and materials sciences and hold the potential to revolutionize the research process. Herein, we discuss the structure of an SDL in terms of the hardware, the coordinator software, and the AI agent and examine how the selection of these elements affects the overall research capabilities. We further look to the application and accessibility of these platforms to better understand their future impact on synthetic and materials chemistry. With the rapid rise in the capabilities of these SDLs, and with the increasing democratization of the space, it is crucial to be aware of both the promises and pitfalls this technology holds.

INTRODUCTION

In the rapidly evolving landscape of scientific research, the integration of automation, particularly in the realm of self-driving labs (SDLs), represents a pivotal shift toward a more efficient, reproducible, and scalable scientific method. This approach leverages advanced algorithms (artificial intelligence [AI] planners), robotics (hardware executors), and control software (system coordinators) for the design, execution, and analysis of experiments in an automated fashion. It marks a significant departure from traditional laboratory practices, promising to accelerate the discovery pace across various fields of science.¹ The potential of SDLs presents the opportunity to enable significant cost-savings, increased safety and the ability to free up scientists' time to partake in more exploratory work.

Automated science is nothing new, with early attempts at computer-controlled processes having been successfully demonstrated in the mid-20th century with the automation of analytical instruments,² experimental apparatus,^{3,4} and due to the sequential and repetitive nature, peptide synthesis.^{5,6} While the automation in these cases focused on a single component of an experimentalist's workflow, these early endeavors laid the groundwork for the more sophisticated SDL systems we see today. The explosive growth of SDLs in the most recent decade can be attributed to several concurrent technological developments. Firstly, from the hardware front, there has been a significant decrease in the costs associated with robotics and high-throughput experimentation platforms.⁷ This decline in prices makes these technologies more accessible to a wider range of research institutions, especially academia, enabling broader adoption, and resulting in a marked increase in academic publications.^{8–10} With the improved technological developments, modern SDLs can now incorporate more complex robotics capable of performing intricate manipulations that were previously either too delicate or too complex.^{11,12} Concurrently, there has also been a remarkable advancement in the capabilities and accessibility of AI and machine learning (ML) workflows as well as the rise of large language models

PROGRESS AND POTENTIAL

Through the integration of AI, automated workflows, and robotics into research processes, self-driving labs (SDLs) have emerged as effective candidates to efficiently accelerate research timelines, increase data output and fidelity, reduce resource consumption, and ultimately liberate researchers from arduous, mundane, and repetitive tasks. These SDLs have already shown significant applications in reaction optimization, property optimization, process optimization, as well as drug discovery, but as the field evolves, more applications are expected to be realized. SDL decentralization, where researchers can connect to SDLs via the internet, and the cheapening of components through 3D printing and open-sourced design have both significantly increased the accessibility of SDLs to the wider scientific community. As the accessibility to SDLs capable of performing autonomous optimization and discovery tasks increases, their impact on the wider scientific community is only expected to increase.



(LLMs). As these technologies have become significantly more powerful and user-friendly, scientists are better able to integrate these sophisticated data analyses and experimental design capabilities into their laboratory operations.^{13,14}

SELF-DRIVING LABS

Hardware

With recent technological advancements already making automation of routine, repetitive, and time-consuming processes common place, the significant challenge facing the use of such hardware lies in the integration of these instruments into a cohesive automated workflow.¹⁵ Many factors contribute to this challenge, but the most significant are the physical transfer of material between stations and the centralized control of automated units that were designed to operate in a stand-alone fashion. To surpass these challenges, and ultimately develop a viable SDL platform, a significant re-conceptualization of the experiment workflow is often required. This has resulted in three archetypical platform designs: the fluidics platforms, the fixed-robotic-arm platforms, and the mobile platforms (Figure 1).¹ These platforms all cater to different needs but are by no means mutually exclusive, with many systems incorporating elements from multiple paradigms into the same platform.^{11,16} While these three strategies have already been observed, as this young field develops further, more paradigms may still emerge inspired by other existing technologies, for example deposition-based platforms resembling current 3D printing technologies.

In the fluidic platforms (Figure 1, purple), the individual experiment stations are linked together via chemically inert tubing, allowing for the sequential flow of material solutions through each unit. As fluidics are capable of material transport by nature, it easily overcomes one of the two central challenges mentioned previously and has therefore been widely adopted in the SDL platforms.^{11,12,17–19} In the simpler fluidics platforms, reagents are individually pumped through a mixing stage and subsequent reactor with an in-/online analytical module(s) (high-pressure liquid chromatography [HPLC],²² infrared [IR],²³ nuclear magnetic resonance [NMR],²⁴ mass spectrometry,²⁵ or ultraviolet [UV]²⁶) directly monitoring the output. The feedback from the analytical module (e.g., yield, e.e., liquid chromatography area percent (LCAP), etc.) can then be returned to the AI agent, which selects conditions for the next experiment in the campaign. While many successful optimizations have been demonstrated in these minimal layouts, the inflexibility of these platforms restricts the chemical space accessible to the SDL and limits the optimization to continuous reaction parameters such as temperature, residence time, or reagent equivalents. Consequently, more flexible systems, such as the RoboChem platform, have sought to introduce liquid handlers (gantry robots) or switch valves that allow the system to autonomously select different reagents, catalysts, additives, or reaction paths between each experiment in the campaign.^{19,27,28} Although this adds more complexity, it allows the systems to optimize reactions based on both categorical and continuous reaction parameters, which significantly increases the accessible chemical space. These platforms have been highly effective for reaction optimization and are relatively straightforward to implement in comparison to more elaborate robotic solutions. However, the fluidic platforms are not universally applicable as they can only be run with reagents where it is feasible to make stock solutions of sufficient concentration and under conditions in which the solutions are flowable/homogeneous.

For scenarios beyond the reach of the fluidic systems, robotic arms play a crucial role (Figure 1, red).^{29–31} In systems that inherently require physical motion such as the

¹University of Amsterdam (UvA), Van't Hoff Institute for Molecular Sciences (HIMS), Science Park 904, 1098 XH Amsterdam, the Netherlands

²These authors contributed equally

*Correspondence: t.noel@uva.nl

<https://doi.org/10.1016/j.matt.2024.06.003>

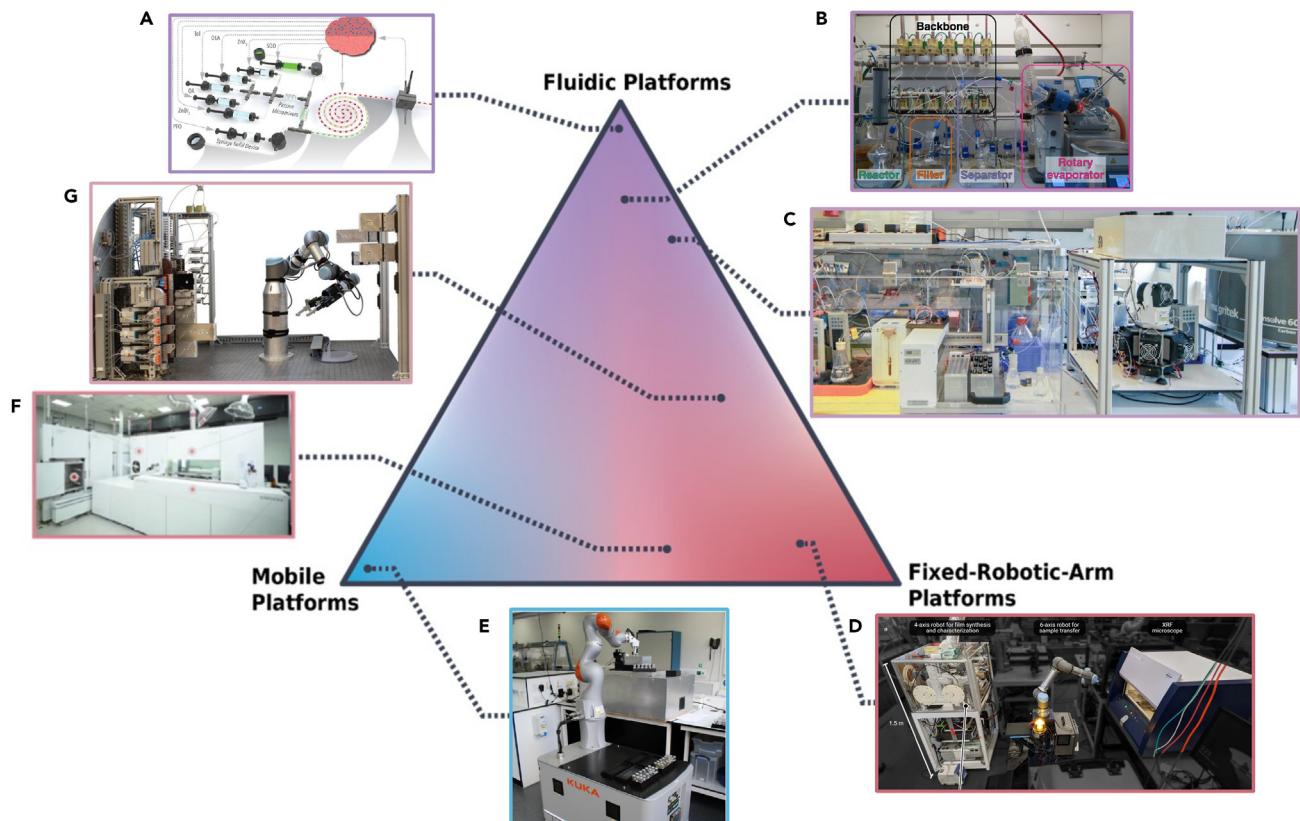


Figure 1. State-of-the-art SDL platforms that highlight the distinctions between the fluidic, the robotic arm, and the mobile robot platform architectures

- (A) The “Artificial Chemist” by Epps et al. (Abolhasani group) designed to generate on-demand quantum dots with a specific band gap.¹⁷ Reprinted with permission from John Wiley and Sons.
- (B) The “Chemputer” by Steiner et al. (Cronin group) designed to perform batch-like chemical reactions with all modules linked via chemically inert tubing.¹⁸ Reprinted with permission from AAAS.
- (C) The “RoboChem” by Slattery et al. (Noël group) designed to optimize single-step flow reactions.¹⁹ Reprinted with permission from AAAS.
- (D) The “Ada” system by MacLeod et al. (Berlinguette group) designed to perform material property optimizations²⁰ Reprinted under the creative commons license 4.0 (CC-BY 4.0).
- (E) The “Mobile Chemist” by Burger et al. (Cooper group) designed to perform chemical reactions and general processes in a human-like manner.²¹ Reprinted with permission from Springer Nature.
- (F) The “SynBot” by Ha et al. (Choi group) designed to perform multi-step chemical manipulations¹⁶ Reprinted under the creative commons license 4.0 (CC-BY 4.0).
- (G) The robotic flow chemistry platform by Coley et al. (Jensen and Jamison groups) designed to autonomously reconfigure the underlying flow platform to perform multi-step flow reactions.¹¹ Reprinted with permission from AAAS.

weighing of solids, measurement of solid-state properties, or relocation of reactor components, robotic arms are employed to perform the dexterous and otherwise unachievable manipulations. Although commonly used to achieve innately physical tasks, robotic arms have also been employed to operate standalone equipment in an automated manner without the need to alter the equipment itself, as was demonstrated by the MAOSIC system.³² This is a particularly significant use for robotic arms, as the integration of human-centered equipment is one of the biggest challenges in SDL hardware development. The robotic-arm and fluidic systems are not mutually exclusive either, with systems such as Coley et al.’s (Jensen and Jamison groups) robotic platform incorporating a central robotic arm capable of autonomously reconfiguring an underlying fluidics platform.^{11,12} While the robotic arms increase the flexibility of a given system, the physical reach of these arms also determines the

platform layouts and dictates its maximum size as well as the placement and accessibility of modules. To address this, larger systems, such as the Synbot by Ha et al. (Choi group) or Adam by King et al. (King group), have sought to integrate multiple robotic arms or robotic stages into a single platform, effectively linking platform modules by overlapping the reach of each robotic element.^{16,33}

Although a multitude of successful studies have been performed with fluidic or fixed-robotic-arm platforms, these layouts still rely on a highly re-conceptualized workflow with all modules either linked together (fluidics) or strategically placed within reach of a robotic system (fixed robotic arm). Aiming to preserve a more human-centered workflow, while still introducing the advantages of automation, a third approach to SDL platforms has been demonstrated in which mobile robotic platforms are equipped with a robotic arm allowing them to operate in a more standardized laboratory environment (Figure 1, blue). Focusing on this approach, the Mobile Chemist from Burger et al. (Cooper group) features a robotic system, supplied by KUKA Robotics, which was enhanced by the researchers to operate standard laboratory equipment within a more traditional lab workspace.²¹ Using this platform, Burger et al. had the robot perform reactions and analytical work in a workspace shared with other human chemists. Despite being built to mimic human-like researchers, and arguably having more difficulty in performing many of the manipulations, the mobile robotic chemist was able to operate for more than 21 h a day, which is notably longer than the workday of a (healthy) human researcher. While this approach maintains a much more human-centered workflow, many operational considerations must still be made, particularly to ensure the safety of the human researchers operating around the mobile robot.

Ultimately, several approaches to platform development have been demonstrated, but trade-offs must always be made in terms of cost, complexity, flexibility, and capability. Not every platform is suited to every task nor is every feature needed in every platform. For instance, the elegantly simple Artificial Chemist from Epps et al. uses a minimalist fluidics platform to achieve its goal without the need for a more complex and costly robotic arm.¹⁷ Conversely, the adhesive SDL from Rooney et al. is reliant on the robotic arm due to the required physical manipulation and would equally have no use for the Artificial Chemist's flow UV/Vis spectrometer.³¹ It is crucial to strike a careful balance between the capabilities and complexity of SDL platforms, particularly in regard to the end goal of the platform, as increased complexity can extend research timelines, introduce more potential sources of error, and generally increase development costs. When assessing increased complexity, it is important to be wary of the perceived "ease" in developing these platforms, as many unexpected problems can arise during assembly (e.g., 32- vs. 64-bit subsystem incompatibility).¹⁹ Many of these issues are not discussed in the literature. Understanding how to approach these challenges and balance complexity with capability is a crucial part of hardware development. For further reading, we recommend the review by Christensen et al. (Hein group).³⁴

Many platform designs, especially fluidics-based platforms, have now been realized. As the field develops, we expect new paradigms to emerge or new combinations that integrate current paradigms. Advanced AI algorithms in many SDLs typically focus on reducing the total number of experiments. In contrast, hardware automation has traditionally been used for high-throughput experimentation (HTE), facilitating exceedingly high reaction counts. Although these goals may seem contradictory, we believe new systems will likely emerge where HTE platforms are incorporated into SDLs. These systems could streamline HTE or, more excitingly, use HTE in tandem with an optimization SDL to perform "hit" testing followed by optimization.

This would enable a single unified platform to perform both hit discovery and optimization in an automated fashion.

Coordinator software

One of the core challenges in SDL development lies in facilitating information exchange between machines, AI, and human users in a universally understood manner. Achieving this within the context of SDLs often requires a coordinator software to provide a (graphical) user interface (UI) for input from the human scientist, a hardware interface to translate experiment parameters into executable tasks for the robotic platform, and an AI interface that translates the data from the platform to an ML-compatible format.¹⁵ Achieving communication between each of these components is far from trivial, as not only do they operate with different inputs and outputs, but on a fundamental level, they each seek to achieve different objectives.

Since the broader chemical and material science community cannot be presumed to have extensive knowledge in ML or software engineering, it is unreasonable to require researchers to encode their experiments for direct execution. Therefore, the coordinator software should have a human-centric interface, ideally a graphical user interface (GUI), allowing users to input their experiment parameters intuitively and straightforwardly. Building user interfaces that provide a simple user experience without losing platform control or creative freedom is not trivial. However, researchers now have access to numerous tools for developing such GUIs. Many of these tools have since found use in existing SDLs, including the LabVIEW visual toolkit³⁵ or Python frameworks, such as Tkinter¹¹ and Streamlit.¹⁷ While a variety of tools are available, developing an effective, appealing, and user-friendly UI is time-consuming and challenging for those without UI development experience. This is particularly true in academic settings, where the value of UI development is not immediately apparent or prioritized. Given the scale of work involved in creating an SDL, time-consuming UI design is not always a priority for developers, especially when the SDL can be operated directly through a command line. Nevertheless, it is imperative to invest the time into developing this crucial element to ensure the platform's use by a wider audience.

At their core, the AI agents are built upon complex mathematical algorithms and used to generate mathematical models that describe the relationship between given variables. As such, these agents are heavily reliant on the numerical nature of their inputs, making their application to chemical and materials science slightly more complex. While translating numerical reaction parameters such as temperature or reaction time into an AI-compatible format is straightforward, encapsulating molecular structures or specific reaction protocols poses a greater challenge. As a consequence, much research has been directed toward chemical featurization—the process of transforming chemical data into a machine-readable format—with the ultimate goal of using chemical structures as inputs for AI agents. Notations such as SMILES, SMIRKS, SMARTS, and Chemical Markup Language as well as data-based formats such as molecular fingerprints, graph-based models,³⁵ or DFT-based methods have now been shown to encode molecular structure information as well as reactivity information with substantial efficacy.³⁶ The development of these featurization methods has opened new avenues for SDLs, allowing the AI agent to be used beyond the optimization of already numerical inputs and allowing them to perform target optimizations using ligands, substrates, or other chemical information as an input. In certain cases, the coordinator software should also filter the raw data fed back to the AI agent to minimize faulty datapoints. Equally, the coordinator software plays a significant role in deciding when the ML training campaign is successful and

when to terminate the campaign based on a number of factors (maximum number of experiments, lack of improvement after a set number of runs, complete consumption of reagents, etc.).

Integrating the operations specified by users or AI agents with the hardware often represents the most formidable challenge in deploying an automated platform.³⁴ Due to the custom-built nature of SDLs, tailor-made software solutions are often required, which grants significant flexibility but often increases the development timelines and leads to platform-specific solutions. To help address the time cost and platform transferability and aid with proprietary hardware integration, a range of ready-built coordinator software packages have emerged including ChemOS, AresOS, or the Robot Operating System, which are designed to provide more versatile and generalized options for the orchestration of the SDL workflows.^{37,38} Additionally, the chemical description language XDL introduced by Mehr et al. (Cronin group) has also been used to convert written synthesis protocols from the literature into unit operations that can be directly executed by an SDL platform, helping to improve the generality of platform inputs.³⁹ While these dedicated software packages can offer many advantages, custom coordinator software is still preferred in many cases as the developers are more familiar with the inner workings of the system, ultimately allowing them to better tailor the software toward the SDL's needs.

The final core function of the orchestration software, the analytical data management and feedback, holds critical importance in the SDL workflow. As many SDL platforms integrate complex analytical equipment such as NMR, HPLC, and IR among others, advanced software to automate the analysis of the complex datasets is required. Many commercial software tools such as MestreNova for NMR spectroscopy or open-source tools such as MOCCA for HPLC analysis⁴⁰ have demonstrated great proficiency in the automated processing and reporting of the data generated by the platforms. Using such tools, the relevant spectral features from the generated data can be identified, normalized, deconvoluted, and quantified to return the necessary numerical data back to the AI agent. By incorporating these existing analysis tools, the coordinator software also plays a key role in translating the raw data output from the platform into viable inputs for the AI agent and thereby "close the loop" of the experiment cycle.

While hardware communication for both command and data retrieval purposes has been demonstrated, the inability to easily realize this with commercial hardware still poses a very significant barrier to lab automation. This is largely due to a lack of documented application programming interfaces (APIs), which allow such communication to occur. In some cases, open-source APIs have been shared by other developers. Occasionally, they are supplied free of charge by the manufacturer. However, in many cases, no immediately accessible API exists for the equipment. There are many reasons for this. One reason is the innate inability of the hardware to be controlled. Another reason is the lack of demand, resulting in no API being developed. In more inhibitive cases, the lack of accessible API is a deliberate decision by manufacturers to force developers to purchase the API separately. To ensure fair democratization of this field and to allow progress to continue, there must be an improvement in the accessibility of these hardware APIs. This improvement can come from the vendors themselves or through open-source solutions provided by other developers in the field.

Akin to hardware developments, great examples of coordinator software have been demonstrated, but there is still a significant amount of research space left to explore.

While ready-made software controllers like ChemOS have already been developed, we hope to see more modular “drop-in” front ends or front-end frameworks. These would not restrict developers’ freedom on the backend but would maintain a consistent and user-friendly interface for the end user. Such consistency would benefit users across different SDLs and save developers’ time, as they would no longer need to spend research time creating user-friendly interfaces. Although financial incentives for restricting API accessibility mean we are unlikely to see significant improvements in free manufacturer APIs, researchers are more focused on democratization. In academic settings, where budgets are limited, there will likely be an increase in user-created APIs for SDL equipment.

AI agent

Hardware automation has long been integral in the chemical realm, but the recent adoption of ML and AI into the decision-making process has been a truly innovative leap. Entrusting ML agents with the role of experimental planner not only streamlines the experimentation process by reducing the total number of experiments needed to map the chemical space, but it also generates valuable datasets for training subsequent AI agents. Although these AI agents are highly capable, care must still be taken to avoid over-fitting the model, where the agent overly conforms to the training data and fails to accurately predict new points—or under-fitting the model, where the agent does not learn enough to have useful predictive ability. It is also crucial to validate the predictions generated by these models to ensure they are not the results of artifacts from the training data. This also underscores the critical role human scientists still play in the initialization and validation of these ML methods.⁴¹ Many different machine learning methodologies have now been applied to the current SDL landscape and can be broadly categorized into optimization-based and discovery-based approaches. Optimization-based methods focus on identifying the best conditions for a predefined system, prioritizing exploitation over exploration. In contrast, discovery-based approaches seek to uncover new interactions, relationships, or chemistries, valuing exploration of new chemical space over the utilization of existing data. The choice of the AI agent for a specific SDL depends on many factors. However, a crude selection can be made based on the platform’s end goal and the available preceding data. For example, for optimization, choose an AI agent from the list of agents compatible with optimizations ([Figure 2](#), left). For a discovery-based SDL with no prior data, use a random sampling agent ([Figure 2](#), right). Although some preselection can be done, multiple AI agents often need to be assessed and their performance compared before identifying the most suitable agent for the SDL.

The area of optimization is the most explored in SDL research, and as such, many different AI agents have been shown to be effective. The classical one-factor-at-a-time (OFAT) approach, which is typically done during reaction optimization, requires an in-depth assessment of every optimizable parameter and is therefore slow, resource intensive, and costly by nature. As such, other methods that yield the desired outcomes with far fewer experiments, including SNOBFIT (stable noisy optimization by branch and fit), gaussian processes (GP), Bayesian optimization (BO), and genetic algorithms, have instead been selected for use in SDLs. While a detailed understanding of the advantages these approaches have over the classical OFAT approach is significant, this has already been reviewed in much detail elsewhere by Taylor et al.⁴² In these more efficient optimization AI agents (SNOBFIT, gradient descent GP, and BO), the initial data are generated through a low-data exploration method, such as random sampling, with the subsequent data then being used to feed the AI algorithm. The algorithms then use these data to map the chemical space

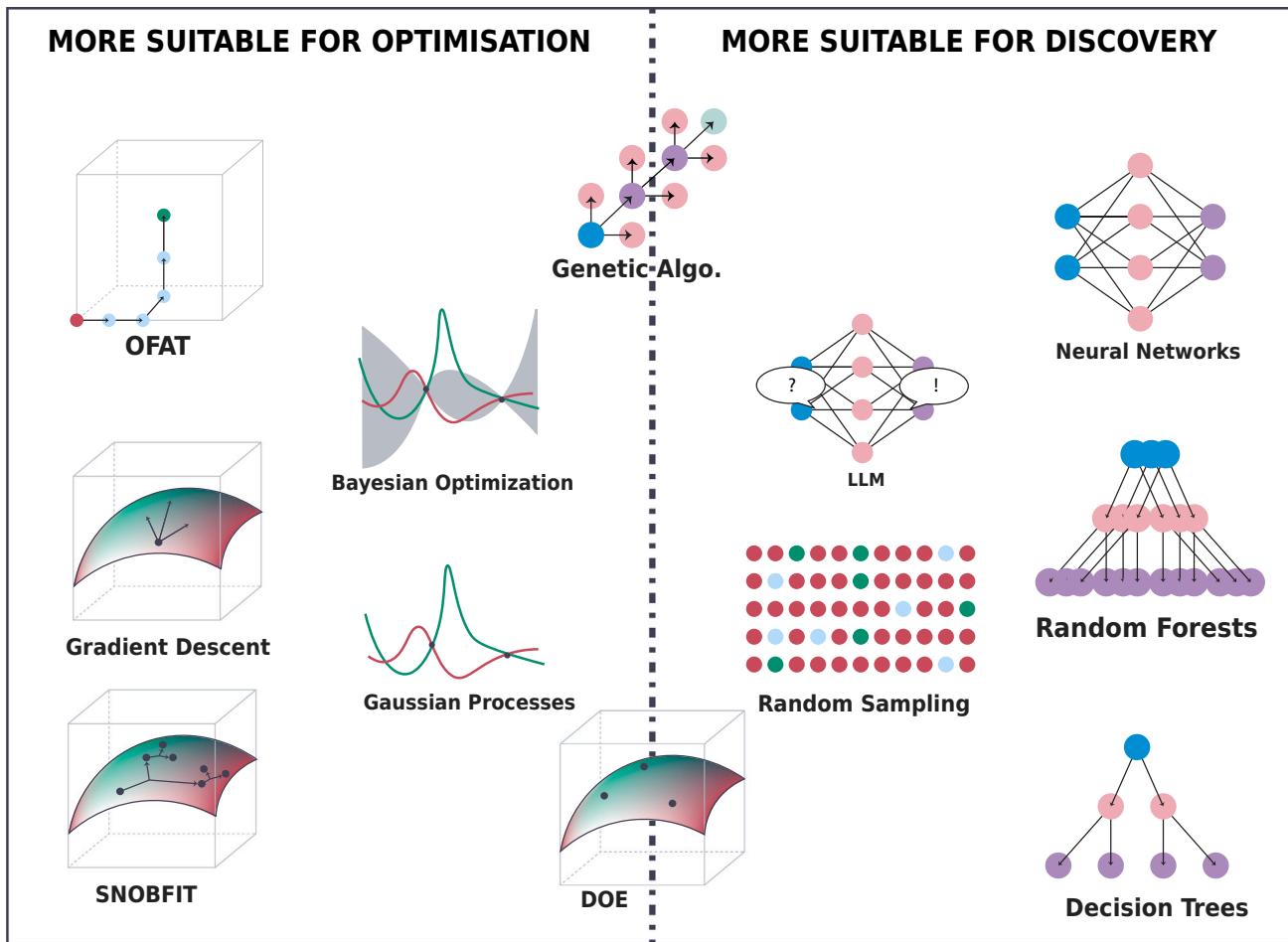


Figure 2. AI agents for SDLs toward optimization or discovery

Overview of machine learning frameworks and their suitability for optimization or discovery tasks.

and decide upon the next required data point based on what will give the algorithm a better understanding of the remaining unmapped space. Each algorithm differs in its approach toward deciding the next data point, each carrying its own strengths and weaknesses, but the AI agent that has quickly become a pseudo-gold-standard for this purpose is BO. BO is notably the most embraced method in reaction optimization as it is renowned for its ability to optimize black box functions, like those found in chemical systems, is able to perform optimizations after very few trials, and can be applied across diverse range of platforms. Its versatility is evidenced by successful implementations in areas such as photochemical transformation, multi-objective optimization, materials discovery, peptide chemistry, and chromatography. BO's widespread adoption underscores its effectiveness in navigating multi-dimensional optimization landscapes, making it a pivotal tool in advancing scientific and engineering challenges.^{17,29,31,43}

In recent years, there has been an increase in SDLs geared toward discovery, particularly with the goal of discovering new materials or chemical compounds. Due to the inherent complexity and unpredictable nature of such discovery, there has been less progress in this area than for optimization processes. Recent work in LLMs,⁴⁴ neural networks,⁴⁵ decision trees,⁴⁶ random sampling,⁴⁷ and random forests (Figure 2,

right), as well as developments outside of AI development in the field of density functional theory (DFT), has given insight into how the discovery process can be transformed.⁴⁸ These AI agents have excelled at discovery tasks due to their inherent ability to use and process large amounts of data. DFT has been particularly effective in providing accurate quantum mechanical descriptions and material properties, which greatly support these discovery tasks. Although the coupling of these techniques with SDLs is still in its infancy, there have been some compelling examples in recent years. Streith-Kalhoff et al. describe a delocalized asynchronous approach for the discovery of organic solid-state lasers using SDLs.⁴⁹ In addition to this, the A-lab developed by Szymanski et al. opens up the path to powder handling in an SDL for the discovery of novel materials.⁵⁰ The A-lab is a notable example of developments in discovery. The system integrates multiple AI agents for studying literature, analyzing data, and selecting experiments. Through its automated platform, the A-lab was able to generate a series of inorganic compounds that were not included in its training set. The characterization and compound novelty claims have faced criticism due to the lack of precision in the automated analysis. Ultimately, they serve as a cautionary tale, highlighting the crucial role of human verification. Furthermore, these technologies are not meant to be human substitutes but rather serve as collaborative robotic systems, where humans and robots work together in synergy.

Final considerations should be made about the availability of data and the overall fidelity of the available literature. Burke, Grzybowski, et al. recently showed that the current landscape of literature is largely insufficient for satisfactorily training ML models.⁵¹ Low reproducibility, human selection bias, lack of negative examples, and relatively small datasets for any given chemistry severely hinder the fidelity of literature data. Therefore, the main innovation and strength of incorporating ML in the experimentation process is the ability to generate its own unbiased data. However, ignoring the wealth of research available in the literature is short sighted. Multi-task models that account for the fidelity of data as an extra parameter can be used to overcome this issue.

Integration

SDLs are growing significantly in prevalence, largely due to advancements in automated hardware and AI. As these improvements are driven by external forces, the most impactful role researchers can play in the SDL field is the integration of different components. For example, the Mobile Chemist by Burger et al. was a highly challenging research feat, not in building the underlying robot (supplied by KUKA Robotics) but in integrating the robot into a chemical research environment and allowing it to seamlessly operate laboratory equipment. The challenging nature of integration should not be underestimated. It often requires new and creative thinking, as well as many rounds of trial and error. Any failure to integrate core elements can result in a significant loss of functionality. For instance, purchasing expensive high-end robotic equipment without the necessary software architecture for executing useful chemical experiments is problematic for the return on investment (ROI). It also means the SDL will not provide a significant advantage over existing human experimentation. Similarly, if the AI agent and hardware are well integrated but the end-user interface is non-existent, the SDL becomes unusable to anyone but the platform developer. This leads to underutilization and a poor return on research investment. Balancing the performance of SDLs is a task in itself, often requiring detailed analysis of various performance metrics, as reviewed by Volk et al.⁵²

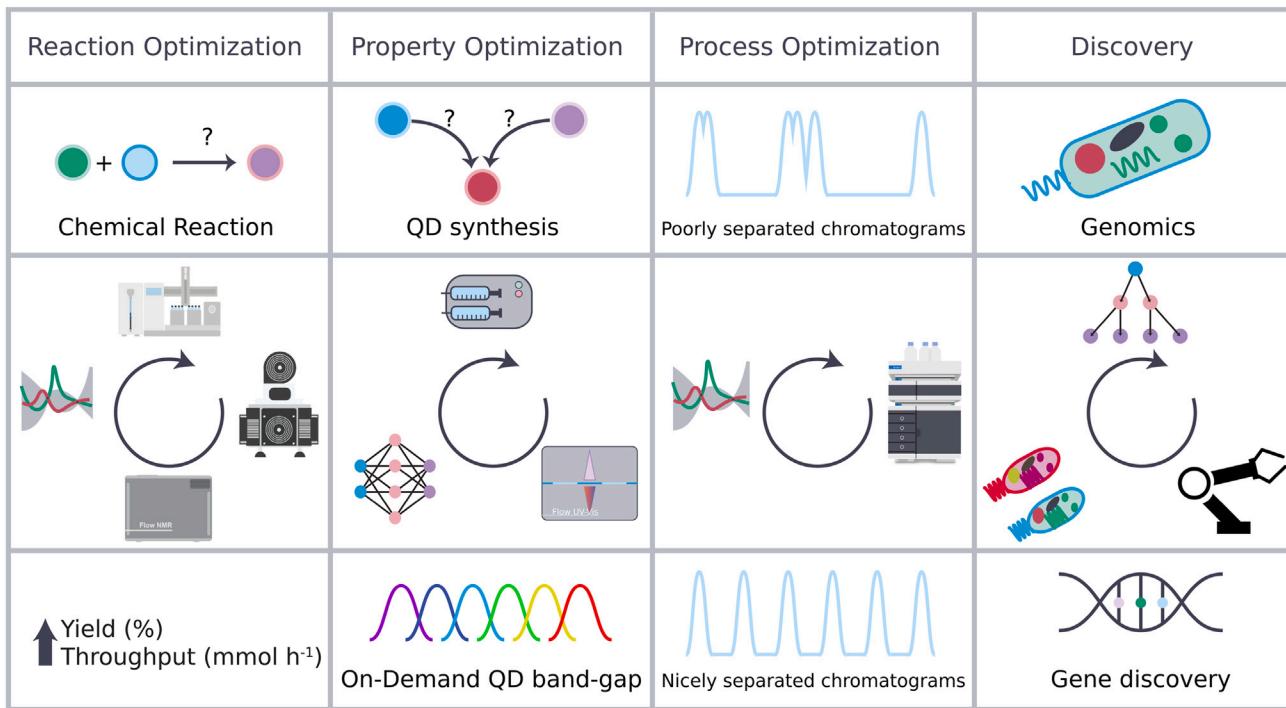


Figure 3. Demonstrated applications of SDLs across different domains, including reaction optimization, material property optimization, process optimization, and targeted discovery

Use cases

Although chemical and materials SDLs have only emerged in recent years, they have already seen significant use in reaction optimization,^{11,12,19} property optimization,²⁹ process optimization,^{53–55} and targeted discovery (Figure 3).^{33,46} Of all these applications, the most prominent use to date has been in reaction optimization for the purpose of identifying conditions that maximize yield or throughput. Although fixed-robotic-arm and mobile robot platforms have also been demonstrated, these SDLs have most commonly been based on fluidic platform layouts and rely on custom coordinator software to interface between the hardware, human user, and AI agent. These types of SDLs have become particularly popular due to both the general appeal of reaction optimization in the wider chemical community as well as the innate compatibility of these objectives with simpler to build, fluidic platforms and numerical-based parameters (e.g., temperature, equivalents, yield, etc.). Many successful reaction-optimizing SDLs have been demonstrated, ranging from single-step continuous parameter optimizations,^{10,23} to multi-parameter single-step optimizations for pareto front optimizations,^{19,56} and even multi-step reaction pathway optimizations.^{11,12}

Although reaction optimization has traditionally been the central focus for optimization-based SDLs, they are not limited to only this goal. As was demonstrated by the formulation of quantum dots by the Artificial Chemist, the adhesive system by Rooney et al., or the organic laser emitter development by the combined groups of Aspuru-Gurzik/Adachi/Gryzbowski/Cronin/Hein and Burke, SDLs can also be applied to refining the intrinsic properties of materials.^{31,49,57} Despite slight variations, these SDLs all follow identical principles to the reaction-optimization SDLs in which an automated hardware platform synthesizes a new material, the material is transported to an analytical unit where a numerical value is calculated to quantify

a specific property, such as shear strength or light emission, and the analytical output is then returned as a numerical input for the AI agent. This methodological consistency strongly highlights the versatility of the underlying SDL approach and demonstrates that the same optimization concept can be more broadly applied beyond chemical reaction optimization and toward the material sciences.

Further highlighting the applicability of optimization SDLs beyond chemical synthesis, the systems have also been applied toward process optimization, most notably the optimization of chromatographic methods. Of particular interest has been the use of BO and retention modeling algorithms by Bos et al., Boelrijk et al., or Molenaar et al. (Pirok group) to optimize 1D and 2D liquid chromatography methods with the goal of improving both peak separation and analytical run time.^{53–55} While still an optimization task, this approach nicely highlights that the closed-loop SDL process can still have a significant research impact even in simpler hardware platforms, in this case a single HPLC unit. This also highlights that SDLs can be applied to tasks beyond chemical synthesis and that they hold application not only to the synthesis and materials communities but also to the analytical field.

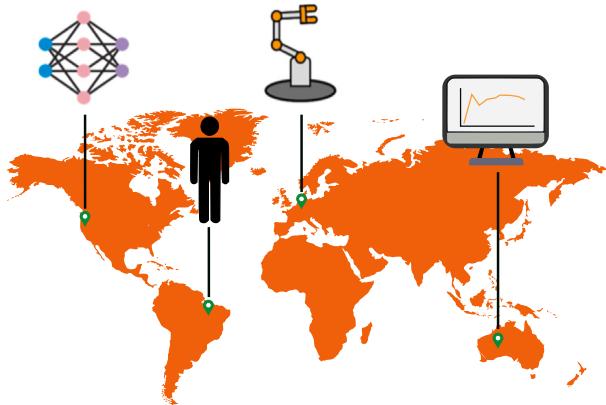
Being highly suited to optimization tasks by no means restricts SDLs to these objectives, and under the right conditions, they can also be efficiently applied to search and discovery through the use of high-throughput platforms and AI agents geared toward exploration. While direct exploration via random combinatorial events has been explored, it is a particularly challenging approach for SDLs due to the open-ended nature, the AI agent's data training requirements, and the analytical challenges in identifying and characterizing novel material. This does not mean efficient exploration cannot be performed by SDLs, but rather that more clearly defined tasks can be executed more efficiently. For example, the "Adam" system from King et al. (King group) was tasked with the identification of genes encoding for a specific enzyme expressed by a yeast species.³³ Using a general bioinformatic database of the yeast species, Adam formulated hypotheses about the genes and enzymes, tested the hypothesis using the automated platform, and after extensive rounds of testing (>6.5 million measurements) managed to identify three genes that encoded for the enzyme. By clearly defining the end target (i.e., "find the gene for ...") rather than leaving the AI to randomly explore space, the system is able to more directly discover previously unknown results or relationships. This nicely highlighted how targeting the discovery allows the SDLs to be employed in chemical discovery and elegantly demonstrates that SDLs are not limited to solely optimization tasks.

CONSIDERATIONS

Accessibility

Thanks to the custom-built nature of SDLs, a variety of design philosophies and approaches have been realized, ultimately highlighting many possible points of entry into field. Despite this, two significant barriers still remain: the high cost of platform components and the need for specific technical skills. Although chemical research is a costly endeavor by nature, the cost of many components implemented in SDLs, such as benchtop NMR, liquid chromatography-mass spectrometry, or X-ray diffraction instruments, can significantly exceed the working costs of a research laboratory.^{19,50} Moreover, the development of SDLs also demands researchers not only have a solid working knowledge of chemistry but also be skilled in programming, electrical engineering, and mechanical engineering. As such, for SDLs to be fairly democratized and adopted by the wider chemical and materials science community, solutions to both the high costs and the skill barrier must be addressed.

Decentralization



Hardware Cost Reduction

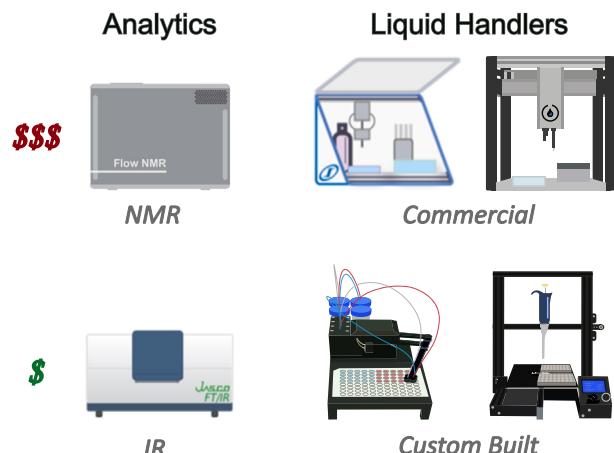


Figure 4. Approaches toward improving accessibility of SDLs for the wider community

(Left) The decentralization approach where elements of SDLs or the users may be located in different physical locations but connected via the internet, allowing use without needing to develop all parts of the SDL.

(Right) The cost-reduction approach where lower performing and cheaper analytics or liquid-handling solutions may be employed to reduce the overall cost of the physical platform, giving a lower cost of entry into the field.

The first approach toward wider SDL adoption lies in the decentralization of SDLs where researchers remotely submit experiments to an established SDL such as the LeyLab or MAOSIC (Figure 4, left).^{32,58} This approach has gained significant traction in both academic and industrial settings as it is cost-effective for the end user, does not require development time, thereby making results more immediate, and does not demand extensive programming or engineering knowledge from the user. This is a particularly appealing option for scenarios in which the specific SDL may only feature lightly in the research, for example having the cloud-SDL perform hit screening that can then be used to inform the researchers' optimization SDL, or an optimization cloud-SDL could be used to achieve a challenging step in a total synthesis. Additionally, user operation of the SDL platform is more facile as there are already developers with existing expertise. This option does, however, limit researchers to fixed, preexisting systems, which may not always suit their specific research needs. Equally, if researchers are looking to routinely use the SDL during their research process, or they need to perform commercially sensitive research, it may be desirable to have an in-house SDL with more private access.

Taking a different approach toward accessibility, others have instead focused on reducing the cost of the hardware either through the use of cheaper components, particularly liquid handlers, or by exchanging more expensive analytical instruments, such as NMR, for cheaper systems such as IR (Figure 4, right). Both the analytical instruments and liquid handlers are particularly good candidates for cost reduction as there are many cases in which the maximum capabilities are not needed, the high-end systems are very costly, and the cheaper substitutes are not exceedingly complex to build or integrate. Cost reduction in SDLs is very promising for future SDL developments, and it is becoming increasingly studied, but as it has already been reviewed elsewhere by Lo et al., the approaches toward cost reduction are not discussed further herein.⁵⁹ Although implementation of cheaper platform components and analytics does reduce the financial cost and thereby improve the accessibility of the SDL platform, it also comes at the cost of research capability. This means careful

considerations must always be made during the cost-cutting process to ensure the SDL remains effective for the desired goal.

Integrating AI and automation in SDLs leads to lower material consumption, faster result generation, and longer operation times (24/7). This creates clear financial incentives for developing SDLs, especially since methodology development and multi-step syntheses are known to offer significant financial benefits in industrial settings. Given these platforms' evident financial upside and broad application, the key financial decision is not whether these systems will provide an ROI but rather the minimum investment needed to create an effective system. The exact ROI depends on the financial value of the end goal, which varies case by case. However, cost-reducing measures such as reduced material consumption or lower hardware costs are always expected to have a positive impact. Developing low-cost hardware alternatives is valuable not only for improving the accessibility of this field but also for enhancing ROI. This, in turn, incentivizes external funders to invest in SDL research.

SDL limitations

AI augmentation has now shown significant promise in many fields, illustrating its transformative potential to solve complex challenges yet. However, as demonstrated by Beker et al. (Grzybowski group), AI still faces many limitations, particularly in regard to its dependency on data quality.⁵¹ If the training data used to build the AI model are biased and lack negative data, the AI's predictive performance is no greater than that of traditional methods. This not only highlights the importance of comprehensive and unbiased datasets but also emphasizes the need for critical evaluation of AI-generated data. Although the data quality dependency may be a weakness of AI, especially when using data for the wider literature, it does highlight the role SDLs can play in the creation of high-quality datasets for training further AI development. As a greater focus on AI development arises in the wider scientific community, the ability of SDLs to generate reliable, high-quality data containing negative results should not be underestimated. This is not only highly valuable for general AI training purposes, but it also opens significant new avenues into predictive AI strategies. This cycle of enhancement—where AI controls the generation of data that is then used to enhance its own performance—illustrates the potential for AI to overcome its current limitations, underscoring the evolving synergy between AI capabilities and the quality of underlying data.

While SDL platforms alleviate researchers from some arduous physical lab work and AI agents expedite research by identifying experiment parameters more rapidly, human researchers still play a vital role in using SDLs. The hardware, coordinator software, and AI agents are all prone to errors, including physical malfunctions, program errors, or data misinterpretation, requiring human oversight to ensure continued operation. This issue was highlighted in the Adam system, where King et al. recommended routine supervision due to its "brittle" nature. Many approaches to addressing this "brittleness" have been demonstrated, with the most common being the selection of reliable and robust research tasks. For instance, many reaction-optimization SDLs choose robust chemistries like imine condensations or palladium cross-couplings, particularly Suzuki reactions. These create fewer unexpected failures and simplify the assessment of the SDL system. While this approach is valuable during the initial phases of SDL development, it is not a long-term solution. Experimental robustness is crucial in research. As SDL development advances, more challenging and error-prone work should be attempted, and a detailed error database should be integrated into the system. Such a database tracks errors for diagnostic and development purposes and distinguishes smaller recoverable errors from significant

non-recoverable errors during runtime. Implementing secondary recovery systems, as demonstrated by Koscher et al.,⁶⁰ allows minor errors to be recovered during runtime, significantly reducing platform brittleness. Although major non-recoverable errors still cause some residual brittleness, integrating technologies like robotic arms and camera-based computer vision into the secondary recovery system will likely reduce the number of non-recoverable errors, further improving system robustness.

Beyond simple supervision, human researchers are essential in ensuring the quality and interpretation of the final data. AI agents identify correlations between variables but do not truly assess the underlying physical phenomena. Therefore, human researchers are crucial for interpreting data to understand the underlying phenomena and for establishing the research direction.

Safe usage

As AI capabilities continue to develop and accessibility to SDLs increases, it is important to consider not only the ways SDLs can revolutionize chemical research positively but also the potential for harm and necessary safeguards to prevent misuse. SDLs, like all delocalized labs, can be connected to the internet and to expensive components. They may host sensitive information and are capable of unsupervised syntheses, making them ideal targets for online hackers. Hackers could hold costly SDL equipment or sensitive information ransom. Since developing SDL platforms often requires dedicated one-off grants, a ransomware attack can be financially devastating for many labs, as they may not be able to afford to replace the ransomed equipment. This is especially true if such an attack occurs before investors see any return on their initial investment.

In more extreme cases beyond ransomware attacks, SDL platforms may be hijacked and used to create dangerous substances without the researchers' knowledge. Although the exploitation of web-based SDLs for large-scale illicit substance manufacture is unlikely, it is conceivable that such SDLs could be hijacked to create explosive or toxic materials, endangering nearby researchers. Given these hacking risks and the fact that SDLs are often built by chemists or engineers without cyber-security backgrounds, particular care must be taken to ensure these tools are cyber-secure. Developing cyber-secure systems for large-scale operations is known to be challenging. However, since the internet-related elements of web-based SDLs are not expansive, implementing basic cyber-security protocols is very achievable. By using low-level communication protocols (e.g., sockets), the points of entry to the system are limited. Additionally, there are many pre-built libraries available for data encryption and authentication, as well as extensive online support for implementing dedicated security features (e.g., IP whitelisting). Given the many resources available to implement basic cyber-security measures and the significant financial risks associated with security breaches, it is imperative that developers commit time and effort to securing their platforms. While this helps prevent security breaches, further good practices, such as storing research data on a remote server, also help reduce the impact of a ransomware attack by ensuring research data security.

Conversely, the development of cheap, open access, and easy-to-use SDLs not only improves access for legitimate scientific use but also increases the risk of exploitation for illicit drug synthesis. While this cannot be actively secured in the same manner as delocalized labs, considerations should be made regarding the publication and distribution of knowledge about creating and using these SDLs, particularly for automated synthesis of illicit materials. We believe that accessibility to scientific

literature should be managed by the scientific community not individual researchers. Therefore, decisions about the distribution of knowledge on SDL development should be left to scientific journals. This means that the publication of SDL developments should remain within scientific journals, where decisions on the safety of knowledge dissemination can be more appropriately made.

While researcher safety is a concern for internet-connected SDLs, it is also one of the significant upsides of SDL usage. These SDLs can perform multiple lab tasks, from weighing materials and loading them into reactors to conducting analytics, reducing the need for physical human intervention during an experiment cycle. Since many chemical compounds are harmful to humans, having machines handle these chemicals significantly reduces researcher exposure, ultimately leading to better health outcomes. Reduced researcher exposure also opens new avenues for more dangerous research that would have been previously avoided due to safety concerns.

OUTLOOK AND FUTURE DIRECTIONS

SDLs have only recently come to the forefront of scientific research, and although they have already seen phenomenal development and application, many avenues remain to be explored. These systems have shown clear potential for optimizations, making them prime candidates for further exploration in tasks such as NMR pulse sequence optimization or gas phase reaction optimization. While SDLs are well suited for optimization, as AI continues to develop through reinforced learning, future SDLs may be better applied to discovery challenges, such as discovering new reactions and materials or even conducting mechanistic investigations. With these AI improvements, a new phase of SDL research could emerge, where self-incentivized SDLs not only perform autonomous research but also guide the research direction without human input. Although the development of self-motivated SDLs may remain unrealized in the near future, there is a clear trajectory toward such a goal as more untouched research space in human-guided SDLs is explored.

SDLs have demonstrated their value in achieving commercially valuable goals, such as performing complete reaction optimization campaigns within a day, developing new materials like organic semiconductor lasers,⁴⁹ or identifying genes that encode for specific enzymes.³³ With clear financial incentives and increasing accessibility, the prevalence of SDL platforms is expected to rise, especially in start-up settings where early research and development costs are crucial. This trend will likely increase the demand for programming- and engineering-related skills within chemical research. As technology-based jobs continue to grow across various fields, the future workforce is likely to have already encountered these skill requirements through education or work experience. The push to increase programming exposure and computational thinking in childhood education has further prepared the upcoming generation with these skills. Modern AI tools, such as ChatGPT, reduce the need to memorize language-specific syntax, allowing students to focus more on computational thinking skills. This approach means future researchers will be better equipped to adapt to SDL development, understanding the necessary thought processes without depending on memorized programmatic syntax.

With a technologically adept future workforce, supportive AI tools, and significant financial incentives, the prevalence of SDLs is likely to increase. Given the strong applications already demonstrated, the future impact of SDLs on the wider scientific field is bound to be significant.

ACKNOWLEDGMENTS

We gratefully acknowledge the generous funding provided by the Dutch Research Council (NWO) under the Open Technology Program (Multi-Modal Photochemistry, no. 18433, to A.S. and T.N.). Furthermore, we extend our gratitude for the generous funding received from the European Union H2020 research and innovation program under the European Research Council program (FlowHAT, no. 101044355, to T.N.) and under the European Innovation Council project (reaCtor, no. 101099405, to T.N. and E.S.).

AUTHOR CONTRIBUTIONS

All authors contributed to researching and writing this perspective.

DECLARATION OF INTERESTS

The authors declare no competing interests.

REFERENCES

- Abolhasani, M., and Kumacheva, E. (2023). The rise of self-driving labs in chemical and materials sciences. *Nat. Synth.* 2, 483–492.
- Spinrad, R.J. (1967). Automation in the laboratory. *Science* 158, 55–60.
- Frediani, H.A. (1952). Automatic Karl Fischer Titration. *Anal. Chem.* 24, 1126–1128.
- Kelley, M.T., Stelzner, R.W., Laing, W.R., and Fisher, D.J. (1959). Automatic Coulometric Titrator for Karl Fischer Determination of Water. *Anal. Chem.* 31, 220–221.
- Merrifield, R.B., and Stewart, J.M. (1965). Automated Peptide Synthesis. *Nature* 207, 522–523.
- Olsen, K. (2012). The First 110 Years of Laboratory Automation: Technologies, Applications, and the Creative Scientist. *J. Lab. Autom.* 17, 469–480.
- Boyd, J. (2002). Robotic laboratory automation. *Science* 295, 517–518.
- Goodell, J.R., McMullen, J.P., Zaborenko, N., Maloney, J.R., Ho, C.X., Jensen, K.F., Porco, J.A., Jr., and Beeler, A.B. (2009). Development of an Automated Microfluidic Reaction Platform for Multidimensional Screening: Reaction Discovery Employing Bicyclo[3.2.1] octanoid Scaffolds. *J. Org. Chem.* 74, 6169–6180.
- Krishnadasan, S., Brown, R.J.C., DeMello, A.J., and DeMello, J.C. (2007). Intelligent routes to the controlled synthesis of nanoparticles. *Lab Chip* 7, 1434–1441.
- McMullen, J.P., and Jensen, K.F. (2010). An Automated Microfluidic System for Online Optimization in Chemical Synthesis. *Org. Process Res. Dev.* 14, 1169–1176.
- 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.
- Nambiar, A.M.K., Breen, C.P., Hart, T., Kulesza, T., Jamison, T.F., and Jensen, K.F. (2022). Bayesian Optimization of Computer-Proposed Multistep Synthetic Routes on an Automated Robotic Flow Platform. *ACS Cent. Sci.* 8, 825–836.
- Boiko, D.A., MacKnight, R., Kline, B., and Gomes, G. (2023). Autonomous chemical research with large language models. *Nature* 624, 570–578.
- Müller, P., Clayton, A.D., Manson, J., Riley, S., May, O.S., Govan, N., Notman, S., Ley, S.V., Chamberlain, T.W., and Bourne, R.A. (2022). Automated multi-objective reaction optimisation: which algorithm should I use? *React. Chem. Eng.* 7, 987–993.
- Seifrid, M., Pollice, R., Aguilar-Granda, A., Morgan Chan, Z., Hotta, K., Ser, C.T., Vestfrid, J., Wu, T.C., and Aspuru-Guzik, A. (2022). Autonomous Chemical Experiments: Challenges and Perspectives on Establishing a Self-Driving Lab. *Acc. Chem. Res.* 55, 2454–2466.
- Ha, T., Lee, D., Kwon, Y., Park, M.S., Lee, S., Jang, J., Choi, B., Jeon, H., Kim, J., Choi, H., et al. (2023). AI-driven robotic chemist for autonomous synthesis of organic molecules. *Sci. Adv.* 9, eadj0461.
- 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, 2001626.
- Steiner, S., Wolf, J., Glatzel, S., Andreou, A., Granda, J.M., Keenan, G., Hinkley, T., Aragon-Camarasa, G., Kitson, P.J., Angelone, D., and Cronin, L. (2019). Organic synthesis in a modular robotic system driven by a chemical programming language. *Science* 363, eaav2211.
- Slattery, A., Wen, Z., Tenblad, P., Sanjóse-Orduna, J., Pintossi, D., den Hartog, T., and Noél, T. (2024). Automated self-optimization, intensification, and scale-up of photocatalysis in flow. *Science* 383, eadj1817.
- 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.
- 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.
- Clayton, A.D., Pyzer-Knapp, E.O., Purdie, M., Jones, M.F., Barthelme, A., Pavey, J., Kapur, N., Chamberlain, T.W., Blacker, A.J., and Bourne, R.A. (2023). Bayesian Self-Optimization for Telescopied Continuous Flow Synthesis. *Angew. Chem., Int. Ed. Engl.* 62, e202214511.
- Moore, J.S., and Jensen, K.F. (2012). Automated Multitrajectory Method for Reaction Optimization in a Microfluidic System using Online IR Analysis. *Org. Process Res. Dev.* 16, 1409–1415.
- Sans, V., Porwol, L., Dragone, V., and Cronin, L. (2015). A self optimizing synthetic organic reactor system using real-time in-line NMR spectroscopy. *Chem. Sci.* 6, 1258–1264.
- Holmes, N., Akien, G.R., Savage, R.J.D., Stanetty, C., Baxendale, I.R., Blacker, A.J., Taylor, B.A., Woodward, R.L., Meadows, R.E., and Bourne, R.A. (2016). Online quantitative mass spectrometry for the rapid adaptive optimisation of automated flow reactors. *React. Chem. Eng.* 1, 96–100.
- Mohapatra, S., Hartrampf, N., Poskus, M., Loas, A., Gomez-Bombarelli, R., and Pentelute, B.L. (2020). Deep Learning for Prediction and Optimization of Fast-Flow Peptide Synthesis. *ChemRxiv*. <https://doi.org/10.26434/chemrxiv-12649343>.
- Reizman, B.J., Wang, Y.-M., Buchwald, S.L., and Jensen, K.F. (2016). Suzuki–Miyaura cross-coupling optimization enabled by automated feedback. *React. Chem. Eng.* 1, 658–666.
- Baumgartner, L.M., Coley, C.W., Reizman, B.J., Gao, K.W., and Jensen, K.F. (2018). Optimum catalyst selection over continuous and discrete process variables with a single droplet microfluidic reaction platform. *React. Chem. Eng.* 3, 301–311.
- Moliner, M., Serra, J., Corma, A., Argente, E., Valero, S., and Botti, V. (2005). Application of

- artificial neural networks to high-throughput synthesis of zeolites. *Microporous Mesoporous Mater.* 78, 73–81.
30. Jiang, Y., Fakhrudeen, H., Pizzuto, G., Longley, L., He, A., Dai, T., Clowes, R., Rankin, N., and Cooper, A.I. (2023). Autonomous biomimetic solid dispensing using a dual-arm robotic manipulator. *Dig. Dis.* 2, 1733–1744.
 31. Rooney, M.B., MacLeod, B.P., Oldford, R., Thompson, Z.J., White, K.L., Tungunyatham, J., Stankiewicz, B.J., and Berlinguette, C.P. (2022). A self-driving laboratory designed to accelerate the discovery of adhesive materials. *Dig. Dis.* 1, 382–389.
 32. Li, J., Li, J., Liu, R., Tu, Y., Li, Y., Cheng, J., He, T., and Zhu, X. (2020). Autonomous discovery of optically active chiral inorganic perovskite nanocrystals through an intelligent cloud lab. *Nat. Commun.* 11, 2046.
 33. King, R.D., Rowland, J., Oliver, S.G., Young, M., Aubrey, W., Byrne, E., Liakata, M., Markham, M., Pir, P., Soldatova, L.N., et al. (2009). The Automation of Science. *Science* 324, 85–89.
 34. Christensen, M., Yunker, L.P.E., Shiri, P., Zepel, T., Prieto, P.L., Grunert, S., Bork, F., and Hein, J.E. (2021). Automation isn't automatic. *Chem. Sci.* 12, 15473–15490.
 35. Coley, C.W., Green, W.H., and Jensen, K.F. (2018). Machine Learning in Computer-Aided Synthesis Planning. *Acc. Chem. Res.* 51, 1281–1289.
 36. Raghavan, P., Haas, B.C., Ruos, M.E., Schleinitz, J., Doyle, A.G., Reisman, S.E., Sigman, M.S., and Coley, C.W. (2023). Dataset Design for Building Models of Chemical Reactivity. *ACS Cent. Sci.* 9, 2196–2204.
 37. Sim, M., Vakili, M.G., Strieth-Kalthoff, F., Hao, H., Hickman, R.J., Miret, S., Pablo-García, S., and Aspuru-Guzik, A. (2024). ChemOS 2.0: An orchestration architecture for chemical self-driving laboratories. *Matter*. <https://doi.org/10.1016/j.matt.2024.04.022>.
 38. Roch, L.M., Häse, F., Kreisbeck, C., Tamayo-Mendoza, T., Yunker, L.P.E., Hein, J.E., and Aspuru-Guzik, A. (2018). ChemOS: Orchestrating autonomous experimentation. *Sci. Robot.* 3, eaat5559.
 39. Mehr, S.H.M., Craven, M., Leonov, A.I., Keenan, G., and Cronin, L. (2020). A universal system for digitization and automatic execution of the chemical synthesis literature. *Science* 370, 101–108.
 40. Haas, C.P., Lübbesmeyer, M., Jin, E.H., McDonald, M.A., Koscher, B.A., Guimond, N., Di Rocco, L., Kayser, H., Leweke, S., Niedenführ, S., et al. (2023). Open-Source Chromatographic Data Analysis for Reaction Optimization and Screening. *ACS Cent. Sci.* 9, 307–317.
 41. de Almeida, A.F., Moreira, R., and Rodrigues, T. (2019). Synthetic organic chemistry driven by artificial intelligence. *Nat. Rev. Chem* 3, 589–604.
 42. Clayton, A.D., Manson, J.A., Taylor, C.J., Chamberlain, T.W., Taylor, B.A., Clemens, G., and Bourne, R.A. (2019). Algorithms for the self-optimisation of chemical reactions. *React. Chem. Eng.* 4, 1545–1554.
 43. Kusne, A.G., Yu, H., Wu, C., Zhang, H., Hattrick-Simpers, J., DeCost, B., Sarker, S., Oses, C., Toher, C., Curtarolo, S., et al. (2020). On-the-fly closed-loop materials discovery via Bayesian active learning. *Nat. Commun.* 11, 5966.
 44. Jablonka, K.M., Schwaller, P., Ortega-Guerrero, A., and Smit, B. (2024). Leveraging large language models for predictive chemistry. *Nat. Mach. Intell.* 6, 161–169. <https://doi.org/10.1038/s42256-023-00788-1>.
 45. Ji, W., Qiu, W., Shi, Z., Pan, S., and Deng, S. (2021). Stiff-PINN: Physics-Informed Neural Network for Stiff Chemical Kinetics. *J. Phys. Chem. A* 125, 8098–8106.
 46. King, R.D., Whelan, K.E., Jones, F.M., Reiser, P.G.K., Bryant, C.H., Muggereton, S.H., Kell, D.B., and Oliver, S.G. (2004). Functional genomic hypothesis generation and experimentation by a robot scientist. *Nature* 427, 247–252.
 47. McNally, A., Prier, C.K., and MacMillan, D.W.C. (2011). Discovery of an α -amino C-H arylation reaction using the strategy of accelerated serendipity. *Science* 334, 1114–1117.
 48. Wang, H., Fu, T., Du, Y., Gao, W., Huang, K., Liu, Z., Chandak, P., Liu, S., Van Katwyk, P., Deac, A., et al. (2023). Scientific discovery in the age of artificial intelligence. *Nature* 620, 47–60.
 49. Strieth-Kalthoff, F., Hao, H., Rathore, V., Derasp, J., Gaudin, T., Angello, N.H., Seifrid, M., Trushina, E., Guy, M., Liu, J., et al. (2024). Delocalized, asynchronous, closed-loop discovery of organic laser emitters. *Science* 384, eadk9227.
 50. Szymanski, N.J., Rendy, B., Fei, Y., Kumar, R.E., He, T., Milsted, D., McDermott, M.J., Gallant, M., Cubuk, E.D., Merchant, A., et al. (2023). An autonomous laboratory for the accelerated synthesis of novel materials. *Nature* 624, 86–91.
 51. Beker, W., Roszak, R., Wołos, A., Angello, N.H., Rathore, V., Burke, M.D., and Grzybowski, B.A. (2022). Machine Learning May Sometimes Simply Capture Literature Popularity Trends: A Case Study of Heterocyclic Suzuki–Miyaura Coupling. *J. Am. Chem. Soc.* 144, 4819–4827.
 52. Volk, A.A., and Abolhasani, M. (2024). Performance metrics to unleash the power of self-driving labs in chemistry and materials science. *Nat. Commun.* 15, 1378.
 53. Boelrijk, J., Ensing, B., Forré, P., and Pirok, B.W.J. (2023). Closed-loop automatic gradient design for liquid chromatography using Bayesian optimization. *Anal. Chim. Acta* 1242, 340789.
 54. Molenaar, S.R.A., Bos, T.S., Boelrijk, J., Dahlseid, T.A., Stoll, D.R., and Pirok, B.W.J. (2023). Computer-driven optimization of complex gradients in comprehensive two-dimensional liquid chromatography. *J. Chromatogr. A* 1707, 464306.
 55. Bos, T.S., Boelrijk, J., Molenaar, S.R.A., van 't Veer, B., Niezen, L.E., van Herwerden, D., Samanipour, S., Stoll, D.R., Forré, P., Ensing, B., et al. (2022). Chemometric Strategies for Fully Automated Interpretive Method Development in Liquid Chromatography. *Anal. Chem.* 94, 16060–16068.
 56. Jablonka, K.M., Jothiappan, G.M., Wang, S., Smit, B., and Yoo, B. (2021). Bias free multiobjective active learning for materials design and discovery. *Nat. Commun.* 12, 2312.
 57. Angello, N.H., Rathore, V., Beker, W., Wołos, A., Jira, E.R., Roszak, R., Wu, T.C., Schroeder, C.M., Aspuru-Guzik, A., Grzybowski, B.A., and Burke, M.D. (2022). Closed-loop optimization of general reaction conditions for heteroaryl Suzuki–Miyaura coupling. *Science* 378, 399–405.
 58. Mateos, C., Nieves-Remacha, M.J., and Rincón, J.A. (2019). Automated platforms for reaction self-optimization in flow. *React. Chem. Eng.* 4, 1536–1544.
 59. Lo, S., Baird, S.G., Schrier, J., Blaiszik, B., Carson, N., Foster, I., Aguilar-Granda, A., Kalinin, S.V., Maruyama, B., Politi, M., et al. (2024). Review of low-cost self-driving laboratories in chemistry and materials science: the “frugal twin” concept. *Dig. Dis.* 3, 842–868. <https://doi.org/10.1039/D3DD00223C>.
 60. Koscher, B.A., Canty, R.B., McDonald, M.A., Greenman, K.P., McGill, C.J., Bilodeau, C.L., Jin, W., Wu, H., Vermeire, F.H., Jin, B., et al. (2023). Autonomous, multiproperty-driven molecular discovery: From predictions to measurements and back. *Science* 382, eadi1407.