Preprint Cambridge Centre for Computational Chemical Engineering ISSN 1473 – 4273

The Digital Lab Framework as part of The World Avatar

Simon D. Rihm[1,2,3], Jiaru Bai[2], Aleksandar Kondinski[1,2], Sebastian Mosbach[1,2], Jethro Akroyd[4], Markus Kraft[1,2,5,6]

released: October 23, 2023

[1] CARES
Cambridge Centre for Advanced Research and Education in Singapore
1 Create Way
CREATE Tower, #05-05
Singapore, 138602

[2] Department of Chemical Engineering and Biotechnology
University of Cambridge
Philippa Fawcett Drive
Cambridge, CB3 0AS
United Kingdom

[3] Department of Chemical & Biomolecular Engineering
National University of Singapore
4 Engineering Drive 4
Singapore, 117585

[4] CMCL Innovations
Sheraton House
Cambridge
CB3 0AX
United Kingdom

[5] School of Chemical and Biomedical Engineering
Nanyang Technological University
62 Nanyang Drive
Singapore, 637459

[6] The Alan Turing Institute
2QR, John Dodson House
96 Euston Road
London, NW1 2DB
United Kingdom

Preprint No. 314

UNIVERSITY OF CAMBRIDGE

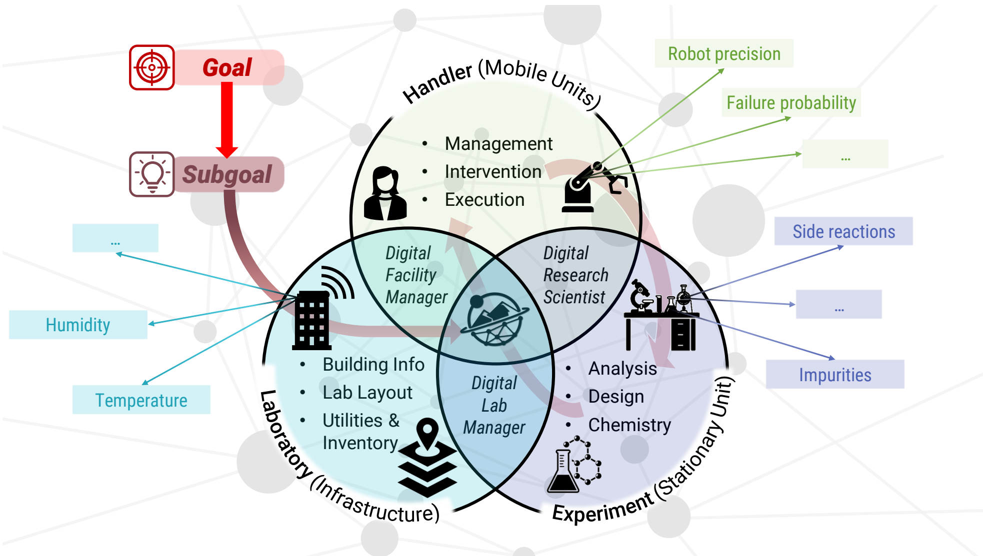
Keywords: holistic laboratory automation, knowledge graph, AI scientist, comprehensive digital twin

Edited by
Computational Modelling Group
Department of Chemical Engineering and Biotechnology
University of Cambridge
Philippa Fawcett Drive
Cambridge, CB3 0AS
United Kingdom

E-Mail: mk306@cam.ac.uk
World Wide Web: `https://como.ceb.cam.ac.uk/`

Abstract

In order to tackle many of humanity's most pressing challenges, scientific discovery needs to be substantially accelerated. The automation of experimental research activities plays a big role in this, from ubiquitous software tools to "self-driving laboratories". Recently, the idea of an "AI scientist" that can make Nobel-worthy discoveries has been introduced in this context. We argue, that current platform-based approaches are insufficient and might even limit further development. Therefore, we introduce the digital lab framework as a holistic approach to laboratory automation. Its hierarchical and semantic structure allows for deep knowledge representation across different domains and scales which is necessary to further interoperability by widening the search and optimisation space to include managerial tasks in research labs as well as information on infrastructure and buildings. This way we can ensure cost effectiveness, improve reproducibility, and bridge the "interim technology gap". To address common challenges related to interoperability and adaptability, this framework is developed as part of "The World Avatar" ecosystem, based on interconnected dynamic knowledge graphs. Viewing at the challenges at hand from a systems engineering perspective, we aim to integrate all aspects of lab work and its automation, contrasting the many isolated solutions available that – amongst others – increase the risk of manufacturer lock-in. The goal-driven architecture enables subsequent design of experiments, and optimal resource distribution according to freely definable research goals.



Highlights

- A holistic lab framework pushes a paradigm shift to include all aspects of experimental research.
- Dynamic knowledge graph technology enables connected lab digital twins.
- A systems engineering approach allows for goal-driven self-driving labs.

Contents

1 Introduction

The world is currently facing several high-level challenges such as sustainability, climate change, and healthcare. Technological advancements have enabled the development of technical solutions to help solve some of these. However, to fulfil critical goals – such as achieving the United Nations' sustainable development goals (SDGs), research and development (R&D) need to be accelerated drastically. One of the biggest levers for such an acceleration is the automation of research laboratories.

Laboratory automation involves the use of technology to perform experiments, analytical procedures, and related activities without direct human intervention. Thereby, the efficiency and productivity of laboratories can be significantly increased while less time and resources are required to conduct experiments, enabling the collection of larger datasets. Automation can furthermore minimise human error, which improves reproducibility and precision. These advantages have made laboratory automation an increasingly popular approach for many research organisations.

In order to analyse the current state of laboratory automation and related challenges from a bird's eye perspective, we adopt a systems engineering approach in this paper. The associated thinking and process [24] guides our work in general, but most importantly the sequence within this introduction: first, we construct an ideal future scenario to derive requirements and define goals in section 1.1 and compare it to current existing research directions. Then a functional analysis of these requirements is conducted in section 1.2 by taking on a systems view going beyond the underlying concepts of existing solutions. Finally, we can synthesise a general system design to accomplish these goals in section 1.3.

1.1 Constructing a scenario: Lab of the future

Imagine a world in which this urgency is taken with adequate seriousness in terms of resource allocation and policymaking, such that the rate of scientific discovery has been increased by orders of magnitude. How would such a world look like?

Human researchers would only need to formulate their research goals, and autonomous software and hardware agents take care of the rest. These agents can define the concrete steps required to achieve set high-level goals, design and execute experiments, analyse and interpret results, manage related inventory and resources, and even suggest new hypotheses or goals for further investigation. Laboratory automation has been perfected, resulting in highly efficient and productive research environments. Artificial intelligence (AI) powered systems analyse challenges and goals to break them down into smaller subgoals which can then be investigated by experimental or theoretical research. For a laboratory, this means that adequate experiments are automatically designed, executed, analysed, and used as a basis for the next experiment to maximise knowledge gain. Not only that, all consumables necessary are ordered automatically according to an optimised experimental schedule which also takes equipment purchases and maintenance into account as well as monitoring of the structural integrity of the building we are in.

It can be easily seen that this scenario is not limited to "total laboratory automation" [65, 94] in the narrow sense of interconnected machinery in an assembly line nor in the broader sense of integrated design, execution, and analysis of experiments – but includes all above aspects related to a (chemical) research laboratory. As humans only define abstract goals, this is going even beyond the RSC scenario of "push-button chemistry" [78]. The closest formulated vision to date is that of an "AI scientist" capable of making Nobel-prize worthy scientific discoveries [55]. This has been discussed in the context of the Nobel Turing challenge, which has been proposed as a Grand Challenge for the Turing Institute [35] in addition to the three active ones under "Towards Turing 2.0" [1]. While this challenge is focused on the software aspect of research automation, one can easily recognise the need for such an AI scientist to be equipped with robotic capabilities for full autonomy. Furthermore, it would need to be embedded in a very broad knowledge model, encompassing everything from quantum mechanics over robotic precision models to energy and resource prices. Next to achieving research goals alone, aspects of sustainability and efficiency also need to be addressed [86, 96]: What is the ideal sequence or timing of experiments to minimise the environmental footprint? Can the setup be changed to reduce power consumption or reactants be swapped to greener alternatives? And is there an optimum cut-off point where diminishing returns in terms of knowledge gain are outweighed by the impact of resources consumed by the experiments? A fully autonomous research laboratory takes these questions into account when making decisions.

We therefore need to consider approaches that go beyond the traditional methods of laboratory automation. While robotic liquid handlers and high-throughput screening platforms are important components of laboratory automation, they only address a limited range of tasks. We also need to automate supporting tasks such as inventory management, resource allocation, and cost reduction. In doing so, we can reduce the time and effort required to manage laboratory operations, enabling researchers to focus on their core scientific activities. Recently an expert panel at SLAS Europe 2023 [81] discussing the "lab of the future" emphasised the necessity of workflow integration as the level of interoperability increases between data, instruments and software and from scheduling to interconnected or even distributed workflows. Going forward, the field needs to put more focus on software to fit in the existing environment and track the provenance of decision-making during the automated workflow [81].

1.2 System analysis: Aspects and requirements

At its core, laboratory operations always involve the interplay of three key aspects: the experiment itself, the handler conducting and monitoring the experiment, and the laboratory infrastructure. These aspects overlap and interact with each other as shown in Fig. 1, and each presents unique challenges for digitisation or automation.

The first aspect, the experiment itself, encompasses everything from the setup and preparation of samples to the measurement and analysis of results. The underlying knowledge occupies a wide spectrum, ranging from quantum chemistry to practical intuitions tubing. Even though there are many digital tools available now aiding chemists in these complex tasks, much of this knowledge is currently considered "tacit" and certain processes impossible to automate as they also require a high degree of intellectual flexibility. For example, automating a high-throughput screening assay for drug discovery requires a different approach than automating a complex synthesis reaction in organic chemistry.

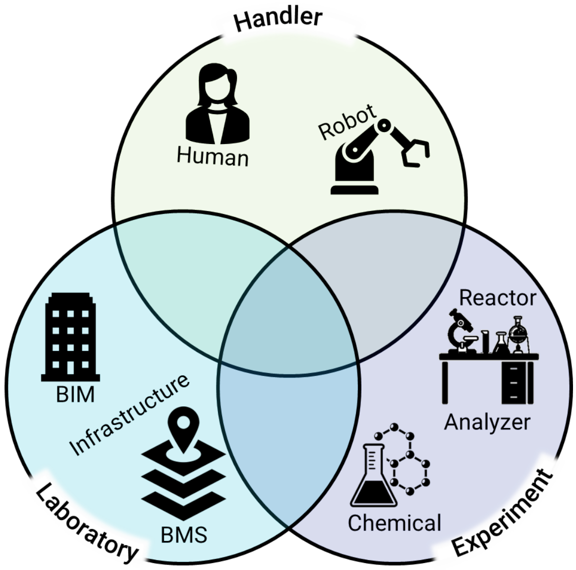


Figure 1: *Our perspective on laboratory tasks and their automation potential.*

The second aspect, the handler conducting and monitoring the experiment, refers to the humans or robots responsible for carrying out the experimental protocols. With the advancement of robotics, the automation of handlers has become widely available to researchers, which can help to reduce error rates, increase throughput, and improve consistency. The challenge here is to ensure that the handlers are able to perform the tasks efficiently and accurately and that they are able to adapt to changing experimental conditions or unexpected results. Many allegedly automated experiments still require manual intervention, such as adjusting reaction conditions or troubleshooting equipment, making complete automation difficult to achieve.

The third aspect, laboratory infrastructure, includes everything from the physical layout of the laboratory to the availability of equipment and resources. Automation of laboratory infrastructure is crucial for achieving the seamless integration of all laboratory activities while providing safe and secure laboratory environments with appropriate levels of access control and monitoring. While great strides have been made in digitising building operations and design, these efforts are currently disconnected from laboratory automation efforts, although the lab of the future needs this information [81]. The design of such a lab could be very different compared to how they currently are if we consider the robots to be part of the regular participants of the lab.

The systemic view shown in Fig. 1 has some similarities with the conceptual view of the "sharework ontology for human robot collaboration" (SOHO) [104]. Particular areas of overlap are the environment context in SOHO and the infrastructure aspect shown here, as well as the behaviour context in SOHO and the handler aspect shown here. Yet, the production context in SOHO only covers objectives and operations and is therefore much more limited than the experiment aspect. Process knowledge is either not integrated or considered part of the environmental context. Furthermore, SOHO is aimed at human-robot collaboration in a more general and simplified context, *i.e.*, on a purely mechanical level for assembly, than needed for holistic and full automation of research laboratories.

The systemic view can be potentially understood better by analogous application to automobiles – especially with regards to increasing digitisation and autonomy. Available infrastructure are analogous roads, junctions, *etc.*, as well as knowledge about expected “traffic”. Stationary units are analogous to components of the car, most importantly sensors and the data they collect. Mobile units are analogous to moving parts such as engines and wheels, as well as controlling entities – be they human or software-based. The same systemic view of experimentation in a laboratory can also be applied to computational experiments but is not the main focus of this work.

1.3 Road to success: Design of solutions

When lab automation is discussed, it usually refers to robots carrying out experimental work [17], even claims about total laboratory automation refer to this [94]. Over the last years, interest in robotics has increased exponentially and overtaken interest in more general automation in research laboratories – as for example indicated by usage frequency in publication titles within biomedicine [44]. Recently, software models of experimental parameter spaces accompanying robotic handler setups have been developed under the term “self-driving laboratory” (SDL) which has become very popular [2, 6, 69]. This term is based on the same analogy to self-driving cars as established in section 1.2. Similar to the vision of truly autonomous cars, SDLs of the future are expected to communicate and coordinate with other SDLs to optimise surrounding traffic and also take care of maintenance, manage resources, and adapt strategy based on underlying goals. To achieve this, a holistic approach will be necessary as shown in Fig. 2: considering all three aspects of lab activities enables automation not only of tasks typically associated with a lab technician but also those of additional stakeholders such as lab managers and facility managers.

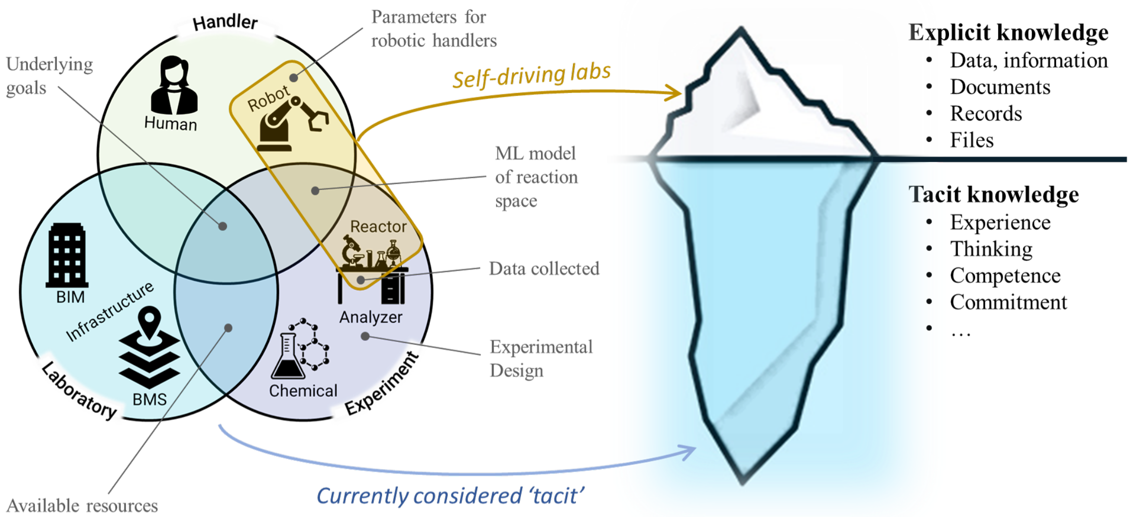


Figure 2: *Self-driving labs only incorporate the “tip of the iceberg” of related knowledge.*

In such a scenario, not only would the experiment itself need to be automated, but the AI scientist would also need to have knowledge of the entire research environment, including the laboratory infrastructure and the expertise of the handler conducting the experiments.

By embracing a holistic approach to lab automation, we can not only accelerate scientific discovery and innovation but also pave the way for the development of truly autonomous research systems. It is apparent, that this requires a much deeper and broader understanding of an experiment to be baked into such systems. This goes beyond the very explicit data fed into AI models which constitutes only a small part of underlying knowledge necessary to successfully design and conduct experiments in pursuit of a greater goal [60]. In this perspective, we propose a paradigm shift in the area of laboratory automation that will enable us to move towards this future scenario. Traditional methods of laboratory automation are too narrow in scope and therefore propose a holistic approach to laboratory automation that takes into account all three aspects of laboratory activities.

2 Background: Current state and existing solutions

Although laboratory automation has been a topic of interest for decades, the field remains highly fragmented, with numerous independent solutions for specific aspects of research activities. While there are a variety of tools and platforms available for note-taking, handler monitoring, and laboratory infrastructure management, few of these solutions are integrated with each other. This fragmentation has resulted in a scattered landscape where laboratories must cobble together disparate systems, resulting in inefficiencies, errors, and a lack of scalability. In this section, we will review current automation efforts and examine the challenges associated with integrating these disparate solutions.

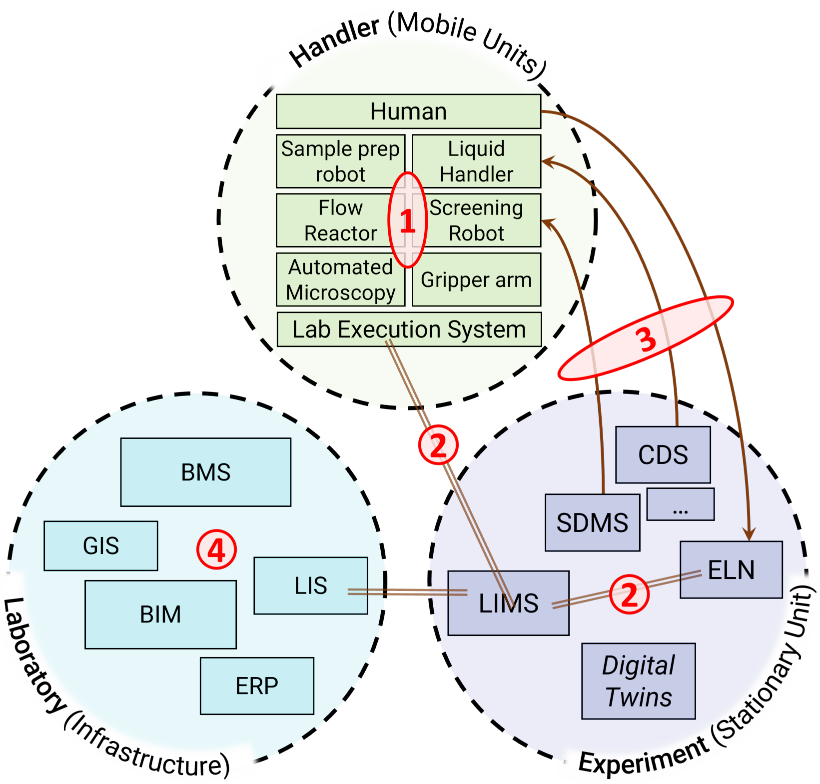


Figure 3: *Current landscape of "island solutions" for lab-related automation challenges.*

2.1 Experiment automation and digitisation: An overview

The experiment aspect of lab work includes a variety of tasks such as analysis, design, and chemistry documentation. As an example, analysis tasks may involve the measurement and interpretation of spectra, while design tasks could involve the use of machine learning (ML) to generate optimised experimental designs. Several software solutions exist for each of these tasks, such as laboratory information management systems (LIMS) for tracking samples and electronic lab notebooks (ELNs) like LabArchives and SciNote are available for chemistry documentation [59]. Meanwhile, the sheer amount of data generated by some instruments necessitates dedicated tools to control and manage, such as scientific data management systems (SDMS) or chromatography data systems (CDS).

Each of these solutions has unique challenges that make holistic lab automation difficult: For example, existing analysis tools may not be compatible with certain equipment or data formats, while ML-based experiment design tools may require significant non-digitised expertise to use effectively. LIMS are often more tailored towards industry environments rather than academic research. The market overview shown in Fig. 4 also shows that some of these platforms include controls of robotic capabilities which are mostly manufacturer-specific though as discussed in section 2.2. ELNs – for which a similar overview is provided by Jablonka et al. [49] – may have limited functionality and may not integrate well with other laboratory software tools. In order to understand new data autonomously, it needs to be automatically converted to an interoperable format which is not the case for most current ELNs [49].

Product	Manufacturer specific?	Open Source	Track samples	Manage workflow	Manage inventory	Includes ELN	SaaS option	Robotics capabilities
BlazeLIMS	-	No	✓	✓	✓	✗	✓	✗
LabWare	-	No	✓	✓	✓	✓	✓	✗
Sample Manager	ThermoFisher SCIENTIFIC	No	✓	✓	✗	✗	✗	✓
LabVantage	-	No	✓	✓	✓	✓	✓	✗
LIMSey	-	No	✓	✓	✗	✗	✓*	✗
SLIMS	Agilent Technologies	No	✓	✓	✓	✓	✓	✓
STARLIMS	-	No	✓	✓	✓	✓	✓	✓**
Occhiolino	-	Yes	✓	✓	✓	✗	✗	✗
Bika LIMS	-	Yes	✓	✓	✗	✗	✓*	✓**

* Exclusively web-based
** SOP enforcement

Figure 4: *Market overview for LIMS solutions and their capabilities.*

The experiment aspect also entails all stationary units related to the experiment, including equipment, sensors, and software solutions, as well as the underlying knowledge models that guide experimental design and analysis. Experimental setups need detailed reporting as they can influence measurement results significantly. Often, these results are treated as part of known base reality (or "ground truth" [13]) but fail to account for uncertainties and ambiguities introduced by specific setups and conditions [23]. For holistic automation, this requires a formalised description of experimental design, methods, and technologies, as well as object models, background knowledge, and reasoning rules for the interpretation of results [98]. Such a formalised description would also allow for a reactionware-based approach to decouple experiment reporting from specific equipment and even enable on-demand 3D-printing of components [18].

2.2 Role of handlers: Humans and robots in the lab

The handler aspect of lab work includes tasks related to the management, intervention, and execution of experiments. Management tasks involve the scheduling and coordination of experiments, while intervention tasks involve making adjustments or providing oversight during experiments. Finally, execution tasks involve physically carrying out the experiment steps. In a nutshell, this aspect entails all mobile units related to the experiment – robotic and human.

The use of robotic handlers in research laboratories enables high-throughput experimentation and has therefore seen a dramatic rise [2, 100] but manual intervention is still necessary for many tasks due to persisting limitations [98]. Latest advancements in the development of different robotic handling systems as well as AI now promise automation on an unprecedented scale [2]: novel ML algorithms can process large data quantities quickly and derive optimal input parameters for the next experiment while high-precision robotic systems efficiently perform diverse tasks – from mixing to controlling flow reactors and carrying out spectral analyses. The culmination of this development is currently self-driving laboratories, which perform closed-loop experimentation (see appendix B.1) to minimise a given objective function [95]. Successful applications have been demonstrated for specific experimental campaigns particularly in the fields of material science [13, 50, 64] and biotechnology [44, 72, 97].

These "closed loops" are quite narrowly defined though and rarely allow for dynamic changes in setup or objective function. Substantial human input is still needed on all levels – particularly around peripheral data and activities, for example maintenance and resource allocation tasks. To reduce this necessity going forward, closer interactions are required between so-called Laboratory Execution Systems (LES) managing robotic handlers and those monitoring the state of devices and reactants, procedural prerequisites and maintenance, as well as chemical or physical data. While LES are often connected with LIMS or ELN to update track workflows, update inventory, and document results of repeating experiments, actual scientific knowledge needs to be consulted and applied by a human researcher to draw conclusions and plan next steps beyond the current experimental campaign.

Furthermore, integration of different functionality and devices can be challenging as drivers might not be readily available or require extensive programming experience. While LES might exhibit predefined routines for standard tasks, they rarely incorporate maintenence or cleaning routines which is left to the user. Lastly, the handler itself might introduce uncertainties and errors into the system as well. These would need to be accounted for by operational models in their digital representations or digital twins. Combining such digital twins with autonomous experimentation platforms (AEP) is therefore discussed as a major milestone going forward [108], see also section 2.4.

2.3 Infrastructure management: Buildings, utilities, and inventory

The laboratory aspect entails all aspects of physical infrastructure, including utilities, building layout, and inventory management, as well as the underlying knowledge models necessary to automate these tasks. The Royal Society of Chemistry (RSC) recently stated in a perspective on sustainable laboratories that a holistic view is needed to consider all activities in a lab, its surrounding building or city, associated storage, and so on [86]. They therefore argue for incorporating life-cycle analysis (LCA) into the design of experiments (DoE). This requires more sophisticated tools to calculate expected footprint as well as implications on cost, health, and safety.

Automation within the laboratory aspect requires the digitisation and integration of various building-related information and laboratory-specific utilities. Various software solutions exist for different tasks already exist: Building Information Modelling (BIM) can be used to create a 3D digital twin of the laboratory that can help optimise lab layout, including equipment placement and routing of utilities such as heating, ventilation and air conditioning (HVAC), water, and electricity. Building management systems (BMS) can be used to monitor and control the laboratory environment, ensuring that conditions are optimal for experiments. Laboratory inventory systems (LIS) or "enterprise resource planning" (ERP) solutions can be used to manage inventory and track the usage of reagents and other supplies. Geographic information system (GIS) also offer 3D representations. These different solutions however are often siloed and not interoperable with each other [84].

Traditionally, the entire infrastructure aspect is neglected in laboratory automation efforts. With rising energy costs and climate awareness, cost and energy efficiency as well as sustainability become more and more important though. Even though lab activities contribute too high consumption of electricity, cooling water, as well as other resources, potential savings are hard to quantify as relevant systems are not part of the typical researcher's considerations. For this reason, different guidelines have been developed to increase resource efficiency and reduce waste production, for example, the Laboratory Efficiency Assessment Framework (LEAF) [86, 103]. Ever-increasing affordability of interconnected sensors and microcontrollers under the umbrella term "internet of things" (IoT) opens up additional possibilities to optimise operations in this regard, from access monitoring to automated HVAC control. A few examples of integration with research labs can be found under the phrase "smart lab" [21, 52] but these are still in their infancy.

These facets are usually considered separate from actual research but the general environment of an experiment has a significant impact on measurements. Only with rigorous reporting of these conditions (*e.g.*, temperature and humidity instead of "standard conditions"), can systems be empowered to re-evaluate their own results regarding uncertainties, potential impurities, *etc.* [23]. Especially in systems striving for autonomy, "quantification of signal vs. background noise [is] necessary to improve signal decomposition and phase mapping" as Montoya et al. [74] aptly put it in a recent article. It is particularly important to collect uncertainties of key parameters that influence subsequent decision-making [107]. Robots also need to have knowledge of the environment to reason about, act in, and modify it [104]. This includes actions by other robots and human operators as mentioned in section 2.2. In this, considering spatial layout and related constraints is critical [44], *e.g.*, via 3D models.

2.4 Standardisation and integration: Data, digital twins, workflows

The need to standardise reporting is obvious in all aspects. In the current landscape of data representation in this space, many concepts are available to describe underlying chemistry, but few for methods and procedures, and none include laboratory infrastructure. For the experiment aspect, deep chemical domain knowledge and data standardisation are crucial. As shown in Fig. 5, significant progress has been made in this area, yet there are still challenges that need to be addressed. One of the main challenges is the difficulty of incorporating human expertise and knowledge into the automation process, as much of this knowledge is considered tacit and not easily captured in data. Moreover, there is a need for better tools and methods to handle the complexity of chemical synthesis and experimentation, including the ability to manage and track large amounts of data and metadata. Lastly, the integration of digital twins of experimental equipment with representations of experimental sequences is necessary to ensure comparability and reproducibility. Often, information given in a publication is not sufficient to confidently reproduce results though [32, 49]. Even though some progress has been made on the standardisation of synthesis step sequences (*e.g.*, XDL [101] or SiLA [10]), a deeper integration with scientific knowledge and additional concepts will be necessary to unambiguously describe reaction conditions and optimise experimental setups based on arbitrary objectives.

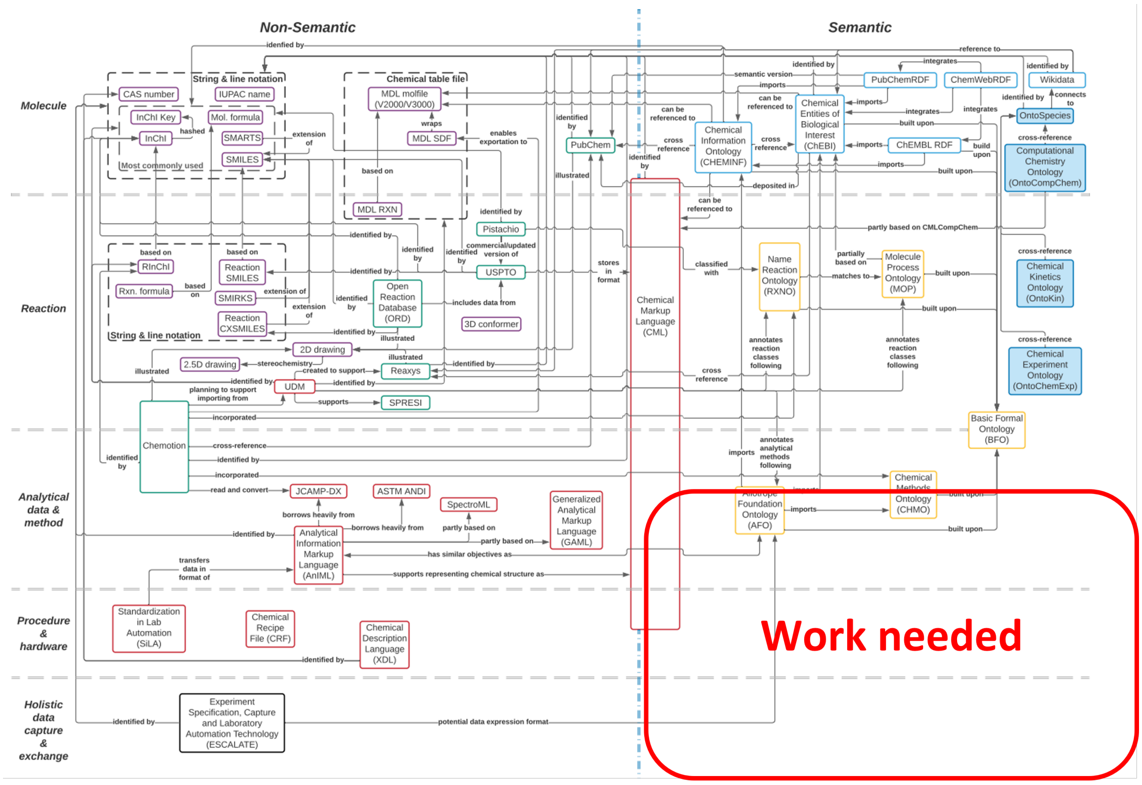


Figure 5: *Landscape of data representation in this space provided by Bai et al. [6]*

There is a need for structured data which has been expressed throughout the literature, including very recently [23, 72]. One potential solution is to adopt ontologies, which has been discussed in this context – mostly for ensuring experimental reproducability [98].

For reporting measurement data, the use of structured AFO ontologies has become standard. Going one step further, "extracting data from experimental workflows and integrating them into a graph database [can ensure] easier access and long-term management" [74]. This can be achieved with semantic representations which are most advanced in the dissemination of data which can then be potentially used for planning new experiments as well [28], *e.g.*, for chemical species. It is still challenging to make different data types accessible, especially as information might be conflicting or based on different assumptions [55, 80]. A reliable and simple system for reporting measurements including metadata is therefore key [28, 80]. The full power of semantic web technologies will only be achieved when the corresponding tools work well across all stages of the experiment, particularly in the laboratory itself [28].

Semantic representations of not only experimental procedure or measurement data are most advanced in the field of biotechnology, see for example EXPO ontology [97] for a very wide range of experiment-related concepts. This also overlaps with the representation of handlers which can be machines or humans. One application is the creation of digital twins, which can simulate the behaviour of a physical system and provide a platform for testing different control strategies and scenarios. In the context of SDLs, such models are usually achieved by ML techniques such as neural nets, leading to essentially behavioural black boxes. While these black boxes can capture the behaviour of a system within the range of conditions it was trained in and even surrogate models can be derived, there are always underlying assumptions [13] which are often times not captured adequately.

It seems imperative to combine these ML models with actual knowledge about the systems in order to turn them into grey or even white boxes. The challenge lies in achieving flexible "human-machine interactions" [108] and interventions. Again, the field of biotechnology leads the charge here with the integration of CORA standard into robotics and automation tasks [104]. There are also efforts to represent a wide range of concepts within the infrastructure aspect, coming from the field of building automation: The Brick scheme integrates building information and management capabilities [9]. These different efforts across the aspects discussed are summarised in Tab. 1. It shows particularly the lack of holistic solutions across aspects as well as automated goal-setting mechanisms, which will be discussed in more detail in section 3.6.

Table 1: *Selection of popular software applications, standardisation efforts, digital twin solutions, methods of workflow integration, and goal setting. These are sorted based on the main aspect of a research laboratory they are targeting.*

	Experiment	**Handler**	**Infrastructure**
Software	LIMS, ELN, CDS, SDMS	LES	BIM, GIS, ERP, LIS
Standardisation	AFO, XDL, …		Brick scheme
Digital twin	*black box (SDL-ML)* *[first principle based]*	*manufacturer specific* *handbooks / documents*	BIM *drawings/documents*
Workflow	SiLA		*platform based*
Goal setting	AHP, DISK	*manual*	*manual*

3 Analysis: Limitations of existing solutions

In the previous section, we have presented the current state of lab automation, highlighting the existing software solutions and approaches for the three crucial aspects of experiments, handlers, and laboratories. However, the current landscape is still fragmented, with limited integration between the different software solutions and challenges related to the automation of complex laboratory workflows. In this section, we will delve deeper into the challenges of current lab automation solutions, focusing on six key areas: interoperability, adaptability, knowledge depth, reasoning, goal derivation, and bridging the "interim technology gap". We will explore each of these challenges in turn and identify potential avenues for overcoming them in the quest for holistic lab automation.

Some of the immediate obstacles were mentioned in section 2 and marked in red in Fig. 3 as follows:

1. Interoperability of hardware, esp. robots
2. Automated execution via ELN-LIMS systems
3. Workflows and provenance tracking
4. Inclusion of infrastructure aspect

Meanwhile, Kitano identified three more abstract challenges in developing a technology platform for an AI scientist [55]: automation, precision and efficiency. While this is a sound analysis, it remains somewhat abstract as these challenges are intertwined and have multiple underlying causes – some of which are related to the idea of a platform *per se*. Looking at the bigger picture and the long-term vision, we identified 7 overarching challenges explained in this section: interoperability, adaptability, knowledge depth, orchestration, reasoning, goal derivation, and bridging the "interim technology gap". These partially correspond to the main factors in the automation of chemistry optimisation as reported by Cronin group [41] but are less specific to optimisation tasks or even chemistry.

3.1 Interoperability: Issues with system integration

Interoperability is a significant challenge in lab automation as it requires seamless integration of equipment, models, and workflows. When mentioning interoperability in this context, classically it refers to challenges associated with equipment of different manufacturers or product lines, interfaces between ELN and LIMS systems or similar software, and data exchange between formats or even research groups. Tackling a different facet in the context of interoperability, the development of SDLs is partially motivated to overcome the siloed nature of conventional research areas, *e.g.*, chemistry and material science [2]. While these challenges are important and considerable progress has been made in solving them, it is still limited to a select number of software and hardware units that are mostly part of the "experiment" aspect in Fig. 6. To achieve full and comprehensive automation of research labs, all aspects need to be connected, broadening the space of possible tasks to automate.

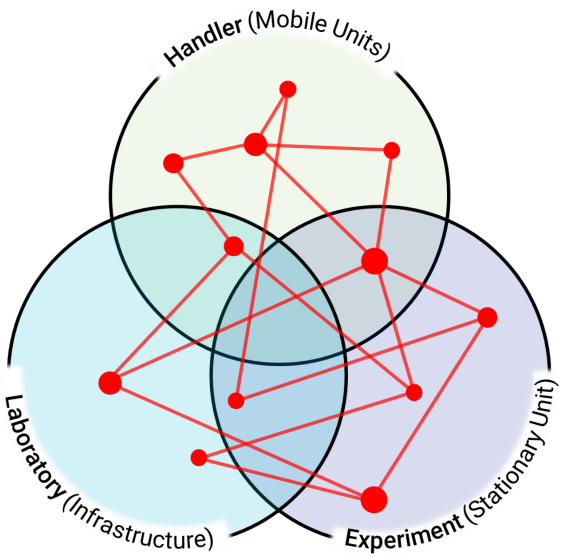


Figure 6: *Illustration of interoperability challenge.*

Full interoperability throughout the system therefore enables automation of peripheral tasks such as resource and space allocation, energy consumption and footprint evaluation, as well as enforcement of safety protocols. This way, not just researchers, but also lab managers, technicians, or facility managers can be potentially assisted or even emulated by a sufficiently advanced AI scientist. This is particularly relevant for tasks that require sufficient knowledge of all system aspects: re-ordering consumables, auditing of assets and inventory, scheduling of predictive maintenance and appointments, managing training schedules and requirements, budgeting for consumables and utilities, and booking of equipment and workspaces.

Overall, the challenge of interoperability is mostly related to cost-benefit or resource allocation tasks. ELN integration can also play a significant role in automating lab processes. While Jablonka et al. [49] identified several key requirements to make data machine-actionable – such as interoperability and an open source infrastructure to avoid "lock-in" effects – they see ELN as the central hub for all chemical research and envision a "platform" solution. Such an approach can result in scalability issues and interoperability problems between such platforms [6] – again posing the danger of some sort of lock-in effect. The same dilemma arises for lab equipment which often comes with easy-to-use, but incompatible tools: while the risk of a lock-in is identified as an issue next to the inherent inflexibility of a single platform [41], platform-based approaches are still the most common due to incentives [19, 70, 71, 100].

3.2 Adaptability: Difficulties in scaling and modifying systems

Dynamic systems that evolve over time are important for lab automation and digital twins because research experiments and infrastructure are inherently dynamic and subject to change. In order to create an efficient and effective automated lab, it is necessary to have systems that can adapt to changes in experimental design, environmental conditions, and equipment availability. Additionally, as new research questions arise, it may be necessary to modify existing experiments or set up new ones, which requires a flexible and adaptable system [67]. Adaptability has therefore been pointed out as a key challenge before [74].

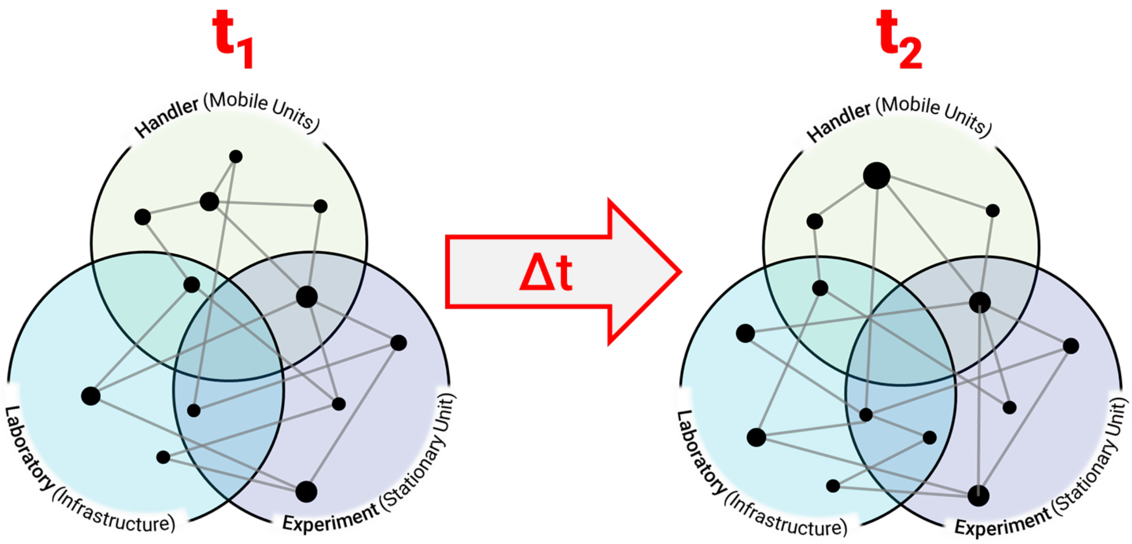


Figure 7: *Illustration of adaptability challenge.*

The problem with many existing lab automation solutions is that they are static and do not allow for easy modification or expansion. For example, software that is designed to automate a specific set of experiments may not be able to accommodate changes to the experimental design or new equipment that is brought into the lab – which touches on the challenge of interoperability discussed in section 3.1. This can lead to inefficiencies and the need for manual intervention, which defeats the purpose of automation. The biggest problem therefore remains adapting generic strategies and reusable software components [74]. A traditional database management system (DBMS) to maintain data integrity takes a lot of effort to contain, is therefore only suitable for large and well-established projects, and falls flat when fundamental changes are introduced [28]. As labs become more interconnected and data-driven, it becomes necessary to have systems that can adapt to changes in the data environment and accommodate new types of data, for example by using semantic web technologies to represent relationships between concepts [28].

In order to address these challenges, lab automation solutions need to be designed with flexibility and adaptability in mind. This means creating systems that can learn from data and adjust to changes in the experimental environment, as well as incorporating modular design principles to allow for easy expansion and modification. This would not only allow for the smooth and continuous acquisition of live data but also increase workflow integrability. Additionally, incorporating digital twin technology can help to create a more realistic simulation of the lab environment and allow for testing of new experimental designs or automation processes [63, 108].

3.3 Knowledge depth: Constraints of existing databases

Deep knowledge models are essential for holistic lab automation because they enable the connection of information between different domains and across multiple scales. This is particularly important for achieving interoperability, which was discussed in section 3.1. To achieve lab automation in a holistic sense, it is necessary to integrate information from various sources such as laboratory equipment, research data, and building systems. However, these sources often use different languages and data formats, making it difficult to integrate them into a unified system. Moreover, deep knowledge models are required to represent past and future states of the lab environment. This is important for enabling adaptability, which was discussed in section 3.2.

A deep knowledge model of a lab system should provide a representation of the current state of the system as well as all the possible future states based on different scenarios. Formalised knowledge representations are crucial to develop the required model of understanding as it is responsible for mapping data from input to output of the system – namely from the newly captured data to an updated model in autonomous experimentation [100]. This would enable the system to be adaptable and respond to changing circumstances, such as equipment failures or changes in research goals. For this reason, virtual laboratories are being developed to simulate and predict the physical world as "digital twins" of real-world laboratories. This does not only help to coordinate instrument sharing and maintenance but also predict faults and procurement needs [63]. More widespread adaptation of such operational models has been seen in industry contexts, where the focus lies more on quality control and LIMS integration [66].

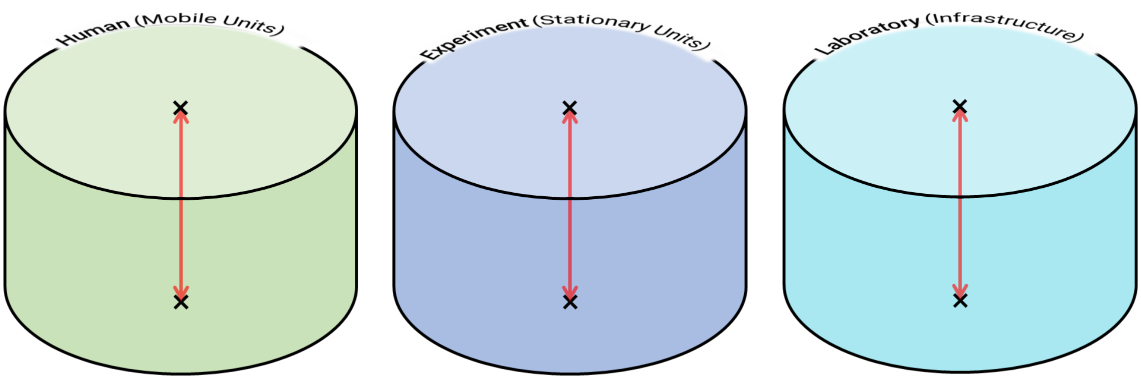


Figure 8: *Illustration of knowledge depth challenge.*

Another crucial aspect related to the challenge of knowledge depth in lab automation is the reporting of data provenance and device information – including metadata such as file formats, conditions, and computations [49]. Ensuring reproducibility and enabling precise post-processing of data requires the creation of actual digital twins of experiments. Such digital twins require deep knowledge of all system aspects as shown in Fig. 8, including metadata. For example, potential sources of error and uncertainty are introduced within all aspects: ambient conditions provided by the laboratory (*e.g.*, dust levels), materials used within the experiment (*e.g.*, vendors or purities), and error rates of specific handlers (*e.g.*, robot precision). Moreover, the sequence of workflows needs to be captured unambiguously [49]. The XDL format for example tackles this problem by providing equipment-independent reaction descriptions [41]. Yet, reporting of results still needs to contain details such as vendors of materials, type of glassware, reactor models, analytical assumptions, and much more to provide a perspective on reproducibility and expected uncertainties. This is particularly important for failed experiments, which usually remain unreported, but can enhance predictive models considerably [12, 85].

To address data provenance, the "event-sourced architecture for materials provenance" (ESAMP) has been introduced. However, the scope of ESAMP is currently limited, and determining which events to capture remains a challenge [74]. Additionally, capturing large amounts of data poses another issue, particularly in terms of data volume and storage requirements. As labs generate vast quantities of data, efficiently managing and storing this data becomes crucial for effective lab automation and analysis. Finding scalable and robust solutions for data storage is essential to support the growing demands of data-driven research [48, 83] and enable integration and retrieval of experiment-related information.

Developing AI scientists or autonomous research systems more generally is heavily referenced in recent publications, see section 1.1. The relevant terms are applied to two different kinds of systems [74]:

1. Robot-assisted / autonomous experimentation (also "Robot scientist" [98, 99])
2. High-throughput computations / materials screening [38]

Apart from the fact that data-driven approaches are applied to both, these system types are currently separated. Even though data science, computation, and high-throughput experimentation are often named as key examples of autonomous experimentation [100], it is hard to couple them intrinsically (on a systems level and not via external human logic imposed) because they exist within vastly different domains and scales. Nonetheless, such coupling would be required to achieve fully integrated autonomous research capabilities [22].

Either way, deep knowledge models are crucial. Deep knowledge models allow us to represent knowledge in a structured way, which enables reasoning and adaptation to new situations. In the field of AI, there is a debate about how to represent knowledge, either as a structured model using semantics or just by feeding enough data to a black box and calling the emerging structure (*e.g.*, weights of neural networks) knowledge. While the latter approach seems to be winning currently, it lacks scientific knowledge and could therefore predict unphysical results – especially beyond the parameter range of training data, see section 2.4. Beyond this, it has three main problems with far-reaching implications.

The first problem is that representing knowledge instead of just data is necessary for reasoning, which is crucial for creating an AI scientist that can learn from data, adapt to new situations, and make decisions. Structured models enable us to encode knowledge about the relationships between different concepts, which makes it possible to perform logical deductions and draw conclusions. This is important in scientific research, where reasoning is often used to make hypotheses or predictions based on existing knowledge (see section 3.5).

The second problem is that alignment (or goal derivation, discussed in section 3.6) is difficult to achieve with a black box. If the knowledge is encoded implicitly in a model's weights, it can be difficult to understand how the model is making decisions or to correct any unconscious biases introduced via the training data [12, 23, 73]. This makes it challenging to ensure that the model is aligned with human values and goals, which is critical in applications where the AI system interacts with humans.

The third problem is that it is unclear if a more general intelligence can emerge from just feeding data to a black box. While these models have achieved impressive results in specific domains, such as playing games or recognising images, they have no representation of the real world or any abstract concepts. This raises the question of whether we can bridge the interim technology gap between current AI capabilities and the long-term vision of holistic lab automation as discussed in detail in section 3.7.

3.4 Orchestration of research campaigns

Current SDLs are somewhat isolated and usually not integrated within a larger system. This means that they will run continuously until some predefined criterion is met (objective function) or human intervention is required (materials run out, devices break, *etc.*). In order to fully automate research laboratories – leveraging increasing autonomy of AI systems – and improve efficiency or sustainability aspects, many experiments within a research laboratory need to be orchestrated: from initial planning to scheduling and execution, management and control of experiments should be based on knowledge about the environment they are in. This not only includes information about shifting research targets, resource limitations, and costs but also the other experiments as well.

As renting or sharing equipment and lab spaces is becoming increasingly important [78], availability of such information to all stakeholders within an organisation can help to reduce downtime and costs. Technological platforms to arrange instrument sharing are therefore being adapted by major organisations [4]. Challenges remain though as databases need to be actively maintained and availability is often not easily predictable [63].

Going further, multiple experiments within the same project or campaign can be orchestrated to work together. This works even across different laboratories. The digital lab framework (DLF) introduced in this paper (see section 5) has been successfully applied to two SDLs in Singapore and Cambridge respectively [6], demonstrating not only remote operation capabilities [2] but forming an effective network of SDLs as envisioned throughout the literature [72, 95]. This is illustrated in Fig. 9 where three formerly independent systems are orchestrated – one of which is shown with multiple handlers within a single experiment and one with multiple experiments within a laboratory.

In the context of the car analogy introduced in section 1.3, this means a smart and interconnected system controlling a majority of vehicles on the road would be able to influence traffic itself. In the context of experimental research, we can go beyond the autonomy of single experiments or tasks to orchestrate whole laboratories and everything going on in them. This would allow us to maximise synergies and optimise intricate objective functions derived from overarching goals (see section 3.6). Ultimately, full interconnectedness will enable the separation of researchers and projects from physical setups, creating

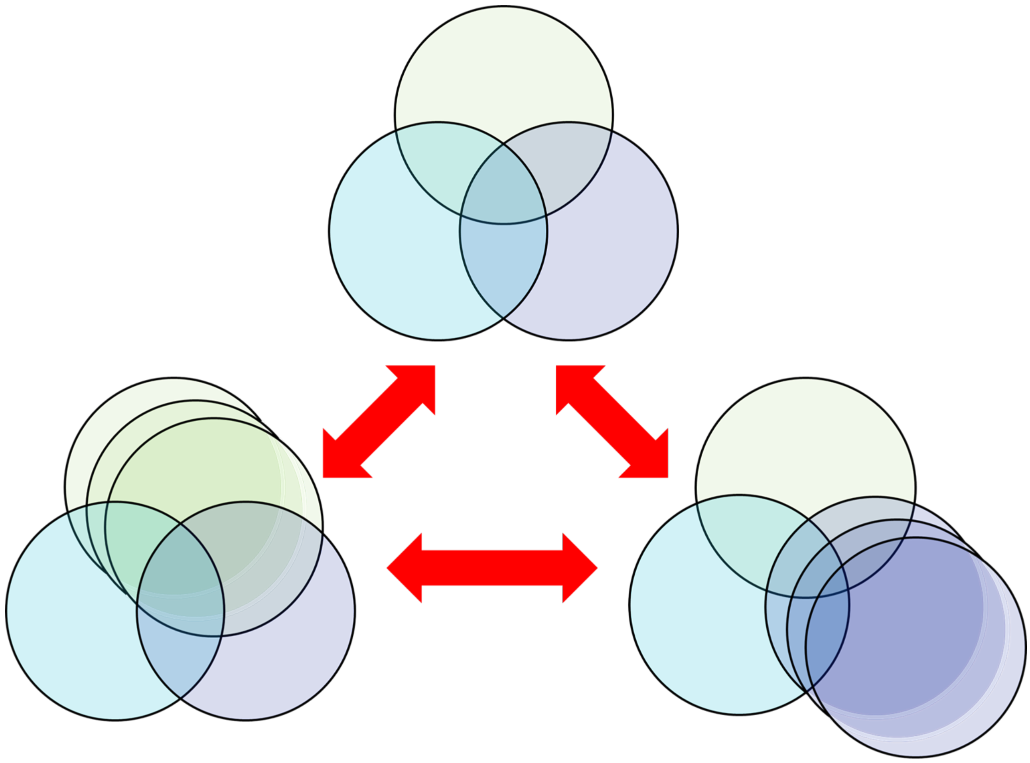


Figure 9: *Illustration of orchestration challenge.*

"meta laboratories" distributed across physical and virtual locations [82]. First steps in this direction have been taken by Roch et al. [90] with the development of "ChemOS" to orchestrate multiple (platform-bound) SDLs.

3.5 Reasoning to close the loop

Another major challenge lies in developing capabilities for reasoning in lab automation. These capabilities are required to infer hypothesised facts from known facts by either deduction, induction, or abduction [99]. This is required to "close the loop" [55] of a fully automated scientific discovery process as shown in Fig. 10, overcoming the human-dependent approach of manual experimentation which is a major driver behind the development of SDLs [2]. These closed cycles or loops are expressed in slightly different ways throughout the literature (see appendix B) but usually boil down to **design**ing an experiment, carrying it out (*e.g.*, **make** a material), **test**ing the results, and **analys**ing the data [6].

The common denominator of these closed-loop experimentation frameworks is the capability to represent hypotheses and then reason on data or models to adapt experiments and hypotheses iteratively, usually using AI-based system [33]. The reasoning involved is thereby usually limited to rudimentary abduction to generate a new hypothesis and deduction to select an appropriate experiment, which can be formalised and automated [54]. If put under scrutiny, many of the available successful systems do not generate and test hypotheses completely autonomously though [19].

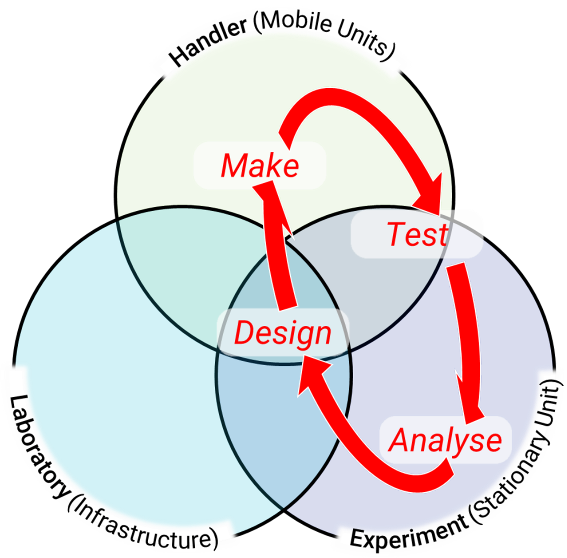


Figure 10: *Illustration of reasoning challenge.*

In current SDLs, this "reasoning" process is usually based on the black box system model generated by some ML algorithm. This is not sufficient to reason on higher levels as required for goal derivation (see section 3.6) as flexibility decreases with the level of autonomy. Therefore, Martin et al. [72] recently suggested to unite classical symbolic reasoning systems with deep learning approaches. Especially in extrapolation tasks, physics-based models outperform statistical black boxes consistently [100]. This again requires accurate and interconnected digital twins with deep knowledge models (see section 3.3).

Reasoning is also important for troubleshooting and awareness of possible impurities or the absence of additives and their implications. For this it would need to access potential influencing factors from all system aspects (see section 3.1). As an example, if an experiment is carried out under an inert atmosphere, a reasoning agent is able to determine if the use of a glovebox vs. a Schlenk line could influence the quality of expected results [23].

The level of autonomy needs some quantification to compare the capabilities of these different systems. Oftentimes, an *acceleration factor* is given to indicate the relative time save in comparison to manual experimentation or automated experimentation without DoE [69, 100]. While this indicator proves practical in the field, it is very specific to the experimental campaign and prone to certain biases which makes it not the most suitable metric for judging technologies more abstractly.

The pyramid shown on the left side of Fig. 11 is often used to illustrate the hierarchy of information processing within the "data, information, knowledge, and wisdom" (DIKW) model, depicting these conceptsas built atop each other. Most of the available lab automation solutions operate on the lower levels of the pyramid, such as data collection and storage, data mining, and optimisation. However, to enable scientific discovery and innovation, lab automation systems must go beyond simply reproducing experiments or finding an optimum within a reaction parameter space.

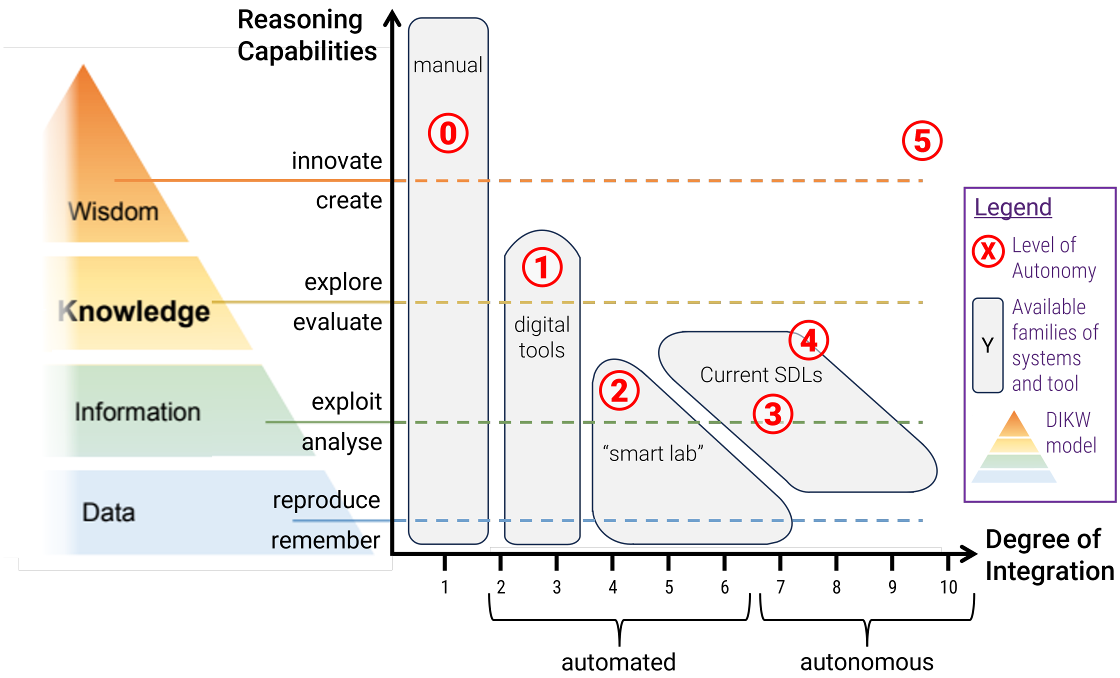


Figure 11: *Reasoning capabilities vs. Integratedness of digital and robotic systems and how these correspond to Levels of Autonomy as introduced by Beal and Rogers [11]. The left side depicts a schematic DIKW pyramid with verbs according to the level of processing above horizontal lines and adapted Bloom's taxonomy below.*

In order to capture the different dimensions of tools' quality, we adapted the pyramid levels to represent potential capabilities, ranging from simple reproduction tasks to exploitation and exploration of experimental parameter spaces to true innovation. These correspond very directly to Bloom's taxonomy of learning objectives which has been used to classify AI systems in the lab context before [23]. The y-axis of Fig. 11 merges these concepts into a qualitative measure of reasoning capabilities. This also corresponds with the versatility of respective systems analogous to human mental flexibility.

The x-axis represents a continuous scale ranging from partially automated to fully autonomous systems. This is supposed to account for the blurry boundary between "automated" and "autonomous" often made throughout the literature [13]. The 10-point scale resembles the "levels of automation" used by Parasuraman et al. [79], yet is supposed to follow a different logic: whereas their model implicitly assumes full capabilities throughout the upper half of the scale which are only limited by permission rights to decide and execute, we rather refer to the level of integratedness within a cyber-physical system. Levels 1 through 7 also correspond loosely to Frohm's "Levels of Automation in Manufacturing" [29] which has been adapted by Holland and Davies [44] to describe the automation level of specific equipment. Going from level 6 to 7, individual machines might now be fully automated but need to be integrated. This is where AE closes the loop and overcomes the "analysis bottleneck" [100].

Fig. 11 attempts to compare the current general approaches to lab automation in terms of the versatility and flexibility they allow for by representing and processing information at a high level at different levels of automation/autonomy. If we consider the 5 levels of autonomy that recently found some adoption in synthetic biology workflows [11, 72], we find that they can be more confidently placed within a multidimensional matrix than along a single axis. As shown in the figure, we see a trend of narrower reasoning capabilities with increasing integration of automated systems into a workflow for the first few levels. Level 3 represents systems with closed-loop experimentation, basically capturing most current SDLs (for which, again, autonomy is ill-defined: while some claim "full autonomy" [50], others are more careful [95]). With level 4, for the first time on this scale autonomy not only increases within a narrowly defined scope, but reasoning on more abstract concepts is required. Based on the exact definition, this still applies to some existing systems [72]. While level 4 broke the observed trend, level 5 seems to represent an even bigger jump if we consider the "machine investigator" [11] as an AI scientist with capabilities as envisioned in section 1.1.

A truly automated lab would therefore be found in the top right corner at LoA 5. This corresponds with the "Nobel Turing Challenge" target zone sketched out by Kitano [55] even though we assume a larger gap between current solutions and the target zone. This can be partially attributed to the different axes used in their figure which introduces a dimension representing task complexity. A similar dimension is used by Parasuraman et al. [79], comparing the level of automation vs. functions derived from human perception. With respect to Fig. 11 these are both considered as part of "degree of integration" in a qualitative sense. We concede that a more-dimensional figure would be more accurate but for simplicity's sake is not given here. All conclusions and observations still stand.

The fact that highly integrated and automated systems are very specific, usually quite isolated, and therefore not flexible, makes it very hard to get to there from here (see section 3.7). This specificity is necessary to boil down complex "thinking" which needs to be done by an AI scientist to single and specific optimisation tasks, which we do not consider to be reasoning in the broader sense. As Kitano suggests, one of the distinguishing features of an AI scientist in comparison to conventional lab automation is its ability to learn, reason and generate hypotheses – this requires not only a high level of autonomous decision-making but also interactions with humans and other parts of the system [55], which refers back to section 3.1. The "increasing scope of applicability" cannot be considered orthogonally as Beal and Rogers [11] suggest because within the current framework, there seems to be a trade-off between breadth and depth.

3.6 Deriving goals and objectives

It can be challenging to determine which experiments to perform to achieve a specific goal or address a particular problem. To derive more specific goals from more general ones, humans typically perform this task, but there is potential to automate it. As the scope of problems and goals becomes more general, the complexity of the derivation process increases, making it more challenging to automate. For example, the derivation process can start with a high-level, abstract world problem, such as the SDGs, and gradually move towards more specific goals and experiments. The process of goal derivation is crucial to enable scientific research to progress, but there is a need to develop automated approaches to address the increasing complexity of this process as we move towards more general and abstract goals. An important part of this process is the derivation of metrics to measure the impact on a specific goal. As an example, there exist 231 unique indicators of SDGs with strong interconnectivity and synergies as well as principles of green chemistry [107].

To tackle a specific goal in the space of sustainability, a single increasingly concise objective would need to be derived in an iterative process, informing experimental design as shown in Fig. 12. Currently, AE will not include this step of iterative goal derivation in their closed-loop cycle as tangible objectives are clearly defined in an "initialisation" stage [100]. This implies that the described process of goal derivation is carried out by a human operator.

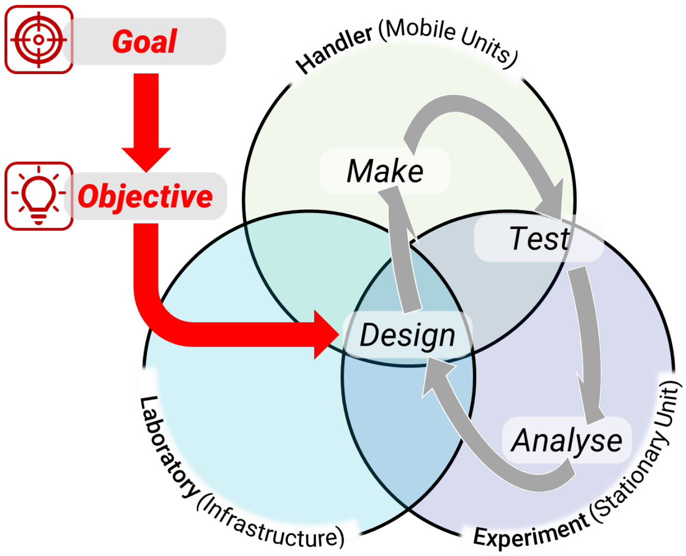


Figure 12: *Illustration of goal derivation challenge.*

To allow for such automated goal derivation, systems would need to accommodate a wider range of definable objective functions for their experiments. This includes data on resource consumption and wider impacts on other systems, which relates back to the challenges of interoperability (see section 3.1) and knowledge depth (see section 3.3). Further derivation then includes planning experiments in a more concrete way for which some methods are available: For example, the Analytic Hierarchy Process (AHP) [77] is a decision-making tool that helps individuals or teams make complex decisions by breaking them down into smaller, more manageable steps. In order to close the loop (see section 3.5), experimental results need to be translated into an updated knowledge model based on which new experiments can be proposed. A similar mechanism is part of the "automated discovery of scientific knowledge" (DISK) framework [34], which focuses on automatically discovering scientific knowledge from large amounts of scientific literature and using that knowledge to propose hypotheses and generate experiments.

Design of experiments is a crucial part of the scientific method, and it is becoming increasingly important in the age of automation and machine learning. Setting new goals is equivalent to DoE at lower levels, which involves choosing a set of input parameters to explore and exploit the parameter space to achieve the best possible output [106]. The choice of input parameters is usually based on a higher-level goal, which needs to be adjusted after the goal is met. As the complexity of the experiment increases, there are several challenges (listed as “DoE agents” in) that need to be addressed, see Fig. 13.

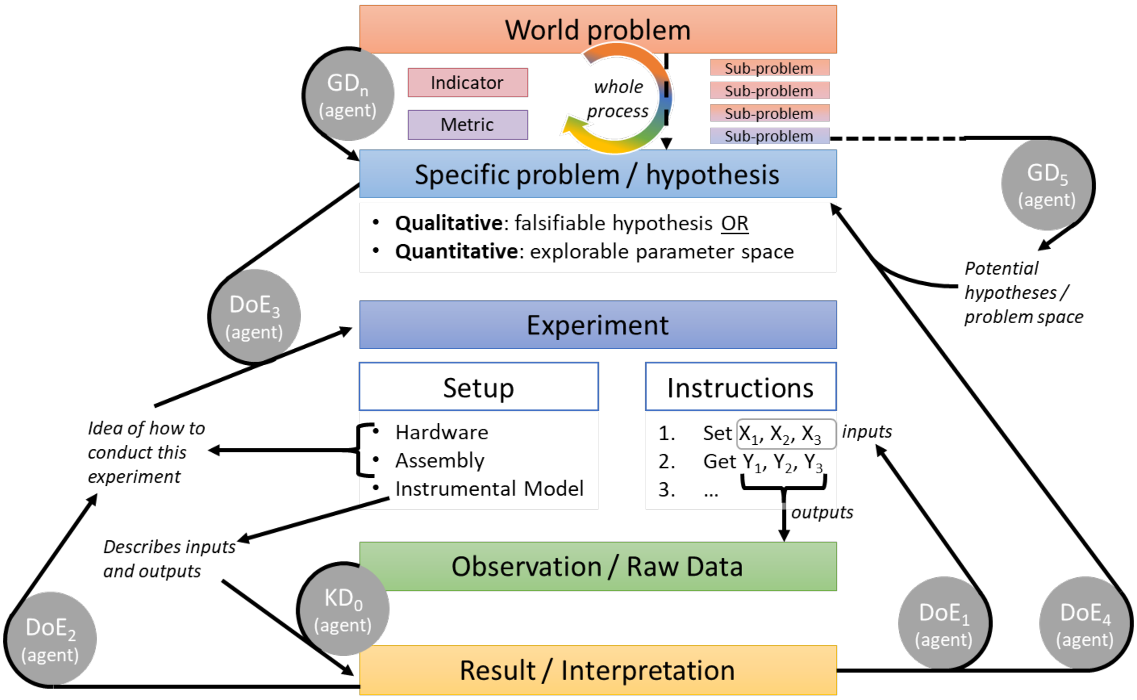


Figure 13: *Goal derivation (GD), Design of Experiment (DoE) and Knowledge Discovery (KD) process cycles in scientific research from abstract goals to specific experimental outcomes.*

Agents shown in Fig. 13 are rated by their level of required goal derivation. The process of knowledge discovery discussed above is technically not goal derivation but still related to the reasoning process where encoded domain and process knowledge need to be utilised [33] for calculating meaningful results based on measurement data and research goals. Then follows the derivation of new goals based on the updated knowledge model.

Firstly, there is a need to choose new input parameters based on outputs of the parameter space as shown in our previous work [6]. Depending on the algorithm used, the next experiment will aim to incrementally improve the previous one or look for a different optimum elsewhere – known as the exploration-exploitation dilemma [100]. It requires an understanding of the existing data and the ability to optimise the experimental parameters for the desired outcome. This most common form of automated experimental design is referred to as DoE$_1$ agent here, which can be used to optimise for a closed-loop objective defined by a human operator [2].

Secondly, changes in experimental setup may need to be made based on the output, to better achieve the desired problem solution or hypothesis falsification, or to circumvent problems that emerged. Constructing an experiment involves making choices about hardware, assembly, and instruction set based on the problem to solve (i.e., the data to gather or the hypothesis to falsify). These choices can have a significant impact on the outcome of the experiment, and so they must be carefully considered. Such a DoE_2 agent has been integrated with some lab automation projects, including our latest work [8].

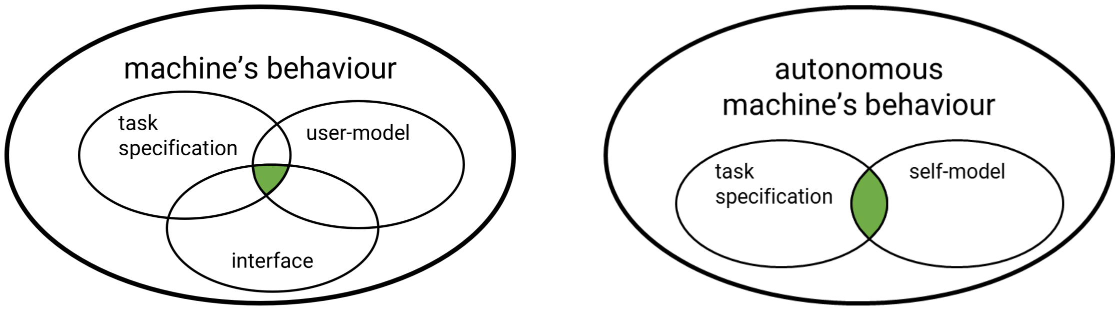
Now, after finishing one experiment (based on some criteria defined as part of the instruction set), the problem might not be fully solved which requires the researcher to construct a new experiment which aims at gathering data or falsifying the hypothesis. Specific choices on hardware, assembly, and instruction set need to be made. To the best of our knowledge, no autonomous system we would describe as a DoE_3 agent has been reported.

Finally, the specific problem is solved which concludes this experimental campaign. There is now a need to formulate new hypotheses based on the knowledge gained. This is a critical step in the scientific method and requires a deep understanding of the underlying principles, data, and more abstract goals. In terms of Fig. 13, a DoE_4 agent would also be a "real" goal derivation agent that provides a space of possible concrete research problems and hypotheses based on a higher-level goal. At the current state, this is the frontier of automated DoE as only first steps have been made recently [8]. Creating general agents deriving sub-goals from higher-order goals up to the highest level requires further advances in the areas of reasoning (see section 3.5) and "natural language processing" (NLP).

The alignment of goals that the system understands with the goals that we humans communicate becomes increasingly relevant when our input is moving up the hierarchy depicted in Fig. 13. It is crucial to ensure the system operates in accordance with the goals and values of its human designers. The higher-order goals that guide the system's operation can be defined by humans, but the process of breaking them down into specific tasks can be done autonomously by the system's agents. While the first one or two levels discussed above can be done without the implementation of actual knowledge by hard-coding iterations and decision trees by domain experts, the more abstract goals require very deep and interconnected knowledge models as well as robust goal derivation agents that are grounded in human-level or above reasoning and logic. With such a system in place, the need for domain experts to hard-code iterations and decision trees can be minimised, as the autonomous agents can use the knowledge model and goal derivation agents to carry out tasks and adjust experimental parameters based on the desired outcome.

However, it is important to note that the involvement of domain experts may still be necessary in certain cases, such as when there is a need for specialised knowledge or when the system encounters unexpected challenges that require human intervention. Ultimately, the key is to strike a balance between human oversight and autonomous operation to ensure that the system remains aligned with its intended goals and values. This problem is known as "AI Goal Alignment" and is usually part of the discussion around the threats of AGI and corresponding legal measures [47]. Usually, this is not a grave concern in the space of laboratory automation as we do not aim to create artificial general intelligence (AGI) and cause "the singularity" [37]. It is still important to monitor alignment very closely for two reasons: Firstly, economic concerns push us to make sure goals are aligned in a way that derived objective functions do not optimise in the wrong way. Secondly, as these systems still include robotics, ensuring human safety is critical. Some argue that robotic systems need to close perception-action loops to achieve AGI [20].

More abstract goal definitions mean that many existing human-machine interfaces (HMIs) will become unnecessary as they control the machine's immediate tasks which can now be derived by the machine itself. At the same time, the machine is no longer constrained by the model the user has about the machine's behaviour (user model) but by the model it has of itself (self model) which can be more restricted if no detailed digital twin is available (see section 3.3). This self model can be extended though and as there is no interface to consider, the space of possible actions can grow proportionally as shown in Fig. 14.



(a) Elements of human-machine interaction as illustrated by Degani et al. [25].

(b) Adapted Venn diagram of relevant elements for machine-derived tasks.

Figure 14: *Comparison of possible machine's behaviour with and without traditional HMI. The green regions correspond with possible correct interactions.*

3.7 Bridging the interim technology gap

Even though AI research could look very different from human research (no labs, no papers, *etc.*), we first need to find a way to integrate AI (or even just digital) systems with current systems. This is analogous to autonomous driving which would be easier with only self-driving cars on the road but this option is unfeasible, so human-driven cars need to be considered and integrated into model and decision-making. As Holland and Davies [44] argue, not all processes have interim labour-saving technologies between manual execution and fully autonomous systems. Meanwhile, many researchers in the field envision a "tipping point" [100] at which globally integrated systems emerge from continued development and deployment of localised SDLs leads.

As illustrated in Fig. 11, the AI scientist is not a logical continuation of current SDL development. This discontinuity is depicted in Fig. 15 and is partially caused by the tendency to automate systems and processes for the sake of automation. Although there are very good reasons for automation to begin with, it does not always continue to make sense – at least not before digitising peripheral elements such as inventory, space constraints, *etc.* as this prioritisation when considering priorities derived from overarching goals and objectives. We cannot go from a fragmented and partially-analog landscape of tools to an AI scientist by adding more and more robots within a narrow closed-loop experiment. This way we risk lock-in effects with regards to technology (potentially even manufacturers, see section 3.1) but also applicability (compare with Fig. 11). Tools should therefore become independent of the specific instrument or technique used [49].

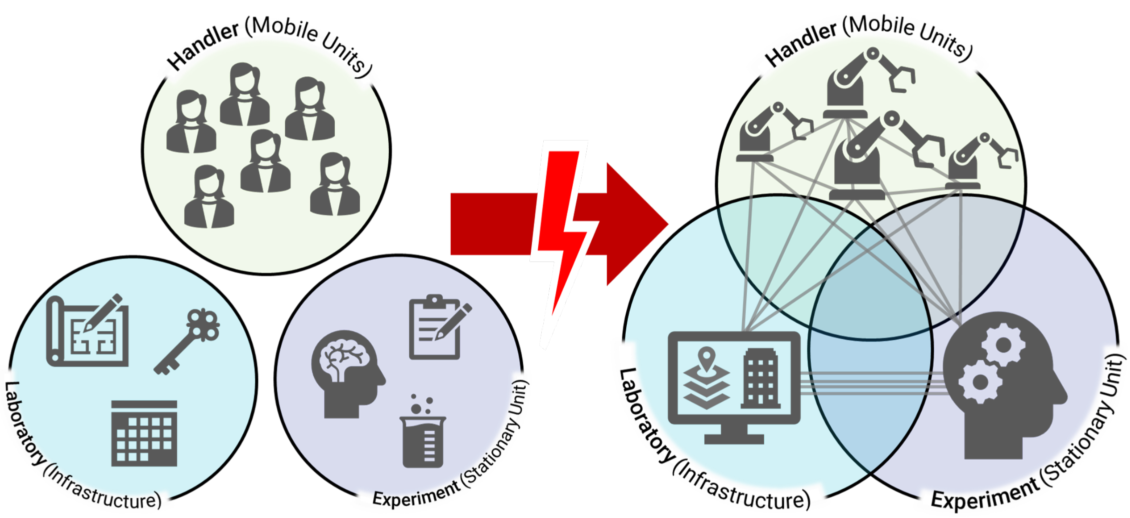


Figure 15: *Illustration of interim technology gap challenge.*

Only David et al. [23] sketch out a different part, starting with accessible and searchable data before constructing virtual laboratory assistants, digital twins, and finally autonomous laboratories. Even though they remain vague regarding implementation, at least they consider a route of intermediate steps. The same applies to building knowledge bases and sharing data: while data of past research is often hard to digitise, a lot of current data is distributed across formats and platforms, It is foreseeable that going forward, interchangeable standards will be used and data automatically generated or even reported according to FAIR principles, some even call for mandates to accelerate this transition [73]. Yet, we have to come up with systems now that can accommodate fragmented or incomplete old data as well as rich datasets with detailed metadata in the future.

Reporting needs to be done in a way that robots as well as humans can execute the instructions and reproduce results. Currently, the only effort to address this specifically is the XDL format [41, 49]. Human intervention is still necessary in many ways, including all aspects of lab operations: robots need to be set up, materials purchased, objective functions clearly defined equipment serviced, errors and uncertainties accounted for, environmental conditions recorded and maintained and many other tasks related to management and maintenance carried out. These human tasks can bottleneck operations [69] as they need to intake observations for interpretation and analysis, then decide on actions to modify, maintain, or configure – all of which can be very time-consuming tasks. MacLeod et al. [69] therefore conclude that "[. . .] time spent on SDL to reliably perform an experiment can far exceed the time saved". This is due to the nature of conventional SDLs in which all actions within a narrowly defined experimental system are automated, while peripheral data and activities – including managerial tasks such as resource allocation – are very much decoupled from actual design and execution of experiments, residing in data silos of often proprietary software if digitised at all.

To bridge the described gap in a practical manner we need to keep the human in the loop: Montoya et al. [74] suggests that a human-in-the-loop is needed for oversight with increasing complexity and to apply knowledge not yet encoded. Beyond that, the concept of a "human-in-the-loop"is only considered within ontological work – *e.g.*, the process chemistry ontology (PROCO) includes an abstract concept of planning process in which humans and robots are mentioned explicitly. This is also part of EXACT for experimental actions by humans and robots used in the biotech space [99]. Moreover, the SOHO ontology for human-robot collaboration exists within a context of mechanical tasks [104].

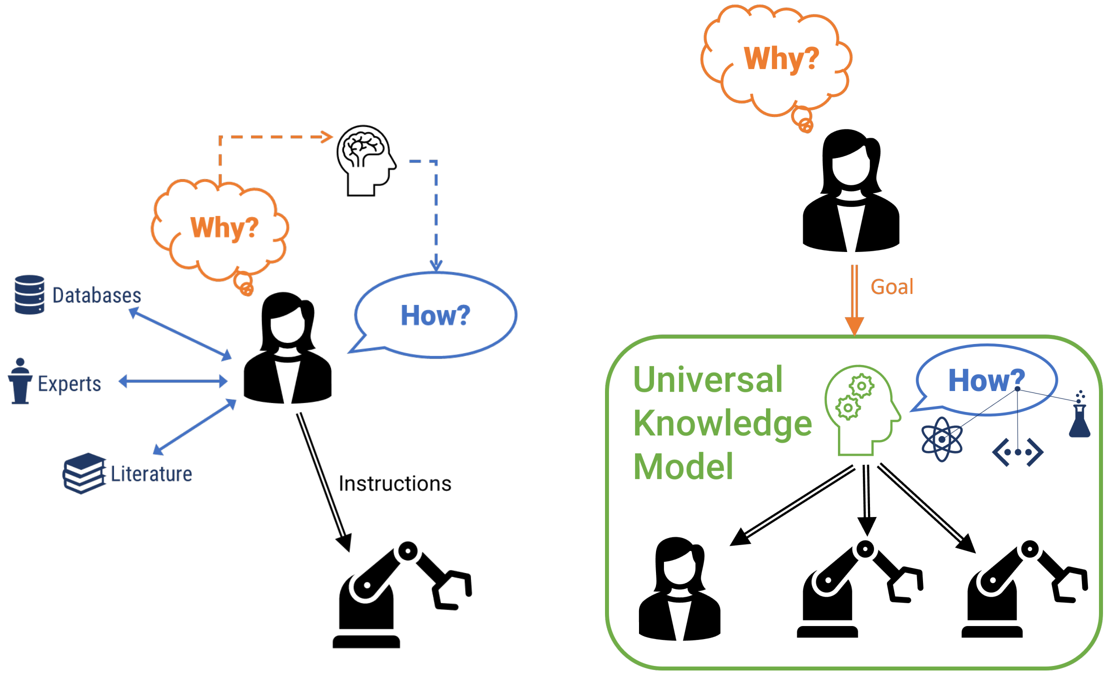
Lastly, there is another disconnect between efforts towards SDLs and the general digitisation of laboratories, often under the IoT tag. While the first is more focused on automating research questions, the second is more focused on smooth operation in production contexts. This becomes apparent at the interface: pilot plants, where scale-up efforts span areas of DoE, computational modelling, as well as design and implementation of control systems. Jones et al. [51] have pointed this out and suggested retrofitting existing equipment via an "Industry 4.0" framework to create digital twins of operational units.

4 A change in perspective

We argue that in order to overcome all the challenges discussed in section 3, create AI scientists that could win the Nobel-Turing challenge, and fully automate research laboratories in their entirety, a change in perspective is needed! Approaching this emerging topic more from an engineering background rather than a fundamental science one, we try to offer a fundamentally different viewpoint. Engineers analyse and develop systems to solve a whole group of problems rather than design of specific solution (*e.g.*, automation of single tasks) [43, 109]. This includes the consideration of research activities across disciplines where experimental setups might be fundamentally different to chemistry.

In the field of biotechnology, for example, systems are more complex and some processes can only be understood from a system dynamics perspective [55] which is formally introduced in section 4.2. At the same time, they provide a unique opportunity as the underlying physical and chemical processes happen within a more narrowly defined and often times accessible parameter space [72]. This led to an accelerated development of approaches to metadata collection and storage in this space. Furthermore, the implicit and explicit inputs of Design-Build-Test-Learn cycles have been analysed and emphasised [11]. This includes protocols and knowledge which need to be provided by digital twins and the goals formulated by humans from which sub-goals need to be derived.

In order to create a framework in which the process of experimental research as a system can be automated, these approaches to the representation of goals, metadata, and processes need to be pursued further. Creating an AI scientist that can plan and carry out all types of experiments beyond simple synthesis or yield optimisation while considering its ever-changing three-dimensional physical environment, requires a shift away from human experts towards a universal knowledge model. This paradigm shift in representing knowledge – including that which is considered tacit or intuition – is illustrated in Fig. 16. It requires a revision of our understanding of a human-in-the-loop which conventionally refers to a researcher that observes the SDL, sets objectives, intervenes and provides domain knowledge [72, 76] as shown in Fig. 16(a). In contrast, Fig. 16(b) depicts the envisioned state of a universal knowledge model that can derive the optimal immediate tasks from abstract goals defined by a human outside the system as well as its knowledge of the world.



(a) Conventional approach. (b) Paradigm shift.

Figure 16: *Human interaction with automated laboratory systems and underlying goals.*

Working towards this vision entails three changes in perspective that are discussed in detail within this section:

1. Automation beyond platform-based SDLs: comprehensive digital twins enable holistic lab automation, see section 4.1.

2. Taking on a systems view instead of defining tasks: a goal-driven approach ensures alignment, see section 4.2.

3. Instead of siloed information and rectangular databases: actual knowledge models are required, see section 4.3.

4.1 From self-driving labs to comprehensive digital twins

As laid out in section 3.7, laying the groundwork for the creation of AI scientists requires broadening the search/optimisation space of SDLs and removing the "predominantly human activities [which are] principal bottlenecks in scientific progress" [33]. To achieve this, all-encompassing digital twins are necessary that also enable researchers to build large-scale precision models (as part of "getting stuff out of the lab" [78]) that include almost every interaction and molecular behaviour as demanded by Kitano [55]. This way, the full pipeline of multi-scale modelling and experimental verification or adjustment can be represented and automated for an iterative improvement of such high-precision models. We argue this for electrocatalytical reduction of CO_2 in a previous publication [89].

In 2022, Zhu et al. [111] reported the first full "AI chemist" that can capture existing knowledge in the literature in an automated manner by reading articles using NLP. Based on this, it can propose experiments and even conduct them autonomously in the laboratory with a mobile robotic system. Two years prior, Hessam et al. [42] reported a similar system specific to chemical synthesis which would read basic literature, translate it into process steps, and recreate these experiments. These albeit impressive efforts remain highly localised and limited by their monolithic nature in the physical dimension (a single robot with certain capabilities in a non-smart lab) as well as the digital space (black box knowledge model with a certain amount of training data).

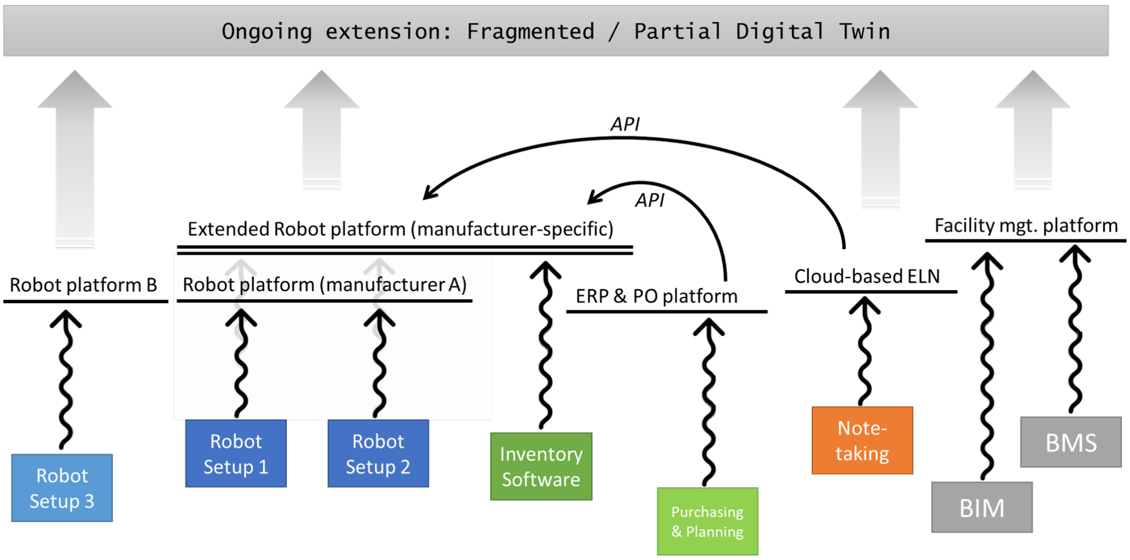
As analysed in some detail, we need digital twins that not only capture behavioural aspects within the default operation range but are able to confidently predict behaviour within novel contexts, including the estimation of possible unknowns leading to errors and uncertainties. This can only be done by integrating different domain models, metadata, and physical shapes while embedding fundamental principles as boundary conditions. A digital twin like this cannot stand alone but has to interact with digital twins on different scales and continuously adapt itself to account for changes in the real world.

This holistic approach to digital twins has gained some attention in recent times, for example as Siemens has introduced their framework "comprehensive digital twins" [39, 61]. They emphasise their approach of model-based systems engineering requiring traceability which is facilitated by their digital twins [61]. This incorporates many of the requirements discussed as challenges in section 3 as they explain their concept as "a precise virtual representation of a physical object, including its mechanical, electrical, and configuration management. This one digital twin evolves across its lifecycle with numerous models used to capture different aspects of the object's physical behavior [...]" [39]. The idea of the Digital Lab Framework is to implement this type of comprehensive digital twin within the space of research laboratories.

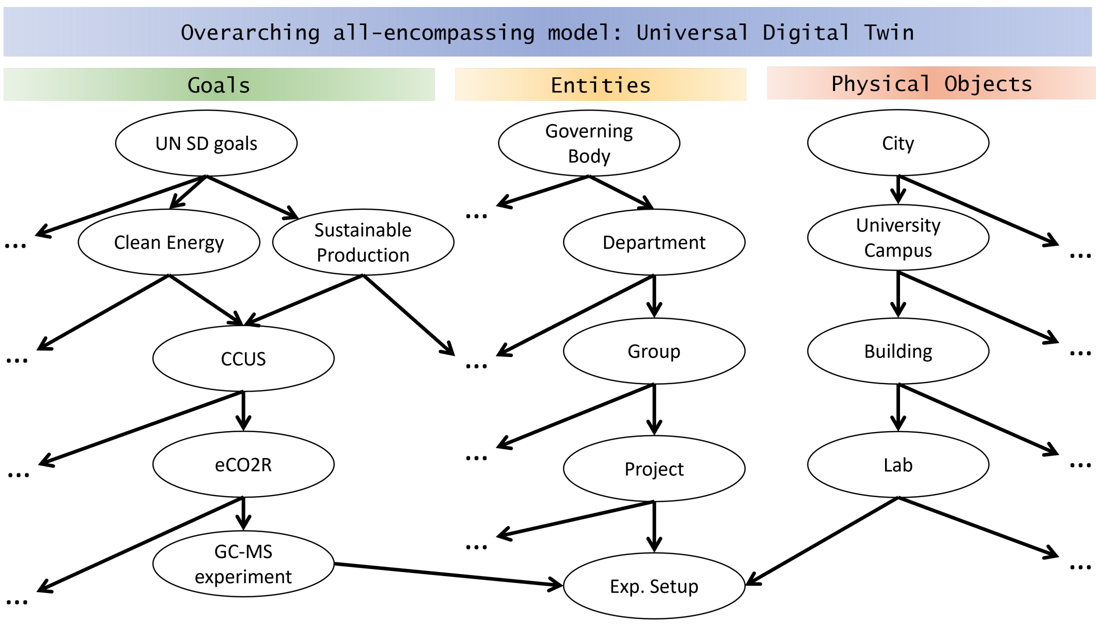
4.2 From task-driven to goal-driven systems

In order to account for the interdisciplinary nature of challenges faced as laid out in section 3, a systemic view should be adopted to help design and integrate solutions across different domains. This systems engineering or systems thinking view has received little attention in the field of laboratory automation so far [108] but allows us – amongst others – to battle island solutions and "manufacturer lock-in" while enabling more efficient resource allocation. Such a perspective shift also opens up the opportunity to establish frameworks to check what should be automated and to what degree [79], coming back to the challenge identified in section 3.7 to not automate for the sake of automation. We, therefore, argue for a goal-driven approach rather than the conventional platform-based approach as shown in Fig. 17.

As illustrated in Fig. 17(a), the conventional platform-based approach aims to solve a given task at hand and then integrate his solution with an existing platform to achieve more universal applicability [5, 27, 62]. This iterative platform extension can only improve the system in a linear fashion at best and will – as pointed out in sections 3.5 and 3.7 – probably not lead to a general system capable of performing tasks it has not



(a) Task-driven platform-centric approach.



(b) Goal-driven holistic approach.

Figure 17: *Comparison of task-driven and goal-driven approaches as strategies to create a digital twin of research laboratories.*

been programmed to do. In contrast, Fig. 17(b) depicts a goal-driven approach of defining overarching goals and concepts that are broken down into more and more granular tasks and objects within an ultimately all-encompassing world model.

Close comparison of the two processes in Fig. 17 reveals a fundamental difference in directionality. This reflects the contrast between the inherently limited "bottom-up" thinking behind the task-driven approach of adding more and more specific capabilities and the “top-down” thinking within the goal-driven approach when deciding whether to execute an experiment and what conditions to use, a technique known as “backcasting” [45, 105] or “inverse design” [92, 108] . Such a top-down thinking enables goal derivation and DoE up to a very high level as analysed in section 3.6.

Top-down vs. bottom-up in lab automation has so far been discussed only in a much more narrow sense [44, 75]. The closest was Kitano about decision-making in autonomous agents [55] but we are talking about a way of thinking and decision-making, not implementing agents or designing ontologies. We therefore refer to these different approaches as goal-driven or holistic vs. task-driven or platform-centric as shown in Fig. 17. Ultimately, a flexible and adaptive lab automation system needs to be able to explore and innovate within the limits of existing scientific knowledge and broadly defined goals while also expanding that knowledge through experimentation. This shows that a common underlying architecture rather than a common platform is required!

To this end, we showed in a recent perspective [6] how next-generation SDLs will evolve beyond single platforms with *ad hoc* data representations towards semantic knowledge representations (*e.g.*, knowledge graphs as discussed in section 4.3). This represents the basis of the DLF introduced in section 5 and the related work in progress on a general goal derivation framework [7, 8]. This goal derivation framework differs from the DISK and AHP approaches in several ways. While DISK uses literature data only [34], our proposed framework incorporates both existing scientific knowledge and new experimental data to derive goals for future experiments. It is similar to AHP in that it also breaks down larger goals into smaller tasks [77], but it differs in that it does so autonomously, without requiring human input for each step (see section 5.3).

4.3 From data science to knowledge engineering

The Semantic Web provides a description logic to represent and link data on the World Wide Web [16] within predefined ontologies, describing classes, object properties and relationships within a certain domain. Ontological objects are instantiated as subject-predicate-object triples that attach data properties and use linked data [14] to encode further information. The key strength of semantic data representation is intelligibility as it allows for actual knowledge accumulation and reasoning within domain models instead of purely data-driven ”black box” models, enabling the often-requested step towards physics-based models [100]. Semantic web technologies can furthermore augment the power of AI search [33] and help to manage large workflows, and data sets, as well as their provenance, permitting systematic tracking and propagation of metadata constraints [53].

A knowledge graph (KG) is a collection of semantic data representations expressed as a directed graph, where nodes denote concepts or instances and edges denote links between related concepts or instances. Unique “internationalised resource identifiers” (IRIs) ensure an unambiguous representation of these concepts and instances. Knowledge graphs allow for a distributed architecture that can be accessed from anywhere and queried for data across multiple connected domains.

The World Avatar (TWA) has been developed within Cambridge CARES as a continuously extendable platform for representing information in a dynamic knowledge graph by integrating real-time data, knowledge, models, and tools mainly related to the process industry. It is basically an attempt to construct the “Giant Global Graph” introduced by Berners-Lee [15]. As illustrated in Fig. 18, two main concept types coexist within the same ecosystem and are both described by modular ontologies: instances representing interlinked data, and agents as autonomous software applications that can act on the data and exchange information via the KG [26]. A knowledge graph can also incorporate time series data either by referencing a local database via IRIs (offline solution) or adding a context element within a so-called “quad store” (online solution) [3]. This capability is essential for progress in holistic lab automation as there currently exist no databases with the ability to store run-time context data collected by related sensors [41].

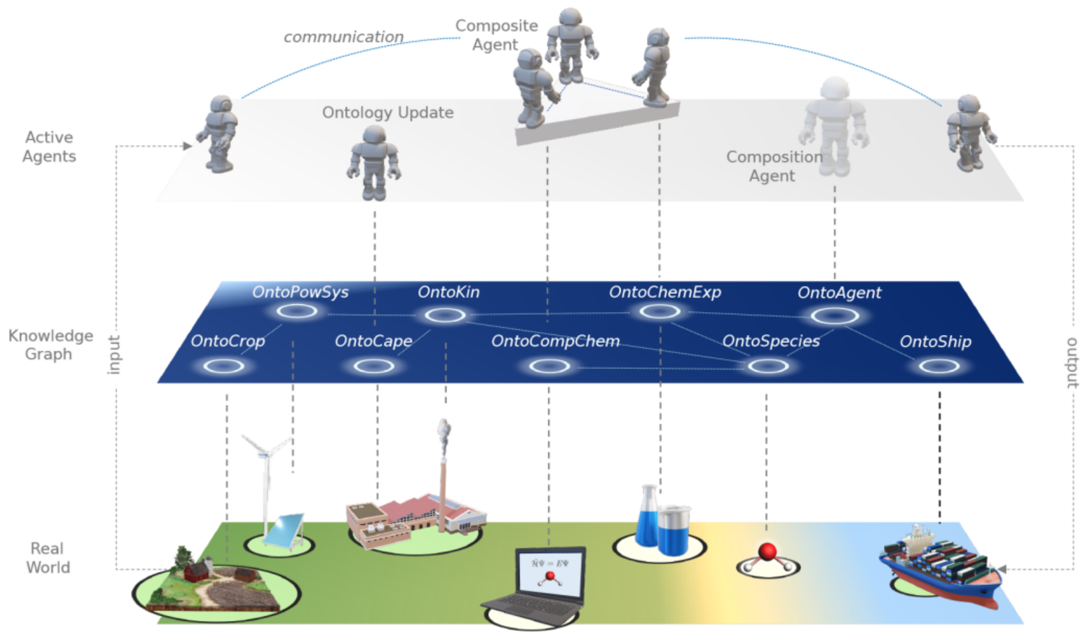


Figure 18: *Three layers of TWA (`www.theworldavatar.com`) digital twin of the real world as shown by Kondinski et al. [57]*

The use of ontologies in this space is important to “make scientific knowledge more explicit, help detect errors, enable sharing and reuse of common knowledge, remove redundancies [in domain-specific ontologies], promote interchange and reliability of experimental methods and conclusions” [98]. It also allows for deep knowledge representation (see section 3.3), linking objects instead of literals and using deep knowledge graphs. Knowledge graph technology slowly finds its way into industrial applications within the broader space of chemistry and biotechnology (*e.g.*, AstraZeneca [93]). While semantic web technologies are somewhat established now for the planning and dissemination phase of scientific experimentation/research, their full capabilities will only be unlocked by connecting processes throughout all phases [28], including execution and goal derivation. First progress has been made in formalising the scientific reasoning process in a more general way [36], potentially unlocking higher levels of DoE agents as discussed in section 3.6. We consider this necessary infrastructure to unleash the envisioned accelaration of scientific research, resembling “Moore’s law” [22, 90].

A lot of work has been done in ontology development across different domains and scales but due to ongoing research and discussions around the correct structures and paradigms, far-reaching technology roll-outs have been rare. We think that it is important to be very practical about this and follow the principles of agile project management to create solutions as quickly as possible and iterate on them. For this reason, we make use of existing domain ontologies whenever possible and do not get caught up in discussions about paradigms or usage of top-level ontologies like SUMO. Therefore, the creation and use of ontologies does not follow a top-down approach, which is limited to selection and derivation of goals as explained in section 4.2.

As more and more standards for specific problems exist, it becomes more important to make existing standards interoperable than creating new ones, as we can assume that the technology to solve most problems exists but needs to be connected [49]. This creation of frameworks is essential for the holistic, goal-driven laboratory automation we have been arguing for in this article. Next to the framework introduced in the coming section, there have been a few efforts in this direction, *e.g.*, the Open Semantic Lab [102] which is currently focused on battery-related research.

5 The Digital Lab Framework

Based on the systems perspective discussed in section 4 we introduce a Digital Lab Framework (DLF) to fully digitise and automate all aspects of research laboratories as part of TWA. The DLF aims to fully incorporate equally the three aspects initially identified by applying a systems perspective in section 1.2: the laboratory (including environmental conditions, building aspects, and available resources) as infrastructure, the experiment (including setup, chemicals, and related knowledge) as stationary unit, and the handler (restricted to tasks that require physical presence) as mobile unit. These three components can best be shown in a Venn diagram such as Fig. 19 (in contrast to Fig. 3). In order to achieve full holistic laboratory automation as envisioned in section 1.1, these aspects cannot be treated in isolation.

The same three aspects can also be derived via overarching goals: The fundamental goal of a research lab is to expand the boundaries of knowledge. As physical setups and actions are required for this, additional boundary conditions and therefore derived goals are imposed. The laboratory is supposed to run as cost-efficient as possible while keeping humans (as well as equipment) safe and happy. Our hypothesis is that currently, automation is mostly done within the three components separately and not at the interfaces. But exactly this is needed to fully automate by combining capabilities of SDLs, smart labs, and DTs as discussed in section 3.7. Exemplary use cases are shown in section 6.

To summarise, within the DLF we now aim to create comprehensive digital twins (see section 4.1) to automate all parts of the scientific research process, driven by overarching goals from which individual tasks are derived (see section 4.2), using dynamic knowledge graph technology developed within TWA project (see section 4.3). When it comes to the implementation of the DLF, we are guided by three key ideas: First, we want to take a holistic view and thereby not only represent and automate narrowly defined but broad systems that can be well-integrated with their environment on all levels. Secondly, we

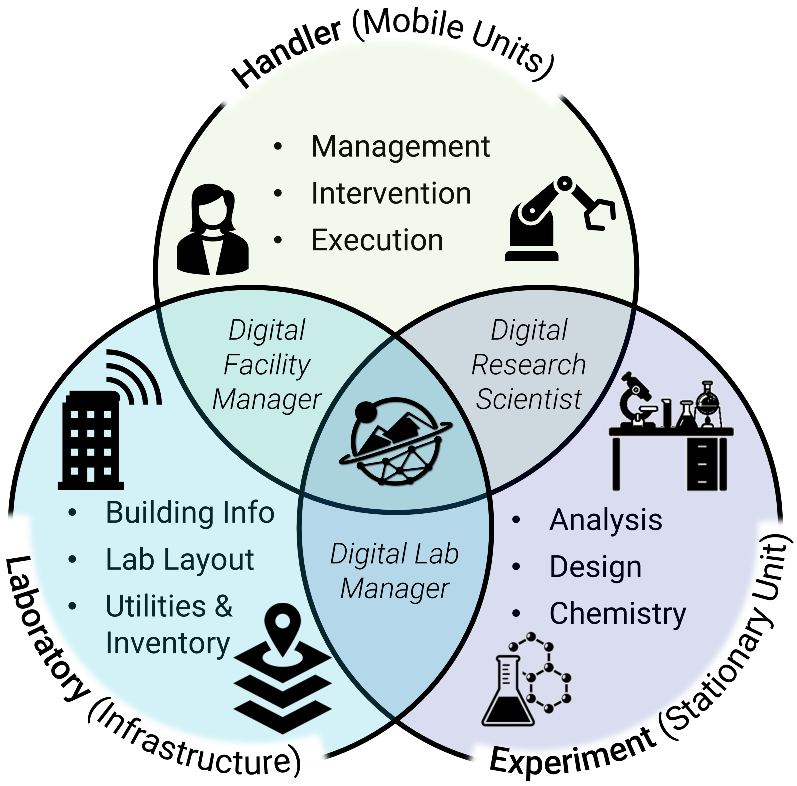


Figure 19: *Lab Automation Framework Venn diagram.*

want to create interconnected digital twins that access distributed deep knowledge from many sources. Lastly, we want seamless yet clearly defined human-machine interactions; particularly, within closed-loop experimentation.

5.1 Holistic view: Integrating experiment, handler, and laboratory

In order to not only run an efficient but also a sustainable laboratory, a holistic perspective needs to be taken [86]. The DLF aims to provide this mainly in two distinct ways: First, by considering all aspects of the research laboratory system as shown in Fig. 19. This enables wider interoperability and lays the foundation for informed reasoning and decision-making. By widening the system boundaries to include all peripheries, we can for example include economic sustainability metrics into the goal derivation process as discussed first in section 2.3. Secondly, the framework is set up to be applicable to the widest possible range of research areas: this goes beyond the common applications in material science and synthesis and includes areas such as biotechnology, combustion research, flow reactors, *etc*. This even extends to theoretical modelling and simulative efforts, i.e. computer experiments as explained further below.

The DLF does not only consider the narrow definition of lab automation between design, analysis, and execution of predefined experimental spaces. Instead, we look at all related activities and required knowledge and break it down into different parts that can be automated separately. This includes the actual laboratory space including environmental conditions and surrounding buildings as well as related inventory, utilities, and other resources that need to be included in representations of a future-proof laboratory [81]. By providing interoperability (see section 3.1) and enabling knowledge depth (see section 3.3) in this way, we can not only automate managerial tasks and ensure better reproducibility. By widening the search space for optimisation tasks to this extreme degree, we might find for example that automating reordering chemicals is a higher priority than the further implementation of liquid handlers for example.

In order to fully represent a (chemical) research laboratory in all its aspects, different ontologies across scales, domains, and representation levels are necessary as shown in Fig. 20. This entails the usage of existing ontologies such as OntoCityGML to describe building information or OntoCAPE to represent the chemistry domain. On the other hand, it is the creation of completely new ontologies to represent experimental setups (OntoLab) such as reactors and processes taking place in them (OntoReaction). Combining these concepts will allow us to implement and connect existing use cases of 3D visualisation, Data Analysis of experimental setups, and chemical models based on first principles.

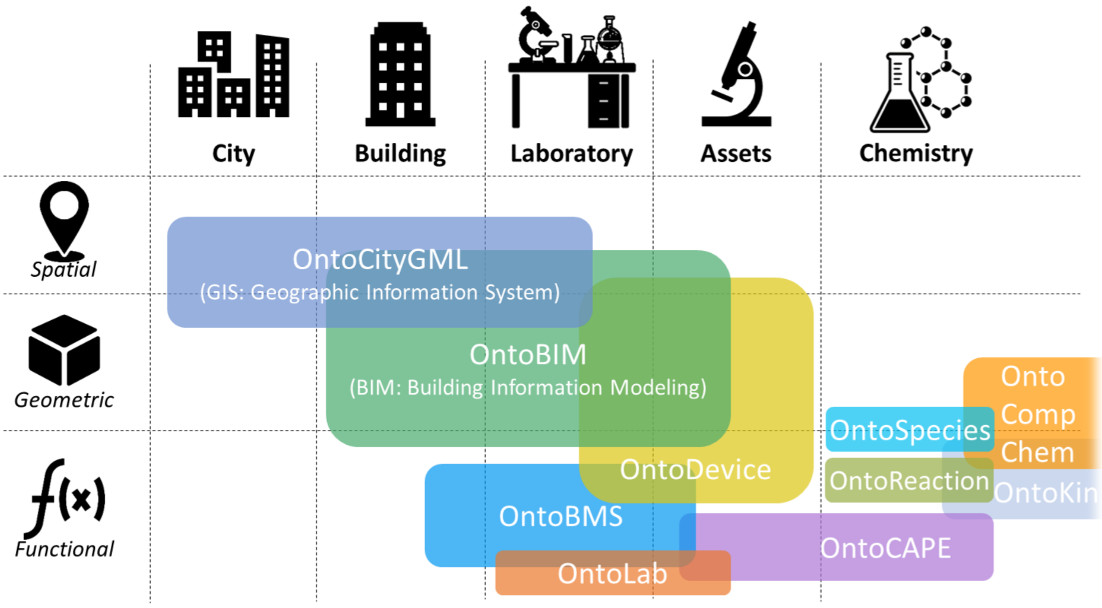


Figure 20: *Representation of concepts across different scales and levels via ontologies in TWA to facilitate the DLF.*

To give a few examples, we integrated BIM and GIS representations in TWA [84], created OntoLab as a base ontology to represent experimental setups for automated execution [8], and established OntoSpecies as a core ontology which is underlying chemical models and materials throughout the system [80]. The use of ontologies in this way removes ambiguities and enables connections across completely different levels of scale and function, *e.g.*, linking a measurement to the city it has been taken in [49]. Going forward, more use cases can be developed easily based on the existing tech stack, filling in the gaps between scales to create a truly interoperable digital scientist. The next steps in this endeavour are laid out in section 6.

5.2 Distributed and connected digital twins

The connection of scale levels in Fig. 20 is necessary to couple the two most prominent instances of research automation – computational simulations (and data screening) on an atomistic level and the experiments carried out at lab scale. These two instances can be approached with the same methodology [100] as the former can be advantageously framed as computer experiments [30]: there is no need to differentiate between measurements or theory as data sources on a fundamental level as they are all based on experiments (physical or computational) and are based on a certain line of reasoning accessible by logic. The lines become blurry in any case: on a fundamental level, data analysis processes resemble multi-step computational methods – including the specific implicit assumptions necessary. These processes can all be captured in the same manner as scientific workflows [32, 53].

The systemic view shown in Fig. 19 can be applied to computational experiments as well – the computer hardware and operating environment (*e.g.*, Linux enterprise system, central and graphics processing units, . . .) functions as the laboratory to conduct the experiment. As the experiment itself we can view the model and simulation details (e.g., quantum-mechanical calculations based on "density functional theory" of carbon monoxide on copper surfaces, . . .). The handler is represented by the actual solver or hardware that is being used (*e.g.*, Vienna *ab initio* simulation package VASP [40]). Even though there are a lot of similarities, the aspects infrastructure and handler seem less relevant as computations are deemed more deterministic. Yet, in the field of computational modelling it is well known that reporting on the computational details and solver parameters is critical. Also when it comes to managing large distributed computations and resources [53] (*e.g.*, planning, scheduling and managing jobs on an high-performance computing cluster), the systemic view becomes quite relevant. Moreover, computer experiments follow a similar cycle that needs to be completely represented for closed-loop automation: Decide – Experiment – Hypothesise – Predict [74] (compare to appendix B).

Apart from increasing practicality of computational studies by including synthesisability aspects [74], these two can now interact beneficial in both directions: Models can inform the experiment by providing material properties needed or even suggesting candidates for condition parameters or materials [100] as discussed in section 3.3. At the same time, experimental measurements can inform the computational model, for example by highlighting deviations from idealised behaviour or parameter spaces that need to be explored further.

In a seemingly contradictory way, we want these digital twins not only to be connected but distributed. This distributedness does not infer isolation though and rather supports the connectedness of concepts and digital twins. There are two main aspects to this distributedness: First of all, data itself is distributed within our approach to dynamic knowledge graphs – namely The World Avatar as introduced in section 4.3. It is inherent to the KG technology and can help solve data volume problems as discussed in section 3.3. Secondly, the experiments or laboratories themselves can be physically separated and even distributed around the globe, enabling complex orchestration as discussed in section 3.4. Distributing experimental research like this in a similar way to distributed computing on the modelling side is needed to accelerate R&D efforts and has been demonstrated as effective in a recent work of present authors [8].

5.3 Human-machine interactions: Inside and outside the loop

As discussed in section 3.6, seamless interfaces between humans and machines are necessary to ensure goal alignment. Moreover, humans will need to stay part of a closed-loop experimentation cycle for the foreseeable future to narrow the interim technology gap (see section 3.7). The phrase human-in-the-loop therefore traditionally refers to humans carrying out two potentially intertwined but fundamentally different tasks [76]:

1. Decision-making, trouble-shooting, result-checking. These tasks mostly arise due to complexities not captured by machines [76]. At high speeds, human intervention becomes effectively impossible [54], causing bottlenecks discussed in section 3.7.

2. Maintenance, manual work, *etc.* that is too costly too automate (yet). This includes humans operating machines of LoA 0 (see Fig. 11) which – with the exception of small handheld tools – is almost gone by now.

Current SDLs report to remove this human factor completely within closed-loop experimentation [68, 91]. On closer analysis, this is only true superficially speaking:

1. The tasks are very specific and closed loop system boundaries are very narrowly defined so that goal derivation only on the level of a DOE_1 agent (see Fig. 13) is actually implemented and further reasoning and strategising about goals is outsourced to humans outside of the SDL.

2. SDLs are known to require lots of maintenance which is often done exclusively manually and without much system-intrinsic troubleshooting support due to relevant aspects not incorporated in digital twins or the inaccessibility of black box systems.

To achieve fully autonomous laboratories and eventually an AI scientist, the ultimate goal seems to imply removing the human-in-the-loop altogether. Considering a systems perspective including economic and sustainability criteria though, the DLF only aims to move the decision-making human slowly out of the loop while keeping technicians *etc.* as potential mobile units. These perform tasks as instructed by TWA, which will naturally fade out over time based on improved cost-efficiency of robotic systems.

In order to bridge the interim technology gap (see section 3.7), as current systems still need plenty of direct instruction and supervision, the process of moving the goal-setting human out of the loop needs to be a continuous process: first, separating the two functions of humans in the loop, then starting to explicitly instruct humans via the system for carrying out plannable work, and finally implementing better and better goal derivation and self-evaluation systems. In the meantime, we need to account for the inevitable interlock with the research ecosystem of human scientists as Kitano articulated [55].

For this, we might make use of existing frameworks such as SOHO because "CORA [and others] do[es] not support the contextualisation and interpretation of behaviours and other autonomous agents (*e.g.*, human operators) with respect to the global production objectives and processess" [104]. Going forward, we see humans as versatile handlers (i.e., mobile units) that act as physical agents based on the information provided to them. In the short term, humans will still synthesise findings and formulate the next investigation as phases of the general scientific data analysis process [36, 82]. These instances of "human-machine teaming" [100] can be achieved *e.g.*, with collaborative robots that allow human researchers to work in the same physical space [19].

For efficient teaming, seamless HMI are required that exhibit low latency. This can be achieved for example via easy-to-use mobile or web.based applications, which have been developed for specific island solutions (see section 2) – *e.g.*, ELN [59]. A framework for mobile applications interacting with TWA would offer a much broader range of potential applications to enable holistic lab automation and is currently under development. Furthermore, recently the rise of large language models (LLM) has revolutionised the way humans can interact with digital systems via natural language. The development of "copilot" systems to interact with SDLs such as CRESt [88] or ChemOS [90] have demonstrated the applicability and relevance in the field of lab automation. We therefore plan to extend the capabilities of Marie [110], a "knowledge graph question answering system" for TWA, to include all concepts related to the DLF (see Fig. 20).

The dual task of humans in the classical as well as the newly proposed approach to lab automation is illustrated in Fig. 16. In Fig. 16(b) we see TWA agents instead of humans breaking down the original abstract goal into smaller sub-goals and eventually deriving concrete tasks and experiments (see section 3.6). Due to its integrated knowledge model, no outside databases, literature, or experts need to be consulted. Furthermore, the world model is constantly updated and kept consistent in TWA via the so-called "derived information framework" [7]. This differs from the network envisioned by Martin et al. [72] where the goal-setting human researcher is part of the loop by giving recommendations to an automated lab. The inclusion of a "traditional lab" with humans still carrying out some of the physical labour shows some similarities but is not as deeply integrated as in this approach.

6 Implications and applications

The successful application of AI and automation in chemistry research depends heavily on the effective embedding of chemistry knowledge. The use of digital twins and dynamic knowledge graphs provides an opportunity to integrate chemical data and knowledge into the digital twin framework proposed in section 5. This framework is, by design, technically able to suffice the "principles of thoughtful AI" formulated by Yolanda Gil [31, 32]:

- Deep semantic knowledge representations – including metadata – incorporate **rationality** by design and enable providing **context**.
- The holistic view of a systems engineer with which we approach the research laboratory and its digitisation suffices the **systems** and **network** principles by design and enables proper **articulation**.
- The goal-driven approach introduced in section 4.2 enables the system to take **initiative** and follow a predefined strict code of **ethics**.

The DLF aims to enable automation across all aspects of a research laboratory. In Fig. 19 the overlaps of these aspects are highlighted and names are given to potential applications for each intersection. The integration of chemical knowledge into our framework allows for the creation of these digital counterparts of typically human responsibility areas to enhance and streamline laboratory processes. Here, we discuss these three use cases: the “Digital Lab Facility Manager”, the “Digital Research Scientist”, and the “Digital Lab Manager”.

6.1 Digital Researcher

Applying the DLF to AE setups enabled us to connect and orchestrate multiple SDLs remotely as detailed in a recent publication [6]. As next steps, we are planning to connect all lab assets to BMS, LIMS, and LIS systems, thereby enabling automated management of environmental conditions as well as procurement, as indicated in Fig. 21. The final goal is full integration of flexible and modular automated experimentation, see appendix A.

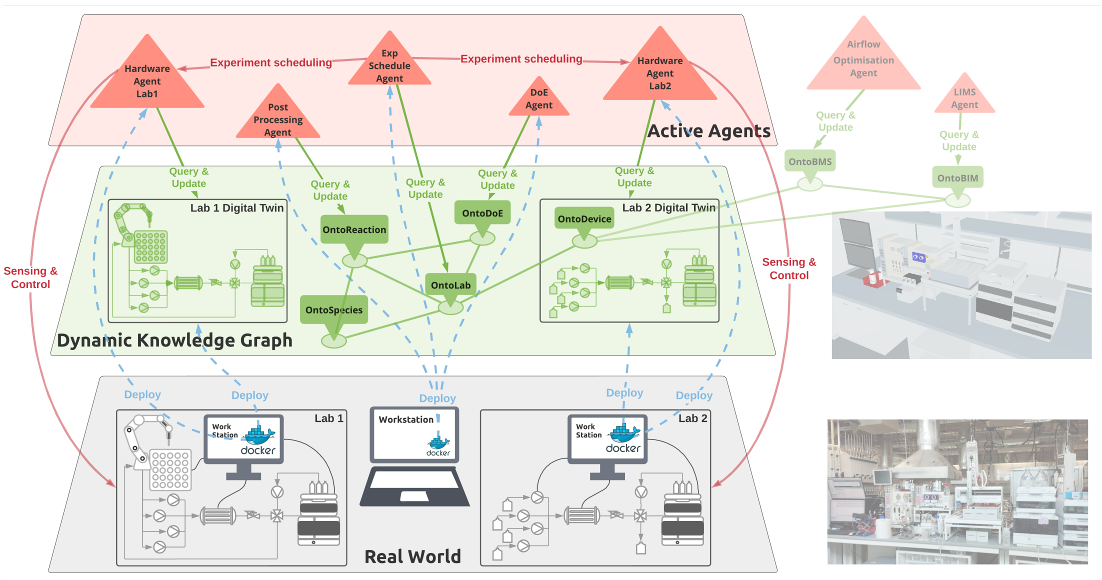


Figure 21: *Potential application of Digital Researcher operating and managing experiments over distributed labs, adapted from Bai et al. [6].*

6.2 Digital Lab Facility Manager

The decarbonisation potential of the building sector is huge, particularly regarding HVAC operations [48]. As chemical research labs require consuming large amounts of makeup air and electricity, loss of energy can be potentially saved. For this, knowledge about building, infrastructure, and behaviour has to be integrated. Within the DLF, we, therefore, aim to connect information and controllability of GIS, BIM, and BMS systems to enable optimisation of temperatures, airflow rates, *etc.* as shown in Fig. 22.

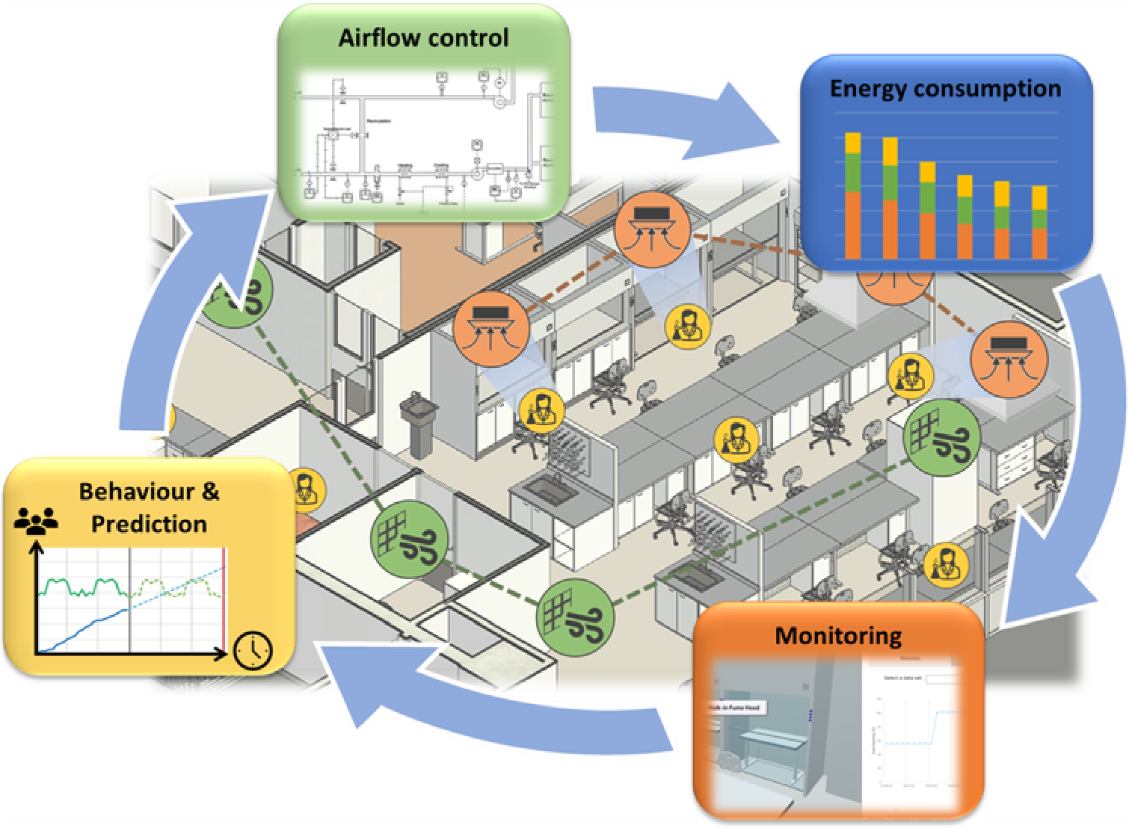


Figure 22: *Application of a Digital Building Manager regulating airflow in a research lab based on operations and live conditions to reduce energy consumption.*

Such a system would not only have major implications with lab sustainability [86] but also enable optimised scheduling and planning of experiments within individual research campaigns. The first steps in this direction have been successful by connecting GIS and BIM data of a chemical research laboratory in Singapore [84].

6.3 Digital Laboratory Manager

The integration of BIM with BMS, assets management, and live data on the other hand enables the automation of many tasks traditionally carried out by a lab manager. As shown in Fig. 23 this encompasses automated asset management, inventory tracking, and resource allocation.

Integration with data about specific experiments and actual chemical knowledge would now even empower scientists to seamlessly choose the types of reactants and setups that suffice certain criteria imposed by lab management. Within the Digital Lab Framework, first steps have been undertaken by showcasing the ability of automated solvent selection [80] based on sustainability-related criteria.

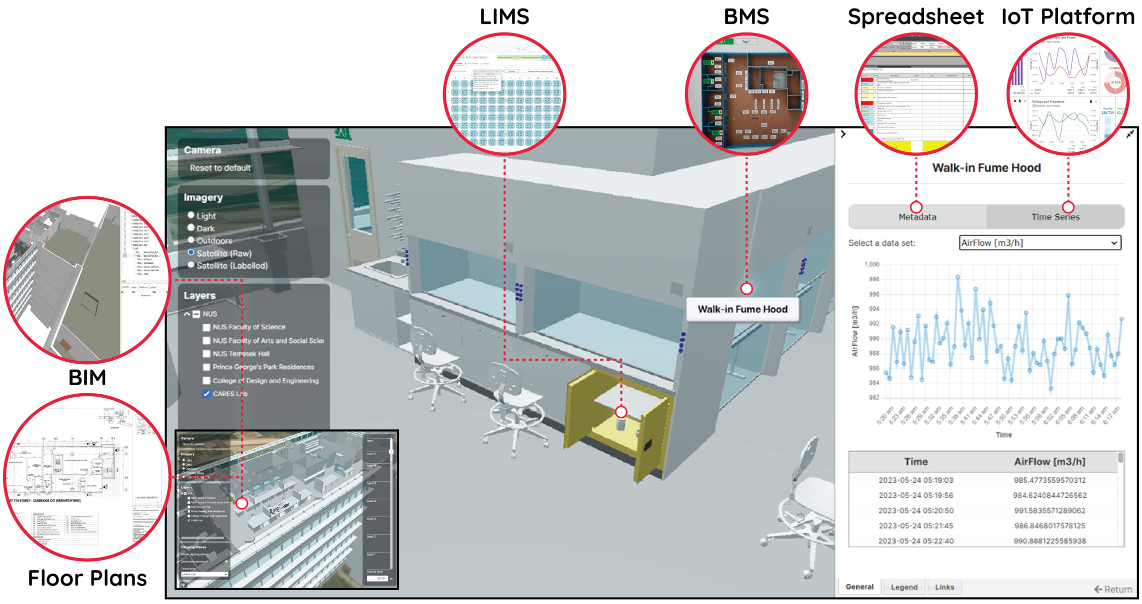


Figure 23: *Envisioned user interface of a Digital Lab Manager, depicting the integration of different tasks and information that are currently spread across individual and siloed applications, adapted from Quek et al. [84].*

7 Conclusion

In the face of mounting global challenges, the urgency for accelerated research has never been more apparent. We imagine a world in which this urgency is taken with adequate seriousness in terms of resource allocation and policymaking, such that the rate of scientific discovery has been increased by orders of magnitude. Such a world would at some point resemble the vision of AI scientists formulated by Hiroaki Kitano which are capable of making Nobel-prize-worthy scientific discoveries. In such a world, humans would only set high-level goals based on which the AI scientist derives appropriate experiments that are designed, executed, and analysed autonomously. Evaluating the current trajectory of automation within the domain of experimental research, we took on the perspective of a systems engineer, considering all aspects of a research laboratory: the experiment itself, the handler conducting and monitoring the experiment, the laboratory infrastructure, and peripheral systems beyond. These aspects overlap and interact with each other, each presenting unique challenges for automation. While specific solutions are available to automate parts of certain aspects, a holistic approach to lab automation is required to move towards the vision of an AI scientist.

Based on a holistic view of these systems, we can identify general challenges that need to be overcome. The first challenge lies in **interoperability**, typically referring to the perceived need for more integrated platforms that amalgamate functionality, such as ELN and LIMS, or standardised protocols for communication and data exchange. It reaches further though as a much broader range of knowledge needs to be available to automate of peripheral tasks usually conducted by managers and technicians, including reordering consumables, auditing assets, scheduling maintenance, budgeting, and enforcing safety protocols. As this knowledge keeps expanding, the need for **adaptability** to effectively manage the dynamic nature of research experiments becomes imminent. Modular design principles are necessary, not only for hardware but also for software components, accommodating live data and enabling fundamental changes in setup or underlying models. It is not enough to account for new data and changing conditions in a reactive way though. A 'system of systems' should also **orchestrate** research activities, ranging from the full cycle of a single campaign to managing complementary experiments or computations. This facilitates a more distributed approach of research activities which has been shown beneficial. The necessary coordination of information across different domains and scales not only requires interoperable components but representations allowing for **deep knowledge** models. This includes underlying processes (*e.g.*, quantum chemistry) as well as potential error sources and uncertainties of methods or equipment used which an AI scientist needs to consider when formulating new hypotheses. Adequate knowledge representation is only a prerequisite here; algorithms need to be capable of advanced **reasoning**. These capabilities are required to infer hypothesised facts from known ones, which will mostly comprise existing knowledge and new measurement data to 'close the loop'. Going beyond closed-loop experimentation, **goal derivation** is essential. We view this as a natural progression of increasingly sophisticated methods for designing experiments based on relevant hypotheses. Taking this to the extreme, an AI scientist would be able to derive potential objectives and ultimately experimental campaigns from very abstract goals such as the SDGs. The underlying model of knowledge and reasoning in most existing SDLs is based on statistical ML algorithms though, making them behavioural black boxes with only implicit knowledge of physics or chemistry of narrowly defined experimental setups. This makes it impossible to autonomously check the alignment of experimental objectives with possibly contradicting higher-level goals that could impel a human researcher to carry out an experiment in a non-ideal manner. The described discontinuity between the current and desired state is called **interim technology gap** and cannot be bridged in a single step of integrating all available solutions into a unified system.

To address the imperative of integrating sustainability deep into the fabric of scientific research and experimentation, systems thinking and systems modelling provide powerful tools. This has been discussed in detail by Weber et al. [107] and the present perspective has been guided by some of their key suggestions: for example, widening system boundaries allows assessment of full impact cradle-gate-crave in life cycle analysis (LCA), which prompted us to rigorously include physical and conceptual peripheries in the defined system. Furthermore, to manage the complexities emerging from connected subsystems across many scales (from market systems to reaction networks), semantic web technology and particularly structured knowledge graphs provide the only promising approach at this time. Leveraging such an approach now enables us – amongst others – to evaluate metrics across different systems and implement graph- or network-based decision strategies more easily.

Based on these assessments, this work offers a change in perspective: instead of building and extending an increasing number of platforms that cater to specific tasks defined by humans, what would be necessary to create a universal knowledge model that derives and allocates tasks based on abstract goals set by humans? We argue that this would require a focus on creating **comprehensive digital twins** instead of self-driving laboratories, shifting from a task-driven to a **goal-driven approach** and using **knowledge graphs** instead of traditional databases. We therefore introduce the TWA Digital Lab Framework for connected and distributed digital twins enabling holistic lab automation while considering human-machine interactions from the very start. In the upcoming series of publications following this, we will introduce applications focusing on the automation of specific sets of tasks within the overall system.

Research data

All data underlying this study is cited in references.

Conflicts of interest

There are no conflicts to declare.

Acknowledgements

This research was supported by the National Research Foundation, Prime Minister's Office, Singapore under its Campus for Research Excellence and Technological Enterprise (CREATE) programme. Part of this work was supported by Towards Turing 2.0 under the EPSRC Grant EP/W037211/1 and The Alan Turing Institute. S. D. Rihm acknowledges financial support from Fitzwilliam College, Cambridge, and the Cambridge Trust. J. Bai acknowledges financial support provided by CSC Cambridge International Scholarship from Cambridge Trust and China Scholarship Council. M. Kraft gratefully acknowledges the support of the Alexander von Humboldt Foundation. For the purpose of open access, the author has applied a Creative Commons Attribution (CC BY) licence to any Author Accepted Manuscript version arising.

Nomenclature

AEP Autonomous experimental platform

AE Autonomous experimentation

AGI Artificial general intelligence

AHP Analytic hierarchy process

AI Artificial intelligence

BIM Building information management

BMS Building management system

CDS Chromatography data system

DBMS Database management system

DIKW Data, information, knowledge, wisdom (model)

DISK (Automated) discovery of scientific knowledge

DLF (TWA) Digital Lab Framework

DoE Design of experiment

ELN Electronic lab notebook

ERP Enterprise resource planning

ESAMP Event-sourced architecture for materials provenance

EXPO The ontology of scientific experiments

GD Goal derivation

GIS Geographic information system

HMI Human-machine interface

HVAC Heating, ventilation, and air conditioning

IoT Internet of things

IRI Internationalised resource identifier

KD Knowledge discovery

KG Knowledge graph

LCA Life cycle analysis

LEAF Laboratory efficiency assessment framework

LES Lab execution system

LIMS Lab information management system

LIS Lab inventory system

LLM Large language model

ML Machine learning

NLP Natural language processing

PROCO Process chemistry ontology

R&D Research and development

RSC Royal Society of Chemistry

SDGs (United Nations') sustainable development goals

SDL Self-driving laboratory

SDMS Scientific data management system

SOHO Sharework ontology for human robot collaboration

TWA The World Avatar (project)

A Necessary concepts

A.1 Conceptual Layers of Experiments

Within the framework of TWA, an experimental setup is a subspace of the base world on which agents can perform actions (experiments) to use a certain set of input parameters to receive a certain set of output parameters. For example, eCO2R experiments by Huang et al. [46, 58, 87] in reference to the DoE process shown in Fig. 13:

- World problem: How can we use carbon dioxide to produce useful chemicals?
  - What is a feasible process? Is electrocatalytical reduction a reasonable candidate?
  - What is the best catalyst? Are copper-based metal electrodes work?
  - How can we tune reaction conditions for better selectivity?
  - ...
- Specific problem: How does applied potential influence Faradaic Efficiency?
- Experiment: see Fig. 24
- Observation / Raw Data: single data-point of U_{ref}, x_j^{GC} and I_{meas}
- Result / Interpretation: single data-point of U_{RHE} and FE_j

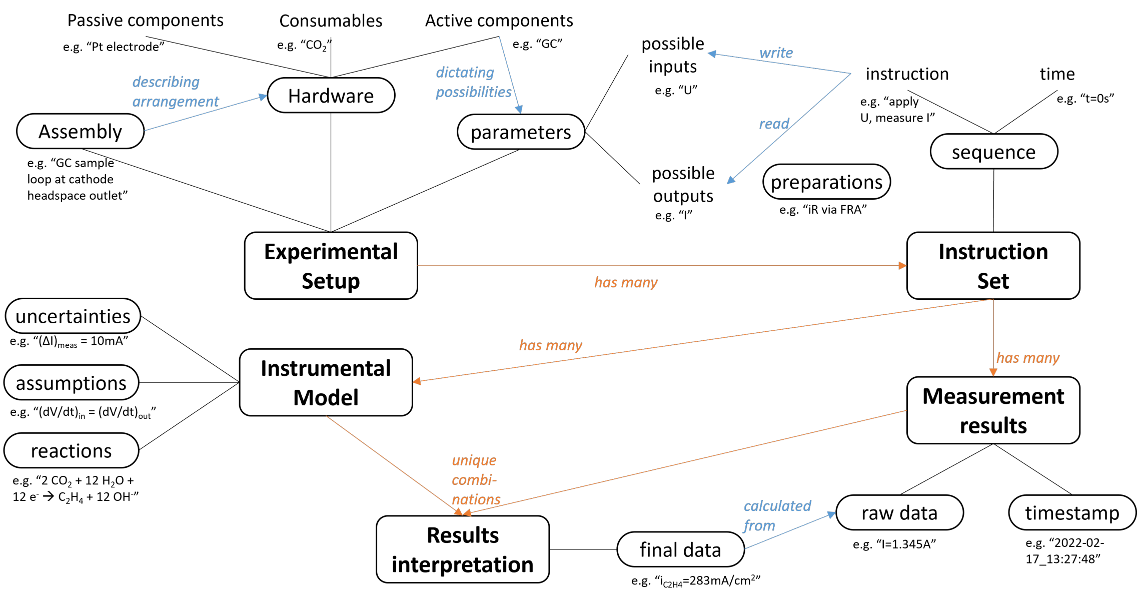


Figure 24: *Relation of concepts to carry out and analyse experiments, for example eCO2R on copper-based catalysts under different constant potentials.*

The conceptual layers shown in Fig. 24 are in detail:

- **Components:** everything needed to run experiment – not prepare!
  - Passive components
    - Two-compartment cell (10 mL each, 0.385 cm^2)
    - Pt electrode (ALS Japan)
    - Anion exchange membrane (Selemion AMV, AGC Asahi Glass)
    - Ag/AgCl reference electrode (sat. KCl, Pine)
    - Cu(100) electrode (99.99%, 10 mm diameter, MTI Corp.)
  - Active components
    - MFC (MC 100SCCM-D, Alicat Scientific)
    - GC (GC-7890A, Agilent)
    - H-NMR (500 MHz, Bruker Avance 500)
    - Potentiostat (Gamry Reference 600)
  - Consumables
    - $KHCO_3$ (0.1M, 99.5%, Sigma-Aldrich)
    - CO_2 (99.999%, Linde)
    - Phenol (99.5%, Scharlau)
    - DMSO (99.9%, Quality Reagent Chemical)
    - D_2O (99.96% deuterium, Merck Millipore)
- **Setup:** putting together components to run experiments
  1. Assemble cell: Pt electrode, anodic compartment (AC), membrane, cathodic compart-ment (CC), Cu electrode (WE)
  2. Put Reference Electrode in cathodic compartment
  3. Fill in electrolyte: 6.4 mL in cath., 8 mL in an. compartm.
  4. Connect CC: inlet to CO_2 via MFC, outlet to GC sample
- **Preparations:** instructions that have to be only done once for X runs
  1. New working electrode
     (a) Mechanical polishing with alumina slurries to mirror-like finishes
     (b) Electropolishing with H_3PO_4/H_2SO_4 solution
     (c) Acid rinsing w/ deionized water and $HClO_4$
  2. Calibrate analytics
     (a) Calibrate GC w/ stan-dards (H_2, CO, C_2H_4)
     (b) Calibrate NMR w/ stan-dards (formate, acetate, ethanol, propanol)
  3. Prepare electrolyte
     (a) Saturated $KHCO_3$ with CO_2 for 10 min

4. Potential compensation
   (a) Interrupt method
   (b) Measure pH

- **Instructions:** every single step I/O to perform a single run
  - $t = 0$ s: Bubble CO_2 through electrolyte at 20 sccm
    * Start stirring, 1500 rpm
    * Apply desired potential
    * Start measuring current
  - $t = 3$ min: Take 2.5 mL headspace aliquot, inject into GC
  - $t = 14$ min: GC sampling as above
  - $t = 26$ min: GC sampling as above
  - $t = 37$ min: GC sampling as above
  - $t = 40$ min: Stop current/bubbling
    * Take 2 mL sample of catholyte & anolyte
    * Mix w/ 25 mM phenol & 5 mM DMSO, take 1 mM
    * Add 0.2 mL D_2O, perform H-NMR

A.2 To develop

Additional concepts are necessary as shown in Fig. 25.

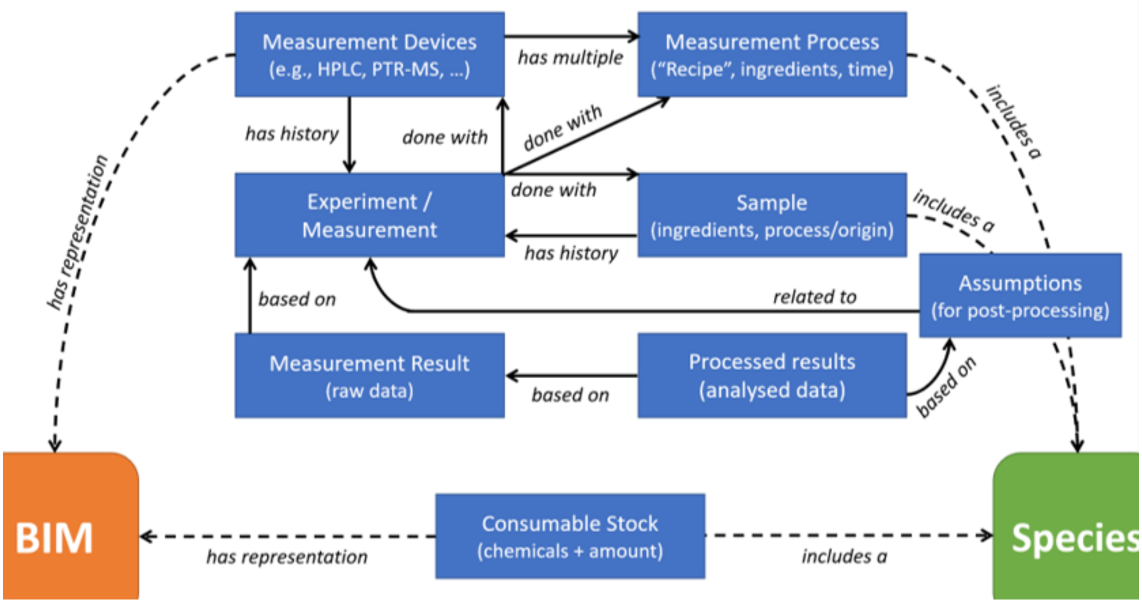


Figure 25: *Relation of concepts to automate general research activities.*

B Nomenclature

The nomenclature used around laboratory automation is inconsistent across the literature.

B.1 ‘Closed loop’ research cylces

When closed loop experimentation is discussed, the specific steps within this loop are often identified and named very differently. All things considerd, most authors refer to the same procedure with some specificities according to their field of application. Some examples are:

- Design - Experiment - Hypothesis - Prediction [74]
- Design - Build - Test - Learn [11]
- Design - Make - Test - Analyse [6]
- Hypothesis - exp. design - exp. execution - exp. observation - analysis [99]
- Hypothesis generation - experimantal planning - experimental execution [55]
- For literature / data-driven knowledge discovery
  - Hypothesis - line of inquiry (retrieve and analyse data) - revised hypothesis [32, 34]
  - Search space - automated experimentation - data analysis - decision making [108]
- More general / summarised cycles
  - (Initialize) - Plan - Experiment - Analyze - (Conclude) [100]
  - Plan - Execute - Disseminate [28]

B.2 Platform-based systems

Most ‘autonomous laboratory’ solutions are platform-centric. Some examples are:

- Self-driving laboratory (SDL) [2, 11, 13, 41, 50, 69, 72, 95]: In use since 2015
- Autonomous Experimental platform (AEP) [108]
- Closed-loop experimentation: in use since 1980s
- Robotic chemist [19], transitioning to the digital researcher theme shown below in appendix B.3. The seamless transition of these phrases indicates again the lack of awareness for the inteim technology gap (see section 3.7) in the broader scientific community.

B.3 Digital or automated researcher

Different names are in circulation for automated systems that can autonomously carry out certain research tasks. In this instance, the differences are mostly reflective of how extensive the envisioned system is, e.g. including robotic capabilities or not. Some examples are:

- AI scientist [32, 33, 35, 36, 55]
- Robot scientist [54, 98, 99]
- Machine investigator [11]
- AI chemist [111]

More towards support of human researchers: "copilot" [88] or "*in silico* colleague" [56]

C Comparisons

Key ideas and details of implementation are shown in section 5. The most important differences to using a more conventional approach are shown in Tab. 2.

Table 2: *Comparison of paradigms between conventional SDL development and the introduced approach of DLF.*

Conventional	**Suggestion**
Bottom-up	Top-down
Platform	Knowledge Graph
Design Thinking	Systems Engineering
Task-driven	Goal-driven
Self-driving labs	Comprehensive digital twins

References

[1] Changing the world for the better with data science and AI. Technical report, The Turing Institute, 2023. URL https://www.turing.ac.uk/sites/default/files/2023-03/turing_2.0_-_institute_strategy_-_final.pdf. Last accessed October 13, 2023.

[2] M. Abolhasani and E. Kumacheva. The rise of self-driving labs in chemical and materials sciences. *Nature Synthesis*, 2(6):483–492, 2023. ISSN 2731-0582. doi:10.1038/s44160-022-00231-0.

[3] J. Akroyd, S. Mosbach, A. Bhave, and M. Kraft. Universal Digital Twin - A Dynamic Knowledge Graph. *Data-Centric Engineering*, 2(4):1–25, 2021. ISSN 2632-6736. doi:10.1017/dce.2021.10.

[4] S. Andereggen, F. A. Zoller, and R. Boutellier. Sharing research equipment to bridge intraorganizational boundaries. *Research Technology Management*, 56(1): 49–57, 2013. ISSN 0895-6308. doi:10.5437/08956308X5601082.

[5] D. Angelone, A. J. Hammer, S. Rohrbach, S. Krambeck, J. M. Granda, J. Wolf, S. Zalesskiy, G. Chisholm, and L. Cronin. Convergence of multiple synthetic paradigms in a universally programmable chemical synthesis machine. *Nature Chemistry*, 13(1):63–69, 2021. ISSN 1755-4349. doi:10.1038/s41557-020-00596-9.

[6] J. Bai, L. Cao, S. Mosbach, J. Akroyd, A. A. Lapkin, and M. Kraft. From Platform to Knowledge Graph: Evolution of Laboratory Automation. *JACS Au*, 2(2):292–309, 2022. ISSN 1520-5126. doi:10.1021/jacsau.1c00438.

[7] J. Bai, K. F. Lee, M. Hofmeister, S. Mosbach, J. Akroyd, and M. Kraft. A Derived Information Framework for a Dynamic Knowledge Graph and its Application to Smart Cities. Submitted for publication; preprint available online, 2023. URL https://como.ceb.cam.ac.uk/preprints/302/.

[8] J. Bai, S. Mosbach, C. J. Taylor, D. Karan, K. F. Lee, S. D. Rihm, J. Akroyd, A. A. Lapkin, and M. Kraft. From Platform to Knowledge Graph: Distributed Self-Driving Laboratories. Submitted for publication; preprint available online, 2023. URL https://como.ceb.cam.ac.uk/preprints/310/.

[9] B. Balaji, A. Bhattacharya, G. Fierro, J. Gao, J. Gluck, D. Hong, A. Johansen, J. Koh, J. Ploennigs, Y. Agarwal, M. Bergés, D. Culler, R. K. Gupta, M. B. Kjærgaard, M. Srivastava, and K. Whitehouse. Brick: Metadata schema for portable smart building applications. *Applied Energy*, 226(September 2017):1273–1292, 2018. ISSN 0306-2619. doi:10.1016/j.apenergy.2018.02.091.

[10] H. Bär, R. Hochstrasser, and B. Papenfuß. SiLA: Basic standards for rapid integration in laboratory automation. *Journal of Laboratory Automation*, 17(2):86–95, 2012. ISSN 2211-0690. doi:10.1177/2211068211424550.

[11] J. Beal and M. Rogers. Levels of autonomy in synthetic biology engineering. *Molecular Systems Biology*, 16(12):1–5, 2020. ISSN 1744-4292. doi:10.15252/msb.202010019.

[12] W. Beker, R. Roszak, A. Wolos, N. H. Angello, V. Rathore, M. D. Burke, and B. A. Grzybowski. Machine Learning May Sometimes Simply Capture Literature Popularity Trends: A Case Study of Heterocyclic Suzuki-Miyaura Coupling. *Journal of the American Chemical Society*, 144(11):4819–4827, 2022. ISSN 1520-5126. doi:10.1021/jacs.1c12005.

[13] J. A. Bennett and M. Abolhasani. Autonomous chemical science and engineering enabled by self-driving laboratories. *Current Opinion in Chemical Engineering*, 36:100831, 2022. ISSN 2211-3398. doi:10.1016/j.coche.2022.100831.

[14] T. Berners-Lee. Linked data - design issues, 2006. URL `http://www.w3.org/DesignIssues/LinkedData.html`. Last accessed October 13, 2023.

[15] T. Berners-Lee. Giant Global Graph, 2007. URL `https://web.archive.org/web/20160713021037/http://dig.csail.mit.edu/breadcrumbs/node/215`. Last accessed October 13, 2023.

[16] T. Berners-Lee, J. Hendler, and O. Lassila. The Semantic Web. *Scientific American*, 284(5):28–37, 2001. ISSN 0036-8733.

[17] R. Brazil. Automation in the Chemistry Lab, 2021. URL `https://www.chemistryworld.com/careers/automation-in-the-chemistry-lab/4012832.article`. Last accessed October 13, 2023.

[18] A. Bubliauskas, D. J. Blair, H. Powell-Davies, P. J. Kitson, M. D. Burke, and L. Cronin. Digitizing Chemical Synthesis in 3D Printed Reactionware. *Angewandte Chemie - International Edition*, 61(24), 2022. ISSN 1521-3773. doi:10.1002/anie.202116108.

[19] B. Burger, P. M. Maffettone, V. V. Gusev, C. M. Aitchison, Y. Bai, X. Wang, X. Li, B. M. Alston, B. Li, R. Clowes, N. Rankin, B. Harris, R. S. Sprick, and A. I. Cooper. A mobile robotic chemist. *Nature*, 583(7815):237–241, 2020. ISSN 1476-4687. doi:10.1038/s41586-020-2442-2.

[20] A. Clark. Conclusions: The Future of Prediction. In *Surfing Uncertainty: Prediction, Action, and the Embodied Mind*. Oxford University Press, 01 2016. ISBN 9780190217013. doi:10.1093/acprof:oso/9780190217013.003.0011.

[21] A. A. da Conceic'ão, L. P. Ambrosio, T. R. Leme, A. C. S. Rosa, F. F. Ramborger, G. P. Aquino, and E. C. Vilas Boas. Internet of things environment automation: A smart lab practical approach. In *2022 2nd International Conference on Information Technology and Education (ICIT&E)*, pages 01–06, 2022. doi:10.1109/ICITE54466.2022.9759899.

[22] Daniel P. Tabor, Loic M. Roch, Semion K. Saikin, Christoph Kreisbeck, Dennis Sheberla, Joseph H. Montoya, Shyam Dwaraknath, Murathan Aykol, Carlos Ortiz, Hermann Tribukait, Carlos Amador-Bedolla, Christoph J. Brabec, Benji Maruyama, Kristin A. Persson, and Alan Aspuru-Guzik. Accelerating the discovery of materials for clean energy in the era of smart automation. *Nature Reviews Materials*, 3:5–20, 2018. ISSN 2058-8437. doi:10.1038/s41578-018-0005-z.

[23] N. David, W. Sun, and C. W. Coley. The promise and pitfalls of AI for molecular and materials synthesis. *Nature Computational Science*, 3(May):362–364, 2023. ISSN 2662-8457. doi:10.1038/s43588-023-00446-x.

[24] Defense Acquisition University Press. *Systems engineering fundamentals*. Fort Belvoir, Virginia, 2001.

[25] A. Degani, N. Ames, and M. View. Formal Verification of Human-Automation Interaction. *Human Factors*, 44(1):28–43, 2002. ISSN 1547-8181. doi:10.1518/0018720024494838.

[26] A. Eibeck, M. Q. Lim, and M. Kraft. J-park simulator: An ontology-based platform for cross-domain scenarios in process industry. *Computers & Chemical Engineering*, 131:106586, 2019. ISSN 0098-1354. doi:10.1016/j.compchemeng.2019.106586.

[27] M. M. Flores-Leonar, L. M. Mejía-Mendoza, A. Aguilar-Granda, B. Sanchez-Lengeling, H. Tribukait, C. Amador-Bedolla, and A. Aspuru-Guzik. Materials Acceleration Platforms: On the way to autonomous experimentation. *Current Opinion in Green and Sustainable Chemistry*, 25:100370, 2020. ISSN 2452-2236. doi:10.1016/j.cogsc.2020.100370.

[28] J. G. Frey. The value of the Semantic Web in the laboratory. *Drug Discovery Today*, 14(11-12):552–561, 2009. ISSN 1359-6446. doi:10.1016/j.drudis.2009.03.007.

[29] J. Frohm, V. Lindström, M. Winroth, and J. Stahre. Levels of Automation in Manufacturing. *Ergonomia - International Journal of Ergonomics and Human Factors*, 30:181–207, 01 2008.

[30] S. S. Garud, I. A. Karimi, and M. Kraft. Design of computer experiments: A review. *Computers & Chemical Engineering*, 106:71–95, 2017. ISSN 0098-1354. doi:10.1016/j.compchemeng.2017.05.010.

[31] Y. Gil. Thoughtful artificial intelligence: Forging a new partnership for data science and scientific discovery. *Data Science*, 1(1-2):119–129, 2017. ISSN 2451-8484. doi:10.3233/ds-170011.

[32] Y. Gil. Will AI Write Scientific Papers in the Future? *AI Magazine*, 42(4):3–15, 2021. ISSN 0738-4602. doi:10.1609/aaai.12027.

[33] Y. Gil, M. Greaves, J. Hendler, and H. Hirsh. Amplify scientific discovery with artificial intelligence. *Science*, 346(6206):171–172, 2014. ISSN 1095-9203. doi:10.1126/science.1259439.

[34] Y. Gil, D. Garijo, V. Ratnakar, R. Mayani, R. Adusumilli, H. Boyce, A. Srivastava, and P. Mallick. Towards Continuous Scientific Data Analysis and Hypothesis Evolution. *Proceedings of the AAAI Conference on Artificial Intelligence*, 31(1): 4406–4414, 2017. ISSN 2159-5399. doi:10.1609/aaai.v31i1.11157.

[35] Y. Gil, R. King, and H. Kitano. AI scientist grand challenge. Technical Report February, The Alan Turing Institute, 2020. URL `https://www.turing.ac.uk/sites/default/files/2021-02/summary_of_discussion_workshop_2020_ai_scientist_grand_challenge_clean.pdf`. Last accessed October 13, 2023.

[36] Y. Gil, D. Khider, M. Osorio, V. Ratnakar, H. Vargas, D. Garijo, and S. Pierce. Towards Capturing Scientific Reasoning to Automate Data Analysis. In J. Culbertson, A. Prefors, H. Rabagliati, and V. Ramenzoni, editors, *Proceedings of the 44th Annual Conference of the Cognitive Science Society*, 2022.

[37] B. Goertzel. Human-level artificial general intelligence and the possibility of a technological singularity. A reaction to Ray Kurzweil's The Singularity Is Near, and McDermott's critique of Kurzweil. *Artificial Intelligence*, 171(18):1161–1173, 2007. ISSN 0004-3702. doi:10.1016/j.artint.2007.10.011.

[38] 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, and A. Aspuru-Guzik. Automatic Chemical Design Using a Data-Driven Continuous Representation of Molecules. *ACS Central Science*, 4(2):268–276, 2018. ISSN 2374-7951. doi:10.1021/acscentsci.7b00572.

[39] D. Greenfield. Defining the Digital Twin. *AutomationWorld*, apr 2022. URL `https://www.automationworld.com/factory/iiot/article/22184905/siemens-defines-the-digital-twin`. Last accessed October 13, 2023.

[40] J. Hafner. Ab-initio simulations of materials using VASP: Density-functional theory and beyond. *Journal of Computational Chemistry*, 29(13):2044–2078, 2008. ISSN 1096-987X. doi:https://doi.org/10.1002/jcc.21057.

[41] A. J. Hammer, A. I. Leonov, N. L. Bell, and L. Cronin. Chemputation and the Standardization of Chemical Informatics. *JACS Au*, 1(10):1572–1587, 2021. ISSN 1520-5126. doi:10.1021/jacsau.1c00303.

[42] S. Hessam, M. Craven, A. I. Leonov, G. Keenan, and L. Cronin. A universal system for digitization and automatic execution of the chemical synthesis literature. *Science*, 370(6512):101–108, 2020. ISSN 1095-9203. doi:10.1126/science.abc2986.

[43] K. Hippalgaonkar, Q. Li, X. Wang, J. W. Fisher, J. Kirkpatrick, and T. Buonassisi. Knowledge-integrated machine learning for materials: lessons from gameplaying and robotics. *Nature Reviews Materials*, 8(4):241–260, 2023. ISSN 2058-8437. doi:10.1038/s41578-022-00513-1.

[44] I. Holland and J. A. Davies. Automation in the Life Science Research Laboratory. *Frontiers in Bioengineering and Biotechnology*, 8(November):1–18, 2020. ISSN 2296-4185. doi:10.3389/fbioe.2020.571777.

[45] J. Holmberg and K.-H. Robèrt. Backcasting—a framework for strategic planning. *International Journal of Sustainable Development & World Ecology*, 7(4):291–308, 2000. ISSN 1745-2627. doi:10.1080/13504500009470049.

[46] Y. Huang, A. D. Handoko, P. Hirunsit, and B. S. Yeo. Electrochemical Reduction of CO2 Using Copper Single-Crystal Surfaces: Effects of CO* Coverage on the Selective Formation of Ethylene. *ACS Catalysis*, 7(3):1749–1756, 2017. ISSN 2155-5435. doi:10.1021/acscatal.6b03147.

[47] O. Inderwildi and M. Kraft. *Synthesis*, pages 243–253. Springer International Publishing, Cham, 2022. ISBN 9783030862152. doi:10.1007/978-3-030-86215-2_32.

[48] O. Inderwildi, C. Zhang, X. Wang, and M. Kraft. The impact of intelligent cyber-physical systems on the decarbonization of energy. *Energy and Environmental Science*, 13(3):744–771, 2020. ISSN 1754-5706. doi:10.1039/c9ee01919g.

[49] K. M. Jablonka, L. Patiny, and B. Smit. Making the collective knowledge of chemistry open and machine actionable. *Nature Chemistry*, 14(4):365–376, 2022. ISSN 1755-4349. doi:10.1038/s41557-022-00910-7.

[50] Y. Jiang, D. Salley, A. Sharma, G. Keenan, M. Mullin, and L. Cronin. An artificial intelligence enabled chemical synthesis robot for exploration and optimization of nanomaterials. *Science Advances*, 8(40):1–12, 2022. ISSN 2375-2548. doi:10.1126/sciadv.abo2626.

[51] M. N. Jones, M. Stevnsborg, R. F. Nielsen, D. Carberry, K. Bagherpour, S. S. Mansouri, S. Larsen, K. V. Gernaey, J. Dreyer, J. Woodley, J. K. Huusom, and K. Dam-Johansen. *Pilot Plant 4.0: A Review of Digitalization Efforts of the Chemical and Biochemical Engineering Department at the Technical University of Denmark (DTU)*, volume 49. Elsevier Masson SAS, 2022. ISBN 9780323851596. doi:10.1016/B978-0-323-85159-6.50254-2.

[52] M. P. Kantipudi and S. Velamuri. Internet of things – based smart campus – a review. In *4th Smart Cities Symposium (SCS 2021)*, volume 2021, pages 490–495, 2021. doi:10.1049/icp.2022.0392.

[53] J. Kim, Y. Gil, and V. Ratnakar. Semantic metadata generation for large scientific workflows. In I. Cruz, S. Decker, D. Allemang, C. Preist, D. Schwabe, P. Mika, M. Uschold, and L. M. Aroyo, editors, *The Semantic Web - ISWC 2006*, volume 4273 LNCS, pages 357–370. Springer Berlin Heidelberg, 2006. ISBN 3540490299. doi:10.1007/11926078_26.

[54] R. D. King, K. E. Whelan, F. M. Jones, P. G. Reiser, C. H. Bryant, S. H. Muggleton, D. B. Kell, and S. G. Oliver. Functional genomic hypothesis generation

and experimentation by a robot scientist. *Nature*, 427(6971):247–252, 2004. ISSN 1476-4687. doi:10.1038/nature02236.

[55] H. Kitano. Nobel Turing Challenge: creating the engine for scientific discovery. *npj Systems Biology and Applications*, 7(1):1–12, 2021. ISSN 2056-7189. doi:10.1038/s41540-021-00189-3.

[56] T. Klucznik, B. Mikulak-Klucznik, M. P. McCormack, H. Lima, S. Szymkuć, M. Bhowmick, K. Molga, Y. Zhou, L. Rickershauser, E. P. Gajewska, A. Toutchkine, P. Dittwald, M. P. Startek, G. J. Kirkovits, R. Roszak, A. Adamski, B. Sieredzińska, M. Mrksich, S. L. Trice, and B. A. Grzybowski. Efficient Syntheses of Diverse, Medicinally Relevant Targets Planned by Computer and Executed in the Laboratory. *Chem*, 4(3):522–532, 2018. ISSN 2451-9294. doi:10.1016/j.chempr.2018.02.002.

[57] A. Kondinski, J. Bai, S. Mosbach, J. Akroyd, and M. Kraft. Knowledge Engineering in Chemistry: From Expert Systems to Agents of Creation. *Accounts of Chemical Research*, 56(2):128–139, 2023. ISSN 1520-4898. doi:10.1021/acs.accounts.2c00617.

[58] K. P. Kuhl, E. R. Cave, D. N. Abram, and T. F. Jaramillo. New insights into the electrochemical reduction of carbon dioxide on metallic copper surfaces. *Energy and Environmental Science*, 5(5):7050–7059, 2012. ISSN 1754-5692. doi:10.1039/c2ee21234j.

[59] R. Kwok. Lab Notebooks Go Digital. *Nature*, 560:269–270, 2018. ISSN 1476-4687. doi:10.1038/d41586-018-05895-3.

[60] D. A. Leins, S. B. Haase, M. Eslami, J. Schrier, and J. T. Freeman. Collaborative methods to enhance reproducibility and accelerate discovery. *Digital Discovery*, 2 (1):12–27, 2023. ISSN 2635-098X. doi:10.1039/d2dd00061j.

[61] J. Leuridan. Comprehensive Digital Twin: An Enabler for Model Based System Engineering, October 2019. URL `https://www.youtube.com/watch?v=JmU5vGl-epI`. Keynote at EclipseCon Europe 2019; online recording available.

[62] J. Li, S. G. Ballmer, E. P. Gillis, S. Fujii, M. J. Schmidt, A. M. E. Palazzolo, J. W. Lehmann, G. F. Morehouse, and M. D. Burke. Synthesis of many different types of organic small molecules using one automated process. *Science*, 347(6227):1221–1226, 2015. ISSN 1095-9203. doi:10.1126/science.aaa5414.

[63] M. Li, Y. Ma, Z. Yin, and C. Wang. Structural Design of Digital Twin Laboratory Model Based on Instruments Sharing Platform. *Proceedings of the 32nd Chinese Control and Decision Conference, CCDC 2020*, pages 797–802, 2020. doi:10.1109/CCDC49329.2020.9164813.

[64] Z. Li, M. A. Najeeb, L. Alves, A. Z. Sherman, V. Shekar, P. Cruz Parrilla, I. M. Pendleton, W. Wang, P. W. Nega, M. Zeller, J. Schrier, A. J.

Norquist, and E. M. Chan. Robot-Accelerated Perovskite Investigation and Discovery. *Chemistry of Materials*, 32(13):5650–5663, 2020. ISSN 1520-5002. doi:10.1021/acs.chemmater.0c01153.

[65] G. Lippi and G. Da Rin. Advantages and limitations of total laboratory automation: A personal overview. *Clinical Chemistry and Laboratory Medicine*, 57(6):802–811, 2019. ISSN 1437-4331. doi:10.1515/cclm-2018-1323.

[66] M. R. Lopes, A. Costigliola, R. Pinto, S. Vieira, and J. M. Sousa. Pharmaceutical quality control laboratory digital twin–A novel governance model for resource planning and scheduling. *International Journal of Production Research*, 58(21): 6553–6567, 2020. ISSN 1366-588X. doi:10.1080/00207543.2019.1683250.

[67] B. P. MacLeod, F. G. Parlane, A. K. Brown, J. E. Hein, and C. P. Berlinguette. Flexible automation accelerates materials discovery. *Nature Materials*, 21(7):722–726, 2022. ISSN 1476-4660. doi:10.1038/s41563-021-01156-3.

[68] B. P. MacLeod, F. G. Parlane, C. C. Rupnow, K. E. Dettelbach, M. S. Elliott, T. D. Morrissey, T. H. Haley, O. Proskurin, M. B. Rooney, N. Taherimakhsousi, D. J. Dvorak, H. N. Chiu, C. E. Waizenegger, K. Ocean, M. Mokhtari, and C. P. Berlinguette. A self-driving laboratory advances the Pareto front for material properties. *Nature Communications*, 13(1):1–10, 2022. ISSN 2041-1723. doi:10.1038/s41467-022-28580-6.

[69] B. P. MacLeod, F. G. Parlane, and C. P. Berlinguette. How to build an effective self-driving laboratory. *MRS Bulletin*, 48(February):1–6, 2023. ISSN 0883-7694. doi:10.1557/s43577-023-00476-w.

[70] T. C. Malig, J. D. Koenig, H. Situ, N. K. Chehal, P. G. Hultin, and J. E. Hein. Real-time HPLC-MS reaction progress monitoring using an automated analytical platform. *Reaction Chemistry and Engineering*, 2(3):309–314, 2017. ISSN 2058-9883. doi:10.1039/c7re00026j.

[71] J. S. Manzano, W. Hou, S. S. Zalesskiy, P. Frei, H. Wang, P. J. Kitson, and L. Cronin. An autonomous portable platform for universal chemical synthesis. *Nature Chemistry*, 14(11):1311–1318, 2022. ISSN 1755-4349. doi:10.1038/s41557-022-01016-w.

[72] 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. S. Sundaram, D. Agarwal, B. A. Simmons, T. Backman, D. Banerjee, D. Tanjore, L. Ramakrishnan, and A. Singh. Perspectives for self-driving labs in synthetic biology. *Current Opinion in Biotechnology*, 79:102881, 2023. ISSN 1879-0429. doi:10.1016/j.copbio.2022.102881.

[73] R. Mercado, S. M. Kearnes, and C. W. Coley. Data Sharing in Chemistry: Lessons Learned and a Case for Mandating Structured Reaction Data. *Journal of Chemical Information and Modeling*, 63(14):4253–4265, 2023. ISSN 1549-960X. doi:10.1021/acs.jcim.3c00607.

[74] J. H. Montoya, M. Aykol, A. Anapolsky, C. B. Gopal, P. K. Herring, J. S. Hummelshøj, L. Hung, H. K. Kwon, D. Schweigert, S. Sun, S. K. Suram, S. B. Torrisi, A. Trewartha, and B. D. Storey. Toward autonomous materials research: Recent progress and future challenges. *Applied Physics Reviews*, 9(1):011405, 01 2022. ISSN 1931-9401. doi:10.1063/5.0076324.

[75] M. Nagata. User-Centered Automation Process in Synthetic Biology Research. Master's thesis, Massachusetts Institute of Technology, 2018.

[76] M. K. O'Malley. Principles of Human-machine Interfaces and Interactions. In M. Zhang, B. Nelson, and R. Felder, editors, *Life Science Automation: Fundamentals and Applications*, pages 101–125. Artech House, Norwood, 2007. ISBN 1596931051.

[77] A. Opoku and J. Y. Lee. The Future of Facilities Management: Managing Facilities for Sustainable Development. *Sustainability (Switzerland)*, 14(3):10–14, 2022. ISSN 2071-1050. doi:10.3390/su14031705.

[78] A. Palermo, D. Chem, and E. Frsc. Future of the Chemical Sciences. *Chemistry International*, 38(6), 2016. ISSN 0193-6484. doi:10.1515/ci-2016-0608.

[79] R. Parasuraman, T. B. Sheridan, and C. D. Wickens. A model for types and levels of human interaction with automation. *IEEE Transactions on Systems, Man, and Cybernetics Part A:Systems and Humans.*, 30(3):286–297, 2000. ISSN 1083-4427. doi:10.1109/3468.844354.

[80] L. Pascazio, S. Rihm, A. Naseri, S. Mosbach, J. Akroyd, M. Kraft, P. F. Drive, and P. F. Drive. A chemical species ontology for data integration and knowledge discovery. Submitted for publication; preprint available online, 2023. URL `https://como.ceb.cam.ac.uk/preprints/306/`.

[81] N. Pattinson, M. Neidhardt, M. Hopkins, and H. Haas. Lab of the Future Panel, May 2023. URL `https://applied.slas.org/products/lab-of-the-future-panel`. Panel discussion at SLAS Europe 2023; online recording available.

[82] X. Peng and X. Wang. Next-generation intelligent laboratories for materials design and manufacturing. *MRS Bulletin*, 48(2):179–185, 2023. ISSN 0883-7694. doi:10.1557/s43577-023-00481-z.

[83] R. Pollice, G. Dos Passos Gomes, M. Aldeghi, R. J. Hickman, M. Krenn, C. Lavigne, M. Lindner-D'Addario, A. Nigam, C. T. Ser, Z. Yao, and A. Aspuru-Guzik. Data-Driven Strategies for Accelerated Materials Design. *Accounts of Chemical Research*, 54(4):849–860, 2021. ISSN 1520-4898. doi:10.1021/acs.accounts.0c00785.

[84] H. Y. Quek, M. Hofmeister, S. D. Rihm, J. Yan, J. Lai, G. Brownbridge, M. Hillman, S. Mosbach, W. Ang, Y.-k. Tsai, D. N. Tran, S. Kang, W. Tan, M. Kraft, P. F. Drive, and P. F. Drive. BIM-GIS Integration : Knowledge graphs in a world of data silos. Submitted for publication; preprint available online, 2023. URL `https://como.ceb.cam.ac.uk/preprints/311/`.

[85] P. Raccuglia, K. C. Elbert, P. D. Adler, C. Falk, M. B. Wenny, A. Mollo, M. Zeller, S. A. Friedler, J. Schrier, and A. J. Norquist. Machine-learning-assisted materials discovery using failed experiments. *Nature*, 533(7601):73–76, 2016. ISSN 1476-4687. doi:10.1038/nature17439.

[86] G. Reid, H. Pain, A. Horan, J. Broderick, C. Dyer-Smith, I. Meazzini, J. Long, N. Sims, J. Anson, L. Marle, W. Niu, P. Stackhouse, H. Armes, S. De Pellegars, E. Eley, R. Holliday, I. Monk, T. Underwood, H. White, E. Ratcliffe, C. Southgate, N. Gibson, and C. Morley. Sustainable Laboratories. Technical report, Royal Society of Chemistry, Cambridge, 2022. URL `https://www.rsc.org/globalassets/22-new-perspectives/sustainability/sustainable-labs/sustainable-laboratories-report.pdf`. Last accessed October 13, 2023.

[87] D. Ren, Y. Deng, A. D. Handoko, C. S. Chen, S. Malkhandi, and B. S. Yeo. Selective Electrochemical Reduction of Carbon Dioxide to Ethylene and Ethanol on Copper(I) oxide catalysts. *ACS Catalysis*, 5(5):2814–2821, 2015. ISSN 2155-5435. doi:10.1021/cs502128q.

[88] Z. Ren, Z. Zhang, Y. Tian, and J. Li. CRESt – Copilot for Real-world Experimental Scientist. jul 2023. doi:10.26434/CHEMRXIV-2023-TNZ1X.

[89] S. D. Rihm, J. Bai, L. Pascazio, and M. Kraft. Fully Automated Kinetic Models Extend our Understanding of Complex Reaction Mechanisms. *Chemie-Ingenieur-Technik*, 95(5):740–748, 2023. ISSN 1522-2640. doi:10.1002/cite.202200220.

[90] L. M. Roch, F. Häse, C. Kreisbeck, T. Tamayo-Mendoza, L. P. E. Yunker, J. E. Hein, and A. Aspuru-Guzik. Chemos: Orchestrating autonomous experimentation. *Science Robotics*, 3(19):eaat5559, 2018. ISSN 2470-9476. doi:10.1126/scirobotics.aat5559.

[91] C. C. Rupnow, B. P. MacLeod, M. Mokhtari, K. Ocean, K. E. Dettelbach, D. Lin, F. G. Parlane, H. N. Chiu, M. B. Rooney, C. E. Waizenegger, E. I. de Hoog, A. Soni, and C. P. Berlinguette. A self-driving laboratory optimizes a scalable process for making functional coatings. *Cell Reports Physical Science*, 4(5):101411, 2023. ISSN 2666-3864. doi:10.1016/j.xcrp.2023.101411.

[92] B. Sanchez-Lengeling and A. Aspuru-Guzik. Inverse molecular design using machine learning: Generative models for matter engineering. *Science*, 361(6400): 360–365, 2018. ISSN 1095-9203. doi:10.1126/science.aat2663.

[93] N. Savage. Tapping into the drug discovery potential of AI. *Biopharma Dealmakers*, (June):37–39, 2021. ISSN 2730-6275. doi:10.1038/d43747-021-00045-7.

[94] R. S. Seaberg, R. O. Stallone, and B. E. Statland. Role of total laboratory automation in a consolidated laboratory network. *Clinical Chemistry*, 46(5):751–756, 2000. ISSN 0009-9147. doi:10.1093/clinchem/46.5.751.

[95] M. Seifrid, R. Pollice, A. Aguilar-Granda, Z. Morgan Chan, K. Hotta, C. T. Ser, J. Vestfrid, T. C. Wu, and A. Aspuru-Guzik. Autonomous Chemical Experiments: Challenges and Perspectives on Establishing a Self-Driving Lab. *Accounts of Chemical Research*, 55(17):2454–2466, 2022. ISSN 1520-4898. doi:10.1021/acs.accounts.2c00220.

[96] Singapore Sustainable Laboratories Group. Sustainable Laboratories, 2020. URL https://www.seas.org.sg/sustainablelaboratories. Last accessed October 13, 2023.

[97] L. N. Soldatova and R. D. King. An ontology of scientific experiments. *Journal of the Royal Society Interface*, 3(11):795–803, 2006. ISSN 1742-5662. doi:10.1098/rsif.2006.0134.

[98] L. N. Soldatova, A. Clare, A. Sparkes, and R. D. King. An ontology for a Robot Scientist. *Bioinformatics*, 22(14):464–471, 2006. ISSN 1367-4811. doi:10.1093/bioinformatics/btl207.

[99] A. Sparkes, W. Aubrey, E. Byrne, A. Clare, M. N. Khan, M. Liakata, M. Markham, J. Rowland, L. N. Soldatova, K. E. Whelan, M. Young, and R. D. King. Towards Robot Scientists for autonomous scientific discovery. *Automated Experimentation*, 2(1):1–11, 2010. ISSN 1759-4499. doi:10.1186/1759-4499-2-1.

[100] E. Stach, B. DeCost, A. G. Kusne, J. Hattrick-Simpers, K. A. Brown, K. G. Reyes, J. Schrier, S. Billinge, T. Buonassisi, I. Foster, C. P. Gomes, J. M. Gregoire, A. Mehta, J. Montoya, E. Olivetti, C. Park, E. Rotenberg, S. K. Saikin, S. Smullin, V. Stanev, and B. Maruyama. Autonomous experimentation systems for materials development: A community perspective. *Matter*, 4(9):2702–2726, 2021. ISSN 2590-2385. doi:10.1016/j.matt.2021.06.036.

[101] S. Steiner, J. Wolf, S. Glatzel, A. Andreou, J. M. Granda, G. Keenan, T. Hinkley, G. Aragon-Camarasa, P. J. Kitson, D. Angelone, and L. Cronin. Organic synthesis in a modular robotic system driven by a chemical programming language. *Science*, 363(6423), 2019. ISSN 1095-9203. doi:10.1126/science.aav2211.

[102] S. Stier. Open Semantic Lab. URL https://github.com/OpenSemanticLab. Last accessed October 13, 2023.

[103] M. Trott, K. Robinson, B. McDowall, R. J. MacKenzie, M. Campbell, and S. Chipper-Keating. *Lab of the Future*. 2022. eBook via Technology Networks.

[104] A. Umbrico, A. Orlandini, and A. Cesta. An ontology for human-robot collaboration. *Procedia CIRP*, 93(March):1097–1102, 2020. ISSN 2212-8271. doi:10.1016/j.procir.2020.04.045.

[105] R. Vinuesa, H. Azizpour, M. Leite, I.and Balaam, V. Dignum, S. Domisch, A. Felländer, S. D. Langhans, M. Tegmark, and F. Fuso Nerini. The role of artificial intelligence in achieving the sustainable development goals. *Nature Communications*, 11(1):233, 2020. ISSN 2041-1723. doi:10.1038/s41467-019-14108-y.

[106] A. A. Volk, R. W. Epps, D. T. Yonemoto, B. S. Masters, F. N. Castellano, K. G. Reyes, and M. Abolhasani. AlphaFlow: autonomous discovery and optimization of multi-step chemistry using a self-driven fluidic lab guided by reinforcement learning. *Nature Communications*, 14(1):1–16, 2023. ISSN 2041-1723. doi:10.1038/s41467-023-37139-y.

[107] J. M. Weber, Z. Guo, C. Zhang, A. M. Schweidtmann, and A. A. Lapkin. Chemical data intelligence for sustainable chemistry. *Chemical Society Reviews*, 50(21): 12013–12036, 2021. ISSN 1460-4744. doi:10.1039/d1cs00477h.

[108] Y. Xie, K. Sattari, C. Zhang, and J. Lin. Toward autonomous laboratories: Convergence of artificial intelligence and experimental automation. *Progress in Materials Science*, 132(December 2021):101043, 2023. ISSN 0079-6425. doi:10.1016/j.pmatsci.2022.101043.

[109] H. Yang, J. Li, K. Z. Lim, C. Pan, T. Van Truong, Q. Wang, K. Li, S. Li, X. Xiao, M. Ding, T. Chen, X. Liu, Q. Xie, P. V. y. Alvarado, X. Wang, and P. Y. Chen. Automatic strain sensor design via active learning and data augmentation for soft machines. *Nature Machine Intelligence*, 4(1):84–94, 2022. ISSN 2522-5839. doi:10.1038/s42256-021-00434-8.

[110] X. Zhou, S. Zhang, M. Agarwal, J. Akroyd, S. Mosbach, and M. Kraft. Marie and BERT – A Knowledge Graph Embedding Based Question Answering System for Chemistry. *ACS Omega*, 8(36):33039–33057, 2023. ISSN 2470-1343. doi:10.1021/acsomega.3c05114.

[111] Q. Zhu, F. Zhang, Y. Huang, H. Xiao, L. Y. Zhao, X. C. Zhang, T. Song, X. S. Tang, X. Li, G. He, B. C. Chong, J. Y. Zhou, Y. H. Zhang, B. Zhang, J. Q. Cao, M. Luo, S. Wang, G. L. Ye, W. J. Zhang, X. Chen, S. Cong, D. Zhou, H. Li, J. Li, G. Zou, W. W. Shang, J. Jiang, and Y. Luo. An all-round AI-Chemist with a scientific mind. *National Science Review*, 9(10):nwac190, 09 2022. ISSN 2053-714X. doi:10.1093/nsr/nwac190.