Review on Virtual Laboratories in Practice

Jarkko Louhi and Juniper Tyree

Final Report for the DATA20039 Seminar on Virtual Laboratories in Natural Sciences

Introduction

Today, scientists are faced with monumental challenges such as modelling the impacts of worsening climate change and increasingly frequent extreme weather events and finding new ways to grow food and produce food sustainably. These tasks have long-surpassed the capabilities of individual research groups working in their own siloed physical laboratories.

Scientists are increasingly adopting Virtual Laboratories (VLs) to virtually replicate laboratories and experiments in their entirety, from the underlying processes to equipment and instruments and the design and steering of the experiment [29, 39]. Virtual Laboratories hold the promise of making research more open, efficient, collaborative, and accessible by pooling resources and combining the strengths of physical experiments, virtual simulations, and their interactions. In particular, VLs let several Physical Entities (PEs) and Digital Twins (DTs) interact and utilise AI and visualisation to accelerate the scientific process [29]. So how are VLs and their components being built to fulfil these goals?

This review tries to answer the question:

What is the status of Virtual Laboratories, and what challenges have to be addressed to reap their benefits for science and society?

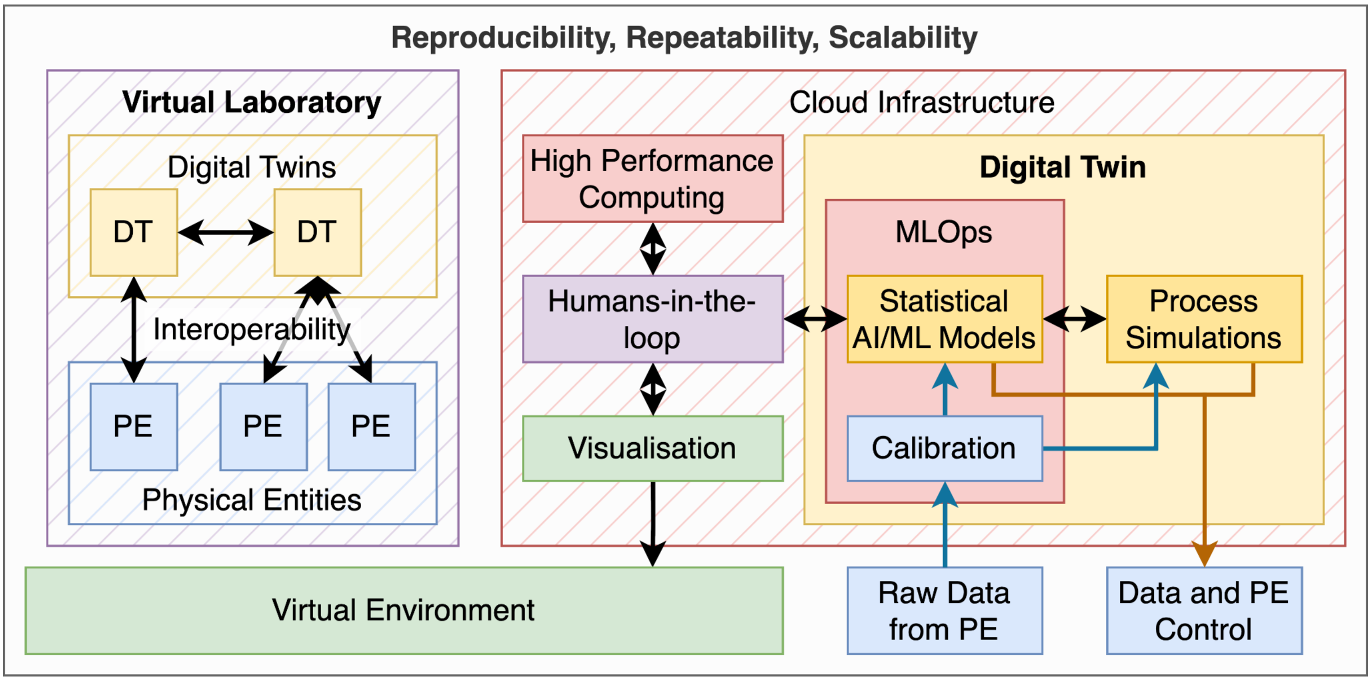

Figure 1: Overview of the architecture, building blocks, and themes of Virtual Laboratories.

In particular, this review first gives an overview of VLs and their underlying components, as sketched out in Figure 1, discusses the infrastructure and process requirements that VLs introduce and presents examples of how researchers are working towards building fully-functional VLs in practice. It also delves into the practical challenges that need to be addressed while designing VLs and discusses how scientific processes are being adapted and standardised for future scientists working with VLs. Throughout, it also explores the key challenges and open questions for VLs in future research. First, we start by introducing the main components of Virtual laboratories (VLs), Digital Twins (DTs).

What are Digital Twins?

A Digital Twin (DT) is a faithful computational representation of a real-world entity or process. A DT can be characterised by its three primary components: the physical entity or process, its virtual representation and the live coupling between them to exchange information [42, 46]. While many working DTs and hundreds of scientific papers on the topic exist, formal standards have yet to be agreed upon. Several recent literature review papers have looked into how to define, create and standardise DTs, e.g. Aheleroff et *al*. (2021) [4], Jones et *al*. (2020) [28], Semeraro et *al*. (2021) [40], VanDerHorn & Mahadevan (2021) [46], and van der Valk et *al*. (2020) [47]. For instance, in van der Valk et *al*. (2020), the authors created a multi-dimensional taxonomy for DTs based on an examination of 233 papers [47], identifying the following vital characteristics of a DT that the majority of (but not all) definitions agreed upon: containing both Human-Machine and Machine-to-Machine interfaces, having a bi-directional data link that automatically exchanges data between multiple sources, processing of raw and pre-processed data, and being bound to a Physical Entity (PE) of which an accurate model is provided.

What separates a Digital Twin from a simulation? The terms Digital Model (DM), Digital Shadow (DS) and Digital Twin (DT) are often misinterpreted as referring to the same concept. However, the difference is how information is transferred between the physical and virtual worlds [4]. In particular, a DM does not have any real-time links, a DS only has a real-time uni-directional link from the physical to the digital domain, but a Digital Twin must provide bi-directional real-time data flow that links the digital to the physical domain [4], e.g. by continuously calibrating a simulation with live measurements, whose output is then used to control the experiment.

What are Virtual Laboratories?

Virtual Laboratories (VLs) are digital, interactive representations of everything that otherwise occurs in a physical laboratory setting. Hence, they provide a readily usable framework to design and perform varying experiments in, just like a physical real-world laboratory does. VLs replicate a laboratory and its experiments in virtual space in their entirety by modelling the physical environments (see Appendix B), processes and instruments. Figures 1+2 show that VLs consist of many interconnected DTs, data, UI, and ML/AI.

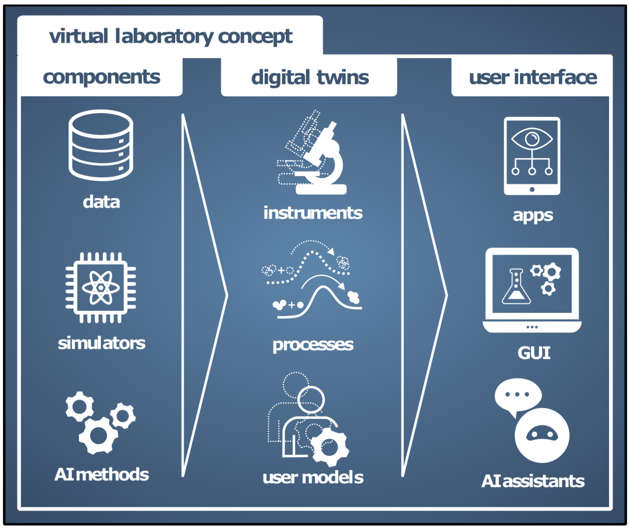

Figure 2: Elements of a Virtual Laboratory. From: Figure 2, Klami et *al*. (2022, preprint) [29]. Licensed under CC BY-NC-SA 4.0.

An essential distinguishing element between a VL and a Digital Twin is that fully functional VLs extend beyond a single domain-specific Digital Twin and instead connect several Digital Twins on a digital ecosystem platform [34]. VLs are also distinct from simulations as their Digital Twins automatically communicate bi-directionally with their physical counterparts.

Via the user interface, which can be a GUI or user-built application that utilises APIs for the Virtual Laboratory framework, the VL's user should be able to interact with information and visualisations, models and simulations, and conduct experiments with different scenarios and forecasts. Data sources for VLs can include traditional databases or live measurements, e.g., observations, GIS, video streams, and measurement devices such as IoT. Preferably, the users should be able to integrate their own data into the VL platform [34]. In addition, using AI-based methods is seen as a crucial element in VLs [29]. A Virtual Laboratory can also optionally integrate a remote, hardware-based laboratory and have remote control and access to actuators in the physical facility [4], going even beyond the bi-directional data exchange requirement.

Some literature, such as Budai et *al*. (2018) [13], defines a virtual laboratory much more generously as just a software-simulated laboratory. In this report, we instead use the strict definition inspired by Klami et *al*. (2022) [29] and Aheleroff et *al*. (2021) [4] that includes the bidirectional data exchange, coupling of several Digital Twins, and usability as a platform to design and conduct novel experiments. We do adopt this stricter definition to follow the evolution of VLs from a historically literal definition (please refer to Appendix A for a brief history of the development of the first virtual laboratories), over a vaguely defined buzzword, to a more clearly defined architecture that has strict criteria which inform where future work is still required to fully realise the potential of VLs. To avoid confusion, we refer to VLs following our definition as capital-case Virtual Laboratories, whilst using lower-case virtual laboratories elsewhere.

Case Studies on the path to fully-functional Virtual Laboratories

What could a fully-functional Virtual Laboratory look like? What architectures could they use? To answer these questions, we first give a higher-level overview of virtual laboratory designs by highlighting three examples on different scales: CROP, WetLands, and Destination Earth.

CROP: The Crop Research Observation Platform

The world's population is growing and increasingly concentrated in urban areas, which has increased the need for more efficient and sustainable food production that is closer to consumers. The "Growing Underground" project in London demonstrates how underground hydroponic farms could help provide more urban farming [25]. In an old Underground tunnel beneath Clapham, many hydroponic arrays grow plants like herbs and salad greens [20, 48].

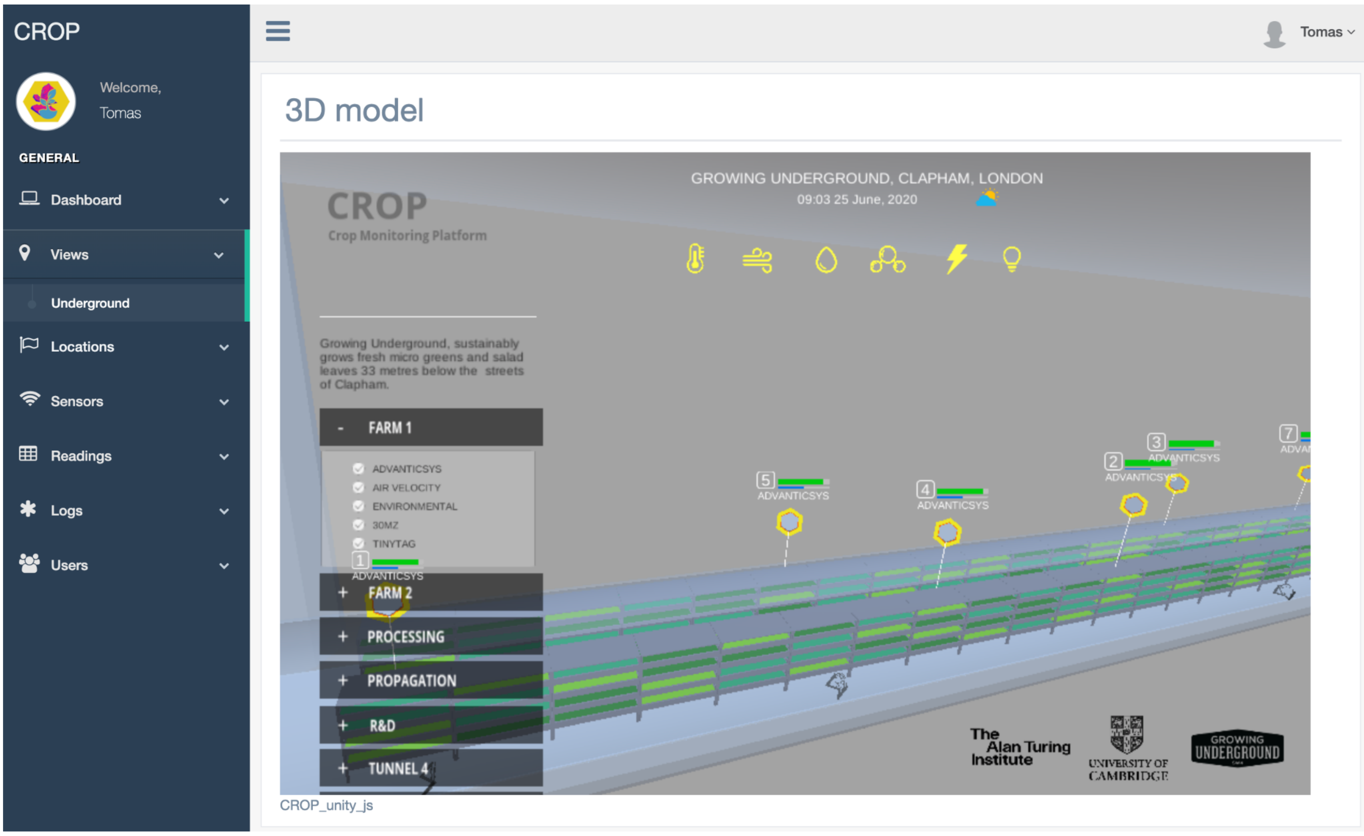

Figure 3: Screenshot of the CROP Monitoring Platform Webapp. From: Alan Turing Institute. (2020) [5]. Screenshot by Tomas Lazauskas. Licensed under the MIT License. Available from: https://github.com/alan-turing-institute/CROP/blob/bcfaccc/media/webapp.png
[Accessed: 4th December 2022]

CROP is a Digital Twin built to optimise the operation of this underground farm, which is still being actively developed for ongoing research [5]. All around the farm, sensors collect data in near-real-time on environmental variables such as temperature, humidity, and air speed, as well as the farm's energy consumption and its live crop growth and yield. These heterogeneous data points are then assimilated in ARIMA, a statistical moving average model, and the Greenhouse Energy Simulation, a physics-based forecasting model [5]. These models, fed with live data, are used to explore what-if scenarios to optimise the placement of crops, the growing conditions and the energy efficiency of the farm by comparing the farm's past and predicted performance.

CROP is fundamentally built as a cloud-based Digital Twin that combines various streams of heterogeneous data sources. It uses serverless Azure functions to connect to the API endpoints of external systems, including Stark ID (energy analytics) and OpenWeather, and collect data from these services into a Postgres database, where the results of scenario forecasts executed on Azure functions are stored as well. All this information is then presented to the end user, the operators of the underground farm, in an HTML5-based web application that visualises the entire farm and its sensors using the Unity 3D game engine. An example screenshot of this web application is shown in Figure 3. All components of CROP are containerised with Docker for easier deployment [5].

The CROP Digital Twin highlights how coupling a virtual model representation to its physical system can integrate heterogeneous data streams in real-time and provide actionable feedback to operate the physical system more efficiently.

WetLands

Wetlands are ecosystems that provide many ecosystem services. For instance, they are important carbon sinks, filter stormwater, and can act as flooding control. In New Zealand, several wetlands have been manually constructed for these reasons. To benefit from these wetlands, they require continuous inspection and maintenance, particularly after heavy rainfall.

WetLands is an industry project that aims to survey these constructed wetlands, e.g. to schedule their maintenance [2, 3, 4]. As real-life wetlands are complex interconnected systems, modelling them using a single monolithic Digital Twin would be insufficient. Instead, the WetLands project deals with this complexity by using a network of communicating Digital Twins, which are provided in a DT-as-a-Service architecture. Each Digital Twin combines weather information from UBIMET, several data sources from across its wetland, such as the water level and various sediment measurements, and is in control of the local pumps. Additionally, all Digital Twins are connected to share water exchange information.

Similarly to CROP, the WetLands Digital Twins fall into the category of predictive Digital Twins, as they are characterised by the bi-directional data exchange between physical sensors, the Digital Twin, and physical actuators such as pumps [4]. This communication is facilitated through the commercial ThingWorx platform, which also provides analytics and machine learning for the Digital Twins. The collected and analysed information about the wetlands is finally presented to different end users in different media. While the operators primarily interact with a typical dashboard, maintenance staff and citizen scientists can access an augmented reality visualisation[1] to monitor their local wetlands with real-time information [2, 3, 4].

The WetLands project highlights how a divide-and-conquer architecture can model both local and system complexity by replicating the structure of the natural world in the digital domain.

Destination Earth

Destination Earth (DestinE) is an ambitious collaboration that aims to create a high-precision, interactive, digital and dynamic simulation model of planet Earth and its systems [17, 18]. This future platform, whose development is funded by the European Union, is aimed at supporting EU evidence-based policy-making by providing decision-makers with even more detailed

[1] Unfortunately, no exemplary diagrams or pictures, e.g. of the augmented reality visualisation of the Digital Twins, can be shared in this report. Aheleroff et *al*. [2, 3, 4] have published their work on WetLands under either closed or no-derivatives open-access licences, thereby going against the spirit of collaborative science in VLs.

predictions on the impact of climate change, the forecasting of natural hazards, and the most effective measures to protect biodiversity in ecosystems.

DestinE is currently still in the preparatory phase. The first core services of the envisioned platform should come around 2024, alongside a showcase of the first two Digital Twins built on them, which will focus on extreme weather events and climate change adaptation. However, the project aims to provide a full replica of the earth's systems by 2030. Overall, the largest future impact of the Destination Earth collaboration might be its drive towards standardising software, model, and data formats, which would greatly simplify future Virtual Laboratory collaborations.

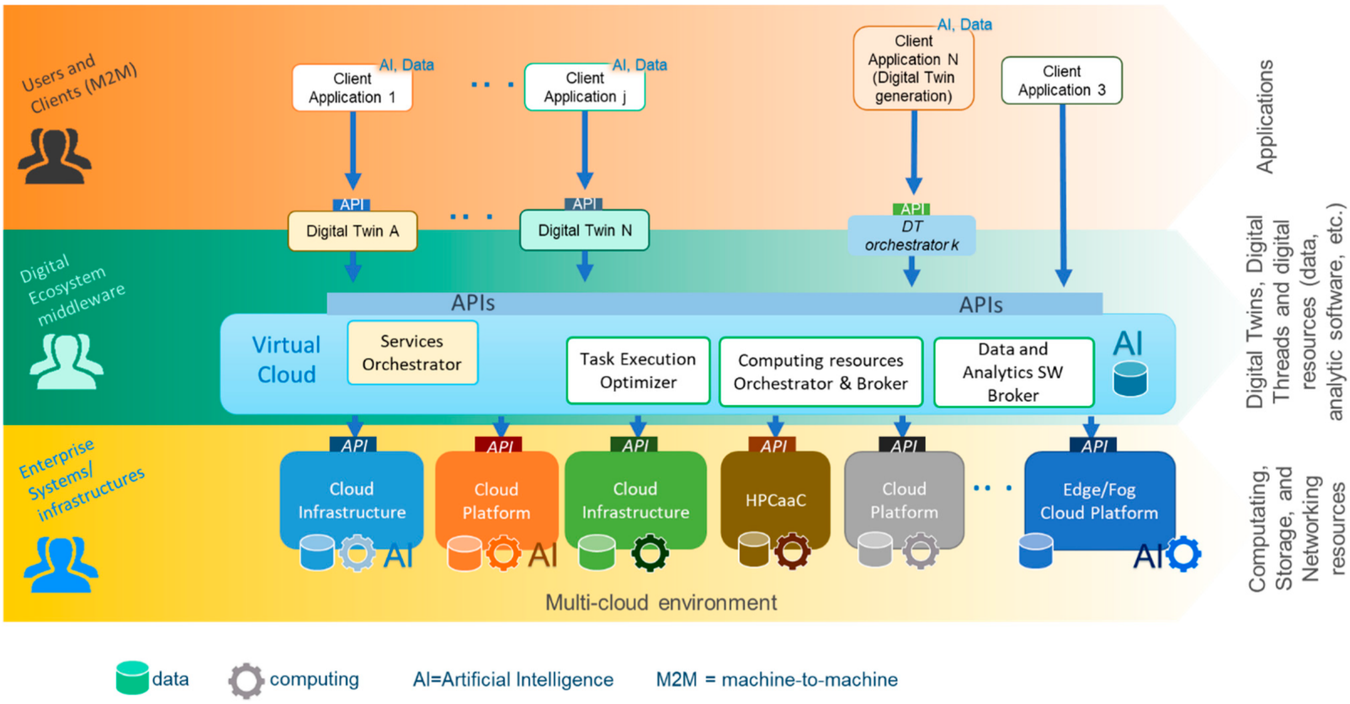

Figure 4: DestinE digital ecosystem & virtual cloud architecture overview. From: Figure 10, Nativi, Mazzetti & Craglia. (2021) [34]. Licensed under CC BY 4.0.

Comparison of the three Case Studies

Since this report aims to give an overview of Virtual Laboratories in practice, an important question is: *Are the presented case studies examples of fully-functional Virtual Laboratories?* The following table summarises the three examples, CROP, WetLands, and Destination Earth, and the virtual laboratory components that they consist of:

Components	CROP [25, 29]	WetLands [2, 3, 4]	DestinE [17,18,34]
Type	Digital Twin	Digital Twin as a Service	VL Digital Ecosystem
Scope	Farm-specific	Domain-specific	Cross-domain + European
Digital Twin Integration	Isolated Digital Twin	Network of identical Digital Twins	Integration of various Digital Twins
Data Connection	Bi-directional	Bi-directional	DT-specific, mostly uni-directional
HPC and cloud	Azure	ThingWorx	EuroHPC
ML/AI	classical models	models and analytics	DT-specific emulators
User Interface	Web app, 3D	Dashboard, 3D AR	Client-application specific
Status	Ongoing	Completed	In planning, by 2024-2030

The CROP and WetLands projects are clearly lacking some important characteristics towards being Virtual Laboratories, as they are isolated and domain-specific Digital Twins. In particular, they lack coupling with other Digital Twins and are not platforms that support conducting new experiments. However, both of these examples showcase some important steps towards a full VL and can be taken as inspiration in their areas of focus. CROP showcases how predictive experiments conducted using the coupled Digital Twins fed from almost real-time data can be used to automatically steer an underground farm to optimise its food growth and yield. WetLands highlights how several instances of a Digital Twin model can be coupled to naturally represent the interconnected complexity of the real-world constructed wetlands in New Zealand.

Destination Earth, on the other hand, will one day become a Virtual Laboratory and a digital ecosystem platform to build new Digital Twins and smaller-scope virtual laboratories. While it is still only a concept at the time of writing, the initial Digital Twins and core platform services in 2024 should already showcase the path towards fully functional Virtual Laboratories. By 2030, Destination Earth would be (one of) the first full Virtual Laboratories, as defined by [29] and [34]. Furthermore, this project and its large scope and funding will be crucial in shaping the standardisation of Virtual Laboratory design to better facilitate the collaboration of different research groups across different disciplines in the future.

Virtual Laboratory Design Space — Practical Challenges

Following on from reviewing three example projects on the path towards Virtual Laboratories, this next chapter delves into the design challenges for Virtual Laboratories in more detail. It gives an overview of the design space, prior work, and the need for future research.

Today, many research groups still build small isolated DTs restricted to a specific research domain [29, 39, 42] and use non-standard technology stacks. Moreover, Niederer et *al*. (2021) state that no cross-domain standards exist to validate DTs and assess their credibility [35]. Due to this, numerical accuracy, stability, uncertainty quantification and propagation tests vary. Hence, a key challenge for VLs lies in standardising formats, protocols and procedures for sharing and evaluating models, software, and data such that they can be easily shared and combined across research groups and disciplines [31, 34, 39].

If the infrastructure of VLs is instead designed from the ground up to support collaborative science, they can enable “acceleration, reproducibility, and scalability of research” [29]. VLs might then allow conducting more complex experiments that would otherwise be too risky or costly with physical equipment. Hence, VL components must be specifically built to keep humans-in-the-loop as active collaborators [29], work on decentralised infrastructure with lots of inter-communication [31, 43, and provide full reproducibility, repeatability and uncertainty quantification [39]. However, while research on individual DTs is well-established, exploring different software infrastructures [43], scientific protocols, and the evolving roles of human scientists in future Virtual Laboratories is still relatively new [29].

Reproducibility, Repeatability, and Scalability

Quintessentially, VLs are all about reproducibility, repeatability, and scalability. Reproducibility and repeatability are fundamental requirements to ensure that researchers can validate, repeat and share the results of a given experiment [14]. Reproducibility is preliminarily achieved by shared data, tools and models, controlled environments, and setups. A controlled technical environment consists of fully specified software and data versions and metadata that are archived alongside any results using version control.

Additionally, shared experiment protocols are also needed. Protocols define data requirements (such as metadata formats and data versions), format (for example, units of measurement, identification of no data values, significant observation period), experiment execution (e.g. selection of a well-documented model code) and outcome analysis (e.g. criteria for judging model performances). While protocol flexibility allows scientists to bring in their personal expert knowledge, not documenting every deviation from the protocol damages reproducibility [14].

Finally, scalability is needed in environments where multiple concurrent users perform research. For instance, when several scientists need to access the same Virtual Laboratory experiments and run different (varying) instances, scalability can be achieved via distributed infrastructures, such as cloud environments with scalable computing and storage, and professional architecture design and orchestration [34].

Interoperable software, models and data

Standard tools and best practice frameworks for optimisation, finite-element methods and Machine Learning exist. Standardisation and interoperability are also particularly needed in scientific research to make collaboration across research groups and disciplines easier, a core aim of Virtual Laboratories. However, scientific software is still often domain-specific, e.g. in DTs [35], and uses custom software stacks and data formats [42]. Moreover, interoperability, security, and hardware performance are often afterthoughts.

Intersect-SDK [43] tries to answer these challenges with a microservices-based, hierarchical system-of-systems infrastructure for Virtual Laboratories. It aims to make containerised components with plugins re-usable and use adapters between specific instruments and standardised APIs and adapters abstracting over specific compute hardware, e.g. GPUs and HPC. Furthermore, components are decoupled and communicate through publish-subscribe messaging. Intersect-SDK also introduces a scientific orchestrator component to guide experiments with domain- and experiment-specific configurations. The project also utilised industry-standard agile principles (DevSecOps) for software design.

In addition to interoperable software, collaborative science also needs standards to share and compare machine learning models and data [42]. One proposal towards sharing models is ONNX (Open Neural Network Exchange), an open, interoperable format to share models between different frameworks, compilers and hardware. For data sharing, the 'FAIR Guiding Principles for scientific data management and stewardship' were published in 2016 [50]. The FAIR data principles emphasise that data should be findable (has rich metadata with a unique and persistent identifier that is searchably indexed), accessible (by a standard and freely implementable protocol that supports authentication), interoperable (uses formal knowledge representation vocabulary and links to other data by their IDs) and re-usable (has a clear and accessible data use licence and records data provenance) [44]. One ambitious example of sharing data is INSPIRE, a European directive for creating a spatial data infrastructure (SDI) and sharing spatial information across the EU to support science and inform policy-making [21, 33]. The INSPIRE directive was passed in 2007 and has since been implemented [21].

For Virtual Laboratories to fulfil their vision of bringing scientists together in a virtual space that fosters accessible and open collaboration, future work on standardising existing scientific software and making them interoperable in VLs will be crucial. However, demand will likely continue to drive more domain-specific innovation and development. Hence, large projects such as Destination Earth, the VL digital ecosystem to model all Earth systems, or ITER, the nuclear fusion reactor project being built by China, the EU, India, Japan, Russia, South Korea, and the US, provide opportunities and funding to surmount these inter-disciplinary challenges.

Digital Twin communication and calibration

A crucial aspect of Virtual Laboratories that separates them from simulations is the bidirectional communication between Digital Twins and their physical representations. This grounding in and calibration with real-world data is especially critical as more research is conducted virtually.

Communication between a Digital Twin (DT) and its Physical Entity (PE) is called intra-twin communication [31]. Raw data is transferred from the PE to the DT, and processed information flows back from the DT to the PE. These data channels must be private and protected and require synchronisation, especially if the DT gives control feedback. The DT can be in the cloud, on the edge, or on the PE itself. Inter-twin communication (between two or more DTs) is more flexible but also sensitive to the availability of an Internet connection between all DTs [16]. One solution for more stable inter-twin communication is an IoT data aggregation and analytics cloud platform called ThingsSpeak. It offers to be an intermediary for inter-twin communication and uses MQTT, a lightweight publish-subscribe IoT protocol.

DTs also need to be calibrated to maintain their model's accuracy over time by integrating new information. Calibration can be accomplished with live data in two phases. In the offline phase, a collection of simulated and real-world data pairs are gathered to a data store to find the differences between both data sets. These error data pairs are then used in offline training of many specialised calibration neural networks to reduce the error between simulations and the real world. In the online phase, predictions are constructed from both the simulation and calibration networks, which are continuously retrained in the background with new data [41].

Another approach for DT calibration is continuous Bayesian calibration (BC), which incorporates prior information about the problem domain. It is an incremental, approximate method to sample the calibration posterior. Posterior samples are refiltered once new observations become available. For instance, Ward et *al*. (2021) calibrated the CROP DT with continuous Bayesian calibration and achieved promising results compared to a static model [49].

MLOps — Machine Learning Operations

VLs combine DTs using vast data and various Machine Learning and Artificial Intelligence (ML/AI) methods, requiring standardised interfaces and workflows. VLs will also increasingly complement or replace physical experiments, necessitating reproducibility and repeatability. In addition, VLs exceed the scale of single research groups, which further requires easy scalability.

Machine Learning Operations (MLOps) is a collection of techniques, tools, and processes aimed at industrialising ML/AI. The framework's main aim is to deploy and maintain ML models in production reliably and efficiently [30]. Systematic ML development, delivery, and monitoring can support thousands of models. The primary benefits of MLOps are efficiency, scalability, and risk reduction [30, 38]. Figure C.1 presents the architecture and phases of the MLOps pipeline.

MLOps will be essential to building Virtual Laboratories to achieve their required reproducibility, repeatability and scalability. There is also an opportunity for further combining the two in a ScienceOps field with open-access, transparent, federated, collaborative and fine-grained attribution. However, the fact that hardly any literature on MLOps specifically for Virtual Laboratories exists highlights that there is a danger of haphazardly designing VLs instead of employing existing software design, deployment, and maintenance expertise.

Big Data Visualisation

Visualisation is a crucial component of Virtual Laboratories, allowing researchers to explore and communicate data more intuitively. With the increasing amount, dimensionality and complexity of data, the use of large screens and VR has increased. These devices can support three different modes of visualisation: (1) the immersive exploration, typically by a single person, (2) the presentation of results to a smaller group, and (3) the interactive investigation of what-if questions with a larger team [45]. Since such visualisations require more and more computing power and faster access to larger data, there is an increasing need for the visualisation to be computed on HPC as well. By computing it as close to the original data source, costly data movement is avoided, and existing infrastructure can be used. Thus, more complex visualisation becomes more accessible to any research group with a good internet connection to an HPC cluster. Such visualisation can be used to visualise results after the fact *and* while a virtual experiment is running. The following sections on Humans in the Loop discuss this idea further.

Humans in the Loop

Today, the increasing amount of data produced in physical and Virtual Laboratories that need to be analysed is far outpacing human processing speeds and our emotion-dependent performance. However, human researchers cannot always be replaced by Machine Learning. While ML/AI has improved remarkably, it still does not match the complex reasoning capabilities of humans. Thus, it is beneficial for scientists to collaborate with AI to harness both a computer's speed and human reasoning, i.e. combining the different skill sets and levels [11, 22].

Human-in-the-loop is a concept that encompasses any activity that requires continuous human input. Flight simulators are a great example of this principle, where a computer performs a complex simulation that is directed by the human pilot's input. Keeping scientists in the loop in Virtual Laboratories brings a multitude of benefits, such as more control over HPC and AI. It also makes decision-making more attributable as both humans and machines are involved.

HPC with Humans in the Loop

The quantity and complexity of big data mean that data analysis today exceeds the processing capacities of a single work machine over a few seconds and thus has to be asynchronous. However, popular tools for data science, such as Jupyter notebooks, still rely on a synchronous processing interface, where code is written for a single processor, not an HPC cluster.

Kale is a tool that was developed to overcome this issue and enable human-in-the-loop interactivity with HPC workflows using the familiar and intuitive Jupyter interface [15]. It provides an HPC job configuration, submission, and monitoring interface that integrates directly into the notebook. Just like single-machine processes can be configured with code and monitored from the notebook, long-running tasks such as machine learning can now be launched from code and monitored in the notebook itself without requiring deep knowledge of the underlying HPC scheduling system. Kale also allows users to pause an asynchronous job at any point, modify its parameters, and then resume it. Lengthy processes such as hyperparameter optimisation can thus be controlled interactively, redirected to a more interesting direction, and stopped early to save processing cycles and energy. Overall, this ensures scientists can make better use of their Virtual Laboratory's resources and obtain a deeper understanding of its processes.

AI with Humans in the Loop

Keeping scientists in the loop not only allows them to take greater control of HPC resources used for e.g. training machine learning models, it also improves the trained models themselves, e.g. by utilising interactive machine learning. IML is an iterative training process where (1) the model is partially trained, (2) the model visually presents partial results to the user, (3) the user provides feedback based on their background knowledge and understanding of the task, and (4) the process repeats with the training the model after integrating the user's feedback [26].

What are the benefits of this interactive process? First, as a human scientist is involved throughout the model training, the model itself and its performance are more transparent. Furthermore, human scientists have existing background knowledge and still perform better at detecting outliers and overfitting. Thus, they can guide the model to become more robust and generalisable. Finally, including a human provides an extra layer of safety for making critical decisions, e.g. when human or environmental safety is involved.

Combining the skills of humans and ML/AI has many different applications. For instance, humans can help ML by guiding the generation of informative features, labelling new examples that the model has deemed critical for its learning process, and tweaking the model parameters to minimise its error residuals. Additionally, ML can find interesting patterns to show to the user in Exploratory Data Analysis or collaborate in identifying a useful projection in interactive dimensionality reduction by tweaking individual point projections or adding constraints. Overall, these examples highlight the manifold benefits of collaborating with AI in a Virtual Laboratory.

Future Vision and Conclusion

In this report, we have defined Virtual Laboratories as platforms for interconnected Digital Twins, which provide researchers with the tools to perform experiments virtually, couple Digital Twins to real-world experiments to more efficiently control their trajectory, and prototype new experiments without the constraints of having to build them physically first. Thus, Virtual Laboratories open up possibilities for designing much riskier experiments. We also highlighted three examples showcasing the path leading to the first fully functional VL. Next, we presented an overview of the design space of VLs and how their different design challenges have been approached thus far. In particular, standardising how software, data and models can be exchanged and how Digital Twins should communicate stand out as significant challenges and open questions in the field, for which MLOps offers some existing solutions. These challenges must be overcome to ensure that Virtual Laboratories can reach across research groups and disciplines to provide a more accessible and collaborative platform that is built for open, reproducible, and repeatable science, helping us produce food, model climate change, and protect our ecosystems.

How will Virtual Laboratories change the field of science, the protocols in use, and the roles of scientists themselves? This is perhaps the most fundamental open question on the path to VLs. We now want to present our own vision for the future role of scientists in Virtual Laboratories.

Science can gain from the collaboration between ML/AI and humans-in-the-loop. In Virtual Laboratories, which aim to make science more accessible, repeatable, reproducible, and scalable, this collaboration will be especially important. However, it will also require a shift in the role of scientists. ML/AI is increasingly good at performing initial data reviews, can discover relationships and postulate plausible theories that might explain enormous data sets. However, deriving new insights, curating and contextualising them within existing research, and explaining the new findings to other researchers and to the broader public still require human scientists [11]. Hence, ML/AI will increasingly fulfil the role of a hardworking research assistant that is especially valuable for repetitive tasks that need more homogeneous performance throughout. Crucially, this will require ML/AI models to explain their conclusions better: both how they arrived at an answer (eXplainable AI) and how certain they are (Uncertainty Quantification).

Virtual Laboratories will help standardise science. Through their worldwide accessibility, greater flexibility, and lower resource consumption VLs could revolutionise the scope of how research can be done: researchers and citizen scientists from all around the world could collaborate on designing and prototyping daring new experiments on shared VL infrastructure without the constraints of physical location or resource investment. In other words, if properly designed and built, Virtual Laboratories will offer exciting new possibilities for Virtual Scientists and our planet.

Acknowledgements

This review was written at the University of Helsinki for the “Seminar on Virtual Laboratories in Natural Sciences” course under the supervision of Jarmo Mäkelä and Kai Puolamäki. As part of that seminar course, several groups researched, presented, and reported on different aspects of virtual laboratories and provided each other with continuous feedback. Hence, we want to take this opportunity to thank our supervisors and colleagues for their valuable input. In particular, we would like to acknowledge our early discussions on this topic with Jonathan Gehret, and the thought-provoking feedback from Anna Brauer, Anu Kirjasuo, and Līva Freimane.

Licensing

Louhi, J. & Tyree, J. (2022) *Review on Virtual Laboratories in Practice*. University of Helsinki.

Reference List

[1] Agarwal, D. et *al*. (1996). Tools for Building Virtual Laboratories. *Computing in High Energy Physics '95*. Available from: doi:10.1142/9789814447188_0001 or https://www.researchgate.net/publication/2595451_Tools_for_Building_Virtual_Laboratories [Accessed: 13th December 2022]

[2] Aheleroff, S., Zhong, R. Y. & Xu, X. (2020). A Digital Twin Reference for Mass Personalization in Industry 4.0. *Procedia CIRP*, 93, 228–233. Available from: doi:10.1016/j.procir.2020.04.023

[3] Aheleroff, S. et *al*. (2020). Digital Twin Enabled Mass Personalization: A Case Study of a Smart Wetland Maintenance System. *Proceedings of the ASME 2020 15th International Manufacturing Science and Engineering Conference. Volume 2: Manufacturing Processes; Manufacturing Systems; Nano/Micro/Meso Manufacturing; Quality and Reliability*. Available from: doi:10.1115/msec2020-8363

[4] Aheleroff, S. et *al*. (2021). Digital Twin as a Service (DTaaS) in Industry 4.0: An Architecture Reference Model. *Advanced Engineering Informatics*, *47*, 101225. Available from: doi:10.1016/j.aei.2020.101225

[5] Alan Turing Institute. (2020). *CROP: a Research Observation Platform*. GitHub. Available from: https://github.com/alan-turing-institute/CROP [Accessed: 18th October 2022]

[6] Afsarmanesh, H. et *al*. (2001). A reference architecture for scientific virtual laboratories. *Future Generation Computer Systems*. Available from: doi:10.1016/S0167-739X(01)00042-5

[7] Alegria, F. et *al*. (1997). A remote controlled automated measurement system. *IEEE Instrumentation and Measurement Technology Conference Sensing, Processing, Networking*. Available from: doi:10.1109/IMTC.1997.612387

[8] Arpaia, P. et *al*. (1996). A distributed measurement laboratory on geographic network. *Proc. IMEKO 8th Symp. on Measurements and Calibration Methods on Electrical Quantities and Instruments*.

[9] Arpaia, P. et *al*. (2000). A measurement laboratory on geographic network for remote test experiments. *IEEE Transactions on Instrumentation and Measurement*, 49(5), 992–997. Available from: doi:10.1109/19.872919

[10] Aziz, A. et *al*. (2006). An Architecture For Virtual Laboratory Experimentation. *2006 Annual Conference & Exposition Proceedings*. Available from: doi:10.18260/1-2--220

[11] Blumauer, A. (2019). Explainable AI: The Rising Role Of Knowledge Scientists. *Forbes*. Available from: https://www.forbes.com/sites/forbestechcouncil/2019/12/30/explainable-ai-the-rising-role-of-knowledge-scientists/ [Accessed: 26th October 2022]

[12] Bohus, C. et *al*. (1996). Running Control Engineering Experiments Over the Internet. *IFAC Proceedings Volumes*. Available from: doi:10.1016/S1474-6670(17)58121-5

[13] Budai, T. et *al*. (2018) Towards a Modern, Integrated Virtual Laboratory System. *Acta Polytechnica Hungarica*. Available from: http://epa.niif.hu/02400/02461/00080/pdf/EPA02461_acta_polytechnica_hungarica_2018_03_191-204.pdf [Accessed: 15th December 2022]

[14] Ceola, S. et *al*. (2015) Virtual laboratories: new opportunities for collaborative water science. *Hydrol. Earth Syst. Sci.*, 19, 2101–2117. Available from: doi:10.5194/hess-19-2101-2015

[15] Cholia, S. et *al*. (2018): Kale: A System for Enabling Human-in-the-loop Interactivity in HPC Workflows. *figshare*. Available from: doi:10.6084/m9.figshare.7067075.v3

[16] Dasbach, T. et *al*. (2019). Digital Twin – Integrating Cloud Services into Communication Protocols. *IFIP Advances in Information and Communication Technology*, 283–292. Available from: doi:10.1007/978-3-030-42250-9_27

[17] Destination Earth. (2021). *Shaping Europe's Digital Future*. Available from: https://digital-strategy.ec.europa.eu/en/library/destination-earth [Accessed: 16th October 2022]

[18] Destination Earth. (2022). *Shaping Europe's Digital Future*. Available from: https://digital-strategy.ec.europa.eu/en/policies/destination-earth [Accessed: 28th October 2022]

[19] de Vries, L. E. & May, M. (2019). Virtual laboratory simulation in the education of laboratory technicians–motivation and study intensity. *Biochemistry and Molecular Biology Education*, 47(3), 257–262. Available from: doi:10.1002/bmb.21221

[20] Digital Twin of Growing Underground. *Energy Efficient Cities Initiative*. Available from: https://eeci.github.io/home/docs/projects/urbanag/digitaltwin/ [Accessed: 18th October 2022]

[21] European Commission. (n.d.). About INSPIRE. *INSPIRE KNOWLEDGE BASE*. Available from: https://inspire.ec.europa.eu/about-inspire/563 [Accessed: 15th December 2022]

[22] Falk, D. (2019). How Artificial Intelligence Is Changing Science. *Quanta magazine*. Available from: https://www.quantamagazine.org/how-artificial-intelligence-is-changing-science-20190311/ [Accessed: 26thOctober 2022]

[23] Fisher-Wilson, J. (1998). Working in a Virtual Laboratory-Advanced technology substitutes for travel for AIDS researchers. *The Scientist*, 12(24), 1.

[24] Hong, B. & Ghanavati, A. (2022), The Virtual Laboratory: A Natural Vehicle for Simulation in Engineering Education. *ASEE-NE 2022*. Available from: https://strategy.asee.org/42213 [Accessed: 30th October 2022]

[25] Jans-Singh, M. (2021). Digital Twin of Growing Underground. *Energy Efficient Cities Initiative*. Available from: https://eeci.github.io/home/docs/projects/urbanag/digitaltwin/ [Accessed: 18th October 2022]

[26] Jiang, L., Liu, S. & Chen, C. (2018). Recent research advances on interactive machine learning. *Journal of Visualization*, 22(2), 401–417. Available from: doi:10.1007/s12650-018-0531-1

[27] Johnston, W. & Agarwal, D. (1995). The virtual laboratory: Using networks to enable widely distributed collaboratory science. *A NSF Workshop Virtual Laboratory whitepaper*. Ernest Orlando Lawrence Berkeley National Laboratory, Report LBL-37466.

[28] Jones, D. et *al*. (2020). Characterising the Digital Twin: A systematic literature review. *CIRP Journal of Manufacturing Science and Technology*, 29, 36–52. Available from: doi:10.1016/j.cirpj.2020.02.002

[29] Klami, A. et *al*. (2022, preprint). Virtual Laboratories: Transforming research with AI. *TechRxiv*. Available from: doi:10.36227/techrxiv.20412540.v1

[30] Kreuzberger, D. et *al*. (2022). *Machine Learning Operations (MLOps): Overview, Definition, and Architecture*. Available from: doi:10.48550/arXiv.2205.02302

[31] Luan, T. H. et *al*. (2021, preprint). The Paradigm of Digital Twin Communications. *arXiv*. Available from: doi:10.48550/arXiv.2105.07182

[32] Machotka, J. et *al*. (2011). The history of developments of remote experiments. *2nd World Conference on Technology and Engineering Education*. Available from: http://www.wiete.com.au/conferences/2wctee/papers/17-12-Machotka-J.pdf [Accessed: 13th December 2022]

[33] Minghini, M. et *al*. (2021). INSPIRE: The Entry Point to Europe's Big Geospatial Data Infrastructure. *Handbook of Big Geospatial Data*, 619–641. Available from: doi:10.1007/978-3-030-55462-0_24 and doi:10.48550/arXiv.2105.12228

[34] Nativi, S., Mazzetti, P. & Craglia, M. (2021). Digital Ecosystems for Developing Digital Twins of the Earth: The Destination Earth Case. *Remote Sensing*, 13(11), 2119. Available from: doi:10.3390/rs13112119

[35] Niederer, S.A., Sacks, M.S., Girolami, M. *et al.* Scaling digital twins from the artisanal to the industrial. *Nat Comput Sci* 1, 313–320 (2021). Available from: doi:10.1038/s43588-021-00072-5

[36] Nimere, K. (1994). Virtual Laboratory for Disabled Students: Interactive Metaphors and Methods. *Proceedings of the Virtual Reality Conference*. California State University, Northridge Center on Disabilities.

[37] Potkonjak, V. et *al*. (2016). Virtual laboratories for education in science, technology, and engineering: A review. *Computers & Education*, 95, 309–327. Available from: doi:10.1016/j.compedu.2016.02.002

[38] Raatikainen, M. et *al*. (2022). Baseline methods and techniques for advanced model engineering (WP3 D3.1). *IML4E & University of Helsinki*. Available from: https://itea4.org/project/iml4e.html [Accessed: 20th October 2022]

[39] Ratner, D. et *al*. (2019). Roundtable on Producing and Managing Large Scientific Data with Artificial Intelligence and Machine Learning. *Office of Basic Energy Sciences*. Available from: doi:10.2172/1630823

[40] Semeraro, C. et *al*. (2021). Digital twin paradigm: A systematic literature review. *Computers in Industry*, 130, 103469. Available from: doi:10.1016/j.compind.2021.103469

[41] Song, H., Song, M., & Liu, X. (2022). Online autonomous calibration of digital twins using machine learning with application to nuclear power plants. *Applied Energy*, 326, 119995. Available from: doi:10.1016/j.apenergy.2022.119995

[42] Tao, F., & Qi, Q. (2019). Make more digital twins. *Nature*, 573(7775), 490–491. Available from: doi:10.1038/d41586-019-02849-1

[43] Thakur, A. M. et *al*. (2022). Towards a Software Development Framework for Interconnected Science Ecosystems. *ResearchGate*. Available from: https://www.researchgate.net/publication/362024652_Towards_a_Software_Development_Framework_for_Interconnected_Science_Ecosystems [Accessed: 22nd October 2022]

[44] The FAIR Data Principles. (2021). Available from: https://force11.org/info/the-fair-data-principles/ [Accessed: 23rd October 2022]

[45] Turner, M. J. et *al*. (2017). Human-in-the-Loop Visualisation Architecture for Monitoring Remote Compute. *Computer Graphics and Visual Computing (CGVC)*. Available from: doi:10.2312/cgvc.20171275

[46] VanDerHorn, E. & Mahadevan, S. (2021). Digital Twin: Generalization, characterization and implementation. *Decision Support Systems*, 145, 113524. Available from: doi:10.1016/j.dss.2021.113524

[47] van der Valk, H. et *al*. (2020). A Taxonomy of Digital Twins. *Americas Conference on Information Systems*.Available from: https://aisel.aisnet.org/amcis2020/org_transformation_is/org_transformation_is/4/ [Accessed: 30th October 2022]

[48] Walsh, L. (2021). Growing Underground: smart farming in the heart of London. *University of Cambridge*. Available from: https://www.cam.ac.uk/stories/growingunderground [Accessed: 1st December 2022]

[49] Ward, R. et *al*. (2021). Continuous calibration of a digital twin: Comparison of particle filter and Bayesian calibration approaches. *Data-Centric Engineering*, 2. Available from: doi:10.1017/dce.2021.12

[50] Wilkinson, M. D. et *al*. (2016). The FAIR Guiding Principles for scientific data management and stewardship. *Scientific Data*, 3(1). Available from: doi:10.1038/sdata.2016.18

[51] Zhang, Z. et *al*. (2018). Collaborative Virtual Laboratory Environments with Hardware in the Loop. *Cyber-Physical Laboratories in Engineering and Science Education*, 363–402. Available from: doi:10.1007/978-3-319-76935-6_15

Appendix A — A Brief History of Virtual Laboratories

The terms virtual laboratory and remote laboratory started to arise in scientific papers in the mid to late 1990s alongside the rise of the Internet. However, these laboratories did not fulfil our definition of a Virtual Laboratory as a platform linking several Digital Twins. Instead, they were focused on specific aspects, such finding solutions to remotely control experimental devices over a distance, or building collaborative tools like video conferencing and file sharing [6].

The first successful implementation of remotely controlled robots was achieved in 1994 at the University of South California in a programme called the Mercury Project [32]. The users could guide a robot to find sand-filled items over the Internet and receive video feedback. By 1996, undergraduate students were able to remotely utilise a fully operational laboratory with an application called ‘Second Best to Being There’ (SBBT) [12, 32]. The first remotely controlled laboratory, which allowed multiple users and sessions for multiple experiments, was developed for measuring semiconductor devices in 1997 [7]. In 1999, the National University of Singapore launched a virtual laboratory for oscilloscope experiments. It was one of the most excessively used remote laboratories during the time [8, 9, 32].

During the early 2000s, along with the development of faster Internet connectivity and computer systems, many educational institutes started implementing remote and virtual laboratories to allow students to conduct experiments remotely [10]. Ever since, the use of virtual laboratories has expanded, particularly in science and engineering education [19, 24, 37].

From the beginning, publications about virtual laboratories have also looked at how they could make science more accessible, e.g. in “Virtual Laboratory for Disabled Students: Interactive Metaphors and Methods” [36], “The Virtual Laboratory: Using networks to Enable Widely Distributed Collaboratory Science” [27], “Tools for Virtual Laboratories” [1], and “Working in a Virtual Laboratory - Advanced technology substitutes for travel for AIDS researchers” [23].

Appendix B — Virtual Laboratories in Virtual Environments

The user interface of VLs can extend far beyond graphs and interactive dashboards. They can also provide imaginative virtual environments that visualise the three-dimensional laboratory space and allow users to explore it [51]. Immersive VLs replicate the interactions a user would have with physical lab equipment, e.g. in the control room, by providing special inputs and feedback sensors. Furthermore, the visualisation of the experiment in a virtual environment can be coupled to a physical simulation of the space, allowing the user to perform simple physical interactions with the equipment as if they were physically there. Virtual Laboratories environments can also benefit from being interactive and allowing multiple users to collaborate (or compete) in the same space. On the other hand, single-user virtual environments can also be beneficial, particularly so for training researchers to use the equipment in a safe environment. Moreover, Virtual Laboratories can also utilise augmented reality, for instance, to overlay diagnostics performed on the Digital Twin onto the real-world experiment. The WetLands project demonstrates the potential of utilising augmented reality to visualise live data and analysis results for citizen scientists and wetland maintenance staff for their local wetlands [4]. Finally, virtual environments can also be used to integrate with virtual classrooms or conferences [10, 37, 51], which have grown in importance over the past few years.

Appendix C — The MLOps Architecture

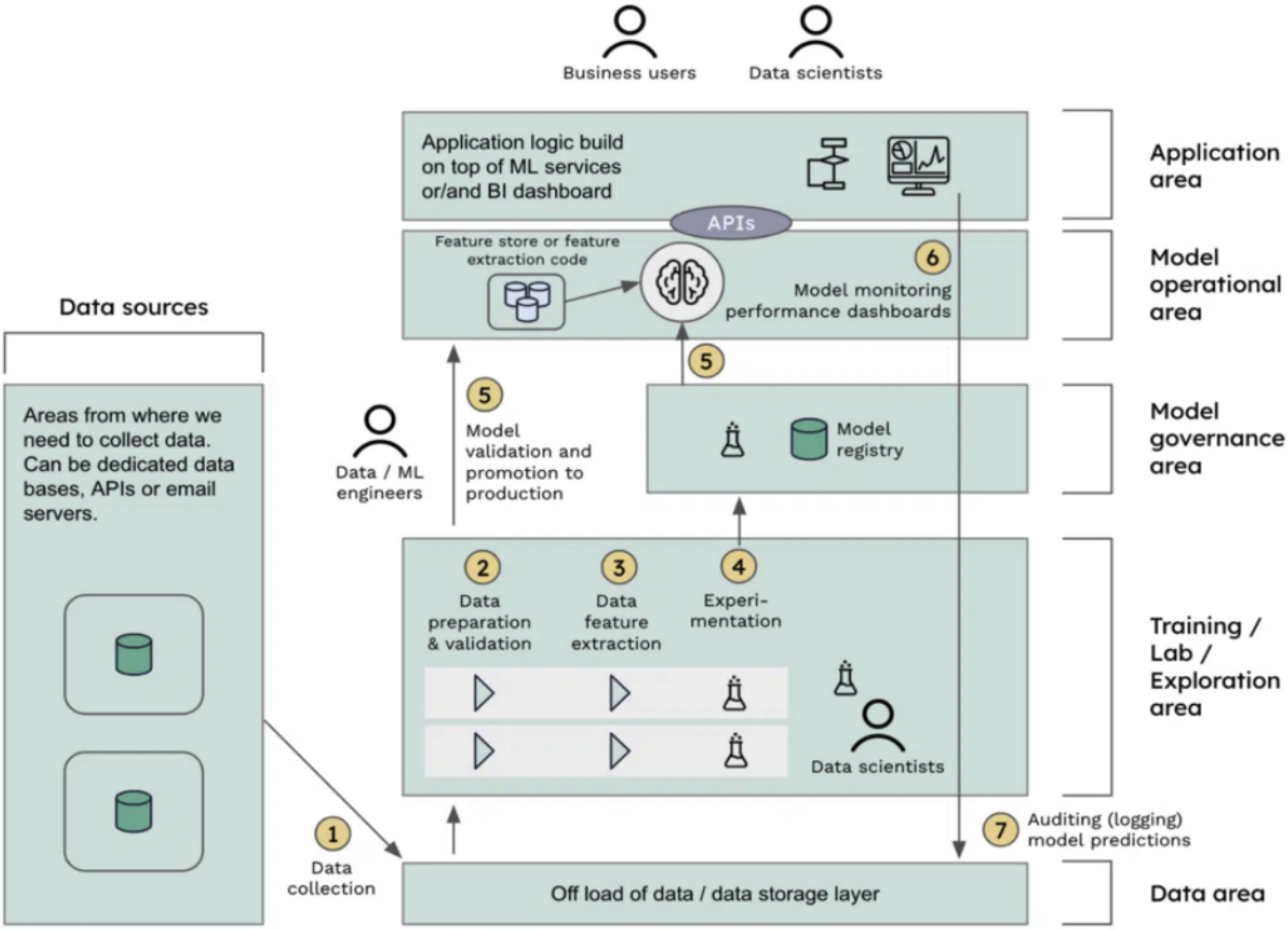

Figure C.1: MLOps pipeline architecture overview. From: Figure 6, Raatikainen, M. et *al*. (2022) [38]. Licensed under CC BY 4.0.