Preprint Cambridge Centre for Computational Chemical Engineering ISSN 1473 – 4273

The Digital Lab Manager: Automating Research Support

Simon D. Rihm[1,2,3], Yong Ren Tan[1], Wilson Ang[1], Markus Hofmeister[1,2,3],
Xinhong Deng[1], Michael Teguh Laksana[1], Hou Yee Quek[1], Jiaru Bai[2],
Laura Pascazio[1], Sim Chun Siong[1], Jethro Akroyd[1,2,4],
Sebastian Mosbach[1,2,4], Markus Kraft[1,2,4,5,6]

released: January 25, 2024

[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
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. 318

UNIVERSITY OF CAMBRIDGE

Keywords: Laboratory Automation, Lab Management, LIMS, RFID, Asset tracking, Dynamic knowledge graphs

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

Laboratory management automation is essential for achieving interoperability in the domain of experimental research and accelerating scientific discovery. The integration of resources and the sharing of knowledge across organisations enable scientific discoveries to be accelerated by increasing the productivity of laboratories, optimising funding efficiency, and addressing emerging global challenges. This paper presents a novel framework for digitalising and automating the administration of research laboratories through The World Avatar, an all-encompassing dynamic knowledge graph. This Digital Laboratory Framework serves as a flexible tool, enabling users to efficiently leverage data from diverse systems and formats without being confined to a specific software or protocol. Establishing dedicated ontologies and agents and combining them with technologies such as QR codes, RFID tags, and mobile apps, enabled us to develop modular applications that tackle some key challenges related to lab management. Here, we showcase an automated tracking and intervention system for explosive chemicals as well as an easy-to-use mobile application for asset management and information retrieval. Implementing these, we have achieved semantic linking of BIM and BMS data with laboratory inventory and chemical knowledge. Our approach can capture the crucial data points and reduce inventory processing time. All data provenance is recorded following the FAIR principles, ensuring its accessibility and interoperability.

Highlights

- A novel framework for digitalising management of research laboratories is demonstrated.
- Knowledge on equipment, chemistry, and operations are integrated via semantic web.
- Enhancing safety via automated tracking of explosive precursors.
- Dynamic knowledge graphs facilitate use of unified user interfaces for managers.
- Mobile asset management app provides instant information access for lab users.

Contents

1 Introduction

Scientific advancement and breakthroughs often depend on the integration and collaboration of various systems, disciplines, and institutions [9, 25]. Conventional workflows therefore entail complex manual coordination of researchers, operations, experiments, and resources across multiple laboratories. To accelerate scientific discovery, those need to be improved. Digitalisation and automation in research laboratories are essential to increase their efficiency and productivity and minimise human error, which improves experimental reproducibility, precision, and accuracy [8, 10].

Conventionally, laboratory automation in scientific research and development (R&D) has been focused on experimental automation and the development of "self-driving laboratories" – specifically to plan, execute, and optimise scientific experiments in an autonomous loop [3]. However, to fully automate processes or even achieve autonomy, a more general view is necessary that includes the influential peripheries of experimental research, often related to managerial duties [31, 37]. The ability of key stakeholders (*e.g.* laboratory managers, technicians, and principal investigators) to focus on core tasks that require sophisticated decision-making, creativity, and problem-solving is pivotal to accelerating scientific discoveries and, more importantly, addressing urgent global challenges through R&D [50].

Realising the digital transformation of R&D in science laboratories requires the adoption of cutting-edge technologies such as artificial intelligence and cloud-based platforms [10]. Automated laboratory data capture systems were recently shown to support harmonising siloed laboratory data from analytical instruments, reporting systems, and operational platforms [46]. However, the sole implementation of digital solutions alone is not sufficient as increasing data complexity and volume requires stored information to be accessible and interoperable. Unfortunately, existing digital tools are limited, isolated, and fragmented [8]. Solutions are constrained by proprietary vendor data formats and hampered by siloed storage platforms that are case-specific.

Semantic Web technology has been identified as a promising solution to address these challenges and enable interoperability of data across scales and domains [2, 14]. The World Avatar, for instance, is an ongoing project that uses the Semantic Web technology for laboratory automation (see Fig. 1) [24, 37] within a "Digital Laboratory Framework" (DLF). This is achieved through the development of a sufficiently generic and all-encompassing digital twin (DT) that reflects every aspect of the physical laboratory accurately and enables seamless orchestration of experimental setups and resources. Software agents are developed and involved in continuously incorporating new concepts and data into the knowledge graph while maintaining connections to existing data. The knowledge graph captures the data provenance of experimental procedures as "knowledge statements" when it expands, thereby functioning as a dynamic representation of the actual world.

This work is part of a series of articles that argue for holistic laboratory automation encompassing all aspects of experimental research within a dynamic knowledge graph [37]. Fig. 1 illustrates these aspects: the laboratory which provides all relevant infrastructure and resources; the moving handlers carrying out the actual work; and the experiment itself. The **purpose of this paper** is to demonstrate a network of synchronised and connected digital twins based on Semantic Web technology, that can transform and automate laboratory operations. Particularly, we aim to overcome challenges posed by the traditionally manual workflows in laboratory management, allowing key stakeholders to focus on essential activities that demand complex reasoning, innovation, and creativity. Common standards and ontologies are used to semantically represent and link laboratory data on chemical properties, asset information, users, and related real-time data in a machine-readable and interoperable format. We also present platform-independent interfaces that allow for digital visualisation and operation of the laboratory, leading to a more connected and intelligent system that enhances the efficiency, quality, and innovation of scientific research.

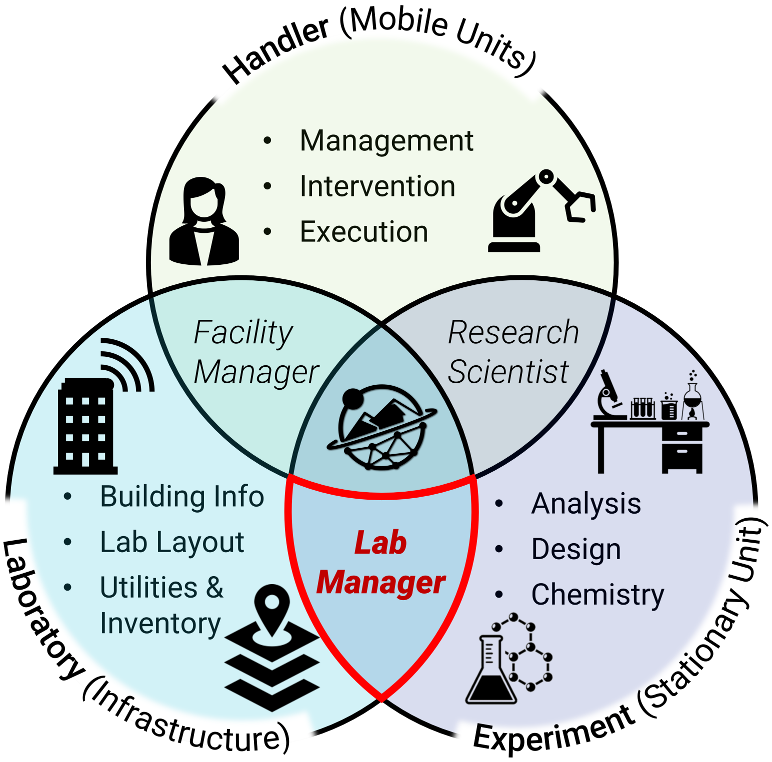


Figure 1: *An illustration of interconnected features of a scientific research laboratory that need to be represented by a comprehensive digital twin (adapted from [37]). The current work will focus on the automation of laboratory management tasks.*

2 Background

The management of research laboratories encompasses a diverse array of tasks that span many domains and scales to provide a safe and efficient work environment for researchers. Many of the tasks are of administrative nature and peripheral to the research itself. They have an enabling effect on users of lab facilities, providing relevant resources and monitoring ongoing research. To the authors' best knowledge, there is a gap in the current literature body on comprehensive reviews of lab managers' responsibilities and their potential for digitalisation and automation. We therefore identified 25 tasks in a strategic assessment of laboratory management duties at the Cambridge Centre for Advanced Research and Education in Singapore (CARES). We consolidated these tasks in Appendix A and categorised resulting problem spaces into asset management, inventory tracking, and resource allocation. For an accurate analysis of suitability for automation, we ranked them in Fig. 2 according to the relative expected cost of automation (ordinate) and the relative expected impact with the implementation of automation (abscissa). For context, a high relative impact means a considerable improvement in productivity and/or reduction of error in the execution of a tasks is to be expected.

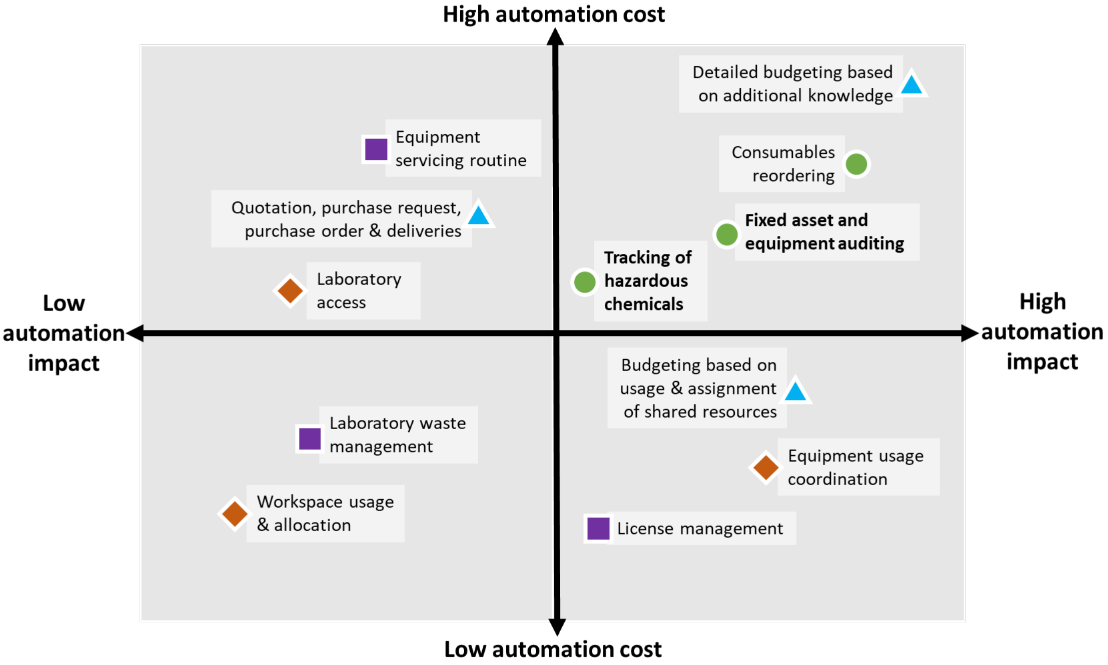


Figure 2: *Cost-impact analysis of challenges in laboratory management. The highlighted problem spaces are the focus of the current work. Legend:* ● *Inventory management;* ▲ *Procurement and finance;* ■ *Maintenance;* ◆ *Access management*

Tasks related to procurement and finance are expected to show a high impact on automation but related systems are usually predefined or highly entangled with other departments. Meanwhile, inventory management is a crucial aspect of laboratory management and operation with a consistently high expected impact of automation. These tasks are highly manual and repetitive which applies in particular to fixed asset and equipment auditing, requiring users to track items and update different systems when new assets are purchased [28]. This can lead to inconsistent and error-prone data, which can compromise the quality and integrity of the experimental metadata. Moreover, the tracking of hazardous chemicals lacks standardisation, and the conventional approach of using a physical log book is prone to human errors. Thus, automating these tasks could potentially enhance the productivity of research facilities and accelerate scientific discoveries by maximising managers' available time for supporting researchers on actual scientific research and discoveries. Whilst the implementation of automation processes using digital tools is crucial, security concerns are of utmost importance as well, especially when managing sensitive research projects. Therefore, an authentication system also needs to ensure that sensitive data or information is safe and secure [26, 33].

2.1 Digital Tools in Laboratory Management

Digital solutions exist in laboratory management to address some of the actionable problem spaces shown in Fig. 2. However, they are often fragmented, and lack interoperability and specificity [37]. This is especially evident in attempts to integrate diverse systems within a unified framework that allows smooth collaboration and communication. The integration of different software, equipment, and monitoring systems from various suppliers, as well as tracking their usage, occupancy, and maintenance schedules remains a critical challenge [28]. Many research laboratories have adopted a plethora of digital tools and platforms in the management of their labs that need to be maintained separately. This section will present some of these solutions and the challenges involved [26].

Laboratory Information Management System (LIMS) can digitalise the management of laboratory inventory such as equipment, chemicals, and consumables. It is used for automating tasks and workflows, handling data entry and import from external sources, facilitating inventory purchase and scheduling, as well as monitoring and tracking inventory levels, usage, location, and status [40]. Meanwhile, Enterprise Resource Planning (ERP) software unifies every element of corporate operations and presents a consolidated overview of the organisation on a single platform, which supports supply chain, human resources, operations, finance, and other key organisational tasks [18]. ERP gives laboratory managers access to crucial information related to stock levels, operating costs, and expenses. It can be coupled with LIMS which allows the synchronisation of internal and external laboratory activities as well as enhanced quality control and resource allocation.

An Electronic Laboratory Notebook (ELN) allows the digital collection and management of laboratory data. A typical ELN offers various functionalities such as workflow automation, documentation, data management, and collaboration tools [34]. Similarly to ELN, the Laboratory Execution System (LES) also facilitates electronic processes and automates the interaction with methods, instruments, and supplies in routine laboratory procedures to ensure compliance and quality control [40]. ELN and LIMS are often integrated to improve laboratory productivity; examples are SLIMS, LabCollector, LabWare, openBIS, and Labguru [4]. A key difference between ELNs and LIMS is that ELNs store and record unstructured research data (*e.g.* R&D stage experimental data), while LIMS stores structured and repetitive data that follow patterns and templates (*e.g.* diagnostic results from a testing lab). LES is also often integrated with LIMS, or ELN when there are parts of the LES workflow that are not supported [30, 40]. As such, LIMS-LES integration was shown to improve the quality of data acquisition and sample turnaround time [40].

The Scientific Data Management System (SDMS) is a centralised document management system, that collects, stores, and exports scientific data generated by and in a laboratory. SDMS is designed to capture, store and manage a variety of unstructured data formats integrated with systems such as LIMS, ELN, LES and ERP [26]. It also allows for direct interaction with laboratory instruments, systems, and databases. With an SDMS, laboratory managers and other laboratory personnel can generate reports on laboratory activities, speed up workflows and approval processes, and optimise collaboration using data from LIMS and ELNs.

Meanwhile, Chromatography Data Systems (CDS) are digital tools that specifically acquire, manage and report test results from experiments involving chromatography [19, 26, 27]. CDS acts as an instrument coordinator with devices to send work schedules, receive measured data, and analyse the data [19]. CDSs are useful for laboratories that perform routinised and regulated testing for quality control, pharmaceutical development, or manufacturing [27]. Mazzarese et al. [27] provided a review of the main CDS providers with some smaller or niche market CDS providers that are often linked to the products of the chromatography products.

Typically, all these software can be classified into commercial and open source. Commercial models are usually platform-based and popular among large corporations because they are particularly designed to be good at managing high-throughput data and preventing data loss. They offer a well-organised support system, frequent maintenance and upgrades, regulatory compliance, and smooth integration capabilities. They do, however, incur significant costs owing to licensing, implementation, customisation, and technical training for users [34]. Open-sourced software is therefore more attractive to small and medium-sized businesses and academic institutions. These solutions have little or no licencing, maintenance, and external service expenses, yet often provide competitive features as their commercial equivalents [6]. However, inactive updates may result in long-term security and maintenance problems [34]. They can be easily customised to suit the unique needs of different laboratories but they may require a higher level of technical expertise to implement and maintain.

It is clear that the laboratory management digital tool ecosystem is quite fragmented, which can hinder collaboration within or between institutions and organisations. These solutions are highly specific and lack **interoperability** hampering **orchestration** of operations. Moreover they are often unable to represent actual **knowledge depth** [37]. Critically, the data from these platforms are often not FAIR (Findable, Accessible, Interoperable, Reproducible) compliant [48]. For example, some software can monitor chemical usage, but not transfer data to other platforms for budgeting or inventory. Similarly, some platforms can manage laboratory assets, but require slow and error-prone manual input and updating. This makes it a challenge to fully utilise the available software solutions as they often need to be consolidated manually via spreadsheets [28]. Furthermore, information on equipment availability, risk assessments, or manuals is often scattered and poorly documented, making it difficult for laboratory users to access and use them. Therefore, a more integrated and automated solution for laboratory management and operation is needed to enable FAIR data practices across organisations.

To address these challenges, an innovative solution lies in the creation of digital twins for lab equipment. These digital replicas comprehensively capture a wide range of relevant information – from financial aspects and purchase history to servicing schedules and measurement data. This approach not only streamlines information access but also aligns with the need for more integrated and automated laboratory management. To ensure accuracy and real-time consistency, these digital twins rely on continuous synchronisation with their physical counterparts through sensor data, facilitating shared instrument use, monitoring operations, and scheduling maintenance [22, 33]. Integrating these elements into an Internet of Things (IoT) framework enhances the overall lab automation and resource management system [22, 45]. Currently, there have been limited efforts toward establishing a "smart lab" environment that leverages these technologies [23].

2.2 Leveraging Semantic Web technology for Lab Management

The adoption of digital solutions, either commercial or open-source, varies across different laboratory settings. Although laboratories in large corporations tend to use them extensively, university-based laboratories exhibit a lower rate of adoption due to the lack of flexibility of most solutions, which are customised for specific workflows and require sophisticated training and investment for any modification [45]. Laboratories based in universities face numerous challenges, such as limited budgets, outdated equipment and software, and heterogeneous vendor equipment. Consequently, they are reluctant to adopt digital solutions that are not interoperable with various vendors. Some vendors have attempted to address this issue by integrating different digital tools (*i.e.* SDMS, LIMS, ELN and LES) to offer more holistic solutions, such as STARLIMS and SampleManager LIMS. However, these solutions still encounter difficulties in terms of vendor-agnosticism, data ingestion, and "lock-in" effects [37].

Semantic Web technology is a promising approach to achieving the required data interoperability and vendor-agnosticism for laboratory automation. It enables the development of ontologies that span across different domains, facilitates the integration of different systems, and more importantly provides human- and machine-readable data [5]. An ontology is a comprehensive and structured representation of knowledge, such as concepts and relationships within a specific domain [1]. These representations provide a shared understanding and a standardised definition of knowledge to align both the human and machine perspectives [5]. Machina and Wild [26] recently suggested a new ELN-centric laboratory informatics tool based on Semantic Web technology that integrates LIMS, ELN, CDS, and SDMS semantically. There have also been attempts to develop an ontology-based ELN for microscopy workflows [35]. However, this domain-specific solution has limited interoperability of data across different experiments and domains and did not address laboratory management issues [35]. Recently, the "Open Semantic Lab" project used Semantic Web technology to build a comprehensive online platform capable of semantically capturing and linking laboratory data and concepts, resulting in machine-readable and human-operable data [43]. While the platform aims to support the integration of complex procedures, software, and data, it is still in the early stages of development and currently does not include many of the peripheral aspects and infrastructure relevant to lab managers.

Given the intrinsic complexity related to the Semantic Web technology, a brief explanation will be provided to describe some of the ideas essential to the subject. The terminological component (TBox) and the assertional component (ABox) are the two basic components of an ontology [47]. Through a taxonomy or classification system for domain knowledge, the TBox establishes concepts, their hierarchical relationships, and associated attributes. The TBox concepts are instantiated by ABoxes to represent real-world entities as instances. These instances are all online resources identified by unique internationalised resource identifiers (IRI) and provide meaningful attributes and relationships about a subject. The data of the instances are typically recorded and stored in the form of a "Resource Description Framework" based on subject-predicate-object triples. By extending the representation capabilities of ontologies through the use of knowledge graphs new cross-domain knowledge can be derived. A knowledge graph consists of "nodes" referring to entities of interest and "edges" representing the relationships between two nodes which eventually form a directed graph with the use of ontologies [17]. Such a representation enables a standardised form of human- and machine-readable data model that can be further processed in applications. Notably, the existing ontologies in the data model can be reused and connected with each other to further encapsulate expanding domain knowledge. Knowledge graphs can therefore play a crucial role in facilitating the creation and management of evolving digital twins, enabling continuous integration of information representing the real-time state of laboratory equipment [37].

2.3 The World Avatar Digital Laboratory Framework

This work is part of the larger "The World Avatar" project. Its objective is to develop an all-encompassing digital twin that can connect data and computational agents in real-time to create a living digital "avatar" of the real world, inclusive of abstract concepts and processes [24]. DTs are realistic digital representations of assets, processes, or systems in the built or natural environment that create the opportunity for providing feedback into the physical world [7]. The TWA approach differentiates itself from traditional Semantic Web implementations through the seamless holistic integration of real-time dynamic data, knowledge, models, and tools (*i.e.* dynamic knowledge graph, dKG) in a distributed architecture through an ecosystem of autonomous agents [24]. By acting on real-time data through the agents, TWA describes the behaviour of complex systems and performs tasks such as updates, analysis, decision-making and control of real-world entities [24]. TWA project was initially developed to address the challenges of decarbonisation in the chemical industry of Singapore [13]. Since then, it has expanded to cover a broad range of domains, such as chemistry, chemical processes, laboratories, power systems, as well as various "smart city" applications [16, 21, 24, 32].

In the following sections, we present the implementation and application of the TWA-based Digital Laboratory Framework (DLF) to supporting the role of laboratory managers and users. DLF uses a multi-agent system to create the comprehensive laboratory DT, which inherently incorporates and considers human and infrastructure aspects [37]. The use of modular ontologies facilitates **interoperability** between scales and domains beyond the capabilities of monolithic ontologies. For example, we recently demonstrated the integration of data collected by distributed IoT sensors with a commercial building management system (BMS) and building information management (BIM) data [36]. This enables TWA to support **orchestration** in monitoring and controlling efforts between different systems. Moreover, TWA includes chemistry domain knowledge [32] which facilitates the **knowledge depth** required for complex lab management tasks.

This offers a holistic approach to goal-driven automation of research laboratories. The framework is aligned with the vision of achieving a fully autonomous "AI scientist" [20].

Within a united framework, tasks relating to laboratory management are incorporated and automated to create a distributed and connected DT. Such a "Digital Laboratory Manager" (DLM) can support, assist, and manage all aspects of the research laboratory which are captured by the dKG. The operation of so-called self-driving laboratories as well as non-automated or hybrid research laboratories is also enabled through the development of dynamic human-machine interfaces, coupled with the underlying use of the dKG within the DLF. Going forward, TWA-based agents can operate with an increasing degree of autonomy by re-evaluating their own goals, performing distributed optimisation tasks [3], considering its direct and indirect environment, and taking over managerial tasks related to inventory and chemical management.

3 Developed Methodology

In order to automate lab management tasks – and particularly those related to to inventory management which were highlighted in section 2 – the challenges faced by currently available solutions need to be overcome. Based on the DLF, we developed appropriate methodology using Semantic Web technology in three areas. The first area is asset tracking where we employed radio frequency identification (RFID) technology to tag hazardous (*e.g.* explosive) chemicals, enabling automated safety measures using **deep domain knowledge**. The second area is asset management where we employed "Quick Response" (QR) code technology to tag assets in the laboratory and uniquely identify them for better planning and **orchestration** of experiments. Lastly, a flexible human-machine interface was established to facilitate constant availability of information, **interoperability**, and applicability of automated procedures: a mobile application that communicates directly with the dKG, enabling real-time access and interaction with underlying data models. The underlying TWA structure follows a trilayer concept as illustrated in Fig. 3.

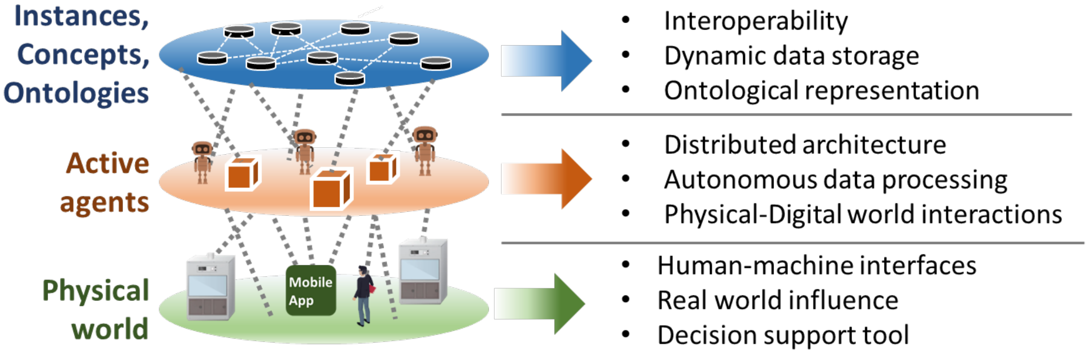


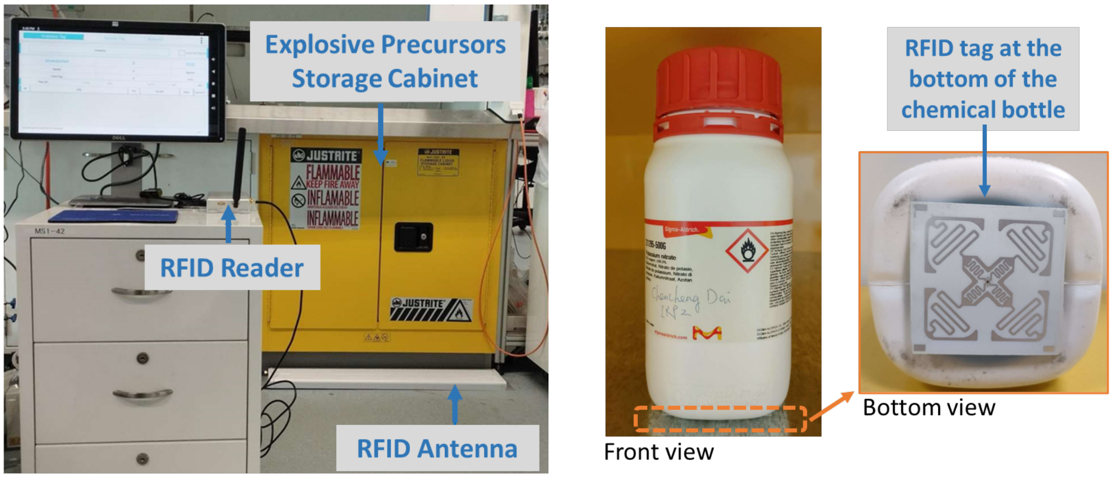
Figure 3: *Trilayer structure of TWA as employed in this work for enhanced management and automation of a research laboratory. The three layers are interconnected as a holistic approach in creating a distributed and connected DT in order to address the challenges in managing research laboratories.*

The top layer represents the ontological models used in the development of the laboratory digital twin. With the development of different ontologies, the resulting dKG can then be used to describe various objects of interest and their connections. The middle layer represents active agents that keep the knowledge graph up-to-date, coordinate data exchange between different servers, and automate various tasks related to lab management. Agents on distributed cloud-based or local servers can interact autonomously with the dKG to instantiate, create, access, and modify data as well as perform simulations, forecast time series, and even control physical objects. The main mode of communication between agents and the dKG is through HTTP requests. The bottom layer consists of physical entities in the real world which interact with and are orchestrated by active agents to carry out certain tasks or achieve certain goal. In the context of the current work, this would include for example fume hoods, temperature sensors, chemical containers, analytical equipment, other laboratory assets, and even laboratory users. Furthermore, multiple digital interfaces and decision support tools are also included in this layer which can assist in communicating and interacting with active agents.

3.1 Tracking assets

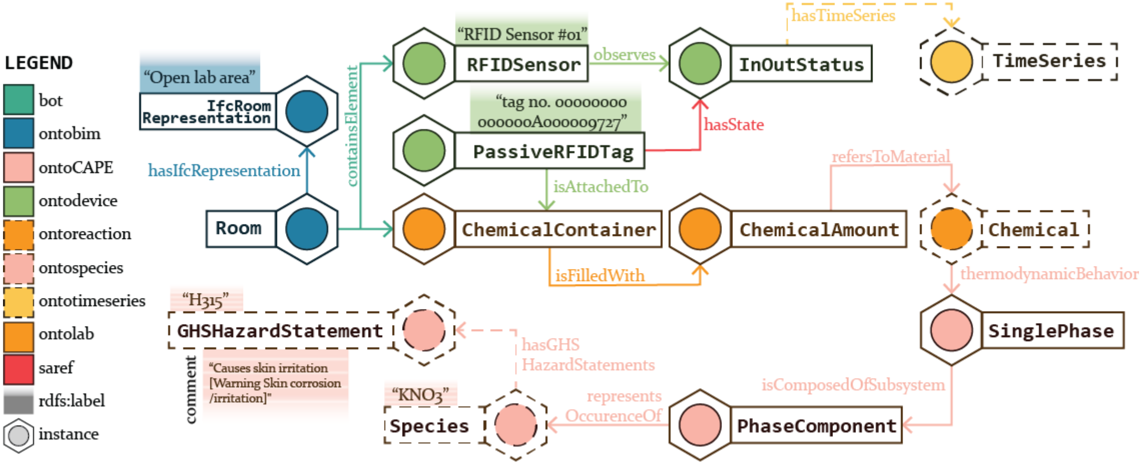
In a research laboratory, many critical assets are often either highly valuable or dangerous and must be closely tracked. As included in Fig. 2, the need to closely track the location of critical assets (especially explosive chemicals) is imperative to ensure the laboratory operates in accordance with safety regulations and compliance. To ensure accurate knowledge on whereabouts of these assets, RFID tags can be employed to track the location of relevant containers. RFID is a key enabler of Industry 4.0 and is a versatile tool for automatic identification and tracking in various domains such as logistics, clothing, agriculture, food, and manufacturing [29, 44]. It is based on a low-cost and accurate sensor that requires minimal or no power consumption and offers wireless power transfer, flexibility, and non-line-of-sight communication [44]. As such, it can be applied to digitise chemical inventory management, resulting in reduced chemical search time and increased efficiency of inventory checks, chemical refills, and safety management [49]. Moreover, integration with a cloud computing platform and a wireless sensor network has been shown to improve efficiency in supervision and utilisation of equipment, and maintain its life cycle accurately [23].

To implement and test location tracking of chemicals via RFID into the DLF, a sample setup was chosen as shown in Fig. 4: An Android `9.0` based 8-port UHF RFID Fixed Reader Writer, along with UHF 902-928 MHz 12dBi RFID Sector Antenna were installed outside of a specially designed metallic cabinet for explosive precursors (Fig. 4(a)). Each of the chemical containers in question were tagged with a UHF 860-960 MHz Frog 3D RFID Tag placed at the bottom of the chemical container (Fig. 4(b)). The reader sends radio frequency signals to the electronic tag through an antenna which receives and returns the signals. In operation, the reader decodes the information embedded in the electronic tag and sends it to the application system for further analysis [23]. If this application system has access to dKG-based digital twins, cross-domain information can be accessed, utilised, and updated in real time. An example is shown in Fig. 4(c): here, existing chemical information on KNO_3 such as the hazard statement (H315) is easily accessed [32] and combined with location data and status information collected by sensors. Additional information on the surrounding facility such as the room the chemical (KNO_3) is located (open lab area), its tag name (RFID Sensor #01), and 3D BIM models of cabinet and facilities [36].



(a) Sample arrangements of the RFID components.

(b) RFID labelling on a chemical container.



(c) ABox snippet for tracking a hazardous chemical container's location with RFID tags.

Figure 4: *Sample setup and corresponding digital twin of the asset tracking use case for explosive chemicals in a laboratory.*

Four key ontologies that were previously developed have been interlinked as part of the development process, namely `OntoLab` [3], `OntoDevice` [36], `OntoBIM` [36], and `OntoSpecies` [32]. This enables the combination and connection of relevant knowledge about location, container, and chemical risk – for example by utilising the previously developed `Email` agent [36] in combination with newly developed RFID agents. These are divided into the "main" agent, "update" agent and "query" agent. The `RFID Update` and `RFID Query` agents are the ones with direct access to the KG. The `RFID Update` agent updates data to the TWA. `RFID Query` agent is used to query and receive data from TWA, depending on the use case.

3.2 Managing assets

In a typical research laboratory, there are a huge number of assets, including laboratory equipment, instruments, chemicals, consumables, and samples. Most of them incur a high cost of acquisition. Additionally, recurring mandatory inventory audit exercises can take many hours or even days to check all assets. To improve productivity and ensure optimal use of resources, lab managers and users need to be able uniquely identify these assets and access relevant information quickly. In fact, as shown in Fig. 2, the topic of inventory management has been identified as a key challenge to potentially benefit from further automation. As the number of assets in the laboratory is high, there is a need to use a low-cost technology to label and identify each asset. We have therefore opted to incorporate QR code technology which is widely used in various settings – including laboratories – to improve management and operation [15].

QR codes have been reported in the literature for similar applications in laboratories, but they have been limited in scope and functionality. For example, Shukran et al. [42] used a QR code to tag chemicals for an inventory system but did not provide information on the location or history of the use of chemicals. In a different publication, researchers reported the use of a QR code to replace the spreadsheet-based inventory checking with a web-based application, but it did not allow data interoperability or new updates [39]. For this work, we combined QR code and Semantic Web technology to address inherent data interoperability issues of previous implementations. An `OntoAssetManagement` has been developed as an ontology for representing asset information such as the assignee, location, identifiers such as serial number, manuals, maintenance information, and related purchase documents. Fig. 5 illustrates an extract of these modular ontologies, represented concepts, and linked domains. An accompanying `AssetManagerAgent` has been developed to perform several key tasks involved in the management of , including instantiation, QR code printing, and data retrieval. For details, see appendix B.2.

In particular, the `OntoAssetManagement` ontology allows the representation of documents involved in the purchase and use of items such as purchase requests, invoices, and manuals. The elements are linked to their physical representations that are located in specific a room or facility within a building, represented by the `OntoBIM` developed previously. The `OntoAssetManagement` ontology extends and interlinks various existing ontologies, such as `Purchase-To-Pay Ontology` (P2P-O) [41] for the related purchase document concepts, `Financial Industry Business Ontology` (FIBO) [12] for concepts relating to persons and organisations, `Time Ontology` [11] for its time concepts, and `Ontology of units of Measure` (OM) [38] for the concepts relating to its measurable quantities. To the best of the authors' knowledge, this holistic interlinking of various domains has not been done before as the existing ontologies were designed discretely and used exclusively in their corresponding domains. For example, the FIBO is mainly used for the financial industry application while the OM is for the formulation of quantitative knowledge in scientific research. In this work, we interlinked these two ontologies to address requirements for representing an R&D laboratory and related managerial tasks: FIBO provides role concepts related to transactions (buyer and seller), roles in an organisation, and independent roles which are crucial to represent the transactional history of lab assets. Meanwhile, most of the assets in a laboratory are used for scientific research, and they will need to have a concept of what they measure and their units.

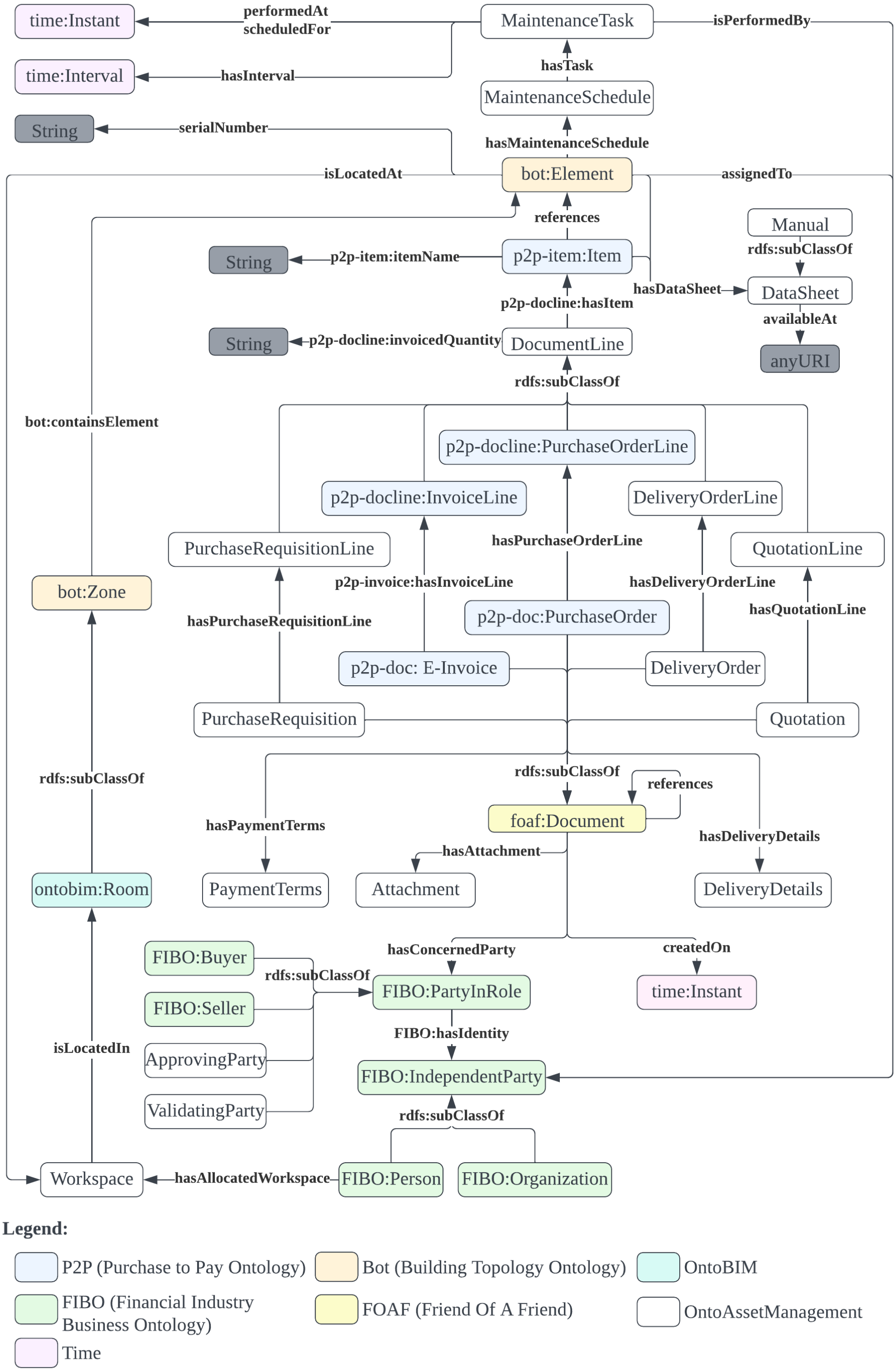


Figure 5: *An extract of the ontology design (TBox) for asset management in a laboratory, integrating geospatial, accounting, and research-related information.*

)

3.3 Mobile interface

In the management of a research laboratory which can span multiple rooms or even buildings, being able to access relevant information on-the-go helps to ensure ease of access and reduce “downtime”. Moreover, many of the manual tasks involved are repetitive such as updating spreadsheets, printing labels, and checking for device information on site. A flexible mobile interface can therefore facilitate user access and interaction, as well as automate repetitive tasks. If this app has access to a dKG such as TWA representing comprehensive digital twins of relevant equipment as introduced earlier in this work, complex tasks and decision-making can be accelerated and supported significantly.

In this work, an Android-based mobile application has been developed within the DLF to facilitate the management of the laboratory. The mobile application has two main functionalities: scanning QR codes to check the asset information and adding new assets along with printing the QR codes. Similarly to the BMS Query mobile application developed in a previous work [36], Keycloak - an open-source identity and access management solution - was used to authenticate and authorise access to the different users for the mobile application. The authorisation code and access token model are implemented by Keycloak based on the OAuth 2.0 authorisation framework (RFC 6749). Keycloak is crucial in the authentication of user login as well as the authentication of users that have rights to access, add, or edit assets, based on their respective role. For this purpose, we extended the ontology used for representing people and their roles based on FIBO [12].

The mobile application has been tested with lab users to ensure that the user experience is seamless, and modification of the user interface has been performed based on users' feedback. The mobile application can be adapted to other types of laboratories or organisations as long as the underlying ontologies ensure data interoperability.

4 Exemplary applications

This section presents specific use cases within the DLF that have been targeted to improve workflows related to laboratory management and, more specifically, the challenges identified in Section 2. The Cambridge CARES research laboratory has been used as the model laboratory for these applications. It supports interdisciplinary and cross-domain research initiatives in a variety of fields, such as combustion research, nanomaterial synthesis, electrochemistry, and pilot plants. Given the complexity of facilities, functions, and requirements, the chosen facilities are suitable to test the TWA methodology and devise strategies to manage laboratories across scales.

A generic and all-encompassing digital twin has been developed and applied to this laboratory, including dedicated agents that interact with both the chemistry and facility domains of the laboratory. This has enabled a number of transformation strategies to enhance operations, namely the orchestration of autonomous experimentation [3] and the integration of facility management aspects to ensure more sustainable operations [36]. The applications presented here are part of this work and extend the DT's capacity to represent knowledge and automate tasks relevant for laboratory management.

We present a system for automated tracking of explosive precursors to enhance safety and a mobile app for asset management to accelerate supporting tasks and increase productivity. Implementing these had significant impact on our lab. Tab. 1 lists some tasks related to lab management that have benefited from the TWA implementation.

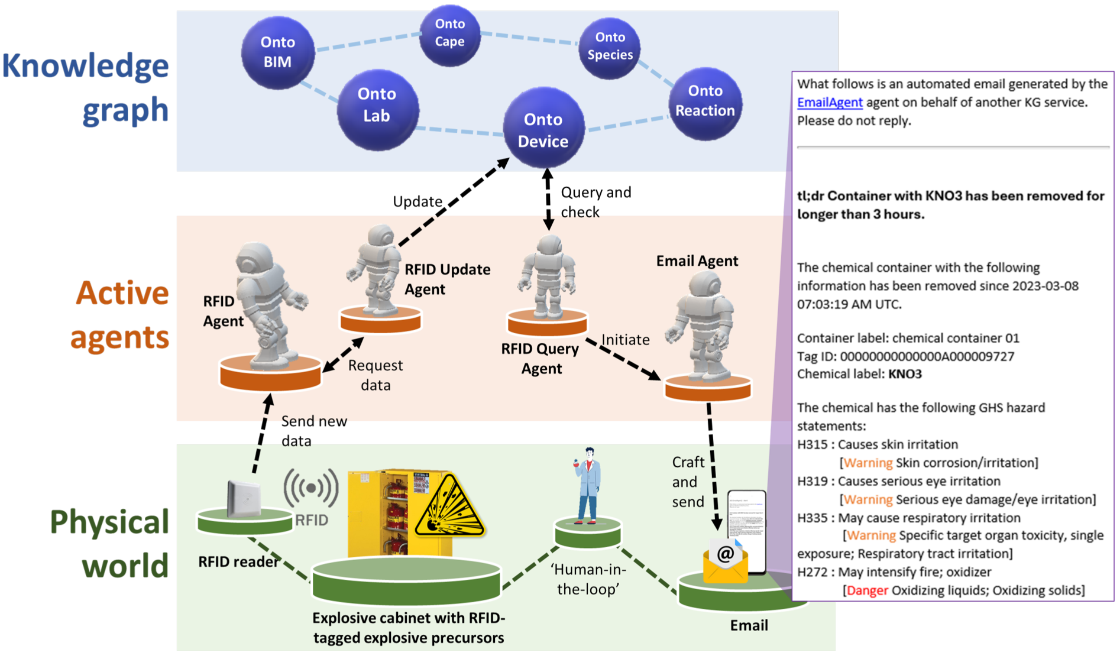
Table 1: *Impact of TWA methodology for three laboratory management tasks.*

Scenarios	**Conventional methods**	**New opportunities**
Registration of a newly purchased item into the system.	New assets are added to an inventory spreadsheet without validation. An identification tag is then manually printed.	New assets can be registered through the mobile app or online interface. After validation, a QR code is automatically printed, enabling users to later retrieve their information immediately.
Monitoring the explosive chemical usage in a storage cabinet.	Users must manually register their usage in a physical log book which is checked in regular intervals.	RFID-enabled automated tracking system sends an email to the lab manager if hazardous chemical is removed from the cabinet for a prolonged time.
Searching for the manual of a certain lab equipment.	Users search online or look through the manual cabinet for the physical copy.	Using the mobile app to scan the QR code on the asset and retrieve a soft copy of the manual.

4.1 Explosive chemicals tracking system

One key responsibility of laboratory management is ensuring the safety of chemicals stored, especially those that are corrosive, flammable, or explosive. To tackle this challenge, a digital explosive chemicals tracking system has been developed and augmented with cross-domain knowledge from chemistry and facility management. This includes information on the explosive limit and conditions. The generic workflow and interaction between the components in the TWA trilayer for this use case is illustrated in Fig. 6.

The connected DTs mirror the real world situation and enable us to observe the agents and their actions autonomously, updating the dKG accordingly. As shown in Fig. 6, when one of the explosive precursor bottles is taken out of the cabinet, it is registered and recorded by the RFID system and the digital twin is updated by the `RFID Update` agent. If the bottle is not returned in due time, the laboratory manager will be notified by email or mobile application, depending on the hazard level of the respective substance. These interactions between `RFID Query` and `Email` agents for example are possible because all relevant hazard and chemical properties are accessible via chemical's `OntoSpecies` representation [32], which is linked to the bottle contents as shown in Fig. 4(c). The RFID tracking method is not limited to the tracking of just explosive precursors but can be used to track any high valued assets that require frequent tracking.



What follows is an automated email generated by the EmailAgent agent on behalf of another KG service. Please do not reply.

tl;dr Container with KNO3 has been removed for longer than 3 hours.

The chemical container with the following information has been removed since 2023-03-08 07:03:19 AM UTC.

Container label: chemical container 01
Tag ID: 00000000000000A000009727
Chemical label: **KNO3**

The chemical has the following GHS hazard statements:
H315 : Causes skin irritation
[Warning Skin corrosion/irritation]
H319 : Causes serious eye irritation
[Warning Serious eye damage/eye irritation]
H335 : May cause respiratory irritation
[Warning Specific target organ toxicity, single exposure; Respiratory tract irritation]
H272 : May intensify fire; oxidizer
[Danger Oxidizing liquids; Oxidizing solids]

Figure 6: *Interactions between TWA components within the RFID application.*

Automating and augmenting lab management tasks in this manner can help improve laboratory safety as it removes some of the potential hazards related to human oversight and forgetfulnes. However, this application also presents some new challenges that require further consideration. For example current safety regulations do not allow the installation of electrical devices inside the cabinet,limiting the use of an omni-directional antenna. Another concern is the possible interference of liquids with the RFID signal. Hence, the introduction of self-correcting agents that can keep track of uncertainties and potential error sources is a key area of future development [37]. For the application at hand, this includes finding the best placement of a tag on the bottle as well as ensuring regular system checks to verify the information presented by the digital twin.

4.2 Mobile app for lab asset management

In the Cambridge CARES research laboratory and office, over 5,000 assets need to be managed, necessitating a comprehensive digital solution to maintain the necessary overview. Furthermore, the processes of adding new equipment to the system and looking up existing equipment in the system need to be streamlined. Hence, we have developed a mobile application, which - together with the underlying `AssetManager` agent and the `OntoAssetManagement` ontology - allows us to manage all assets via QR codes. Three key functionalities are present in the mobile application, *i.e.* instantiate device, print code, and scan code. As shown in Fig. 7, all the functionalities for the management of the assets were enabled through the collaboration of the `AssetManager` agent, the `OntoAssetManagement` and the Asset Management mobile application. Fig. 7 illustrates the interactions between the involved components of each TWA layer when executing these key functionalities.

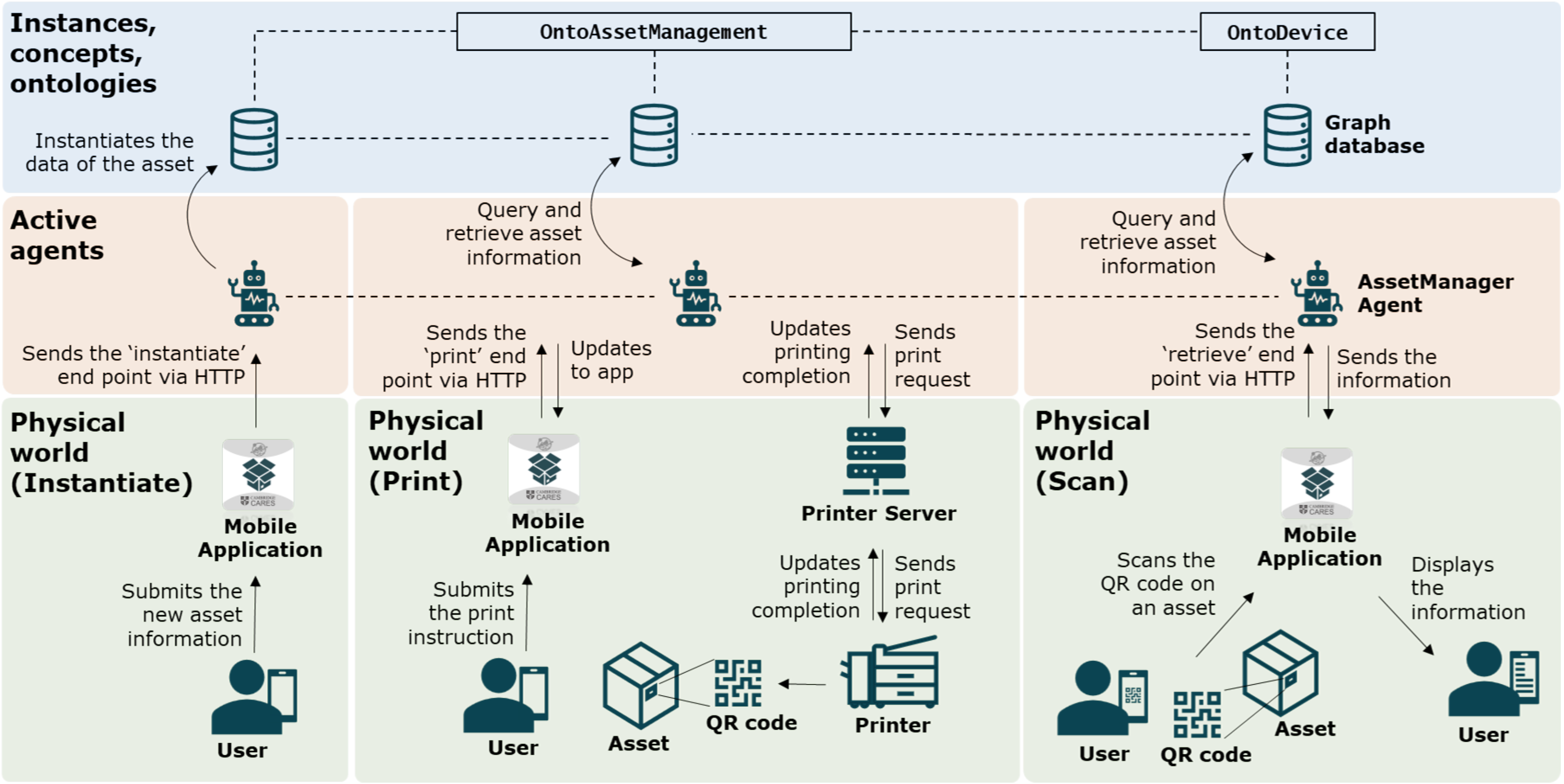


Figure 7: *Interactions between TWA components within the asset management application. The different coloured rows represent the TWA layers as introduced in Fig. 1 while columns represent a key functionality each.*

These features constitute a significant improvement compared to the existing Business-As-Usual (BAU) processes: As is common for a medium-sized research laboratory, assets were mostly tracked within a spreadsheet. While some platform-based solutions for special use cases such as equipment booking or maintenance planning were employed, the "master list" of all existing equipment needed to be maintained separately. When acquiring new equipment for example, a lab manager or user would need to manually add a row to the database with all relevant information, determine a new identification number based on naming conventions, accurately key in this code, and print a label. Within this process, numerous potential error sources exist: the rigid structure of a rectangular database might prevent recording an important property specific to an equipment; an edit in the "master list" might not be enough as the same asset needs to be added or updated across different systems; the unique identifier on the label might be erroneous; *etc.*. With the implementation of the presented solution, error sources are reduced and process time accelerated. For example, the dKG structure inherently ensures consistency of new or changed equipment properties throughout different applications and an automated agent takes care of printing a unique label.

One of the most tedious BAU processes was retrieving asset information such as purchase history, manufacturer specifications, or usage instructions. As shown in Fig. 8, this often included performing searches on the internet, manually checking the inventory list,asking other people, or check physical documents in a cabinet to obtain information. With the DLF asset management mobile application, it requires the user to scan the QR code and all the relevant information is made available instantly.

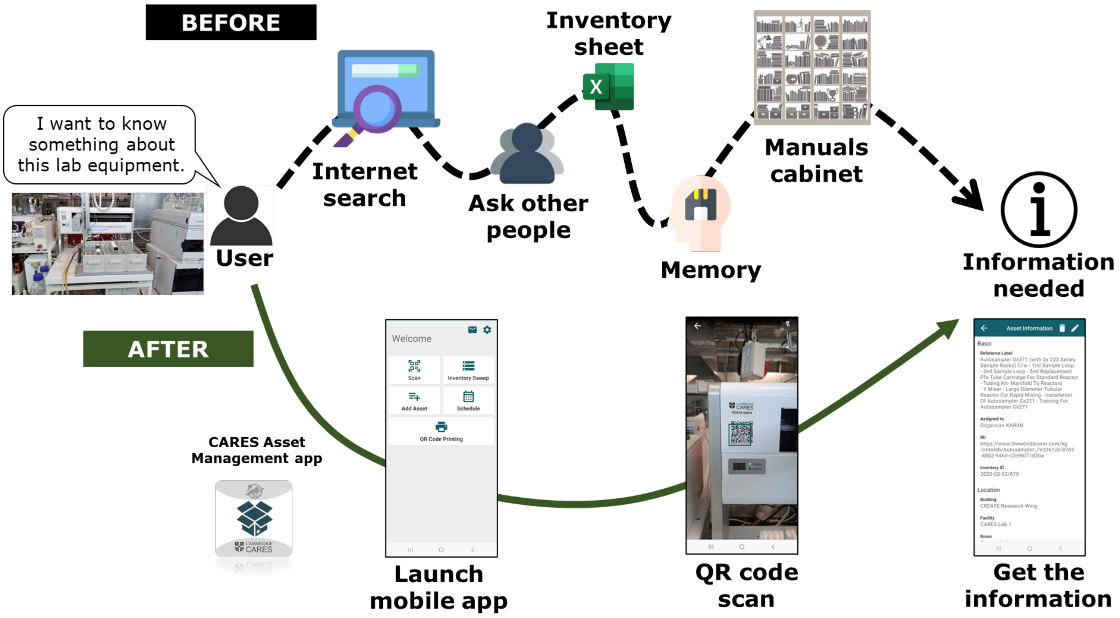


Figure 8: *A comparison of user flow for looking up equipment information before and after the implementation of the DLF-based asset management mobile application.*

5 Conclusions

The advancement of automation in laboratory management is crucial to ensure lab users' safety while providing optimal research support. However, available digital solutions are often fragmented and kept separated from information and data around actual research. To and ultimately realise the vision of an "AI scientist"[20], we harnessed the power of the Semantic Web and developed ontologies, agents, and digital tools for the digitalisation and automation of laboratory management as part of The World Avatar, an all-encompassing digital twin. The modular flexible ontologies developed are able to link knowledge across domains and support data granularity across different scales from chemistry and molecules to lab equipment, room, and facility level. We have shown two use cases that address crucial challenges in the management of research laboratories, *i.e.* keeping track of hazardous chemicals such as explosive precursors and managing laboratory assets via a mobile application offering a unified user interface.

The gradual implementation of the proposed "Digital Lab Framework" [37] in our model laboratory demonstrates the value of interoperability between different technologies and systems that not only address existing challenges but also enhance the accessibility and convenience for laboratory management operations. Software systems and information can be connected within the agent ecosystem to enable new modes of interaction. The integration of RFID tracking with an autonomous email agent serves to automate tracking of critical assets, which is often still based on manual paper-based procedures. For the asset management use case, the unique resource identifiers of the dKG were coupled to QR codes with which relevant attributes, digitised documents, and environmental information of assets can be made instantly accessible. The accompanying mobile application provides a one-stop platform to support laboratory managers and users in their daily routines.

Further implementation of the DLF will continue to address different challenges of laboratory management such as access management, maintenance, and procurement. Consequently, key stakeholders in laboratory management can better focus on core tasks that require sophisticated attention such as decision making, innovation, and problem solving. Expanding the underlying knowledge model with more comprehensive and interoperable digital twins will facilitate collaboration between facility managers [36], laboratory managers, researchers, technicians, and robots [3] to create synergies that can increase the productivity and operation of the research facility. Ultimately, a fully connected and intelligent system for holistic lab automation will enhance the efficiency, quality, and innovation of scientific research with Semantic Web technology.

Declaration of Generative AI and AI-assisted technologies in the writing process

During the preparation of this work the authors used ChatGPT in order to enhance the readability and language of the manuscript. After using this tool, the authors reviewed and edited the content as needed and take full responsibility for the content of the publication.

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. 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

ABox	Assertional Component (of an ontology)
BAU	Business-as-usual (process)
BIM	Building Information Model
BMS	Building Management System
CARES	Cambridge Centre for Advanced Research and Education in Singapore
CDS	Chromatography Data System
CREATE	Campus for Research Excellence and Technological Enterprise
dKG	Dynamic Knowledge Graph
DLF	Digital Laboratory Framework
ELN	Electronic Laboratory Notebook
ERP	Enterprise Resource Planning
IoT	Internet of Things
IRI	Internationalised Resource Identifier
LES	Laboratory Execution System
LIMS	Laboratory Information Management System
QR	Quick Response
R&D	Research and Development
RFID	Radiofrequency Identification
SDMS	Scientific Data Management System

TBox	Terminological Component (of an ontology)
TWA	The World Avatar (project)

A Assessment of automatable lab management tasks

Collection and analysis of use cases for subsequent cost-benefit analysis was done in collaboration with researchers, lab managers, and program managers. From the primary research conducted, there are four key topics of concern in a typical research facility:

- Inventory management;
- Access management;
- Maintenance;
- Procurement and finance.

The details of the challenges can be found in **Table 2**. The order of the challenges do not represent the importance of the challenges.

Table 2: *Collection and cost-benefit analysis of potential use cases.*

Nr	Topic	Challenges	Proposed automation solutions
1	Inventory management	Managing explosive precursor usage for safety reasons	Automated tracking and logging of item location spatially using RFID tags with a receiver in front of the cabinets
2	Inventory management	Re-order solvents when running low	Automated notification to the lab manager when stock reaches a certain threshold; Predict the stock will run low based on historical data and real-time usage and then have an automated alert for replenishment
3	Inventory management	Re-order gases when running low	Automated tracking of gas storage pressures using webcam; Track individual consumption using mass flow meters on specific work stations and sends notification to the lab manager when gas supply reaches a certain threshold
4	Inventory management	Re-order gases when running low	Predict stock or when to order based on the last order and lac activity since; Camera-based tracking of flammable liquids filling levels
5	Inventory management	Re-order cleaning or hygiene supplies when running low	Automated placing of orders based on the number of activity in the laboratory (*e.g.* scientist or day since last order); Automatic ordering of new gloves based on laboratory entry records
6	Inventory management	Audit of assets: Full up-to-date asset inventory list of the laboratory, find specific pieces of equipment requested by auditors	Automated checklist generation; 3D representation or image of all equipment

Continued on next page

Table 2 – continued from previous page

Nr	Topic	Challenges	Proposed automation solutions
7	Access management	Find out status of equipment or workspace: Location and owner	Representation of full laboratory, all pieces of equipment with their location and assignment to a workspace or researcher or project; Full representation of single workspace with multiple users
8	Maintenance	Equipment servicing schedule and possibility prediction	Represent the servicing status and schedule for every asset, possibly based on the type or frequency of use and time interval; Representation of heavy-maintenance equipment with servicing intervals included
9	Access management	Plan temporary usage of shared equipment in advance	Equipment booking system; Representation of full laboratory, all pieces of equipment with their location and assignment to a workspace or researcher or project; Full representation of single workspace with multiple users
10	Assess management	Find out the status and occupancy of work-spaces	Representation of full laboratory, all workspaces with their location and assignment to a researcher or project
11	Procurement and finance	Management of purchase orders	Representation of purchase orders and their status for querying. Representation of ordering history with single supplier, directly using an existing API from ordering portal
12	Maintenance	Overview of current amounts of waste	Representations of toxic waste and their containers
13	Maintenance	Management of waste collection requests	Represent waste collection like purchase requests, trigger based on waste amount
14	Inventory management	Yearly stock-taking of inventory	Create a list of all chemicals currently in the laboratory, their amounts and locations to verify and adjust if needed
15	Inventory management	Keeping track of radioactive material	Automated tracking and logging of item location spatially using RFID tags with a receiver in front of the cabinets. Tracking beacon for radiation source
16	Procurement and finance	Management of incoming deliveries	Represent all purchase orders including their (planned) delivery status. Live APU to update delivery status
17	Procurement and finance	Management of problematic deliveries	Represent all purchase orders including their (planned) delivery status. Live APU to update delivery status. Mechanism to track the type of laboratory activities and risk assessment
18	Access management	Making sure everyone has received adequate training before gaining access to laboratory and/or equipment or starting any work	Representation of all researchers and their training, check against assignments to workspaces, equipment *etc.* Mechanism to update researchers training status. Check if everyone on weekly list has safety inductions

Continued on next page

Table 2 – continued from previous page

Nr	Topic	Challenges	Proposed automation solutions
19	Access management	Manage laboratory access	Representation of all researchers and their assigned laboratory access slots. API to update access details for each week
20	Procurement and finance	Overview of requested quotations	Representation of of quotations similar to purchase orders to query. Mechanism to instantiate quotations when received
21	Procurement and finance	Verifying received quotations	Check details of quotations automatically. Representation of of quotations similar to purchase orders to query. Mechanism to instantiate quotations when received. Feature extraction for PDF quotations
22	Maintenance	Licence management	Representations of entities and activities that require licence, their corresponding expiry date and renewal requirements
23	Procurement and finance	Consumable budget overview	Represent relevant aspects of the purchase system (Synergix) *e.g.* cost centre, IRP, budget line, PO, PR and delivery status
24	Procurement and finance	Yearly laboratory budget setting	Automated budget creation based on the usage and assignment of assets for the next year and historic budgeting and usage data of consumables. Budget planning for single item based on historic budget and current usage data
25	Procurement and finance	Assign gas usage to different IRPs	Separate tracking of gas usage between different researchers or projects and automated billing accordingly. Tracking of gas usage for different workstations outlets and mechanism to always assign these to a researcher or project based on date and time of consumption (possibly API to Synergix)

B Ontologies and Agents

B.1 Namespaces

The ontological representation involves several namespaces, some of which have been reused with or without adaptations and some of which were developed entirely as part of The World Avatar project.

bot: <https://w3id.org/bot#>

om: <http://www.ontology-of-units-of-measure.org/resource/om-2/>

ontoapplication: <https://www.theworldavatar.com/kg/ontoapplication/>

ontoassetmanagement: <https://www.theworldavatar.com/kg/ontoassetmanagement/>

ontobim: <https://www.theworldavatar.com/kg/ontobim/>

ontocape: <https://www.theworldavatar.com/kg/ontocape/>

ontodevice: <https://www.theworldavatar.com/kg/ontodevice/>

ontolab: <https://www.theworldavatar.com/kg/ontolab/>

ontoreaction: <https://www.theworldavatar.com/kg/ontoreaction/>

ontospecies: <http://www.theworldavatar.com/ontology/ontospecies/OntoSpecies.owl#>

rdfs: <http://www.w3.org/2000/01/rdf-schema#>

saref: <https://saref.etsi.org/core/>

timeseries: <https://www.theworldavatar.com/kg/ontotimeseries/>

p2p-invoice: <https://purl.org/p2p-o/invoice#>

p2p-item: <https://purl.org/p2p-o/item#>

p2p-doc: <https://purl.org/p2p-o/document#>

p2p-docline: <https://purl.org/p2p-o/documentline#>

foaf: <http://xmlns.com/foaf/0.1/>

time: <http://www.w3.org/2006/time#>

fibo-organization: <https://spec.edmcouncil.org/fibo/ontology/FND/Organizations/Organizations/>

fibo-parties: <https://spec.edmcouncil.org/fibo/ontology/FND/Parties/Parties/>

fibo-people: <https://spec.edmcouncil.org/fibo/ontology/FND/AgentsAndPeople/People/>

B.2 Agent details

All major agents developed and deployed for the applications detailed in this manuscript are listed in Tab. 3. Their source code is available under `https://github.com/cambridge-cares/TheWorldAvatar/tree/main/Agents`.

Table 3: *A summary of the agents and their tasks deployed for the current work.*

Agents	Task
`AssetManagerAgent`	Instantiate assets and their information into the KG, print QR codes, retrieve asset information from the KG.
`RFIDAgent`	Receive and parse data from RFID reader. Saves data to a postgreSQL database and provides an API for retrieving it.
`RFIDUpdateAgent`	Retrieve data from postgreSQL database and instantiate as timeseries in the KG.
`RFIDQueryAgent`	Query KG for tagged item status and relevant information, evaluate and determine whether to send a notification email and trigger the `email` agent.
`EmailAgent`	Receive email content, craft emails and send them out to specific parties.

The RFID agents are illustrated in Fig. 9 with details of their interactions. The `RFID` and `RFID Update` agents collaborate to constantly update the item status from the RFID reader. Meanwhile, the `RFID Query` agent obtains the information of the item status from the TWA, evaluates if an email needs to be sent regarding the status and lastly, notify the `Email` agent to craft and send an email to the laboratory manager.

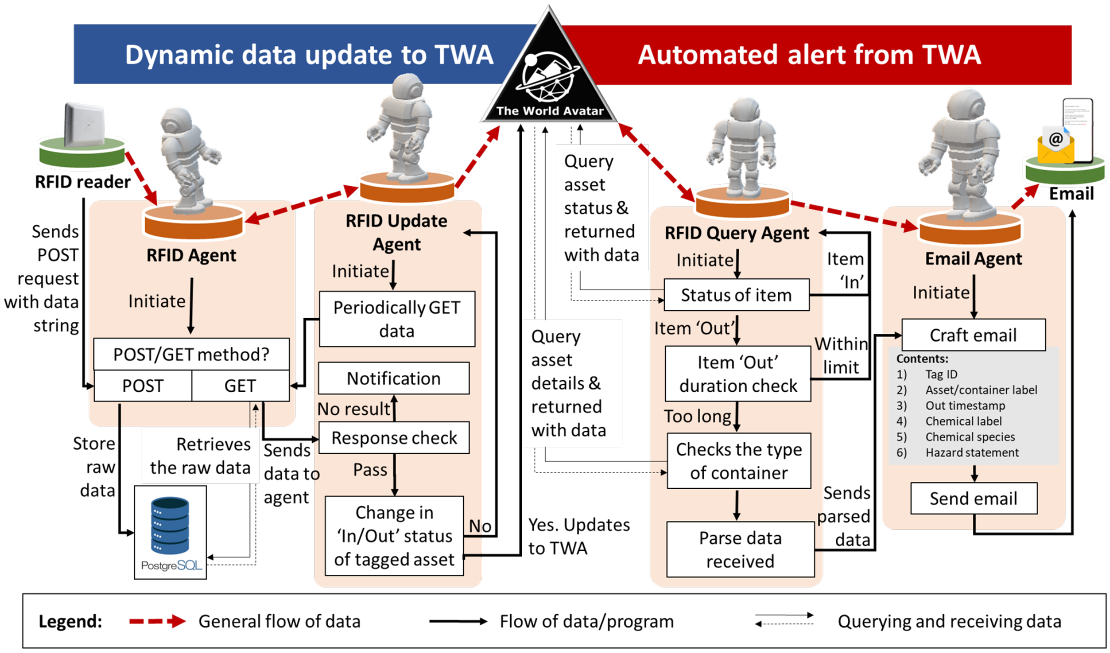


Figure 9: *Unified Modelling Language (UML) diagram for RFID agents.*

The AssetManager agent has been developed to perform several key tasks involved in the management of assets such as the instantiation of the assets, printing of QR code and retrieval of data. The UML diagram of the agent is as shown in Fig. 10.

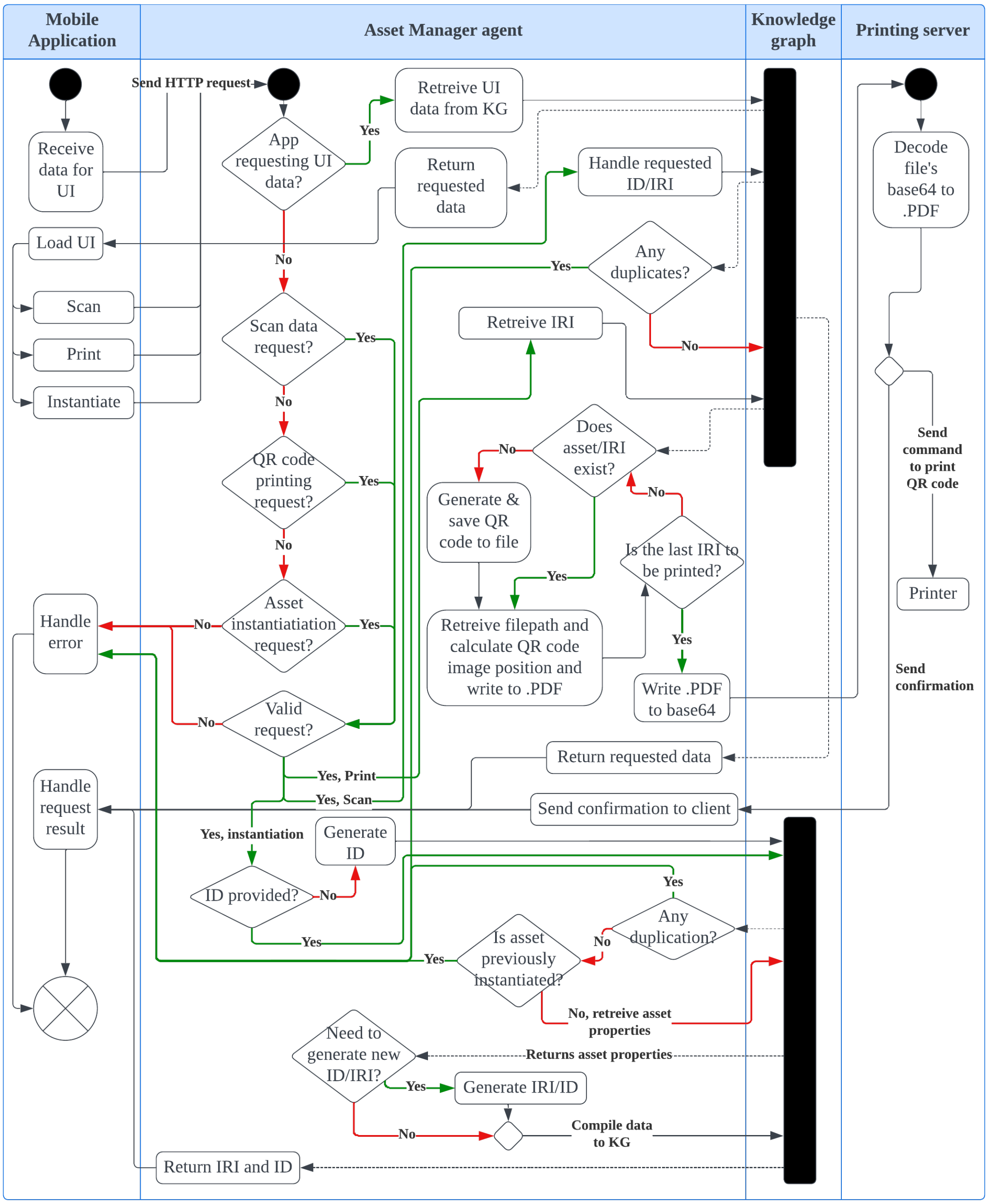


Figure 10: *UML diagram for AssetManager agent.*

C Mobile app workflows

C.1 Authentication implementation workflow

An Open Sourced Identity and Access Management solution (Keycloak) has been deployed in the mobile application which can support authorisation services, standard Protocols (*e.g.* OIDC, Oauth 2, SAML) and *etc.* This is to ensure that only authorised users are able to make changes to the asset information. As shown in Fig. 11, Keycloak is crucial in the authorisation of login of the user to log into the mobile application, as well as verifying if the user has the right access token to add or edit the asset.

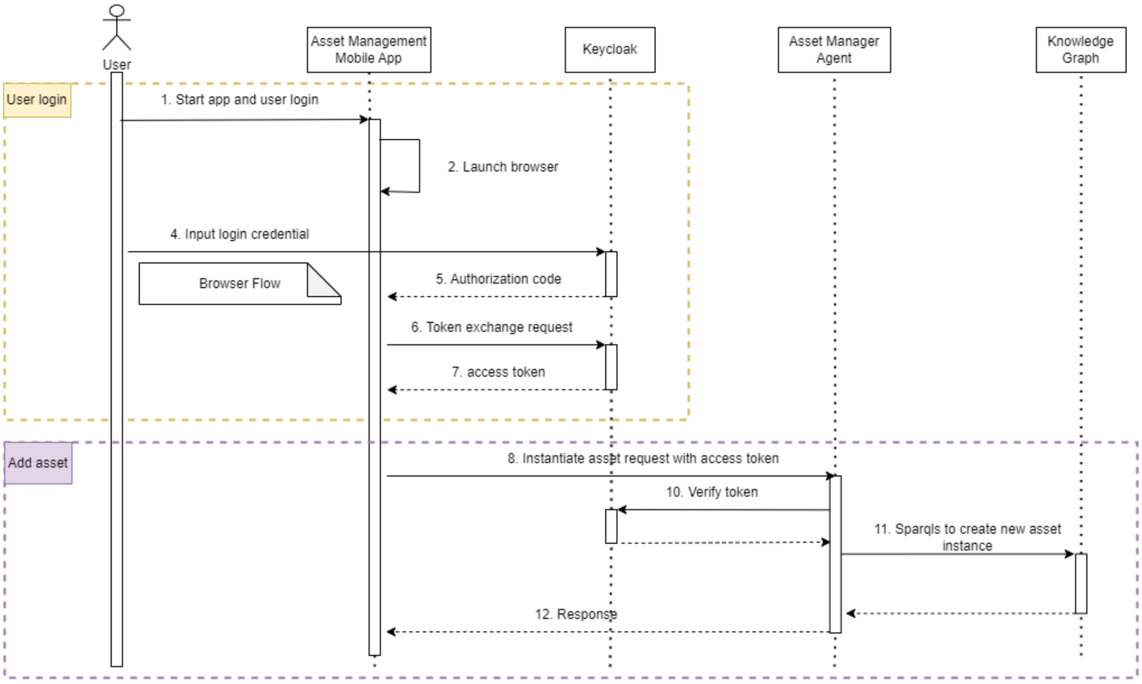


Figure 11: *UML Sequence diagram for Keycloak implementation for the Asset Management mobile application.*

Keycloak employs diverse entities, a Realm is a collection of users, credentials, roles, groups, OIDC clients, and additional data. Each Realm has its own collection and can only authenticate users belonging to the Realm. Regarding roles, Keycloak offers three different types of roles: Realm Roles which are associated with a certain realm, accessible anywhere and may be assigned to any user, Client Roles which are only available in that client and are only assigned to users of that client, and at last Composite Roles which combine realm and client roles.

Fig. 12 shows that the Keycloak Realm Configuration, *i.e.* the set of users, credentials, roles, and groups, which are isolated from one another. This allows only the ones within the same realm to manage and authenticate the users that they control. For instance, there are two clients (`AssetManager` agent and Asset Management mobile application) that are registered with Keycloak. For the `AssetManager` agent client, there are authorisation rules that are implemented while for the Asset Management mobile application client, there are no authorisation rules implemented. This is because the `AssetManager` agent acts an interface between the knowledge graph and the mobile application. Once the `AssetManager` agent has performed the authorisation process through Keycloak authorisation rules based on the user roles and user information, the Asset Management mobile application can access to the knowledge graph within the limits of the authorisation.

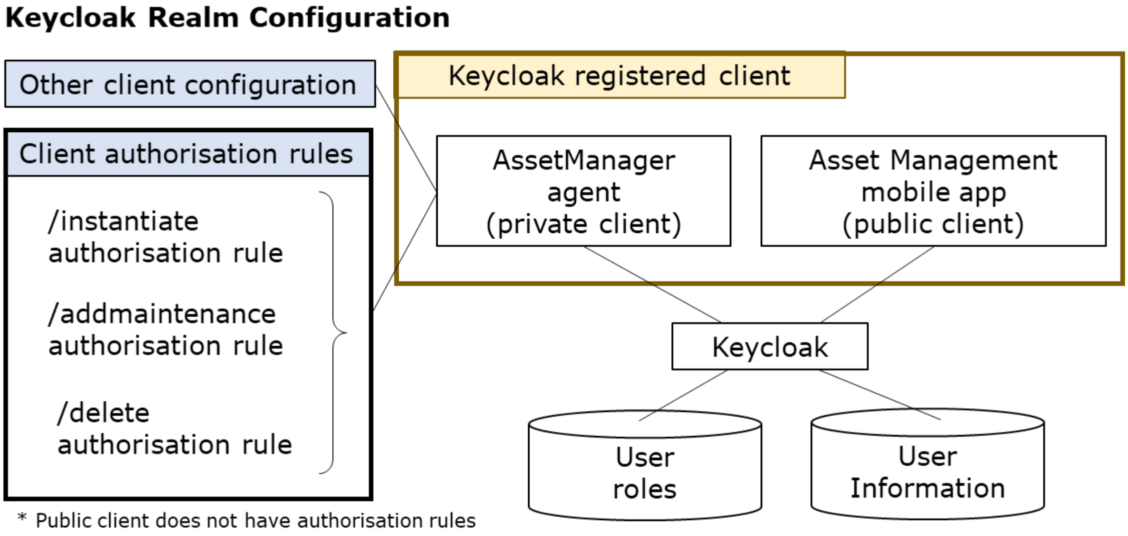


Figure 12: *The underlying architecture of the Keycloak realm configuration.*

Fig. 13 and Fig. 14 show the detailed flows for user login and asset instantiation using the Asset Management mobile application. For both the flows, Keycloak is the key to provide access token and authorisation code, as well as to verify access token and evaluate authorisation rules. The Keycloak accesses the 'user information' and 'user roles' data from the knowledge graph to perform the verification and assessment processes.

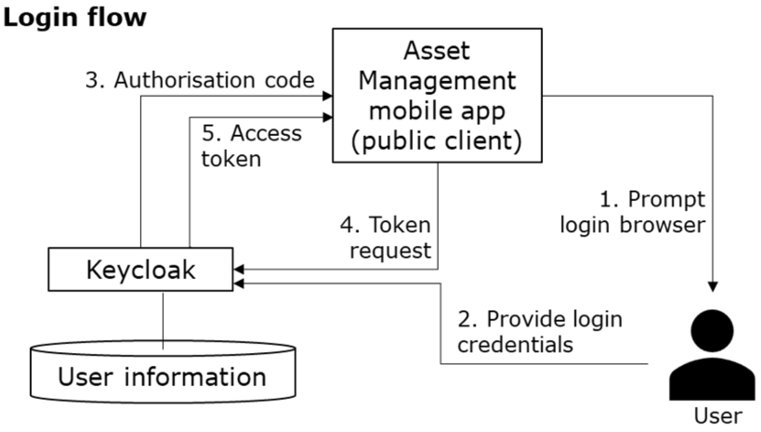


Figure 13: *UML diagram for user login.*

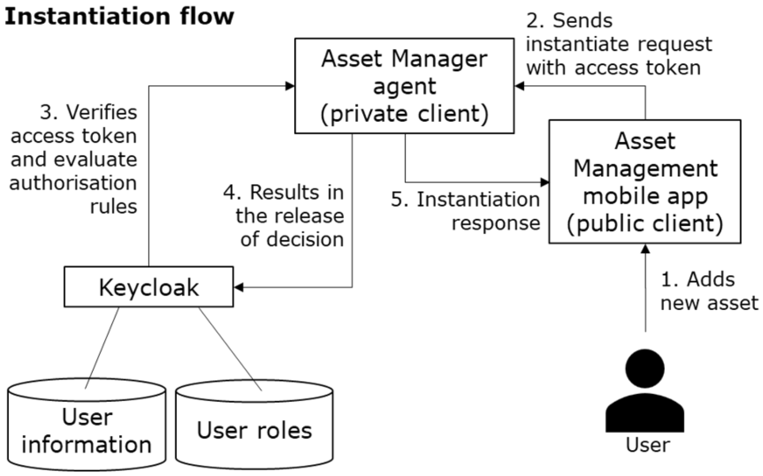
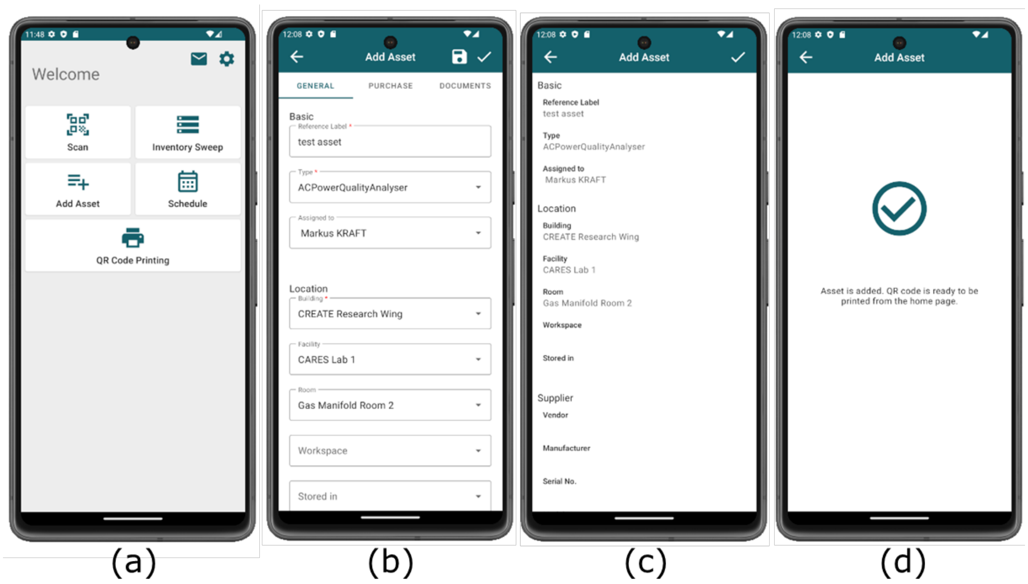


Figure 14: *UML diagram for asset instantiation.*

C.2 Asset Management Mobile Application User Flow

Fig. 15 shows the typical user flow adding assets using the mobile application. Starting at Fig. 15(a) the "Add Asset" tile on the home screen will direct the user to the next screen, Fig. 15(b), where the user can enter or select the relevant information regarding the new asset. In the next screen, Fig. 15(c), the user will check and confirm the information before the user can submit the information. Lastly, the user will see a confirmation page (Fig. 15(d)) and at the same time, the QR code generated will be sent as a print task (covered in the next subsection). The user can then continue to add more new assets until they are ready to print the QR codes. From the introduction of this feature, the time taken to obtain add a new asset has been reduced from the conventional 5–10 minutes to 3 minutes.



(b)

Basic
Reference Label *: test asset
Type *: ACPowerQualityAnalyser
Assigned to: Markus KRAFT

Location
Building *: CREATE Research Wing
Facility: CARES Lab 1
Room: Gas Manifold Room 2

(c)

Basic
Reference Label
test asset
Type
ACPowerQualityAnalyser
Assigned to
Markus KRAFT

Location
Building
CREATE Research Wing
Facility
CARES Lab 1
Room
Gas Manifold Room 2
Workspace
Stored in

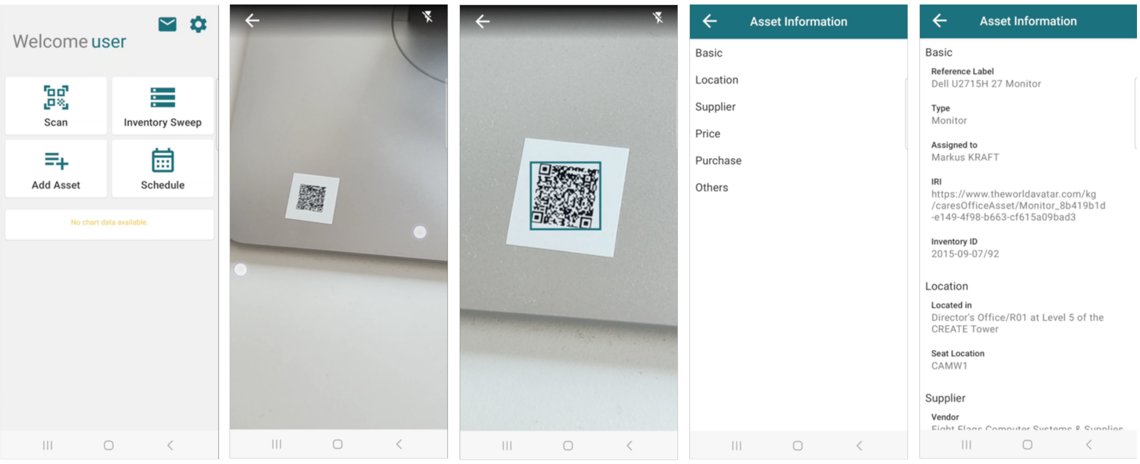
Supplier
Vendor
Manufacturer
Serial No.

(d)

Asset is added. QR code is ready to be printed from the home page.

Figure 15: *Exemplary user flow of the asset adding process using the mobile application.*

Fig. 16 shows the user flow of checking assets using the mobile application. In short, the user can use the application to scan the QR code on the asset and all the information of the asset can be retrieved instantly. On the home screen in Fig. 16(a), the user will select the "Scan" tab in order to initiate the feature. It will then activate the camera of the mobile phone with zoom in function, as shown in Fig. 16(b). When the QR code has been zoomed in to a detectable size, in Fig. 16(c), it will then direct the user to the next screen in Fig. 16(d) and (e) where the detailed information of the asset will be shown to the user. From the introduction of this feature in the Cambridge CARES laboratory, the time taken to obtain information on any asset has been reduced from the conventional 2–3 minutes to 5 seconds.



Basic

Reference Label
Dell U2715H 27 Monitor

Type
Monitor

Assigned to
Markus KRAFT

IRI
https://www.theworldavatar.com/kg/caresOfficeAsset/Monitor_8b419b1d-e149-4f98-b663-cf615a09bad3

Inventory ID
2015-09-07/92

Location

Located in
Director's Office/R01 at Level 5 of the CREATE Tower

Seat Location
CAMW1

Supplier

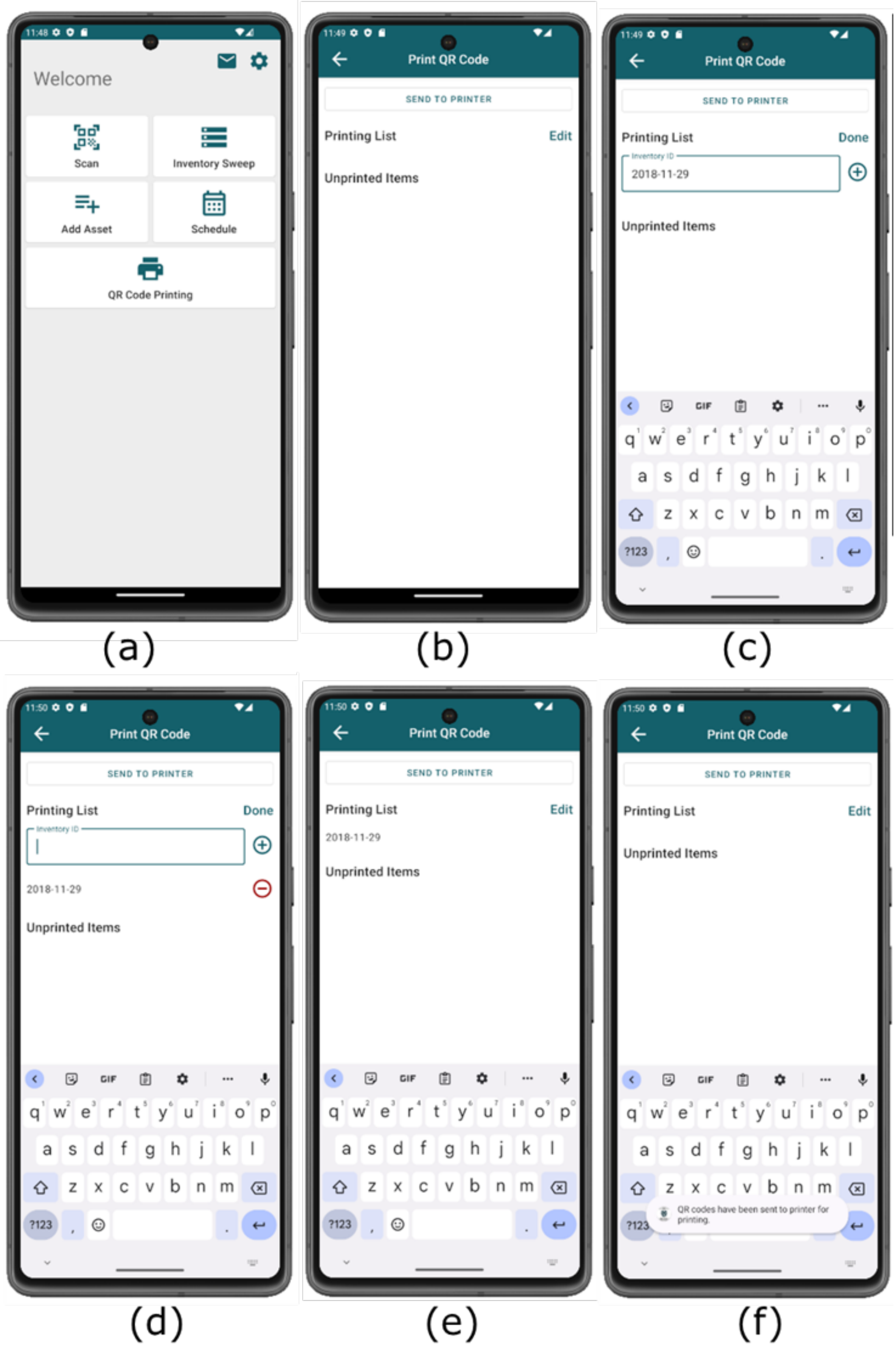
Vendor

(a) (b) (c) (d) (e)

Figure 16: *Exemplary user flow of the asset checking process using the mobile application.*

Fig. 17 shows the screenshots of the steps to follow for the printing of QR code labels using the TWA asset management mobile application. Before the implementation of the TWA DLF, the user will need to determine manually the identification number of the new asset, do to a printing machine, key in the code and then print the label. With the implementation of the current solution, the printing of the labels are automated. With just a tap on the mobile application, the printing task will be executed and the label will be ready to be collected by the user.

To access the printing feature, the user can select 'QR Code Printing' in Fig. 17(a). Then, it will bring the user to page Fig. 17(b). The user then can select "Edit" to allow the user to manually insert the "Inventory ID" of the asset, as shown in Fig. 17(c). Users can also print several QR codes by adding a few more "Inventory ID" of the asset. Once the list is complete, the user can select 'Done' to confirm the 'Printing List', as shown in Fig. 17(d). In Fig. 17(e), there will be a list of assets ready for the QR code printing. When a new asset has been instantiated, the printing task will be sent to the printer server as a printing task and appear in the list in in Fig. 17(e). When the user is ready to print the QR codes of assets, they can select "Send to Printer" in Fig. 17(e) to send the printing job to the printer and a message "QR codes have been sent to printer for printing" will appear as shown in Fig. 17(f).



QR codes have been sent to printer for printing.

Figure 17: *Exemplary user flow of the asset printing process using the asset management mobile application.*

References

[1] G. Antoniou, E. Franconi, and F. van Harmelen. *Introduction to Semantic Web Ontology Languages*, volume 3564 of *Lecture Notes in Computer Science*, pages 1–21. Springer Berlin Heidelberg, 2005. doi:10.1007/11526988_1.

[2] 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. doi:10.1021/jacsau.1c00438.

[3] J. Bai, S. Mosbach, C. J. Taylor, D. Karan, K. F. Lee, S. D. Rihm, J. Akroyd, A. A. Lapkin, and M. Kraft. A Dynamic Knowledge Graph Approach to Distributed Self-Driving Laboratories. *Nature Communications*, 15(1):462, 2024. doi:10.1038/s41467-023-44599-9.

[4] C. Barillari, D. S. Ottoz, J. M. Fuentes-Serna, C. Ramakrishnan, B. Rinn, and F. Rudolf. OpenBIS ELN-LIMS: An open-source database for academic laboratories. *Bioinformatics*, 32(4):638–640, 2016. doi:10.1093/bioinformatics/btv606.

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

[6] P. Blazek, K. Kuca, J. Krenek, O. Krejcar, and E. Nepovimova. Comparative evaluation of open source laboratory information and management systems. In *2018 IEEE Conference on Open Systems (ICOS)*, pages 19–24. IEEE, 2018. doi:10.1109/ICOS.2018.8632810.

[7] A. Bolton, L. Butler, I. Dabson, M. Enzer, M. Evans, T. Fenemore, F. Harradence, E. Keaney, A. Kemp, A. Luck, N. Pawsey, S. Saville, J. Schooling, M. Sharp, T. Smith, J. Tennison, J. Whyte, A. Wilson, and C. Makri. Gemini principles. Technical report, Centre for Digital Built Britain, Cambridge, 2018.

[8] C. L. Borgman. The conundrum of sharing research data. *Journal of the American Society for Information Science and Technology*, 63(6):1059–1078, 2012. doi:10.1002/asi.22634.

[9] S. A. Brown, R. Sparapani, K. Osinski, J. Zhang, J. Blessing, F. Cheng, A. Hamid, M. B. MohamadiPour, J. C. Lal, A. N. Kothari, P. Caraballo, P. Noseworthy, R. H. Johnson, K. Hansen, L. Y. Sun, B. Crotty, Y. C. Cheng, G. Echefu, K. Doshi, and J. Olson. Team principles for successful interdisciplinary research teams. *American Heart Journal Plus: Cardiology Research and Practice*, 32, 2023. doi:10.1016/j.ahjo.2023.100306.

[10] M. L. Comeaga. Digital transformation of the laboratories. *IOP Conference Series: Materials Science and Engineering*, 1268(1):012001, nov 2022. doi:10.1088/1757-899x/1268/1/012001.

[11] S. J. Cox and C. Little. Time Ontology in OWL, 2017. URL `https://www.w3.org/TR/owl-time/`. Last accessed January 8, 2024.

[12] EDM Council. The Financial Industry Business Ontology (FIBO), 2020. URL `https://spec.edmcouncil.org/fibo/`. Last accessed January 8, 2024.

[13] A. Eibeck, M. Q. Lim, and M. Kraft. J-Park Simulator: An ontology-based platform for cross-domain scenarios in process industry. *Computers and Chemical Engineering*, 131, 2019. doi:10.1016/j.compchemeng.2019.106586.

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

[15] C. Hernandez, E. Aslankoohi, P. Frolikov, H. Li, S. Kurniawan, and M. Rolandi. Implementing QR codes in academia to improve sample tracking, data accessibility, and traceability in multicampus interdisciplinary collaborations. *PLoS ONE*, 18(4 April):1–12, 2023. doi:10.1371/journal.pone.0282783.

[16] M. Hofmeister, S. Mosbach, J. Hammacher, M. Blum, G. Röhrig, C. Dörr, V. Flegel, A. Bhave, and M. Kraft. Resource-optimised generation dispatch strategy for district heating systems using dynamic hierarchical optimisation. *Applied Energy*, 305: 117877, Jan. 2022. doi:10.1016/j.apenergy.2021.117877.

[17] A. Hogan, C. Gutierrez, M. Cochez, G. de Melo, S. Kirrane, A. Polleres, R. Navigli, A.-C. Ngonga Ngomo, S. M. Rashid, L. Schmelzeisen, S. Staab, E. Blomqvist, C. D'Amato, J. E. Labra Gayo, S. Neumaier, A. Rula, J. Sequeda, and A. Zimmermann. *Knowledge Graphs*. Synthesis Lectures on Data, Semantics, and Knowledge. Springer Nature Switzerland AG, 1 edition, 2022. doi:10.1007/978-3-031-01918-0.

[18] H. Jo and D. H. Park. Mechanisms for successful management of enterprise resource planning from user information processing and system quality perspective. *Scientific Reports*, 13(1):12678, 2023. doi:10.1038/s41598-023-39787-y.

[19] Y. Kalambet. Data acquisition and integration. In C. F. Poole, editor, *Gas Chromatography (Second Edition)*, Handbooks in Separation Science, pages 505–524. Elsevier Inc., 2021. doi:10.1016/b978-0-12-820675-1.00038-1.

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

[21] A. Kondinski, A. Menon, D. Nurkowski, F. Farazi, S. Mosbach, J. Akroyd, and M. Kraft. Automated Rational Design of Metal-Organic Polyhedra. *Journal of the American Chemical Society*, 144(26):11713–11728, 2022. doi:10.1021/jacs.2c03402.

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

[23] S. Li, X. Gao, W. Wang, and X. Zhang. Design of smart laboratory management system based on cloud computing and internet of things technology. *Journal of Physics: Conference Series*, 1549(2):022107, 2020. doi:10.1088/1742-6596/1549/2/022107.

[24] M. Q. Lim, X. Wang, O. Inderwildi, and M. Kraft. The World Avatar—A World Model for Facilitating Interoperability. In O. Inderwildi and M. Kraft, editors, *Intelligent Decarbonisation: Can Artificial Intelligence and Cyber-Physical Systems Help Achieve Climate Mitigation Targets?*, pages 39–53. Springer International Publishing, 2022. doi:10.1007/978-3-030-86215-2_4.

[25] M. M. Little, C. A. St Hill, K. B. Ware, M. T. Swanoski, S. A. Chapman, M. N. Lutfiyya, and F. B. Cerra. Team science as interprofessional collaborative research practice: A systematic review of the science of team science literature. *Journal of Investigative Medicine*, 65(1):15–22, 2017. doi:10.1136/jim-2016-000216.

[26] H. K. Machina and D. J. Wild. Laboratory Informatics Tools Integration Strategies for Drug Discovery: Integration of LIMS, ELN, CDS, and SDMS. *Journal of Laboratory Automation*, 18(2):126–136, 2013. doi:10.1177/2211068212454852.

[27] R. P. Mazzarese, S. M. Bird, P. J. Zipfell, and M. W. Dong. Chromatography data systems: Perspectives, principles, and trends. *LCGC North America*, 37:852–865, 2019. URL `https://www.chromatographyonline.com/view/chromatography-data-systems-perspectives-principles-and-trends`.

[28] A. Moser, A. E. Waked, and J. Dimartino. Consolidating and Managing Data for Drug Development within a Pharmaceutical Laboratory: Comparing the Mapping and Reporting Tools from Software Applications. *Organic Process Research and Development*, 25(10):2177–2187, 2021. doi:10.1021/acs.oprd.1c00082.

[29] A. Mostaccio, G. M. Bianco, G. Marrocco, and C. Occhiuzzi. RFID Technology for Food Industry 4.0: A review of solutions and applications. *IEEE Journal of Radio Frequency Identification*, 7:145–157, 2023. doi:10.1109/JRFID.2023.3278722.

[30] S. Neubert, B. Göde, X. Gu, N. Stoll, and K. Thurow. Potential of Laboratory Execution Systems (LESs) to Simplify the Application of Business Process Management Systems (BPMSs) in Laboratory Automation. *SLAS Technology*, 22(2):206–216, 2017. doi:10.1177/2211068216680331.

[31] S. Neubert, X. Gu, B. Göde, T. Roddelkopf, H. Fleischer, N. Stoll, and K. Thurow. Workflow Management System for the Integration of Mobile Robots in Future Labs of Life Sciences. *Chemie Ingenieur Technik*, 91(3):294–304, 2019. doi:10.1002/cite.201800007.

[32] L. Pascazio, S. Rihm, A. Naseri, S. Mosbach, J. Akroyd, and M. Kraft. Chemical Species Ontology for Data Integration and Knowledge Discovery. *Journal of Chemical Information and Modeling*, 63(21):6569–6586, 2023. doi:10.1021/acs.jcim.3c00820.

[33] A. S. Pillai and T. B. Isha. A power-aware multiprocessor based laboratory automation and resource management system. In *2014 IEEE International Conference on Computational Intelligence and Computing Research*, pages 1–4. IEEE, 2014. doi:10.1109/ICCIC.2014.7238523.

[34] J. Potthoff, P. Tremouilhac, P. Hodapp, B. Neumair, S. Bräse, and N. Jung. Procedures for systematic capture and management of analytical data in academia. *Analytica Chimica Acta: X*, 1:1–7, 2019. doi:10.1016/j.acax.2019.100007.

[35] I. Rees, E. Langley, W. Chiu, and S. J. Ludtke. EMEN2: An Object Oriented Database and Electronic Lab Notebook. *Microscopy and Microanalysis*, 19(1):1–10, 01 2013. doi:10.1017/S1431927612014043.

[36] S. D. Rihm, Y. R. Tan, W. Ang, H. Y. Quek, X. Deng, M. T. Laksana, S. Mosbach, and M. Kraft. The Digital Lab Facility Manager : Using dynamic knowledge graph technology to automate operations of research laboratories. Submitted for publication; preprint available online, 2023. URL `https://como.ceb.cam.ac.uk/preprints/316/`.

[37] S. D. Rihm, J. Bai, A. Kondinski, S. Mosbach, J. Akroyd, and M. Kraft. Transforming Research Laboratories with Connected Digital Twins. *Nexus*, 2024. URL `https://como.ceb.cam.ac.uk/preprints/314/`. Article in press; preprint available online.

[38] H. Rijgersberg, M. Van Assem, and J. Top. Ontology of units of measure and related concepts. *Semantic Web*, 4(1):3–13, 2013. doi:10.3233/SW-2012-0069.

[39] N. Rochmawati, I. G. Buditjahjanto, R. E. Putra, and A. Y. Wicaksono. A Responsive Web-Based QR Code for Inventory in the Laboratory of Informatics, UNESA. *IOP Conference Series: Materials Science and Engineering*, 288(1):012109, jan 2018. doi:10.1088/1757-899X/288/1/012109.

[40] F. Schmieder, C. Polk, F. Gottlöber, P. Schöps, F. Sonntag, R. Deuse, A. Jede, and T. Petzold. Universal LIMS based platform for the automated processing of cell-based assays. *Current Directions in Biomedical Engineering*, 5(1):437–439, 2019. doi:10.1515/cdbme-2019-0110.

[41] M. Schulze, M. Schröder, C. Jilek, T. Albers, H. Maus, and A. Dengel. P2P-O: A purchase-to-pay ontology for enabling semantic invoices. In *The Semantic Web: 18th International Conference, ESWC 2021*, pages 647–663, 2021. doi:10.1007/978-3-030-77385-4_39.

[42] M. A. M. Shukran, M. S. Ishak, and M. N. Abdullah. Enhancing chemical inventory management in laboratory through a mobile-based qr code tag. *IOP Conference Series: Materials Science and Engineering*, 226(1):012093, aug 2017. doi:10.1088/1757-899X/226/1/012093.

[43] S. Stier. Open Semantic Lab. URL `https://github.com/OpenSemanticLab`. Last accessed January 8, 2024.

[44] S. Suresh and G. Chakaravarthi. RFID technology and its diverse applications: A brief exposition with a proposed Machine Learning approach. *Measurement: Journal of the International Measurement Confederation*, 195(February):111197, 2022. doi:10.1016/j.measurement.2022.111197.

[45] A. Tayi. The Internet of Things Is Digitizing and Transforming Science. *SLAS Technology*, 23(5):407–411, 2018. ISSN 24726311. doi:10.1177/2472630318788533. Special Issue: The Internet of Things in the Life Sciences Laboratory.

[46] G. A. Van Den Driessche, D. Bailey, E. O. Anderson, M. A. Tarselli, and L. Blackwell. Improving protein therapeutic development through cloud-based data integration. *SLAS Technology*, 28(5):293–301, 2023. doi:10.1016/j.slast.2023.07.002.

[47] W3C. Semantic web, 2015. URL `https://www.w3.org/standards/semanticweb/`. Last accessed January 8, 2024.

[48] M. D. Wilkinson, M. Dumontier, I. J. Aalbersberg, G. Appleton, M. Axton, A. Baak, N. Blomberg, J. W. Boiten, L. B. da Silva Santos, P. E. Bourne, J. Bouwman, A. J. Brookes, T. Clark, M. Crosas, I. Dillo, O. Dumon, S. Edmunds, C. T. Evelo, R. Finkers, A. Gonzalez-Beltran, A. J. Gray, P. Groth, C. Goble, J. S. Grethe, J. Heringa, P. A. t Hoen, R. Hooft, T. Kuhn, R. Kok, J. Kok, S. J. Lusher, M. E. Martone, A. Mons, A. L. Packer, B. Persson, P. Rocca-Serra, M. Roos, R. van Schaik, S. A. Sansone, E. Schultes, T. Sengstag, T. Slater, G. Strawn, M. A. Swertz, M. Thompson, J. Van Der Lei, E. Van Mulligen, J. Velterop, A. Waagmeester, P. Wittenburg, K. Wolstencroft, J. Zhao, and B. Mons. The FAIR Guiding Principles for scientific data management and stewardship. *Scientific Data*, 3:1–9, 2016. doi:10.1038/sdata.2016.18.

[49] N. Xu, M. Li, W. Kong, Y. Li, Y. Wang, H. Yu, and J. Yang. RFID with Multi-Sensing & Blockchain Empowered Digitalization of Chemical Inventory Management. In *Proceedings - 3rd International Conference on Next Generation Computing Applications, NextComp 2022*, pages 1–6. IEEE, 2022. ISBN 9781665469548. doi:10.1109/NextComp55567.2022.9932171.

[50] M. Zamiri, J. Sarraipa, and R. J. Goncalves. A Reference Model for Interoperable Living Labs Towards Establishing Productive Networks. In B. Archimède, Y. Ducq, B. Young, and H. Karray, editors, *Enterprise Interoperability IX*, pages 183–199. Springer International Publishing, 2023. doi:10.1007/978-3-030-90387-9_16.