Preprint Cambridge Centre for Computational Chemical Engineering ISSN 1473 – 4273

The Digital Lab Facility Manager: Using dynamic knowledge graph technology to automate operations of research laboratories

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

released: December 20, 2023

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

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

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

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

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

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

Preprint No. 316

UNIVERSITY OF CAMBRIDGE

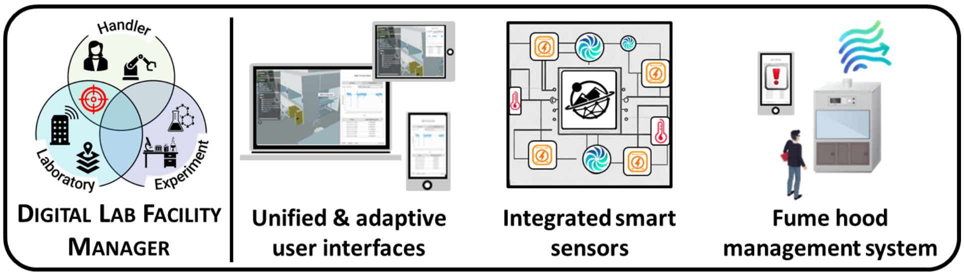
Keywords: Facility management , Laboratory automation, Air flow control, Internet of Things, Dynamic knowledge graphs

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

Automation in facility management is critical for enhancing efficiency and sustainability. However, significant challenges in interoperability and adaptability of digital tools are prevalent, particularly for specialised facilities like research laboratories. This paper presents a novel framework for digitalising and automating the management of specialised research laboratories using The World Avatar, a general all-encompassing dynamic knowledge graph. This framework is a versatile tool empowering users to effectively utilise data across various systems and formats without restricting them to a single software or protocol. Employing agents and ontologies, it enables seamless data sharing, computational reasoning, and gradual automation of tasks, addressing obstacles in interoperability and adaptability. Showcasing the capabilities of this approach, we tackle some common challenges in lab facility management, including cost-effective IoT sensor network integration with an existing BMS system and efficient airflow management for fume hoods via 'human-in-the-loop' interventions. We also developed unified platform-agnostic interfaces that complement the framework. These advancements represent a significant stride in the holistic digitalisation and automation of facility management, setting a foundation for future research facilities to achieve operational excellence and sustainability.



Highlights

- A novel framework for digitalising management of research facilities is demonstrated.
- BIM, BMS, assets, and domain knowledge are integrated via semantic web technology.
- Effective deployment and integration of IoT sensors in distributed control systems.
- Energy efficiency in fume hood air supply by deploying a human-in-the-loop approach.
- Dynamic knowledge graphs facilitate use of unified user interfaces for managers.

Contents

1 Introduction

In recent decades, cities have received increasing impetus to pursue sustainable development [15, 57]. This is unsurprising, considering that buildings contributed 39% of total global carbon emission in 2019 [1]. Consequently, the field of facility management has come under mounting pressure to achieving long-term sustainability [40, 43]. To address sustainability challenges and realise the concept of "smart buildings", a range of digital technologies, including the Internet of Things (IoT), Building Information Modelling (BIM), and digital twins (DT), have been adopted [4, 30, 48].

However, substantial challenges need to be overcome. One such challenge is achieving interoperability between domain-specific data from different technologies to support and facilitate knowledge discovery [50]. This requires integrating and optimising data usage to enable collaboration between systems. Additionally, there is a need to expand the functionalities of existing digital solutions, from monitoring and displaying information only to being able to produce valuable insights that encompass safety and controllability [33]. Furthermore, stakeholders in facility management often resist the adoption of digital technologies due to the cost and time required for implementation [21, 43]. These financial and human capital investments required can present obstacles to widespread adoption and implementation of digital solutions in smart buildings.

Existing research on facility management has predominantly focused on the management of general facilities, overlooking the specific requirements and challenges posed by specialised facilities. Research laboratories, for example, present unique challenges for facility managers due to irregular equipment usage, the disposal and re-acquisition of costly one-use resources (*e.g.*, test tubes and reagents), and safety considerations. Laboratories have also been identified as significant energy consumers within universities, often consuming up to four times the energy of equivalent office buildings [28]. One notable aspect of laboratories is that they are part of larger research facilities that usually have diverse buildings, functionalities and equipment over a large land area. One example is the Sandia National Laboratories Albuquerque campus in the USA which spans an area of $67\,\mathrm{km}^2$ with gross $0.6\,\mathrm{km}^2$ of general office, engineering, research & development, and manufacturing facilities [7]. Given the scale of the infrastructure and facilities, the resulting complexity of managing various aspects such as safety, ventilation, airflow control, and energy consumption necessitate the integration of an IoT system to achieve their sustainability goals [7]. Consequently, there is a pressing need to integrate various technologies in a seamless and structured way to manage these specialised and complex facilities and address their sustainability challenges.

The use of Semantic Web technology, such as knowledge graphs (KGs) has emerged as a promising digital solution to meet sustainability and efficiency goals in buildings while addressing the challenges associated with investments in financial and human capital to achieve the vision of smart buildings [16, 46]. This technology enables the abstraction of both resources and data using a common language, facilitating task allocation and the sharing of results among participants. One of these Semantic Web projects is The World Avatar (TWA) [34] which aims to digitise and automate complex processes through a holistic and goal-driven approach. Software agents are developed to allow dynamic and continuous incorporation of new concepts and data into the knowledge graph while maintaining connections to existing information. As the KG expands, it captures the data provenance of experimental and managerial processes, effectively acting as a living representation of the real world.

This work is part of a series of articles that argue for holistic laboratory automation encompassing all aspects of scientific research laboratories within a dynamic knowledge graph [54]. 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 proof of concept for a digitalisation and automation framework that focuses on the management of research facilities using TWA technology at the intersection of underlying infrastructure and moving parts. While the focus is on research laboratories, this methodology is relevant to any specialised facility with a diverse range of goals and stakeholders. We present a unified user interface that can monitor and manage laboratories through automation to achieve operational optimisation and cost reduction. This has the potential to elevate classical smart buildings with embedded automated tasks, ultimately supporting the realisation of a fully-autonomous "AI Scientist" including reasoning and decision-making capabilities [31, 54].

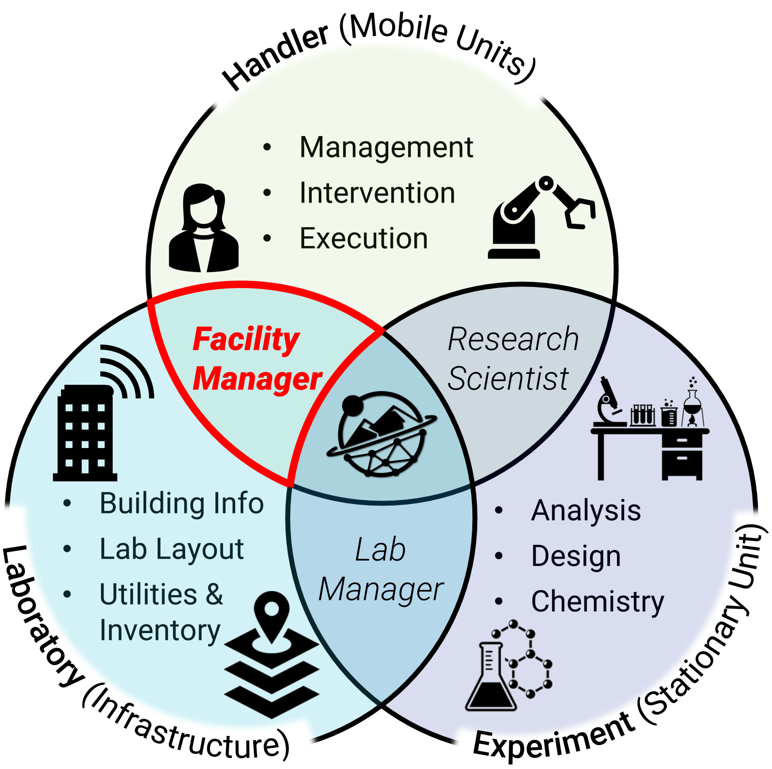


Figure 1: *An overview of intertwined aspects of a chemical research laboratory that need to be represented by a connected lab digital twin, adapted from [54]. This paper focuses on the automation of tasks related to facility management.*

2 Background

Facility management involves the maintenance and monitoring of physical infrastructure and assets in a facility to ensure compliance with business, hygiene, safety, and sustainability requirements [33, 43]. The responsibilities of a facility manager (FM) encompass a wide range of tasks, from day-to-day building operations such as maintenance, repairs, and inspections, to managing vendors, tenants, and other stakeholders. Their role becomes more complex when managing niche facilities like as laboratories [2]. In this situation, the scope of a FM must be expanded beyond the traditional focus on people and physical infrastructure to incorporate research-specific knowledge. FM in chemical laboratories, for example, must ensure the optimal conditions for safe handling of hazardous chemicals - including temperature, humidity, and air flow [7, 42]. Furthermore, laboratories are typically clustered in a larger research campus, confronting FMs with the complexity of managing diverse facilities in terms of layout and function. Several key challenges have been identified by Aishah Kamarazaly et al. [2], including emergency management and business continuity planning, sustainability, compliance, operational cost and efficiency, and asset and space utilisation. These considerations are critical to ensuring personnel safety and the productivity of scientific research.

These tasks encompass not only immediate responsibilities related to day-to-day building operations such as maintenance, repairs, and inspections, but also managerial activities involving vendors, tenants, and additional stakeholders such as janitorial staff, security personnel, maintenance personnel, engineers, lab managers, supervisors, group leaders, and program managers. As illustrated in Fig. 2, particularly those tasks with strong cross-domain requirements and involvement of different stakeholders – such as challenges around sustainability and energy efficiency – tend to have the largest impact.

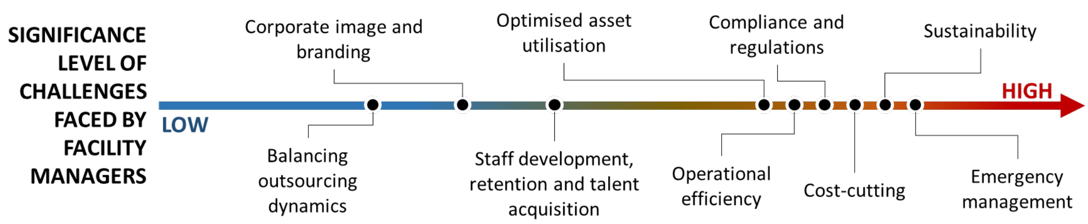


Figure 2: *Significance level of challenges faced by a university facility managers [2].*

In research facilities, regulating environmental factors like ventilation, temperature, humidity, air quality, power, and lighting is essential. Employing smart sensors within an IoT framework offers an effective way to track these elements. Yet, merging these sensors with existing building management system (BMS) poses significant challenges. A notable example is the optimisation of fume hoods' airflow which contributes heavily to the total energy consumption [25, 28, 52] but also strongly affects overall ventilation and pollutant concentrations [27, 36]. The overarching challenge lies in effectively controlling and maintaining a comprehensive overview of these diverse factors, as they often need to be managed using a variety of different tools.

2.1 Automation in (laboratory) facility management

In managing the increasing complexities and challenges faced by FM, digitalisation offers a promising solution. Despite the diverse nature of FM tasks, many of them are repetitive and amenable to automation. Technologies like BIM, BMS, IoT, and reality capture technologies have been increasingly employed to digitise and automate facility management tasks in various capacities to enhance interoperability, efficiency and sustainability [51, 61]. For example, FMs rely on BMS for monitoring electrical, plumbing, heating, ventilation, and air conditioning systems; BIM for accessing physical asset and material information; and other software for tracking readings from diverse temperature and gas sensors provided by different vendors and spread throughout the facility. In addition, access control systems and occupancy sensors are necessary to manage facility access and optimise space utilisation [18, 30]. Through the automation of FMs' tasks, digitalisation supports the broader shift of the operational FM role towards a more strategic one [2, 33]. In particular, automation can free up human resources for creative and strategic endeavors, including engaging in high-level decision-making, implementing innovative solutions for non-automatable issues, and managing stakeholders and professional development.

Digital solutions are often fragmented across different platforms, each serving specific needs and equipment. For instance, FMs need to access information from a BIM model and/or a spreadsheet to locate equipment for planning, coordination and resource allocation. For special facilities with safety-critical activities like research laboratories, additional tools are necessary to ensure compliance with corresponding regulations [52]. The fragmented storage of data across disparate user interfaces is an obstacle [16, 49] to a comprehensive, integrated and real-time view of resource utilisation, which limits the capacity of FMs to optimise resource allocation at the building level. Moreover, research facilities must adapt to the dynamic demand for various layouts, equipment, and research activities to support the latest scientific developments while ensuring sustainability. When scaling these requirements across laboratories, smart buildings become essential [37, 62] to automate these tasks and improve the productivity and efficiency of research facilities . To achieve this automation, BMS and IoT platforms must be integrated to enable communication among various systems and sensor networks in a unified interface to assist FM. While some standards exist, such as Industry Foundation Classes (IFC) for BIM, most current integration strategies are vendor-specific. They typically use plugins or tools designed for a specific (commercial) software, without integration to other software. Commercial solutions also restrict the access to these solutions, by requiring investments for learning and implementation. Therefore, these solutions fail to be generalised and scaled to other technologies or use cases.

Previous studies have explored the potential of vendor-agnostic digital solutions. For example, Da Conceic'ao et al. [18] presented a platform for IoT-based automation particularly for smart labs, including access management and lighting control. Valinejadshoubi et al. [61] developed an IoT-BIM integrated alert system for buildings without BMS, which is restricted to low-rise and mid-rise buildings. Lin and Su [35] proposed a mobile-and BIM-based facility maintenance system that showed the 3D model and meta information of each facility, but it did not involve real-time data and required manual updates of the BIM models. Another study [51] combined a BIM-based AR application with a location-based management system (LBMS) to enhance the construction performance, but it faced alignment errors between the virtual model and the real environment, and needed a case-specific mechanism to define locations according to LBMS. A hybrid model based on IoT and mobile applications to reduce electricity consumption in a smart space has also been introduced [45]. Although it has low development and deployment cost, the conditions were hard-coded and inflexible to a dynamic environment. Moreover, it could not incorporate BIM or BMS, which are essential for large-scale buildings such as a research facility. Similar limitations have been reported by Khan et al. [29] and Da Conceic'ao et al. [18], making it difficult for these IoT models to scale up and cater to larger spaces with more devices.

The landscape of solutions for facility management is diverse yet fragmented, both in terms of tasks and accessibility. Despite the prevalence of mobile and web-based applications in research [18, 45], surprisingly few cater specifically to this field, with existing options often lacking in comprehensive control functions. For instance, mobile applications for BIM typically offer only viewing capabilities, and BMS apps tend to be vendor-specific and operationally limited. While mobile apps for smart spaces and IoT do offer control functions, they generally do not integrate with BIM or BMS. A fundamental challenge in digitising facility management is the poor **interoperability** between existing data and systems [54]. 'Interoperability' refers to the capability to operate systems and connect information horizontally among similar data as well as vertically across functions and scales. This is particularly evident in the development of Digital Twins for research facilities, laboratories, and experiments [6, 41, 48], where integrating varied environmental monitoring systems into facility infrastructure poses significant difficulties. Key technologies like BMS and IoT platforms often lack generalisability and scalability in their current integration approaches. The adoption of Semantic Web technologies has been proposed, aiming to overcome these challenges faced by FMs [16, 58].

2.2 Semantic Web in (laboratory) facility management

The Semantic Web is heralded as next evolutionary step of the World Wide Web providing human- and machine-readable data enabled by ontologies as a Web of Data [12]. An ontology is a comprehensive and structured representation of knowledge such as their concepts and relationships within a specific domain [5]. These representations provide a shared understanding and standardised definition of knowledge to align both human and machine perspectives [12]. By annotating data through ontologies, the Semantic Web can facilitate seamless data sharing and address the challenge of interoperability today.

Given the inherent complexity related to this technology, we will briefly explain some of the concepts relevant to the discussion. An ontology consists of two main components: the terminological component (TBox) and the assertional component (ABox) [63]. The TBox defines concepts, their hierarchical relationships, and associated properties through a taxonomy or classification system for domain knowledge. The ABox instantiates the TBox concepts to represent real-world entities as instances. These instances are all resources available on the internet identified by unique internationalised resource identifiers (IRI). When looked up, these instances provide meaningful information about individual things, their attributes, and relationships. The data is usually represented and stored in the form of a subject-predicate-object triples based Resource Description Framework (RDF).

Some applications extend the representation capabilities of ontologies using KGs to derive new cross-domain knowledge and support larger applications. A KG is a directed graph making use of ontologies, where the nodes refer to entities of interest and the edges represent their relationships with each other [26]. One of the distinguishing contribution of the Semantic Web lies in enabling a new form of standardised human- and machine-readable data model through ontologies and KGs that can be retrieved and further processed in applications.

In the facility management field, a wide variety of specific ontologies for everything from smart appliances to building materials exist, but the heterogeneity of information remains a significant challenge [23]. Existing Semantic Web applications therefore predominantly employ either the Brick or the RealEstateCore (REC) ontology. Both ontologies are designed to represent buildings and their subsystems inclusive of HVAC, lighting, security and their IoT network [10, 24]. Despite their far-reaching design ambitions, we argue that such a monolithic approach to ontologies can be impractical to maintain or scale for three reasons: First, the comprehensive design and strict definitions of these ontologies are an obstacle for further extension to more domains like laboratories. Both aforementioned ontologies currently contain at least 100 classes covering buildings and building elements as well as BMS and sensor concepts, with partially conflicting or incompatible definitions that require major efforts to align and reuse. For example, developing the REC ontology from scratch proved more efficient than extending the BMS and IoT representation within the Brick ontology to include the perspective of real estate stakeholders and their interaction with these systems [10, 24]. In future, when extending or scaling beyond their intended design scope, such an approach will require greater and greater effort. Second, the geometry concepts for both ontologies are insufficient to accommodate the various geometry representations used in the industry, specifically for BIM and geographic information system (GIS) technologies. The Brick ontology is missing geometry representation, whereas the geometry representation in the REC ontology is coupled tightly with specific GIS formats. Last, compared to smaller ontologies, these monolithic ontologies also require a steeper learning curve in time and resources for both new and existing users. Consequently, the current monolithic and cumbersome design of these ontologies are restrictive in addressing the technological challenges faced by FMs.

With the emergence of smart building technologies which allow FM to monitor and control operations via various platforms, these Semantic Web approaches are now also confronted with the challenge of **adaptability**. While there have been attempts to represent buildings semantically and integrating BMS, BIM, and IoT [16], no well-accepted standards have been established. Handling the data abundance provided by the ubiquity of large sensor networks can be a serious issue for knowledge graphs. Especially live data is complicated due to the constant stream of new data at different sampling frequencies and their sheer size which becomes too large to store as individual triples. Furthermore, to actually leverage the inherent interoperability of the Semantic Web, flexible user interfaces are required for access of information stored in the KG, ultimately replacing the various specific platforms and softwares available today. While some applications in the space of Semantic Web have dedicated graphical user interfaces (GUIs), they usually remain complicated to navigate as knowledge of programming or query languages is necessary. In particular, the development of appropriate mobile apps has been underutilised, barring a few examples in the fields of food science [39] and clinical medicine [38].

2.3 Digital Lab Framework within The World Avatar

The World Avatar (TWA) has been developed as a potentially 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 [34]. DTs are realistic digital representations of assets, processes, or systems in the built or natural environment that create the opportunity for positive feedback into the physical world [13]. The basis of its system is the dynamic knowledge graph (dKG) technology, which continuously updates and restructures the KG and its data based on the real-time status of the physical world through an ecosystem of autonomous agents [34]. TWA facilitates a distributed architecture that enables access from any location and facilitates data querying across multiple connected domains. Within this ecosystem, two main concept types coexist and are both described by modular ontologies: instances represent interlinked data, while agents are autonomous software applications capable of acting on the data and exchanging information through the KG.

By acting on real-time data through computational agents, the dKG can describe the behavior of complex systems and perform tasks such as updates, analysis, decision-making and control of real-world entities [34]. Notably, the TWA approach distinguishes itself from traditional Semantic Web implementations through the seamless integration of real-time dynamic data, knowledge, models, and tools in a distributed architecture. 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 BIM and GIS [49]. As a dynamic knowledge graph, TWA also keeps evolving over time, partially by autonomous agents deriving new knowledge from live data [9], to address **adaptability** challenges. Also to this end, efficient ways to represent time series within the KG for monitoring and forecasting have been implemented [9].

TWA originated from its initial application in addressing decarbonisation challenges within the chemical industry of Singapore [22, 65]. Over time, it has evolved to encompass a wide range of concepts, including but not limited to chemistry, chemical processes, laboratories, power systems, urban planning, and environmental planning [9, 34, 44, 66]. Recently, we introduced the "Digital Lab Framework" (DLF) [54], offering a holistic approach for goal-driven automation of research laboratories. The framework involves the deployment of comprehensive digital twins, in which humans and infrastructure are inherently considered. In contrast to most laboratory automation projects, we focus on integrating the periphery of actual experiments, including building, layout, resources, and utilities [49, 54].

Working towards the vision of a fully autonomous AI scientist [31], tasks traditionally related to facility management need to be incorporated and automated within the same framework. Such a "Digital Lab Facility Manager" (DLFM) is anticipated to support FMs in managing so-called self-driving laboratories as well as non-automated or hybrid research facilities. Analogous to the experimental research conducted, a DLFM can act with a increasing degree of autonomy by deriving its own goals, executing optimisation tasks in a distributed manner [8], considering its direct and indirect environment, and taking over managerial tasks related to inventory and energy consumption, which would allow results to be reproduced and human resources to be freed up for creative tasks.

3 Developed Methodology

In order to automate the challenges faced by FMs in research facilities as described in section 2, the obstacles posed by the current fragmented landscape of available solutions need to be overcome. These include interoperability of BIM, BMS, and IoT systems as well as adaptability of interfaces to allow for automation and scale-up. To this end, we developed methodology using Semantic Web technology based on the key ideas behind the DLF [54]. In this section we describe the methodology developed to establish a digital twin for enhanced management and automation of research laboratories. The underlying structure provided by TWA consists of three layers as illustrated in Fig. 3.

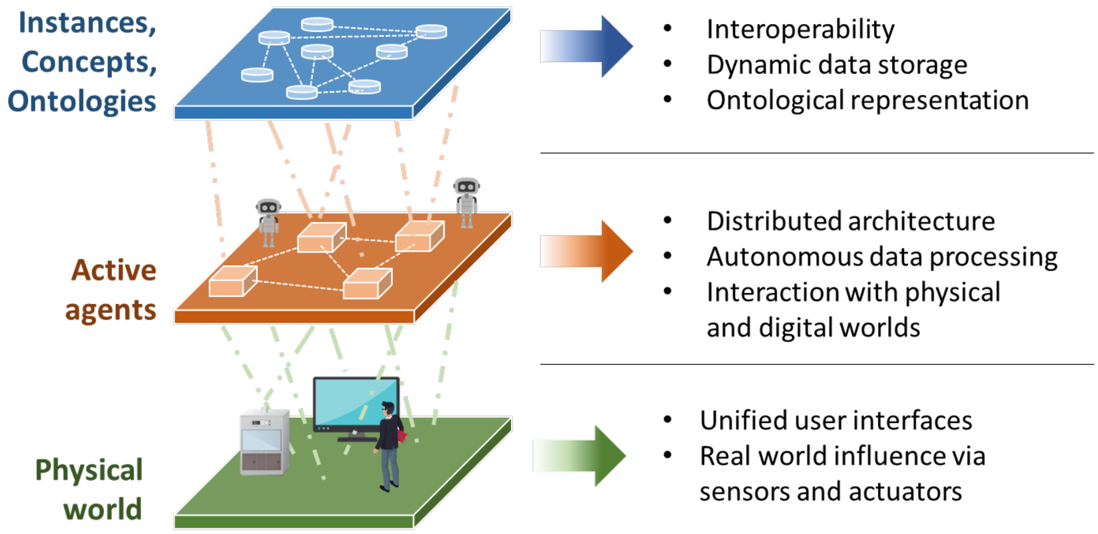


Figure 3: *Three-layer TWA structure as applied to enhanced management and automation of research facilities.*

The top layer represents the ontological representations used in the development of the laboratory digital twin. The interlinked ontologies are described in detail in section 3.1. The middle layer represents active agents which keep the KG up to date, coordinate data exchange between different servers, and automate various tasks related to FM. Server architecture and agent development are described in section 3.2. The bottom layer represents the physical world which interacts with its digital twin in multiple ways – *e.g.*, sensor networks perceive changes to trigger KG updates. Most importantly, a variety of human-machine interfaces facilitate support for FM and other stakeholders, including web-based 3D laboratory visualisations and mobile applications as shown in section 3.3.

3.1 Ontological integration

In TWA, domain ontologies are developed as modules that can be linked together to represent and establish connections between diverse concepts in disparate domains. In accordance with the DLF philosophy of taking on a holistic view of the laboratory environment [54], the ontologies used are supposed to offer a balanced approach between generality and specificity across different scales and domains. Fig. 4 illustrates an extract of how these modular ontologies were integrated to represent interactions between the laboratory environment, laboratory building, and building occupants by connecting domain knowledge with sensor and managerial information at the building level. Notably, this ontological design achieves the integration of specialised knowledge (*i.e.* chemistry with OntoCAPE) with managerial information at the building level, including static BIM and dynamic BMS data, using modular design as suggested by Gilani et al. [23]. Moreover, sensor data and controllable values from decentralised IoT systems and monolithic BMS are represented and thereby accessible within the same conceptual structure, including time series. This represents a work in progress which is extended continuously, currently focusing on fume hoods and smart sensors to support measures aimed at resource efficiency and sustainability as described in section 4.

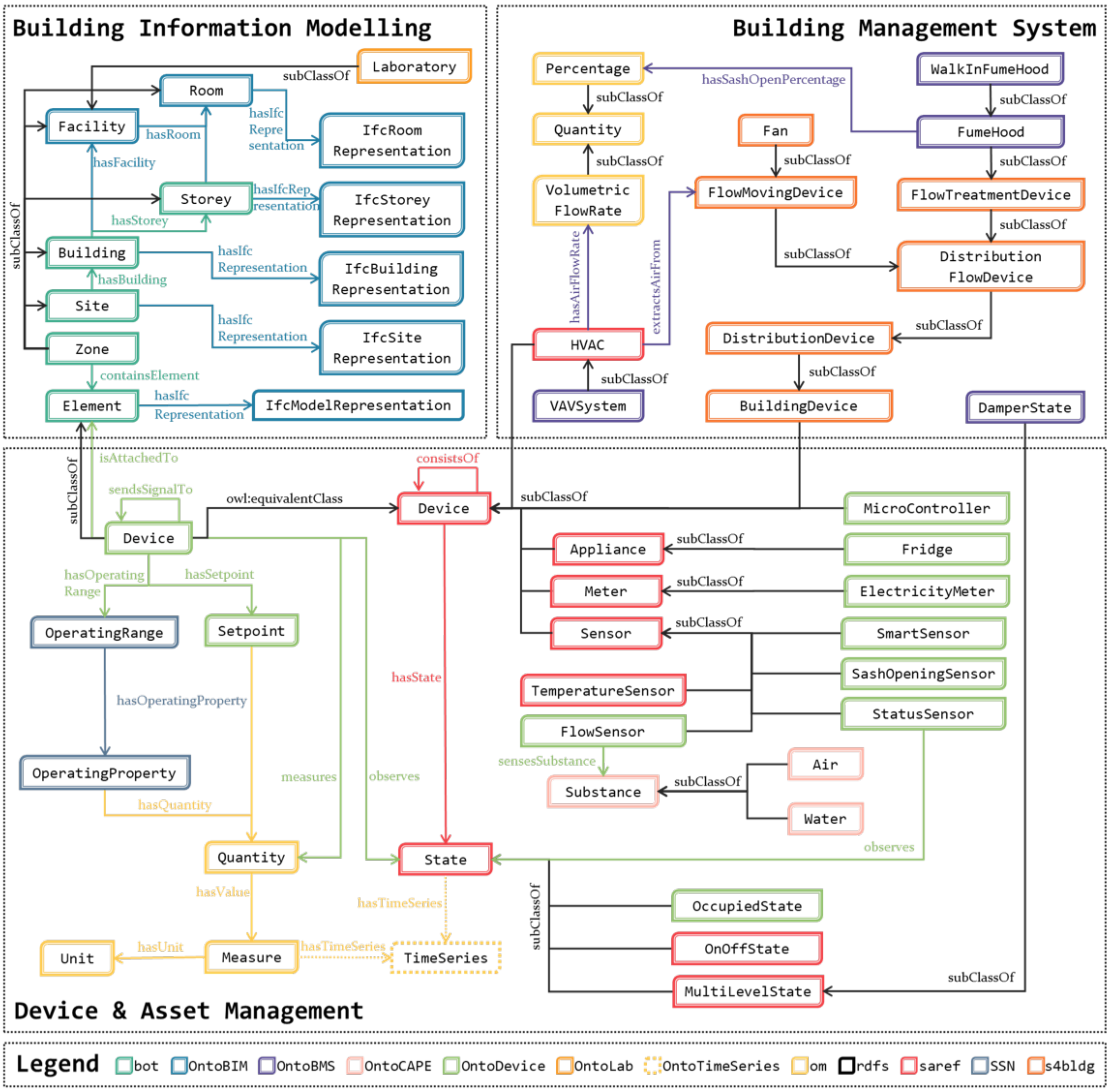


Figure 4: *An extract of the ontology design (TBox) to integrate building, BMS systems, devices, and sensors for cross-domain interoperability. This is only a small part of all concepts and relations developed and used, see appendix A.2.*

A detailed description of the developed ontologies, namely "OntoBIM", "OntoBMS", and "OntoDevice", and their roles can be found in Tab. 1. Briefly, this standardised ontological framework attaches semantic descriptions to the physical, logical, and virtual assets in a typical research facility and enable cross-domain interoperability among the building, facility, assets and their static and dynamic data to support holistic laboratory automation. Although the current study focuses on a specific research laboratory, these ontologies can be reused and applied to almost any research laboratory.

Table 1: *Description and role of the core ontologies used in this work.*

Ontology	**Specification**
`OntoBIM`	This ontology defines a topological hierarchy connecting spatial entities – including zones, sites, buildings, stories, and rooms – to physical elements. Each of these entities can have a corresponding IFC representation containing geometry and other metadata if available [49]. Its role is to integrate different domain ontologies for various assets to their spatial location, facilitating the interoperability of buildings and assets.
`OntoDevice`	As an extension of the "Smart Applications REFerence" (SAREF) [19, 20], this ontology defines general concepts and properties that are applicable to most devices. It can represent and link multiple subdevices to primary devices, states, and roles (*e.g.* a sensor can measure a quantity or observe a state). Its role is to represent the general static properties of devices and their dynamic measured quantities, enable links to the topological hierarchy in OntoBIM and specialised domains like BMS. Measured quantities are represented using concepts from the "Ontology of units of Measure" (OM) [55, 56] and "OntoTimeSeries".
`OntoBMS`	This ontology extends OntoDevice and "SAREF extension for building" (S4BLDG) [47] to include additional concepts and properties pertaining to the BMS and its devices. Its role is to represent BMS-specific knowledge and link it to the topological hierarchy through OntoDevice.

3.2 Servers, protocols, and agents development

Driven by the DLF design philosophy of "Distributed and connected digital twins" [54], parts of the KG can be hosted on different servers as their ontological links shown in section 3.1 are secured by using IRIs. TWA also deploys various computational agents on distributed cloud-based or local servers to provide Semantic Web services that act upon the KG or real world to fulfill its tasks and goals. Namely, agents are responsible for instantiating, retrieving and modifying data, running simulations, forecasting time series and controlling real-world entities. The primary communication mode between agents and the dKG is through "Hypertext Transfer Protocol" (HTTP) requests.

The instantiation of static data is usually done directly by dedicated agents without any middleware involved, using data from different sources, types, and formats to effectively create ABoxes. For example, the `Ifc2OntoBIMAgent` instantiates BIM models in the standardised IFC format in the knowledge graph, including metadata. Now, a second set of agents can retrieve this data from the KG to generate outputs, *e.g.* for the visualisation interface. To enable real-time information access, a third set of agents continuously runs in the background to update the KG based on the latest real-world status, which is then propagated to the web visualisation interface. To minimise the number of active agents, ensure consistency across protocols, and enable precursory data sampling and filtering, so-called middleware is used. The dKG's distributed architecture is able to address the poor interoperability and scalability as well as security risks of middleware solutions in the implementation of IoT systems [60].

IoT play a key role to support FM tasks by generating dynamic data as a time series, which enables trends to be monitored and analysed. For example, sensors can measure the building and its spaces' environmental conditions, based on which actuators and controllers then act automatically to either maintain optimal environmental conditions or notify stakeholders if an abnormality is observed. However, often small heterogeneous devices with no standard architecture and protocols are used [60]. Given their small size, these devices also do not have sufficient memory capacity to host the massive time series data generated as well as device-specific libraries required to operate [60]. A middleware solution such as the open-source ThingsBoard platform or a commercial BMS is typically required to communicate and transfer data between these devices to their application programs [60]. For example, ThingsBoard can store time series data collected from Wi-Fi enabled microcontrollers via the MQTT (MQ Telemetry Transport), HTTP or Constrained Application Protocol (CoAP). Meanwhile, commercial BMS solutions often use the Building Automation and Control Networks (BACnet) protocol, complicating the integration with IoT devices. Additionally, the memory and energy constraints of these small devices is a hindrance to the adoption of stronger security techniques requiring more processing power [60]: as IoT applications generally directly access these middleware solutions to retrieve data, applications become vulnerable to security threats.

The flexible design of the dKG's agent ecosystem complement existing workflows to transfer the data from the middleware solutions to the end user application. First, existing middleware are crucial to store and remove time series data collected within the IoT devices, addressing their limited memory and energy capacities. During downtimes, these middleware serves as backups to temporarily store data, which can be retrieved by agents once the server is available. Second, these agents can instantiate time series data into a standardised format in spite of the distinct data communication protocols. End user applications merely need to retrieve these data from the KG instead of directly interfacing with the middleware or IoT devices. Third, the interoperability achieved by the dKG enables customised sensor deployment for enhanced measurement values in one interface. Sensors from different vendors and systems can be deployed together to generate data in the same server through the ontologies and agent ecosystem. Last, the agents anonymise and protect the access to any data within the KG, middleware and devices from end user applications. Applications simply retrieve the data from the KG through an agent but can be blocked from influencing these devices and middleware based on user privileges.

3.3 Human-machine interactions

While the deployment of DLF aims to eventually automate all tasks related to FM, developing seamless human-machine interactions is still key where either humans can carry out a task more efficiently than a robot or managers need to monitor and control critical operations [54]. TWA can fill out this dual role, offering active and passive interactions: On one hand, autonomous agents deployed in the KG can actively alter the real world based on inputs by controlling robotic systems, *i.e.* actuators. Often times, facilities that are in the process of automating their processes step-by-step will not meet the technical requirements for this kind of autonomy. To bridge this "interim technology gap", notification agents (*e.g.*, `EmailAgent` in appendix A) can be deployed as a first step to instruct a "human-in-the-loop" to carry out the necessary actions in the real world [54].

On the other hand, (passive) digital interfaces can be customised to visualise the real-time information stored within the dKG. TWA user interfaces in general are designed to represent and visualise the geospatial, geometric, and semantic aspect of any physical element, ranging from assets to rooms, facilities, buildings, and cities [49]. To ensure ease of access for all stakeholders with varying prior knowledge and technical equipment, we developed two dedicated applications to support decision-making processes: a browser-based web application and a mobile application. These interfaces are supposed to monitor facilities through enhanced identification, collection, and visualisation of sensor and BIM data within the facility and were developed based on literature reviews, expertise in handling BMS, and feedback by FMs.

A web-based visualisation of BIM-GIS building models for the CARES laboratory on the National University of Singapore (NUS) campus is shown in Fig. 5. BIM enables the accurate representation of smaller scale entities and their properties at the building level such as the work spaces, BMS systems, sensors, and equipment like fume hoods or a synthesis robot [8]. BIM also provides a topological relationship that can spatially link devices and sensors to each other and the building, enabling more accurate representation and complex control of object interactions in physical vicinity. Consequently, this creates a comprehensive representation of both the building and laboratory embedded with functional information for navigation, monitoring or interventions. The incorporation of GIS models facilitate assessments at larger scales, ranging from neighbourhoods and cities to regions. This integration extends system boundaries by incorporating surrounding buildings within the research institution, thereby expanding the potential use cases to include campus- and city-level considerations. Although GIS has not been fully utilised within this work, its availability is crucial to scale our work, allowing managers to monitor operations and allocate resources efficiently at various scales.

At present, 3D visualisation of urban models is only available on the web-based interface. Given the hardware limitations and small screen sizes of mobile devices, the layout of the mobile application has been simplified to offer a clean user experience. Users are still able to access the 3D visualisation on their mobile devices through the browsers if they wish. Notably, while the BIM-GIS geometry is not visualised, their topological relationships are still available within the application. Users are able to find and select any asset within the building from a dropdown menu. Depending on the device type, related metadata or real-time measurements are also available for visualisation. Furthermore,

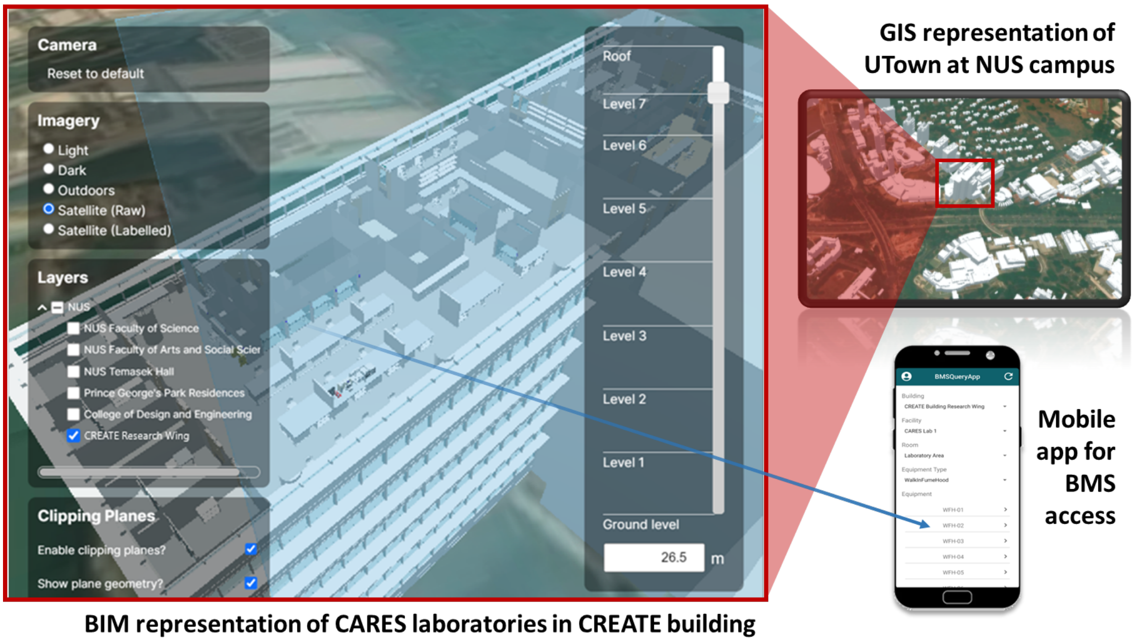


Figure 5: *An example of the web-based 3D visualisation of multi-scale urban models for the CARES lab at UTown within platform-agnostic user interfaces.*

users with appropriate access rights can change corresponding set points directly within the application.

4 Exemplary applications

This section details three TWA applications enabled by the developed methodology to enhance the interoperability and adaptability of digital systems embedded within laboratories to support sustainable facility management functions. Based on the challenges identified in section 2, we present the deployment of integrated smart sensors, the reduction of energy consumption via effective airflow management, and the introduction of flexible user interfaces for facility management.

The laboratory under investigation in this paper is the CARES lab in University Town (UTown) on the NUS campus which supports multi-domain and multi-disciplinary research activities, including combustion research, electrochemistry, and pilot plants. At the campus scale, the NUS campus has a blend of building types and spans a reasonably large area of 170 ha. This laboratory is embedded within the CREATE building's topological hierarchy as illustrated in Fig. 6. Given the complexity of facilities, functions and requirements across scale, the chosen facilities are suitable to test the TWA methodology and devise strategies to manage laboratories across scales. Moreover, Singapore's sustainability policy for energy intensive buildings sets the scene for a recent push for energy conservation efforts in labs across the country [14, 53, 59].

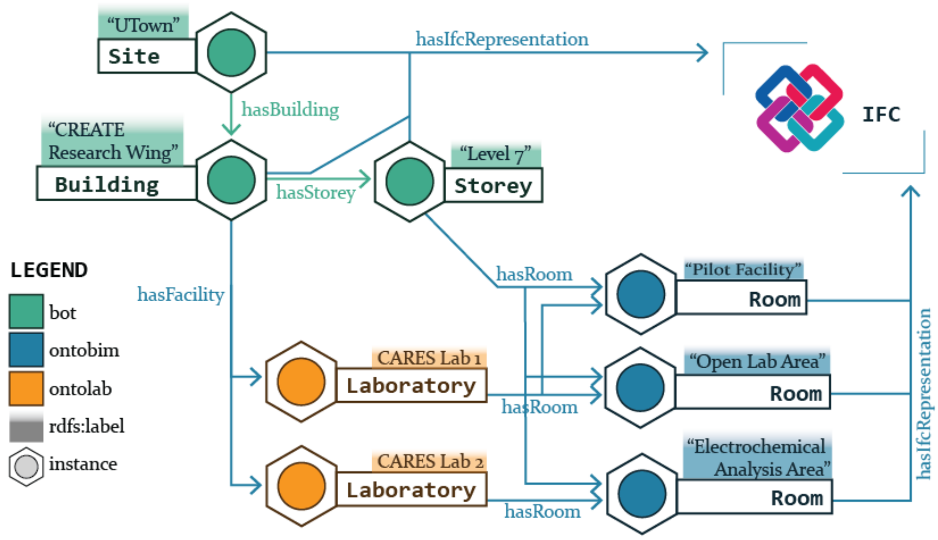


Figure 6: *ABox snippet for building and facility topological hierarchy: Linking assets to their physical location for navigation and inventory management.*

Implementing the three applications mentioned at our facility had significant impact. Tab. 2 lists some FM tasks that have benefited from the TWA implementation in terms of grater data granularity, convenience, and productivity.

Table 2: *Impact of TWA-DLF applications for three facility management tasks.*

Tasks	Conventional methods	New opportunities
Monitoring BMS and IoT sensor networks within the facility.	A laptop must be carried when inspecting different locations. Adoption of diverse softwares and platforms.	A single interface accessible through any computer or mobile devices that display all information even during site inspections.
Monitoring the sash opening of the fume hoods.	Daily walk around the research facility to physically check.	Automated email alerts for unoccupied fume hoods with open sashes.
Controlling airflow and other BMS set points for thermal comfort or energy savings.	A dedicated desktop computer in a restricted area needs to be accessed for any changes.	Set points can be accessed and changed via the mobile app based on the privileges associated with the FM's user account.

4.1 Deployment of smart sensors for facility management

The first application involves the deployment of customised smart sensor setups that can measure and optimise the environmental conditions of laboratory equipment. The specific system at hand consists of a DHT-22 temperature and humidity sensor, an electrical meter, and a cooling fan. It was deployed to monitor the electrical consumption of a chemical cooling system and maintain an optimal internal temperature to reduce energy consumption and ensure the sensor has a robust operating capacity. This application showcases three core capabilities of the DLF applied to facility management: First, the seamless integration of IoT devices with different architectures, protocols etc. within a single framework including BMS. Second, the representation of self-knowledge within digital twins that allows agents to find and control variables autonomously. And third, setting up decentralised control systems within a distributed data architecture.

The first two key capabilities are illustrated by the instantiated ontological representation of the smart sensor set up and their measurement time series shown in Fig. 7. At the sensor level, the representation links the sensor's subdevices and their time series and properties to the smart sensor concept itself. At the laboratory level, the integration of BIM-GIS adds a geo-spatial topological aspect to link the location of an asset within a building and the sensors attached to this asset. By linking the sensors and their measurements alongside their geo-spatial context, TWA provides a rich semantic representation of the real world with cross-sector data interoperability across scales.

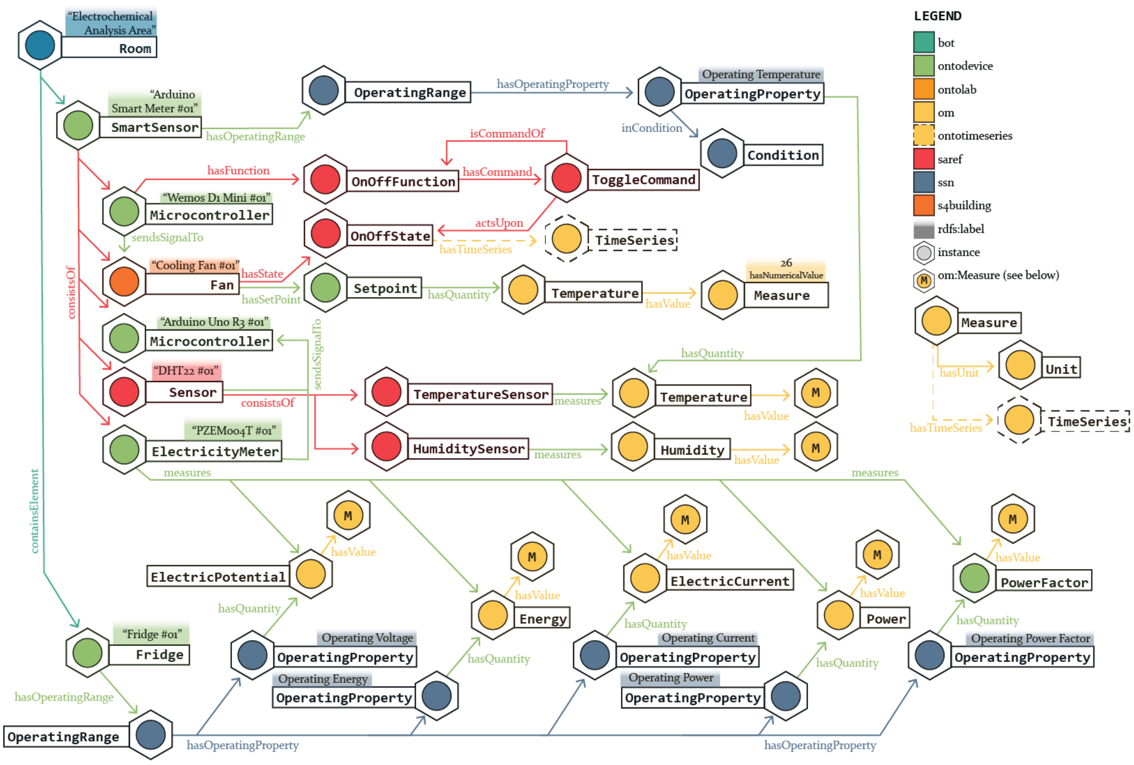


Figure 7: *ABox snippet for a smart sensor system deployed in the CARES lab.*

The workflow of a sample customised smart sensor is illustrated in Fig. 8 using the agent ecosystem and ontologies described in section 3 enabling real-world interactions to support automatibility in FM functions. For instance, the smart sensor is able to monitor and optimise its internal temperature through the `ThingsBoardAgent` and `ESPHomeAgent`. The ThingsBoard agent monitors and updates the temperature data from the thermistor in the KG. This data is then queried by the `ESPHomeAgent` to check if internal temperature exceeds a predefined threshold temperature ($T_{\text{threshold}}$). If so, the agent sends a request to the cooling fan to again reduce the system temperature. This function is also available within the human-machine interfaces at an authorised user's discretion to introduce flexibility into the system for human interventions. By reducing the temperature, the sensor can reduce its own energy consumption, increase its estimated service life, and mitigate safety risks created by perpetually overheated equipment.

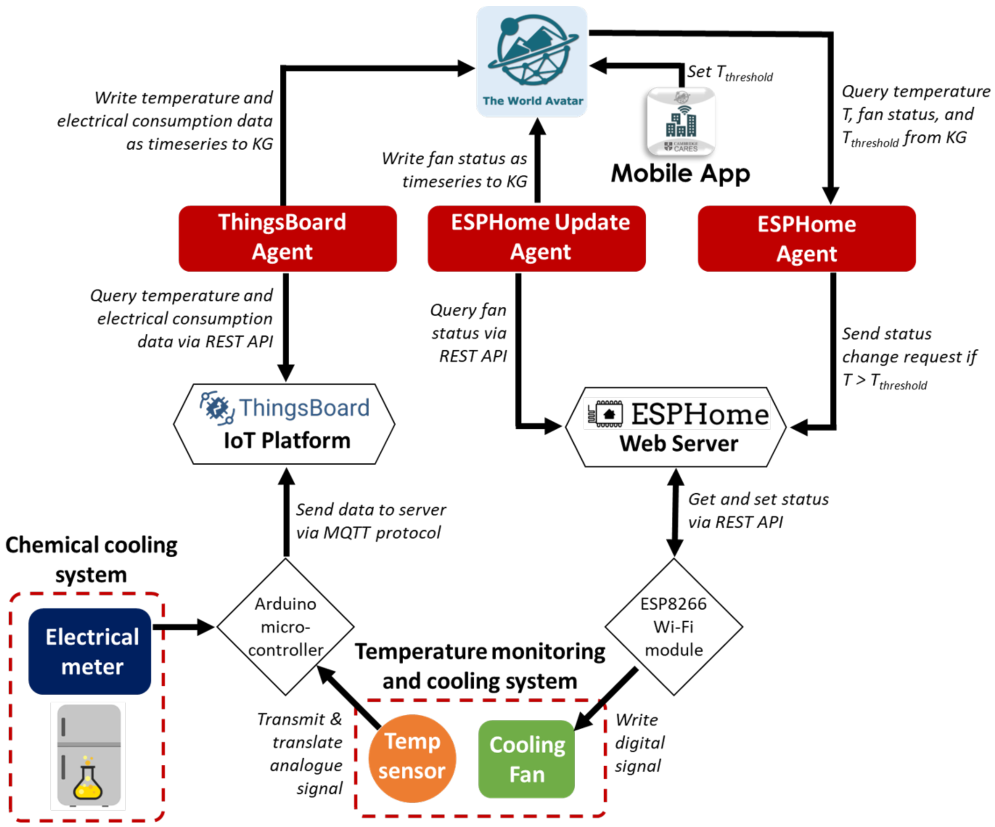


Figure 8: *Workflow of the servers and agents involved for described smart sensor system.*

This is of course a trivial task in itself which was carefully chosen and designed to clearly demonstrate the aforementioned third capability: while temperature sensor and cooling fan could be easily controlled by a single microcontroller, they are intentionally separated and connected to different microcontrollers. This way, agents within TWA assume the control function independent of hardware and its physical distance. The heterogeneity of protocols, middleware, signal types, and microcontrollers used attests to the interoperability and adaptbility of the proposed framework.

Although the scope of this work is localised within one sensor set up, the automated control functions demonstrated can be further replicated and expanded to more meaningful applications such as autonomous optimisation of chemical cooling systems and even buildings. This modular approach further allows for accelerated and automated deployment of such smart sensor systems, independent of underlying hardware and software. In appendix B.3 we introduce a set of tools to automate instantiation of digital twins as depicted in Fig. 7 as well as deployment of relevant agents as shown in Fig. 8.

4.2 BMS fume hood management for laboratories

The second application involves the development of a laboratory DT managing fume hood airflow. In a typical laboratory, fume hoods are one of the most significant energy consuming equipment [25]. When running experiments, lab users require their sash to be open which causes an increased airflow in the connected exhaust system and thereby increased demand of fresh air to be supplied back to the laboratory after energy-intense cooling. As lab users may forget to close the sash when not in use, a number of research projects have aimed to reduce the energy consumption of fume hoods through audibly alerting lab users when unoccupied [11, 32]. Some research employs sensors to determine the sash position and occupancy through motion detection [11, 32], while others rely only on the sash position [64]. However, there are concerns about the disruptive impact of audible alarms on laboratory workflows [3, 11, 64], sensor sensitivity causing erroneous readings [32], and upfront investment costs hindering scalability [11]. Thus, such solutions are usually standalone and difficult to integrate in a comprehensive facility management approach.

To address these challenges, we integrated comprehensive monitoring and control of fume hood occupancy and sash height into a unified system of BMS, customised sensors, and BIM-GIS models. While the sash height is accessible via our Schneider Electric BMS, to monitor occupancy we had to design a smart sensor system based on a proximity sensor and an ESP32 microcontroller. Fig. 9 demonstrates our integration approach to a laboratory DT supporting the comprehensive management of fume hoods and optimising energy efficiency.

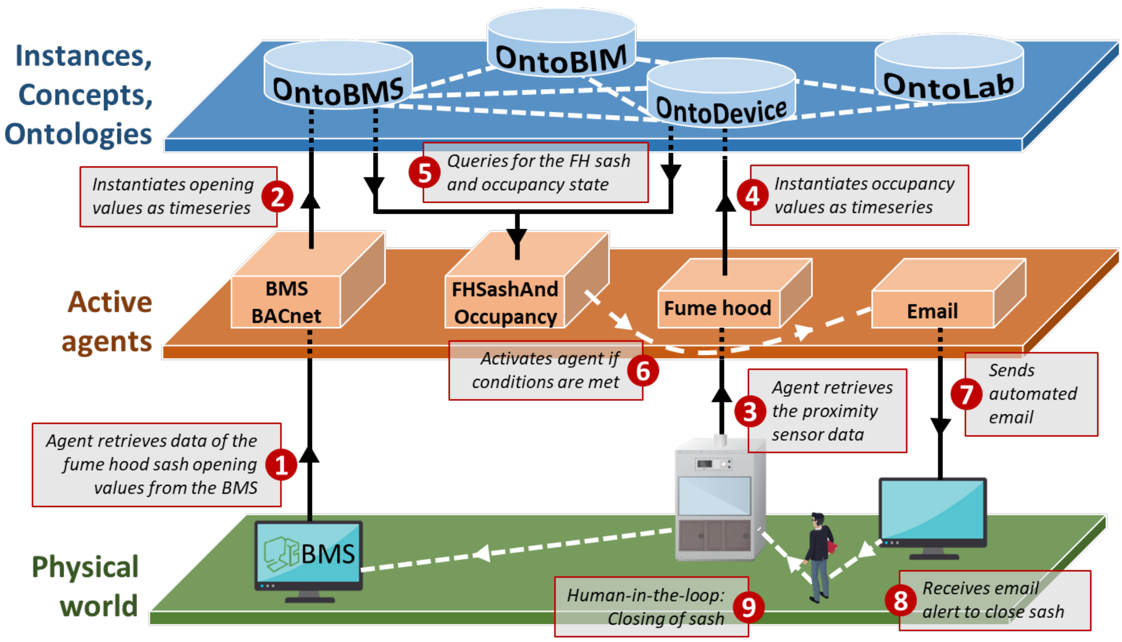


Figure 9: *Workflow of the automated fume hood monitoring and alert system.*

The first function of this DT is real-time data monitoring enabled by the deployment of smart sensors similar to the ones used for the application in section 4.1 and their integration with BMS. The second function of the DT is a fault alert system that issues email warnings if the sashes of unoccupied fume hoods are open after a certain time period.

The sash opening and occupancy data is collected from the BMS and proximity sensors respectively and processed in the dKG. Although it is technically possible to automate the closure of the sash with additional actuators, we opted for proactive manual intervention from users and managers to assess the situation before taking corrective measures (*i.e.* closing the fume hood sash). Such manual interventions coupled with data monitoring systems have reduced the instances of unoccupied open fume hoods and demonstrated significant energy savings through feedback for lab users to close unused fume hoods in other studies [3, 64]. Moreover, it prevents invasive alerts or closures through false positives if proximity sensors have inaccurately determined that the fume hood is unoccupied. Incorporating a human element is not only crucial for assessing and determining the appropriateness of a required intervention, but also as human-in-the-loop systems will be necessary to transition from semi-automated to eventually autonomous laboratories [54].

The value of this approach lies in the rich cross-domain representation of concepts and data as shown in Fig. 10. While the smart sensors technically measure distance, the dKG is able to infer occupancy via the "Derived Information Framework" as recently introduced by Bai et al. [9]. As time series data on airflow, occupancy, and sash opening are connected to a specific instance of a fume hood which also has geometric and geospatial representations, live and historic monitoring of these variables is possible by referencing location or hierarchical structure. Moreover, it enables accessibility of data from a commercial BMS and open-source low-cost IoT devices within the same system, allowing agents to carry out energy-saving tasks in an automated way.

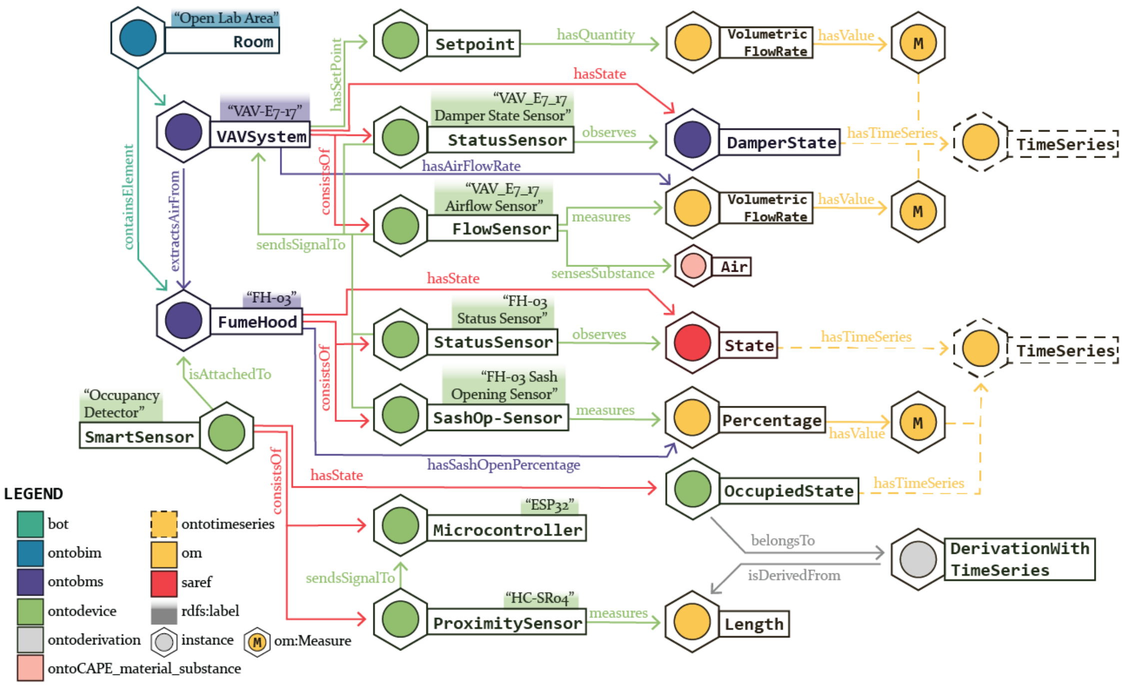


Figure 10: *ABox snippet of the fume hood monitoring and alert system.*

This can serve as a blueprint for other research facility challenges related to energy consumption, maintenance routines and regulatory compliance. For example, the process of "control banding" can help eliminate excessive airflow while ensuring safety by reviewing current information of chemicals, equipment, and procedures in use to align ventilation controls with specific hazards [42]. Using TWA methodology this can be automated through integration of cross-domain data including chemical molecular structure and hazard statements [44] which could lead to significant reductions in annual energy consumption [42]. Thus, a comprehensive interoperable laboratory digital twin can effectively monitor, optimise, and streamline buildings, assets, and operations to ensure automated, efficient and safe management of laboratory facilities.

4.3 Unified platform-agnostic interfaces

The third application involves the development of platform-agnostic user interfaces to create new modes of human-machine interactions and support fault detection, diagnostics and maintenance processes in facility management. As detailed in section 2, many currently available solutions are specific to equipment, operating system, or facility environment. By developing two freely expandable interfaces to monitor and control the use cases discussed in sections 4.1 and 4.2, we demonstrate the adaptability of a dKG approach to facility management. To meet changing demands in user-friendliness and accessibility, one interface shall be browser-based and the other one a mobile application – both accessible either within a local network or the internet, based on configuration.

Fig. 11 shows the browser-based interface, providing visualisation of static and dynamic data stored in the dKG embedded in a 3D model based on the Cesium geospatial engine as envisioned in a previous paper [49]. This comprehensive representation of the real world for monitoring real-time processes supports FMs in their role. For example, historical data can aid in identifying anomalies in airflow rates or equipment performance to detect faults. Given that some control operations can be performed from these interfaces, FMs are also able to perform their tasks on-the-go acting as mobile units and resolve multiple incidents simultaneously at greater efficiencies.

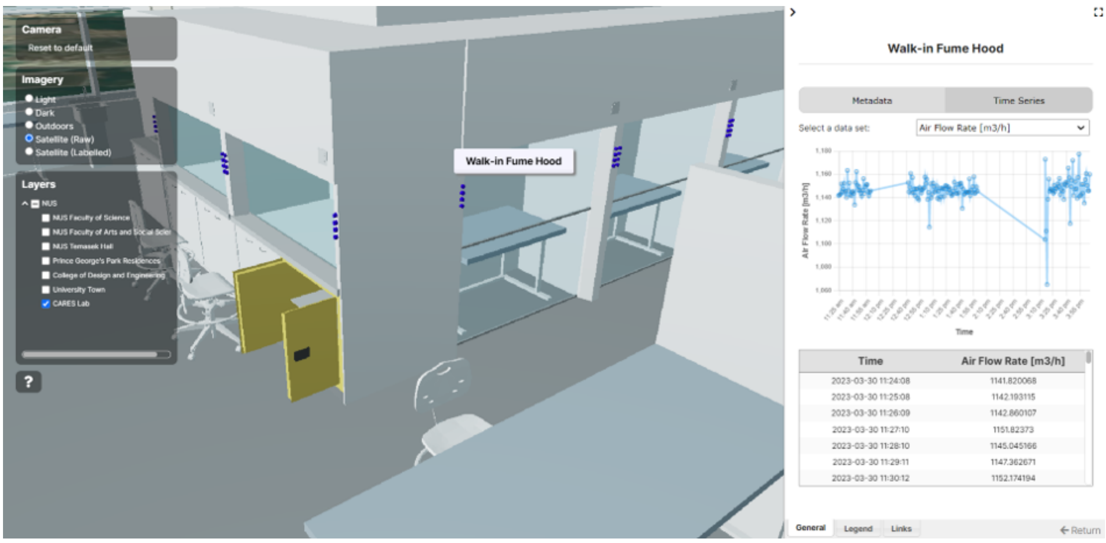
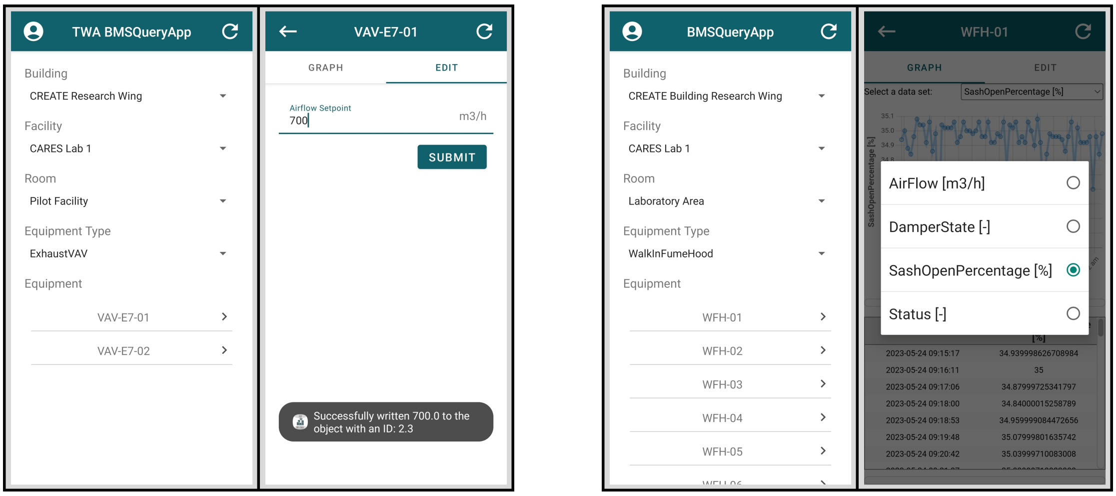


Figure 11: *Live monitoring of fume hoods via the browser-based interface.*

The integration of BIM-GIS into these interfaces is significant for three reasons. First, the availability of spatial information enables the tagging of assets to specific locations. Considering the extensive size of institutional research facilities, BIM-GIS streamlines their management processes, optimising space utilisation and localised airflow volume within all facilities while complying with occupational health and safety regulations [17]. Shared resources between buildings and faculties, such as solar power generation, can also be managed efficiently. Second, precise information on the assets' location improve the efficiency of maintenance and fault detection work. Currently, this work can be challenging as FMs have to look at multiple physical documents or labels with complex and sometimes inaccurate asset description, that may degrade over time. Last, the availability of cross-domain data provides emergency personnel with time-critical information on the building and facility layout as well as potential hazards from chemicals, blockade or other causes during crises such as a fire outbreak or gas leakage, enabling safer and more effective rescue and containment operations.

Fig. 12 depicts the mobile app interface with exemplary screenshots of hierarchical asset selection via BIM-GIS and monitoring, or adjustment of relevant BMS variables. Interactions with these interfaces are customised according to the assigned user profile, providing different user experiences and data access as detailed in appendix B.2. For example, lab users can check on the live status of specific assets as shown in Fig. 12(b). This knowledge supports them in assessing the suitability of the laboratory for conducting specific experiments. In addition, the availability of the equipment's location and occupancy information help optimise the use of resources. Meanwhile, laboratory managers are able to control airflow in the laboratory up to individual rooms as shown in Fig. 12(a).



(a) Control for canopy hood airflow set point.

(b) Monitoring of fume hood airflow.

Figure 12: *Exemplary screenshots of device selection, monitoring and control of environmental variables via mobile application.*

With the development of these broadly applicable user interfaces we aim to reduce the entry barrier for lab users and managers alike to access relevant information by providing ease of access. The open architecture of the dKG and design principles of DLF guarantee true platform independence.

5 Conclusion

The advancement of automation in facility management is pivotal for enhancing efficiency and sustainability, particularly in specialised environments like research laboratories. However, significant challenges persist in achieving interoperability and adaptability of the diverse tools employed in these settings. Our work introduces the Digital Lab Facility Manager, a novel integration at the nexus of facility and laboratory management with direct engagement in experiments and research processes. This integrated approach addresses the specific needs of resource-intensive niche facilities and underscores the potential of comprehensive automation in facility management to support knowledge discovery and perform more strategic functions.

In realising this vision, we have harnessed the power of Semantic Web technology to develop dynamic knowledge graphs as part of The World Avatar, an all-encompassing Digital Twin. In the current work, ontologies for digitising and automating the management of specialised facilities have been established and linked, including integration of 3D models with the research facility's digital twin. Through the further development of domain-spanning agents, we have tackled specific challenges such as the effective integration of IoT sensor networks, efficient airflow management in laboratories, and the accessibility of monitoring and control over critical variables. This ontological approach not only enhances the ability to manage diverse data and systems but also paves the way for significant advancements in laboratory automation and facility management.

In our work, we have demonstrated three key applications of the Digital Lab Facility Manager. Firstly, the integration of smart sensors with BMS for measuring electricity consumption of appliances and solving a simple control problem in a distributed manner. This showcases the power of dynamic knowledge graphs and the ability to deploy and integrate sensors in a straightforward and unified way. Secondly, we implemented a fume hood monitoring system to alert researchers when sashes of unoccupied fume hoods are open. This "human-in-the-loop" approach helps decrease excess airflow and energy consumption by fostering researchers' accountability. Lastly, we developed two platform-agnostic user interfaces: a browser-based UI which allows for navigating and considering the three-dimensional space when managing the lab facility, and a mobile app designed for efficient, on-the-go device navigation based on BIM hierarchy and control of BMS/IoT.

These applications showcase a digitalisation and automation framework for the management of research facilities, leveraging TWA technology. This framework is not intended as a new standard or monolithic software solution to adapt. Rather, it is a versatile system designed to overlay existing software and hardware infrastructures, utilising various protocols and data types to facilitate interoperability among diverse systems. This approach is not only applicable to research laboratories but also to other specialised facilities with varied objectives and stakeholders. Future developments aim to transform smart buildings into autonomous entities with decision-making capabilities. This necessitates the clear representation of overarching goals to inform specific tasks. Lastly, the overlapping tasks and challenges of lab managers as well as lab users call for the further exploration of a "digital lab manager" and a "digital research scientist", pushing the holistic approach even further towards an AI scientist.

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

3D Three-Dimensional

ABox Assertional Component (of an ontology)

AI Artificial Intelligence

BACnet Building Automation and Control Networks

BIM Building Information Model

BMS Building Management System

CARES Cambridge Centre for Advanced Research and Education in Singapore

CoAP Constrained Application Protocol

CREATE Campus for Research Excellence and Technological Enterprise

dKG Dynamic Knowledge Graph

DLFM Digital Lab Facility Manager

DLF Digital Lab Framework

DT Digital Twin

FM Facility Manager

GIS Geographic Information System

GUI Graphical User Interface

IFC Industry Foundation Classes

IoT Internet of Things

IRI Internationalised Resource Identifier

KG Knowledge Graph

LBMS Location-based Management System

MQTT MQ Telemetry Transport

NUS National University of Singapore

RDF Resource Description Framework

REC RealEstateCore (ontology)

SAREF Smart Applications REFerence (ontology)

TBox Terminological Component (of an ontology)

TWA The World Avatar (project)

UTown University Town (at NUS)

A Ontologies and agents

A.1 All agents

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
`AndroidStatus`	Logs the status of the mobile app shown in appendix B.2.
`BMSBACnet`	Reads and instantiates the data from BMS using the BACnet protocol as a time series in the dKG.
`BMSQuery`	Queries for equipment instances and the related zones from the dKG for the BMS Query mobile application (see Section 4.3).
`BMSUpdate`	Updates the temperature set point for the cooling fan in Fig. 8 through the BMS Query mobile application. In future, this agent will be extended to update set points found within the BMS.
`DevInst`	Instantiates sensors, microcontrollers, and devices in a knowledge graph. See details in appendix B.3.
`Email`	Sends an automated email to the intended person when specific events occur.
`ESPHome`	Queries data from the dKG and sends a status change request to device if a certain conditions are met.
`ESPHomeUpdate`	Queries and instantiates data from the ESPHome server as timeseries in the dKG.
`FeatureInfo`	Retrieves real-time static and dynamic data from the dKG.
`FumeHood`	Retrieves, derives, and instantiates the occupancy of the fume hoods as time series in the dKG.
`FHSashAndOccupancy`	Queries for the fume hoods' occupancy and sash opening values from the dKG. An email will be sent to alert the lab personnel if the fume hood is unoccupied and sash opening is greater than the threshold.
`Ifc2OntoBIM`	Instantiates IFC models and their metadata into RDF triples that are stored in the dKG.
`SmartMeter`	Handles requests to retrieve latest reading for the current time from a database storing smart meter readings and uploads the data to instantiated time series.
`ThingsBoard`	Queries and instantiates data from the ThingsBoard IoT platform as time series in the dKG.

A.2 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/>

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

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

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

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

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

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

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

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

ssn: <https://www.w3.org/TR/vocab-ssn/>

s4bldg: <https://saref.etsi.org/saref4bldg/>

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

B Frameworks and tools

B.1 Cesium-based visualisation interface

The visualisation of the combined information of the integrated geo-spatial and non-spatial data, which are semantically represented, is as shown in Fig. 5. This visualisation highlights the capability for the representation of the CARES Laboratory within UTown at the NUS.

Upon closer inspection of the BIM model of the CARES laboratory in Fig. 5 (left panel), specific work spaces within the laboratory, such as the fume hoods and the synthesis robot [8], are accurately represented within the model. Different areas and rooms of a research laboratory can be shown within the same model enriched with information on devices as well as live data from BMS and additional sensors. As all devices and sensors are now in spatial and functional relations to each other, this enables modelling and control of object interactions that are in physical vicinity to each other or functionally connected. Consequently, the building and laboratory can be visualised not only as physical structures but also with underlying functional information, providing a comprehensive representation of the laboratory.

In order to facilitate assessments from a broader perspective, encompassing the city-level view, a GIS model has been incorporated into the framework (see the right panel of Fig. 5). Notably, the GIS model encompasses an industry park located in Singapore, *i.e.* Jurong Island, which has been reported previously [22, 49]. This integration extends the system boundaries by incorporating surrounding buildings within the research institution, thereby expanding the potential use cases to include campus, city, and even national-level considerations. This integration proves relevant not only in the context of construction planning, but also for managers of institutions, such as those overseeing universities, enabling them to monitor operations and statuses across various scales and allocate resources efficiently.

B.2 Mobile application

An open-sourced prototype for an Android mobile application has been developed. The features of the application are based on the primary research conducted through interviews with Facility Managers and secondary research conducted through literature review, aiming to fulfill their needs for a mobile solution to monitor facilities under their management. This includes enhanced identification, collection, and visualisation of sensor data from different components of the facility, as well as the visualisation of the BIM/IFC data. Currently, the mobile application allows users to monitor live laboratory airflow, canopy hood air flow, smart sensors readings, and fume hood sash opening information. Also, it facilitates user account management, including defining users roles and access levels to facility controls. After a successful login, users are directed to the device selection page, as shown in Fig. 13. User information can be viewed and edited on the user profile page.

To optimise user experience and account for limited screen size of mobile devices, clean and simple user interface using dropdown menus instead of a 3D model for device selection has been implemented. This design choice aims to maintain spatial information while

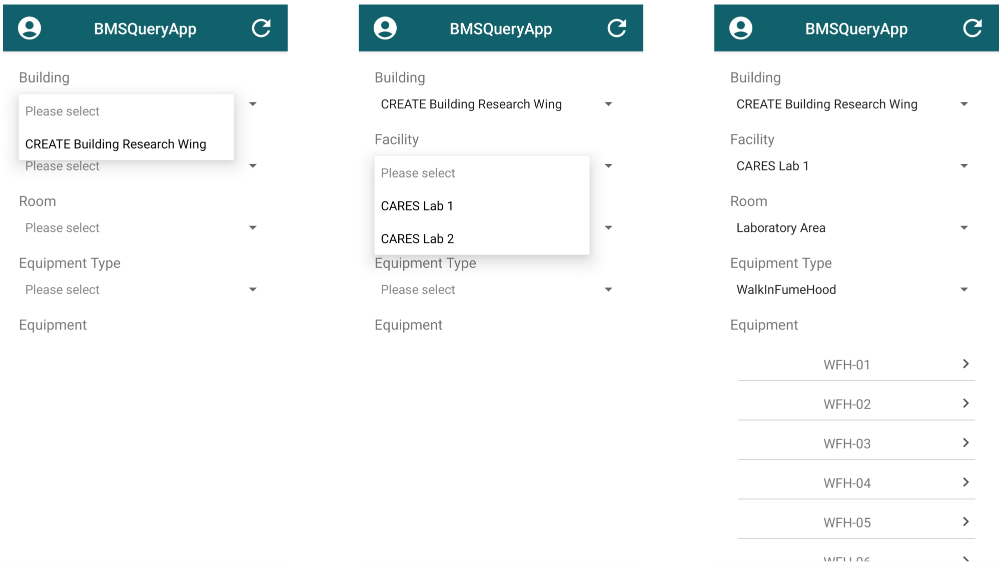


Figure 13: *Mobile Application interface for spatial locations and asset selection*

providing ease of use. It was created with the aim of minimising clutter. Users can select any device from the dropdown menu to access its data and control its state. Depending on the device type, various measurements are available for visualisation, which can be chosen by the user within the application, as shown in Fig. 14.

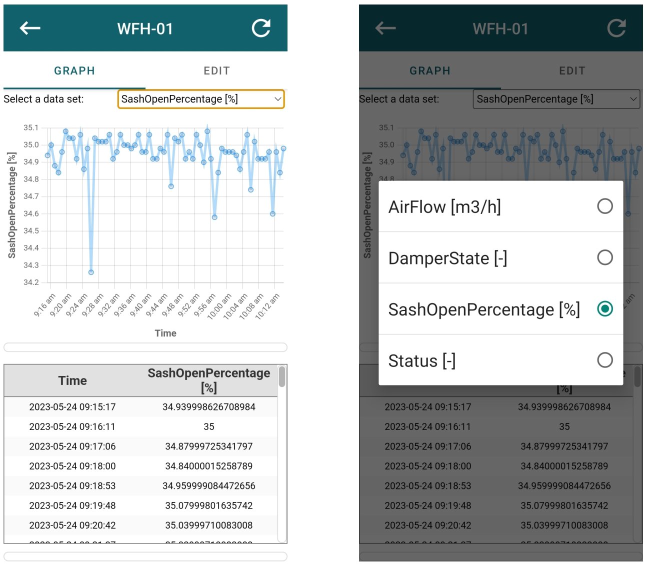


Figure 14: *Visualisation of time series measurements on the Mobile Application*

Role-based access control: A dedicated login system based on Keycloak has been developed to authenticate and authorise access to these human-machine interfaces depending on the user role and authorisation level. An example of this system for the mobile application is illustrated in Fig. 15. After successfully signing in, users can choose to view and edit their profile information accordingly.

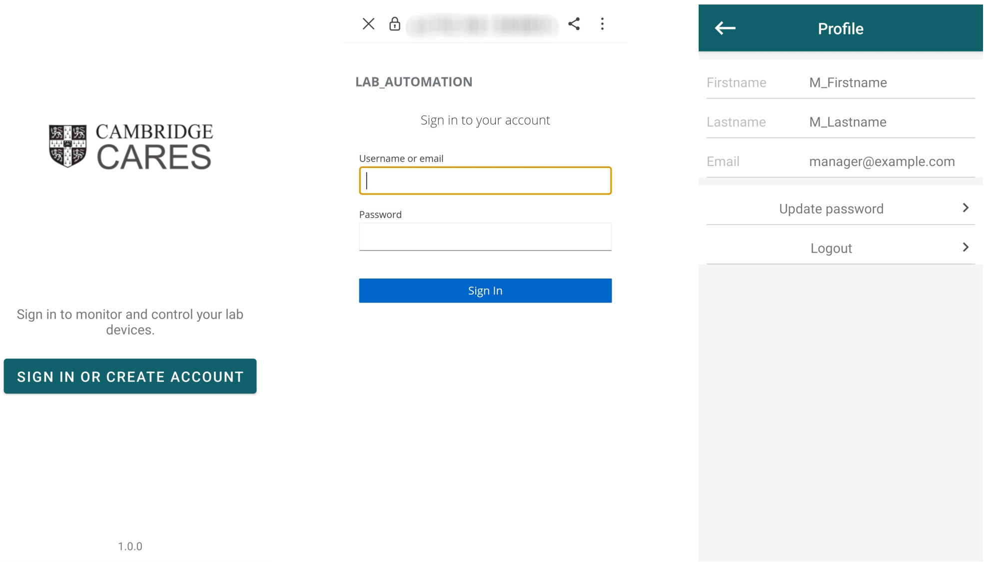


Figure 15: *The World Avatar's login and account management system for mobile applications*

Role-based restrictions are also placed on any operations in these interfaces that influence the physical state of the facility and its assets. Only authorised individuals such as managers are allowed to control the physical state. If unauthorised users attempt to perform these operations, they will be blocked and receive an in-app notification, as shown in Fig. 16.

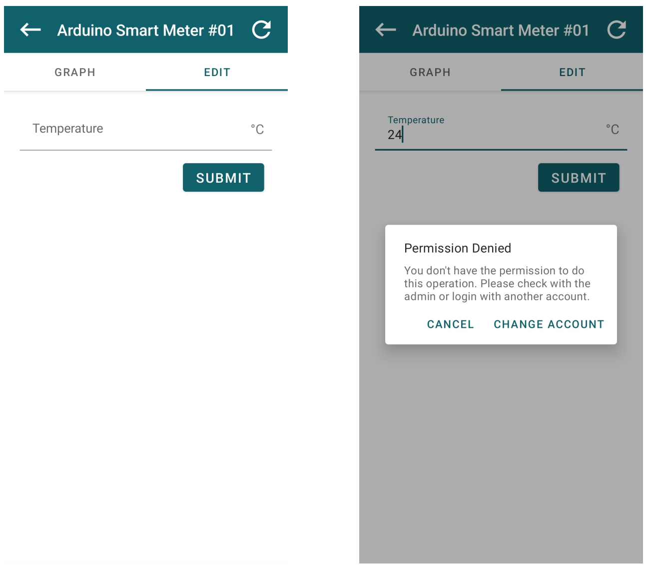


Figure 16: *Restricted permission for sensitive operations in the mobile application*

B.3 Sensor Deployment framework

In order to easily deploy sensors and other IoT devices, a Device Instantiation Framework was developed. It allows fast and consistent sensor instantiation on the knowledge graph to enable seamless deployment and development. The process is summarised in Fig. 17.

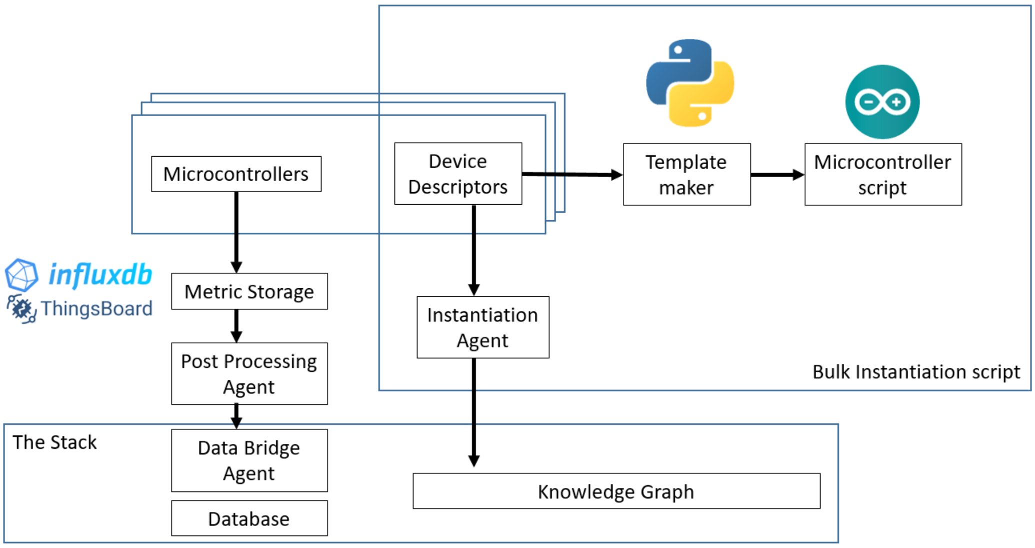


Figure 17: *Illustration of the instantiation process in the sensor deployment framework.*

Sensors will be connected to a microcontroller which will send data to an IoT database - *e.g.*, ThingsBoard. The post-processing agent will pull the data from the database and instantiate a timeseries of the data on the knowledge graph. The post-processing agent will also be responsible for instantiating agent and derivation instances when needed as well as creating the timeseries instances. Hence, for different projects, different post-processing agents may be needed. An example of this post-processing agent is the fume hood agent described in appendix A.1. The device instantiaion agent will be responsible for instantiating the sensors and devices in the knowledge graph using a device descrpitor file written by the user. This description must be provided as JSON file the structure of which is shown in Fig. 18.

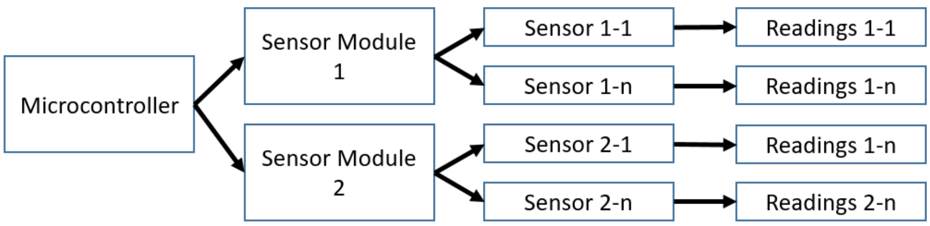


Figure 18: *Structure of the JSON input file for sensor device instantiation.*

The file itself is composed of IDs which will be used to find the IRI in the IRI Mapper. Each device will require one descriptor. A descriptor can be composed of a single microcontroller and several sensor modules. Each sensor modules can be composed of several sensors. Each sensor will have one type of sensor reading. The relevant source code and documentation is available under https://github.com/cambridge-cares/TheWorldAvatar/tree/main/Agents/DevInstAgent.

References

[1] M. Adams, V. Burrows, S. Richardson, J. Drinkwater, C. Gamboa, C. Collin, X. Le Den, L. Ostenfeld Riemann, S. Porteron, and A. Qvist Secher. Bringing embodied carbon upfront: Coordinated action for the building and construction sector to tackle embodied carbon. Technical report, World Green Building Council, London, 2019. URL `https://worldgbc.org/advancing-net-zero/embodied-carbon/`. Last accessed December 11, 2023.

[2] M. Aishah Kamarazaly, J. Mbachu, and R. Phipps. Challenges faced by facilities managers in the Australasian universities. *Journal of Facilities Management*, 11(2): 136–151, 2013. doi:10.1108/14725961311319755.

[3] K. Aldred Cheek and N. M. Wells. Changing Behavior Through Design: A Lab Fume Hood Closure Experiment. *Frontiers in Built Environment*, 5(January):1–13, 2020. doi:10.3389/fbuil.2019.00146.

[4] K. Amin, G. Mills, and D. Wilson. Key functions in BIM-based AR platforms. *Automation in Construction*, 150(November 2022):104816, 2023. doi:10.1016/j.autcon.2023.104816.

[5] G. Antoniou, E. Franconi, and F. van Harmelen. Introduction to semantic web ontology languages. In N. Eisinger and J. Małuszyński, editors, *Reasoning Web*, volume 3564 of *Lecture Notes in Computer Science*, pages 1–21. Springer Berlin Heidelberg, 2005. ISBN 978-3-540-31675-6. doi:10.1007/11526988_1.

[6] V. A. Arowoiya, R. C. Moehler, and Y. Fang. Digital twin technology for thermal comfort and energy efficiency in buildings: A state-of-the-art and future directions. *Energy and Built Environment*, 2023. doi:10.1016/j.enbenv.2023.05.004. In press.

[7] C. J. Backlund, C. E. Hjorth, R. D. Armijo, R. M. Jones, C. A. Quinn-Vawter, and T. C. Smith. The Benefits and Challenges of Implementing Smart Labs in a Multipurpose Research Laboratory Building: Undertaking a Pilot Project at Sandia National Laboratories. *ACS Chemical Health & Safety*, 29(4):344–349, 2022. doi:10.1021/acs.chas.2c00011.

[8] 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.

[9] J. Bai, K. F. Lee, M. Hofmeister, S. Mosbach, J. Akroyd, and M. Kraft. A derived information framework for a dynamic knowledge graph and its application to smart cities. *Future Generation Computer Systems*, 152:112–126, 2024. doi:https://doi.org/10.1016/j.future.2023.10.008.

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

[11] L. L. Becerra, J. A. Ferrua, M. J. Drake, D. Kumar, A. S. Anders, E. N. Wang, and D. J. Preston. Active fume hood sash height monitoring with audible feedback. *Energy Reports*, 4:645–652, 2018. doi:10.1016/j.egyr.2018.09.010.

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

[13] 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.

[14] S. Building and C. Authority. Singapore green building masterplan 4th edition, 2022. URL `https://www1.bca.gov.sg/buildsg/sustainability/green-building-masterplans`. Accessed June 2023.

[15] S. Campbell. Green Cities, Growing Cities, Just Cities?: Urban planning and the contradictions of sustainable development. *Journal of the American Planning Association*, 62(3):296–312, 1996. doi:10.1080/01944369608975696.

[16] L. Chamari, E. Petrova, and P. Pauwels. A web-based approach to BMS, BIM and IoT integration: a case study. In *CLIMA 2022 conference*, May 2022. doi:10.34641/clima.2022.228.

[17] C. S. Committee. Code of practice for fire safety for laboratories using chemicals. Standard SS 641:2019, Singapore Standards Council, Singapore, 2019. URL `https://www.singaporestandardseshop.sg/Product/SSPdtDetail/e35e9a81-eb82-4c84-a94a-a610a9b40eee`.

[18] A. A. Da Conceic'ao, L. P. Ambrosio, T. R. Leme, A. C. Rosa, F. F. Ramborger, G. P. Aquino, and E. C. Vilas Boas. Internet of Things Environment Automation: A Smart Lab Practical Approach. In *2022 2nd International Conference on Information Technology and Education (ICIT&E)*, pages 322–327. IEEE, 2022. ISBN 978-1-6654-9433-5. doi:10.1109/ICITE54466.2022.9759899.

[19] L. Daniele, F. den Hartog, and J. Roes. Created in Close Interaction with the Industry: The Smart Appliances REFerence (SAREF) Ontology. In R. Cuel and R. Young, editors, *Formal Ontologies Meet Industry*, pages 100–112. Springer International Publishing, 2015. ISBN 978-3-319-21545-7. doi:10.1007/978-3-319-21545-7_9.

[20] L. Daniele, R. Garcia-Castro, M. Lefrançois, and M. Poveda-Villalon. SAREF: the Smart Applications REFerence ontology, 2020. URL `https://saref.etsi.org/core/`. Accessed June 2023.

[21] S. Durdyev, M. Ashour, S. Connelly, and A. Mahdiyar. Barriers to the implementation of Building Information Modelling (BIM) for facility management. *Journal of Building Engineering*, 46:103736, apr 2022. doi:10.1016/j.jobe.2021.103736.

[22] 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.

[23] S. Gilani, C. Quinn, and J. J. McArthur. A review of ontologies within the domain of smart and ongoing commissioning. *Building and Environment*, 182(January): 107099, 2020. doi:10.1016/j.buildenv.2020.107099.

[24] K. Hammar, E. O. Wallin, P. Karlberg, and D. Hälleberg. The RealEstateCore Ontology. In *The Semantic Web – ISWC 2019*, pages 130–145. Springer International Publishing, 2019. ISBN 978-3-030-30796-7. doi:10.1007/978-3-030-30796-7_9.

[25] R. K. Haugen. Laboratory Hood Energy Savings: The Low-Hanging Fruit. *ACS Chemical Health & Safety*, 27(2):125–128, 2020. doi:10.1021/acs.chas.9b00013.

[26] A. Hogan, E. Blomqvist, M. Cochez, C. D'Amato, G. D. Melo, C. Gutierrez, S. Kirrane, J. E. L. Gayo, R. Navigli, S. Neumaier, A. C. N. Ngomo, A. Polleres, S. M. Rashid, A. Rula, L. Schmelzeisen, J. Sequeda, S. Staab, and A. Zimmermann. Knowledge graphs. *ACM Computing Surveys*, 54(4), jul 2021. doi:10.1145/3447772.

[27] M. Jin, F. Memarzadeh, K. Lee, and Q. Chen. Experimental study of ventilation performance in laboratories with chemical spills. *Building and Environment*, 57: 327–335, 2012. doi:10.1016/j.buildenv.2012.04.022.

[28] M. D. Kaplowitz, L. Thorp, K. Coleman, and F. Kwame Yeboah. Energy conservation attitudes, knowledge, and behaviors in science laboratories. *Energy Policy*, 50: 581–591, 2012. doi:10.1016/j.enpol.2012.07.060.

[29] M. A. Khan, I. A. Sajjad, M. Tahir, and A. Haseeb. IOT Application for Energy Management in Smart Homes. *Engineering Proceedings*, 20(1), 2022. doi:10.3390/engproc2022020043.

[30] H. Kim, H. Kang, H. Choi, D. Jung, and T. Hong. Human-building interaction for indoor environmental control: Evolution of technology and future prospects. *Automation in Construction*, 152(March):104938, 2023. doi:10.1016/j.autcon.2023.104938.

[31] 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.

[32] J. Kongoletos, E. Munden, and J. Ballew. Motion and Sash Height (MASH) alarms for efficient fume hood use. *Scientific Reports*, 11:21412, 2021. doi:10.1038/s41598-021-00772-y.

[33] J. Y. Lee, I. O. Irisboev, and Y. S. Ryu. Literature review on digitalization in facilities management and facilities management performance measurement: Contribution of industry 4.0 in the global era. *Sustainability (Switzerland)*, 13(23), 2021. doi:10.3390/su132313432.

[34] 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. ISBN 978-3-030-86215-2. doi:10.1007/978-3-030-86215-2_4.

[35] Y. C. Lin and Y. C. Su. Developing mobile- and BIM-based integrated visual facility maintenance management system. *The Scientific World Journal*, 2013, 2013. doi:10.1155/2013/124249.

[36] W. Liu, D. Liu, and N. Gao. CFD study on gaseous pollutant transmission characteristics under different ventilation strategies in a typical chemical laboratory. *Building and Environment*, 126(August):238–251, 2017. doi:10.1016/j.buildenv.2017.09.033.

[37] K. Lynn and A. Fowler. Trends in Facilities Management in the UK. Technical report, CBRE, Inc., London, 2023. URL `https://www.cbre.co.uk/insights/articles/trends-in-facilities-management-in-the-uk`. Last accessed December 11, 2023.

[38] F. S. Maikore, G. Selenge, A. Olayinka, P. Abbott, and L. Soldatova. An ontology for clinical laboratory standard operating procedures. In *CEUR Workshop Proceedings*, volume 2050, 2017.

[39] R. Maimone, M. Guerini, M. Dragoni, T. Bailoni, and C. Eccher. PerKApp: A general purpose persuasion architecture for healthy lifestyles. *Journal of Biomedical Informatics*, 82(April):70–87, 2018. doi:10.1016/j.jbi.2018.04.010.

[40] E. Marsh, S. Allen, and L. Hattam. Tackling uncertainty in life cycle assessments for the built environment: A review. *Building and Environment*, 231(December 2022): 109941, 2023. doi:10.1016/j.buildenv.2022.109941.

[41] H. G. Martin, T. Radivojevic, J. Zucker, K. Bouchard, J. Sustarich, S. Peisert, D. Arnold, N. Hillson, G. Babnigg, J. M. Marti, C. J. Mungall, G. T. Beckham, L. Waldburger, J. Carothers, S. S. Sundaram, D. Agarwal, B. A. Simmons, T. Backman, D. Banerjee, D. Tanjore, L. Ramakrishnan, and A. Singh. Perspectives for self-driving labs in synthetic biology. *Current Opinion in Biotechnology*, 79:102881, 2023. doi:10.1016/j.copbio.2022.102881.

[42] J. F. McCarthy, M. A. Fragala, and B. J. Baker. Analyzing the Risk: Balancing Safety and Efficiency in Laboratory Ventilation. *ACS Chemical Health & Safety*, 29 (5):434–440, 2022. doi:10.1021/acs.chas.1c00095.

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

[44] L. Pascazio, S. D. 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*, 2023. doi:10.1021/acs.jcim.3c00820.

[45] M. A. Patil, K. Parane, S. Poojara, and A. Patil. Internet-of-things and mobile application based hybrid model for controlling energy system. *International Journal of Information Technology (Singapore)*, 13(5):2129–2138, 2021. doi:10.1007/s41870-021-00667-1.

[46] T. Petkova. Smart Buildings Are Built of Smart Data: Knowledge Graphs for Building Automation Systems, 2022. URL https://www.ontotext.com/blog/knowledge-graphs-for-building-automation-systems/. Last accessed December 11, 2023.

[47] M. Poveda-Villalón and R. Garcia-Castro. SAREF extension for building, 2020. URL https://saref.etsi.org/saref4bldg/. Accessed June 2023.

[48] G. Qiang, S. Tang, J. Hao, L. Di Sarno, G. Wu, and S. Ren. Building automation systems for energy and comfort management in green buildings: A critical review and future directions. *Renewable and Sustainable Energy Reviews*, 179(January): 113301, 2023. doi:10.1016/j.rser.2023.113301.

[49] H. Y. Quek, S. Rihm, M. Hofmeister, J. Yan, S. Mosbach, W. Ang, D. N. Tran, and M. Kraft. BIM-GIS Integration: Knowledge graphs in a world of silos. Submitted for publication; preprint available online, 2023. URL https://como.ceb.cam.ac.uk/preprints/311/.

[50] H. Y. Quek, F. Sielker, J. Akroyd, A. N. Bhave, A. Von Richthofen, P. Herthogs, C. V. D. L. Yamu, L. Wan, T. Nochta, G. Burgess, M. Q. Lim, S. Mosbach, and M. Kraft. The conundrum in smart city governance: Interoperability and compatibility in an ever-growing ecosystem of digital twins. *Data and Policy*, 5(287), 2023. doi:10.1017/dap.2023.1.

[51] J. Ratajczak, M. Riedl, and D. T. Matt. BIM-based and AR application combined with location-based management system for the improvement of the construction performance. *Buildings*, 9(5):118, 2019. doi:10.3390/buildings9050118.

[52] G. Reid, H. Pain, A. Horan, J. Broderick, C. Dyer-Smith, I. Meazzini, J. Long, N. Sims, J. Anson, L. Marle, W. Niu, P. Stackhouse, H. Armes, S. De Pellegars, E. Eley, R. Holliday, I. Monk, T. Underwood, H. White, E. Ratcliffe, C. Southgate, N. Gibson, and C. Morley. Sustainable Laboratories. Technical report, Royal Society of Chemistry, Cambridge, 2022. URL https://www.rsc.org/globalassets/22-new-perspectives/sustainability/sustainable-labs/sustainable-laboratories-report.pdf. Last accessed December 11, 2023.

[53] S. B. Review. Singapore to enact energy audits on 'energy-intensive' buildings, 2023. URL https://sbr.com.sg/building-engineering/news/singapore-enact-energy-audits-energy-intensive-buildings. Accessed June 2023.

[54] S. D. Rihm, J. Bai, A. Kondinski, S. Mosbach, J. Akroyd, and M. Kraft. The Digital Lab Framework as part of The World Avatar. Preprint available online, 2023. URL https://como.ceb.cam.ac.uk/preprints/314/.

[55] H. Rijgersberg, M. Van Assem, and J. Top. Ontology of units of Measure (OM) 2.0. https://github.com/HajoRijgersberg/OM, 2013.

[56] 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.

[57] D. Satterthwaite. Sustainable Cities or Cities that Contribute to Sustainable Development? *Urban Studies*, 34(10):1667–1691, 1997. doi:10.1080/0042098975394.

[58] D. Shkundalov and T. Vilutienė. Bibliometric analysis of building information modeling, geographic information systems and web environment integration. *Automation in Construction*, 128(March), 2021. doi:10.1016/j.autcon.2021.103757.

[59] Singapore Sustainable Laboratories Group. Sustainable Laboratories, 2020. URL `https://www.seas.org.sg/sustainablelaboratories`. Last accessed December 11, 2023.

[60] C. C. Sobin. A Survey on Architecture, Protocols and Challenges in IoT. *Wireless Personal Communications*, 112(3):1383–1429, 2020. doi:10.1007/s11277-020-07108-5.

[61] M. Valinejadshoubi, O. Moselhi, A. Bagchi, and A. Salem. Development of an IoT and BIM-based automated alert system for thermal comfort monitoring in buildings. *Sustainable Cities and Society*, 66(November 2020):102602, 2021. doi:10.1016/j.scs.2020.102602.

[62] A. Verma, S. Prakash, V. Srivastava, A. Kumar, and S. C. Mukhopadhyay. Sensing, Controlling, and IoT Infrastructure in Smart Building: A Review. *IEEE Sensors Journal*, 19(20):9036–9046, 2019. doi:10.1109/JSEN.2019.2922409.

[63] W3C. Semantic web, 2015. URL `https://www.w3.org/standards/semanticweb/`. Accessed June 2023.

[64] D. Wesolowski, E. Olivetti, A. Graham, S. Lanou, P. Cooper, J. Doughty, R. Wilk, and L. Glicksman. The use of feedback in lab energy conservation: Fume hoods at MIT. *International Journal of Sustainability in Higher Education*, 11(3):217–235, 2010. doi:10.1108/14676371011058523.

[65] X. Zhou, A. Eibeck, M. Q. Lim, N. B. Krdzavac, and M. Kraft. An agent composition framework for the J-Park Simulator - A knowledge graph for the process industry. *Computers and Chemical Engineering*, 130:106577, 2019. doi:10.1016/j.compchemeng.2019.106577.

[66] X. Zhou, D. Nurkowski, S. Mosbach, J. Akroyd, and M. Kraft. Question Answering System for Chemistry. *Journal of Chemical Information and Modeling*, 61(8):3868–3880, 2021. doi:10.1021/acs.jcim.1c00275.