Data Management for Building Information Modelling in a Real-Time Adaptive City Platform

Justas Brazauskas  
jb2328@cam.ac.uk  
University of Cambridge  
Cambridge, United Kingdom

Rohit Verma  
rv355@cam.ac.uk  
University of Cambridge  
Cambridge, United Kingdom

Vadim Safronov  
vs451@cam.ac.uk  
University of Cambridge  
Cambridge, United Kingdom

Matthew Danish  
mrd45@cam.ac.uk  
University of Cambridge  
Cambridge, United Kingdom

Jorge Merino  
jm2210@cam.ac.uk  
University of Cambridge  
Cambridge, United Kingdom

Xiang Xie  
xx809@cam.ac.uk  
University of Cambridge  
Cambridge, United Kingdom

Ian Lewis  
jll20@cam.ac.uk  
University of Cambridge  
Cambridge, United Kingdom

Richard Mortier  
rmm1002@cam.ac.uk  
University of Cambridge  
Cambridge, United Kingdom

ABSTRACT

Legacy Building Information Modelling (BIM) systems are not designed to process the high-volume, high-velocity data emitted by in-building Internet-of-Things (IoT) sensors. Historical lack of consideration for the real-time nature of such data means that outputs from such BIM systems typically lack the timeliness necessary for enacting decisions as a result of patterns emerging in the sensor data. Similarly, as sensors are increasingly deployed in buildings, antiquated Building Management Systems (BMSs) struggle to maintain functionality as interoperability challenges increase. In combination these motivate the need to fill an important gap in smart buildings research, to enable faster adoption of these technologies, by combining BIM, BMS and sensor data. This paper describes the data architecture of the Adaptive City Platform, designed to address these combined requirements by enabling integrated BIM and real-time sensor data analysis across both time and space.

CCS CONCEPTS

- Computer systems organization → Real-time system architecture;  
- Information systems → Information integration.

KEYWORDS

smart buildings, IoT, BMS, BIM, real-time, API

1 INTRODUCTION

Smart infrastructure projects, e.g., building automation, are increasingly dependent on BIM and BMS for metadata and data collection, analysis and activation components. BMS’ often have the ability to monitor lighting, heating, ventilation and air conditioning (HVAC) as well as electricity consumption. Typically proprietary software, most BMS vendors provide closed system products with industrial interfaces for activation, control, and basic data visualisation without the additional contextual data from BIM. Simultaneously, considerable effort is being expended on the deployment of IoT devices to increase sensor density, but unfortunately the industry remains highly fragmented [10, 30].

Additionally, legacy BMS’ and much current research focuses on in-building sensor data collection, storage and presentation platforms, rarely emphasising the challenges and benefits of being able to analyse and respond to data in real-time [16, 18, 45]. BMS’ have historically dealt with low-volume low-velocity data and metadata, so the adoption of IoT devices poses substantial network and system challenges in dealing with real-time data analysis, event recognition, prediction and action planning [40].

In this paper we focus on the real-time aspects of spatio-temporal data available from IoT sensing. We define a real-time platform as an asynchronous system capable of processing high-volume, high-heterogeneity data with minimal latency to collect, analyse, predict and adapt to changes in a timely manner.

A real-time data architecture is only part of the puzzle though: despite the increasing deployment of IoT devices, there are still no canonical means to join BIM and deployed sensors in a single unified system. While numerous attempts exist in the form of creating ontologies (e.g., BRICK) [8, 9] to unify static metadata management for use by building automation systems, industry recognition for metadata standards is limited [10, 11, 30]. Also, as a result augmenting BIM with IoT devices, building and facility management software must be adapted [28]. Highly siloed BMS software must become able to handle an increased amount of contextual building data in a timely manner to comply with the use of edge computing to for accident and emergency management [33] and smart home initiatives resulting in the creation of safer and more resilient smart spaces [23, 46]. New approaches that combine BIM, BMS and sensor data are thus needed.

To meet these important challenges, we propose the Adaptive City Platform (ACP), a system for collecting, processing and visualising building information and sensor data in real-time. Our contributions are: (i) the design of the ACP, a real-time software architecture combining BIM, BMS and IoT data into a unified low-latency, high-throughput building monitoring system; (ii) description of our prototype implementation of the ACP, deployed in our department building; and (iii) initial evaluation of our platform, in
Today, building systems monitoring and actuation is managed by highly encapsulated, siloed BMS systems [10, 30]. Usually employed to monitor building-scale mechanical and electrical systems (e.g., HVAC, electricity consumption), the majority of BMS’ are unfit for use with IoT as they typically lack the contextual information provided by the BIM software as well as the capacity to access and process low-latency, high-bandwidth IoT data [31, 40]. Though there are numerous benefits for further integration of these technologies, BIM-IoT fusion remains in its infancy.

Successful BIM and IoT device use for energy, health, safety and building management, and construction monitoring have been demonstrated [37, 47]. Furthermore, combined with BMS, BIM-IoT fusion allows for a new way to manage, control and ultimately automate building systems to achieve lower electricity consumption and reduce the CO₂ footprint [32]. While much smart infrastructure research focuses on overcoming the standardization problems and crafting ontologies for static reference data, BIM-IoT incorporation is becoming an important topic in its own right [45].

While real-time analysis of data processing is often overlooked, several papers have proposed interlinking and contextualising sensor data using BIM tools [45]. Nevertheless, most attempts remain relatively small scale and often lack good (non-blocking, asynchronous) support for real-time data. A common pitfall with data storage-centric models that rely on repeated querying is that sensor data is collected and archived while sensor metadata is made an extension of the BIM model. This results in smart building platforms that are dependent on static BIM software with no real capacity to provide a low-latency data flow, as the examples below show.

Penna et al [38] describes a system where sensor readings are linked with a BIM database using Revit and Dynamo (a visual scripting language for Revit). The proposed system integrated LoRaWAN-powered environmental sensor data such as temperature, CO₂, humidity with people’s presence, storing it in a SQL database. The data was then loaded to Revit using a Dynamo script to visualise readings. However, the system lacked a real time component, resulting in transmission latencies of up to 15 minutes.

Rasmussen et al [26, 42] presents an integrated web system using their proposed Linked Building Data ontology with sensor and actuator readings. Although their demo featured a well described web interface and detailed API reference, it lacked a real-time component as a result of their storage-centric model, as well as also relying on Revit. Similarly, Chevalier et al [16] propose a reference architecture for Digital Twin buildings, using BIM and IoT technologies. Focused on ontology definitions, the presented system architecture and examples also feature a database-centric model reliant on queries.

Dave et al [18] proposed a platform that integrates IoT data and real-time web visualisation environment, Otaniemi3D, that showed the University of Aalto’s campus on several scales. They describe the design criteria, system architecture, and workflow, along with examples of how their system would operate in a real-world environment. However, this also assumed a “collect, store, query” model for sensor data which fundamentally limits its real-time capability. Additionally, the static nature of their XML-based data storage makes it unsuitable for spatio-temporal data.
More widely, Tang et al [45] provides an overview of research covering both BIM and IoT, and highlighting several major challenges in research regarding smart buildings. Key overlooked aspects identified include practical implementability – many described systems were conceptual rather than usable. Moreover, prototypes were often tested under lab conditions, limiting the assessment of the ability of those systems to withstand the challenges of real-world deployments. Finally, they assert that most research they survey fails to achieve real-time information queries, and lacks actuation elements.

3 REAL-TIME REQUIREMENTS

As deployed IoT sensors become the standard for applications such as periodic monitoring of temperature and air quality in buildings, we propose that the next iteration of connected environments in-world deployments. Finally, they assert that most research they survey fails to achieve real-time information queries, and lacks actuation elements.

3.1 Key Concepts and Data Types

A real-time focused system architecture is important because connected environments software need event-driven data to provide responsiveness without introducing arbitrary programmer-determined delays. Instead of being solely reliant on data querying using APIs, our proposed system architecture also enables instantaneous data flow, allowing for rapid event detection. By implementing this architecture we can overcome some of the hard limits related to IoT sensor deployment and BMS, such as the dependence on periodic data querying and static BIM systems.

In this section we outline key concepts related to an effective real-time data management platform for buildings. We distinguish different data types flowing through the platform, whilst simultaneously declaring that all data be treated as spatio-temporal data. We consider building data as spatial data, while all time-series sensor readings are temporal data, split into periodic and event-based data.

Many BIM systems assume spatial data is static by ignoring building metadata changes once construction is complete. We contend that this is false: buildings change over their life-cycle and so does their spatial data, either in the form of changes to floor plans or movement of building contents. A proper real-time platform integrating sensors, BMS and BIM must have the capacity to update spatial data properties as they change. This is particularly important in cases where building equipment often changes location, e.g., hospital beds, as coordinate data becomes (perhaps low-rate) real-time data subject to stream processing.

3.2 Status Reading vs Events

One of the fundamental challenges in real-time software platforms for sensor data is the lack of distinction between event data and periodic readings. This issue arises with the off-the-shelf sensor software which is often based on periodic status reporting rather than publish-subscribe and stream processing architectures.

While periodic readings are useful for sensor status reporting or sensing data that changes very gradually (e.g., air humidity), a platform solely based on periodic queries cannot be defined as real-time due to arbitrary blackout time-periods where no data is reported. For this reason, we believe a proper real-time architecture must focus on the concepts of timeliness and events, where messages are sent instantaneously after readings indicate a significant change in state.

We define timeliness as a characteristic of an event, in which it is appropriate to act on changed readings, e.g. the duration from the beginning of the end of an event. In other words, it is a timeframe during which our event is at its most relevant. While for some events, e.g. interrupt-triggered sensor readings, latency and timeliness are almost identical time-wise, however, for more sophisticated events that follow a sequence of multiple sensor reading, minimizing latency becomes difficult due to uncertainty. As a result, we propose that every incoming event comes in with a probability value, determining the how likely the event-recognition is to be true.

Furthermore, the focus on events rather than raw data reporting allows for a flexible and easily scalable approach to data collection by having the ability to group raw sensor events into more informative, less frequent events. As the number of sensors increases,
events can occur due to a single sensor being triggered, or as a result of a sequence of triggers passed down from multiple sensors. We distinguish two such categories of events: Simple and Derived.

**Event types.** Depicted in Figure 3, we derive events based on their readings, e.g., when sensor readings reach a specific threshold or an interrupt occurs, for Sensors A and D. After a trigger occurs, the sensor sends an event to the real-time platform. We define such incoming data as Simple Events. Sensors B and C are programmed to follow event detection based on a sequence of readings, e.g., sensor B’s activation being followed by sensor C’s. In such cases, events are detected as a combination of sensor readings and we consider these Derived Events.

Derived events are only possible when the timeliness of individual sensor readings overlap (Figure 2). Our platform permits us to look for such sequences of individual asynchronous events that result in more complex events being tracked, thus creating a richer building information environment.

Our intent is not to collect and solely rely on periodic readings and then send the data back to the platform, but rather to use stream processing to detect derived events as they happen. Real-time stream processing allows for data to be received and analysed instantly; therefore, we are able not only to recognize events from a single sensor but also from a combination of multiple incoming data sources. This event-focused approach to real-time system architecture allows for non-blocking data processing with minimal additional latency and minimal network load.

### 3.3 Programmatic Access to BMS

In addition to event detection, another critical aspect of smart building infrastructure is the BMS-in-the-loop integration for building technical systems monitoring and actuation. While IoT sensor deployment is instrumental in smart buildings, programmatic access to BMS software for real-time usage suffers from the same lack of ontological standards that hinders deeper BIM-BMS integration.

The combination of a real-time platform that has access to both BMS and in-building sensors then allows for timely decision making by collecting, analysing, predicting and adapting to the sensed environment. As it stands, BMS software is limited to low granularity data (e.g., electricity consumption on a per-floor basis) and closed control loop mechanisms that would benefit from richer IoT data sources. While there are products on the market that claim to augment BMS capabilities by implementing MQTT drivers and deeper IoT integration [44], the benefits of such platforms remain to be seen. In addition to well-known building automation benefits (e.g., decreased energy usage or HVAC optimisations), deeper BMS integration with the BIM software and IoT could be particularly useful in (i) detecting accidents and emergencies, (ii) microgrid energy management, (iii) domain-specific applications.

(i) **Accident and emergency management.** The use of a real time platform for accident and emergency response can form the basis for effective immediate risk management. Broader IoT integration within BMS would play a critical role in recognising such anomalies in two ways. First, deployed sensors would be capable of detecting changes and anomalies in the behaviour of critical assets in real-time, avoiding further damage e.g., by monitoring pump vibrations and sensing any deviation from what is observed to be normal.

Furthermore, the use of real time event recognition could be crucial in recognising gas or oil leaks in industrial settings where such accidents could be handled before becoming serious [33]. Second, it would permit a general system-wide tracking across numerous subsystems deployed in buildings e.g., HVAC. A real-time BMS with access to highly localised sensor data would be especially useful in large scale infrastructure projects e.g., airports where numerous subsystems create a complex whole. Therefore, the resulting outcome becomes not only the provision of time-critical response systems but also a generally lowered complexity that makes large infrastructure projects easier to manage.

(ii) **Microgrid energy management.** Event-based occupancy tracking from a multitude of sensors deployed in the built environment allows for more efficient energy usage. Having the ability to combine sensor readings coming from environmental factors, like CO₂ concentration [22] and presence-detection indoors [52] presents numerous possibilities for heating optimisation and electricity consumption on a finer granularity a per room basis [39, 51].

(iii) **Domain-specific use cases.** Real time event detection can benefit built environments where building inhabitants are at an increased risk, e.g. in hospitals, for patient tracking and fall detection in wards [21, 23]. Similarly, in a smart home context, real time accident recognition could help the elderly population receive timely help if unable to live independently [3, 46]. Lastly, real time-based people density estimation to avoid overcrowding and violent behaviour detection in crowds can be useful for managing people flow in busy areas like stations or airports and stopping crowd crushes [24, 35].

### 4 DESIGN

The core of our contribution is our real-time system, the Adaptive City Platform (ACP), as well as the accompanying example application showing our data pipeline in a real-world scenario. We connected our platform to sensors deployed in a non-BIM native building. The platform uses an ontology that we created to illustrate how real-time mechanisms allow for more versatility in BIM-IoT fusion.

Figure 4 depicts the overall system. The ACP accepts real-time data flow from in-building sensors over the MQTT protocol. The back-end consists of a non-blocking real-time platform with a set of APIs accessing the file system and a PostgreSQL database. Any incoming data is passed through our real-time platform where it is restructured, stored, and ultimately propagated to our front-end
visualisation. The collected data is saved in a PostgreSQL database and the file system along with the BIM metadata.

We next describe our real-time platform architecture, APIs, and the front-end interface.

4.1 Storage Structure
In contrast to traditional XML based schemata used in BIM and other open data models like IndoorGML [48] or Brick [9], the data inside the Adaptive City Platform is propagated using JSON objects. There were several reasons for this. First, it enables the data to be easily readable to humans and provides numerous benefits when working on both front-end and back-end languages [29]. Secondly, we envision the data flowing through the platform to utilise a Linked Data system, and we found JSON-LD to be the most compliant with our needs.

We annotate our data with auxiliary metadata properties, such as unique name identifiers, building descriptions and location data. This is done by appending metadata properties to the propagated data rather than replacing them, resulting in self-contained data structures that do not require additional API queries to be contextualised in use. For example, a sensor data object does not rely on querying the BIM to obtain information on the sensor’s location, because that data is already embedded within the sensor metadata.

Our data structures can generally be separated into three distinct types: (i) building data, (ii) sensor metadata and (iii) sensor readings. This separation provides flexibility in data processing, as well as making it more usable for visualisation. Finally, sensor readings are stored in a data repository, where the data is sorted by date, as well as saved individually for each sensor and then sorted by date. While the data is duplicated this way, it does become useful when working with API queries that ask for data over a time period rather than from a specific sensor. Since we focused on a single building deployment, we found that saving our data in the file system worked sufficiently well, however, with a bigger deployment the use of a different data repository, e.g., a SQL database, would be a viable option instead.

4.2 Platform Modules
The ACP real-time platform is based on the underlying architecture of the SmartCambridge framework [43] which is implemented using Eclipse Vert.x [17], to allow for real-time event-based data processing. The platform consists of multiple Vert.x modules termed verticles responsible for asynchronous, non-blocking data traversal. Figure 4 illustrates these modules as Data Collection Architecture, Stream Processing, and Real-Time Monitor (RTMonitor) blocks.

Data Collection Architecture. Incoming sensor data is captured by our Data Collection verticles that listen for incoming messages over MQTT. After a message is received, the modules timestamp binary sensor data, archive it in the file system and further propagate sensor readings down the platform as JSON files. Using the SmartCambridge framework allowed us to reach an average latency of 160ms from sensors sending the data to its use in the sample application.

Stream Processing Modules. Stream Processing is comprised of modules that further parse the received JSON files. These modules decode and write the data to the file system as well as inserting relevant metadata in the database and passing data to the RTMonitor for low-latency data visualisation. Stream Processing modules are versatile because they are able to sort the incoming data and separate it into different directories per event basis for future API use.

RTMonitor. The RTMonitor is a module allowing client web pages to issue subscriptions to the Stream Processing verticles and receive updates via websockets on the specified URI the moment sensor readings are passed down the pipeline. The RTMonitor then manages these subscriptions using a token-based system, where users can issue subscriptions by specifying the sensor data they would like to receive.

However, while the RTMonitor is an integral part of the real-time platform, we also use regular API queries to fetch historical data and BIM-related metadata from our database.

4.3 APIs
We created four APIs to fetch the sensor metadata, sensor readings, BIM metadata and rendered SVG BIM data. These API endpoints are used in our example application to provide the scaffolding that permits the non-blocking real-time data flow straight to the front-end visualisation.

The front-end is maintained using the API endpoints that interface with the PostgreSQL database and the file system to query the BIM and sensor data. Three of the four API endpoints return JSON-formatted historical data that is easily manipulated by JavaScript or Python. The /space API endpoint is unique in that it returns nested XML SVG objects that are rendered on screen for the building visualisation. The diagram shown in Figure 5 and the Table 1 describes the API in more detail.

### Table 1: API description.

| Endpoint  | Description                              |
|-----------|------------------------------------------|
| /api      |                                        |
| /bim      | returns building metadata                |
| /space    | returns rendered SVG objects based on the BIM data |
| /sensors  | returns sensor metadata                  |
| /readings | returns sensor readings                  |

![Figure 5: API structure.](image)
4.4 Building Reference Data Platform

In order to evaluate the Adaptive City Platform, we deployed and tested IoT sensors in a non-BIM native building, that did not have any Revit files, other than 2D floor plans. Initially we considered adopting a BRICK ontology for the building, however, after testing it we decided to craft a JSON-LD compliant proprietary ontology with the ability to be converted to BRICK. We did this for the following reasons.

First, BRICK’s focus is to capture information on building operations and improve application portability – not to create another building design model [9], like the IFC standard. This was problematic as our system was deployed in a non-BIM native building that lacked any IFC files for contextual visualisation. Thus, the ability to embed precise location coordinates and boundary data, in addition to relational information, was important for us. While achieving this using the BRICK ontology was technically feasible by saving the additional properties as literals, we found it to be impractical. The impracticality was caused by our metadata properties existing as dictionaries, and hence saving them as strings would (a) reduce human readability and (b) slower data retrieval time, as string literals would have to be converted back to objects.

Second, we found that including additional metadata properties like timestamps for recognising patterns were key. For our application area, spatio-temporal data is highly relevant, but BRICK mostly ignores such information by disregarding time, making it unsuitable for true real time system applications.

Third, we found that our existing JSON (and further JSON-LD) data structures, whilst providing more flexibility than RDFa used by BRICK, could be easily converted to BRICK ontology, providing the possibility of portability when needed. Therefore, we crafted a sample ontology to reference our building and sensor metadata in the platform.

For our building reference data we used a universal naming convention for buildings, floors and rooms called crates. A crate is an object (however, usually a building or part of a building) with a defined enclosed boundary that denotes its perimeter. Crates can also be nested by defining the outer crate as a parent crate. For example, we can define a crate that denotes a building, and within that building we have other crates that are its children – e.g., floors, that in return also have their own children – e.g., room crates. This parent-child relation in our building data allows for hierarchical data management that is useful for efficient API queries.

We use the nested tree-like structure in our BIM database to retrieve data associated with our building. For this reason, the API call /bim/get/<crate_id>/<children> has the optional <children> property that allows us to specify if any children should be included for that particular crate.

Additionally, every crate has a crate_type property to select specific object types, e.g., floor, building, etc., as shown in Figure 6. Further metadata associated with crates includes location, boundary and crate type, as well as a UNIX timestamp that indicates the date any information had been updated. Overall, using this metadata structure permits us to index our objects in the PostgreSQL database using any of the properties defined below to allow for fast and flexible metadata retrieval and data management. A complete example of the BIM metadata is given in Table 2, and a sample response to the BIM API query is shown in Listing 1.

4.5 Sensor Reference Data Platform

Similarly to crate_id, sensors also have unique identifiers, the acp_id property, defined using the manufacturer’s name and a six character-long unique identifier. Since different types of sensors often have varying attributes, our sensor metadata table is defined by two columns containing: (i) the acp_id identifiers and (ii) another containing the other auxiliary properties. Such auxiliary metadata includes information on sensor properties like the location and type. Complete metadata description can be found in Table 3, as well as a queried sample sensor metadata object from /sensors/get/<acp_id> in Listing 2.
Data Management for Building Information Modelling in a Real-Time Adaptive City Platform

| crate_id | parent_crate_id | crate_type | location | boundary |
|----------|-----------------|------------|----------|----------|
| WGB      | -               | "building"| { "system":"GPS", "acp_lat":-27.116667, "acp_lng":-109.366667, "acp_alt":0.0 } | { "system":"WGB", "boundary":[[0,0],[......],[......],[245,56]] } |
| GF       | WGB             | "floor"    | { "system":"WGB", "x":36.5, "y":39, "f":0, "zf":0 } | { "system":"WGB", "boundary":[[0,0],[......],[......],[245,56]] } |
| FE11     | FF              | "room"     | { "system":"WGB", "x":22.06, "y":34.67, "f":1, "zf":0 } | { "system":"WGB", "boundary":[[0,0],[0,78],[73,78],[73,0]] } |

Table 2: BIM metadata example, extracted from our database, showing how data can be indexed based on range of properties.

| key       | definition                                                                 |
|-----------|-----------------------------------------------------------------------------|
| acp_id    | sensor identifier, globally unique e.g., elsys-eye-049876.                   |
| acp_ts    | epoch timestamp most relevant to data reading or event, e.g., "1586461606.465372". |
| acp_type  | sensor type, determines data format e.g., elsys-eye.                         |
| acp_event_value | qualifier or data reading for event, e.g., open.                         |
| acp_event | event type, for time-stamped events, e.g., open/close.                     |
| acp_location | location using a custom coordinate system e.g., { "system": "WGB", "x": 12, "y": 45, "f": 1 } |
| acp_confidence | a value 0-1 indicating the reliability of the sensor reading. |

Table 3: Sensor metadata properties.

Listing 2: Sensor object metadata, received after querying the API endpoint /sensors/get/elsys-co2-041ba9.

Finally, we expanded the sensor API to be capable of retrieving sensor metadata by their physical deployment location, specifying the parent crate in the API call /sensors/bim/get/<crate_id>.

4.6 Historical Sensor Data Platform

We attach UNIX timestamps to all metadata and incoming sensor readings as soon as they enter the ACP. Depending on the type of data received, JSON packets further propagate to the PostgreSQL database or the file system, where they become historical data. Historical sensor readings can be fetched by querying the API with a sensor’s acp_id property. The API call then returns the following JSON file from the most recent entry in the database, as shown in Listing 3.

4.7 Spatial Coordinates Data Platform

In both building and sensor metadata we refer to our spatial coordinate system as acp_location. As the ACP needs to provide data for both relative (in-building) and global (WGS84) positioning references for sensors and crates alike, we introduce three parallel location reference systems.
Global. The definitive common reference system constituting of latitude, longitude and altitude. The global system is necessary for outdoor sensors as latitude and longitude coordinate system is used while interacting with the site template view in the sample application (Figure 7A). We define this in our BIM model by setting the acp_location parameter as GPS, shown in Table 2.

In-building coordinates. We use a spatial coordinate system unique to each building, typically when interacting with in-building floor plan or 3D views of sensors or data. Sensors that transmit their position in the building (particularly relevant for sensors that move around) may use this system in their sensor data. We set in-building coordinates by specifying the acp_location’s parameter as the building’s name, as illustrated in Table 2 and Listing 1.

Building object hierarchy. Hierarchies are often used in BIM software, and in the ACP the hierarchy is defined with the parent_crate parameter. It is reasonable for a sensor (or other monitored device) to be recorded as being in location based on a crate i.e. a room/office, which relates to the BIM data structured as site → building → floor → room → window, etc. This hierarchy is often natively used when collating or browsing in-building information, such as electricity consumption in specific rooms or floors.

By annotating our metadata with a combination of these three location systems, we achieve high flexibility to manage and visualise the BIM and sensor data in multiple ways. This is exemplified in the following section where we showcase how the real-time platform can be effectively used in a real-world scenario.

4.8 Visualisation
To best illustrate how the ACP can perform in the wild, we conceived a data visualisation application to monitor and visualise the BIM-IoT fusion data. The visualisation acts as an example of how we can develop software based around our platform and APIs to facilitate the BMS and IoT integration.

The front-end consists of five templates showing the BIM-IoT fusion at different levels. We use multiple templates for our data visualisation, allowing us to select the granularity at which the information is displayed. The templates are loaded in a hierarchical order, where the first template (Figure 7A) represents the site view, while the last one (Figure 7D) shows the crate-level view, followed by a sensor readings template (Figure 8) with time-series plots.

4.8.1 Site Template. The site-level visualisation (Figure 7A) displays buildings and sensors in aggregated groups on the campus. The map is rendered using Leaflet and OpenStreetMap [36], an open-source JavaScript library for interactive maps.

We query our database to return the boundary data for the buildings we would like to render on the screen. After receiving the boundary coordinates in latitude and longitude we then render polygons over the original building locations on the map. Finally, users are able to navigate to the other templates by clicking on individual buildings.

4.8.2 Building Template. In order to render the building-level template we query the API to return individual floor and room boundaries based on the information we have in the BIM database. After fetching the boundary data, we then render the building in 3D using the Three.js library [49].

The building-level template (Figure 7B) allows for a general building view to inspect individual crates, or proceed to any of the floor plan templates by clicking on the floor icon.

4.8.3 Floor Plan Template. The floor-level template (Figure 7C) has several important features. Primarily, it has the ability to display heat-maps of sensor readings, as well as the most recent sensor data. Users can inspect sensor readings by hovering over the sensor icon on the floor plan. Upon doing so, an API call is executed fetching the most recent data for the particular sensor in question.

We query the /space/get_bim_floor_number/<floor> API endpoint to acquire all crates that are on the queried floor. The API returns an SVG object along with the metadata necessary to render the floors and individual rooms. We then use D3.js [12] to render the SVG on screen and make it interactive.

With this approach we eliminate the need to possess static SVG files in the file system, and instead rely on our predefined BIM data structures that can be easily updated and generated into SVG files on demand.

Users proceed to enter the floorspace template by clicking on a crate on the floor-level template.

4.8.4 Floorspace Template. This final BIM-generated template (Figure 7D) is used to display the finest granularity spatial data, including the crate’s BIM description and metadata for all the sensors deployed in that crate. The selected crate of type room is zoomed in on, allowing to more accurately display the sensor location.

4.8.5 Sensor Template. As users travel down the spatial hierarchy of the web application, the data becomes increasingly more granular, allowing detailed time-series data analysis. The sensor-level template is accessed by clicking on a sensor on either the floor plan or floor space views. The user is then redirected to the sensor template where detailed spatio-temporal time-series data is shown, as illustrated in Figure 8. This template allows to visualise the data provided by every in-building sensor. Users are able to select the timeframe they wish to inspect. The visualisation changes in real-time as new data readings propagate through the platform.
5 DISCUSSION

During the development of the ACP and the accompanying visualisations, we identified four major themes associated with BIM-IoT fusion: (i) scalability, (ii) sensor-derived constraints, (iii) BMS actuation and (iv) privacy.

5.1 Scalability

While the ACP does provide real-time sensor information, it remains a single building deployment with a limited scope. We anticipate using our modular architecture to add additional buildings simply by appending the BIM files to the current PostgreSQL database.

One challenge in doing so is to acquire the BIM data in the first place. In our example we used a non BIM-native building and as a result had no access to 3D building models and other key BIM data. Instead we parsed the SVG floor plans for our building to extract the boundary data for individual crates. While such CAD file parsing could also be adapted to other 2D-friendly formats such as DXF or DWG, the differences between the way building data is stored remains an issue. As use of BIM in the design phase of buildings increases, a potential solution would be to write a Revit plugin that would automatically generate and export the requested data in a suitable ontological standard [41], or utilise IFC-BRICK converter [9].

In terms of the back-end system performance, we have no reason to suspect that any further scalability efforts would affect the real time platform itself. Since our real-time architecture is based on the Vert.x system used in the SmartCambridge project [43] that had ~1000 sensors sending data every 20 seconds, we have encountered no slowdown with asynchronous modules processing data with low-latency.

5.2 Sensor-produced constraints

Event-driven data is crucial for responsiveness but the majority of off-the-shelf sensors, with the exception of interrupt-based devices, are only capable of sending periodic readings, e.g., CO$_2$ every 5 minutes. Furthermore, such sensors are rarely reconfigurable to allow for dynamic data transmission based on user-defined threshold values, making it impossible to implement timely event-driven behaviour to track sensor readings in real-time. Finally, many sensors lack control over where data is delivered – while we chose to rewrite factory sensor firmware in some cases, the complexity and the required man-hour efforts may not necessarily be worth the desired outcome.

Overall, sensor choice and their detailed capabilities plays a key role in determining the timeliness factor of any real-time platform, therefore considerable attention should be paid to what sensors are integrated within the IoT ecosystem.

5.3 BMS Implementation and Actuation

As research progresses towards the adoption of a universal ontological standard to incorporate BIM and BMS with IoT technologies, machine learning-based approaches to building automation and actuation will become more important. According to Tang et al, many current BIM-IoT fusion attempts fail to achieve the enclosed loop system capability by not enacting on actuation and only being capable of sensing [45].

5.3.1 BMS in-the-loop

BMS implementation is often a non-trivial task due to siloed software systems, as well as some buildings not having actuation capabilities implemented and accessible by their BMS [10]. While beyond the scope of this paper, such attempts to normalise BMS data are made through the creation of building ontologies, providing unified metadata schemas.

However, a substantial number buildings have neither BMS nor BIM-enabled facilities management systems, and therefore lack basic programmatic access. While deploying IoT sensor would make such buildings ‘smarter’, the lack of actuation equipment means that we would not be able to reach the full potential of modern, BIM-native buildings.

While implementing actuation components in such buildings would likely be the next step after the deployment of sensors...
and machine learning tools, it would also mean a substantial financial commitment to AEC stakeholders. Nevertheless, it has been reported that building automation does indicate a substantial energy savings [15, 53], thus potentially making expenses cost-effective, even ignoring the benefits that such systems could bring in risk/disaster management and anomaly detection.

### 5.3.2 Data Analysis, Prediction and Actuation

Even though HVAC systems in buildings have been around for a considerable time and have shown positive results in increasing energy efficiency [53], the mass deployment of IoT sensors in the built environment will provide much more granular data feeds. This increase in the amount and detail of data is what drives the need for fast and efficient real-time platforms like the ACP, as they permit the analysis and decision making to be done on the spot, once the right machine learning tools are implemented in the data pipeline.

The ongoing COVID-19 pandemic prevented us progressing our deployment to the point where we had dense enough deployment of sensors and rich enough data from those sensors, building automation and data analytics remains a topic of great importance for future research.

### 5.4 Privacy

Privacy in IoT-backed smart buildings has grown in importance over the last decade [2, 15, 27], but more work is needed as privacy issues remain a minefield and may hinder further technology adoption [25]. It is important to note that while sensing in the presence of people in itself causes privacy concerns, real-time tracking further amplifies these concerns, as data integrity may not always be guaranteed [45]. Even though none of the sensors are used to track individuals directly, latent information may still cause ethical issues, as it may be possible to, e.g., infer information about office workers by analysing CO2 data and electricity consumption [5].

The primary concern about deploying sensors in the built environment has to do with the ownership of and the access to the collected data. Sensors are relatively easily locatable objects in the workplace and a possible source of anxiety due to the collection of data. Currently there are few IoT deployment guidelines covering office spaces and data ownership, other than the GDPR, focused on data collection and storage [50]. We envision a privacy-oriented system provided by a capability-based platform where people are given access to data based on a combination of factors such as their day-to-day proximity to sensors (you should have means to examine data collected about your behaviour), and role in the company (a building manager may need to see aggregate data for the whole building).

For example, people with sensors deployed in their offices would have the access and control of the high granularity (or frequency) time series data that they produce. However, such access would be limited to the rest of the building, providing a decreasing level of granularity (or frequency) across time and space domains. An example of this in practice could be the building facility manager having the access to entire building data on a weekly scale without specific information on individual offices.

In addition to personal privacy in the workplace, in a case study presented by Cascone et al, security of data was the second most common concern after privacy [15]. The deployment of similar web applications to ours comes with a risk of data being leaked to third parties. Therefore, considerable care should be taken to secure both the BIM data, and the sensor data so that it would only be accessible to authorised personnel. The same capability-based platform could be employed in cases where third parties are given access to some of the low frequency, low resolution spatial data.

### 6 CONCLUSION

The lack of standardisation in BIM-IoT fusion means that the AEC industry is unprepared for mass IoT sensor deployment in the built environment. Our contribution to research on BIM-IoT fusion has been to describe a high-level real-time software architecture for processing data, and to show how it can be used to visualise in-building sensor data with useful features for facility managers and building inhabitants alike.

As it stands, our real-time Adaptive City Platform is currently unique, playing a key role in the way the IoT sensors are integrated with BIM. By using the concept of events in capturing time-sensitive data in the ACP we were able to achieve low-latency stream processing capable of instantaneous updates to our in-house visualisation.

We have further shown how hierarchical data visualisation can be used to display the collected data at different levels of granularity. Furthermore, we have described how metadata is described and our data is stored, both buildings (crates) and sensors.

While we have demonstrated that the ACP works, there are still substantial challenges in how to tackle mass sensor deployment and further data analytics. As workplaces and buildings eventually reopen as the current phase of the COVID-19 pandemic subsides, we anticipate being able to increase the density and spread of our deployment, enabling us to collect and process more and richer data to explore how the ACP can support prediction and actuation.

### REFERENCES

[1] 2016. The Internet of Things in Smart Commercial Buildings 2016 to 2021. http://memori.com/portfolio/internet-things-smart-commercial-buildings-2016-2021/

[2] Yuvraj Agarwal and Anind K. Dey. 2010. Toward Building a Safe, Secure, and Easy-to-Use Internet of Things Infrastructure. IEEE Computer 44, 4 (2010), 88–91.

[3] Ridwan Alam, Nutta Homdee, Sean Wolfe, James Hayes, and John Lach. 2019. Besi: behavior learning and tracking with wearable and in-home sensors-a dementia case-study. In Proceedings of the International Conference on Internet of Things: Design and Implementation. 281–282.

[4] Sofia Antonopoulou and Paul Bryan. 2017. BIM for heritage: developing a historic building information model. Historic England.

[5] Hany F. Allam and Gary B Wills. 2020. IoT security, privacy, safety and ethics. In Digital Twin Technologies and Smart Cities. Springer, 123–149.

[6] Salman Azhar. 2011. Building information modeling (BIM): Trends, benefits, risks, and challenges for the AEC industry. Leadership and management in engineering 11, 3 (2011), 241–252.

[7] Salman Azhar, Michael Hein, and Blake Sketo. 2008. Building information modeling (BIM): benefits, risks and challenges. In Proceedings of the 46th ASC Annual Conference 2–5.

[8] Bharathani Balaji, Arka Bhattacharya, Gabriel Fierro, Jingkun Gao, Joshua Gluck, Dezhi Hong, Aslak Johansen, Jason Koh, Joern Ploennigs, Yuvraj Agarwal, et al. 2016. Brick: Towards a unified metadata schema for buildings. In Proceedings of the 3rd ACM International Conference on Systems for Energy-Efficient Built Environments. 41–50.

[9] Bharathani Balaji, Arka Bhattacharya, Gabriel Fierro, Jingkun Gao, Joshua Gluck, Dezhi Hong, Aslak Johansen, Jason Koh, Joern Ploennigs, Yuvraj Agarwal, et al. 2018. Brick: Metadata schema for portable smart building applications. Applied energy 226 (2018), 1273–1292.

[10] Burcin Becerik-Gerber, Farrokh Jazizadeh, Nan Li, and Gulben Calis. 2012. Application areas and data requirements for BIM-enabled facilities management. Journal of construction engineering and management 138, 3 (2012), 431–442.
