A linked-data paradigm for the integration of static and dynamic building data in Digital Twins

Dimitris Mavrokapnidis
University College London
d.mavrokapnidis@ucl.ac.uk

Kyriakos Katsigarakis
University College London
k.katsigarakis@ucl.ac.uk

Pieter Pauwels
Eindhoven University of Technology
p.pauwels@tue.nl

Ekaterina Petrova
Eindhoven University of Technology
e.petrova@tue.nl

Ivan Korolija
University College London
i.korolija@ucl.ac.uk

Dimitrios Rovas
University College London
d.rovas@ucl.ac.uk

ABSTRACT

Digital Twins is an emerging field of research, mainly because they span the entire building lifecycle promising to uncover hidden inefficiencies and deliver data-driven applications. Broadly defined as real-time digital representations of physical assets, Digital Twins require a connection between static and real-time data. However, building information is usually stored in different formats across the lifecycle, making data integration a challenging task. We hereby often rely on linked data technologies, yet overall system integration approaches with multiple types of data sources. In this work, a data linking methodology is proposed to combine static building design data from Industry Foundation Classes (IFC) and dynamic data using the Brick Schema; a domain ontology which configures data analytics applications during the operational phase. To facilitate this integration, we develop a tool to facilitate the linking of building topology, product, and sensor data using the two schemata. The implementation of our methodology in a real test case demonstrates its significance in combining diverse data sources which can be an important step for the delivery of Digital Twin applications.

CCS CONCEPTS

• **Information systems → Graph-based database models**.

KEYWORDS

Ontologies, Linked Data, Semantic Web Technologies, BOT, IFC, Brick, Metadata, OWL, RDF, Smart Buildings, Digital Twins

ACM Reference Format:

Dimitris Mavrokapnidis, Kyriakos Katsigarakis, Pieter Pauwels, Ekaterina Petrova, Ivan Korolija, and Dimitrios Rovas. 2021. A linked-data paradigm for the integration of static and dynamic building data in Digital Twins. In *The 1st ACM International Workshop on Big Data and Machine Learning for Smart Buildings and Cities (BuildSys '21), November 17–18, 2021, Coimbra, Portugal*. ACM, Coimbra, Portugal, 4 pages. https://doi.org/10.1145/3486611.3491125

BuildSys '21, November 17–18, 2021, Coimbra, Portugal

ACM ISBN 978-1-4503-9114-6/21/11...$15.00
https://doi.org/10.1145/3486611.3491125

1 INTRODUCTION

With the advent of laser scanners, drones and IoT sensors, it is now feasible to collect both as-built and operational data which can serve to create a Digital Twin of the physical building; what is generally defined as a *"real-time virtual representation of what has been produced"* [5]. In practice, a Digital Twin can fulfil various purposes, has different data requirements and thus, needs a combination of building data. However, building information is siloed in different formats, mainly due to different requirements and complexities across the lifecycle, which impede data exchange processes, causing information loss and limiting data reuse.

Building Information Modelling (BIM) was devised to provide seamless information exchange across the entire building lifecycle; still, BIM has not been widely applied in the operational phase, mainly due to the complex nature of the underlying open standard data schema, Industry Foundation Classes (IFC). This complexity lies in the primary focus of IFC on geometry representations which limits its extensibility, modularity and cross-domain linking capabilities.

On the other hand, Brick Schema is considered as the most expressive and complete ontology-based data model that captures the operational aspects of buildings. More specifically, Brick is designed to represent the entire domain of smart buildings and thus focuses on HVAC, Lighting and Building Management Systems [1]. Given that smart buildings do not need editable envelope characteristics, representing design-related information such as topology and building product data are considered out of Brick's scope.

On the grounds that IFC and Brick mainly capture building information in two different phases, there is an opportunity in the integration of data from both sources. This paper aims to provide specific guidelines and a set of tools for the integration of static data from the building design phase (using IFC) with dynamic data from building operation phase (using Brick).

2 BACKGROUND

IFC's complex structure and limitations in representing operational aspects of buildings has led to the adoption of semantic web technologies and linked data [8]. Aiming to decentralise building information and link it with other domains, EXPRESS (IFC's underlying language) has been translated into an owl version, i.e. ifcowl [7]. Although ifcowl makes available IFC data in RDF format under the W3C standards, and thus, able to link it with other domain information, the complexity of IFC leads to a large number of triples

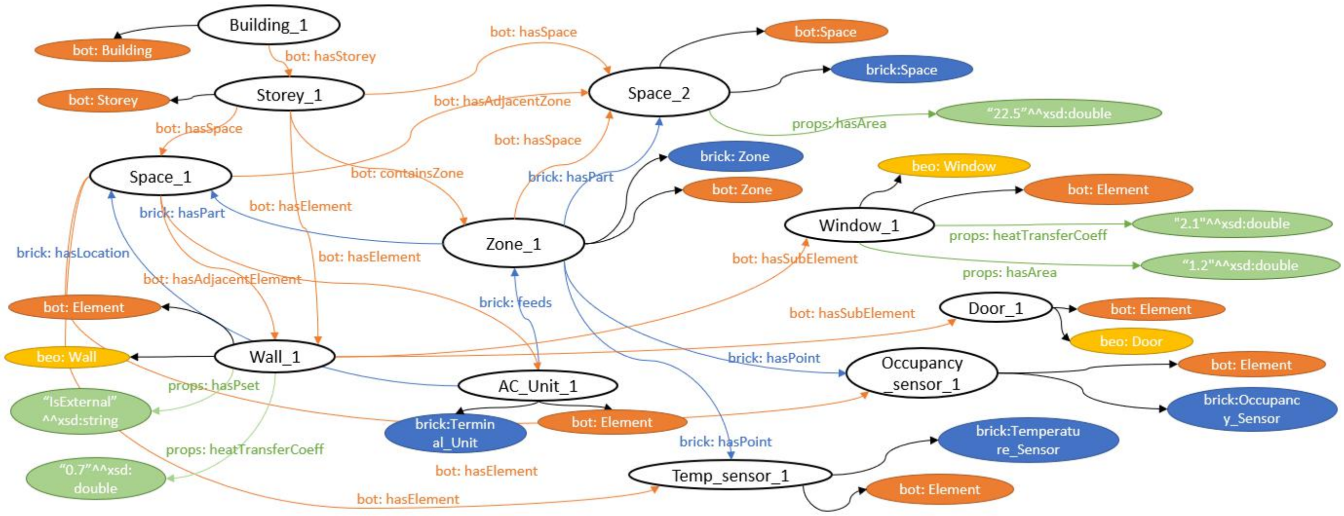

Figure 1: Conceptual building data model using BRICK (blue), BOT (orange), PROPS (green) and BEO (yellow) ontologies

and thus still creates storage inefficiencies in terms of querying the graph. For example, a recent study introduced an architecture for a Digital Twin which integrates static and dynamic data based on ifcowl, leading to a large unified graph which would be both time-consuming and impractical to use [3].

To meet the requirements for a more simple, modular and extensible data storage model, the Linked Building Data (LBD) group devised the Building Topology Ontology (BOT) to be used as foundation in combination with other ontologies, including building elements (BEO), product catalogues (BPO), IoT sensor observations (SSN) and other ontologies [11]. Through he IFC-to-LBD converter [2], the user can extract IFC data and transform them into RDF format, using modular ontologies (i.e. BOT, PROPS, BEO etc.), making it possible to store data and make queries more easily compared to ifcowl-enabled graphs.

In contrast to general-purpose ontologies of the LBD group, Brick consists of multiple instances, aiming for completeness of the building operation domain rather than flexibility, and thus can lead into more constrained representations which limits its use in a variety of building energy applications [9]. Despite Brick is considered as the most expressive and complete data schema compared to other ones in the same domain, it still lacks in completeness (i.e. achieved 58%) based on a recent study [10], highlighting the need for integration with other ontologies, such as BOT, SAREF and BACS.

In a recent study, Fierro et al. [4] introduced an algorithm which automatically converts the common entities from different data models (including IFC) across the lifecycle into a Brick model. Although the goal of this study is to maintain building data over the course of the entire lifecycle, this approach leads to a graph which can only include data represented by the Brick Schema. Instead, to create a foundation data model for Digital Twin applications in buildings, a combination of static and dynamic data can be realized by linking IFC and Brick schemata.

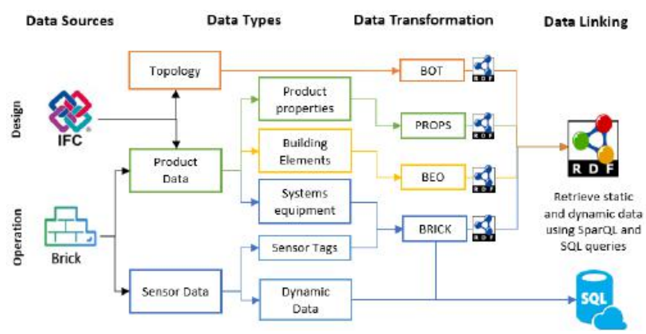

Figure 2: Proposed IFC-Brick integration framework

3 METHODOLOGY

In figure 1, we present a conceptual example of a building data model that illustrates how different instances of a building (i.e. Storey_1, Space_1, Wall_1, Temp_Sensor_1) are represented using the proposed set of ontologies. Following the design principles of BOT, every instance in our model is either defined as *bot:Zone* (or subclass of it, i.e. bot:building, bot:storey) or *bot:element*. In addition, BEO ontology is used as an additional layer of information, enhancing the *bot:element* instances with the type of the building element. Although PROPS is still in conceptual stage [2], it is used here to demonstrate how we can assign properties in building elements such as room area and heat transfer coefficients of walls.

The proposed integration framework in Figure 2 demonstrates how IFC and Brick Schema can be used to extract three diverse categories of building data; topology, building products, and sensor data. This architecture builds upon the IFC to LBD converter [2] to extract and transform IFC data into RDF format. More precisely, BOT is used to represent the topology of the building while PROPS and BEO ontologies are applied to represent the building data properties and the type of building elements respectively. In addition, systems equipment and sensor tags are represented using the Brick

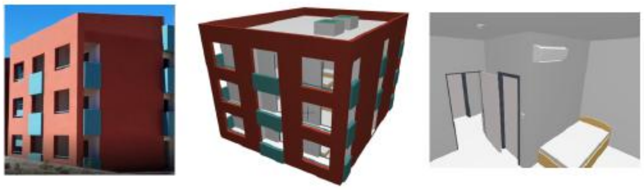

Figure 3: Test Case - Physical vs Digital representation

Schema, while time series data corresponding to the sensor tags are stored in a historical time-series database. Overall, static building data are stored in RDF format while dynamic data are stored in a SQL database and Brick Schema provides the connection between the two databases.

Brick instances are generated in accordance with the existing alignment instructions between BOT and Brick [12]. In our model as illustrated in Figure 1, spaces, zones and elements are used for alignment between Brick and BOT. The relationship between *ifc-Space*, *bot:Space* and *brick:Space* is central in the proposed solution. Although Brick would be able to represent additional spatial information (i.e. Building_1), there is not practical value in doing so, since we expect the user to use Brick for the retrieval of operational information and BOT for spatial characteristics.

To facilitate the RDF data generation, from IFC to BRICK, BOT, BEO, and PROPS, an online knowledge graph generator is developed in JAVA [1] as part of this study. It builds on top of the existing efforts of the IFC to LBD converter [2], by extracting IFC data to create additional BOT entities such as bot:ContainsElement, bot:hasAdjacentElement which represent the relationship between spaces and bounding elements. Using our tool, the user is able to upload an IFC file and automatically transform it into RDF data following the concepts of Figure 1.

Overall, the resulting unified knowledge graph can only represent the physical Brick entities (i.e. Temp_Sensor, AC_unit, etc) as generated from IFC. Non-physical Brick entities related to the operation of BMS, mechanical systems and sensors (e.g. Setpoints, Signals, etc.) cannot be represented in IFC and thus, cannot be generated from it. As a result, the output of this tool should be considered as a "ready to connect with Brick" graph. To connect the unified graph with a full Brick, the serial number of the devices is included in IFC as "Device_ID", using property sets (i.e. IfcPropertySet).

4 TEST CASE

To demonstrate the practical value of the proposed integration approach, we use a real-world building that serves as a student hall at Technical University of Crete (TUC). This is a three storey building, consisting of en-suite student rooms and a communal kitchen per floor. Each space contains a split air conditioner supporting heating and cooling requirements, a thermostat to control indoor climate and an occupancy sensor. The real-world building along with its digital representation are illustrated in Figure 3.

We create a 3D parametric model in Revit from which IFC data are exported and enriched with second level space boundaries using the CBIP tool [6], an algorithm that applies geometric processing to address issues from the IFC exportation process. The proper authoring of IFC's space boundaries is essential for creating a full representation of the topology, including spaces' adjacency and their bounding building elements. Similar processing is required to create the missing semantic link of the mechanical and sensor equipment with the space on which they are installed.

For this particular building, a Building Management System (BMS) system has already been in use, monitoring occupancy, temperature, energy consumption and other operational parameters of the building and storing real-time dynamic data in a SQL server. To organise and retrieve data from the server, a Brick model has already been generated and used to represent in-depth information about sensor data points and the related mechanical systems.

5 RESULTS

The implementation of our proposed method occurs in two steps: (1) the transformation of IFC into a unified knowledge graph and (2) its connection with the Brick representation of the BMS system. The first part is automatically accomplished thanks to our knowledge graph generator tool; whereas, the connection with the Brick graph that includes the sensor tag information is performed manually for demonstration of the proposed integrated system. The IFC file, the corresponding results from the knowledge graph generator and the final integrated knowledge graph are available on Github [2]. To demonstrate the utility of the proposed unified graph, we form a simple SPARQL query (Listing 1) to extract static data from the graph, including the window and space areas as well as the IDs of the air zone temperature sensors for each space in the building. These are then used to retrieve time-series data from the historical database which are processed to calculate the daily mean air temperature. The results of this process in Table 1 demonstrate that it is possible to extract diverse building data. Similar queries can also return additional data including window properties, bounding wall elements, adjacent spaces, room area as well as dynamic data from temperature and occupants sensors. These results demostrate how the proposed approach can be used for the retrieval of both static and dynamic data and potentially support a various Digital Twin use cases in buildings.

```
prefix schema:<http://schema.org#>
prefix rdf:<http://www.w3.org/1999/02/22-rdf-syntax-ns#>
prefix owl:<http://www.w3.org/2002/07/owl#>
prefix bot:<https://w3id.org/bot#>
prefix xsd:<http://www.w3.org/2001/XMLSchema#>
prefix rdfs:<http://www.w3.org/2000/01/rdf-schema#>
prefix brick:<https://brickschema.org/schema/1.1/Brick#>
prefix beo:<https://pi.pauwel.be/voc/buildingelement#>
prefix om:<http://openmetrics.eu/openmetrics#>
prefix props:<https://w3id.org/props#>
SELECT DISTINCT ?sp_a ?win_a ?temp_id
/*temp_id:temperature_sensor_id*/
/*sp_a:space_area*/
/*win_a:window_area*/
WHERE {
    ?sp a bot:Space, brick:Space ;
        props:hasArea ?sp_a .
    ?sp bot:adjacentElement ?win.
    ?win a beo:Window ;
```

[1]https://github.com/kyriakoskatsigarakis/openmetrics/blob/master/openmetrics-kgg/examples/101/101_Enhanced_Detailed.ttl

[2]https://kgg.openmetrics.eu/

```
        props:hasArea ?win_a .
    ?th a brick:Thermostat ;
        brick:hasLocation ?sp .
    ?tset a brick:Zone_Air_Temperature_Setpoint ;
        brick:isPointOf ?th ;
        brick:timeseries [brick:hasTimeseriesId ?tset_id
     ].
    }
```

Listing 1: SPARQL query on the unified knowledge graph

SPARQL Query results			SQL Query results
sp_a	**win_a**	**temp_id**	**Daily Mean Temp**
"18.68"	"8.16"	"TUC.245.77.R440"	28.65
"17.12"	"8.16"	"TUC.245.77.R460"	29.81
"18.68"	"8.16"	"TUC.245.77.R240"	27.68
"17.92"	"8.16"	"TUC.245.77.R60"	27.49
"17.12"	"8.16"	"TUC.245.77.R260"	28.33
"18.32"	"8.16"	"TUC.245.77.R200"	28.0
"18.68"	"8.16"	"TUC.245.77.R100"	30.46
"18.32"	"8.16"	"TUC.245.77.R20"	29.79
"15.75"	"8.16"	"TUC.245.77.R40"	29.31
"15.75"	"8.16"	"TUC.245.77.R280"	29.69
"15.75"	"8.16"	"TUC.245.77.R80"	30.34
"17.12"	"8.16"	"TUC.245.77.R420"	28.38
"17.92"	"8.16"	"TUC.245.77.R220"	29.04
"18.32"	"8.16"	"TUC.245.77.R380"	27.61

Table 1: SPARQL and SQL results (where sp_a: space area (m2), win_a: window area (m2), temp_id: temperature sensor ID, and Daily Mean Temperature (C))

6 DISCUSSION AND CONCLUSIONS

This study claims that the integration of two widely applied data models, adopting the linked data principles, will be a key enabler of Digital Twin applications. As a first step towards this direction, we focus on linking data from IFC and Brick, two diverse data models from different lifecycle phases. To facilitate this integration the knowledge graph generator has been developed to transform IFC data into a knowledge graph automatically. This tool enhances the existing capabilities of the IFC to LBD converter [2], representing additional building information, including advanced topological relationships (i.e. adjacent spaces). The resulted unified graph is then ready to link with a full Brick graph representation (either of a BMS or IoT system). Our test case demonstrated that we can create a single source of building data, able to retrieve both static and dynamic data using a combination of SparQL and SQL queries with a view to support the data requirements for potential Digital Twin use cases.

An important limitation that needs to be considered is the availability of properly defined IFC data. Despite IFC is a rich data model, extracting IFC data from BIM authoring tools (such as Revit) is not always fully correct. For example, it was not possible to extract space boundaries correctly from Revit's IFC exporter. For this reason, CBIP tool [6] was applied to enhance IFC with correct space boundaries. As a result, our knowledge graph generator will not be able to extract information about adjacent spaces without proper authoring of space boundaries in IFC. Thus, an adequate level of expertise with IFC as well as modelling guidelines in BIM authoring tools (i.e. Revit) are required to overcome these limitations.

Future research aims to implement a full Digital Twin use case using our proposed data model. We aim to incorporate the strengths of Brick in representing operational aspects within the 3D visualization environment offered by IFC data. This integration could offer a practical virtual representation of the building, in which the user would be able to locate different building and system elements within the building, optimise space usage and visualise various simulation scenarios.

ACKNOWLEDGMENTS

The research leading to these results has been partially funded by the CBIM-ETN funded by the European Union's Horizon 2020 research and innovation programme under the Marie Skłodowska-Curie grant agreement No 860555 and by the European Commission H2020 "COnstruction-phase diGItal Twin mOdel (COGITO)" under contract #958310.

REFERENCES

[1] Bharathan Balaji, Arka Bhattacharya, Gabriel Fierro, Jingkun Gao, Joshua Gluck, Dezhi Hong, Aslak Johansen, Jason Koh, Joern Ploennigs, Yuvraj Agarwal, Mario Bergés, David Culler, Rajesh K. Gupta, Mikkel Baun Kjærgaard, Mani Srivastava, and Kamin Whitehouse. 2018. Brick : Metadata schema for portable smart building applications. *Applied Energy* 226 (2018), 1273–1292. https://doi.org/10.1016/j.apenergy.2018.02.091

[2] Mathias Bonduel, Jyrki Oraskari, Pieter Pauwels, Maarten Vergauwen, and Ralf Klein. 2018. The IFC to Linked Building Data Converter - Current Status. , 10 pages. http://urn.fi/URN:NBN:fi:aalto-201810165365

[3] Zoé Chevallier, Béatrice Finance, and Benjamin Cohen Boulakia. 2020. A reference architecture for smart building digital twin. https://hal.archives-ouvertes.fr/hal-02876610

[4] Gabe Fierro, Anand Krishnan Prakash, Cory Mosiman, Marco Pritoni, Paul Raftery, Michael Wetter, and David E. Culler. 2020. Shepherding Metadata Through the Building Lifecycle. In *Proceedings of the 7th ACM International Conference on Systems for Energy-Efficient Buildings, Cities, and Transportation* (Virtual Event, Japan) *(BuildSys '20)*. Association for Computing Machinery, New York, NY, USA, 70–79. https://doi.org/10.1145/3408308.3427627

[5] Michael Grieves. 2014. Digital twin: manufacturing excellence through virtual factory replication. *White paper* 1 (2014), 1–7.

[6] Georgios N. Lilis, Georgios I. Giannakis, and Dimitrios V. Rovas. 2017. Automatic generation of second-level space boundary topology from IFC geometry inputs. *Automation in Construction* 76 (2017), 108–124. https://doi.org/10.1016/j.autcon.2016.08.044

[7] P. Pauwels and W. Terkaj. 2016. EXPRESS to OWL for construction industry: Towards a recommendable and usable ifcOWL ontology. *Automation in Construction* 63 (2016), 100–133.

[8] Pieter Pauwels, Sijie Zhang, and Yong-Cheol Lee. 2017. Semantic web technologies in AEC industry : a literature review. *AUTOMATION IN CONSTRUCTION* 73 (2017), 145–165. http://dx.doi.org/10.1016/j.autcon.2016.10.003

[9] Marco Pritoni, Drew Paine, Gabriel Fierro, Cory Mosiman, Michael Poplawski, Avijit Saha, Joel Bender, and Jessica Granderson. 2021. Metadata Schemas and Ontologies for Building Energy Applications: A Critical Review and Use Case Analysis. *Energies* 14, 7 (2021), 1–32. https://doi.org/10.3390/en14072024

[10] Caroline Quinn and J.J. McArthur. 2021. A case study comparing the completeness and expressiveness of two industry recognized ontologies. *Advanced Engineering Informatics* 47 (2021), 101233. https://doi.org/10.1016/j.aei.2020.101233

[11] Mads Holten Rasmussen, Maxime Lefrançois, Georg Ferdinand Schneider, and Pieter Pauwels. 2021. BOT: The building topology ontology of the W3C linked building data group. *Semantic Web – Interoperability, Usability, Applicability* 12, 1 (2021), 143–161. https://doi.org/10.3233/SW-200385 Publisher: IOS Press.

[12] G. Schneider. 2019. Automated Ontology Matching in the Architecture , Engineering and Construction Domain-A Case Study.