

AN ONTOLOGY-BASED REASONING FRAMEWORK

Towards Multi-level and Data-efficient Building Material Stock Modelling

WANYU PEI¹, SHUYAN XIONG², GUILLAUME HABERT³ and
RUDI STOUFFS⁴

^{1,4} *National University of Singapore; Singapore-ETH Center; ^{2,3}ETH
Zurich*

¹ *peiwanyu@u.nus.edu, 0000-0003-2822-6418*

² *xiong@ibi.baug.ethz.ch, 0000-0002-4563-730X*

³ *habert@ibi.baug.ethz.ch, 0000-0003-3533-7896*

⁴ *stouffs@nus.edu.sg, 0000-0002-4200-5833*

Abstract. The materials stored in existing urban buildings represent a significant share of globally accumulated resources, the composition and quantity of which should be tracked for management and reuse purposes. Due to the coarse-grained nature of building data at the city level, the description of building material stock (BMS) is usually limited to the material intensity (MI) level of several key materials, omitting the component-level analysis of construction elements, and building devices. Hence, a flexible and compatible modelling framework is needed for BMS modelling to adopt different levels of detailed building data. This study proposes an ontology-based framework, which sets the characteristics of available building data as context and makes reasoning for a feasible modelling level. An ontology is developed to capture context knowledge and define the BMS concepts and their properties. A reasoning algorithm is designed to query and categorise building instances with the same attributes into an archetype, to integrate their various granularity of property data, and to calculate the material stocks at appropriate levels. Some Singapore buildings are used for ontology instantiation and explanation. This framework is anticipated to be a new paradigm for multi-level BMS modelling and contribute strategies for urban circularity design.

Keywords. Circular City, Building Material Stock, Domain Ontology, Multi-level Modelling, Missing-data Imputation.

1. Introduction

The energy-intensive production of building materials, such as cement, steel, glass, etc., causes a heavy environmental load during urban expansion. Exploring the potential of material recycling and reuse from the urban building stock becomes an urgent problem for developing a resilient urban environment. Under the circular

economy, quantifying and tracking materials stored in existing buildings to model BMS is essential for mining urban construction resources and preparing them for reuse. Urban-level BMS modelling widely uses the “bottom-up” approach, which allows for a more granular analysis of BMS by analysing individual buildings and their components and indicating the precise locations where materials are utilised. Existing BMS modelling mainly starts from data collection. However, due to the wide variety and quantity of city buildings, applying “bottom-up” to describe the total BMS in countries/cities that face building material data limitations is challenging.

At the urban level, the building data required for BMS modelling has some particularity. First, the data required covers a wide range and types, including various geometric (e.g., building height, footprint area, components thickness, etc.) and attribute (e.g., building age, function, lifespan, material intensity, etc.) “features” of buildings, components, and materials. Besides, the data are heterogeneous, with different formats, semantics, and sources in the temporal and spatial dimensions. Hence, collecting and processing data to form a relatively complete dataset before conducting analysis required much time and labour. Third, data of various grains available needs to be evaluated and processed manually by researchers and the level (in MI or component levels) for modelling is identified (Fig.1). At the MI level, the total material mass can be calculated by multiplying the building’s Gross Floor Area (GFA) or Gross Volume (GV) by an appropriate MI in kg/m^2 or kg/m^3 . The MI coefficients vary depending on the building’s age, function, and structural system. At the component level, the mass of the materials stored can be calculated by multiplying the volume (m^3) of various components and material density (kg/m^3).

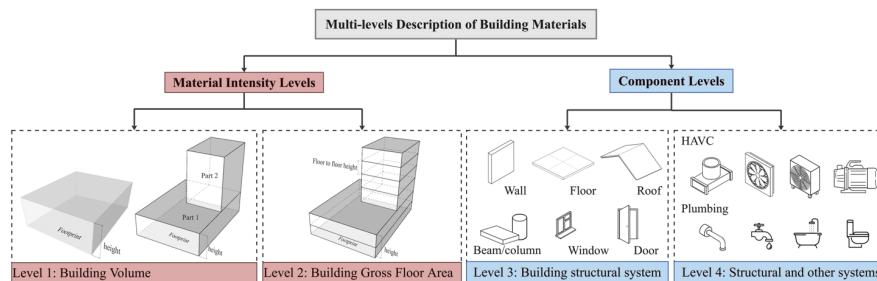


Figure 1. Description of BMS at different levels of detail.

The component-level BMS analysis aids in a more efficient return of materials into the resource loop. It allows for distinguishing between the recovery paths of materials retrieved from various components. Also, even though materials stored in other building systems are limited, their recycling capabilities can be commendable. For instance, copper, aluminium, and wood from wires and pipes can be recycled, making them deserving attention (Yang et al., 2022). However, because of the lack of corresponding detailed data, the materials in components of such building systems, such as HVAC, plumbing, electricity systems, etc., have always been neglected in existing research. It is challenging to obtain “features” data of all systems’ components and materials of one building in a short period. Researchers often have to rely on various assumptions due to missing information, particularly concerning data such as

MI that can mostly be calculated from a previous demolition project. These assumptions cover multiple data types, making them difficult to track and interpret when expanding the material description results to the urban level. Because the link between raw data and domain knowledge is hidden in the implementation, the BSM and its analysis process can be seen as a “black box” with little interpretability. Most existing research also concentrates solely on modelling BMS at a singular level (material or component) and one kind of building system. It is difficult to form a complete dataset that includes “features” of all building systems due to the inconsistent availability of varying buildings, which results in some data being abandoned and not being used efficiently. The added value of the data itself is limited if it is not translated into knowledge to support BMS modelling across levels.

It has been recognised that the “features” of various buildings, systems, and components are highly interconnected when modelling material stock at different levels of detail with corresponding calculation rules (Yuan et al, 2023). For instance, the buildings with same function and completed in the same period in a city should have the same structural system and material properties. These buildings can quantify the mass of their material stock using the same indicator value (such as MI). Suppose that the relationship between these “features” can be defined based on expert knowledge and applying a set of reasoning rules; the multi-level BMS modelling, data collection, and processing can then be conceived as a holistic system. In knowledge representation, an ontology refers to a formal and explicit specification of concepts within a domain and the relationships between these concepts. In recent years, ontologies have demonstrated significant potential. It has been widely used in the construction industry, including for data integration, domain model development, and application ontology creation (Zhou et al., 2016). From domain knowledge, it can capture and organise the complex elements’ relationships and underlying logic within “multi-level BMS” models.

Hence, this research proposes a new paradigm for “multi-level BMS” modelling (the domain of this study) in a data-efficient way. This new paradigm introduces ontology as a foundational structure to organise the domain’s information, knowledge, and calculation rules. The goals of this paradigm are: 1) Automatically identify buildings with potentially similar materials stored according to “building attributes”, such as building age and function; 2) Supplement missing data by incorporating newly added building instances if their material properties are identified as similar to previously input ones; 3) Determine the applicable level for modelling BMS and execute corresponding calculation rules according to available data.

This paper mainly includes the following five parts. Section 2 reviews the related research to introduce the background of this study. Section 3 proposed the method of ontology and reasoning function development and tested the ontology-based reasoning paradigm suggested using several building instances with different-grained data. The results are summarised and discussed in section 4. In section 5, the conclusion of this research is provided.

2. Background

BMS modelling is one of the critical domains of sustainable city development, as it involves intricate variables affecting the composition, mass, and proportion of

materials in buildings. Its analysis relies heavily not only on information technology (data and information) but also on knowledge and human wisdom. To formalise and share complex knowledge within the domains, ontologies are valuable, yet their application in solving urban analysis problems is relatively new. To solve specific urban questions, ontologies have been proposed to support the definition of the concepts, their properties, and the relationships of domain knowledge from two aspects: 1) knowledge management: representing knowledge in a standard way, and 2) automated data reasoning: reasoning about knowledge in a machine-interpretable way. For instance, a domain ontology was used for urban building renovation (Daneshfar et al., 2022) and built cultural heritage conservation (Zalamea et al., 2016), including knowledge about how urban-related features impact building renovation and conservation activities. Allan et al. (2021) enhanced an Urban Energy Simulation (UES) ontology to identify objects, classes, and properties for defining the UBEM archetypes based on data reasoning. An ontology can be coupled with algorithms or structural knowledge graphs to form a semantic model. Some researchers developed applications to operate on top of the ontology. To our knowledge, no ontology has been proposed for modelling multi-level BMS. However, integrating an ontology into BMS modelling shows promising potential for 1) supporting “multi-level” modelling by representing all related domain knowledge and calculation rules; and 2) developing an application algorithm for “data efficient” BMS modelling by processing data at different levels of granularity automatically, based on the reasoning of ontology.

3. Method

The proposed framework, outlined in Fig. 2, involves several key steps. Firstly, domain knowledge, calculation rules, concepts, and relationships about BMS descriptions are extracted to create a domain ontology. Secondly, building data in diverse formats within the data source layer is semantically linked to the domain ontology using a set of mapping assertions. In this way, the combination of ontology and mapping, known as an OBDA specification, reveals the underlying data source as a virtual RDF knowledge graph, enabling access via SPARQL during query execution. Thirdly, a reasoning algorithm is developed to query information from the virtual RDF graph, enabling analysis and filling the data gap based on building features. Finally, this framework selects model levels and performs BMS modelling automatically, reasoning from available data. This section will comprehensively introduce the detailed steps involved in developing this framework.

3.1. ONTOLOGY DESIGN FOR MULTI-LEVEL BMS MODELING

Various methods for ontology development have been proposed in different domains, and there is no single “correct” way. This paper follows the guidelines and steps introduced by existing research (Darlington & Culley, 2008). The development phases mainly include 1) ontology specification, determining the domain and scope of domain ontology; 2) knowledge acquisition: listing important terms (concepts), relationships between concepts (class hierarchy), and class properties; 3) ontology conceptualisation; and 4) ontology implementation: create instances.

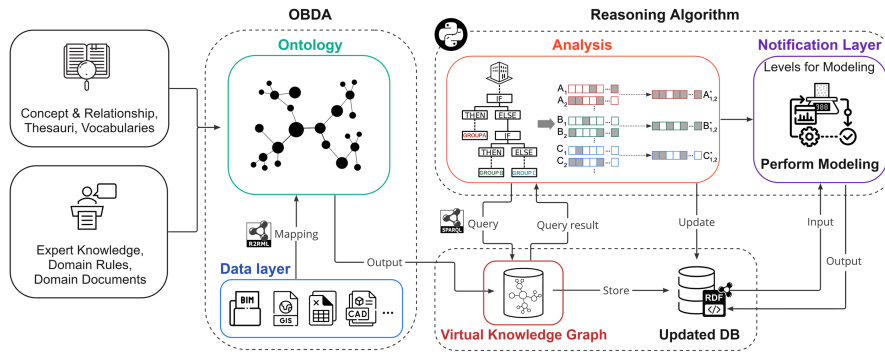


Figure 2. The structure of the ontology-based reasoning framework.

3.1.1. Ontology Specification

Ontology specification is the first step of ontology development by answering questions regarding purpose, scope, intended end-users, and intended use of the ontology. Tab. 1 summarises the information on ontology specification in this study.

Table 1 Information related to ontology specification.

Questions	Answers
Q1: What is the domain that the ontology will cover?	Urban building material stock modelling.
Q2: For what we are going to use the ontology?	Facilitate urban BMS modelling at a corresponding level based on available data granularity.
Q3: For what types of questions, the information in the ontology should provide answers?	Answer queries related to material stock quantification rules aligned with data scenarios to optimise information utilisation.
Q4: Who will use and maintain the ontology?	Decision-makers; Urban; Researchers.

3.1.2. Knowledge Acquisition

This study first conducted a systematic literature review of 99 papers (recent ten years) on urban BMS modelling, which is the basis for knowledge acquisition. Knowledge acquisition aims to define a list of related concepts and terminologies. This study refers to the “*object*” and “*process*” concepts of ontology development and composes the BMS quantification as a set of “*processes*”. The “*resources*” artefact set includes the material stock quantification-related objects, which are inputs of the process part for mass/volume calculation. The “*actors*” constitute the researchers, urban designers, and related government departments involved in identifying the calculation rules and processing quantification. The calculation results are outputted as data items and generate material mass. The concepts and terminologies are extracted from domains following our designed ontology structure (Fig. 3) and expressed in Web Ontology Language (OWL). This multi-level BMS stock ontology is compacted as “*bms:*”.

For the “*object*” concept, we investigated the materials in the buildings, various

building systems, and corresponding components to categorise these entities to account for the mass (volume) of material stock. This study reused the *bot:Building*, *bot:Storey*, *bot:Zone*, *bot:Space*, and *bot:Element* in Building Topology Ontology (BOT) (Rasmussen et al., 2021) to represent the topological structure of buildings. The *mat:Material* and related entities in Building Material Ontology (Fenz et al., 2021) are reused to describe materials and their properties in the components. The “*process*” in this ontology is mainly related to calculating material mass/volume based on the properties of the building, components, and material. For other building systems’ components (devices) we refer to the SAREF4BLDG extension ontology (Poveda-Villalón & Garcia-Castro, 2018). For the design of the “*process*” part, we refer to the concepts related to data quantitative method proposed by the Ontology of Units of Measure and Related Concepts (OM) (Rijgersberg et al., 2013).

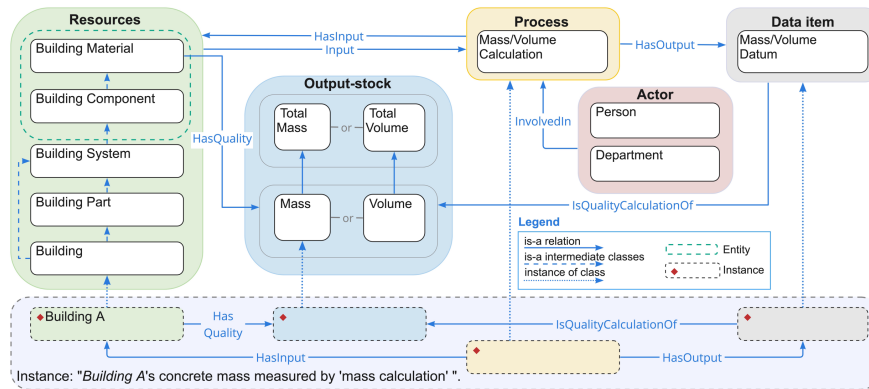


Figure 3. Multi-level BMS ontology structure.

3.1.3. Ontology Conceptualisation and Implementation

To organise the related concepts and terminologies and develop the conceptual ontology, some attributes and properties are associated with “*objects*” and “*processes*” (Fig. 4). The ontology conceptualisation is achieved using *diagrams.net* (JGraph, 2021), which allows the building of ontologies in a visualised and collaboratively modified way.

Fig.4 shows the “*objects*” part in the proposed ontology, which includes the concepts and concept relationships related to “MI-level” and “component-level” BMS modelling. For the “MI level” calculation, the *bms:Archetype* is defined to populate each building instance’s parameters, such as MI, with matching categories. We adopted *bms:SubPropertyOf* to define the archetype based on the properties from *bms:Ifcbuilding*. The properties are building use (function), building age, and building location, which are typically archetype categories. To get the MI coefficients of various materials stored in archetype buildings, we will use *bms:IdentifierInDatabase* to link a global MI database in future versions under *bms:MaterialEntity* and drive MI retrieving. The MI databases were developed by Heeren and Fishman (2019), including 300 MI data points (in kilograms per gross floor area) of 30 kinds of materials from 33

studies. The MI coefficients can be obtained from databases based on building properties, including building age, function, and location.

The “component level” ontology describes all construction systems’ elements and building devices, their materials, and coordinating properties. To calculate the mass/volume of material stored in these building elements/devices, this ontology reused the artefacts of the SAREF ontology. The material and component properties, such as component thickness and material density, are grouped by *mat:PropertySet* and measured *mat:Measurement*. In the “process” part, the materials and component properties are inputted for material mass calculation. The implementation of the whole ontology aims to make the conceptual ontology computer-readable, which can be achieved in the *Protégé* tool (Tudorache et al., 2008), an open-source ontology editor.

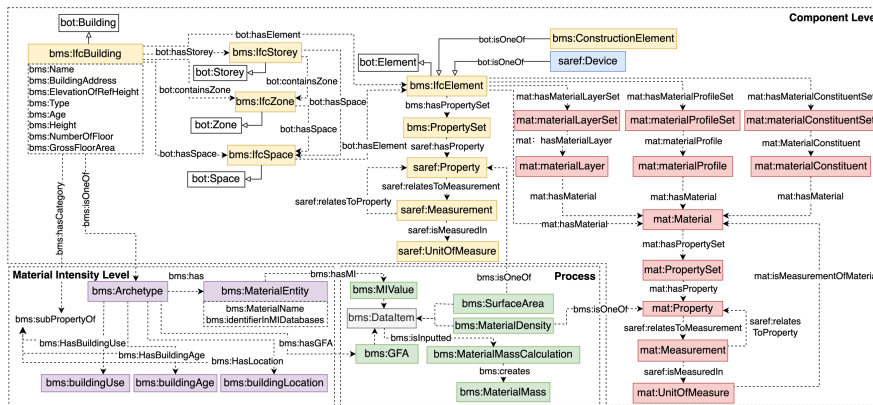


Figure 4. Conceptualisation of BMS Ontology.

3.2. REASONING ALGORITHM DEVELOPMENT

Based on the ontology developed, this study designed some algorithms (Fig.5) to query data through the ontology from data sources and conduct multi-level BMS modelling. When inputting a building instance, the first algorithm queries building attributes data (age, function, etc.) and compares it with the attributes of existing buildings stored in the database using Algorithm 2 marching rules. This step divides building instances with the same features into several groups and uses a virtual building’s (archetype) material properties to stand for the properties of all the buildings in this group. Algorithm 3 integrates the different granularity-level property data in various building instances and forms a complete property set for archetypes. For example, if the newly added building instance has component property data that the existing property set does not have, then the new data is added, and the vector updated. Finally, Algorithm 1 executes BMS modelling, adapting corresponding material mass/volume calculation rules for the final property set.

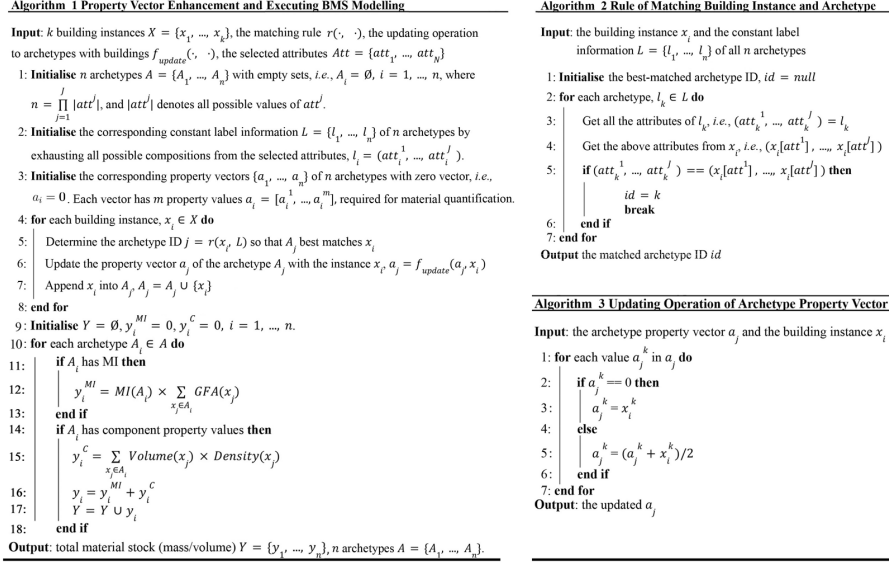


Figure 5. The pseudocode of the reasoning algorithms.

3.3. CASE STUDY AND FRAMEWORK INSTANTIATION

This study uses four different building projects in Singapore (as instances x_1, x_2, x_3 and x_4) to conduct a case study and examine the conceptual framework proposed. Tab.2 summarises the property data of the four building instances collected according to their published information and BIM models. First, x_1 , a non-residential building built during “1990-2000”, is inputted in the framework. Second, Algorithm 1 queries the attributes, including the year of construction and building function of x_2 and matches with x_1 , noticing x_1 and x_2 are both non-residential buildings completed during the same period. Hence, the algorithm matches them into archetype A_1 , which can share their material property vector. Then, algorithm 3 conducts a property update operation for archetype A_1 . For instance, the updated MI will be the average value of the two instances (0.145), while the window-wall ratio (0.32) and window thickness (2.2 mm) will be added based on property data of x_2 . The result forms an updated database of archetype A_1 which can be used to fill the missing values for all the building instances in the same group as A_1 , and calculate concrete stock at the “MI level” and glass stock at the “window component level”.

Similarly, x_3 and x_4 are residential buildings completed in different periods compared with x_1 and x_2 . In this way, x_3 and x_4 will be identified as another archetype A_2 ; their material property data will contribute together to the properties vector of archetype A_2 , supporting to quantify concrete at MI level, using the MI from x_4 , as well as steel in the air duct. If another instance with “window-wall ratio” data could be available as inputs and it is identified as belonging to A_2 , the glass stored in the windows of buildings in A_2 can then be calculated.

Table 2 The information of four building instances.

Instances	Country	Year	Function	Footprint Area (m ²)	GFA (m ²)	Height (m)	MI level	Component level (parts of elements)				
							MI-Concrete (m ³ /m ²)	Window-Wall ratio	Window thickness (mm)	Material Density-Glass (kg/m ³)	Air Duct Thickness (mm)	Density Air Duct-Steel (kg/m ³)
x ₁	SG	1994	Office	1202.84	5216.0	14.3	0.13	-	-	2510.0	-	-
x ₂	SG	1995	Academic	1202.84	2181.6	13.0	0.16	0.32	2.2	-	-	-
x ₃	SG	2013	Residential	567.35	6240.8	34.0	-	-	-	2531.0	1.2	7850.0
x ₄	SG	2016	Residential	1692.10	42302.4	80.0	0.38	-	3.0	-	-	-

4. Result and Discussion

The framework designed by this study provides a new paradigm for urban BMS quantification. However, at the current stage, this research has some limitations. First, as the ontology design depends on expertise, it is essential to carefully document and validate the framework to ensure its effectiveness and applicability. While reusing parts of existing ontologies, this study introduces new artefacts based on specific needs. At this stage, the ontology proposed in this study has yet to undergo evaluation by relevant domain experts. Subsequent efforts will involve organising workshops and inviting industry experts and scholars to evaluate the ontology and validate its design. Additionally, although this paper describes and presents the conceptual design of the framework, it has not yet been made computer readable. Also, the current ontology was instantiated by analysing some real building projects, but the data volume is small.

The subsequent research will complete the implementation of the ontology and algorithms, including mapping data sources in OBDA, generating a virtual knowledge graph, and establishing links with structural databases. Once the ontology implementation is completed, we will gather more practical cases to explain and demonstrate this framework. Also, we will optimise the current “updating operation algorithm” by introducing a collaborative filtering technique. With increasing instances, collaborative filtering can assist in comparing the attributes of building instances within an archetype group and perform data integration, ultimately achieving calculation in a “building-by-building” way. This process could retain the diversity of buildings during BMS modelling. Simultaneously, collaborative filtering can facilitate the flow of information and supplement data between different archetypes, thus avoiding excessive lack of material attribute data of a certain archetype group.

5. Conclusion

In conclusion, this research proposes an innovative framework integrating ontology and reasoning algorithms for urban BMS modelling. Driven by an ontology model, this framework can support BMS modelling at various levels of detail and fully use currently available building data without requiring the preparation of a complete dataset. This framework addresses building material data limitations some regions face, which restricts the modelling and analysis of building stocks at the urban level. With the multi-level modelling capability of this framework, building data, including component-level data, can be fully utilised, enabling the statistical determination of building stocks within components. The potential for recyclability of building components can be better described, providing essential information and strategies for designing recyclable buildings. Therefore, this proposed framework offers a more flexible approach to choosing the level of detail in data utilisation and BMS modelling.

Acknowledgements

The research was conducted partly at the Future Cities Lab Global at the Singapore-ETH Centre, which was established collaboratively between ETH Zurich and the National Research Foundation Singapore. This research is supported by the National Research Foundation Singapore (NRF) under its Campus for Research Excellence and Technological Enterprise (CREATE) programme. The financial support the China Scholarship Council (CSC) provided to the first author Wanyu Pei is acknowledged.

References

- Allan, J., Fierz, L., Bollinger, A., & Orehounig, K. (2021). Linked data ontology for urban scale building energy simulation. *eSIM 2020 Conference Proceedings*, (pp.8). IBPSA.
- Daneshfar, M., Hartmann, T., & Rabe, J. (2022). An ontology to represent geospatial data to support building renovation. *Advanced Engineering Informatics*, 52, 101591. <https://doi.org/10.1016/j.aei.2022.101591>.
- Darlington, M. J., & Culley, S. J. (2008). Investigating ontology development for engineering design support. *Advanced Engineering Informatics*, 22(1), 112-134. <https://doi.org/10.1016/j.aei.2007.04.001>.
- Fenz, S., Bergmayr, J., Plattner, N., Chávez-Feria, S., Poveda-Villalón, M., & Giannakis, G. (2021). Integration of building material databases for IFC-based building performance analysis. In ISARC. *Proceedings of the International Symposium on Automation and Robotics in Construction* (Vol. 38, pp. 182-189). IAARC Publications.
- Heeren, N., & Fishman, T. (2019). A database seed for a community-driven material intensity research platform. *Scientific data*, 6(1), 23. <https://doi.org/10.1038/s41597-019-0021-x>.
- JGraph. (2021). Diagrams.net (v22.1.8) [Computer software]. <https://app.diagrams.net/>.
- Poveda-Villalón, M., & Garcia-Castro, R. (2018). Extending the SAREF ontology for building devices and topology. *Proceedings of the 6th Linked Data in Architecture and Construction Workshop (LDAC 2018)*, Vol. CEUR-WS, 2159, 16-23.
- Rasmussen, M. H., Lefrançois, M., Schneider, G. F., & Pauwels, P. (2021). BOT: The building topology ontology of the W3C linked building data group. *Semantic Web*, 12(1), 143-161. <https://doi.org/10.3233/SW-200385>.
- Rijgersberg, H., Van Assem, M., & Top, J. (2013). Ontology of units of measure and related concepts. *Semantic Web*, 4(1), 3-13. <https://dl.acm.org/doi/10.5555/2595053.2595055>.
- Tudorache, T., Noy, N. F., Tu, S., & Musen, M. A. (2008). Supporting collaborative ontology development in Protégé. In *The Semantic Web-ISWC 2008: 7th International Semantic Web Conference, ISWC 2008, Karlsruhe, Germany, October 26-30, 2008*. Proceedings 7 (pp. 17-32). Springer Berlin Heidelberg.
- Yang, X., Hu, M., Tukker, A., Zhang, C., Huo, T., & Steubing, B. (2022). A bottom-up dynamic building stock model for residential energy transition: A case study for the Netherlands. *Applied Energy*, 306, 118060. <https://doi.org/10.1016/j.apenergy.2021.118060>
- Yuan, L., Lu, W., Xue, F., & Li, M. (2023). Building feature-based machine learning regression to quantify urban material stocks: A Hong Kong study. *Journal of Industrial Ecology*, 27(1), 336-349. <https://doi-org.libproxy1.nus.edu.sg/10.1111/jiec.13348>.
- Zalamea Patiño, O., Van Orshoven, J., & Steenberghen, T. (2016). From a CityGML to an ontology-based approach to support preventive conservation of built cultural heritage. In *Proceedings of the 19th AGILE International Conference on Geographic Information Science*. Springer; Belgium.
- Zhou, Z., Goh, Y. M., & Shen, L. (2016). Overview and analysis of ontology studies supporting the development of the construction industry. *Journal of Computing in Civil Engineering*, 30(6), 04016026. [https://doi.org/10.1061/\(ASCE\)CP.1943-5487.000059](https://doi.org/10.1061/(ASCE)CP.1943-5487.000059).