SEMANTIC WEB BASED INTEGRATION BETWEEN BIM COST AND GEOMETRIC DOMAINS

Jacopo Cassandro & Claudio Mirarchi & Alberto Pavan
Polytechnic of Milan, Italy

Andrea Zamborlini & Carlo Zanchetta
University of Padua, Italy

ABSTRACT: *In the architecture, engineering, construction, and facilities management (AEC/FM) industry methodologies are needed to ensure the interoperability of data and effective management of information from different sources. Integration of the cost domain and cost estimation within the Building Information Model (BIM) in the AEC/FM sector is still an unresolved problem and one of the most critical tasks due to the lack of a standardised cost domain, especially in the tendering phase.*

To ensure interoperability between cost data and geometric data, this research aims to address this gap by analyzing methods of converting cost data into Linked Building Data, thereby defining a cost domain in the Semantic Web, by collecting them into a graph database. This allows for structuring a cost domain, translating an IFC based structure previously developed by the research group, visualizing it using a graph system, and connecting it to the BIM geometric domain. Furthermore, it is possible to extend the cost ontology previously identified in the IFC model and facilitate the queries and analysis of cost data currently fragmented and based on unstructured data.

The results show how Semantic Web technology can be used to improve data interoperability, develop a cost ontology, and join both cost data and BIM models.

KEYWORDS: *Semantic Web, Linked Building Data, IfcOWL, cost ontology, IFC, RDF, graph system.*

1. INTRODUCTION

The construction process is complex, dynamic and requires numerous interactions between the different actors involved. The sharing of different information, including information on the amount of physical components of the building, the planning plan, and the consumption of resources and costs is essential for accurate time and cost management.

Nowadays, in the Architecture, Engineering, and Construction (AEC) industry, the standard format used for information exchange is the Industry Foundation Classes (IFC), a neutral and open ISO standard by buildingSMART. Currently, the most recent IFC scheme is IFC4 ADD2 and contains about 1200 classes. BIM software developers can implement an exporter to convert respectively their native BIM format to the neutral IFC format.

Despite the clear advantages of data interoperability, which would otherwise only be feasible in proprietary BIM software, the IFC data model still has some limitations (Pauwels & Terkaj, 2016). These include, for example, the impossibility of extending it in a user-friendly way, the difficulty of developing applications with this template, due to the complexity of the schema expressed in EXPRESS format (Krijnen & Beetz, 2018), or the impossibility of relating data between the IFC file and other cloud historicized files. The IFC scheme is used as an interoperable format for sharing information, with the aim of being a supplier-independent exchange format, but not a fully integrated and comprehensive description of a project. This leads to an immense untapped potential of data in the AEC industry (Krijnen & Beetz, 2018). In comparison, the Semantic Web (SW) uses the Resource Description Framework (RDF) triple schema to store data and ontologies to enhance the semantic structure to make information machine-readable (Berners-Lee et al., 2001). It may also contain multiple ontologies, a formal and explicit specification of a shared conceptualisation (Gruber, 1993), covering specific areas not included in the IFC scheme (Rasmussen et al., 2017) and enabling the visualization and improvement of the interoperability of the IFC information model.

In order to overcome these limitations, this research is focused on the potential of SW. This research focuses on the conversion of the architecture of cost items, assumed and implemented in the IFC data model in a previous study carried out by the research group (Cassandro et al., 2023), to RDF using the emerging Linked Building Data

Referee List (DOI: 10.36253/fup_referee_list)
FUP Best Practice in Scholarly Publishing (DOI 10.36253/fup_best_practice)
Jacopo Cassandro, Claudio Mirarchi, Alberto Pavan, Andrea Zamborlini, Carlo Zanchetta, *Semantic Web Based Integration Between BIM Cost and Geometric Domains*, pp. 813-823, © 2023 Author(s), CC BY NC 4.0, DOI 10.36253/979-12-215-0289-3.81

(LBD) modular ontologies as proposed by the W3C LBD CG (Bonduel et al., 2018). To validate what has been done in the previous study, the IFC-to-LBD converter presented by Bonduel et al. (2018) has been used. Information from IFC building models is extracted and transformed into Abox RDF graphs suited for usage in Linked Data applications.

The graph system will contain the relevant information of both the geometric model and the cost architecture with the related properties. Data translation within the SW will allow to query the model, associate the two different domains studied (geometric and cost) within a single environment, and view cost data and related architectures.

This paper develops following these steps. First, a detailed analysis of the literature to deepen the themes on the SW. Secondly, the identification of the conversion tools available to date for the transition from IFC to RDF and the subsequent validation and analysis of the results within a graph database. Finally, the conclusions and future works will be set out.

2. BACKGROUND

2.1 Literature review

The first research on the application of the SW in the AEC Industries dates back to the early 2000s; since then, their use has spread to more and more areas of the industry, producing interesting results in terms of number of publications and significance of results. Beetz et al. (2015) and Pauwels et al. (2017) analysed the reasons for the spread of SW in the AEC industry, identifying three main reasons: (1) Interoperability, (2) Liking across domains and (3) Logical inference and proofs.

The possibility of improving the Interoperability (1) relies on the SW structure which provides a way to store information in a computer-understandable manner, making possible the comprehension of the information involved in the process both by a human being and a machine.

The linking across domain (2) relies on the opportunity to create a unique web of linked data with information from all the different areas of the AEC process, (e.g. GIS, costs, energy, facility management, and so on).

The third motivation is the logical inference and proofs (3), which relies on the OWL language used for the semantic meaning. Correct use of the language offers the opportunity to infer more information from the original input, improving the web and allowing us to do more complex queries.

As we previously said, the usage of SW in the AEC involves different areas of the industry; a review of some of the most significant utilisations is reported below.

H. Abanda et al. (2011) proposed a SW based decision support system, which helps the government to speed up and automate the bureaucratic processes. This approach shows how SW could be a powerful ally for the PA to manage all the applications for a licence. In Karan et al. (2015) the SW is used to overcome heterogeneity problems, which came from the traditional methods of heterogeneous data sharing, generating IFC from semantic web query results. The IFC structure lends itself to a translation in linked data, which is why Zhao et al. (2020) suggested a method for IFC data merging based on miming the IFC structure with nodes and edges. The IFC graphs are merged and then restored in the starting files, implementing the information. Merging information through SW can be done not only using the same type of input, like different IFC files; in fact, Malinverni et al. (2022) used an approach that merges information from GIS and BIM models, producing an enriched model that has many benefits for the entire project life cycle. Also, the safety issues are affected using SW, and Zhang et al. (2015) developed a prototype of a new approach to organize, store, and re-use construction safety knowledge to produce a framework that supports automated job hazard analysis in BIM. Also, the safety issues are affected using SW, and Zhang et al. (2015) demonstrated a prototype of a new approach to organize, store, and re-use construction safety knowledge to produce a framework that supports automated job hazard analysis in BIM.

The damage and degradation analysis can also be positively affected by linked data and an interesting example is reported by Jung et al. (2021); they discussed an automatic approach to infer the causes of concrete cracks starting from information about pattern, location, and penetration status. The possibility to express human thinking and make a logical inference by machines thanks to ontologies can remove the issue of complex qualitative analysis during the process of identifying the cause of a crack.

The most meaningful domain analysed for this paper is 5D planning, for which can be found different approaches.

F. H. Abanda et al. (2011) developed an ontology-based technology in modelling information about labour costs that aims to facilitate decision-making among building developers in Cameroon; Vakaj et al. (2023) proposed a new domain ontology called Offsite Housing Ontology to support cost estimation about resources, products, and production processes. A further reference for cost estimation is Fürstenberg et al. (2021); they studied how semantic web technology can support BIM-based automated cost estimation and the related challenges, focused on Norwegian road projects.

The need to find ways of translating IFC into ontologies led to different approaches. The first interesting attempts can be found in Beetz et al. (2005), where two different approaches convert the EXPRESS schema of an IFC into an ontology in OWL notation: one using an intermediate step with an XSLT file between the XML file and the OWL, and a second which derives the OWL notation directly from the original EXPRESS schema format of the IFC. After that the approach evolved, leading to Beetz et al. (2009) which presented a semi-automatic way of lifting EXPRESS schemas into ontologies. Pauwels et al. (2015) analysed the correct ways to translate EXPRESS language into OWL. Hoang and Torma (2015) present the IFC2LD converter, a Java application with a Web interface, for converting IFC schemas into OWL2 ontologies and IFC data into RDF graphs aligned with the ontologies. Pauwels and Deursen (2015) present their online RDF to IFC conversion service, which converts an IFC into RDF triples. Continuing chronologically, Ismail et al. (2017) show a workflow for the automatic transformation of IFC into an object graph database; this method is based on a dynamic EXPRESS parser and a web script console that creates a meta graph inside Neo4j. Bonduel et al. (2018) developed a conversion of IFC to RDF using W3C Linked Building Data modular ontologies. The graphs are structured with three types of ontologies: BOT (building topology), PRODUCT (classification of building elements), and PROPS (building-related properties); the result is a more user-friendly graph than the ifcOWL Abox graphs.

2.2 Cost definition in IFC domains focus on IfcCostItem

Nowadays cost items are associated as attributes to geometric entities. However, to correctly return the analysis and economic evaluation processes, it is necessary to have cost architectures configured as more complex systems than a simple attribute associated with a geometric object. The IFC standard, through the cost class (*IfcCostItem*), offers the possibility of structuring a cost data model.

IfcCostItem is a non-geometric entity, a subclass of *IfcControl*, within IFC. *IfcCostItem* describes a cost or financial value with descriptive information that describes its context (BuildingSMART, 2022). It represents the cost of assets and services, the execution of works by a process, lifecycle cost, cost estimates, budgets, and more in the IFC standard.

IfcCostItem is characterized by its own attributes (PredefinedType, CostQuantites, and CostValue) and others inherited. An *IfcCostItem* has the possibility to instantiate one or more cost values (*IfcCostValue*). Other key features are that every single *IfcCostItem* can be nested to create cost assemblies through the *IfcRelNests* report, can be assigned to an *IfcProduct* through the *IfcRelAssignsToControl* report, may have an associated product through the *IfcRelAssignsToProduct* report or a resource through the *IfcRelAssignsToResource* report.

2.3 IfcOWL and Linked Building data (LBD)

In recent years, the area of cross-domain linking has received increasing attention. This area aims to combine data from various sources with construction data, management of information based on ontology, and analysis of the performance of buildings (Pauwels, Krijnen, et al., 2017). According to W3C, the Web Ontology Language (OWL) is a language designed to represent complex knowledge about objects, relationships between objects, and groups of objects in a way that can be exploited by computers (*IfcOWL - BuildingSMART Technical*, 2023.). BuildingSMART has developed the IfcOWL ontology based on these definitions, providing an OWL representation of the Industry Foundation Classes (IFC) schema, maintaining the same status as the Express schema. OWL concepts (OWL - Semantic Web Standards, 2023) can be used to construct RDF graphs, called OWL ontologies (Pauwels & Terkaj, 2016), enabling easy linking between the building data and material data, cost data, GIS data, and so forth. However, due to the complex structure of the IFC data model, the ifcOWL representation of geometric data is difficult to manage (Pauwels et al., 2017a).

Achieving interoperability between domains is the main purpose of the Linked Building Data (LBD) Community Group in the World Wide Web Consortium (W3C) (*The Linked Building Data Community Group*, 2021). LBD allows for storing construction data sources separately and processing them through digital and computer systems (Curry et al., 2013). This results in a set of data that can be utilized and interconnected. Expandability is a key aspect of AEC where most projects are fragmented, complex, and diverse. Using the LBD different ontologies can be mapped and enhanced each other, facilitating a more comprehensive and integrated approach to handling diverse data sources and formats within the AEC industry.

3. RESEARCH AIM & METHODOLOGY

The research aims to verify the effective possibility of associating the new cost domain with the geometric one through the SW. This will allow to relate different domains within the same environment to improve data sharing and interoperability. In addition, it will be possible to manage cost items no longer as simple attributes attached to a geometric object, but as real cost architectures, more complex, ensuring in the future also the ability to verify and validate the associated data.

The first study to examine the possibility of structuring a cost domain using Industry Foundation Classes (IFC) has already been addressed in Cassandro et al. (2023). In this work, based on the assumptions and limitations of previous research in Cassandro et al. (2023), has been translated the ontology previously developed in the Linked Building Data (LBD) format. This will allow us to relate the cost domain to the existing domains (in the specific case the geometric domain); in fact, SW technology is well-suited to link knowledge stored in different domains (Beetz et al., 2015; Pauwels, Zhang, et al., 2017).

The methodology adopted is characterized by the following steps presented in Figure 1:

1. Study of the State of the Art of the current practices and research connected to SW and graph database;
2. Analysis of IFC entities (*IfcCostItem* identified in the standard to manage the cost information) and how to translate it into LBD;
3. Translation of cost ontology, developed in Cassandro et al. (2023) from IFC to LBD through a tool developed by Bonduel et al. (2018);
4. Information validation in a graph database such as Neo4j;
5. Results of experimental research

A way to represent cost information by using SW was formulated, developed, and validated. This could provide the basis for the information exchange resources among information systems for the more user-friendly query of data and linking different domains. The next sections describe the mechanisms used to implement and test the method.

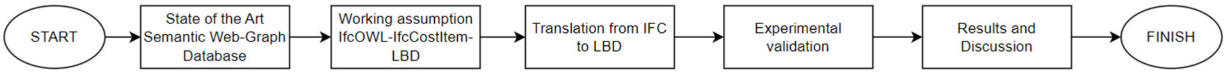
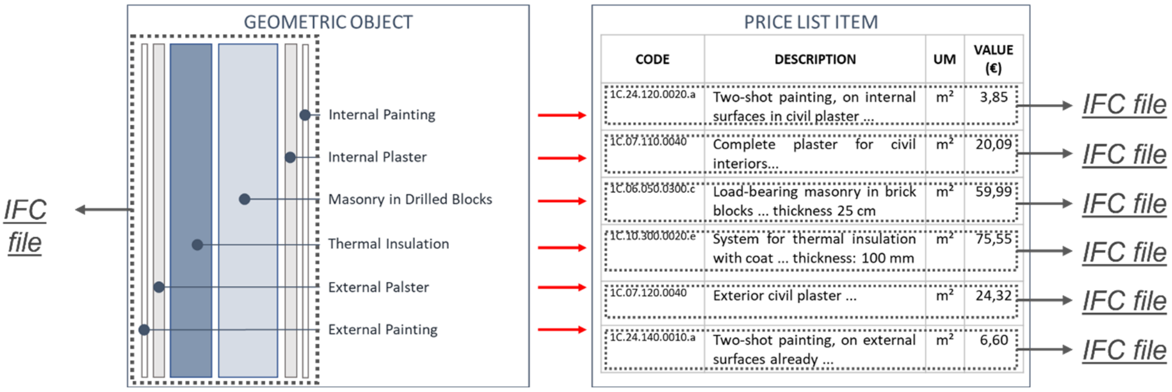


Fig. 1: Research Methodology

4. RESULTS

In this research, the case study is the same as that analysed by Cassandro et al. (2023). This allows to fully understand the differences between the methodologies adopted and to compare the results obtained during the two research. The case study is a wall composed of six layers; each layer corresponds to a different cost item within the regional price list (Lombardy Region) which must therefore be associated with the related geometric object, see Figure 2. These cost items have already been structured within the IFC data model.



CODE	DESCRIPTION	UM	VALUE (€)
1C.24.120.0020.a	Two-shot painting, on internal surfaces in civil plaster ...	m^2	3,85
1C.07.110.0040	Complete plaster for civil interiors...	m^2	20,09
1C.06.050.0300.c	Load-bearing masonry in brick blocks ... thickness 25 cm	m^2	59,99
1C.10.300.0020.e	System for thermal insulation with coat ... thickness: 100 mm	m^2	75,55
1C.07.120.0040	Exterior civil plaster ...	m^2	24,32
1C.24.140.0010.a	Two-shot painting, on external surfaces already ...	m^2	6,60

Fig. 2: Example layers of masonry and relative price items

Then the individual IFC files were converted into RDF format using the tool "IFCtoLBD[1]" developed by Jyrki Oraskari and Mathias Bonduel. This tool allows to convert an IFC file to a Turtle file described using BOT and optionally PRODUCT, PROPS, and GeoSPARQL (Bonduel et al., 2018).

The correct export specifications have been set thanks to Bonduel et al. (2018). For the correct output, it is necessary to activate the PROPS module of the ontology structure, which includes three levels of complexity. For the case study, Level 3 was selected (i.e., the most complete level). Thereby, the Blank node option was activated, which decreased the file size, exported RDF, and improved the readability. The output is a Terse RDF Triple Language (Turtle); a format made to express RDF data. This format uses triples made by subject, predicate, and object to represent information.

At this point the new file in Turtle format is loaded inside a graph database (in this specific case Neo4j is used); in this way, you can view the information and links between these (Figure 3). As visible from Figure 3 every entity is associated with a series of intrinsic attributes of the same one. IFC information is intrinsically interconnected and can naturally be represented by graphs. Figure 3 shows the IFC file for a painting cost item at the top and the corresponding data representation in a graph system at the bottom (nodes and edges). As it is visible, the graph representation is more intuitive in revealing the relationships between instances than the text based IFC.

Translating the IFC data model into a graphical system can lead to a simplified representation of the construction information and its relationships, as well as improving data query.

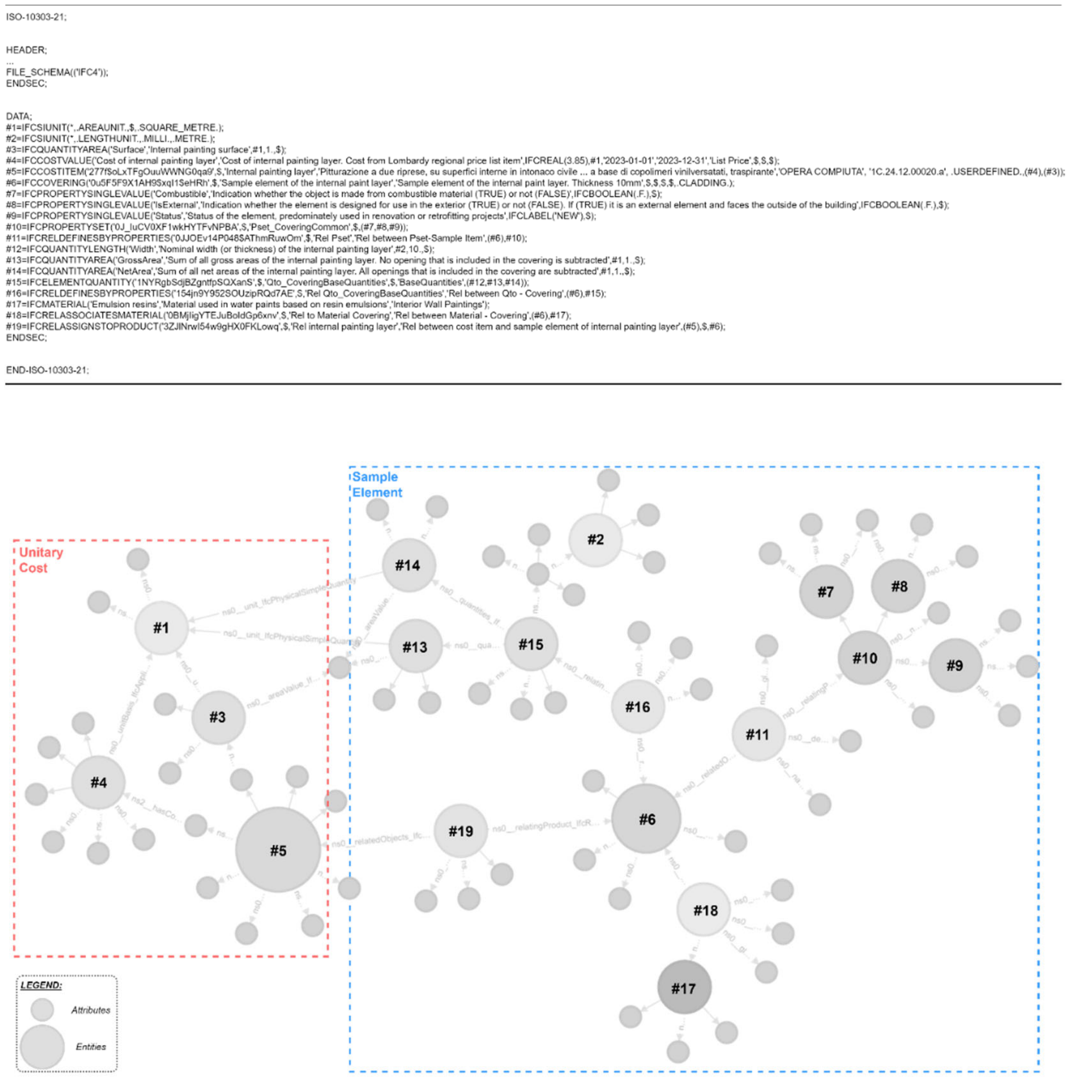
The developed methodology relies on an IFC, which contains both the geometrical information and cost information, inserted into the file using the appropriate classes previously studied in Cassandro et al. (2023). The creation and compilation of *IfcCostItem* classes has been implemented in Python using IfcOpenShell, as shown in Cassandro et al. (2023).

The conversion from IFC to an LBD was carried out using a tool developed by Bonduel et al. (2018). The correct exporter setting has been set after several attempts to get the type of LBD needed. Figure 4 shows all the cost item files, in RDF format, imported into the graph database and related to the geometric object (wall system). After that, the latter was also imported and displayed, visible in Figure 5.

The Neo4j graph database has been used to visualize the data. The Neosemantics plugin (n10s) was used to load RDF data and its associated vocabularies, including OWL, RDFS, SKOS, and others, into Neo4j. This plugin extends Neo4j's capabilities to work with semantic data in RDF format, allowing users to import, store, query, and analyze RDF data within the Neo4j graph database. It facilitates the integration of RDF-based knowledge graphs and linked data into Neo4j, enabling more comprehensive and semantic data modeling and analysis.

---

[1] https://github.com/jyrkioraskari/IFCtoLBD



```
ISO-10303-21;

HEADER;
...
FILE_SCHEMA(('IFC4'));
ENDSEC;

DATA;
#1=IFCSIUNIT(*,.AREAUNIT.,$,.SQUARE_METRE.);
#2=IFCSIUNIT(*,.LENGTHUNIT.,.MILLI.,.METRE.);
#3=IFCQUANTITYAREA('Surface','Internal painting surface',#1,1.,$);
#4=IFCCOSTVALUE('Cost of internal painting layer','Cost of internal painting layer. Cost from Lombardy regional price list item',IFCREAL(3.85),#1,'2023-01-01','2023-12-31','List Price',$,$,$);
#5=IFCCOSTITEM('277f$oLxTFgOuuWWNG0qa9',$,'Internal painting layer','Pitturazione a due riprese, su superfici interne in intonaco civile ... a base di copolimeri vinilversatati, traspirante','OPERA COMPIUTA', '1C.24.12.00020.a', .USERDEFINED.,(#4),(#3));
#6=IFCCOVERING('0u5F5F9X1AH9Sxql1SeHRh',$,'Sample element of the internal paint layer','Sample element of the internal paint layer. Thickness 10mm',$,$,$,$,.CLADDING.);
#7=IFCPROPERTYSINGLEVALUE('Combustible','Indication whether the object is made from combustible material (TRUE) or not (FALSE)',IFCBOOLEAN(.F.),$);
#8=IFCPROPERTYSINGLEVALUE('IsExternal','Indication whether the element is designed for use in the exterior (TRUE) or not (FALSE). If (TRUE) it is an external element and faces the outside of the building',IFCBOOLEAN(.F.),$);
#9=IFCPROPERTYSINGLEVALUE('Status','Status of the element, predominately used in renovation or retrofitting projects',IFCLABEL('NEW'),$);
#10=IFCPROPERTYSET('0J_luCV0XF1wkHYTFvNPBA',$,'Pset_CoveringCommon',$,(#7,#8,#9));
#11=IFCRELDEFINESBYPROPERTIES('0JJOEv14P048$AThmRuwOm',$,'Rel Pset','Rel between Pset-Sample Item',(#6),#10);
#12=IFCQUANTITYLENGTH('Width','Nominal width (or thickness) of the internal painting layer',#2,10.,$);
#13=IFCQUANTITYAREA('GrossArea','Sum of all gross areas of the internal painting layer. No opening that is included in the covering is subtracted',#1,1.,$);
#14=IFCQUANTITYAREA('NetArea','Sum of all net areas of the internal painting layer. All openings that is included in the covering are subtracted',#1,1.,$);
#15=IFCELEMENTQUANTITY('1NYRgbSdjBZgntfpSQXanS',$,'Qto_CoveringBaseQuantities',$,'BaseQuantities',(#12,#13,#14));
#16=IFCRELDEFINESBYPROPERTIES('154jn9Y952SOUzipRQd7AE',$,'Rel Qto_CoveringBaseQuantities','Rel between Qto - Covering',(#6),#15);
#17=IFCMATERIAL('Emulsion resins','Material used in water paints based on resin emulsions','Interior Wall Paintings');
#18=IFCRELASSOCIATESMATERIAL('0BMjligYTEJuBoldGp6xnv',$,'Rel to Material Covering','Rel between Material - Covering',(#6),#17);
#19=IFCRELASSIGNSTOPRODUCT('3ZJlNrwl54w9gHX0FKLowq',$,'Rel internal painting layer','Rel between cost item and sample element of internal painting layer',(#5),$,#6);
ENDSEC;

END-ISO-10303-21;
```

Fig. 3: Comparison between cost items in IFC format and RDF format

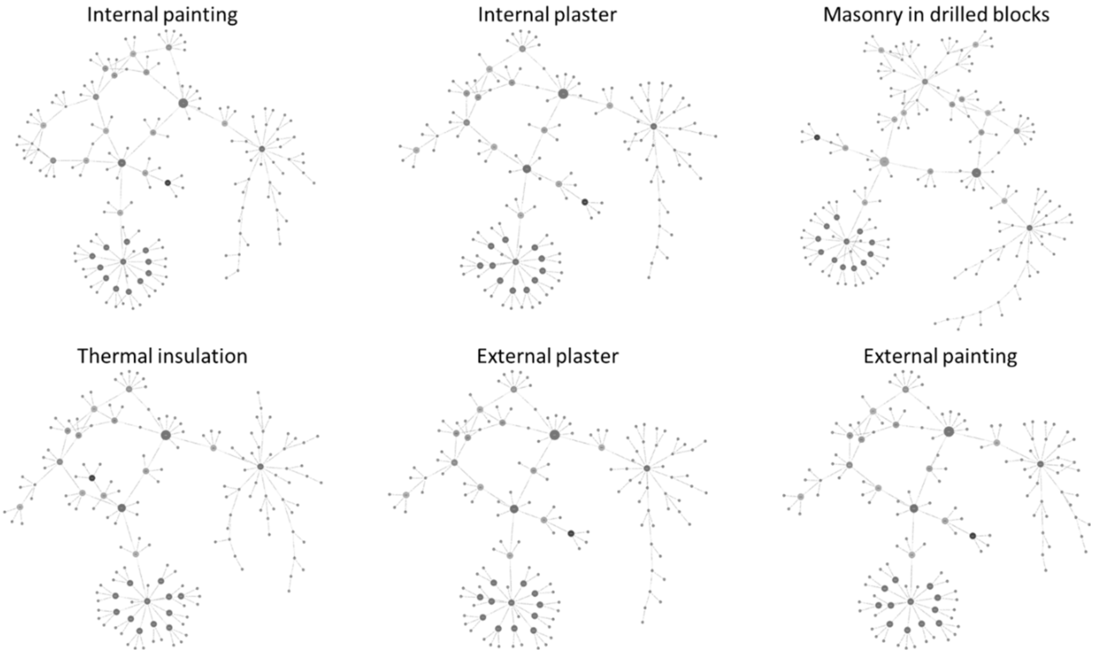


Fig. 4: Representation in the graph system of the six different cost items starting from IFC files

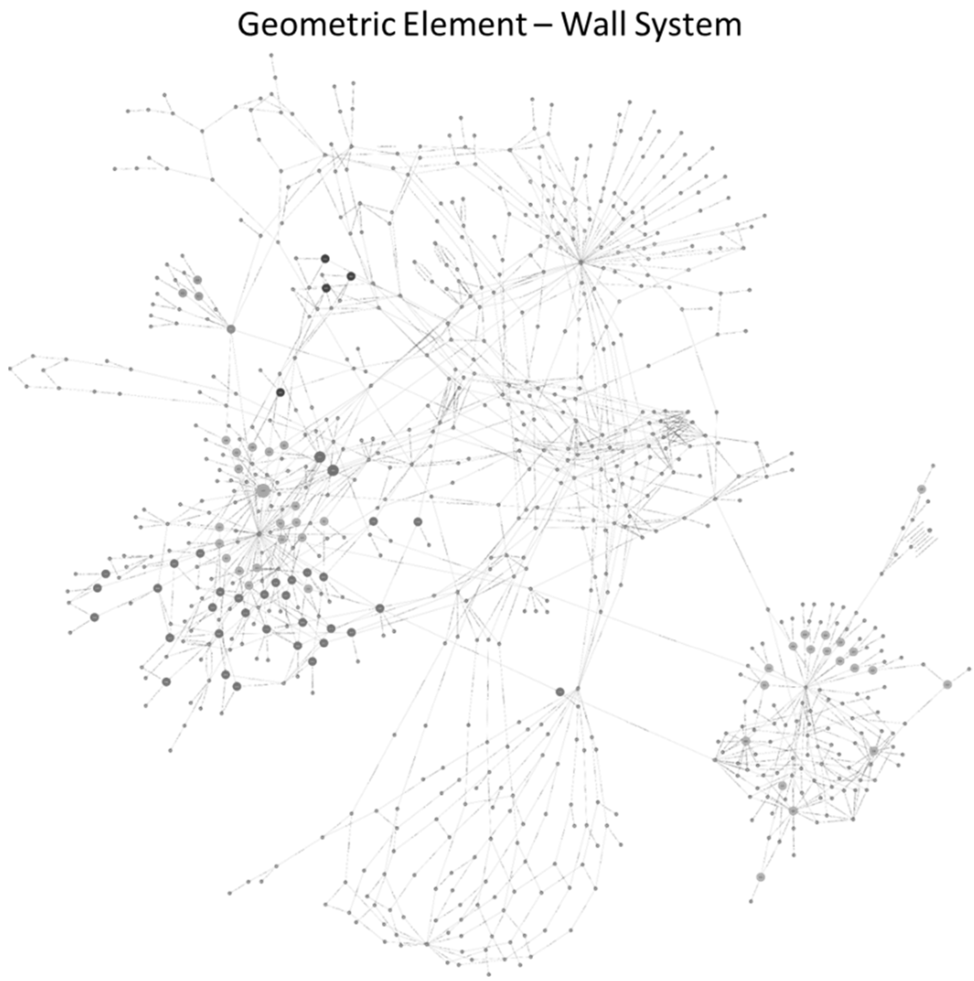
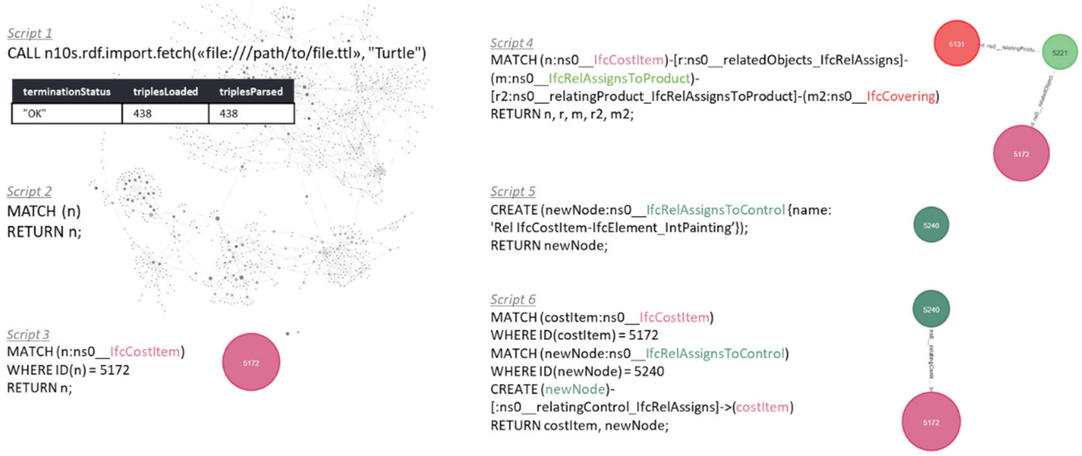


Fig. 5: Representation in the graph system of the geometric object (wall system) starting from IFC files

Through the Cypher script language, it was possible to work with the different Turtle format files available. Cypher is a declarative query language created specifically for working with graphs and interacting with the Neo4j database. Cypher queries are very expressive and readable and allow operations such as creating, editing, and querying data within the Neo4j database.

The first step was to load the data of the different files (Figure 6 - Script 1, Script 2). Subsequently, the data files were queried, and two node-entities were identified: *IfcCostItem (*Figure 6 – Script 3) which gathers all the architectural data related to individual cost items (within the cost domain) and *IfcElement* (Figure 7 - Script 7) which represents the geometric domain to which the cost item must be associated.



Script 1

```
CALL n10s.rdf.import.fetch(«file:///path/to/file.ttl», "Turtle")
```

terminationStatus	triplesLoaded	triplesParsed
"OK"	438	438

Script 2

```
MATCH (n)
RETURN n;
```

Script 3

```
MATCH (n:ns0__IfcCostItem)
WHERE ID(n) = 5172
RETURN n;
```

Script 4

```
MATCH (n:ns0__IfcCostItem)-[r:ns0__relatedObjects_IfcRelAssigns]-
(m:ns0__IfcRelAssignsToProduct)-
[r2:ns0__relatingProduct_IfcRelAssignsToProduct]-(m2:ns0__IfcCovering)
RETURN n, r, m, r2, m2;
```

Script 5

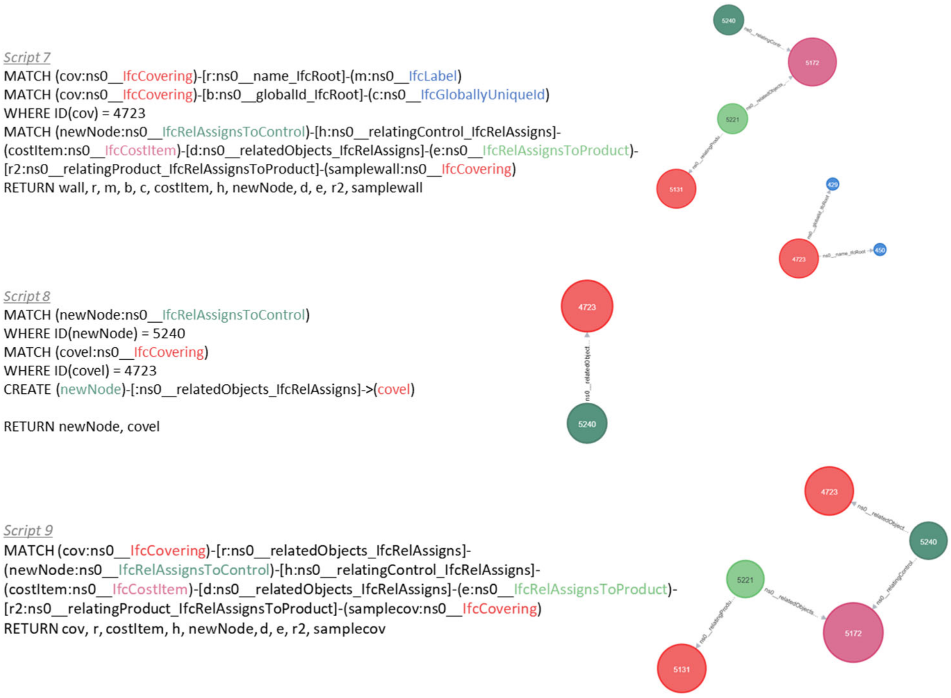
```
CREATE (newNode:ns0__IfcRelAssignsToControl {name:
'Rel IfcCostItem-IfcElement_IntPainting'});
RETURN newNode;
```

Script 6

```
MATCH (costItem:ns0__IfcCostItem)
WHERE ID(costItem) = 5172
MATCH (newNode:ns0__IfcRelAssignsToControl)
WHERE ID(newNode) = 5240
CREATE (newNode)-
[:ns0__relatingControl_IfcRelAssigns]->(costItem)
RETURN costItem, newNode;
```

Fig. 6: First script sequence for *IfcCostItem-IfcCovering* connection (cost-layer painting)

Finally, the relationship between the two nodes-entities contained in the two domains has been created (Figure 7 – Script 9). As we can see in Figure 7, the two entities belonging to the two different domains (costs and geometry) have been connected using the logic intrinsic to the IFC data model. A node ("5240 - *IfcRelAssignsToControl*") has been created that corresponds to the exact IFC entity that ensures the connection between geometric entities and cost entities. This has led to the connection of the two different domains (cost and geometry).



Script 7

```
MATCH (cov:ns0__IfcCovering)-[r:ns0__name_IfcRoot]-(m:ns0__IfcLabel)
MATCH (cov:ns0__IfcCovering)-[b:ns0__globalId_IfcRoot]-(c:ns0__IfcGloballyUniqueId)
WHERE ID(cov) = 4723
MATCH (newNode:ns0__IfcRelAssignsToControl)-[h:ns0__relatingControl_IfcRelAssigns]-
(costItem:ns0__IfcCostItem)-[d:ns0__relatedObjects_IfcRelAssigns]-(e:ns0__IfcRelAssignsToProduct)-
[r2:ns0__relatingProduct_IfcRelAssignsToProduct]-(samplewall:ns0__IfcCovering)
RETURN wall, r, m, b, c, costItem, h, newNode, d, e, r2, samplewall
```

Script 8

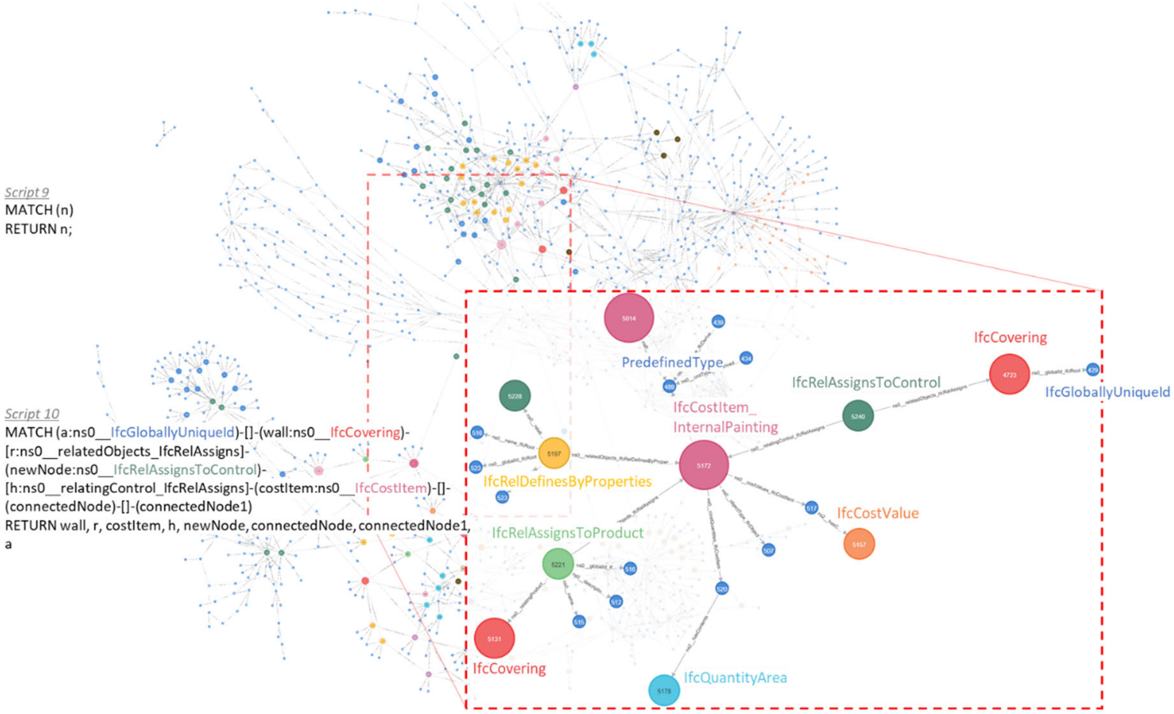
```
MATCH (newNode:ns0__IfcRelAssignsToControl)
WHERE ID(newNode) = 5240
MATCH (covel:ns0__IfcCovering)
WHERE ID(covel) = 4723
CREATE (newNode)-[:ns0__relatedObjects_IfcRelAssigns]->(covel)

RETURN newNode, covel
```

Script 9

```
MATCH (cov:ns0__IfcCovering)-[r:ns0__relatedObjects_IfcRelAssigns]-
(newNode:ns0__IfcRelAssignsToControl)-[h:ns0__relatingControl_IfcRelAssigns]-
(costItem:ns0__IfcCostItem)-[d:ns0__relatedObjects_IfcRelAssigns]-(e:ns0__IfcRelAssignsToProduct)-
[r2:ns0__relatingProduct_IfcRelAssignsToProduct]-(samplecov:ns0__IfcCovering)
RETURN cov, r, costItem, h, newNode, d, e, r2, samplecov
```

Fig. 7: Second script sequence for *IfcCostItem*-*IfcCovering* connection (cost-layer painting)

This procedure has been replicated for the remaining cost items to associate with the respective geometric objects to obtain a new system to graph containing the data of the geometric domain and the cost domain. Figure 8 shows the final output and a zoom on the association of the cost of the layer of internal painting (IfcCostItem_InternalPainting) to its geometric object (*IfcCovering*).



Script 9

```
MATCH (n)
RETURN n;
```

Script 10

```
MATCH (a:ns0__IfcGloballyUniqueId)-[]-(wall:ns0__IfcCovering)-
[r:ns0__relatedObjects_IfcRelAssigns]-
(newNode:ns0__IfcRelAssignsToControl)-
[h:ns0__relatingControl_IfcRelAssigns]-(costItem:ns0__IfcCostItem)-[]-
(connectedNode)-[]-(connectedNode1)
RETURN wall, r, costItem, h, newNode, connectedNode, connectedNode1,
a
```

Fig. 8: Final output and zoom on the association of the cost to its geometric object (*IfcCovering*)

5. DISCUSSION

The study seeks to optimize and make the data of the cost items more user-friendly. Currently, these are displayed only as informational inputs, lines of code, or even simple attributes. The aim of the research is to focus on a subject of great interest and still cause numerous legal disputes. In the research, the possibilities and the limitations of the proposed method are highlighted based on the association of different domains within a graphic system. This would allow you to link different data from files even in different formats. This study presented how price elements can be instantiated as a graph and how their visualization allows us to better understand the logic of connection and relationship between entities. It has been studied the possibility of visualizing a new domain of cost for the price list of the Lombardy Region based on a graphical system to standardize and regulate the prices and the relative information.

Converting the IFC data model into a graph system can bring many advantages (Pauwels, Zhang, et al., 2017), (Zhu et al., 2022). The main reasons why converting IFC to a graph is an objective to be pursued are:

- a simplified representation of construction information (Silvescu & Caragea, 2019);
- clearer object relationships (Figure9);
- improved data integration and interoperability (Rodriguez & Neubauer, 2010), (Mazairac & Beetz, 2013);
- improved query and processing information (Pérez et al., 2010).

However, even if this methodology makes the data more visually intuitive, due to the limitations of the current tools available, it is not yet possible to convert the information in the IFC data model into RDF in a simplified way. A further limitation is due to the use of a programming language to query, create, and associate data belonging to different domains. This causes problems in a sector, the AEC, which is only beginning in recent years to interface with these new technologies.

This research has led to technological attempts made through the writing first in IfcOpenShell of individual cost items and then in RDF format for their association to the geometric domain and their simplified visualization; in this way, it is not necessary to understand the logic behind the IFC data model. The achievement of results that are real, effective, and scalable confirms the scalability of the method as it can also be implemented for other list items.

6. CONCLUSION

The results show how SW technology can be used to show and relate the cost domain for construction projects to different domains, such as the geometric domain. Starting from the IFC data model, the cost domain has been translated into LBD ontology thanks to the IFCtoLBD converter developed by Bonduel et al. (2018). In fact, due to the complex structure of the IFC data model, the LBD representation makes it easier for stakeholders to visualize and manage the data contained in the models. The IFCtoLBD converter developed by JURI uses the smallest BOT, PRODUCT, and PROPS ontologies to better separate and represent data (Bonduel et al., 2018).

In this study, the proposed architecture for the new cost domain, developed and validated by Cassandro et al. (2023), can be easily visualized within the graphical database. As a result, individual cost items and related information become readily accessible in the graphical database. Moreover, these elements can be queried individually, associated with their corresponding geometric objects, or even extended by creating new nodes and relationships.

During the study, several limitations were identified. Firstly, the complexity of data in the AEC industry can be more difficult due to the variety of information involved, demanding a deep understanding to integrate data into LBD format. Secondly, the AEC domain includes numerous standards and ontologies, which makes it difficult to ensure compliance with all relevant standards (e.g., IFC) when integrating data into LBD. Finally, interoperability remains a concern as not all software platforms are equipped to handle LBD, posing difficulties in achieving seamless data integration and interoperability across different platforms.

For future work, it is essential to prototype and extend the concepts explored in this study. In addition, the development of an extension for the converter to improve data translation should be considered.

REFERENCES

Abanda, F. H., Tah, J. H. M., Pettang, C., & Manjia, M. B. (2011). An ontology-driven building construction labour cost estimation in Cameroon. *Electronic Journal of Information Technology in Construction*, *16*(January), 617–634. http://www.itcon.org/2011/35

Abanda, H., Ng'Ombe, A., Tah, J. H. M., & Keivani, R. (2011). An ontology-driven decision support system for land delivery in Zambia. *Expert Systems with Applications*, *38*(9), 10896–10905. https://doi.org/10.1016/j.eswa.2011.02.130

Beetz, J., Coebergh Van Den Braak, W., Botter Eindhoven, R., Zlatanova, S., & De Laat, R. (2015). Interoperable data models for infrastructural artefacts - A novel IFC extension method using RDF vocabularies exemplified with quay wall structures for harbors. *EWork and EBusiness in Architecture, Engineering and Construction - Proceedings of the 10th European Conference on Product and Process Modelling, ECPPM 2014*, 135–140. https://doi.org/10.1201/B17396-26

Beetz, J., Van Leeuwen, J., & De Vries, B. (2009). IfcOWL: A case of transforming EXPRESS schemas into ontologies. *Artificial Intelligence for Engineering Design, Analysis and Manufacturing: AIEDAM*, *23*(1), 89–101. https://doi.org/10.1017/S0890060409000122

Beetz, J., van Leeuwen, J. P., & de Vries, B. (2005). An Ontology Web Language Notation of the Industry Foundation Classes. *Proceedings of the 22nd CIB W78 Conference on Information Technology in Construction, January 2005*, 193–198.

Berners-Lee, T., Hendler, J., & Lassila, O. (2001). THE SEMANTIC WEB. *Scientific American*, *284*(5), 34–43. http://www.jstor.org/stable/26059207

Bonduel, M., Oraskari, J., Pauwels, P., Vergauwen, M., & Klein, R. (2018). The IFC to linked building data converter - Current status. *CEUR Workshop Proceedings*, 34–43.

BuildingSMART. (2022). *IfcCostItem*. IFC4_ADD2_TC1 - 4.0.2.1 [Official]. https://standards.buildingsmart.org/IFC/RELEASE/IFC4/ADD2_TC1/HTML/schema/ifcsharedmgmtelements/lexical/ifccostitem.htm

Cassandro, J., Donatiello, M. G., Mirarchi, C., Zanchetta, C., & Pavan, A. (2023). Reliability of IFC classes in ontology definition and cost estimation of public procurement. *2023 European Conference on Computing in Construction 40th International CIB W78 Conference Heraklion, Crete, Greece.*

Curry, E., O'Donnell, J., Corry, E., Hasan, S., Keane, M., & O'Riain, S. (2013). Linking building data in the cloud: Integrating cross-domain building data using linked data. *Advanced Engineering Informatics*, *27*(2), 206–219. https://doi.org/10.1016/J.AEI.2012.10.003

Fürstenberg, D., Wikström, L., Laedre, O., & No, O. L. (2021). Enabling automation of BIM-based cost estimation by semantic web technology. *ITC Digital Library*, 937–945.

Gruber, T. R. (1993). A translation approach to portable ontology specifications. *Knowledge Acquisition*, *5*(2), 199–220.

Hoang, N. V., & Torma, S. (2015). Implementation and Experiments with an IFC-to-Linked Data Converter. *Proc. of the 32nd CIB W78 Conference 2015, 27th-29th October 2015, Eindhoven, The Netherlands*, *1*, 285–294.

ifcOWL - buildingSMART Technical. (2023). Retrieved June 23, 2023, from https://technical.buildingsmart.org/standards/ifc/ifc-formats/ifcowl/

Ismail, A., Nahar, A., & Scherer, R. (2017). Application of graph databases and graph theory concepts for advanced analysing of BIM models based on IFC standard. *Digital Proceedings of the 24th EG-ICE International Workshop on Intelligent Computing in Engineering 2017*, *July*, 146–157.

Jung, S., Lee, S., & Yu, J. (2021). Ontological approach for automatic inference of concrete crack cause. *Applied Sciences (Switzerland)*, *11*(1), 1–20. https://doi.org/10.3390/app11010252

Karan, E., Irizarry, J., & Haymaker, J. (2015). Generating IFC models from heterogeneous data using semantic web. *Construction Innovation*, *15*(2), 219–235. https://doi.org/10.1108/CI-05-2014-0030

Krijnen, T., & Beetz, J. (2018). A SPARQL query engine for binary-formatted IFC building models. *Automation in Construction*, *95*, 46–63. https://doi.org/10.1016/J.AUTCON.2018.07.014

Malcolm, A., Werbrouck, J., & Pauwels, P. (2021). LBD Server: Visualising Building Graphs in Web-Based Environments Using Semantic Graphs and GlTF-Models. *Advances in Science, Technology and Innovation*, 287–293. http://dx.doi.org/10.1007/978-3-030-57509-0_26

Malinverni, E. S., Naticchia, B., Lerma Garcia, J. L., Gorreja, A., Lopez Uriarte, J., & Di Stefano, F. (2022). A semantic graph database for the interoperability of 3D GIS data. *Applied Geomatics*, *14*, 53–66. https://doi.org/10.1007/s12518-020-00334-3

Mazairac, W., & Beetz, J. (2013). BIMQL – An open query language for building information models. *Advanced Engineering Informatics*, *27*(4), 444–456. https://doi.org/10.1016/J.AEI.2013.06.001

OWL - Semantic Web Standards. (2023). Retrieved June 23, 2023, from https://www.w3.org/OWL/

Pauwels, P., & Deursen, D. Van. (2015). *IFC-to-RDF : Adaptation , Aggregation and Enrichment IFC / RDF : Adaptation , Aggregation and Enrichment*. *October*, 2–5.

Pauwels, P., Krijnen, T., Terkaj, W., & Beetz, J. (2017). Enhancing the ifcOWL ontology with an alternative representation for geometric data. *Automation in Construction*, *80*, 77–94. https://doi.org/10.1016/j.autcon.2017.03.001

Pauwels, P., & Terkaj, W. (2016). EXPRESS to OWL for construction industry: Towards a recommendable and usable ifcOWL ontology. *Automation in Construction*, *63*, 100–133. https://doi.org/10.1016/j.autcon.2015.12.003

Pauwels, P., Terkaj, W., Krijnen, T., & Beetz, J. (2015). Coping with lists in the ifcOWL ontology. *EG-ICE 2015 - 22nd Workshop of the European Group of Intelligent Computing in Engineering*.

Pauwels, P., Zhang, S., & Lee, Y. C. (2017). Semantic web technologies in AEC industry: A literature overview. *Automation in Construction*, *73*, 145–165.

Pérez, J., Arenas, M., & Gutierrez, C. (2010). nSPARQL: A navigational language for RDF. *Journal of Web Semantics*, *8*(4), 255–270. https://doi.org/10.1016/J.WEBSEM.2010.01.002

Rasmussen, M. H., Pauwels, P., Lefrançois, M., Schneider, G. F., Hviid, C. A., & Karlshøj, J. (2017). *Recent changes in the Building Topology Ontology*.

Rodriguez, M. A., & Neubauer, P. (2010). Constructions from dots and lines. *Bulletin of the American Society for Information Science and Technology*, *36*(6), 35–41. https://doi.org/10.1002/BULT.2010.1720360610

Silvescu, A., & Caragea, D. (2019). Graph Databases. *Computer Science - Encyclopedia of Big Data Technologies*, 835–835.

The Linked Building Data Community Group. (2021). https://w3c-lbd-cg.github.io/lbd/#

Vakaj, E., Lecturer, S., Environment, B., Cheung, F., Environment, B., Cao, J., Management, I., Tawil, A. H., Environment, B., Patlakas, P., & Environment, B. (2023). *An ontology-based cost estimation for offsite construction*. *28*(March), 220–245. https://doi.org/10.36680/j.itcon.2023.011

Zhang, S., Boukamp, F., & Teizer, J. (2015). Ontology-based semantic modeling of construction safety knowledge: Towards automated safety planning for job hazard analysis (JHA). *Automation in Construction*, *52*, 29–41. https://doi.org/10.1016/j.autcon.2015.02.005

Zhao, Q., Li, Y., Hei, X., & Yang, M. (2020). A Graph-Based Method for IFC Data Merging. *Advances in Civil Engineering*, *2020*. https://doi.org/10.1155/2020/8782740

Zhu, J., Chong, H., Zhao, H., Wu, J., & Tan, Y. (2022). *The Application of Graph in BIM / GIS Integration*.