BIM4EEB

BIM based fast toolkit for
Efficient rEnovation in Buildings

D3.1 A BIM-based framework for building renovation using the linked data approach and ontologies – state-of-the-art, use cases, and high-level architectural specifications

This project has received funding from
European Union's H2020 research and innovation
programme under grant agreement N. 820660

Programmes	H2020
Call for Proposal	LC-EEB-02-2018 Building information modelling adapted to efficient renovation
Project Title	BIM based fast toolkit for Efficient rEnovation in Buildings
Acronym	BIM4EEB
Project Grant Agreement	820660

D3.1 A BIM-based framework for building renovation using the linked data approach and ontologies – state-of-the-art, use cases, and high-level architectural specifications

Work Package	WP3
Lead Partner	VTT
Contributing Partner(s)	Visualynk, TUD, PoliMi, Suite5, UCC, Rise
Dissemination Level	Public
Type	R: Report
Due date	31/03/2019
Updated version 4.0 submission date	07/02/2022
Version	4.0

DOCUMENT HISTORY

Version	Date	Comments	Main Authors
0.1	13.02.2019	ToC created	Markku Kiviniemi (VTT)
0.2	15.03.2019	Draft version for 1. review	Seppo Törmä (Visualynk), Markku Kiviniemi, Jari Shemeikka, Teemu Vesanen, Teemu Mätäsniemi, Ala Hasan (VTT), Kostas Tsatsakis (Suite5), Karsten Menzel, Janakiram Karlapudi (TUD), Andriy Hryshchenko (UCC), Dag Glebe (Rise), Alberto Pavan, Claudio Mirarchi (PoliMi)
0.3	22.03.2019	Comments from reviewers	Karsten Menzel (TUD), Alberto Pavan, Claudio Mirarchi, Meherun Lucky, Bruno Daniotti (PoliMi),
0.4	28.03.2019	Refinements according to comments and additions	Seppo Törmä (VTT), Claudio Mirarchi (PoliMi), Markku Kiviniemi (VTT)
1.0 FINAL	29.03.2019	Final version submitted	Karsten Menzel (TUD)
1.1	15.11.2019	Revised version created	Seppo Törmä (VisuaLynk)
1.2	22.11.2019	Integrated the occupant behaviour background	Kostas Tsatsakis (Suite5), Seppo Törmä (VisuaLynk)
1.3	24.11.2019	Broader discussion about ifcOWL	Seppo Törmä (VisuaLynk)
1.4	25.11.2019	Version for internal review	Seppo Törmä (VisuaLynk)
1.5	28.11.2019	Language check	Seppo Törmä (VisuaLynk)
2.0	20.12.2019	Second version	Seppo Törmä (VisuaLynk)
2.1	01.12.2020	Review comments added for implementing revisions. Review Annex added.	Markku Kiviniemi (VTT)
2.2	09.12.2020	Terms and definitions table to clarify meanings	Seppo Törmä (VisuaLynk)
2.3	12.03.2021	Revision of the Section 2 with term definitions. Added a table about the framework. Rewording of text in Chapter 3.	Seppo Törmä (VisuaLynk)
2.3.1	01.05.2021	A description of ontologies, conceptual models, vocabularies, and classifications in Section 2	Seppo Törmä (VisuaLynk)
2.3.2	02.05.2021	Collection of requirements into a Section 3. Moved LBD text to BIM ontology section. Revised the section/subsection division throughout. Moved Suggestions just before the Framework section.	Karsten Mentzel (TUD), Seppo Törmä (VisuaLynk)
2.5	03.05.2021	Completion of the text about framework.	Markku Kiviniemi (VTT), Seppo Törmä (VisuaLynk)

2.6	04.05.2021	Added a diagram of the relations to other tasks in Section 1. Made corrections and improvements to several parts of the text. Moved use case descriptions to Appendix II.	Teemu Vesanen (VTT), Seppo Törmä (VisuaLynk)
2.7	31.05.2021	Condensed and streamlined state of the art. Rewrote the Section 2 about challenges. Revised formatting. Review of PoliMi's team including PC. Plagiarism check.	Markku Kiviniemi (VTT), Teemu Vesanen (VTT), Bruno Daniotti (PoliMi), Karsten Menzel (TUD), Janakiram Karlapudi (TUD), Prathap Valluru (TUD), Kostas Tsatsakis (Suite5), Tiziana Toniatti (PoliMi), Seppo Törmä (VisuaLynk)
2.8	21.06.2021	Detailed quality check. Checking of references.	Karsten Mentzel (TUD), Markku Kiviniemi (VTT)
3.0	01.07.2021	Revisions based on the quality check. Approval of WP3 leader. Approval of PC.	Markku Kiviniemi (VTT), Karsten Menzel (TUD), Janakiram Karlapudi (TUD), Seppo Törmä (VisuaLynk)
3.1	16.11.2021	New version created. Ontology section 2.3 condensed to one page, as requested. Section 5.1 renamed to "Objectives and high-level architecture", with a diagram and description added.	Seppo Törmä (VisuaLynk)
3.3	26.11.2021	Section 1.2 revised with a figure that includes all WPs and text that summarizes the challenges described in Section 3.	Seppo Törmä (VisuaLynk)
3.4	14.01.2022	Revisions suggested by the internal review. Typographical and layout corrections. As a result, the WP Leader approved the version developed.	Prathap Valluru (TUD), Karsten Mentzel (TUD), Meherun Lucky (PoliMi), Seppo Törmä (VisuaLynk)
4.0 FINAL	31.01.2022	Final version, after final and formal review and plagiarism check. Annex with review tables was kept separated from the deliverable.	Bruno Daniotti, Martina Signorini (PoliMi), Gabriele Canzi (FPM/PoliMi)

Statement of originality: This deliverable contains original unpublished work, except where clearly indicated otherwise. Acknowledgement of previously published material and the work by others has been made through appropriate citation, quotation, or both.

BIM4EEB action has received funding from the European Union under grant agreement number 820660.

EXECUTIVE SUMMARY

To improve the overall productivity and quality of renovation projects, it is more important to focus on the information flows between different tasks than on the efficiency or quality of individual tasks. Linked Data and ontologies are technological enablers for interlinked management and utilization of information across different tasks. In renovation projects an increasing volume and variety of information is available, covering the existing buildings, requirements of renovation, designs of renovation solutions, performance of the building, construction plans, and the progress of execution. There is also increasing amount of high-velocity data, coming from sensors, mobile devices, and information systems. The central objective of BIM4EEB is to develop a framework that enables interoperation among systems.

To achieve that goal, Linked Data and the ontology approach need to be adapted to the renovation and construction domain by identifying and developing appropriate ontologies and vocabularies. This report begins with an overview of Linked Building Data technologies, covering central concepts of Web of Data, Semantic Web representations, Linked Data principles, and ontologies.

The fundamental ontology identified to address the data sharing needs of renovations is ifcOWL, the ontology version of the IFC meta-data schema. It provides a standard way – supported by existing tools – to access BIM models as Linked Data thus giving a starting point for the interoperability framework in BIM4EEB. Additional ontologies are needed to connect BIM objects to other aspects of renovation projects: construction processes (time periods, activities, resources), organizational entities (labour crews, companies), information objects (BIM models, drawings, issues), sensor observations (positioning, temperature, humidity), occupant behaviour and comfort, energy efficiency, indoor air quality, acoustics, building performance, and equipment and material properties. The goal of WP3 is to identify existing ontologies in these domains and when needed, develop new ontologies based on existing standards and models utilized in these areas. The resulting set of ontologies will be developed into a harmonized and modular ontology suite that enables the connection of entities across these different areas.

This report gives a high-level architectural specification for the BIM4EEB framework based on the Linked Building Data approach and focusing on building performance. The framework will maintain a catalogue of relevant ontologies to provide their addresses and alignments with each other. The body of the report consists of a state-of-the-art analysis of ontologies in domains relevant for renovation ranging from occupant comfort, indoor air quality and acoustics to construction management, cost management, energy efficiency, building performance, equipment and material properties, and connections to urban energy systems. In the last section, the report specifies the high-level architecture of the Linked Data and ontology framework for renovation.

PUBLISHING SUMMARY

Accurate and efficient renovation projects require better management and utilization of information about existing buildings, objectives of renovation, designs of end results and performance of the building construction plans, and progress of execution. BIM4EEB uses the Linked Data approach as a glue between different data sources at technical and syntactic levels to make data interlinked and accessible. The goal of WP3 is to adapt the Linked Data approach at the semantic level to the renovation domain, through the identification and development of appropriate vocabularies. Task 3.1 introduces the overall Linked Data framework, outlines the vocabulary requirements in the renovation domain, analyses existing ontologies covering specific areas of the required vocabulary, and identifies gaps where new ontologies should be developed, or existing ones be extended or refined. To facilitate the shared ontology work, it also specifies the tools and practices for ontology development, publication and documentation.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

TERMS AND DEFINITIONS

1 Introduction

1.1 Objectives and overview

The goal of BIM4EEB is to improve the performance and quality of renovations by developing digital solutions that can efficiently utilize the various kinds of information that are increasingly produced in construction projects, such as BIM models, point clouds, project plans, energy simulation data, sensor data, and product data. Diverse datasets are produced by different parties over a renovation project and over the lifecycle of the renovated building. A major reason for the inefficiency of current renovation practices is the fragmentation of work, which is both reflected in and maintained by the lack of connections between different datasets and insufficient sharing of information among relevant parties and systems.

Consequently, the foundation for BIM4EEB is an information sharing framework that enables the interlinked sharing of information across different tasks and parties in a renovation project. WP3 (Linked Data and Ontologies) provides the specification of the framework – initially a high-level architectural specification in Deliverable 3.1 and in detail in Deliverable 3.6 – influenced by the parallel requirement analysis carried out in WP2 (Requirements for an efficient renovation process). The framework influences the development of the open data sharing platform (WP4) for sharing the results of scanning (WP5) or HVAC design and renovation scenarios (WP6) with renovation process planning (WP7 and WP8).

The approach chosen for the information sharing framework in BIM4EEB is based on the use of Linked Data technologies together with appropriate ontologies, to provide a common language that different systems during a renovation project can use to communicate. Linked Data provides the technical and syntactical levels of the communication solution; these technologies already exist and the task in BIM4EEB is to apply them in the data sharing platform and renovation projects. Ontologies, however, provide domain specific vocabularies that need to cover all essential topics in renovations; although there are many applicable ontologies available that cover important parts of relevant terminology, there are also aspects of renovations where terminology can be lacking. Additionally, this deliverable is to identify existing ontologies that could be used in renovations and find out the gaps where new definitions are needed.

The specification of the framework relies heavily on the previous research on Linked Building Data where BIM models have been brought into the domain of Linked Data using ifcOWL, an ontology version of Industry Foundation Classes (IFC) that is the standard representation of BIM models supported in practice by most BIM tools. There are particularly relevant ontologies for construction management covering construction objects, organizations, planning, and monitoring produced by the Diction project, and for sensor data produced by W3C. There are also many standards and computational models for energy efficiency, acoustics, indoor air quality, product data, and so on. In the body of this report these ontologies and other conceptual models are reviewed.

After the review of state of the art, a Linked Data and ontology framework for renovation is presented. The framework applies and specializes the existing practices in the Semantic Web, Linked Data, and ontology domain to the management of renovation projects. The framework is designed to be open in the sense that any system that implements the required functions and uses the specified conventions will be able to achieve semantic interoperability with other framework-compliant systems in the renovation domain.

1.2 References to other activities and to the state of the art

This report is a state-of-the art study concerning Linked Data technologies and ontologies for BIM-based renovation: it reviews the related research and development concerning the technologies, ontologies, and other conceptual models relevant for renovation.

Based on the state-of-the-art review, this report defines an open Linked Data and ontology framework for BIM-based renovation and makes suggestions about ontologies to be adopted in the subsequent tasks in BIM4EEB project. Within BIM4EEB this work directly affects the subsequent tasks in WP3 and indirectly the technologies and tools developed in other work packages.

Task 3.1 started in the beginning of the project and the work was carried out in parallel with the Task 2.1 that analyses the renovation processes (Figure 1). The interactions between these tasks have influenced the work carried out in this task, even though Task 2.1 did not produce direct input to be utilized in the Task 3.1.

Task 3.1 has directly influenced the other tasks in WP3 by providing them the initial definition of the Linked Data and ontology framework and the initial state-of-the-art study to guide their work. Especially important is the Task 3.6 that has produced the integrated ontology framework to be used for data representation and sharing by the subsequent tasks on the development of the data sharing platform (BIMMS) and the renovation tools.

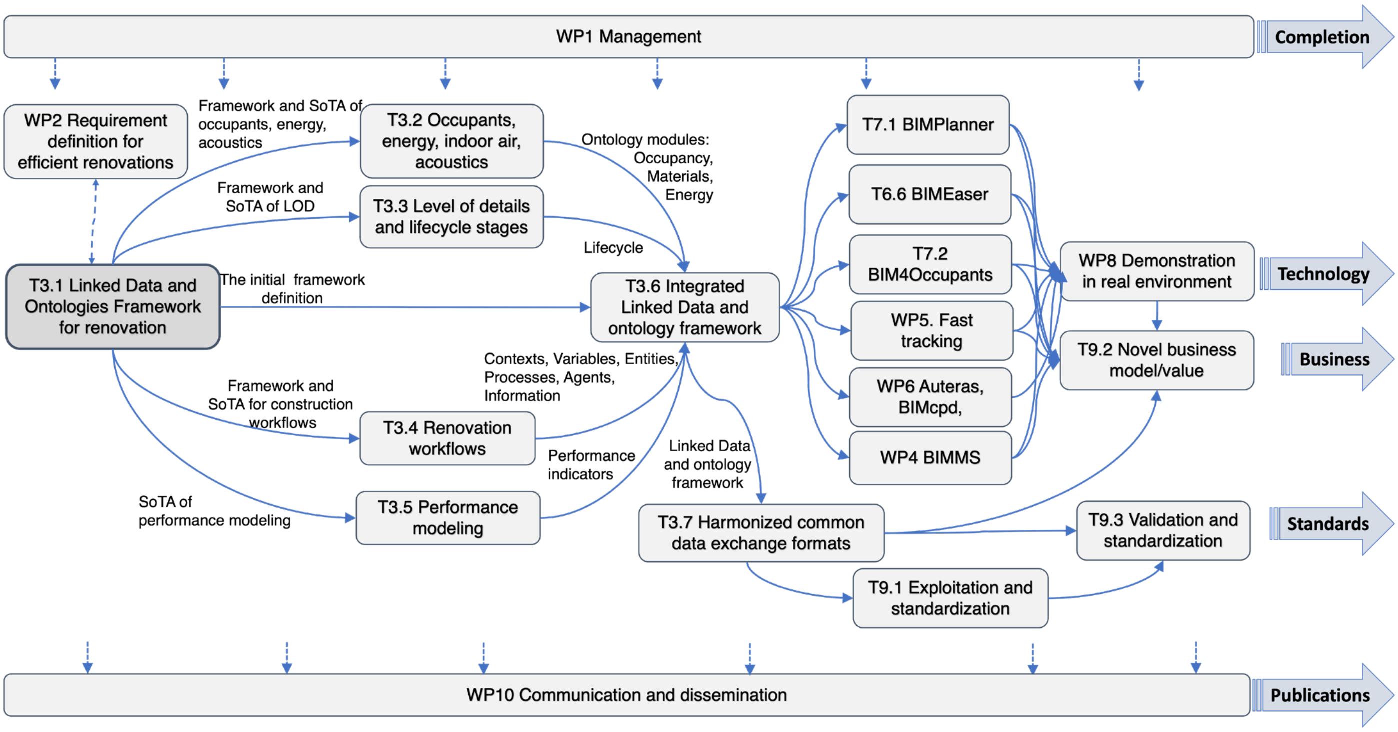


Figure 1: Relationships with other work packages and tasks

The work in Task 3.1 aims to utilize and build upon the work done on Linked Data and ontologies in different organizations including buildingSmart, W3C, ISO and ETSI, and in relevant research projects such as the DiCtion project[1] on digitalized construction workflows.

The main challenge of Task 3.1 is to provide terminology to connect the multiple conceptual domains

[1] DiCtion - Digitalized Construction Workflows, a research project funded by Business Finland, 2018-2020

where the tools and systems created in WP4, WP5, WP6 and WP7 operate, including building design, energy efficiency, occupant profiles and requirements, construction processes, sensor data, product and material data, and information management. The challenges of decentralized nature renovation projects and the evolution of information during the execution had a major impact on the work. The goal was to define the proper terminology by utilizing the established ontologies developed for different conceptual domain, and to organize the information management in accordance with relevant standards. The challenges addressed by Task 3.1 are elaborated in the Section 3 and summarized in the Subsection 3.6.

1.3 Innovative results and progresses

This report makes advances by

- presenting an introducing to Linked Data and ontologies and their application in the renovation domain,
- providing a state-of-the-art study about Linked Data and ontologies and other conceptual models in renovation domain, and
- presenting a high-level architectural specification of the framework for the renovation management systems in BIM-based building renovations using Linked Data and ontologies.

1.4 Structure of the deliverable

Section 2 gives an overview of basic concepts covering interoperability, Web of Data, Linked Data, and ontologies. In Section 3 the challenges of renovation projects are outlined. Section 4 contains a review of applicable ontologies and other conceptual models, together with suggestion for subsequent work. Each of the information domains are studied for the existing ontologies, standards or computational models, and the gaps for ontology development in WP3 are identified. Finally, a high-level architectural specification for Linked Data and ontology framework for BIM-based renovations are presented.

2 Technical approach and basic concepts

This section describes the technical approach and related concepts (e.g. Web of Data, Linked Data, ontologies) adopted in BIM4EEB to achieve interoperability between different tools in the renovation domain.

2.1 Interoperability

The results of the ongoing digitalization of the construction and renovation industry are demonstrated by many impressive point solutions for individual tasks, with some examples shown in Figure 2. Each of the task-specific solutions can dramatically improve the productivity and quality of the task in question.

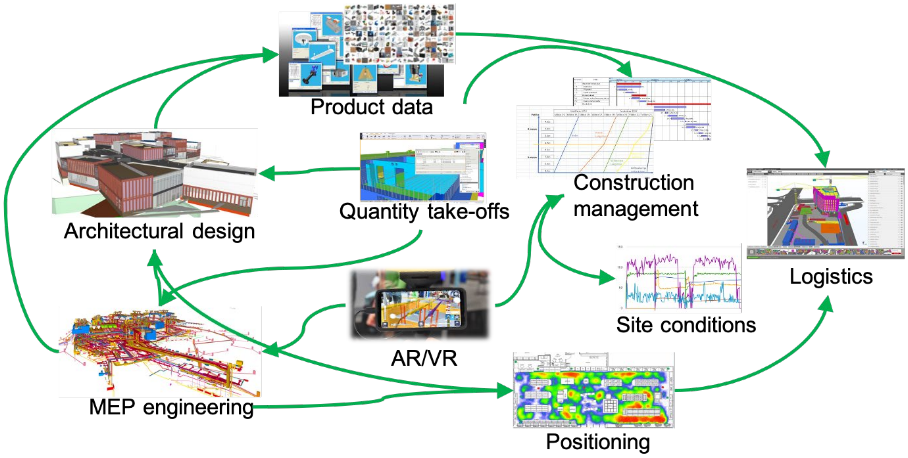


Figure 2: Point solutions versus information flows

However, the *overall productivity and quality* of construction and renovation depends crucially on *what happens between the individual tasks*: that is, how different point solutions *share information and interoperate with each other*. In recent years the questions of information sharing, data flows and interoperability have become more pressing, since the variety, volume and velocity of data creation have grown all the time. Consequently, the traditional mode of working, that is, manual exchange of information between systems, has become an increasingly restricting bottleneck: people cannot work at the rate, for instance, in which sensor data or imaging information can be produced, and can even find it difficult to retain high quality of work in such repetitive cognitive tasks. Consequently, solutions are needed to reduce or even remove manual information exchange by *automatizing the way different systems interoperate*. The following levels of interoperability between systems can be identified (Singh, 2005):

- *Technical interoperability*: At the lowest level there must be a *connection* between the systems, and an *interface* through which bits and bytes can be transferred from one system to another. This can nowadays be achieved by connecting the systems to a common communication network (for instance, the Internet) and have APIs in the systems.
- *Syntactical interoperability*: There should be common understanding regarding the format of transferred data, so that the recipient can parse the structure of the data. This can nowadays be solved by using standard data formats (e.g., JSON, XML, CSV, or STEP Physical File Format).
- *Semantic interoperability*: The terms used in the data (types of entities, their properties, datatypes, and identifiers) should be understood in a sufficiently similar manner by the systems to make their practical operations successful. The solutions for semantic interoperability are still in development with approaches based on standards, ontologies, wrappers, and mediators (Abukwaik, 2014). In the construction domain also classification systems address aspects of semantic interoperability.

- *Pragmatic interoperability*: There are several issues related to processes, security, adaptation to dynamic changes, organizational arrangements, and even legal considerations where the systems may need to be in an agreement to successfully work together. Occasionally some of these aspects are regarded as additional layers of interoperability (EIF, 2017) (Abukwaik, 2014).

Term	**Interoperability**
Definition	“The ability of computer systems or software to exchange and make use of information“. (Oxford Language) “A characteristic of a product or system, whose interfaces are completely understood, to work with other products or systems, at present or in the future, in either implementation or access, without any restrictions” (https://interoperability-definition.info)
Demarcation	If contrasted with *integration*, interoperability refers to loosely-coupled systems working together, whereas integration means a more tightly-coupled setting. Interoperability thus enables more open and multi-directional collaboration, while integration suits better for more closed and static settings. As elaborated by Singh and Huhns (2005): “*We often hear discussion of integrating schemas, databases, workflows of services. Integration refers to the idea of putting diverse concepts together to create an integrated whole. Having an integrated model would facilitate the services working well together. This contrasts with interoperation, which refers to making services work together by sharing appropriate messages and using narrow, agreed-upon, interfaces, but without any single conceptual integration. In general, interoperability is what we desire, and integration is often the wrong way to go about trying to obtain it. This is because integration can be expensive to achieve. Also, integration is fragile, meaning that when one of the integrated services changes, it may affect the integrated whole. Therefore, it is wiser to motivate service composition from the perspective of integration.*” (Singh, 2005)

Definition 1: Interoperability

As can be seen, the solutions to tackle technical and syntactic interoperability, at least in ordinary domains, are already available, and to achieve interoperation at those levels is generally a matter of willingness and effort of implementation. Semantic interoperation is still an area of active development, with promising technologies and initial results. However, even after the problems of semantic interoperability have been tackled, issues still remain to be solved at the pragmatic level to make systems work seamlessly together.

2.2 Web of Data

Web was born in the beginning of 1990’s as a decentralized platform for publication of hypertext documents for human readers. It soon caused an almost explosive growth of the global space of interlinked documents. An important enabler for the growth were Web browsers, visual and easy-to-use end-user tools that allowed people to easily explore the vast sphere of Web pages. Web has since become a central part of information and communication technologies. Over the time it has been extended with many new capabilities: advanced media, services, mobility, real-time communication, and so on.

The extension that is especially relevant to interoperability is *Web of Data*: the support for *publication and access of structural data on the Web*. It provides an open framework for data sharing in decentralized environments, where different pieces of data can originate from different actors. The goal is that data shared will be mutually understandable for participating actors and ready for machine consumption.

Web is the most successful *decentralized* publication system ever. With Web of Data, the obvious question is how to support *decentralized publication of data*. How could any user publish data, to make her data available for others to access? How to do that so that data will be available also in machine-understandable formats since data is primarily meant to be used by applications and tools, not manually by users.

Term	**Semantic interoperability**
Definition	"[In semantic interoperability] the semantic aspect refers to the meaning of data elements and the relationship between them. It includes developing vocabularies and schemata to describe data exchanges, and ensures that data elements are understood in the same way by all communicating parties" (EIF 2017) "Shared reference model for information exchange and clear data meanings" (Abukwaik 2014) "The ability of computer systems to exchange data with unambiguous, shared meaning". (Wikipedia)
Demarcation	Semantic interoperability must be distinguished from technical and syntactic interoperability that are its prerequisites, as well as from pragmatic aspects of interoperability that it does not address. That is, semantic interoperability requires the deployment of solutions for lower levels (e.g., connectivity, APIs, common data formats) but it is not the final level in the pursuit for complete interoperability. Issues can remain in many areas, such as establishing single sign-on, or tackling problems caused by different privacy regulations at different jurisdictions. Examples: "Agreements on reference data, in the form of taxonomies, controlled vocabularies, thesauri, code lists and reusable data structures/models are key prerequisites for achieving semantic interoperability. Approaches like data driven design, coupled with linked data technologies, are innovative ways of substantially improving semantic interoperability." (EIF 2017)

Definition 2: Semantic interoperability

The goal to support decentralized publication of data on the global scale of the Web has created a range of challenges. The solutions and practices that apply in local data sharing contexts – for instance, within one company or country – are not directly applicable. Among the users in a global community there should be a way to refer to objects using unique global identifiers: they can be used to retrieve objects or create links to objects. The identifiers of data objects are the same as Web pages: URIs (Uniform Resource Identifiers) or nowadays actually IRIs (Internationalized Resource Identifiers). They have the required properties: (1) global uniqueness and (2) retrievability.

URI is a generalization of URL (Uniform Resource Locator or a Web address): it can identify also non-information resources (a building, a pylon, a wall, a person) that are obviously non-retrievable over the Web. However, URIs can be used with little confusion to retrieve a *description* of a non-information object: instead of an object itself, retrieval returns *information about* the object. As the published datasets typically contain many URIs from the same domain, a shorthand notation called CURIes (Compact URI) is used, also in examples below. The format of a CURIe is: prefix:reference.

Web of Data uses a graph-based data representation schema called Resource Description Framework (RDF) (Cyganiak, 2014). It allows the representation of elementary facts about any subject as *triples*, that is 3-place tuples of the form <subject, predicate, object>. There are two simultaneous perspectives to triples. A triple <:Floor3, :hasSpace, :Room31> can be considered as a:

1. *Statement*: A triple can be read as an elementary fact: *"Floor3 hasSpace Room31"*
2. *Link*: A link from the subject to the object labeled with the predicate: 

Together these two perspectives define a graph-based representation with clearly understandable meaning. The information can be given as a list of triples that forms a graph (Figure 3).

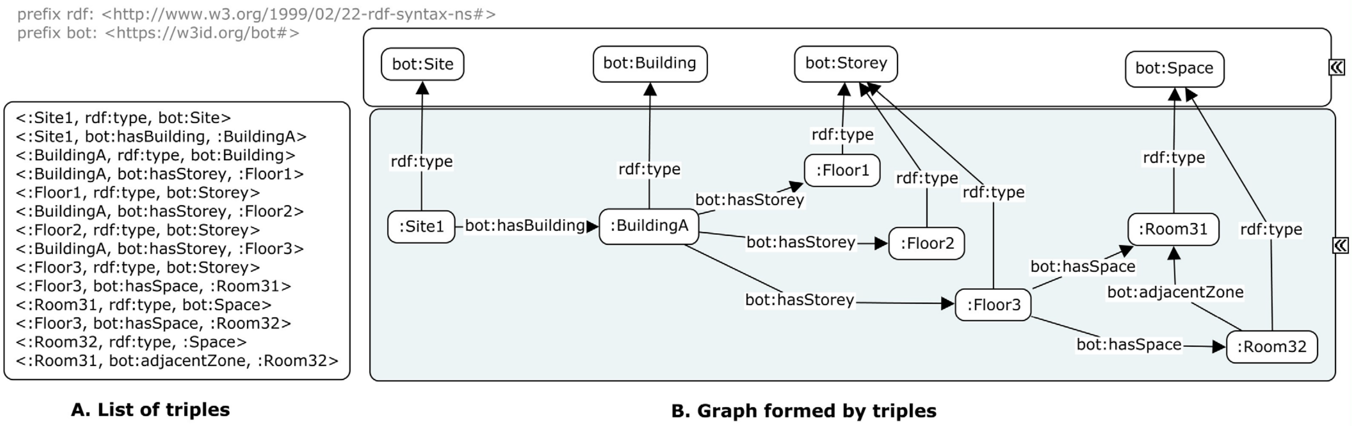


Figure 3: RDF triples represented (A) as a list and (B) as a graph

The advantages of RDF arise from the uniform and flexible graph representation that can cover various kinds of information created in an independent and decentralized manner by many different parties. It supports the following functionalities that are not present in object-oriented representation approaches:

1. *Merging* – combining information from multiple sources. If there are other data sources about BuildingA – for instance, about occupant comfort or energy efficiency - their triples sets can be combined to form a larger graph. The resulting triple set is the union of those of the input graphs.
2. *Reasoning* – adding inferred statements to a graph. Based on existing data and inference rules, a reasoning engine may be able to derive conclusions. Conclusions are additional statements, represented also as triples. They can be inferred either in advance and added to the graph (as materialized reasoning) or at query time and returned as part of the query result.
3. *Linking* – non-local references across datasets can be made in a uniform manner. Since URIs are globally unique and retrievable, the subject and object of a triple can be entities represented in different hosts on the net. The overall graph can thus span multiple different hosts.
4. *Standardization* – a standardized data format. RDF is a standardized graph-based data format and RDF databases are standard-based NoSQL databases, with standard serialization formats, conceptual models, and query languages. This improves tool chain interoperability and prevents the vendor lock-in in data management solutions.

RDF databases are graph databases that support the storage, management and querying of RDF data. Multiple *repositories* can be created into an RDF database, and each repository contains an *RDF Dataset*. An RDF Dataset is a set of RDF graphs, consisting of one *default graph* and an unlimited number of *named graphs* (Cygniak, 2014). Each named graph is identified by its own URI. The possibility to store and use data from multiple different graphs creates an additional dimension to RDF. Instead of triples, the statements are actually represented as *quads* in the form <graph, subject, predicate, object>.

When RDF data is stored in a file or transferred between systems, standard formats to serialize the RDF graphs to strings are needed. There are multiple different serialization formats available, and usually they are all supported by different RDF databases and RDF libraries. Currently, the most popular ones are:

- Turtle (Terse RDF triple language) (Prud'hommeaux, 2014)
- JSON-LD (JSON for Linking Data)

Figure 4 shows the data of Figure 3 serialized both in Turtle and in JSON-LD. The prefix declarations are shown in blue colour, references to external concepts and properties in green colour and the new definitions in the model in red colour. The black colour shows the boilerplate part of the serialization: in Turtle this consists mostly of punctuation marks while in JSON-LD there are several different keywords.

```
@prefix :    <https://ex.com/> .
@prefix rdf: <http://www.w3.org/1999/02/22-rdf-syntax-ns#> .
@prefix bot: <https://w3id.org/bot#> .

:Site1 rdf:type bot:Site ;
  bot:hasBuilding :BuildingA .

:BuildingA rdf:type bot:Building ;
  bot:hasStorey :Floor1, :Floor2, :Floor3 .

:Floor1 rdf:type bot:Storey .

:Floor2 rdf:type bot:Storey .

:Floor3 rdf:type bot:Storey ;
  bot:hasSpace :Room31, :Room32 .

:Room31 rdf:type bot:Space ;
  bot:adjacentZone :Room32 .

:Room32 rdf:type bot:Space .
```

Turtle

```
{ "@context": { "bot": "https://w3id.org/bot#",
                "rdf": "http://www.w3.org/1999/02/22-rdf-syntax-ns#" },
  "@graph":  [ { "@id": "https://ex.com/Site1",
                 "@type": "bot:Site",
                 "bot:hasBuilding": { "@id": "https://ex.com/BuildingA" } },
               { "@id": "https://ex.com/BuildingA",
                 "@type": "bot:Building",
                 "bot:hasStorey": [ { "@id": "https://ex.com/Floor2" },
                                    { "@id": "https://ex.com/Floor1" },
                                    { "@id": "https://ex.com/Floor3" } ] },
               { "@id": "https://ex.com/Floor1",
                 "@type": "bot:Storey" },
               { "@id": "https://ex.com/Floor2",
                 "@type": "bot:Storey" },
               { "@id": "https://ex.com/Floor3",
                 "@type": "bot:Storey",
                 "bot:hasSpace": [ { "@id": "https://ex.com/Room31" },
                                   { "@id": "https://ex.com/Room32" } ] },
               { "@id": "https://ex.com/Room31",
                 "@type": "bot:Space",
                 "bot:adjacentZone": { "@id": "https://ex.com/Room32" } },
               { "@id": "https://ex.com/Room32",
                 "@type": "bot:Space" } ] }
```

JSON-LD

Figure 4. Example data serialized in Turtle and JSON-LD

Figure 5 presents the similarities of the regular Web and the Web of Data. In a similar manner as there can be hyperlinks from one document to another in the regular Web, there can be links to objects (represented in a graph-oriented manner) in the Web of Data.

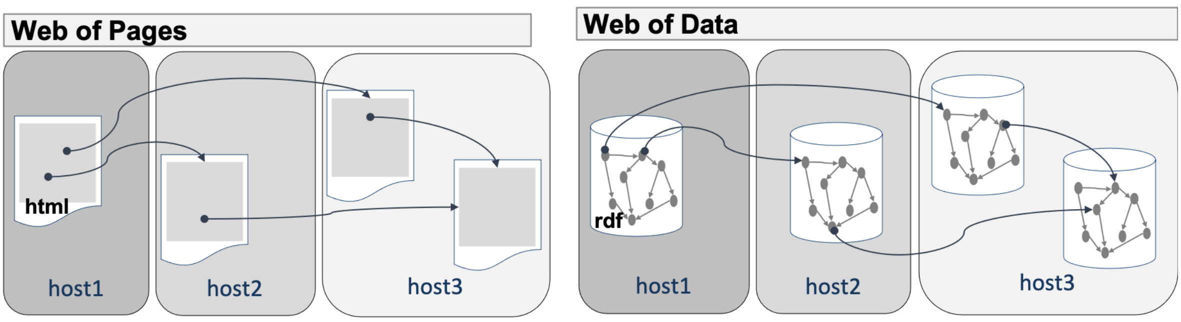


Figure 5: Web of Data versus the regular Web

When data objects are retrieved, their efficient use of them requires almost always that their types and types of their properties are known. In a global environment it is not possible to rely on the use of simple labels for type names, since there is a potentially unlimited set of users that interpret same terms to mean different entities and different terms to mean same entities. There are a range of issues such as *synonymy* (different terms have same meanings), *homonymy* (same term has multiple unrelated meanings), and *polysemy* (same term has multiple related meanings) to be solved. All of these are exacerbated by the problem of machine understandability, since for a machine no term has any meaning as such; whatever meaning a term has arises from its formal relations with other terms.

In the Web of Data problems of concepts are tackled with *ontologies*: controlled vocabularies that provide formal specifications of conceptualizations. They specify which concept should be used, what is its unique global identifier, what are its properties and relations, and what constraints apply to it.

Table 1: Levels of interoperability related to corresponding technical solutions

<table>
<tr><th>Level of interoperability</th><th>Corresponding technical solution</th></tr>
<tr><td>4. Pragmatic interoperability
- process and organizational aspects
- security, dynamics, and legal issues</td><td>Applications
- based on Linked Data and ontologies
- manage other aspects of interoperability</td></tr>
<tr><td>3. Semantic interoperability
- terms used in data</td><td>Ontologies
- definitions for classes, properties, and restrictions
- alignment to other ontologies</td></tr>
<tr><td>2. Syntactic interoperability
- data formats
- structure of data</td><td rowspan="2">Linked Data technologies
- data publication and granular access (Linked Data principles)
- graph-based data model (RDF)
- graph databases
- serialization formats (Turtle, Trig, JSON-LD)
- query language (SPARQL)
- HTTP(S) protocol, REST interface</td></tr>
<tr><td>1. Technical interoperability
- connectivity
- access interfaces</td></tr>
</table>

The Web of Data involves Linked Data technologies at the technical and syntactic levels of interoperability (Table 1). Ontologies provide concepts and properties at the level of semantic interoperability. Linked Data applications can utilize data that is based on commonly used concepts and formats, but they will still need to tackle the interoperability issues at the pragmatic level, including the interaction patterns between data providers and clients.

The full power of Semantic Web technologies comes into play only in decentralized environments where network effects increase the value of data: publishing of data creates new opportunities for incoming links from other datasets, adding an outgoing link to a data item enriches its contents by additional details and context, and new links establish information access paths and information flows between datasets.

Tim Berners-Lee (2006) suggested four principles to observe when data is published:

1. *“Use URIs as names for things”*
2. *“Use HTTP URIs so that people can look up those names.”*
3. *“When someone looks up a URI, provide useful information, using the standards (RDF*, SPARQL)”*
4. *“Include links to other URIs. so that they can discover more things.”* (Berners-Lee, 2006)

These principles suggest that data objects should be published on the Web, not just in an inhouse RDF database with local identifiers. The data object will then be accessible and retrievable using its URI. When retrieved, a structural description of the object is received, possibly containing URIs of further objects.

It should be stressed that from the perspective of a data publisher, Linked Data principles aim to advance the following complementary practices to *allow incoming links* – publish your own data in a manner that others can link to it – and *provide outgoing links* – enrich your data with links to other relevant data. Wide observance of these principles would significantly increase the interconnectedness and network effects in the sphere of the Web of Data.

Term	Linked Data
Definition	Structural data in which the represented entities are identified with HTTP URIs and which is published in a manner that the representations of the identified entities are retrievable over HTTP using their URIs so that the retrieved representations can include URIs of further entities. Linked Data Principles determine the rules of how data should be published on the Web to be regarded as Linked Data (Berners-Lee, 2006): *1. Use URIs as names for things* *2. Use HTTP URIs so that people can look up those names.* *3. When someone looks up a URI, provide useful information, using the standards (RDF*, SPARQL)* *4. Include links to other URIs. so that they can discover more things.*
Demarcation	Linked Data is a data level concept and specifies how data should be published. It does not take any position with respect to semantics of data (concerning the use ontologies or schemas). However, it is quite compliant with the principles – and generally recommendable – to include also type information about the data items exchanged, and this type of information can be defined in ontologies.

Definition 3: Linked Data

Term	Ontology
Definition	"An explicit specification of a conceptualization (the concepts and entities that are assumed to exist in an area of interest and the relationships that hold among them). That is, an ontology is a description (like a formal specification of a program) of the concepts and relationships that can exist for an agent or a community of agents." (Gruber, 1999).
Demarcation	As a subdomain of philosophy, ontology studies the questions of what exists and under which conditions can something be regarded as existing. Computational ontology is a field of applied ontology that defines what kinds of entities can exist in a particular domain, and how they may relate to other entities. Ontologies can be contrasted with data models (Spyns, 2002): data models consider particular use cases or technological solutions while ontologies are domain models to capture what exists in the domain without reference to use cases or technologies. Therefore, ontologies work as bridges between use cases and support interoperability between different systems.

Definition 4: Ontology

2.3 Ontology

An ontology (Definition 4. Ontology) defines the classes, properties and constraints for entities in a domain[2]. The languages to define ontologies have evolved over the time. Currently the following are used:

- RDFS: RDF Schema (Brickley, 2014a) enables the definition of class hierarchies (rdfs:Class, rdfs:subClassOf), property hierarchies (rdfs:Property, rdfs:subPropertyOf), and their connections (rdfs:domain, rdfs:range), literals, datatypes, collection types and reification.

[2] The relation of ontologies and classification systems prevalent in construction industry (OmniClass, Uniclass, CoClass, ETIM, Talo2000, etc.) is discussed in the Deliverable 3.4 (BIM4EEB-D34, 2022).

- OWL: Web Ontology Language (W3C OWL WG, 2012) provides an extensive set of representational primitives for classes, properties, and datatypes based on description logics (Baader, 2003).

OWL supports logical reasoning about an ontology and data represented according to the ontology. Some decision problems in restricted description logics are tractable but they become intractable when expressive power is increased (Levesque, 1987): there is a trade-off between the expressive power and computational complexity. As none of the trade-offs is generally optimal, OWL2 provides alternative profiles (EL, QL, DL and RL[3]) to choose from.

Term	**Ontology alignment**
Definition	Specification of correspondences between the terms in different ontologies. The terms can denote concepts, properties, or individuals. (Ardjani, 2015)
Demarcation	The correspondences mostly consist of definitions that two classes are equivalent, one is subclass of another, or two properties are equivalent or one is a subproperty of another. However, it is possible to align terms with complex classes, property restrictions or property chains as well.

Definition 5: Ontology alignment

Two systems can semantically interoperate by using the terms from the same ontology. If they use terms from two different ontologies, semantic interoperation may still be possible if the two ontologies have been aligned with each other. Ontology alignment (Definition 5) is the specification of correspondences between the terms of different ontologies.

2.4 Querying and managing data

RDF data can be queried using SPARQL (W3C SPARQL WG, 2013), a standard query language resembling SQL but adapted to RDF graphs. A SPARQL query contains a body whose purpose is to match a triple pattern and bind different parts of the triples to variables and a head that allows to return the matching bindings as the result. For example, a query that returns all distinct subjects of all triples:

```
SELECT DISTINCT ?s WHERE { ?s ?p ?o }
```

⇨ “http://ex.com/Site1”, “http://ex.com/BuildingA”, “http://ex.com/Floor1”, “http://ex.com/Floor2”, “http://ex.com/Floor3”, “http://ex.com/Room31”, “http://ex.com/Room32”

Another example: select all classes used in a model with the number of instances of each class:

```
SELECT ?type (COUNT(?subject) as ?c) WHERE {
  ?subject a ?type.
} GROUP BY ?type
```

⇨ "https://w3id.org/bot#Site", "1" ,
"https://w3id.org/bot#Space", "2" ,
"https://w3id.org/bot#Building", "1" ,
"https://w3id.org/bot#Storey", "3"

In most queries it is necessary to refer to classes or properties defined in ontologies. The following query selects all the buildings in the model where elevation of reference height is smaller than elevation of terrain:

```
PREFIX ifc: <http://www.buildingsmart-tech.org/ifcOWL/IFC4_ADD2#>
SELECT ?building ?erh ?et WHERE {
  ?building a ifc:IfcBuilding ;
        ifc:elevationOfRefHeight_IfcBuilding ?erh;
        ifc:elevationOfTerrain_IfcBuilding ?et
  FILTER (?erh < ?et) }
```

[3] See also: https://www.cambridgesemantics.com/blog/semantic-university/learn-owl-rdfs/flavors-of-owl/

SPARQL is a large and complex query language with many advanced constructs, including optional matches, negation, subqueries, and different kinds of filters. In addition to SELECT queries that return variable bindings matching the query body, SPARQL also supports other query types: ASK (return true or false), DESCRIBE (returns a description of resource) and CONSTRUCT (return a new RDF graph). CONSTRUCT form allows the query to create an RDF graph using its variable bindings. Therefore, SPARQL queries stay in the realm of RDF: take RDF as input and produce RDF as output. For instance, it is possible to map existing RDF graphs into new graphs using just SPARQL queries, or to use queries in stream processing to match conditions (in RDF) and to produce complex events (in RDF) (Rinne, 2012).

SPARQL supports federated queries, where part of a query can be delegated to another SPARQL endpoint by using the SERVICE form. Federated queries can be used to enrich queries with additional data available on the Web. However, due to the network delays - and all the steps to prepare a remote query - their performance is significantly lower than that of regular queries. An example is the following:

```
PREFIX foaf:   <http://xmlns.com/foaf/0.1/>
SELECT ?name FROM <http://example.org/myfoaf.rdf>
WHERE { <http://example.org/myfoaf/I> foaf:knows ?person .
      SERVICE <http://people.example.org/sparql> {
        ?person foaf:name ?name } }
```

In addition to the query capabilities, there is the SPARQL UPDATE language[4] that contains graph update operations (simple INSERT DATA and REMOVE DATA actions, query-based INSERT and DELETE actions, LOAD and CLEAR data) and graph management operations (CREATE, DROP, COPY, MOVE and ADD graphs). SPARQL UPDATE language has been used, for instance, in the INSTANS Complex Event Processing system[5] to maintain the system state based on the incoming events (Rinne, 2012). Finally, SPARQL connects to the ontologies and reasoning through entailment regimes that specify the kind of reasoning done during a query execution. With an entailment regime, some triples can be derived during the query execution, for instance, resulting from transitive or inverses properties declarations.

[4] https://www.w3.org/TR/sparql11-update/
[5] https://github.com/aaltodsg/instans

3 Interoperability challenges in renovation projects

This section outlines the interoperability challenges faced in renovation projects. The purpose is to identify what is needed for the interoperation of tools: that is, how can they share data in a manner that other tools could utilize it without manual interpretation or translation by human personnel. The relevant characteristics of renovation projects are analysed, covering the multiple information domains, multiple parties and datasets, and evolution of data. The need for standards compliance it elaborated. The challenges are outlined, at instance level, type level, and social level.

Terms	**Project** **Construction project** **Renovation project**
Definition	Project is a temporary endeavour undertaken to create a unique result. (PMI, 2021) Construction project is a project whose result is a building conforming to requirements. Renovation project is a construction project in which the building already exists, and the requirements represent an improvement from its state preceding the project.
Demarcation	Renovation project is a part of the overall facility management (FM) process of a building. FM consists mostly of an incremental, day-to-day improvement process while renovation means an extensive repair of the building at one time. Renovations can be disruptive to the operations of the building and usually only few renovation projects will be carried out during the whole construction lifecycle.

Definition 6: Project, construction project, and renovation project

3.1 Special characteristics of renovation projects

Renovation projects are special kind of construction projects that happen in the context of existing buildings (Definition 6). From the management perspective, most of the characteristics of renovation projects are similar to those of new construction projects, ranging from the large number of parties, contractual structures, and multiplicity of datasets to the nature of design and planning. There are, however, certain differences in the emphasis of renovations when compared to new construction:

- Existing building:
    - Detailed survey of the existing building is costly and may still leave a lot of uncertainty;
    - Uncertainty of existing structures often causes surprises when they are opened;
    - Existing performance profile of a building creates more strict success criteria for the project;
    - More constrained surroundings require more detailed planning;
    - Generally, better infrastructure services are available.
- Larger stakeholder groups of owners and occupants:
    - Decision making is more complex;
    - Coordination of activities with the occupants is needed if they stay in the building;
    - Privacy of occupants need to be respected.
- Work content:
    - Generally, less foundation work and frame construction and more indoor work.

The discussion overleaf focuses on renovation projects, even though a large part of the interoperability challenges mentioned are common to all construction projects.

3.2 Connecting multiple conceptual domains

Multiple conceptual domains meet in the design, management, and execution of renovation projects:

- Building design: Current design is increasingly based on BIM. It is divided into several design disciplines: architectural design, structural engineering, MEP engineering, etc.
- Processes: Planning, execution, and management of the project take place at multiple hierarchical levels and use distinct approaches for each phase.
- Energy: Data gathering about energy consumption, and design and simulation of energy efficiency solutions and energy systems.
- Occupancy: Coordination with occupants is required. Data about occupant behavior, comfort, and preferences concerning user comfort and occupant actions are gathered.
- Sensor and metering data: Data gathered about the positions and trajectories of workmen, equipment, and materials at the site, about space occupancy and environmental conditions, including consumption figures.
- Products and supply chains: Types and brands of selected products and materials, specification data, vendor relations, and coordinated deliveries to the site.
- Classification systems: Categories of products and activities defined by the vocabularies used in the project.

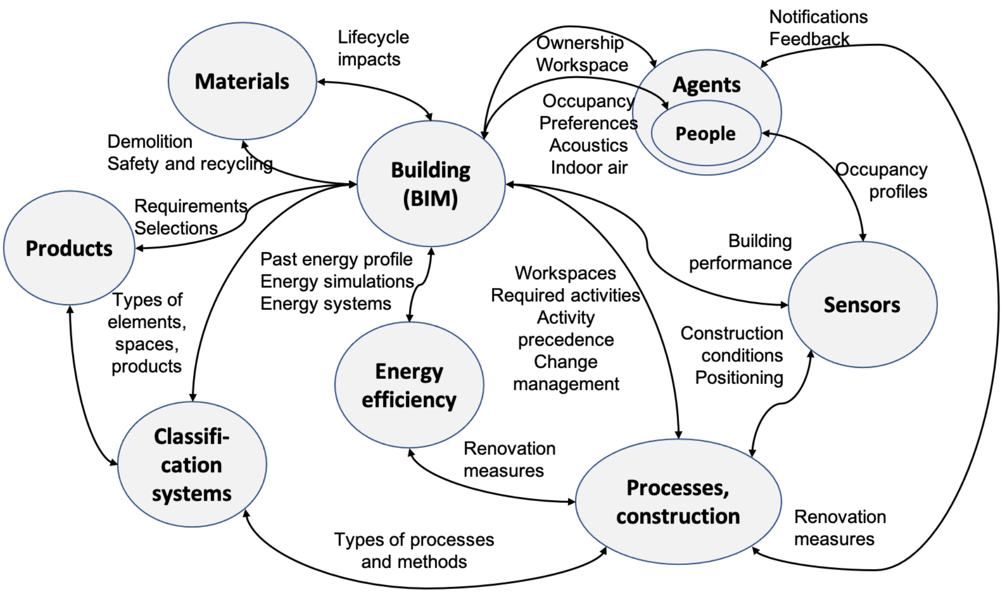


Figure 6: Different conceptual domains and their interactions

Some interactions between these domains are illustrated in Figure 6. For a successful management of a renovation project, there is a need to associate information from one domain with another. For instance, how to express in the terms of an ontology that some zone of the building is a residential unit, that some agent is the owner of that unit and another is the occupant, that there is an activity that causes electric shutdown affecting that unit and that the owner and occupant should be notified, that the energy consumption in that unit is much above the average, that according to data from temperature sensors, the temperature fluctuates more than usual, suggesting problems in insulation?

3.3 Multiple parties, tools, and datasets

The work in construction projects is fragmented to numerous smaller work efforts carried out by different actors. The analysis in D2.1 about the high-level processes of renovation projects identifies around thirty parties that can produce information in a project, still excluding the multitude of different subcontractors and supply chain actors. Moreover, a new project consortium is generally established for each project. Since the collaboration relationships change from project to project, the parties do not have harmonized systems or processes with each other: every party has chosen its tools, systems, and practices independently, causing heterogeneity at the level of tools. This situation emphasizes the importance of commonly agreed formats, enabling all tools to share data and interoperate with each other.

Over the lifecycle of a building, large numbers of tools are used in the production and utilization of building data. Data is typically generated in a domain-specific or discipline-centered manner (Figure 7): The Architect creates an architectural model containing all architectural design data in one package, Structural Engineer creates a structural model that is a package of structural data, Project planners create the master plan maintained in a separate system, and IoT systems generate data that resides in IoT databases.

Apart from the ontology-level connections between different domains, there is another problem related to the disconnection of the different datasets at the instance level. For example, the various IFC models of one particular building do not typically have any references to each other; they simply reside in similar geometric coordinates. They can be visualized on top of each other or can be checked against each other for geometric collisions using special tools. Even worse, the plans created in a project seldom have any formal, machine-understandable references to the identifiers of the parts of the building – neither to spaces nor to elements. In such case the connections must be interpreted by the people using the models.

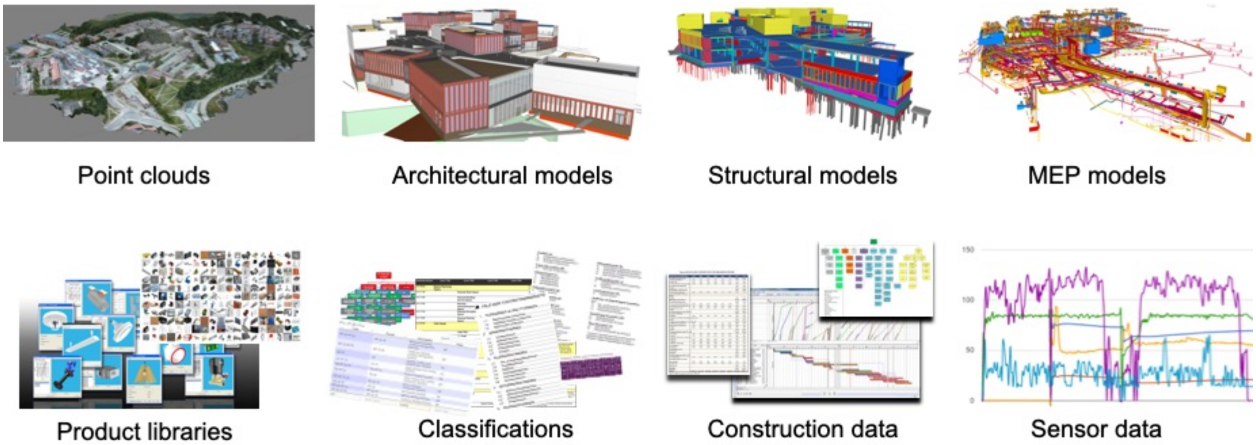


Figure 7: Examples of different kinds of building data

Different parties in a large renovation project can create hundreds of datasets in total, using various distinct tools and systems. To enable the interoperability between the systems within the renovation process, it would be essential that the datasets refer to the same entity instances in a consistent manner: for instance, the floors of a building should be the same instances or declared to be same, and not represented as disconnected instances in different datasets created by different tools.

The challenges for the interoperability at the dataset level are the following:

- Enable the consuming tools to access the information produced (technical)
- Share data in formats that the tools can understand and parse (syntactic)
- Use the terms that refer to entity instances in a same way across tools (semantic)

3.4 Evolution of information during execution

Over the execution of a construction or renovation project, new datasets containing BIM models, scenarios, plans, procurement data, or sensor data will be produced. The datasets can contain refinements of previously published data – such as more detailed models or plans – or new versions that override previous datasets. These datasets refer to the same entities than previous datasets and the references should be managed in a proper way.

In the design area the refinement of information has been conceptualized with the concept of LOD that can be "level of detail", "levels of development", or "level of definition" depending on the source and that can have additional dimension of refinement attached to it, such as "level of information", "level of geometry", and so on. There are established frameworks of LOD levels defined, for instance, in the USA (BIMForum, 2018), UK (BS 11292-1 and PAS 1192-2 and 3), Italy (UNI 11337, 2017).

These LOD framework specifications define requirements for the information contents that entities specified at those levels must satisfy; the level is intended to be specified for each entity in the model separately, which means that not all entities need to be described at the same LOD level. There are also more recent standards, such as ISO 19650 and CEN TC/442 that define the concept of a LOIN, "level of information need". The LOIN approach does not define global levels of design progression with level-specific requirements, but instead aims to provide a framework to locally specify what information contents is needed for the execution of particular activities.

In the planning domain, the specification of the work to be done generally progresses in a hierarchical manner. The overall, high-level plans (a project plan and/or a master plan) are created first, and then as the execution of the project starts, more specific and detailed plans will be created. The exact practices can vary significantly depending on the region, the companies involved, and the project in question. There can be phase plans (e.g., separately for design phase, demolition, frame construction, or indoor work), look-ahead plans that elaborate the tasks for a couple of months, week plans defined separately in a rolling manner for each week, and in some cases even daily plans for exact coordination between trades.

In addition to the planned refinement of information based on LOD/LOIN requirements or agreed-on planning levels, the *unplanned and unexpected events and changes* often cause the need to create new, revised versions of models and plans. These kinds of events or changes are ubiquitous in construction and renovation projects: new customer requirements, failed activity, delayed delivery, wrong type of resource, resource breakdown, an unnoticed conflict in different models, and so on.

For interoperability and automation purposes the datasets that refine or override previous datasets should maintain the same identifiers of entities that remain common to the datasets. This will enable

(i) smooth activation of new version of designs and plans,

(ii) linking of the common entities in successive versions or refinements to each other,

(iii) comparison of differences between versions, and management of changes in a targeted manner, and

(iv) automatic identification of modified parts of datasets.

If the identities are not maintained – or mapped to each other across versions – human interpretation and manual work is needed.

Consequently, the challenges from the information evolution perspective are the following:

- Refer to individuals in a same way in refined or overriding datasets (semantic)
- Maintain shared awareness what are the active datasets at each moment (pragmatic)

3.5 Relevant standards to comply with

The following standards are pertinent to information sharing and interoperability in renovation projects, and establish the background for the framework development in BIM4EEB:

- **ISO 16739 IFC**: The central data definition for exchanging BIM data is the Industry Foundation Classes (IFC) developed by buildingSMART International and published as a standard by ISO in committee: "*ISO/TC 59/SC 13 Organization and digitization of information about buildings and civil engineering works, including building information modelling (BIM)"* (ISO16739-1, 2018). IFC contains an elaborate schema to capture the details of the building design to the extent that geometric collisions between entities included in different models can be computed. IFC contains an object-oriented backbone structure of a building with an extensive set of property sets that make it possible to capture rich information contents in addition to the geometry of the design. The IFC schema has parts supporting the representation of other aspects of construction projects. Examples are representations for resources, processes, sensors, approvals, and so on. In practice, however, IFC is primarily used to capture the design of the building as expressed in a "conventional" BIM model. Most of the existing BIM authoring tools support the export of building designs to IFC files. It remains to be seen whether there will be BIM systems or related practices that allow useful specification of information about resources, processes, sensors, or approvals within BIM models, or whether existing systems for resource management, construction management, sensor data or document management would import or export IFC data.
- **ISO 19650 BIM/M**: The ISO 19650 standards series is named *"Organization and digitization of information about buildings and civil engineering works, including building information modelling (BIM). Information management using building information modelling.*" (ISO19650-1, 2018; ISO19650-2, 2018). The purpose and content of ISO19650 is different when compared to ISO16739 (IFC) While IFC is a meta-data definition, ISO19650 is a business process standard that defines requirements for the organization of information-related operations in construction processes. Processes are considered from the information management and sharing perspective. The standard defines a procedure where the *appointing party* sets requirements to *appointed parties* for information delivery and management. Requirement definitions cover both project information delivery and asset information management. The *lead appointed party* with other *appointed parties* in the *delivery team* should plan and implement information deliveries. The standard introduces concepts for the delivery team capability and capacity review, responsibility matrix, information status management, and specifies a common data environment. The concepts are refined in ISO19650-2 Part 2: Delivery phase of the assets (ISO19650-2, 2018) defining a BIM execution plan, a master information delivery plan and a task information plan. On the other end the ISO19650 refers to higher-level business process standards like ISO55000 Asset management or ISO9000 Quality management. In ISO19650 the information delivery contents are defined as a named *information container* that is a "*persistent set of information retrievable from within a file, system or application storage hierarchy"*. An *information model* is defined as a federated set of information containers and requirements.
- **ISO 21597 ICDD**: The ISO 21597 "*Information Container for Linked Document Delivery (ICDD)"* (ISO21597, 2020) defines an ontology for the representation of packages of several datasets and linksets that connect identifiers across the datasets. Such interlinked multi-model containers are needed in the exchanges of complex contents between project parties as well as in the documentation of contractual obligations and milestones in a project. ICDD also defines an approach for linking identifiers across different documents.

- **CEN/TC 442 BIM**: Within the European Committee for Standardization (CEN), the technical committee *CEN/TC 442 "Building Information Modelling (BIM)"* is working on the "*Standardization in the field of structured semantic life-cycle information for the built environment*." CEN/TC 442 is actively cooperating with ISO/TC 59/SC13.

The need to comply with relevant standards creates the following challenges:

- Utilize the BIM models as IFC data according to ISO 16793
- Organize information management according to ISO 19650
- Support exchange of containers of interlinked datasets according to ISO 21597
- Follow the standardization process of CEN/TC 442

3.6 Summary of the challenges

When reflected against the interoperability levels discussed in Section 2, the summary of the interoperability challenges of renovation projects outlined above is the following:

Technical level (addressed by WP4):

- Enable the consuming tools to access the information produced.

Syntactic level (addressed by WP3/WP4):

- Share data in formats that the tools can understand and parse.
- Support exchange of containers of interlinked datasets according to ISO 21597.

Semantic level (addressed by WP3):

- Types (ontologies)
  - Find or define the sufficient terminology (classes and properties) for different domains relevant to renovations, and the axioms to align the domains with each other
  - Utilize the BIM models as IFC data according to ISO 16793 and ifcOWL ontology
  - Organize information management according to ISO 19650
  - Represent interlinked datasets of ISO 21597
  - Follow the standardization process of CEN/TC 442
- Identifiers (cross-dataset/cross-version linking)
  - Use the identifiers in the same way across tools
  - Use the identifiers in the same way in refined or overriding datasets

Pragmatic (addressed by WP3/WP6/WP7)

- Organize information management according to ISO 19650
- Maintain shared awareness what are the active datasets at each moment (pragmatic)

4 Existing ontologies and other conceptual models

This section reviews existing ontologies that are relevant to cover the representational needs of renovation projects. There are existing ontologies for fundamental categories, reference ontologies for specific well-understood domains (such as time, units of measure, and sensor data) and domain ontologies that relate to the contents of renovation projects. Furthermore, there are many existing conceptual models that – even though not ontologies themselves – are directly relevant for potential development of new ontologies.

4.1 Existing general-purpose ontologies

4.1.1 BIM ontologies

The primary motivation behind the Linked Building Data research and development has been to increase the interoperability of various built environment data by mapping to Web of Data representations, with the added benefit of gaining access to advanced reasoning functionalities and tools.

Linked Building Data research has been based on BIM models exported into IFC, and to make IFC data accessible to Linked Data tools, ifcOWL, the ontology version of IFC has been specified together with the method to convert IFC to RDF conforming to ifcOWL (sometimes called ifcRDF). As there are different versions of IFC in use – for instance, IFC2x3 TC1 and IFC4 ADD2 TC1 - there are also multiple versions of ifcOWL. Rather than being a single ontology, ifcOWL is a set of principles for deriving an OWL version from any IFC version.

The ifcOWL conversion of the IFC schema from EXPRESS format to OWL is based on the following criteria (Pauwels, 2016a). The ifcOWL ontology (1) must belong to the OWL2 DL profile, (2) should match the original EXPRESS schema as closely as possible, and (3) aims primarily to support the conversion of IFC instance files into equivalent RDF files. It is thus less important that an ifcOWL-based RDF file can be written from scratch with an ontology editor.

There is an extensive list of conversion patterns described in Pauwels (2016a). The basic rules are straightforward: IFC entities to OWL classes, IFC properties to OWL properties, and so on. However, since EXPERSS and OWL are so completely different languages, there are some non-obvious conversion rules:

- The IFC properties whose values are simple datatypes (boolean, integer, and so on) are not converted to OWL datatype properties since that would push ifcOWL outside of the OWL DL profile. Instead, a “boxing” approach is used to capsulate data values within objects (Pauwels, 2016a).
- The procedural constraint types in EXPRESS – such as FUNCTION and RULE – do not have a natural counterpart in OWL and are ignored in the conversion. This is not as serious as it may sound, because the RDF files are created from properly exported IFC files, and these procedural constraints should already be satisfied by the original IFC data. However, if ifcOWL is newly created, it should be validated also against FUNCTION and RULE constraints. SHACL or SHACL Rules (Knublauch, 2017a, 2017b) are used to implement such constraints.
- Because all names in OWL are global and the property names of EXPRESS entities are local, it is not possible to directly convert a property name from IFC to the same property name in OWL. Rather, each property in IFC is converted to corresponding property of ifcOWL using the IFC name of the property appended by the name of the defining IFC entity. An IFC property with the name “p” defined in class with name “c” would have the name “p_c”. For example: timeStep_IfcRegularTimeSeries, or transparency_IfcSurfaceStyleShading.

The IFC schema contains areas that are not specific to BIM models, such as classes of IfcResource, IfcActor, IfcProcess and IfcApproval, and some that are only partially design-related such as IfcSensor. There does not seem to be any software available that would produce IFC files with resource, actor, or process models – that is, files without a BIM model but including those other entities. However, it would also be problematic to include these kinds of information into a same IFC file with a BIM model. There are going to be multiple BIM models of any building, typically at least an architectural model, a structural model, and an MEP model. Which one of these files is the correct one to include the descriptions of resources, actors, or processes? Moreover, approvals, process status or sensor data from the construction stage will be produced a long time after the design models have been completed. Should a new BIM model including this fast-changing data be re-exported periodically? For these reasons, in BIM4EEB the ifcOWL is used only to represent design BIM data.

So far ifcOWL has not gained widespread use. The reason has to do with the unfamiliarity of Semantic Web technologies for the software developers in general, but also with the complexity of ifcOWL itself:

1. Typographical complexity: The long property names makes the ontology difficult for humans to understand, and difficult to write SPARQL queries, for instance. This may sound as a somewhat trivial critique, when combined with the problem of structural complexity discussed below, it creates a steep learning curve and a practical barrier of entry for new users – a similar phenomenon that affects the success of programming languages as well.

2. Structural complexity: IFC4 ADD2 TC1 has around 800 entity types, 400 other type definitions, 500 property and quantity sets, and 1700 individual properties. The IFC-to-RDF conversion makes the resulting structures even more complex as different types of collections in IFC are represented with linked lists in RDF (Pauwels, 2016b). Farias (2014, 2015) presents the approach of simplify RDF graphs resulting from the IFC-to-RDF conversion. The simplifications address the representation of collections in RDF as sets – which are simple and natural to represent in RDF – when the order does not matter (Farias, 2014), and different kinds of property and relationship representations. The approach materializes shortcut properties over both the objectified relationships belonging to the class IfcRelationship, and over the IfcPropertySet/IfcProperty structures using the new ifcWoD ontology (Farias, 2015) as an extension to ifcOWL to capture these shortcut relations. The approach thus takes the semantics of the relations into account and enables more readable queries and faster query execution time.

SimpleBIM (Pauwels, 2016a) studies the simplification of BIM data as a more generic problem and reviews various approaches to simplify BIM models. It especially contrasts the Model View Definitions of IFC with Semantic Web technologies and suggest their further integration. The statistics about implemented simplification process indicate that it is possible to reduce the size of a model by an order of magnitude.

Zhang (2018) develops a BimSPARQL query language, i.e. a version of SPARQL adopted to the specifics of ifcOWL. It provides solutions to make SPARQL queries on IFC data more natural. First is RDF data materialization concerning schema-level shortcut properties and relationships (direct properties over IfcRelationship objects, such as schm:isContainedIn), instance-level shortcut properties for property sets and quantity sets (direct properties capturing the IfcPropertySet/IfcProperty data, e.g. pset:loadBearing), and lower-level geometry library for materializing geometry data and geometrical computations producing condensed geometries as Well Known Text strings similarly as in GeoSPARQL (Battle, 2011).

Second is the extensions functions of SPARQL queries, providing functions for a single product based on geometry (e.g., pdt:hasSpaceArea, pdt:hasGrossWallArea, pdt:hasOverallHeight, pdt:hasVolume), and functions for properties and relationships based on geometry data of multiple products (e.g., spt:touches, spt:intersects, spt:contains, spt:distance). The first two materializations are similar to those already discussed by Farias (2015). The third one is an adaptation from GeoSPARQL. The main contribution of BimSPARQL is in the area of the SPARQL extension functions.

Many of the above-described simplifications are presented in a comparison of different Semantic Web rule systems (Pauwels, 2017): SPIN/Jena, EYE and Stardog. In addition to showing that the only commercial solution, Stardog, outperformed the open-source tools, the comparison also helped to identify important questions when designing rule-based reasoning functionalities for ifcOWL (Pauwles, 2017):

1. How is data indexed, can query rewriting be applied, and what rule execution strategies are used?
2. How to combine forward chaining and backward chaining reasoning in an optimal manner?
3. How are the rules dependent on what kind of data is exported to the models?
4. What is the effect of using a triple store compared to main memory execution?
5. How does the performance depend on the number of output results produced?

Finally, Farias (2018) presents a rule-based method to extract model views from ifcOWL. This method provides a possibility to produce a subset of the ifcOWL-schema based on the application needs. Rule-based model views are more dynamic and simpler to define than the Model-View Definitions of IFC. A prototype of the view extraction system was implemented and successfully validated with a number of real IFC models. Further development needs to address the scalability to large models and large views.

3. Computational complexity: The ifcOWL ontology belongs to the OWL2 DL profile which means that various ontology reasoning tasks – such as Ontology Consistency, Class Expression Satisfiability, Class Expression Subsumption, and Instance Checking – are decidable. However, theoretically these tasks are still intractable – their complexity is according to Motik (2012) at least in the class NExpTime complete. However, the structural simplifications described above do not transform ifcOWL to a profile of lower computational complexity: it would require significant changes also to the way datatypes are used in ifcOWL (Hoang, 2015). For example, OWL2 EL and QL profiles do not allow the use of the xsd:boolean datatype. Even with structural simplifications, ifcOWL would remain in OWL2 DL, with associated computational complexity.

4. Multiplicity of ifcOWL ontologies: Currently, there is a separate ifcOWL ontology for each different version of the IFC schema. If there is a context where different IFC versions are used simultaneously (e.g., a renovation project), and if there is a need to link or aggregate data from converted ifcOWL-based RDF graphs, a problem arises. All the concepts in the different versions of ifcOWL ontologies are different from each other. IfcObject in ifcOWL based on IFC2x3 TC1 is a different concept than IfcObject in ifcOWL based on IFC4 ADD2 TC1. Despite the same label, these two concepts have a different URI. So far there has not been any alignments of these different ifcOWL ontology versions available.

Motivated by the problems of ifcOWL, and to simplify the utilization of building data in use cases less dependent on the geometric design, there is an active development within the Linked Building Data Group of W3C to create small and modular ontologies about different aspects of buildings. The central ontology in this modular suite is BOT - Building Topology Ontology (Figure 8). BOT contains only few classes and properties related to the spatial structure of a building (Rasmussen, 2017):

- *Spatial entities*: zone, site, building, floor, and space, where zone is a superclass of the other spatial entities.
- *Containment relations*: hasZone, hasBuilding, hasFloor, hasSpace, where hasZone is a superproperty of all other containment relations.
- *Elements*: All other things that can be contained in zones.
- *Inclusion relations*: hasElement which connects spatial entities to elements.

BOT is intentionally kept very simple as its goal is to be a linking point for other information. It is supposed to work together with other ontologies that enrich the spatial entities with other information.

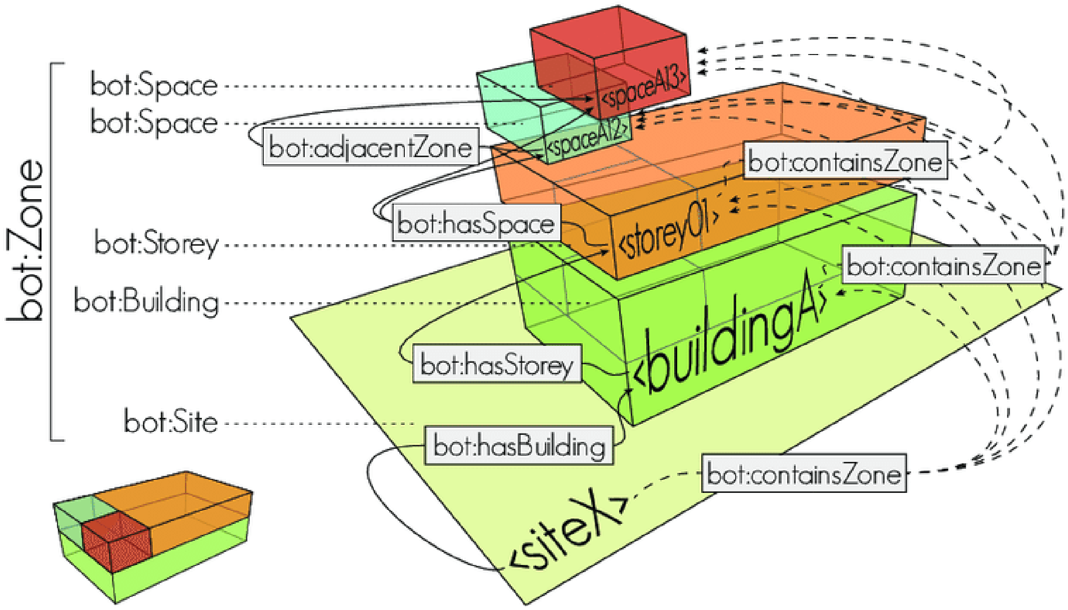


Figure 8: The concepts of Building Topology Ontology (BOT)

There are other modular ontologies in development in LBD Group at W3C:

- PROPS: Ontology generation from the property set definitions
- OPM: Ontology for Property Management (properties whose values change over the design)
- PRODUCT: Building product ontology
- FLOWS: An ontology of building systems (the concepts of flow systems)
- OMG: Ontology for managing geometries (linking objects with a geometry in any geometric format)
- FOG: File format ontology for geometries (to specify additional geometric formats)

It is worth noting that the BIM data can be represented simultaneously according to both ifcOWL and BOT ontology. That is, there can be simple view to the data (through BOT) and at the same time access to the full complexity and all details of a BIM model (through ifcOWL) if needed.

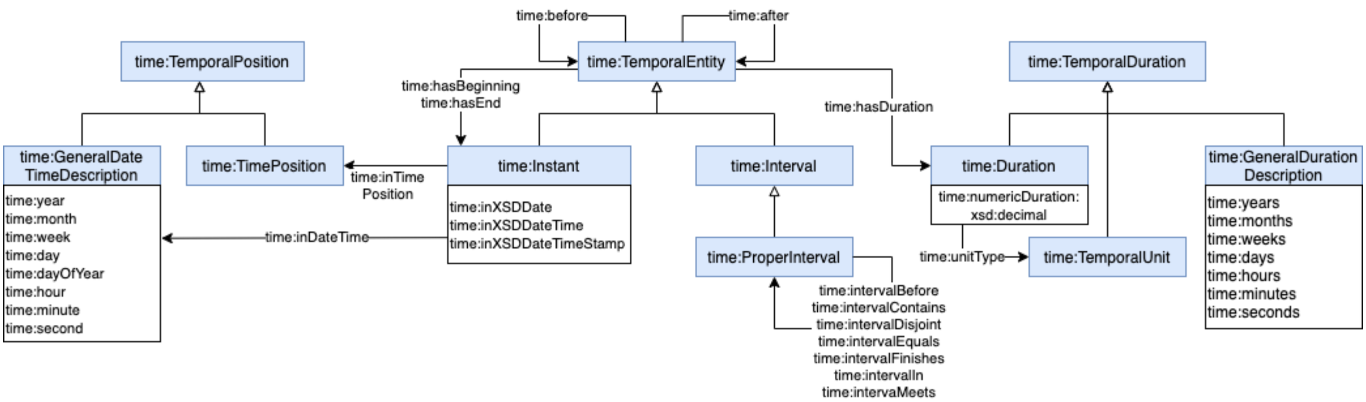


Figure 9: Basic temporal concepts in OWL-Time

4.1.2 Time and temporal relations

OWL-Time (Cox, 2017) is a time ontology that is a W3C candidate recommendation. It defines the concepts for representing ordering relations between instants and intervals (before, during, and so on), information about durations, temporal positions, as well as general date-time information. Time positions and durations can be given with respect to conventional calendar and clock, using alternative calendar systems, or even in Unix-time.

OWL-Time is especially useful for representing time information about human-scale activities; it allows to express time at different levels of granularity, for instance, in an hour or day granularity.

The underlying technical idea in OWL-Time is to represent all temporal entities as intervals with start and end points. Instants are just intervals whose start and end points are similar (Figure 9). Time intervals can relate with each other with 13 qualitatively different ways based on the interval algebra of Allen (Allen, 1983). OWL-Time provides the basic representation for temporal reasoning related to activity precedence, sub-activity relations (work breakdown structures), and potential concurrency of activities. As OWL-Time is the W3C Recommendation and only widely known time ontology, it is the obvious choice to align the temporal concepts of BIM4EEB.

4.1.3 Units of measurement

The representation of numeric values augmented with units of measure is a cross-cutting need in decentralized systems where technical data is exchanged between parties and constituent systems. There have been many large efforts to create a common representation for units of measure:

- QUDT (Quantities, Units, Dimensions and Types) has an extensive class-based representation of units, whose definition has been led by NASA and TopQuadrant. QUDT 2.1 was released in the fall 2019 and contains comprehensive semantic descriptions of quantity kinds (over 800), units (over 900), disciplines and dimension vectors.
- OM (Ontology for units of Measure) is another unit ontology with OWL2 definition. It is comparable to QUDT but has different modelling choices and coverage of application areas than QUDT. OM has been defined around 2010 but is still maintained.
- UCUM (Unified Code of Units of Measure) (Schadow, 2017) is not an ontology but a language for specifying unit names and relation in a simple syntax and covering all existing unit systems.

UCUM can be used in the combination of with CDT (Custom datatypes) which is a proposal of a way to attach UCUM strings to RDF literals: the UCUM strings encode the units and CDT is used to attach that string to RDF literals. Since units are not represented explicitly in RDF, the possibilities to reason about units or make conversions between them will not be available.

After the publication of version 2.1, QUDT provides extensive vocabularies for units and quantity kinds, and is the most advanced and comprehensive ontology of unit information. It is therefore the obvious alignment choice in BIM4EEB.

4.1.4 Provenance of data

The provenance of a digital object represents its origin, describing the people, tools and activities involved in producing and delivering an object. PROV-O (Lebo, 2013) is a W3C recommendation for a provenance ontology. Provenance can be used to represent the relations of information entities with each other (what is an information entity based on) and with actors (who produced the information entity), and relation of physical entities to information entities. This kind of information is essential for modelling the design stage of a renovation/construction project and the relations to later stages of the project.

The central concepts of PROV-O are shown in Figure 10. The goal is to capture the relations between information entities, activities that process them and the actors involved in those activities.

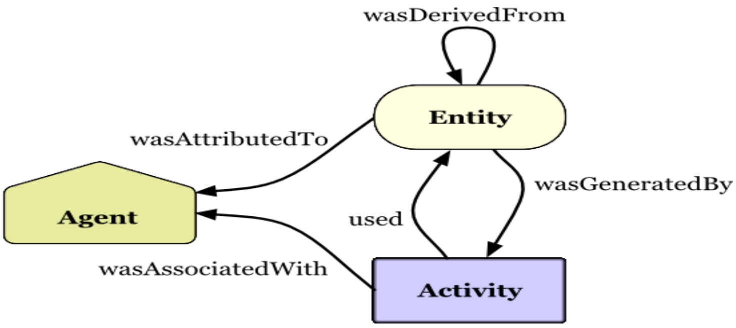


Figure 10: Core concepts of the PROV-O ontology

Provenance information can be used to represent the derivation of BIM models from each other - e.g., creation of structural or MEP models using the architectural model as reference, or creation of higher level of detail models from previous ones. It can also represent the information origin of physical entities in the construction stage. The activities in provenance are part of the information process of a construction project, covering the design, procurement, and monitoring activities.

4.1.5 Sensor data and sensor systems

There are many existing interfaces, conceptual models and ontologies to represent sensor data gathered from IoT or BAS systems: SensorML (OGC, 2000) focusing on measurement processes and transformations, and provides XML-based interoperability API; Semantic Sensor Network Ontology (W3C, 2005, 2017) that defines essential concepts of SensorML, such as sensor and actuator systems and observations; IfcSensor, the sensor classes in the IFC (bSI, 2013) such as IfcSensor, IfcController, and IfcPerformanceHistory; OMI-ODF, the Open Messaging Interface/Open Data Format (OpenGroup, 2014) that provide a uniform XML/SOAP (or REST/JSON) API to access or subscribe to sensor data; SAREF, Smart Applications Reference Ontology (ESTI, 2015) that defines sensing/actuating for smart appliances and applications initially in the home domain; SensorThings (OGC, 2016) that provides a REST/JSON-interface for sensor data based on earlier OGC models; Brick Schema, A Uniform Metadata Schema for Buildings (2017), giving ontology of entities and systems of commercial buildings from the perspective of BAS; WoT, Web of Things (W3C, 2018-2019) that defines uniform access interfaces to IoT devices (sensors/actuators) in different deployment architectures, providing JSON-based APIs and scripting interfaces focusing on JavaScript.

There is an important difference between interfaces that allow access to sensor information and ontologies that can be used to create a sensor model. A sensor model is a description of a sensor system including metadata about units of measurement, warning, and alarm limits, and positions of sensors which is linked to BIM models. The interfaces could be used in addition of the sensor model to provide access both to real-time values or streams, and sensor histories over time.

The most established ontology in the sensor area is SSN/SOSA. It is based on SensorML. It has been jointly prepared by W3C and OGC.,Its second version has already reached the status of a W3C-recommendation. Also, Saref has a large community and an established standardization process within ETSI. These two ontologies should be the primary targets for alignment within BIM4EEB.

4.1.6 Actors and organizations

There are numerous stakeholders involved in renovation projects whose relations to activities, information objects and so on need to be represented to properly manage data sharing, operational relations, and communication actions between them.

Friend-of-a-friend (FOAF) (Brickley, 2014b) is a well-known ontology for the basic concepts of agents, persons, and organizations. It was originally targeted as an open ontology for representing social networks in a decentralized manner. That is, each user could describe his or her social relationships in an own FOAF profile, a RDF file utilizing the FOAF ontology. FOAF has been used in other ontologies such as PROV-O or W3C Organization ontology (Org) (Reynolds, 2014) to represent agents and persons. In addition, FIBO (Financial Industry Business Ontology) defines legal concepts that are needed for the definition of ownership and corporations.

The combination of FOAF, Org, PROV-O and FIBO – as shown in Figure 11– can form the alignment basis for actor and organization terminology in BIM4EEB.

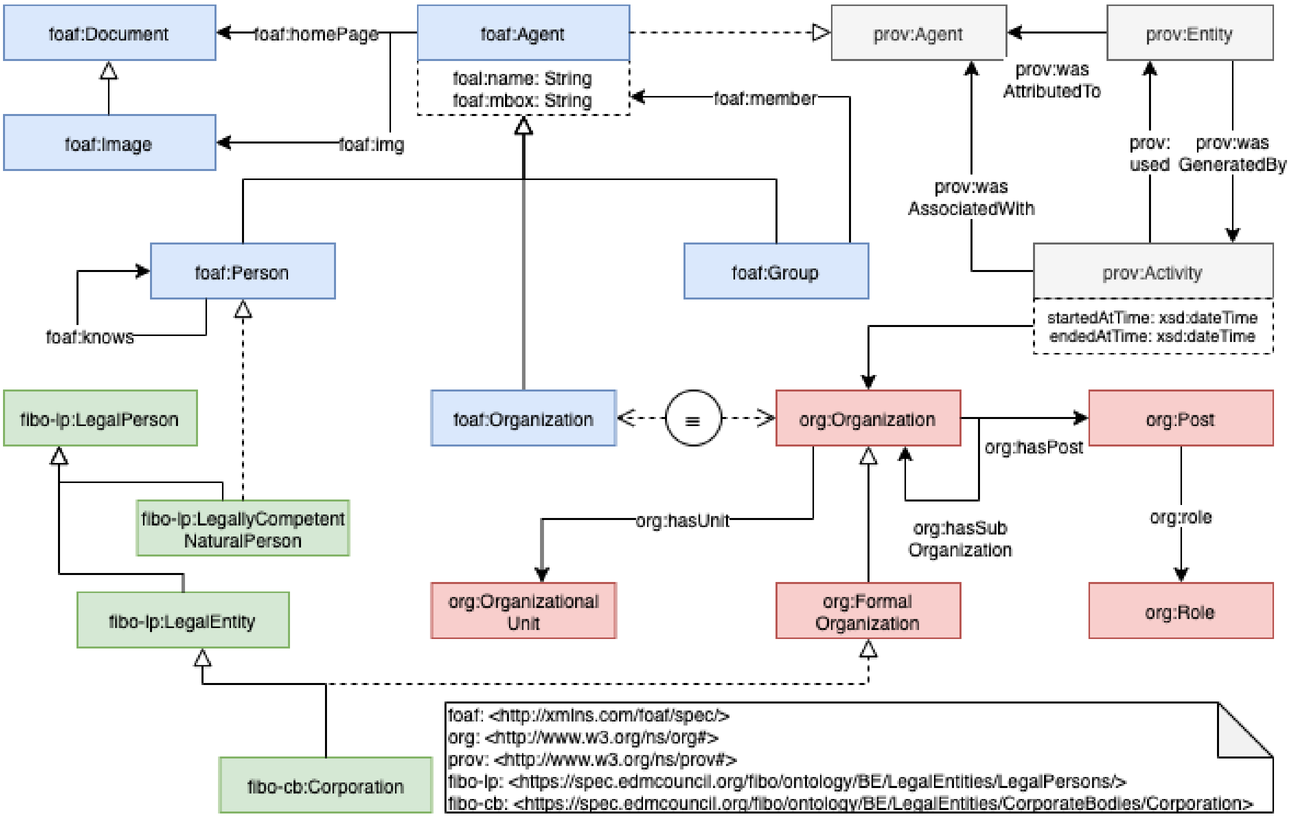


Figure 11: The relevant concepts of FOAF, Org, PROV-O and FIBO

4.1.7 Fundamental categories

The use of a top-level ontology is a way to integrate many different ontologies together in a principled manner, as it can provide the fundamental categories of entities that are represented in different reference ontologies or domain ontologies. ISO/IEC:21838 defines a *top-level ontology* as:

> *top-level ontology = def. an ontology [of] categories shared across a maximally broad range of domains where ontology = def. a collection of terms, relational expressions and associated natural-language definitions together with [ ] formal theories designed to capture the intended interpretations of these definitions.*

There are three well-known top-level ontologies:

1. BFO – Basic Formal Ontology (Smith, 2015) (Arp, 2015).
2. DOLCE – Descriptive Ontology for Linguistic and Cognitive Engineering (Borgo, 2009).
3. SUMO – Suggested Upper Merged Ontology (Pease, 2002).

Since SUMO is not a pure top-level ontology but also contains domain level concepts, and since BFO and DOLCE are more widely used, the subsequent discussion is limited to the latter ones. Both BFO and

DOLCE have much in common in their models. They both are based on three fundamental dichotomies: (1) *universals* and *instances,* (2) *dependent* and *independent* entities, and (3) *continuants* and *occurrents* (or in DOLCE, respectively, *endurants* and *perdurants*).

The philosophical difference between DOLCE and BFO is that BFO is based on the view of *ontological realism*: ontology development is considered similar to the development of scientific theories, which aim to discover universal laws of nature. They study "what is it that is general in nature" (Arp, 2015). Furthermore, BFO adopts the view that so-called *universal categories* (e.g. person, ball, redness, growth) exist in the external world. The implication of ontological realism is the focus to identify regularities (concepts and relations) in the external world, and not in the data describing the external world. Data is always based on some perspective to the phenomena in the external world. Therefore, an ontology should rather be based on the phenomena that data is describing. In comparison, DOLCE aims to capture the ontological categories lying behind natural language and human common sense. Its categories are regarded as "cognitive artefacts ultimately depending on human perception, cultural imprints and social conventions" (Oltramari, 2002). While the philosophical validity of this view can be argued, Smith (2019) points out that it can be "detrimental to the goal of using a top-level ontology to promote interoperability among domain ontologies, since the cultural imprints and social conventions vary so widely".

As a part of the Industrial Ontologies Foundry development effort, an evaluation of top-level ontologies was carried out (Smith, 2019), focusing on BFO and DOLCE, and in the end BFO was selected. Main reasons related to the broad and active use and maintenance of BFO (Smith, 2019). It is the top-level ontology of many successful ontology efforts: the Gene Ontology (GO) and the Open Biological and Biomedical Ontology (OBO) Foundry build around GO (Smith, 2007). There are at least 300 different ontology efforts using BFO as the top-level ontology. The concepts of BFO are shown in Figure 12.

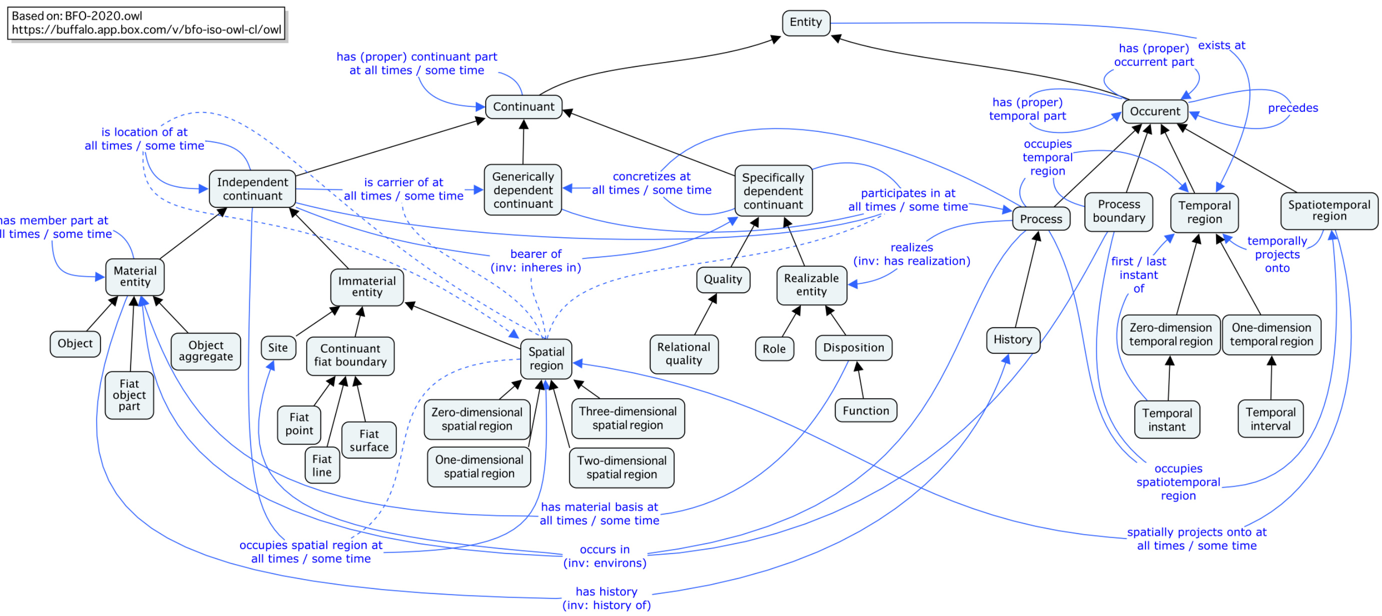


Figure 12: The IS-A structure of BFO-ISO

At the first level all entities that may exist are divided into *continuants* (things that continue to exist) and *occurrents* (things that have a start and end). Continuants are divided into *independent continuants* (that can exist independently), further refined to *material entities* (having subclasses *object*, *object aggregate*, and *fiat object part*) and *immaterial entities; specifically dependent continuants* (whose existence depend on some specific other continuant) such as *qualities* and *realizable entities* such as *roles* (external realizations) and *dispositions* (internal realizations), the latter of which can be used to represent *capabilities* and *functions;* and *generically dependent continuants* (whose existence depends on multiple, but at least one other continuant); examples are *information content entities* (e.g., a BIM model) whose existence depends on at least one *material information bearer* (e.g., a computer file). Occurrents can be divided to *processes* that can represent activities, tasks, and histories, and *temporal regions*, that are further refined to *temporal instants* and *temporal regions*.

The ISO/IEC ISO/IEC:21838 standardization process of BFO includes its axiomatization in OWL2 and in Common Logic, a computer-friendly format for first order logic.

One large application of BFO is in the Industrial Ontologies Foundry (IOF, 2019), a collaborative effort to create interoperable ontologies for industrial and engineering domains. It was initiated by the National Institute of Standards and Technology (NIST) in the USA and held its first workshop in 2016. The objective is to create core and reference ontologies for the whole domain of digital manufacturing to advance data interoperability. The effort has a large number of participating companies and organizations, mostly from the USA and Europe. From the perspective of BIM4EEB, IOF is an interesting effort since there are clear overlaps between manufacturing and engineering, on one hand, and construction and renovation, on the other. The increasing use of prefabrication in construction is bringing more and more manufacturing activities into the realm of construction and renovation projects. Moreover, the problems of planning and scheduling, while having somewhat different emphasis and different pain points, also bear resemblance across these domains. In 2019 the IOF Reference Ontology for Industrial Maintenance (ROMAIN) has been published (Karray, 2019). The Ontology of Commerical Exchange (OCE) has been drafted (Vajda, 2019), to represent commercial exchanges as activity patterns between agents, also utilizing the Ontology of Document Actions (d-acts) (Almeida, 2012) and Information Artifact Ontology (IAO) (Ceusters, 2015) to represent the documents involved in those exchanges (request for tenders, tenders, contracts, invoices, receipts, reclamations) and actions carried on the documents, such as signing.

4.1.8 Multi-context data and property metadata

During a renovation project many important areas of data – designs, construction plans, resource management plans, progress data, and so on – will evolve over time. This evolution can be gradual refinement of information that has been planned in advance or can happen as a reaction to unexpected events, such as changes in requirements, or discovery of unusual structures at the site or in the existing building. In both cases, new and revised versions of designs will need to be created. From the perspective of data management, this creates a situation where the same properties of one distinct object can have different values associated with different contexts. For instance, geometric coordinates related to a routing of a pipe can be different after a design change than what they were before the change. Moreover, the run of a pipe may also be different, perhaps due to the addition of new flow segments to the pipe.

RDF or OWL do not directly support the representation of dynamic evolving multi-context data. It should be noted that while in an RDF triple <subject, predicate, object> the subject and object refer to instances, the predicate refers to a type. The result is that there is no direct way to represent additional triple-specific information about the predicate. This creates challenges in situations where there can be different values in different contexts or time points, or where data has different origins, as is typical for situations when information from multiple different systems is integrated.

Renovation projects are an example of this kind of domain. Even a superficial understanding of construction and renovation management reveals that the execution times of activities have typically planned and actual values, and one most important functionality is to compare them to each other. A planned value cannot simply be overwritten by a corresponding actual value when execution progresses. Instead, both values need to be preserved and their relationship maintained. Similarly, in renovation projects the design-related variables can have values in as-is, as-designed, and as-built contexts. Moreover, depending on the design management practices employed in a project, the designs can go through several distinct levels of detail or development. Furthermore, in the planning domain the situation can be even more complex, since there can be many partly overlapping lower level plans.

A short overview of the representational options for multi-context data is in the following.

Extension of the ontology with definitions based on Qualified Relation pattern: This approach means that the ontology defines two versions of each property: a context-insensitive property (usually simple) and a context-sensitive property with objectification. This is called a *Qualified Relation* pattern (Dodds, 2012). An example of an ontology that uses the Qualified Relation pattern is PROV-O (Lebo, 2013):

```
@prefix xsd:  <http://www.w3.org/2001/XMLSchema#> .
@prefix prov: <http://www.w3.org/ns/prov#> .
@prefix :     <http://example.com/> .

:post7654 a prov:Entity;
  prov:wasGeneratedBy :publicationActivity0123;
  prov:qualifiedGeneration [ a prov:Generation;
    prov:activity   :publicationActivity0123;
    prov:atTime     "2019-11-22T12:34:56Z"^^xsd:dateTime;
    prov:atLocation <http://dbpedia.org/resource/Madrid> ].
```

Here the generation information is presented both with the unqualified property prov:wasGeneratedBy and the qualified property prov:qualifiedGeneration, whose value is an objectified relation object of type “a prov:Generation”. The generation object can provide a richer representation of the generation event, such as specifying its time and location. Even though the PROV-O examples just address property metadata, it is easy to see that once a relationship is objectified using the Qualified Relation pattern, there is a possibility to include all kinds of context dependent values to the relationship objects.

BFO-ISO uses a similar approach to model temporary predicates, that is, predicates whose values can change over time. The advantages of this modelling approach are: (1) simple (unqualified) relations can be supplemented with richer (qualified) relations, (2) the modelling works with OWL, and (3) the model is based on a well-known pattern. Disadvantages are (1) the solution is conceptually verbose; the same underlying problem is solved over and over again by defining additional properties and classes, (2) users need to learn which unqualified and qualified relations correspond to each other, and (3) the size of the ontology grows significantly and becomes more difficult to understand.

General objectification of properties: This approach is defined in the Ontology for Property Management (OPM), (Rasmussen, 2018). Instead of defining a separate qualified version of each unqualified property, there is just one uniform mechanism to provide qualified properties. Each object can have a opm:hasProperty property containing all qualified property objects (Figure 13). An opm:Property object refers to the unqualified property which contains static metadata about the property, and a number of opm:PropertyState objects. Each property state can be related to a particular context or time period, and it can also include information about the origin of the value.

OPM allows to describe the properties in three different levels: (1) direct (unqualified) property, (2) objectified property, allowing the description of additional static metadata of a property, and (3) objectified property states, allowing the description of the values of the property in different contexts, possibly with additional metadata such as their origin.

Some features of OPM limit its general use in recording context-sensitive data. Firstly, the model is defined only for datatype properties. However, a component may be subject to structural changes from one levels of detail to the next, and it can have new parts defined at the next level. Secondly, OPM puts emphasis of maintaining the current value of a property that has multiple property states. This notion is limited to specific use cases where properties go through a sequence of changes, out of which the last one is the relevant one. However, when there are multiple contexts that are relevant simultaneously – such as planned value and actual value – the concept of current value does not make sense.

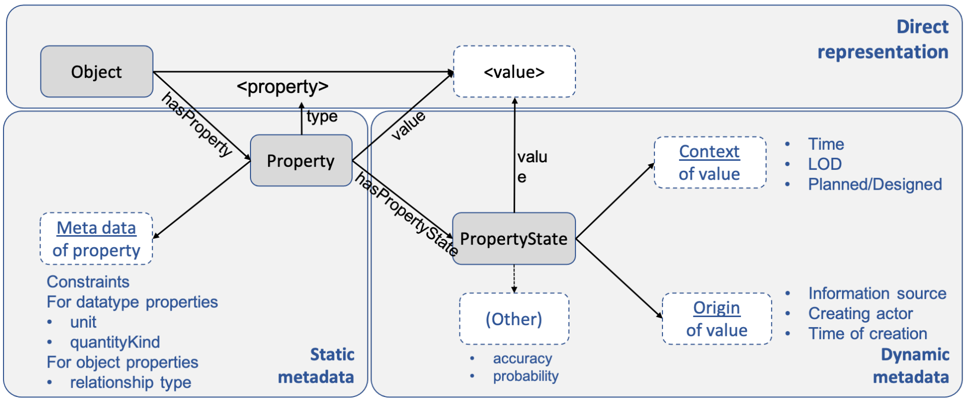


Figure 13: General objectification of properties

The advantages of the approach: (1) Flexible approach that can work with any property and any number of static and dynamic contexts, (2) combines simple and complex representations, (3) can be used with existing ontologies, where any unqualified properties can be extended with qualified counterparts, and (4) does not clutter the ontology with verbose definitions. Disadvantages: (1) OPM is only defined for data properties, and (2) OPM's current value maintenance is a mechanism that is not relevant in all use cases.

Context-related sub-properties: Possibly the simplest approach to deal with different contexts is to define different properties to the ontology to represent values in them. For example:

```
:startTime a rdfs:DatatypeProperty .
:plannedStartTime rdfs:subPropertyOf :startTime .
:actualStartTime  rdfs:subPropertyOf :startTime .
```

In this approach values can be represented in multiple predefined contexts but without any other metadata about the property (e.g. timestamp or origin). However, there is a need to define a new property for each combination of interesting properties. Moreover, in case of dynamic context (i.e. allcontexts are not known in advance) this approach will not work. The advantages are: (1) Simple and efficient. Disadvantages are: (1) Allows no other metadata and (2) possibly a large number of new property definitions would be needed.

Context-related singleton properties: This is an instance-level version of the previous approach. Instead of defining subproperties of a property, a new identifier is created and declared as a singleton property of *p*. Singleton property only applies between a particular subject-object pair. Since the singleton property has its own identifier (URI), all other information can be associated with it. For example:

```
:startTime a rdf:Property .
:plannedStartTime-A1-T1 rdf:singletonPropertyOf :startTime .
:actualStartTime-A1-T2  rdf:singletonPropertyOf :startTime .

:A1 :plannedStartTime-A1-T1 :T1 ;
:plannedStartTime-A1-T1 :hasContext :planned ;
:plannedStartTime-A1-T1 :hasOrigin :weekplan-26 .
```

Singleton properties have been proposed by Nguyen (2014), and the extensions to the RDF semantics are outlined. Singleton properties solve many problems of the subproperty approach. As singleton properties are not defined in an ontology, they work with dynamic set of contexts, and they increase the size of the RDF graph only modestly. Advantages: (1) Simple, flexible, and efficient approach and (2) does not clutter the ontology with additional definitions. Disadvantages: (1) Non-standard representation and (2) does not separate between static property metadata and dynamic property state metadata.

Context-related property namespaces: Similar properties can be defined in different namespaces corresponding to different contexts. This approach has been used in Common Core Ontologies (CCO)

(Rudnicki, 2019). One of the ontologies in CCO is called the Modal Relations Ontology (MRO) whose sole purpose is to redefine in a new namespace all properties of the other ontology modules in CCO. The properties redefined in MRO represent the planned or as-designed values, and the properties in the original ontologies represent the actual values. Specifically, there are two namespaces:

@prefix cco: <http://www.ontologyrepository.com/CommonCoreOntologies/>
@prefix mro: <http://www.ontologyrepository.com/CommonCoreOntologies/ModalRelationOntology/>

As an example, there could be two properties mro:hasEnd and cco:hasEnd to represent the end time of an activity. The former property would be the planned end time, and the latter the actual end time. This approach would require static knowledge about what the possible "modes" or contexts are. If there will be more contexts, a separate namespace would need to be created for each one of them and all properties defined also in other ontologies should be redefined there. For instance, if there is a need to represent values for different LOD levels, a separate namespace for each level must be defined and all properties redefined in each of these namespaces. Moreover, information about a context is hidden in the URI of an object and general queries such as "what values in each context differ from those of the previous context" cannot be answered easily. Either there needs to be queries that encode the relations of contexts ("differences from LOD100 to LOD200", or "differences from LOD200 to LOD300") or the inspection of the structure of URIs, which is not considered as a good practice.

Object for a statement: This is the classical approach used in RDF called reification, that is, making an object corresponding to a statement. The reification vocabulary is described in the RDF Schema specification (Brickley, 2014a). Reification resembles the objectified properties approach in Figure 13 with the difference that instead of having one object for each property, there would be a statement object for each property-value pair of an object, and therefore separate PropertyState objects are unnecessary. Reification has been considered a more wasteful approach for adding metadata to properties (Ngyuen, 2014) than for instance singleton properties, and it has gained only limited popularity even though being presented in the first RDF specifications.

Graph for statements: The basic idea is to encapsulate a statement or a set of statements into a named graph. Since the named graph has its own URI, it is possible to make other statements about the graph to record its context, origin, or other metadata. There are various proposals about graph-based metadata called nanopublications (Groth, 2010) or Ovopub (Callahan, 2013). The main idea is shown in Figure 14.

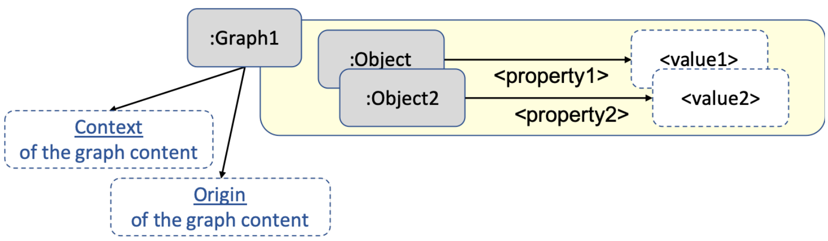


Figure 14: Graph-based metadata

The metadata can be represented for larger fragments of a graph, which makes this approach space efficient. The structures can also be nested which increases the expressiveness of the approach.

4.2 Existing ontologies or conceptual models for D3.4

The execution of a renovation or construction project requires careful management of activities. There are questions of how activities can be decomposed into more elementary activities. Further questions are: (1) What are the resource requirements and durations of activities? (2) What are the preconditions to start an activity (availability of information, labour, location, materials, equipment)? (3) Who is responsible for the execution? (4) When is an activity completed? (5) What effects or products does the activity have? The preconditions and effects of activities create temporal precedence constraints between them. The decomposition and precedence relations establish a network structure amongst the various activities.

4.2.1 Construction management

Smooth and efficient management of the construction workflow has huge potential to improve the productivity of renovation projects. It is estimated that less than one third of work on construction site is value-adding. Furthermore, based on results for construction site indoor positioning, it appears that the actual work has little relation to construction plans. Recently, the advances in technology – such as BIM modelling and sensor-based monitoring and planning methods, such as Location-Based Management System (Kenley, 2010, Seppänen, 2014b), Takt-time planning (Seppänen, 2014a), or lean construction management practices (Koskela, 1999, 2000, Seppänen, 2010) or the Last Planner™ (Ballard, 2000) carry big promises to improve construction management.

BIM models provide more detailed information about the goals of construction, individual elements with identifiers, relations of elements, and geometries that can be used in visualization and user interaction.

Sensor data can help in accurate and automated progress monitoring of renovation activities. The connection of activities to elements defined in BIM models is important since physical elements (positioning), spaces (occupancy/motion detection) and resources (positioning and motion/activity detection) can be directly monitored and progress of activities can be derived from these observations.

Processes in renovation projects usually cross organizational boundaries. When faster execution and shorter throughput times are pursued, more coordination between different actors are required and more accurate information about activities need to be exchanged. As more data can be acquired from BIM models, sensor data, and so on, the information about activities needs to be machine understandable.

The use cases of activity and workflow information range from planning to coordination of activities:

1. *Master planning*: Creation of the top-level plan and schedule for the project;
    - Input: Renovation measures, planning principles and strategies, budget constraints, data from past projects;
    - Output: Master plan including a work breakdown structure (WBS), target schedule, team roles, and responsibilities.
2. *Task planning*: Defining the concrete activities and their ingredients;
    - Input: Design models, existing plans, team composition, estimates of activity size;
    - Output: Location breakdown structure, executable task definitions, including activity duration and costs, and connections to building objects, locations, equipment, and agents.
3. *Progress coordination*: Guiding the execution, and gathering data, and updating models;
    - Input: Existing plans, status/sensor/procurement data, task details;
    - Output: Enhanced plans and designs, issues, deviations, needs for replanning.
4. *Plan revisioning*: Creating new plans and activating them for execution;
    - Input: Existing plans, evolving data about situation, issues, deviations, task details;
    - Output: New plan version and updated coordination structures to activate it.
5. *Week planning*: Creation and activation of a plan for the next week;
    a. Input: Existing plans, evolving data about situation, issues, deviations, task details;
    b. Output: New plan version and updated coordination structures to activate it.

6. *Stakeholder coordination*: Notifying stakeholders about activities impacting them;
    a. Input: Week plans, location breakdown structure, location-stakeholder connection.

Applicable ontologies, models and languages are reviewed in the following.

Project management ontologies and models. There have been some early efforts to define ontologies for project planning and scheduling (Tate 1996, Smith 1996) as well as to construction management (Wetherill 2003). More recent proposals for project management ontologies are PROMONT (Abels 2006) and IT-Code (Lui 2009). Even though all these ontologies contain concepts for activities, resource, assignments, and so on, none of these efforts appear as a directly useful basis for a construction management ontology. The ontology definitions are not available online or they are based on ontology languages preceding the definition of OWL. Moreover, there is a lot of conceptual package that comes from their specific viewpoints in many cases combined with some idiosyncratic terminology, that would make the ontology unnecessarily complex and difficult to adapt to varied environments. An example is the tendency to use different concepts for activities at different levels - such as project, phase, task, activity - which bring little additional value to the ontology but may create unnecessary incompatibilities with other conceptualizations.

Outside of the realm of ontologies, the de facto model for exchanging project information is the XML schema of the MS Project Data Interchange (MSPDI). Another similar XML schema is PMXML defined by Pacific Edge and implemented by Primavera, although it has not managed to challenge the position of MSDPI. The basic concepts of MSPDI are shown in Figure 15. All the objects shown in Figure 15 have tens of properties, also related to data management and user interface aspects. Full support for those properties is outside the scope of an interoperable ontology. It is important to notice that the only activity flows that can be represented with these concepts are the resource-type flows. The schema lacks the concepts and relations through which the activities can be connected to building objects, locations, materials, documents, models, etc. Since sensor observations can typically only be made from these ingredients of activities, the lack of connections to the ingredients also limits the possibilities to associate sensor data to activities themselves. For these reasons the MS Project schema cannot be regarded as an applicable basis for a construction management ontology; it is better suited to activities at aggregate levels, such as to the representation of a master plan of a renovation project.

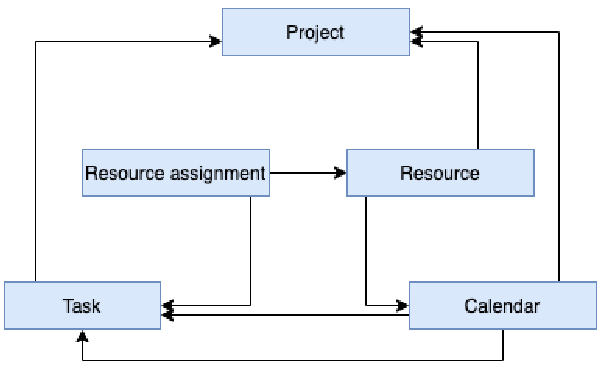


Figure 15: Concepts of the MS Project model

Planning ontologies, models and description languages. There is a long research tradition within the field of AI and robotics to develop automatic planning systems, the first important results were the STRIPS system (Fikes, 1971). Planning is regarded as generation of plans as activity sequences or more generally as sets of activities bound together by a set of constraints for activity orders and variable bindings.

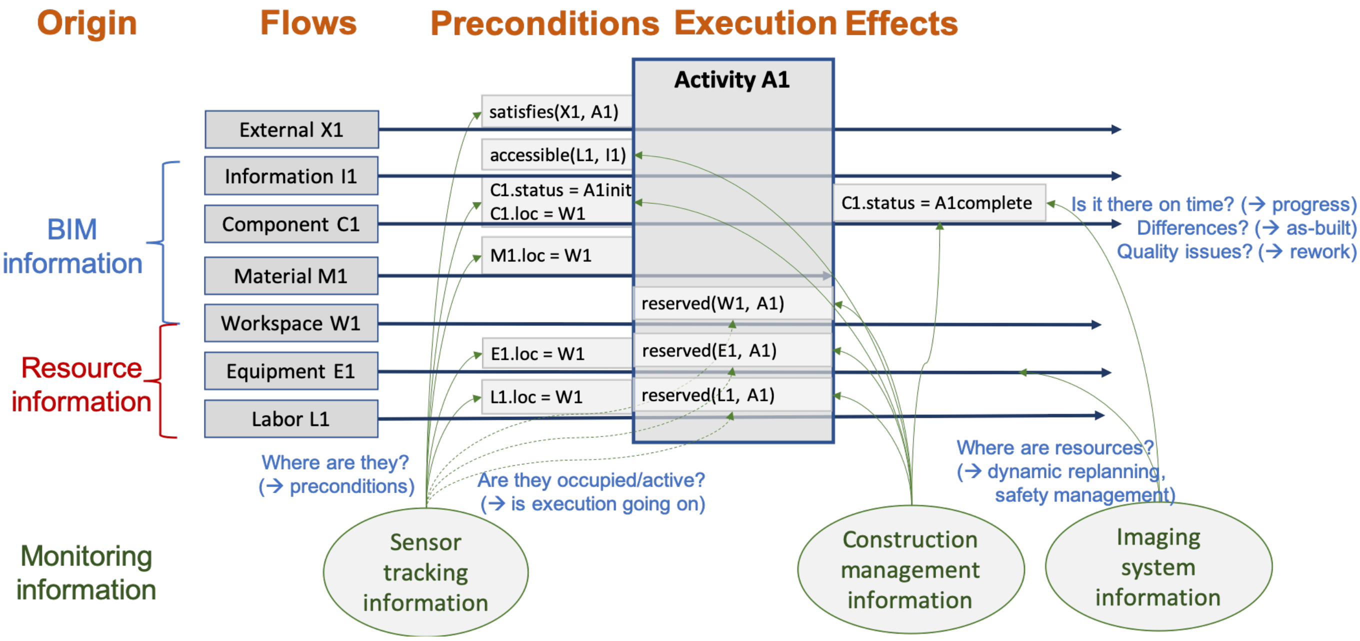


Figure 16: Activity flows with the preconditions and effects

Over the time AI planning research has led to the development of a broad variety of planning algorithms, including required representational constructs (documented in (Ghallab, 2016)). The planning system can be roughly divided into state-based planners, partial-order planners and hierarchical task network (HTN) planners. Additionally, there are many variations and algorithms related to each of those approaches. The preconditions and effects connect the activities to other domain entities, such as connecting the transport activity to the shipment. Moreover, proper states of the shipment are specified, i.e. before starting the transport activity the shipment needs to the in the location of origin (?from) and in the end of the shipment process it is in the target location (?to). Figure 16 shows how different activity-flows (Koskela, 2000) (Garcia-Lopetz, 2017) relate to activities.

Numerous ingredients need to be available for any activity to become executable. These are referred to as flows. The information about flows comes from BIM models or other management systems. Physical ingredients (e.g. labour, materials, equipment) need to be at the location of activity execution before it can start. Preconditions and effects can be modelled based on AI planning formalisms. It should be noted that states are a more natural representation for the relations between activities and domain entities than the input-output model – adopted in many manufacturing-oriented models and also in the IfcProcess entities of IFC (Figure 21) – since it is a more general representation that applies to all types of activities. Input-output representation mainly applies to assembly type activities and to closed settings such as in a manufacturing plant (where the representation of state information such as the location of a component can be ignored from the process point-of-view). For instance, it is unnatural to represent the transport activity (above) with input and output relations, since its output is the same shipment than its input. The essential semantics of the transformation that the transport activity makes, is a state change to the shipment, that is, a change in its location.

The encoding of preconditions is also useful in the monitoring of the construction process execution. Using status tracking tools or indoor position systems – or ultimately even imaging data. Thus, it can be automatically predicted or checked whether different preconditions hold at the start of the activity execution. If the prediction shows that it is not going to hold, an issue can be raised to take care of the problem. Consequently, although the precondition/effect representation has been adopted as a basis of automatic planning systems, it has additional uses in e.g. plan monitoring or interactive plan repair.

Numerous publications document activities dedicated to the development of a shared planning ontology (Tate (1996)) and common planning domain description languages (e.g. PDDL1.0 (Ghallab, 1998) or PDDL2.1 (Fox, 2003)). The implementation of relations between activity representations from AI planning and activity-flow modelling in lean construction (Garcia-Lopetz, 2017) need to be also considered for the development of the workflow ontology in BIM4EEB.

Construction management ontologies. The Diction project (Digitalization of construction workflows, Business Finland, 2018-2020) has developed ontologies for construction management, with the objective to maintain shared situational awareness of the progress of a project amongst stakeholders. The Diction ontology work aimed at the conceptual integration of BIM models and sensor-based data gathering with new planning methods such as the Last Planner practice (Ballard 2000) for developing weekly plans, and Takt time planning (Heinonen 2016) for developing fast tracking construction plans.

The ontology work in the Diction project was based on close participation of construction companies, interested in a more integrated management of BIM models, sensor data and their connections to advanced planning methods. The work proceeded through several industry workshops. To achieve the linking of BIM models and sensor data with activities and to support the Last Planner™ practice, the connections of activities with different activity ingredients have been represented. The term *flow* has been used to specify the ingredients that are essential to make an activity executable. In the detailed modelling, the relation between activities and flows is defined more accurately with *preconditions* (the state in which a flow should be at the start of the activity) and *effects* (the state of a flow after the execution of an activity). The following categories to classify flows were introduced (Koskela, 2000):

1. *Prerequisite* - The state of the entity that the activity is going to transform
2. *Workspace* - The location for performing the activity
3. *Labour* - The labour crews that execute the activity
4. *Equipment* - The equipment needed in the execution
5. *External* - Humidity, temperature, dustiness, etc.
6. *Material* - Material batches consumed by the activity
7. *Information* - Models, drawings, instructions, permissions, etc.

Any activity can have multiple independent flows in each of these categories. A precast installation can have (i) multiple labour crews, such as crane operator and precast installers, (ii) multiple equipment such as the crane itself and precast supports and so on, (iii) materials such as grating concrete and precast embeds, and so on. The flows depend on the type of activity and the selected work method.

Since introduced by Koskela (1999, 2000), the concept of a flow has been refined into an Activity Flow Model (Garcia-Lopez 2017). It studies the flow model in closer detail and presents concept diagrams that outline an informal ontology.

The construction management ontology of Diction has been divided into four modules:

- **Construction objects**: The representation of the domain objects that are subjects of transformations within a construction project: physical objects (building elements, workspaces), information entities (BIM models, drawings, messages), and object groups (procurement package, inspection area, transportation package). The domain objects can be decomposed and can have multiple identifiers, classifications, and level of detail/development assignments.
- **Construction organization**: The actors (agents that carry out activities) and their organizational relations, as well as the contracts between them (contract obligations such as deliveries, services, payments, and constraints among them).
- **Construction planning**: Activities (at multiple levels of work breakdown), resource requirements and resource allocations. Temporal associations and activity precedencies. Relations with flows through conditions (states of flows at given periods of time).

- **Construction data gathering**: Information gathered from sensor systems, status events, and scanning.

The overview of these modules is shown in Figure 17.

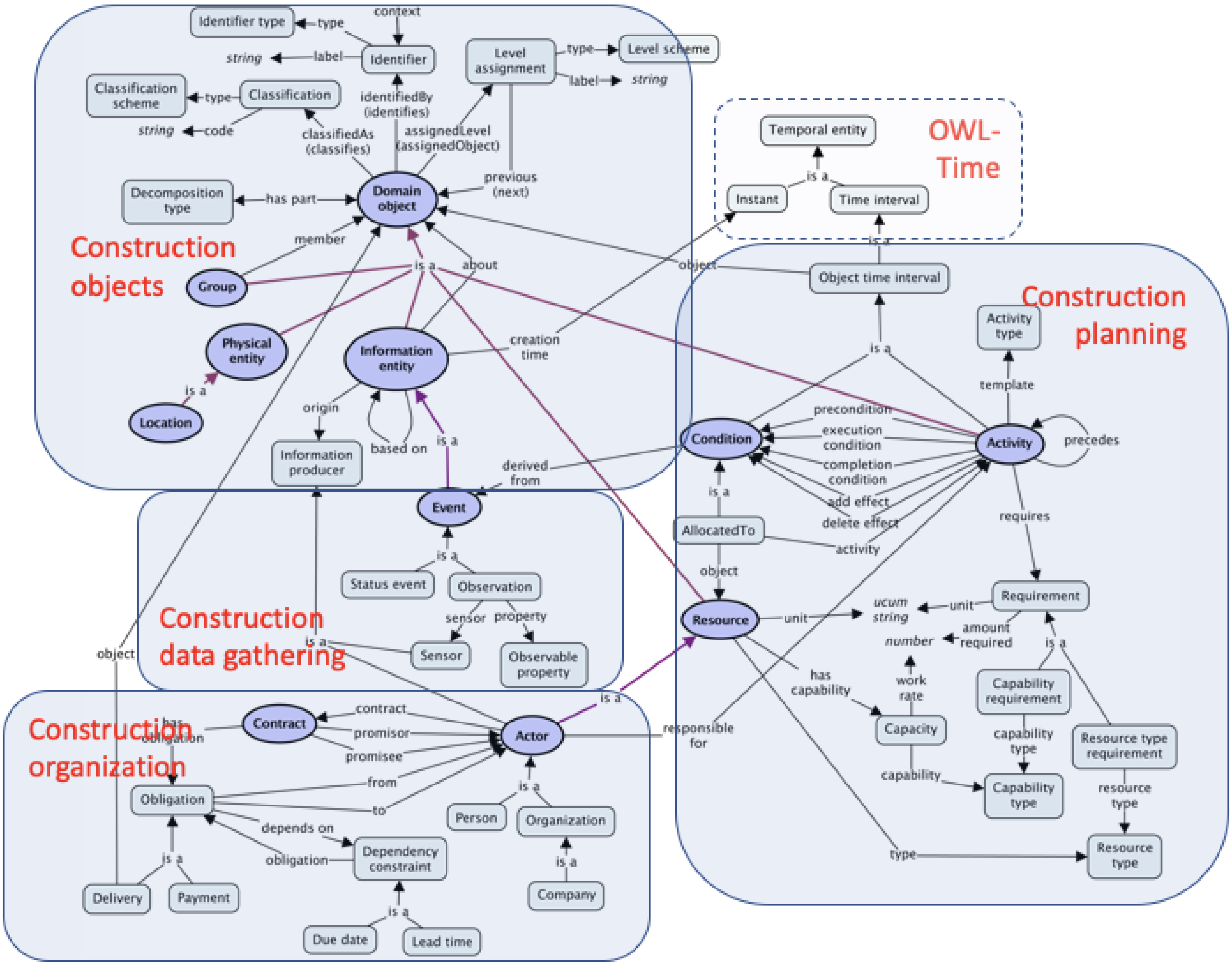


Figure 17: Construction management ontology of Diction

4.2.2 Cost management

The interpretation of cost in the context of BIM for energy-efficient renovation can be manifold, such as:

- Initial, monetary cost for the building renovation.
- Operational cost for the building before and after renovation. In this case, one important factor for consideration is the scope for the operational cost, including what systems are included and what time frame is considered.
- Total lifecycle cost: In this case the consideration of monetary cost is complemented by the consideration of "environmental cost". The scope of the determination and analysis of environmental cost can also vary substantially, e.g., are "support processes" for the manufacturing of building material included in the life-cycle analysis (LCA) or not.

The cost of renovation activities and the related return of investment are important factors during the decision process if a renovation shall go ahead or not. In the following chapter we prefer to talk about cost-effective renovations instead of profitable renovations.

Modelling cost is involved in numerous activities or scenarios which are relevant in the BIM4EEB project.

1. Estimation of renovation cost
2. Specify approximate renovation cost
3. Specify exact quantities and costs
4. Specify environmental cost and full lifecycle cost

In IFC a cost item (IfcCostItem) is a subclass of IfcControl (Figure 18). Its primary use case is to document planned and accrued cost. Based on this documented cost either (i) simulation tools or (ii) control software can make informed decisions how to progress with the execution of (construction) work. As per the IFC documentation:

"*An* IfcCostItem *describes a cost or financial value together with descriptive information about its context that enables its use within a cost schedule. An* IfcCostItem *can be used to represent the cost of goods and services, the execution of works by a process, lifecycle cost and more.*

Each instance of IfcCostItem *may have a name and a description. Depending on the use for which the cost is intended, these values should be asserted based on an agreement such as the name attribute could be used to provide a common value that enables distinct instances to be brought together in a nesting arrangement while the Description attribute may be used to provide text used for item description in a cost schedule."* (ISO16739-1, 2018)

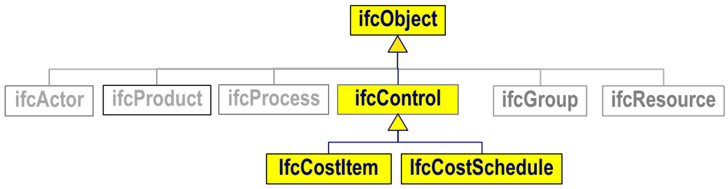


Figure 18: Inheritance tree for cost-related elements

In an IFC-based open BIM, the basic concept to model cost is the *cost value*. The concept of a cost value is explained in the documentation of IfcCostValue as "*a value that affects an amount of money*" (ISO16739-1, 2018). It can be interpreted as the "unit cost" and the total cost can be calculated as "unit cost times number of units".

"An IfcCostItem *can link one or many* IfcCostValue *objects representing a unit cost, total cost, or a unit cost with one or many quantities used to generate the total cost. The quantities can be given as individual quantities or [] as element quantities by one or many building elements. The* IfcCostValue.CostType *attribute indicates the [cost category]. For nested cost items (based on* IfcRelNests *relationship,* Figure 19*),* IfcCostValue.CostType *is significant such that* IfcCostValue.AppliedValue *is calculated as the sum of all nested costs having the same* IfcCostValue.CostType *or if set to an asterisk ('*'), then the sum of all nested costs of all cost types. An* IfcCostValue *can represent an original value or a value derived from formulas using* IfcAppliedValueRelationship*. For example, taxes may be calculated as a percentage of a subtotal."* (ISO16739-1, 2018)

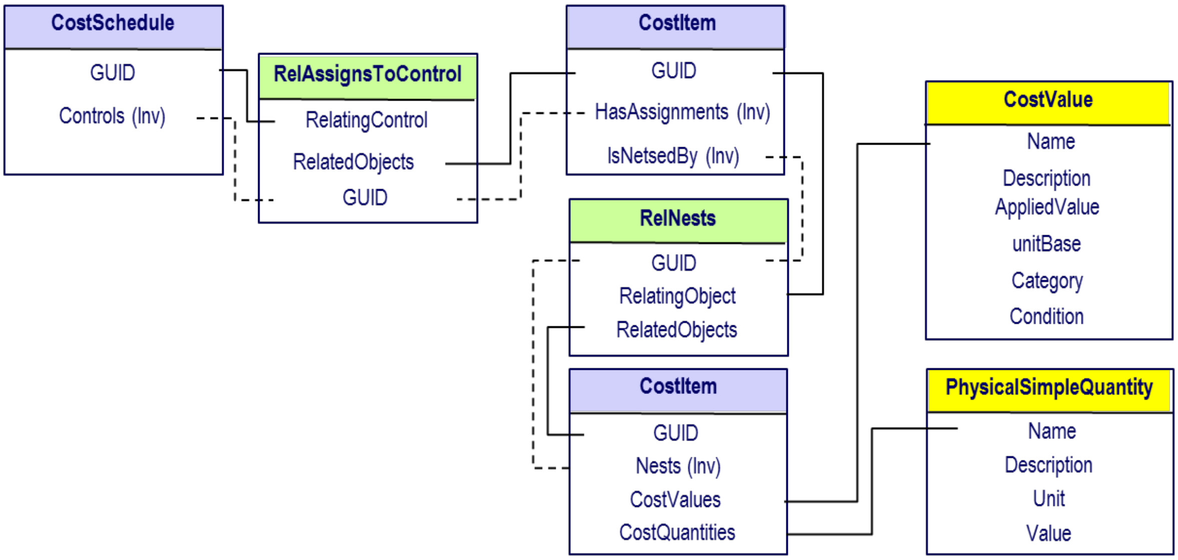


Figure 19: Breakdown of CostSchedules to single cost values and quantities (simplified UML)

Cost items can be composed of other cost items (Figure 20). Furthermore, the attribute "arithmetic operator" allows the specification how individual cost values shall be combined with each other.

In the IFC modelling approach cost can be linked to numerous elements, depending on the process model (or contractual model) that needs to be supported. Figure 21 represents the fundamental process-oriented concept of modelling construction projects in IFC.

The horizontal flow represents the material transformation processes

(1) starting with a specification of all input products which are assigned to the process;
(2) The material transforming process is specified through the description of
    (2.1) cost required for the execution of the process plus or
    (2.2) resources required for the execution of the process and subsequently the related costs;
(3) a process is connected to the final product through the IfcRelAssignsToProduct relationship.

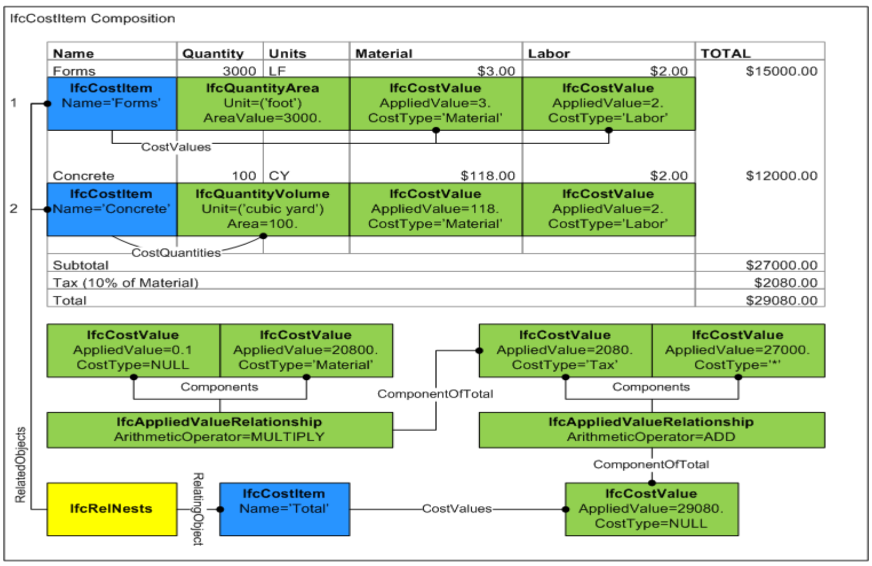


Figure 20: Composition of multiple cost items to calculate a total cost (taken from BuildingSmart)

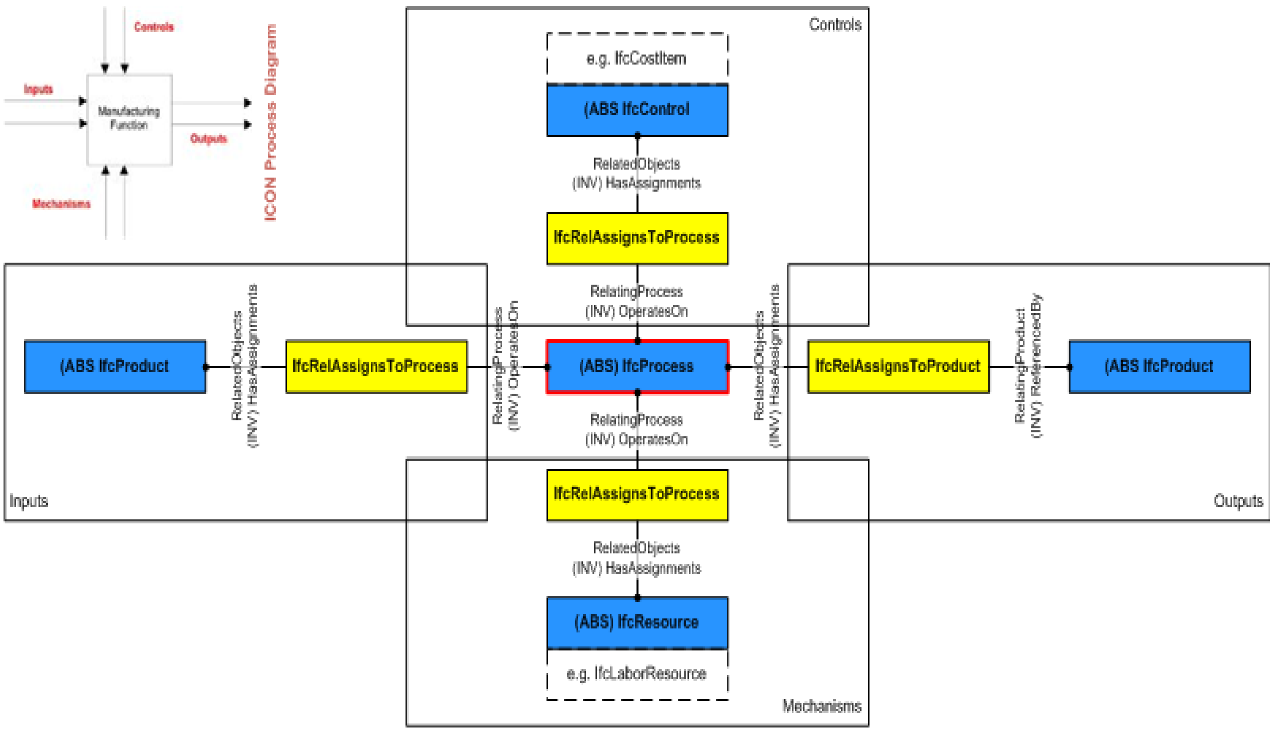


Figure 21: Linking Control Elements to Products, Processes and Resources (BuildingSmart)

In summary, on can state that IFC supports holistically and comprehensively the modelling of monetary cost. A full representation in ifcOWL is available and thus a modelling with ontologies is possible. At its current stage IFC supports the management of LCA-related data to a limited extend only. This makes sense, since LCA-related data are usually made available through national databases[6]. An important step towards a seamless, robust integration and improved interoperability is the usage of standardised material classifications (or ontologies), which is currently not available on an international level.

[6] A list of such databases: https://www.oneclicklca.com/support/faq-and-guidance/documentation/database/.

4.3 Existing ontologies or conceptual models for D3.5

The goal of task 3.5 is defined as follows: “To provide the basis for the BIM4EEB Measurement and Verification (M&V) protocol.The protocol will enable more accurate energy performance assessment thus further tackling causal factors of limited renovation market growth in Europe and increasing the competitiveness, attractiveness and reliability of renovation projects, while reducing associated failure risks or change requirements. The task will develop local energy performance models allowing accurate that will be linked to the BIM4EEB Linked Data framework (thus enabling semantic interoperability and access to data already described in existing ontologies and models configured in the previous tasks of this WP) in this task.” (BIM4EEB, 2018).

The use cases specified for the framework are divided into three categories:

- energy efficiency,
- occupant’s behaviour and comfort and
- building performance

Energy efficiency has been defined as *the ratio of output of performance, service, goods or energy, to input of energy* (EU-EED 2018). In the context of buildings, the main *output* is an *indoor environment* that affects both user comfort and technical maintenance of the building. Energy *input*, again, is needed by the *technical building systems* to meet their energy demand. Here the goal is on finding existing ontologies for *representing the input data* that is required for *energy simulation* especially for the *building permit* purposes and also to show that other requirements concerning the building are met during the *lifecycle of the building*.

The occupant behavior that has direct energy impacts are activities such as adjusting a thermostat for improving thermal comfort, switching lights on and off, opening and closing windows, pulling window blinds up or down, and moving between spaces. They are the key focus on building design optimization, energy diagnosis, performance evaluation, and building energy simulation due to the significant impact on real energy use and indoor environmental quality in buildings.

It is essential in design and operation of low energy buildings to have a deep understanding of – typically over-simplified or under-recognized – occupant behavior and being able to model and quantify its impact on the use of building systems and on the energy performance of buildings. There should be a clear connection between occupant models and building information models to satisfy the requirements of fine-grained simulation and building operation.

Building performance (and diagnostics) is a well-established field in the AEC and FM sector (Hartkopf, 1986). One major goal of research in this area is to harvest benefits from systems integration during the operational phase of a building in the interest of tenants, operators, and owners. The research field got new momentum with the need to convert buildings from “consumers” to “power plants” (Matthew, 1998). Most recently, the requirement for “Near Zero Energy Buildings”, i.e., the “Energy Performance of Buildings Directive” (EU-EPBD, 2010) requires that new buildings are close to zero-energy by 2021.

Typically, the requirements of the directive can only be achieved if energy generation from non-fossil fuel energy sources are widely integrated into the buildings. Thus, in previous research methodologies were developed which allow to support the holistic, integrated evaluation of multiple KPIs (Menzel, 2013; Menzel, 2016). In comparison to other approaches the goal is to define a methodology which allows the comparison of KPIs across multiple evaluation criteria, in different buildings, and in different climatic zones. This can – to the largest extent – be achieved through normalization (Menzel 2015).

4.3.1 Main use cases and data requirements for energy efficiency

Nine use cases for the energy efficiency assessment, related to different phases of the renovation project are listed below. The detailed descriptions are in ANNEX II (energy efficiency use cases).

- Initiative: Preliminary feasibility assessment from the energy efficiency perspective
- Initiation: Target setting for the energy efficiency indicators to support OPRs
- Concept Design: Quick calculation to find the conceptual design alternatives to meet OPRs
- Preliminary Design: Preliminary energy simulations of design alternatives with rough models
- Developed Design: More detailed energy simulation of design alternatives using BIM models
- Detailed Design: Detailed simulation of the design alternatives based on a digital twin
- Construction: Implementation of the renovation measures ensuring their compliance with plans
- Building Use: Evaluation of OPRs at the operational phase of the renovated building
- End of Life: Recycling of the products and materials of the renovated building

A starting point for defining the requirements of energy efficiency data are the Energy Performance of Buildings (EPB) standards (M/480 given to CEN). The EPBD article 3 / Annex I requires that each member state must describe their national calculation methodology in the national annex of the standard. The idea is to take steps towards a harmonized European way of energy calculation and construction requirements. Therefore, European projects about construction should use the framework to describe the information that is related to building permits and fulfilling the requirements concerning buildings.

4.3.2 Main use cases, data requirements and applicable ontologies for occupant's behaviour and comfort

The use cases for occupant's behaviour and comfort are presented with the focus about the definition of initial data requirements aligned with the use cases, namely identification of producer/consumer of data, timeline for data generation, data availability, etc. The renovation process flow as presented in D2.1 along with the detailed analysis of occupants and owners' requirements in D2.5 is considered. The use cases are listed below and the details are shown in ANNEX II (Occupant's behaviour and comfort use cases):

- Early identification of occupants' behavioural and comfort parameters - Establishing a comfort preserving framework for inhabitants
- Establishing a comfort preserving framework during the renovation process
- Renovation tasks and processes that take into account occupants' comfort conditions
- Establishing a comfort preserving framework following the renovation process

Relevant existing ontologies are reviewed in the following section in order to further evaluate the potential of covering the data requirements with already existing models and standards. The aim is to identify existing formal definitions (ontologies and other conceptual models) of the needed data as specified in the aforementioned use cases.

DNAs Ontology

This ontology is an OWL definition of the DNA concept (Drivers, Needs and Actions), an outcome of "International Energy Agency's Energy in Buildings and Communities Programme (IEA-EBC) Annex 66 WG" working for the "Definition and Simulation of Occupant Behavior in Buildings". This is the result of the EU project HIT2GAP on the way to incorporate occupant's behaviour parameters in the energy simulation process. As stated by the consortium:

> *This part of the ontology is aimed at representing building occupant's DNA (Drivers, Needs and Actions). The model represents the user-building interaction not only in form of activities but also considering comfort parameters and other aspects that motivate users to perform building actions.*

ThinkHome Ontology

The ThinkHome Ontology[7] is the result of a project with the focus on occupant's behaviour integration in the building environment. There are two main models considered for behaviour profiling and of interest within the context of the BIM4EEB project. Occupancy patterns form a core point of the project as the objective is to incorporate occupants' profiles and activities in the overall management framework. This is the scope of the actor ontology, providing the details from the occupant end point.

There are additional ontologies available in the field; both generic models with low level of expressivity: e.g. PROV-O ontology (Lebo, 2013), Ontology Patterns for Complex Activities Activity Pattern Ontology (Meditskos, 2013), and more detailed models such as Adapt4ee Occupancy Ontology (Adapt4ee, 2014) mainly as results of different research projects. A more detailed presentation of the different ontological models focusing on occupants' behaviour modelling in the building environment will be reported in Deliverable 3.2 along with an in-depth analysis to define the basis of the BIM4EEB modelling work. While there are several EU projects examined in this part, the focus of the analysis is on the two main WGs of IEA, "Definition of Occupant Behaviour in Buildings" (IEA-EBC Annex 66), and "Occupant-Centric Building Design and Operation" (IEA-EBC Annex 79).

Annex66 defines "Energy related occupant behaviour in buildings" with the goal of creating obXML that is a standard representation of occupant behaviour to support building performance simulation. obXML follows the DNA's principles with a library containing occupants' behaviours that have energy impacts. Annex 79 started in June 2018 and is a continuation of Annex 66. It aims to integrate and implement occupant's presence and behaviour in the design process and building operation so that both the energy performance and the comfort conditions in the building can be improved. The scope of this WG is close to the objectives of BIM4EEB about occupant's behaviour and comfort modelling. A more detailed analysis of the above two initiatives will be presented in D3.2.

4.3.3 Main use cases and data requirements and applicable ontologies for building performance

The four main pillars for building performance are: (i) Comfort Performance, (ii) Systems' Performance, (iii) Energy-related Performance, and (iv) Occupation Density related Performance. Usually, evaluation methods consider single aspects, such as energy consumption or user comfort. However, so called multi-step calculation methods are required to provide owners, tenants and operators with instruments to specify in a flexible way Service Level Agreements. In order to make data better comparable, measured performance data is normalised. Well-known examples for normalization are e.g. Degree Days, or the energy consumption per usable area. Property sets for normalised data are not available in ifcOWL. However, the inclusion of instructions for the calculation of normalised data is one example for a possible extension of ifcOWL. The basic instruments for an extension are available, e.g. by using the objectified relationship type "Association" links to external data sources (e.g. Degree Days) could be easily added. In case of normalised energy consumption values all data is already available in ifcOWL. Thus, the extension would be moderate, since only the calculation instructions need to be added.

To support the aggregation of KPIs, further constraints must be satisfied. One major constraint is that the range of the KPIs must be identical. In case of building performance those KPI can be calculated by *comparing the satisfaction of constraints in fixed time intervals*. One example is to calculate temperature satisfaction as values out of agreed range divided total number of values. Such KPIs can be defined for different sensor types, actuator types but also for different types of smart meters. Thus, an "average, aggregated performance" can be calculated as the "average over all KPIs". Irrespectively of the numbers of installed types of monitoring devices the average, aggregate performance KPIs is always between 0 and 1 and remains comparable. More information of performance criteria is in the standards in Table 2.

---

[7] https://www.auto.tuwien.ac.at/downloads/thinkhome/ontology/

Table 2: Selection of standards specifying performance criteria

ID	Title
ISO EN 7730 (2007)	Analytical determination and interpretation of Thermal Comfort
EN 12792	Mechanical Ventilation of Buildings
EN 13 306	Maintenance - Maintenance terminology;
EN 15240 (2011)	Ventilation for buildings; Energy performance - Inspection of air-conditioning systems
EN 15251	Indoor environmental parameters for design and assessment of energy performance of buildings for indoor air quality, thermal environment, lighting and acoustics
EN 15316	Heating Systems in Buildings - Calculation of System energy requirements and efficiencies
EN 15341:2007-06	Maintenance - Maintenance Key Performance Indicators
EN 15603	Energy Performance of Buildings
DIN 31051	Fundamentals of maintenance
ISO 50001 (2011)	Energy Management System

A comprehensive overview about currently valid and maintained performance modelling ontologies is provided at http://smartcity.linkeddata.es (Ses, 2015):

- Simulation Information Model (SIM) Ontology
- Performance Information Model (PIM) Ontology
- CASCADE Generic Facility ontology
- CASCADE Fiumicino Airport ontology
- Energy in Buildings Ontology
- SAREF: the Smart Applications REFerence Ontology
- EnOcean ontology

In addition, SSN represented in the earlier section, has to be taken into account as well.

- SSN - Semantic Sensor Network Ontology (W3C, 2005, 2017)

The list of use cases is the following. For full details, see ANNEX II (building performance use cases).

- Documentation of user comfort
- Documentation of system usage
- Documentation energy consumption:
- Documentation of occupation density

The Systems' Performance indicators are used to document with what intensity services were used. By measuring for how long certain building services devices were used, it can be determined what effort was required to achieve a certain level of user comfort. For example, a comfortable temperature can be achieved by reducing the heating demand or by operating the radiator whilst the window is open. In the latter case, energy is obviously wasted.

So far, energy consumption cannot be metered on the level of individual consumers, but only on apartment level. This means that tenants will receive limited feedback how to improve their behaviour.

The knowledge when and how "dense" a space was populated may impact the "control scenario" for certain spaces, e.g., to decide about pre-heating or pre-cooling periods.

4.3.4 Discussion

The tools developed in BIM4EEB will be used in the real-life demonstrations in Italy, Finland, and Poland. While the tools are developed based on general European standards, it is important to assess how compatible they are with the national methodologies. Due to national differences, the tools should be designed to allow national extensions. These should be made at least for the demonstration countries.

The scope of BIM4EEB is to have the building occupants at the heart of the process; therefore, a thorough analysis of occupants' behaviour profiles and models is required. Considering existing ontologies, a need for a unified framework to incorporate the different layers of occupants' behaviour modelling arises. The selected ontologies or models are retrieved from the standardization or from different working groups in the domain towards establishing a standardized approach for occupants behavioural modelling. Furthermore, it is evident that the incorporation of additional attributes in existing ontologies is a key task in BIM4EEB to address project specific needs and requirements. Overall, the scope of this process is twofold: a) to merge the attributes of different ontologies in the domain and b) to extend existing ontologies with new attributes considering the specificities of the BIM4EEB project. The details of this work will be reported in D3.2.

Building performance evaluated with performance indicators can be represented as properties of a building or its systems. The properties need to be objectified in order to add sufficient metadata to the indicators, including the quantity kinds, units of measure, time of measurement, time of validity, and so on. Each property object should also contain a reference to the definition of the related performance indicator. That is, a property of an object has two perspectives:

1) Property of something: for example, the annual energy consumption of a building
2) Property description: what does the energy consumption mean and how it is measured.

The properties of objects (a quality) can refer to the property description (and information entity) that can again refer to the primary source document (e.g., a document containing the CE specification).

4.4 Existing ontologies or conceptual models for D3.2

Data requirements with respect to Equipment and Materials exist in numerous cases – nearly through all lifecycle phases. Building materials and HVAC equipment are the pre-dominant technical artefacts which determine the quality of “Indoor air quality” and “acoustic comfort”. It is important to notice that certain technical parameters may contribute to good performance of one aspects (e.g. natural ventilation) but may cause a degradation of other aspects (e.g. acoustic comfort).

4.4.1 Equipment parameters

In this section we discuss ontologies to model equipment and material parameters. Since both materials and equipment are included in the ifcOWL schema – converted from IFC – we will start the discussion with an analysis of IFC regarding these areas.

Figure 22 illustrates the set of building automation systems. We assume that during a building renovation a major part of the work will be the installation of advanced building automation components. In order to achieve higher efficiency, it is recommended to improve the level of control (e.g. per room instead per apartment).

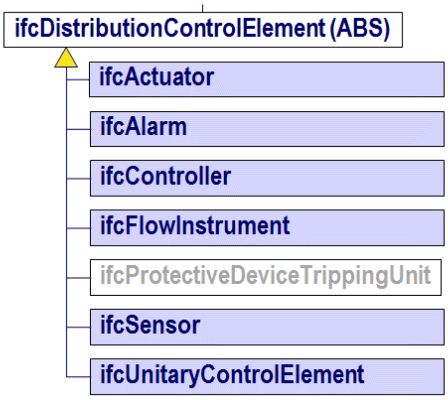


Figure 22: Building Automation Domain

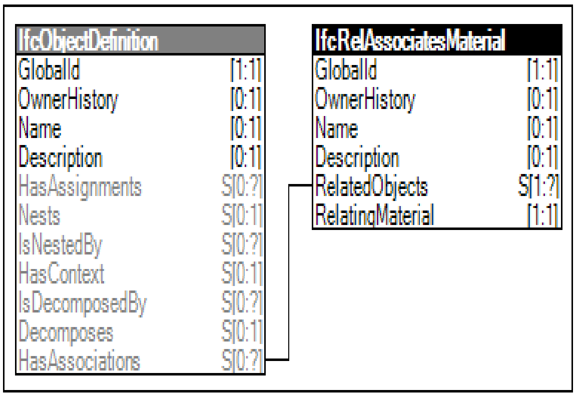
Scope: The required scope for modelling seems to be rather limited, since the number of information objects is limited in the currently available standards (ifcOWL). However, usually in early planning stages much less information is compiled in the design and documentation of automation systems. This is currently a major obstacle in the process management.

4.4.2 Material Parameters

One prominent example for existing developments in the area of ontologies for construction materials is the Eurobau Ontology (Radinger, 2013), covering 81 Manufacturers / Brands, 19 Resellers, 183 Warehouse locations, 56360 Product types, including over 88 million triples of real business data with a high domain density.

Information about building materials and equipment is relevant in several steps in a renovation process, especially (i) in the identification of the existing state of the building (i.e., the “as-built-status”) and (ii) for the preparation of documents during the planning and evaluation of the building renovation. This includes in the specification of input parameters for any type of simulation systems, such as (a) energy simulations, (b) acoustic simulations, or (c) workflow simulations.

Finally, (iii) the documentation of the construction work executed should be as precise and complete as possible, to support advanced, automated operation of the building after renovation. So far, the emphasis of modelling is in most cases on the planning phases. An initial SOTA is based on the Material Definition provided by IFC (ISO16379). Material can be associated with any subtype of IfcObjectDefinition. For the purpose of this document, we focus on the association of material definitions to IfcElement and IfcTypeElement (Figure 23), since children of these elements deliver the constituent parts of either the “core and shell” or the “building services systems” of a building.




```
ifc:IfcRelAssociatesMaterial
 rdfs:subClassOf          ifc:IfcRelAssociates         ;
 rdfs:subClassOf
  [rdf:type owl:Restriction                            ;
   owl:allValuesFrom ifc:IfcMaterialSelect             ;
   owl:onProperty
   ifc:relatingMaterial_IfcRelAssociatesMaterial]      ;
 rdfs:subClassOf
  [rdf:type              owl:Restriction               ;
   owl:qualifiedCardinality  "1"^^xsd:nonNegativeInteger
;
   owl:onProperty
   ifc:relatingMaterial_IfcRelAssociatesMaterial       ;
   owl:onClass      ifc:IfcMaterialSelect        ]     ;
 rdf:type  owl:Class .
```

Figure 23: Association of Material Properties to IfcElements

4.4.3 Indoor air quality

According to Shree (2019): "*In the field indoor environmental quality, the study of indoor air pollution (IAP) requires knowledge in several diverse areas including principles of fluid mechanics, heat transfer, species transport, and systems engineering. Moreover, the complexity of buildings systems provides high levels of electronic control capabilities embedded in the structures. Of particular concern are issues involving contaminants that routinely enter or lie dormant within building interiors, and their effects upon human health. Efforts to define and describe pollutant transport within buildings and interiors have become complex, requiring nowadays computational tools and techniques that were used only in research laboratories a few years ago. Knowledge of principles of ventilation and building systems must now be coupled with computational fluid dynamics techniques to accurately assess human health and predict contaminant exposure*". Therefore, it is very important to address the modelling of indoor air pollution and quality in building premises as a key concept of the renovation process.

This section presents the main BIM4EEB use cases related to indoor air quality. Once again, the analysis is covering the different phases of a housing renovation process including the phases during and after a renovation project. The definition of these use cases is in line with the outline of the renovation processes in D2.1 and with the analysis of occupants and owners' requirements in D2.5. The following use cases are presented in this section with the focus about the definition of initial data requirements aligned with these use cases. (Full details in ANNEX II)

1. Early identification of parameters for an Indoor Air Quality model
2. Establishing a health and Indoor Air Quality environment during the renovation process - Renovation tasks/processes to address Indoor Air Quality standards
3. Establishing an Indoor Air Quality preserving framework for post renovation processes

In the subsequent sections we aim to identify existing formal definitions of the data needed. A non-exhaustive list of ontologies and other conceptual models applicable in the specific domain (IAQ analysis) is investigated. The emphasis of the investigation is on 1) IEA-EBC Annex 66 occupancy behaviour model, 2) Indoor air quality standardization/legislation, 3) ISO 45001 (ISO 45001, 2018) for health and safety conditions. A brief description of each model/ontology follows:

IEA-EBC Annex 66 - Definition of Occupant Behaviour in Buildings

The framework was already presented in section 4.3 as part of Occupant Behaviour and Comfort analysis. Complementary to visual and thermal comfort model definition, IAQ parameters are incorporated as part of the occupants' behaviour modelling. More specifically, IAQ comfort has already been prescribed as part of the obXML model as part of the initial model of NDAs. Along with the incorporation of IAQ as part of the overall Occupant Behaviour and Comfort framework as examined in the project, we have to take into account the domain specificities, namely the standards and guidelines that specify the IAQ KPIs and the associated values for boundaries.

Indoor air quality standardization/legislation

Concerning indoor air quality, BIM4EEB envisions to support the evaluation of indoor hygienic and health/well-being conditions. The Environmental Protection Agency (EPA) list of VOCs helps to identify and reduce hazardous substances. The list has been adopted from **ASHRAE** through the **2016- 62.1 IAQ Standard**. Although threshold limit values for VOCs vary between countries and organizations, the EPA outlines several common VOCs and substances with threshold limit values:

- **PM2.5**: Particulate matter with the diameter less than 2.5 μm is a dangerous form of pollution since small particles can reach lungs causing numerous adverse effects.
- **CO**: Carbon monoxide is a colourless, odourless and lethal gas, one of the most dangerous compounds in indoor environments. Threshold limit for an 8-hour workday is 25 ppm[8] or 35 ppm[9].
- **CO2**: Carbon monoxide CO2 has the average outdoor concentration of 300-400 ppm. Indoor levels are higher than that. Adverse health effects can be observed at levels over 7,000 ppm. Occupational limits set by ACGIH are 5,000 ppm TLV-TWA* and 30,000 ppm TLV-STEL**.
- **Radon**: Radon is a carcinogenic radioactive gas formed by the decay of uranium in the soil with no safe levels of exposure. Still, the EPA has set an action level of 4 pCi/L.
- **Formaldehyde**: A common VOC which is emitted from different sources, like furniture, incense burning, and cooking. Threshold limit is 0.1 ppm TLV-TWA* and 0.3 ppm TLV-STEL**.
- **Methylene chloride**: Methylene chloride (dichloromethane) can be found, for instance, solvents. Odour threshold is 250 ppm. Long-term exposure can have effects to the central nervous system.
- **NO2**: Nitrogen dioxide has adverse health effects. Threshold in 1-hour exposure is 100 ppb.

When focusing on the management of renovation activities from an organization point of view, we have to consider specific limitations about IAQ conditions during the renovation process. The ISO 45001 for health and safety has been designed to be integrated into an organization’s existing management processes to handle issues related to the health conditions in premises; this means IAQ conditions among others. Therefore, a thorough evaluation of the ISO 45001 standardization is required in order to address the requirement for establishment of a high (indoor air) quality environment in building premises during the renovation process.

The scope of the project is to clearly address aspects related to IAQ conditions in a building environment. The analysis is twofold addressing both requirements during the renovation process and requirements following the renovation process (building operational phase).

[8] The American Conference of Governmental Industrial Hygienists (ACGIH)
[9] The National Institute for Occupational Safety and Health (NIOSH)

Therefore, the analysis in this section is focusing on the review of existing work in the field in order to:

a. Incorporate IAQ as part of the holistic behavioural framework defined in the project
b. Extensively review existing standards and methodologies in order to:
    a. Identify the way to incorporate IAQ parameters in the modelling framework (from an operational and organizational viewpoint)
    b. Evaluate IAQ KPI values in order to set the boundaries for the modelling process.

Furthermore, it is evident that the incorporation of additional attributes in existing ontologies is a key task in the BIM4EEB to address project specific needs and requirements. Overall, the scope in the project is twofold: a) to merge the attributes of the different ontologies in the domain in a unified framework and b) to extend the existing ontologies with new attributes considering BIM4EEB project specificities. The details of this work will be reported in D3.2.

4.4.4 Acoustics

The field of building acoustics is primarily targeting and dealing with system properties, which are often hard to change after the system is finalized. Therefore, it is of paramount importance that acousticians will be able to predict and analyse in an early stage what acoustic properties and performance can be expected from a building as a result of a planned renovation program. The acoustic performance needs to be planned with respect to applicable standards – although, in building renovations the criteria is often just to match the existing acoustic performance. Today, there exist various numerical tools that can handle complex structures, for instance, tools based on the EN ISO 12354 standard series. However, two important preconditions for accomplishing a satisfactory acoustic design are that (1) all relevant acoustic properties of the building components are available, and (2) all structural transmission paths are set and documented.

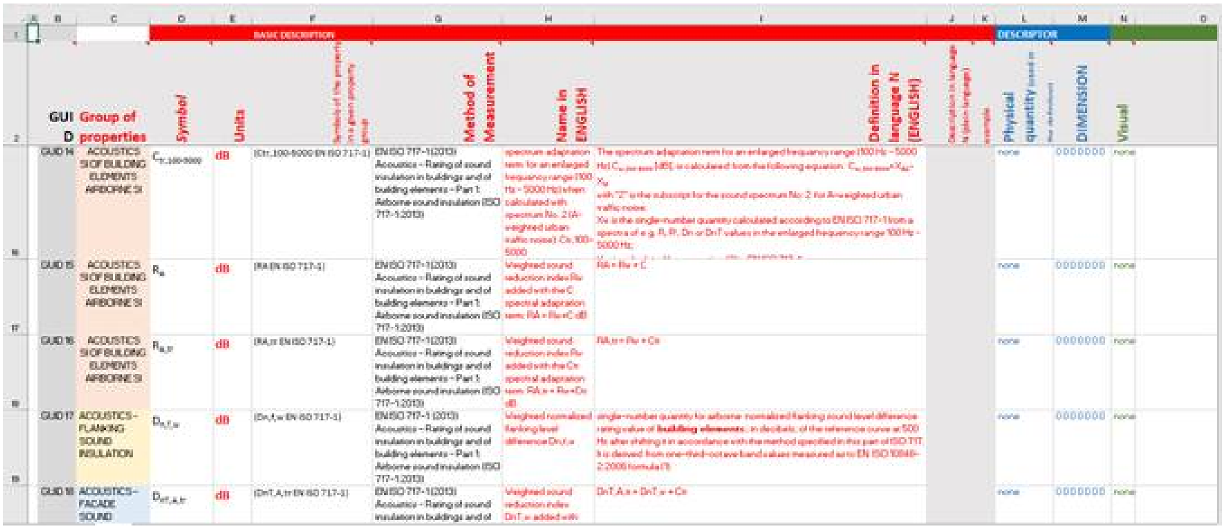


Figure 24: Excel-sheet of adaption of prEn 23 386 from CEN TC 126 – WG 2 (unofficial document).

Renovation projects often involve improvements of the energy efficiency by increasing the thermal insulation of the building envelope. This does not mean, however, that a similar improvement on the sound insulating properties can be expected. On the contrary, efficient light weight thermal insulating systems are often poor sound insulators, particularly in the low frequency region and if acoustical properties have not been considered in the design.

The following use cases are presented in this section with the focus about the definition of initial data requirements (data modes) aligned with these use cases. (Full details in the ANNEX II)

1. Early identification of goals of acoustical properties to define requirements (Concept design)
2. Early identification of acoustical parameters for planning, specifications, and design of solutions (Preliminary/Developed/Detailed Design iterative process)
3. Implementing solutions for the acoustic performance during the construction phases
4. Maintaining acceptable acoustical indoor conditions.

In this section, the aim is to identify existing formal definitions (ontologies and other conceptual models) of the needed data as specified in the aforementioned use cases. On international levels these projects are in progress:

- LOIN (Level of Information Need) in acoustics. SubWG of CEN TC 126–WG 2
- CEN TC 126 – WG 4 in the preparation of the prEN 23 386 (Figure 24).
- Development of international acoustic classification scheme for residential buildings ISO/TC43/SC2/WG29.
- Ongoing work in CEN TC126/WG12 on BIM Acoustics.

Example of suggested framework and is primarily being developed based on available models and parameters in relevant acoustic ISO standards

Since there is on-going work in working groups of international standardisation bodies that are focusing on acoustical applications of BIM in agreement with what is needed in this project, the acoustical formulations should in this project primarily be adaptations of the models and ontologies that currently are being developed, and selection of a relevant sub-set. A collaboration has already been established with representatives of the LOIN in Acoustics and the prEN 23 386 projects, respectively.

4.5 Existing ontologies or conceptual models for D3.3

During renovation projects – and in other construction efforts as well – several complex structural models are produced about physical entities, information entities, and activities. Over the engineering and management work, these models can grow extremely large. In big projects their production is therefore organized as a succession of increasingly refined and focused models. This information refinement process has been traditionally conceptualized through the concept of LOD.

There is no common, unambiguous definition of the term LOD. According to the context, it can be interpreted to mean Level of Development, Level of Detail, or Level of Definition; each of these terms has a slightly different meaning. Since 2004, different countries have developed dedicated LOD standards generating a complex framework at the world level. The main specifications are the following:

- USA BIMForum Specification (BIMForum, 2018),
- UK BS 11292-1 and PAS 1192-2 and 3, and
- Italian UNI 11337 in its part 4 (UNI 11337, 2017).

USA – In the terminology used by the U.S. legislators since 2013, LOD has assumed the meaning of Level Of Development. In the USA context, there is no formal difference between geometric and non-geometric information (while it is included in the UK context, Level of Detail – LOD and Level of information – LOI). Nevertheless, this distinction is embedded in the two reference documents of the BIMForum Specification where Part I identifies the element geometry and Part II identifies the attribute information. The LOD levels present in the US references are named and ordered progressively according to the following scale:

- LOD 100, LOD 200, LOD 300, LOD 350, LOD 400, LOD 500.

UK – According to the UK standards, LOD has the general meaning of Level of Definition, which includes the two distinct parts of the Level of Detail (LOD) and Level Of Information (LOI). The level of detail represents the description of the graphic contents at each stage, while the level of information represents the description of non-geometric contents. The UK standardization system defines the following scale from 1 to 6:

- LOD: 2 – Concept stage; 3 – Developed Design; 4 – Technical Design; 5 – Construction
- LOI: 2 – Concept stage; 3 – Developed Design; 4 – Technical Design; 5 – Construction; 6 – Operation and maintenance

ITALY – Simultaneously, the Italian norm UNI 11337:2017 defines the LOD as the Level of Development of the objects, according to the quality and quantity of the attributes needs related to the phase of the process. The LODs are divided into LOG, Level of development of objects – Geometric Attributes, and LOI, Level of development of the object – Informational Attributes. The scale of the Italian LOD is organized in seven steps ordered from the letter A to the letter G, as follows: LOD A: Symbolic object; LOD B: Generic object; LOD C: Object-defined; LOD D: Detailed object; LOD E: Specific object; LOD F: Object executed; LOD G: Object updated.

Comparison of LOD systems – In each of the three systems introduced above, there is a relationship between the phases of the building process and the classes defined by the LOD. Figure 25 shows how the different systems cover the construction lifecycle differently, with particular attention to the reference periods of the last LOD levels of the three analysed systems. For example, the LOD Class 5 (UK) covers a broader period than the US and Italian counterparts. The Italian system provides the LOD category G – Updated object, comparable to the categories defined by the as-built expression, but without a corresponding category in the other systems.

To this type of "horizontal" comparison between the different classifications, it is possible to add a "vertical" control to match different contents between the different LOD systems. There are not many differences in geometric descriptions between the corresponding levels of the three systems; also, the non-geometrical contents are similar in the first levels but significant differences are introduced in the later levels.

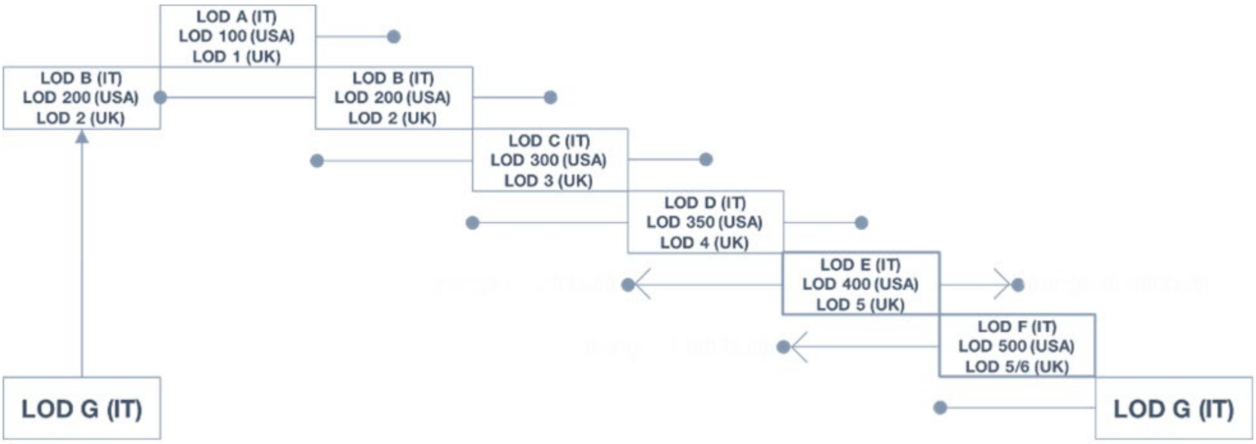


Figure 25: Comparison between the existing LOD scales

In different countries, based also on the correspondence with the phases of the construction process, the number of the attributes required for each LOI rises progressively from a few units for the generic description of the object up to several dozens in detailed levels. For instance, the LOD 500 (USA) for a steel pillar requires 31 parameters, while the corresponding LOI 6 (UK) requires a total of 64. Indicatively, the contents of the LOI from the third level onwards are constructive in nature and no longer affect the technical performance, as otherwise happens only for the UK division. The US System already expects sustainability requirements for LEED certifications from the outset. The last major exception of the Italian system is the LOI G level addressing the temporal and qualitative parameters of maintenance of the objects, which is absent in the USA system.

In conclusion, it is possible to find similarities between the different classification systems and it is possible to conclude that, even with a different subdivision, the amount of information presented covers all phases of the building process with regard to objects, while providing distinct insights of different standardization bodies. However, even though there are no specifications available for LOD levels, there is still a problem that none of these specifications defines about what are the acceptable contents of individual data fields, creating an obstacle for interoperability due to the different interpretations of the inputs of the same parameters.

LOIN standards

With the publication of ISO 19650 Parts 1 and 2 the consolidated structure of the LOD, as illustrated by the three systems described above, is being disrupted through the introduction of the concept of LOIN (Level of Information Need). LOIN basically aims to erase the concert of predefined scales such as the UK (Level or definition 1, 2, 3, etc.), USA (Level of Development 100, 200, 300, etc.), ITA (Level of Development A, B, C, etc.) in order to switch into a gradual, incremental, milestone-free system (relevant points) established beforehand.

LOIN does not include any predefined levels. The idea is that the definition of the quantity and quality of information produced at specific tasks should be based on actual necessity present in the related information exchange situations. Any additional information produced beyond that necessity can be considered as a form of waste. The LOIN in ISO 19650 is introduced but not defined in detail, the reason why the CEN has developed its own norm for the European LOIN.

4.6 Suggestions for further actions in BIM4EEB

There is significant amount of work done ontologies relevant to renovation and construction projects. The suggestions for further activities in BIM4EEB are the following:

1. Utilize the ifcOWL and related converters to be able to access existing BIM models as Linked Data.
    - Limit the use of ifcOWL to the BIM related data, not directly for the representation of resources, actors, approvals or other data external to BIM models.
    - Use the version of IFC (and ifcOWL) standardized as ISO 16739.
2. Design the ontologies to be compatible with the concepts of ISO 19650.
3. Align with the ISO 21597 ICDD Container ontology for importing interrelated datasets.
4. Utilize other commonly used ontologies as follows:
    - BFO (ISO/IEC21838) for fundamental categories;
    - OWL-Time for temporal phenomena;
    - QUDT for units (units of measure as well as currency units);
    - FOAF, Org, and PROV-O for agents (people and organizations);
    - FIBO for legal roles of agents;
    - SSN/SOSA and Saref Core for sensor data, and
    - PROV-O for provenance data.
5. Use the named graphs of RDF dataset to implement the information containers of ISO 19650.
6. Use objectified properties to manage the temporal evolution of data, aligned with OPM, QUDT, SSN/SOSA and Saref.
7. Develop the ontologies in a modular fashion and use the vertical and horizontal segmentation approach of SSN/SOSA.
8. Define a representation that allows the specification of different LOD frameworks.

There are no existing established construction management or project management ontology available. Since it is not advisable to create ontologies from scratch, it has been considered best to continue the initial ontology work of the DiCtion project on construction management ontologies – based on a significant industry input – by refactoring, refining, replacing, and supplementing those initial ontologies with more appropriate, detailed, and carefully thought of definitions of construction management concepts (D3.4) and implementing several additional modules for concepts related to materials, energy, occupancy (D3.2) and construction lifecycle (D3.3).

It has been agreed with Diction project that the work on the common ontologies will be carried out, and so that the results can be shared among both projects. The ontology refactoring and definition work starting at spring 2019 will be done in BIM4EEB, but DiCtion or one of its continuation projects is later on going to implement use cases related to construction site logistics and supply new ontology modules concerning semantic image interpretation, image-based progress monitoring, and construction process libraries. Such additional development and use will also give a bigger future impact to the work done in BIM4EEB.

The alignment of Digital Construction Ontologies with the ontology work done in DiCtion is provided in the Deliverable 3.6, Appendix I (BIM4EEB-D3.6, 2022).

5 High-level architectural specification for Linked Data and ontology framework

To support the building performance-based renovation planning and management, this section presents an information management framework based on the technological basis of Linked Data and ontologies. The basic terminology and structure of information management is motivated by the approach specified in ISO 19650, parts 1-3 (ISO 19650-1, 2018).

The Linked Data and Ontology framework is designed to be extensible in a sense that it should support semantic interoperability between standalone systems that can implement the functionalities defined in the framework and adopt the conventions specified in it. Furthermore, the framework defines the practices that it itself will be developed and deployed.

5.1 Objectives and high-level architecture in BIM4EEB

The objective of the framework is to enable semantic interoperability between independent systems in the renovation domain that implement the functionalities and comply with the conventions specified in the framework. The framework is open in a sense that no restrictions are placed on which tool can participate in it. This is a decision that can be made separately by each tool.

A tool does not need to be implemented originally according to the framework. It is expected that future solutions for renovation management are likely to be implemented as *systems of systems* (Maier 1999). There will be independently developed and used systems serving their specific purposes that will produce and consume information. The overall information system environment in any specific renovation project will be built from the existing systems of the parties involved. The essential goal is to have interoperability between different systems because it allows to connect them together to serve the goals of the project.

The constituent systems must be connected with each other for the duration of a renovation project to act in complementary roles in the project. Detailed data remains inside each system. Systems can share relevant common data with others. To achieve interoperability, the constituent systems must have interfaces for the required interactions and semantic interoperability relevant to the renovation domain.

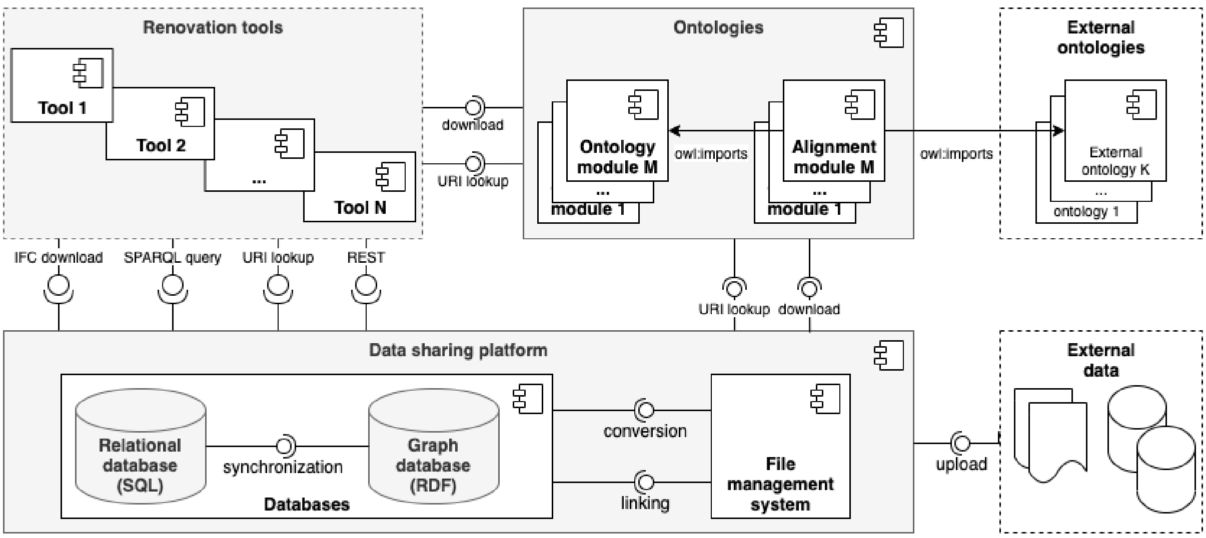


Figure 26: High-level architecture in BIM4EEB

The content of the framework are definitions of the common functions and conventions – interfaces, data formats, and ontologies – of constituent systems for semantic interoperability.

To illustrate the role of Linked Data and ontology framework, Figure 26 contains a high-level architectural diagram as realized in the BIM4EEB project. There is a set of renovation tools that interoperate by sharing data through a common data sharing platform. The tools do not interact with each other directly; all interactions happen through shared data. According to the framework, data sharing is based on the Linked Data principles and the use of ontologies that enable the semantic interoperability among tools. Both the tools and the data sharing platform access the ontologies either by downloading whole ontology modules or through URI lookup of individual terms. Downloading of whole ontologies is needed for performance reasons and to support reasoning functionalities. The data sharing platform stores data in files and in databases. There is both a relational database (SQL) and a graph database (RDF) between which the data contents are synchronized. The tools can access the data sharing platform in several different ways: by downloading IFC files and through SPAQRL queries, URI lookup and ordinary REST interface. Finally, the ontology modules are aligned with external ontologies through the alignment modules that import both the external and internal ontologies, and connect the terms across them with alignment axioms.

The characteristics of this architecture can be considered from different perspectives:

1. Data-centred interoperation approach avoid tight coupling: it is crucial to make the architecture open to new tools and to support the independent evolution of each tool. The Linked Data interfaces are based on standards and commonly agreed on and publicly available ontologies.
2. While the realization of the framework in BIM4EEB is based on a central data sharing platform, the framework itself would also work in a decentralized setting: there could well be multiple data providers that the tools could access. For instance, each of the tools could act as a data provider to other tools. Already in the current approach in BIM4EEB, each tool also stores its own data, either in a native format or according to the ontologies. The ontologies are primarily used in data sharing to provide semantic interoperability between tools.
3. Although decentralized approach is more flexible and open – and the direction of the future – a centralized platform has advantages at this early state of adoption of linked data technologies. At this stage, it is more straightforward to provide a satisfactory security and user management in a central platform with proven solutions. While the corresponding technologies for decentralized settings are available, they are still not in a wide-spread use or known by developers.
4. The BIM4EEB architecture also has the advantage of provided a smooth transition between the more established REST API based applications and Linked Data applications by synchronizing the data between SQL and RDF databases. Traditionally, this would need to be done inside an application; however, in BIM4EEB the data sharing platform takes care for it for the applications.

The characteristics of the framework are presented in Table 3. Each aspect of the framework shown in the table is further elaborated in subsequent subsections.

5.2 Required functions

The tools that want to be compliant with the framework need to understand the basic Semantic Web representations – such as RDF, SPARQL, and OWL – and implement the basic capabilities needed for Semantic Web aware software, as detailed below.

Framework-compliant tools can act in two different roles:

- *Data provider*: A data producer or publisher, anyone who makes data accessible for others. In Figure 26 the tools can upload the data they have produced to the data management platform, and thereby delegate the data provider role to it.
- *Data consumer*: anyone who accesses data provided by others.

Table 3: Linked Data and ontology framework for renovation

Objectives	To enable semantic interoperability between independent systems in the renovation domain that implement the functionalities and comply with the conventions specified in the framework	
Required functions	Each system that **provides data** to other systems must: P1. Implement the Linked Data Principles (Berners-Lee, 2006) for data sharing P2. Represent shared data (maybe converted from native data) as RDF graphs with links to additional data P3. Support the specified interfaces (SPARQL endpoint, URI Lookup) P4. Use the specified set of shared ontologies to define the types/properties of entities in shared data	
	Each system that **consumes data** from other systems must: C1. Be able to access the data using specified identifiers, query language, and interfaces C2. Be able to parse the received data and query results C3. Use the specified set of shared ontologies to interpret the types/properties of objects in shared data	
Conventions	**Identifiers**	HTTP URIs, using the HTTPS protocol If an object has a GUID, the URI should be based on the GUID
	Data formats	Turtle for ontologies JSON-LD, TriG and Turtle for data
	Interfaces	SPARQL Endpoint URI Lookup REST API GraphQL – For complex queries on the REST API
	Existing ontologies	ifcOWL – BIM models (building objects, their relations and properties) ICDD Container ontology – Exchange of interlinked datasets OWL-Time – Time concepts (time intervals, instants, temporal relations) PROV-O – Provenance metadata QUDT – Units of measure and quantity kinds BFO – Fundamental categories FOAF, Org and PROV-O – Agents and organizations FIBO – Legal concepts for assets and ownerhsip SSN/SOSA – Sensor data and sensor networks Saref – Devices, sensors and sensor datas
	Conceptual gaps that need to be covered by additional ontologies	G1. Main entities interlinked with renovation activities, external systems, and classifications G2. The types of agents and their roles, capabilities, and production rates G3. Information entities: models, plans, renovation measures, indicators, notifications G4. Different contexts of information: planned/actual, as-designed/as-built, scenarios, LODs G5. Variables and constraints to capture management knowledge and evolving designs/plans G6. Occupant behavior and profiles for requirements and evaluation of renovation scenarios, including occupant comfort, indoor air quality, and y G7. Building acoustics, a property affected by renovation measures G8. Energy efficiency and energy systems, central aspects of energy renovations G11. Building materials and their layering, which affects the energy efficiency G11. Building lifecycle and levels of detail, as concrete frameworks
Practices	Ontology modularization and reuse: The overall ontology is divided into logical modules based on the vertical and horizontal segmentation approach. All external references made explicit in separate alignment modules.	
	Ontologies definition: OWL2 DL profile	
	Ontology metadata provided: License CC-BY, preferred prefixes	
	Drafting of ontologies: CMapTools	
	Editing of ontologies: Protégé and a text editor	

	Documenting of ontologies: pyLODE
	Publishing of ontologies: w3id.org/digitalconstruction/
	Hosting of ontologies: GitHub Pages
	Maintenance of ontologies: Support the continuous evolution of ontologies with versioning

Many tools act in both roles: they consume data from others and provide some data for others.

Data providers must:

P1. Implement the Linked Data Principles (Berners-Lee, 2006) for data sharing.
P2. Represent the shared data as RDF graphs with links to additional data.
P3. Support the specified interfaces (SPARQL endpoint, URI Lookup).
P4. Use the specified set of shared ontologies to define the types/properties of entities in shared data.

When providing data to other, a tool must provide it in some serialization of RDF. If the data is hosted in the native format, it must be first converted to RDF.

Each system that consumes data from other systems must:

C1. Be able to access the data using specified identifiers, query language, and interfaces
C2. Be able to parse the received data and query results
C3. Use the specified set of shared ontologies to interpret the types/properties of objects in shared data

The requirements for tools accessing the data from others is to be able to work with RDF: to make SPARQL queries, to be able to process the results, whether in SPARQL result format or in a fragment of an RDF graph. Moreover, it should be aware of the ontologies used in the framework to be able to interpret the meaning of the data exchange sufficiently well to achieve semantic interoperability.

5.3 Conventions adopted by framework-compliant tools

Identifiers. The independent entities in the datasets published according to the framework must use HTTP URIs as their primary identifiers. If an entity has been associated with a GUID before its publication (such as in IFC datasets), the URI should be derived from the GUID.

The URIs should not be version or file specific but remain the same from version to version, complying to the approach in which GUIDs are used in IFC.

If a new identifier must be created for the same object for technical reasons, the old and new identifiers can be linked to each other with the owl:sameAs relation maintained in the dataset where the new object first appears.

Data formats. The ontologies compliant with the framework are accessible at least in the Turtle serialization. The RDF data shared according to the conventions of the framework can be in following serializations: JSON-LD, Turtle, or TriG. At least one graph-sensitive serialization (e.g., JSON-LD or TriG) must be supported. The results of SPARQL queries can be exchanged in any SPARQL result formats.

Interfaces. The data providers must support two interfaces: SPARQL endpoint and URI lookup. URI lookup is supposed to return, in the minimum, the content of SPARQL DESCRIBE query. Additionally, they can support GraphQL interface that maps to RDF data.

Ontologies. Data can be exchanged using the terminology defined in

- Digital Construction Ontologies – covering the overall structure and the gap areas in Table 3

complemented with one of the aligned ontologies:

- ifcOWL – BIM models (building objects, their relations and properties);
- ICDD – Containers for the exchange of multiple interlinked datasets;
- DCAT2 – Data catalogs;

- BOT – Building topology;
- OWL-Time – Temporal concepts (time intervals, instants, temporal relations);
- PROV-O – Provenance metadata;
- QUDT – Units of measure and quantity kinds;
- OPM – Property objectification;
- BFO – Fundamental categories;
- FOAF, Org and PROV-O – Agents and organizations;
- FIBO – Legal roles of agents, to define assets and ownership;
- SSN/SOSA – Sensor data and sensor networks;
- Saref, Saref4Bldg – Devices, sensors, and sensor data.

The tools participating in the system need to understand these ontologies to the extent that is relevant for the data that they consume and provide.

Alignment. These ontologies have been aligned with the Digital Construction Ontologies (see BIM4EEB-D3.6, 2021): they can be used together with each other by loading the appropriate alignment module. For instance, a renovation project and its information model can be represented according to Digital Construction Ontologies but the details of BIM models according to the ifcOWL, and specific units according to QUDT. As an example, the object :ti1 below represent a time point in the class dice:TimeInstant of DiCon Entities module. In the OWL-Time ontology there is a class time:Instant that is the domain of a datatype property time:inXSDDateTimeStamp. In the alignment module the classes dice:TimeInstant and time:Instant have been declared equivalent. Now it is possible to create the following kinds of entities where all ontological relations are understood correctly by an ontology reasoner.

```
@prefix dice: <https://w3id.org/digitalconstruction/0.5/Entities#> .
@prefix time: <https://www.w3.org/2006/time/> .
@prefix :     <http://ex.com/id/>

:ti1 a dice:TimeInstant ;
     time:inXSDDateTimeStamp ”2021-02-01T09:00Z” .
```

Once this code, the two ontologies and the alignment module have been loaded into an RDF databasse, the semantic connections are established and can be used in ontology reasoning.

5.4 Practices for the development and publication of the framework

Ontology definition. The Digital Construction Ontologies are defined using OWL2, with the objective to stay within the OWL2 DL profile. The primary source code of the ontologies would be maintained in Turtle serialization.

Ontology license. Ontologies will be published under the Creative Commons Attribution 4.0 International License (CC-BY 4.0).

Modularization and reuse. The Digital Construction Ontologies will be defined as a set of modules to keep them manageable. The modules would be defined according to the vertical and horizontal segmentation approach used by the Semantic Sensor Network ontology.

The following principles will be followed:

1) All external ontology references are made explicit through separate alignment modules. However, references to external vocabularies (instances) that will be allowed in the ontologies themselves.
2) The ontologies can be segmented vertically as follows:
    - Ontology B imports ontology A.

- Ontology B *refines the terms* already defined in ontology A by creating additional subclasses, properties, restrictions, or alignments.
- The ontology B will have a narrower user base than ontology A (a subset).
- The motivation is to support selected use cases better.

3) The ontologies can be segmented horizontally as follows:
    - Ontology B can import ontology A.
    - Ontology B *defines complementary terms* to those of ontology A by creating new classes and properties that connect them to the classes of ontology A.
    - The ontologies A and B together will have a broader user base.
    - The motivation is to expand the set of supported use cases.

Ontology alignment is always based on vertical segmentation. The alignment module will:

- Import all the ontologies to be aligned.
- Define additional axioms that refer to the terms defined in the aligned ontologies.

Benefits which will result from reusing existing, well-known ontologies through alignment are:

- *Saving work* – The duplication of work that has already been done can be avoided.
- *Broader perspective* – Since many well-known existing ontologies have a larger developer community behind them, they naturally tend to incorporate a richer variety of perspectives in their conceptualizations.
- *Higher quality* – Well-known existing ontologies have been tested and validated in practice and are therefore less likely to contain obvious errors.
- *Tool support* – There can be existing tools or tools emerging from well-known, existing ontologies:

Reuse and alignment can also create complications:

- *Different perspective* – Different ontologies can approach the domain from remote enough perspectives to render their alignment impossible, or they can employ incompatible modelling assumptions and approaches.
- *Inconsistency* – The ontologies to be aligned may not be compatible with each other, and the result of the attempted reuse is an inconsistent ontology.
- *Evolution* – The aligned ontologies may be under development, therefore requiring resources to keep alignments up-to-date.

Ontology metadata. The following metadata pattern for the ontologies will be used:

```
@prefix :       {namespaceURI} .
@prefix dc:     <http://purl.org/dc/elements/1.1/> .
@prefix owl:    <http://www.w3.org/2002/07/owl#> .
@prefix rdf:    <http://www.w3.org/1999/02/22-rdf-syntax-ns#> .
@prefix xml:    <http://www.w3.org/XML/1998/namespace> .
@prefix xsd:    <http://www.w3.org/2001/XMLSchema#> .
@prefix rdfs:   <http://www.w3.org/2000/01/rdf-schema#> .
@prefix vann:   <http://purl.org/vocab/vann/> .
@prefix terms:  <http://purl.org/dc/terms/> .

{ontologyURI} rdf:type   owl:Ontology ;
        owl:imports     {otherOntologyURI}, {otherOntologyURI}, ... ;
        dc:date {publicationDate} ;
        dc:title  {ontologyName} ;
        terms:created   {creationDate} ;
        terms:creator   {creator} , {creator}, ... ;
        rdfs:comment    {description} ;
        rdfs:label      {ontologyName} ;
        vann:preferredNamespacePrefix        {prefix} ;
        vann:preferredNamespaceUri           {namespaceURI} ;
        terms:license   <https://creativecommons.org/licenses/by/4.0/> ;
        rdfs:seeAlso    <https://w3id.org/digitalconstruction/> .
```

Tools for ontology development. Only freely available tools for ontology development will be used. The drafting of ontologies can be done with a concept mapping tool CMapTool. Ontologies will be edited using the Protégé ontology editor and with suitable text editors (such as Emacs), when there are changes that are difficult to do with Protégé. When an ontology is defined, consistency checking is performed using OWL reasoners (such as Hermit or Pellet), that work as plug-ins in Protégé. Ontology is documented with pyLODE documentation generation tool, and with other associated tools.

Hosting and publishing ontologies. Digital Construction Ontologies will be versioned and the version numbers will be visible in the URIs. The version will always be the same for all ontology modules. The ontologies will be published as presented in Table 4.

Table 4: Addresses of the Digital Construction Ontologies

Name	Digital Construction Ontologies	Comments
Acronym	DiCon	
Address	https://w3id.org/digitalconstruction/	Addressed of the ontologies; Redirection and content negotiation
Versioning	https://w3id.org/digitalconstruction/0.5/	Global versioning of the whole ontology suite
Code repository	https://github.com/digitalconstruction	Hosting and version control the ontologies at development stage
Publication page	https://digitalconstruction.github.io/	Web-publication of ontologies

The code repository is maintained in the GitHub in organisation called ‘digitalconstruction’ (address: https://github.com/digitalconstruction) that contains several GitHub repositories, one for each ontology module (such as https://github.com/digitalconstruction/Entities for the Entities module) and one for all alignments (https://github.com/digitalconstruction/Alignment). Ontology developer interacts with the code repositories using git commands by checking repositories out to a local repository, modifying them, and checking them back to GitHub.

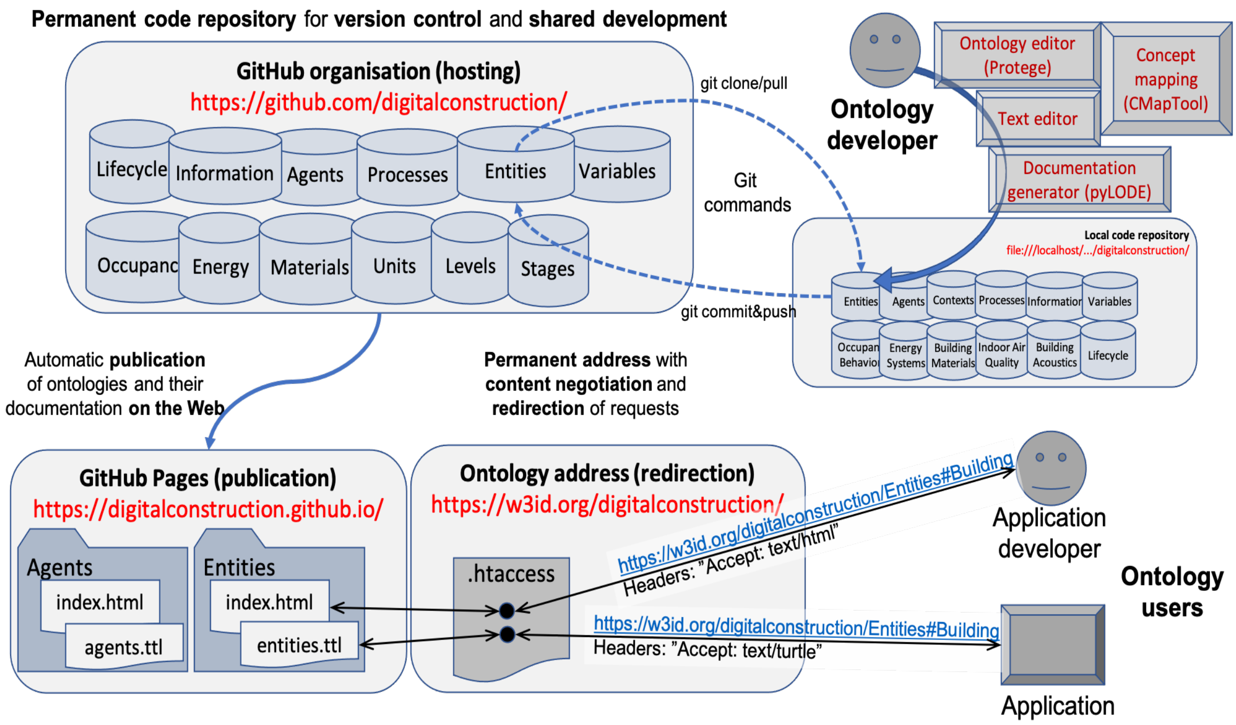


Figure 27: Hosting and publication of Digital Construction Ontologies

The content updated to the GitHub repositories is automatically published in the GitHub Pages in the address https://digitalconstruction.github.io/. For example, the Entities module will be published in the address https://digitalconstruction.github.io/Entities.

Since only static content can be published at the GitHub Pages and ontology publishing requires content negotiation (a dynamic capability), a permanent identifier for ontologies have been acquired from W3C service at w3id.org and content negotiation with related redirections is placed in a .htaccess file there. The permanent address is https://w3id.org/digitalconstruction/. The .htaccess file also takes care of the proper management of versions.

The overall setup is shown in Figure 27. There are two user roles: ontology developer and ontology user. There can be two kinds of ontology users: application developer and application. After the publication of the Entities ontology, the concept Building in that ontology would be available at the following address:

- https://w3id.org/digitalconstruction/Entities#Building,

or when referring to the concept in a particular version of the DiCon ontologies, at the address:

- https://w3id.org/digitalconstruction/0.5/Entities#Building.

The .htaccess file in the address https://w3id.org/digitalconstruction/.htaccess is written in a manner that when a user types the URI above in a browser, the content negotiation redirects the call to documentation file

- https://digitalconstruction.github.io/Entities/v/0.5/index.html#Building.

The html-file is returned and rendered in the browser, allowing user to explore the ontology in a human-readable form. However, when an application program accesses the same URI, the content negotiation will return a machine-readable version of the ontology in the file entities.ttl.

Ontology URI patterns. The ontology URIs are formed according to the patterns in Table 5.

Table 5: URI patterns of Digital Construction Ontologies

https://w3id.org/digitalconstruction https://w3id.org/digitalconstruction/index.html	The home page of the latest published version of the ontology
https://w3id.org/digitalconstruction/0.5 https://w3id.org/digitalconstruction/0.5/index.html	The home page of a particular version of the ontology
https://w3id.org/digitalconstruction/Entities	The documentation page of the latest published version of an ontology module (Entities)
https://w3id.org/digitalconstruction/0.5/Entities	The documentation page of a particular version (0.5) of an ontology module (Entities)
https://w3id.org/digitalconstruction/Entities/entities.ttl	The definition file of the latest published version of an ontology module (Entities)
https://w3id.org/digitalconstruction/0.5/Entities/entities.ttl	The definition file of a particular version (0.5) of an ontology module (Entities)
https://w3id.org/digitalconstruction/Entities#Activity	The latest published version of a term definition (Activity) in the given ontology module (Entities)
https://w3id.org/digitalconstruction/0.5/Entities#Activity	The term definition (Activity) in a particular version (0.5) of the given ontology module (Entities)
https://w3id.org/digitalconstruction/Alignment/SSN/	The latest published version of the alignment with a given external ontology (SSN)
https://w3id.org/digitalconstruction/0.5/Alignment/SSN/	A particular version (0.5) of the alignment with a given external ontology (SSN)
https://w3id.org/digitalconstruction/Alignment/SSN/dicon-ssn.ttl	The definition file of the latest published version of the alignment with an external ontology (SSN)
https://w3id.org/digitalconstruction/0.5/Alignment/SSN/dicon-ssn.ttl	The definition file of a particular version (0.5) of the alignment with an external ontology (SSN)

6 Bibliography

(Allen, 1983) Allen, J.F., 1983. Maintaining knowledge about temporal intervals. Communications of the ACM, 26(11):832-843.

(Almeida , 2012) Almeida, M. B., Slaughter, L., & Brochhausen, M. (2012). Towards an ontology of document acts: Introducing a document act template for healthcare. In *OTM Confederated International Conferences" On the Move to Meaningful Internet Systems"* (pp. 420-425). Springer, Berlin, Heidelberg

(Abels, 2006) Abels, S., Ahlemann, F., Hahn, A., Hausmann, K., & Strickmann, J., 2006. PROMONT – A project management ontology as a reference for virtual project organizations. In OTM Confederated International Conferences" On the Move to Meaningful Internet Systems" (pp. 813-823). Springer, Berlin, Heidelberg.

(Adapt4ee, 2014) Adapt4ee Occupancy Ontology. URL: http://www.adapt4ee.eu/adapt4ee/results/ontologies.html

(Arp, 2015) Arp, R., Smith, B., & Spear, A. D. (2015). *Building Ontologies with Basic Formal Ontology*. Mit Press.

(Abukwaik, 2014) Abukwaik H., Taibi D., Rombach D. (2014) Interoperability-Related Architectural Problems and Solutions in Information Systems: A Scoping Study. In: Avgeriou P., Zdun U. (eds) Software Architecture. ECSA 2014. Lecture Notes in Computer Science, vol 8627. Springer, Cham. https://doi.org/10.1007/978-3-319-09970-5_27

(Ardjani, 2015) Ardjani, F., Bouchiha, D., & Malki, M. (2015). Ontology-Alignment Techniques: Survey and Analysis. International Journal of Modern Education & Computer Science, 7(11).

(Baader, 2003) Baader, F., Calvanese, D., McGuinness, D., Patel-Schneider, P., & Nardi, D. (Eds.). (2003). *The description logic handbook: Theory, implementation and applications*. Cambridge university press.

(Ballard, 2000) Ballard, G., 2000. The Last Planner System of Production Control, Doctoral dissertation, University of Birmingham.

(Battle, 2011) Battle, R., & Kolas, D. (2011). Geosparql: enabling a geospatial semantic web. *Semantic Web Journal*, *3*(4), 355-370.

(Berners-Lee, 2006) Berners-Lee, T., 2006. Linked Data – Design Issues, W3C Note, URL: http://www.w3.org/DesignIssues/LinkedData.html.

(BIM4EEB, 2018) BIM4EEB project (2018). *BIM based Fast Toolkit for Efficient Renovation in Buildings*. Grant Agreement 820660 BIM4EEB. European Commission.

(BIM4EEB-D3.6, 2022) BIM4EEB project, Integrated Linked Data Modeling and Sharing Framework, Grant Agreement 820660 BIM4EEB. European Commission.

(BIMForum, 2018) BIMForum, 2018. Level Of Development (LOD), Specification Part I & Commentary.

(Borgo, 2009) Borgo, S., & Masolo, C. (2009). Foundational choices in DOLCE. In Handbook on ontologies (pp. 361-381). Springer, Berlin, Heidelberg.

(Brickley, 2014a) Brickley, D., Guha, R. V., & McBride, B. (2014). RDF Schema 1.1. *W3C recommendation*, *25*, 2004-2014. URL: https://www.w3.org/TR/rdf-schema/

(Brickley, 2014b) Brickley, D., & Miller, L. (2007). FOAF vocabulary specification 0.99. W3C Namespace Doc, W3C.

(Callahan, 2013) Callahan, A., & Dumontier, M. (2013). Ovopub: Modular data publication with minimal provenance. *arXiv preprint arXiv:1305.6800*.

(Ceusters, 2015) Ceusters, W., & Smith, B. (2015). Aboutness: Towards foundations for the information artifact ontology. URL: https://philpapers.org/archive/CEUATF.pdf

(Cox, 2017) Cox, S. and Little, C., 2017. Time Ontology in OWL. W3C Recommendation. URL: https://www.w3.org/TR/owl-time/.

(Cyganiak, 2014)	Cyganiak, R., Wood, D. and Lanthaler, M., 2014. RDF 1.1 Concepts and Abstract Syntax, W3C Recommendation. URL: https://www.w3.org/TR/rdf11-concepts/.
(Dodds, 2012)	Dodds L. &; Davis I. (2012). *Qualified Relation*. Modified 31 May 2012, accessed 2019-11-25. URL: http://patterns.dataincubator.org/book/qualified-relation.html
(EIF, 2017)	European Union (2017). New European Interoperability Framework, Promoting seamless services and data flows for European public administrations. Publications Office of the EU, Luxembourg.
(EU-EED, 2018)	Directive 2012/27/EU of the European Parliament and of the Council of 25 October 2012 on energy efficiency (updated 2018). Document 02012L0027-20180709.
(EU-EPBD, 2010)	Directive 2010/31/EU of the European Parliament and of the Council of 19 May 2010 on the energy performance of buildings. (2010). Document 32010L0031.
(Farias, 2014)	De Farias, T. M., Roxin, A., & Nicolle, C. (2014). A rule based system for semantical enrichment of building information exchange. RuleML 2014, Prague, Czech.
(Farias, 2015)	de Farias, T. M., Roxin, A., & Nicolle, C. (2015). IfcWoD, semantically adapting IFC model relations into OWL properties. *arXiv preprint arXiv:1511.03897*.
(Farias, 2018)	de Farias, T. M., Roxin, A., & Nicolle, C. (2018). A rule-based methodology to extract building model views. *Automation in Construction*, *92*, 214-229.
(Fikes, 1971)	Fikes, R. E., & Nilsson, N. J. (1971). STRIPS: A new approach to the application of theorem proving to problem solving. *Artificial intelligence*, *2*(3-4), 189-208.
(Fox, 2003)	Fox, M., & Long, D. (2003). PDDL2. 1: An extension to PDDL for expressing temporal planning domains. *Journal of artificial intelligence research*, *20*, 61-124.
(Garcia-Lopez, 2017)	Garcia-Lopez, N., 2017. An Activity and Flow-based Construction Model for Managing On-site Work, Doctoral dissertation, Stanford University.
(Ghallab, 1998)	Ghallab, M., Howe, A., Knoblock, C., McDermott, D., Ram, A., Veloso, M., ... & Wilkins, D. (1998). PDDL—the planning domain definition language. *AIPS-98 planning committee*, *3*, 14.
(Ghallab, 2016)	Ghallab, M., Nau, D., & Traverso, P. (2016). *Automated planning and acting*. Cambridge University.
(Groth, 2010)	Groth, P., Gibson, A., & Velterop, J. (2010). The anatomy of a nanopublication. *Information Services & Use*, *30*(1-2), 51-56.
(Hoang, 2015)	Hoang, N. V. (2015). IFC-to-Linked Data Conversion: Multilayering Approach, in: the Third International Workshop on Linked Data in Architecture and Construction, Eindhoven, Netherlands.
(IOF, 2019)	Industrial Ontologies Foundry (2019). URL: https://www.industrialontologies.org/, at 2021-05-31.
(ISO12006-2, 2015)	ISO (2015) Building construction - Organization of information about construction works - Part 2: Framework for classification
(ISO16739-1, 2018)	ISO (2018). Industry Foundation Classes (IFC) for data sharing in the construction and facility management industries -- Part 1: Data schema. URL: https://www.iso.org/standard/70303.html
(ISO19650-1, 2018)	ISO (2018). Organization and digitization of information about buildings and civil engineering works, including building information modelling (BIM) -- Information management using building information modelling. Part 1: Concepts and principles.
(ISO19650-2, 2018)	ISO (2018). Organization and digitization of information about buildings and civil engineering works, including building information modelling (BIM). Information management using building information modelling. Part 2: Delivery phase of the assets.
(ISO1675, 2018)	ISO (2018). Data structures for electronic product catalogues for building services, Part 1: Concepts, architecture and model, Part 2: Geometry. ISO.
(ISO21597, 2020)	ISO (2020). Information container for linked document delivery — Exchange

	specification — Part 1: Container. URL: https://www.iso.org/standard/74389.html.
(ISO/IEC21838, 2021)	ISO (2021). Information technology — Top-level ontologies (TLO) — Part 2: Basic Formal Ontology (BFO). URL: https://www.iso.org/standard/74572.html.
(ISO 45001, 2018)	ISO (2018). Occupational health and safety management systems — Requirements with guidance for use
(Karray, 2019)	Karray, M.H., Ameri, F., Hodkiewicz, M., & Louge, T. (2019). ROMAIN: Towards a BFO compliant reference ontology for industrial maintenance. *Applied Ontology*, 1-24.
(Kenley, 2010)	Kenley, R. and Seppänen, O., 2010. Location-based Management System for Construction: Planning, Scheduling and Control. Spon Press. London, New York.
(Knublauch, 2017a)	Knublauch, H., & Kontokostas, D. (2017). Shapes constraint language (SHACL). *W3C Recommendation*, *20*.
(Knublauch, 2017b)	Knublauch, H., & Allemang, D. & Steyskal, S. (2017). SHACL Advanced Features. W3C Working Group Note 08 June 2017, W3C.
(Koskela, 1999)	Koskela, L. J., 1999. Management of production in construction: a theoretical view. 7th Annual Conference of the International Group for Lean Construction Berkeley, CA. 241–252.
(Koskela, 2000)	Koskela, L., 2000. An exploration towards a production theory and its application to construction. VTT Publications 408, Espoo, Finland.
(Lebo, 2013)	Lebo, T., Sahoo, S. and McGuinness, D., 2013. PROV-O: The PROV Ontology, W3C Recommendation. URL: https://www.w3.org/TR/prov-o/.
(Levesque, 1987)	Levesque, H. J., & Brachman, R. J. (1987). Expressiveness and tractability in knowledge representation and reasoning 1. *Computational intelligence*, *3*(1), 78-93.
(Meditskos, 2013)	Meditskos, G., Dasiopoulou, S., Efstathiou, V., & Kompatsiaris, I. (2013). Ontology patterns for complex activity modelling. In *International Workshop on Rules and Rule Markup Languages for the Semantic Web* (pp. 144-157). Springer, Berlin, Heidelberg.
(Menzel, 2013)	Menzel, K.; Weise, M., Liebich, T.; Valmaseda, C. 2013 "Capabilities of IFC4 for Advanced Building Performance Management": in: Proceedings of 2nd Central European Symposium on Building Physics, ed. Mahdavi and Martens (Vienna University of Technology, 2013, pages 467 to 475).
(Menzel, 2015)	Menzel, Karsten; Browne, Donal; Deng, Siyu 2015, "Performance Indicators to evaluate buildings' systems' performance": in: eWork and eBusiness in AEC; Martens, Mahdavi and Scherer (Taylor and Francis Group, London, pages 751 to 759).
(Menzel, 2016)	Menzel, Karsten; Hryshchenko, Andriy, 2016, "Business Modelling for Smart Continual Commissioning in ESCO Set-ups": in: Collaboradion in a data-rich world ed. L. Camarinha-Matos, H. Afsarmanesh, R. Fornasciero (Springer, Heidelberg, pages 307 to 314).
(Miles, 2009)	Miles, A., & Bechhofer, S. (2009). SKOS simple knowledge organization system reference. *W3C recommendation*, *18*, W3C. URL: https://www.w3.org/TR/skos-reference/
(Motik, 2012)	Motik B., Grau B.C., Horrocks I., Wu Z., Focoue A. & Lutz C. (2012). OWL 2 Web Ontology Language Profiles (Second Edition), W3C Recommendation, URL: https://www.w3.org/TR/owl2-profiles/.
(Nguyen, 2014)	Nguyen, V., Bodenreider, O., & Sheth, A. (2014). Don't like RDF reification?: Making statements about statements using singleton property. In *Proceedings of the 23rd International Conference on World Wide Web* (pp. 759-770). ACM.

(Pauwels, 2016a) Pauwels, P., & Walter, T., 2016. EXPRESS to OWL for construction industry: Towards a recommendable and usable ifcOWL ontology. Automation in Construction, 63, pp.100–133.

(Pauwels, 2016b) Pauwels, P., & Roxin, A. (2016). SimpleBIM: From full ifcOWL graphs to simplified building graphs. In *Proceedings of the 11th European Conference on Product and Process Modelling (ECPPM)*.

(Pauwels, 2017) Pauwels, P., de Farias, T. M., Zhang, C., Roxin, A., Beetz, J., De Roo, J., & Nicolle, C. (2017). A performance benchmark over semantic rule checking approaches in construction industry. *Advanced Engineering Informatics*, *33*, 68-88.

(Pease, 2002) Pease, A., Niles, I., & Li, J. (2002, July). The suggested upper merged ontology: A large ontology for the semantic web and its applications. In *Working notes of the AAAI-2002 workshop on ontologies and the semantic web* (Vol. 28, pp. 7-10).

(PMI, 2021) Project Management Institute (2021). What is project management? URI: https://www.pmi.org/about/learn-about-pmi/what-is-project-management (viewed 2021-05-02)

(Prud'hommeaux, 2014) Prud'hommeaux, E. and Carothers, G., 2014. RDF 1.1 Turtle – Terse RDF Triple Language. W3C Recommendation. URL: https://www.w3.org/TR/turtle/.

(Radinger, 2013) Radinger, Andreas; Rodriguez-Castro, Bene; Stolz, Alex; Hepp, Martin: BauDataWeb: The Austrian Building and Construction Materials Market as Linked Data, in: Proc of the 9th Int'l Conf on Semantic Systems (I-SEMANTICS 2013), Graz, Austria, September 4-6, 2013. New York: ACM, 2013.

(Rasmussen, 2017) Rasmussen, M. H., Pauwels, P., Hviid, C. A., & Karlshøj, J. (2017). Proposing a central AEC ontology that allows for domain specific extensions. In *2017 Lean and Computing in Construction Congress*.

(Rasmussen, 2018) Rasmussen, M., Lefrançois, M., Bonduel, M., Hviid, C., & Karlshøj, J. (2018). OPM: An ontology for describing properties that evolve over time. In proc. of 6th *Linked Data in Architecture and Construction Workshop*. London, UK.

(Reynolds, 2014) Reynolds, D. (2014). The Organization Ontology. W3C Recommendation. URL: https://www.w3.org/TR/vocab-org/

(Rinne, 2012) Rinne, M., Nuutila, E., and Törmä, S., 2012. INSTANS: high-performance event processing with standard RDF and SPARQL. In Proc. of the 2012 International Conference on the Semantic Web, Posters & Demonstrations Track. Volume 914, pp. 101-104.

(Rudnicki, 2019) Rudnicki R. (2019). *An Overview of the Common Core Ontologies*. CUBRC Inc, Buffalo, US. URL: https://www.nist.gov/document/nist-ai-rfi-cubrcinc004pdf.

(Schadow, 2017) Schadow, G. and McDonald, C. J., 2017. The unified code for units of measure (UCUM). Version 2.1. URL: http://unitsofmeasure.org/ucum.html, Regenstrief Institute and Indiana University School of Informatics.

(Singh, 2005) Singh, M. P., & Huhns, M. N. (2005). Service-oriented computing (pp. 410-495). Chichester: Wiley.

(Smith, 1996) Smith S., Lassila O., and Becker M. , 1996. A Scheduling Ontology: Informal Concept Descriptions, ICLL Working Paper, Carnegie Mellon University.

(Smith, 2007) Smith, B., Ashburner, M., Rosse, C. *et al.* (2007). The OBO Foundry: Coordinated evolution of ontologies to support biomedical data integration. *Nat Biotechnol* **25,** 1251–1255, doi:10.1038/nbt1346.

(Smith, 2015) Smith, B., Almeida, M., Bona, J., Brochhausen, M., Ceusters, W., Courtot, M., ... & Hogan, W. Basic Formal Ontology 2.0 – Specification and User's Guide (2015). *URL: https://github.com/BFO-ontology/BFO/raw/master/docs/bfo2-reference/BFO2-Reference.pdf*, *755*.

(Smith, 2019)	Smith, B. (2019). Ontology as Product-Service System: Lessons Learned from GO, BFO and DOLCE. *Proceedings of the 10th International Conference on Biomedical Ontology* (ICBO), Buffalo, NY.
(Seppänen, 2010)	Seppänen, O., Glenn, B. and Pesonen, S., 2010. The combination of Last Planner System and Location-Based Management System, Lean Construction Journal.
(Seppänen, 2014a)	Seppänen O., 2014. A comparison of Takt time and LBMS planning methods. Proc. 22nd Annual Conf. of the Int'l. Group for Lean Constr. (IGLC 22), Oslo, Norway.
(Seppänen, 2014b)	Seppänen, O., Evinger, J. and Mouflard M., 2014. Effects of the Location-Based Management System on Production Rates and Productivity. Construction Management and Economics, 32(6).
(Ses, 2015)	READY4SmartCities (2015) “Repository for Semantic Web and ontologies”, CSA Horizon 2020; http://smartcity.linkeddata.es/index.html (last accessed: March 2019).
(Shree, 2019)	Shree, V., Marwaha, B. M., & Awasthi, P. (2019) Assessment of Indoor Air Quality in Buildings using CFD: A Brief Review. *International Journal of Mathematical, Engineering and Management Sciences* Vol. 4, No. 5, 1154–1168. https://dx.doi.org/10.33889/IJMEMS.2019.4.5-091.
(Spyns, 2002)	Spyns, P., Meersman, R., & Jarrar, M. (2002). Data modelling versus ontology engineering. *ACM SIGMod Record*, 31(4), 12-17.
(UNI 11337, 2017)	UNI 11337- 4 (2017) - Building and civil engineering works - Digital management of informative processes - Part4: Evolution and development of information in models, documents and objects
(Vajda, 2019)	Vajda, J., Merrell, E., & Smith, B. (2019). Toward an Ontology of Commercial Exchange. URL: https://philarchive.org/archive/VAJTAO
(Wetherill, 2003)	Wetherill, M., 2003. Knowledge management for the construction industry: the e-cognos project. *Journal of Information Technology in Construction (ITCon)*, *7*(12), 183-196.
(Zhang, 2018)	Zhang, C., Beetz, J., & de Vries, B. (2018). BimSPARQL: Domain-specific functional SPARQL extensions for querying RDF building data. *Semantic Web*, (Preprint), 1-27.

7 ANNEX

Construction management use cases

Use Case	1 Master planning
Relevant requirements	• Continuation plan from the initiative phase • Renovation measures • Planning principles and execution strategies • Expected budget requirements • Data from past projects
Information entities	• The development of the master plan of a project taking into account the general planning principles, execution strategies (e.g., types of contracts, subcontracting, prefabrication), renovation measures, and budget constraints. • Initial estimation based on some measure of the size of the project, in later versions on quantity surveys.

Information exchange	• Master plan, including:
	  • Project team structure
	  • Work breakdown structure
	  • Target schedule for activities
	  • Allocation of responsibilities

Use Case	2 Task planning
Relevant requirements	• Design information: BIM models • Master plan / week plan • Project team composition • Estimates of activity size
Information entities	• Development of a location breakdown structure • Creation of location-specific activity decomposition • Requirements/assignments for exchangeable resources (types/capabilities of equipment and agents) • Connections of building objects to activities

Information exchange	• Concrete, executable activities including:
	  ▪ Connections of activities to building objects, locations, equipment, and agents
	  ▪ Activity durations and costs

Use Case	3 Progress coordination
Relevant requirements	• Master plan / week plan • Status updates / sensor data • Procurement data (types/instances) • Data about product-process connections
Information entities	• Providing operational guidance to actors based on the plans (what to do next/ what to do today) • Monitoring the execution of activities: start/end, positions of entities, issues encountered • Detection of deviations from plans • Gathering and recoding enhanced data during execution
Information exchange	• Enhanced product and process information • Detected issues, deviations, or other needs for plan revision

Use Case	4 Plan revisioning
Relevant requirements	• Master plan / week plan • Evolving data about entities (new versions/levels of BIM models, procurement data, appointments) • Previously planned revision schedule (e.g., time to create the next week's plan) • Detected issues, deviations, or other needs for plan revision
Information entities	• Creation of a new version of a plan • Approving the plan for execution • Managing the information so that the content of the new plan becomes basis of execution
Information exchange	• The new plan content • New information container for plan • Updated information model to activate the plan

Use Case	5 Week planning
Relevant requirements	• Detailed designs for construction • Location-breakdown structure • Master plan and previous week plans • Monitoring data about current week's activities • Connections of activities to building objects, locations, equipment, and agent • Activity durations
Information entities	• The activities to be executed during the next week are selected and scheduled • Optionally ensuring that the preconditions of the activities are predicted to be satisfied • The week plan is approved and activated to execution at the proper time
Information exchange	• Week plan for the next week • A new information container for the week plan • An updated information model where the new week plan is activated to execution at proper time

Use Case	6 Stakeholder coordination
Relevant requirements	• Week plans • Location breakdown structure • Product-process connection • Location-stakeholder connection
Information entities	• The execution of activities at different locations and specified times is recorded • Occupants and/or owners are notified about the activities
Information exchange	• Notifications and alerts to stakeholders

Energy-efficiency use cases

Use Case	**0 Initiative:** Preliminary decision for the renovation (go / no go decision)
Relevant requirements	• Financial feasibility requirements of the owner. • Monitoring data about the energy performance history. • Surveys about building conditions • Assesments as well as energy audit reports from facility management. • Background knowledge of the applicable renovation measures for the type of building to be renovated (for instance, based on national databases of renovation costs and savings).
Information entities	• Financial feasibility data from the owner. • Environmental target data from the owner. • Owners' Project Requirement (OPR) • Historical energy performance data • Building condition survey data • Building energy audit data • Typical renovation measures associated with cost and expected savings
Information exchange	The preliminary need for the renovation, **go / no go** -decision is made. The metrics are typically related to the **financial feasibility** of the renovation measures needed. The energy and cost efficiency is checked without detailed simulation according to the common knowledge of available renovation measures against the historical energy performance. The outcome of the phase is: • Go / no go -decision • The strategic environmental targets as part of Owners Project Requirements (OPR) are set for the initiation (these high level targets can be common with the 3.5-3.10).

Use Case	**1 Initiation:** Renovation project initiation
Relevant requirements	• The strategic environmental and financial targets of the owner (OPR). • Targets from the "3.5. Occupant's behavior and comfort" available • Targets from the "3.7 Indoor air quality" available • Monitored historical energy performance data is available
Information entities	• Owners Project Requirements (OPR) with numerical values • Historical energy performance data • Simplified representative model of the building for the energy simulations
Information exchange	The initiation consists of **target setting for the energy efficiency indicators to support the OPR** and their numerical values for the building to be renovated. The indicators are derived from the strategic environmental targets of owners. The first feasibility simulations against the environmental targets are made using energy simulation tool with simplified model geometry. The outcome of the phase is: • Numerical energy efficiency indicators (probably kWh/m²,a) as part of Key Performance Indicators (KPI), are agreed for the **heating, electricity and cooling** to be applied in the concept phase

Use Case	**2.1 Concept Design:** Quick calculation to find the conceptual design alternatives
Relevant requirements	• Owner's strategic financial and environmental goals (OPR). • Targets from the "3.5. Occupant's behavior and comfort" available • Targets from the "3.7 Indoor air quality" available • Architectural design alternatives concerning the energy efficiency available • A database of conceptual renovation measures (HVAC, structural and electrical)
Information entities	• Owners Project Requirements (OPR) with numerical values • Architectural model of the design alternatives • Conceptual HVAC, structural and electrical renovation measure models
Information exchange	The concept design identifies the initial idea in reaching the OPRs in the energy efficiency of the building. This phase drills into conceptual details of the possible design alternatives and their combinations (architect, HVAC, structures, electricity) to reach the financial feasibility and the expected energy efficiency, which is set by the owner. The outcome of the phase is: • The set of suitable renovation concepts (1-5) fulfilling the OPR about energy efficiency and financial constraints. The details are on the conceptual level.

Use Case	**2.2 Preliminary Design:** Preliminary energy simulations of design alternatives with a rough model
Relevant requirements	• The strategic environmental and financial targets of the owner (OPR). • Targets for the occupant comfort available • Targets for indoor air quality available • General technical energy efficiency data of the products used • The preliminary architectural design(s) related to the concepts available
Information entities	• Owners Project Requirements (OPR) with numerical values • Fast mapping model of the building • Architectural model of the design alternatives • Technical energy efficiency data of the products used
Information exchange	The preliminary design identified the concepts and requirements for the products that enable the targetted energy efficiency level. The design team members configure energy efficiency design solutions further based on the outputs from concept design phase. Renovation scenarios that conform with **the energy related OPR** for **the review to the owner**. The renovation scenario approved by the owner. The outcome of the phase is: • The renovation alternative approved by the owner • Required technical design documents to the building permit authorities. • Updated energy performance with conceptual selections to be used in renovation. • The simulated indoor environment conditions. Both energy performance and indoor environment conditions measured in terms of selected KPI.

Use Case	**2.3 Developed Design:** More detailed energy simulation of design alternatives with BIM models
Relevant requirements	• Owner's strategic environmental and financial goal setting (OPR). • Targets for Occupant's behavior and comfort by space type and purpose • Targets for Indoor air quality by space type and purpose • OPR's compiled to specific numerical energy efficency indicators (kWh/m²,a) • Specific BIM's (architecture, HVAC, structures, electricity) related to the renovation alternative approved by the owner. The BIM contains the space layout according to their purpose of use. • The general technical energy efficiency data of the products used
Information entities	• Owners Project Requirements (OPR) with numerical values • Specific BIM's (architecture, HVAC, structures, electricity) related to the renovation alternative • Technical energy efficiency data of the products used
Information exchange	The preliminary design consists of finding the real product candidates to show the selected efficiency. The energy efficiency is simulated more detailed according to the space and system layout of the building. The design team (architecture, HVAC, structures, electricity) is making together the space and system level design selections to fulfill the energy efficiency and indoor climate targets. The outcome of the phase is: • The breakdown of the technical energy efficiency details for the product selection (SFP, heat recovery efficiency, lighting efficiency, U-values, solar shading, air tightness, heat pump COP, etc.). • Updated energy performance with solutions to be applied in the renovation in terms of KPIs. • The simulated indoor environment conditions by space types and by the location (floor, facade orientation) in terms of KPIs.

Use Case	**2.4 Detailed Design:** More detailed simulation of the design alternatives based on a digital twin
Relevant requirements	• The strategic environmental and financial targets of the owner (OPR). • Targets for the Occupant's behavior and comfort by space type and purpose • Targets for the Indoor air quality by space type and purpose • OPR's compiled to more specific numerical energy efficency indicators (kWh/m²,a). • The developed design available. • The benchmark database of the realised operation and maintenance costs and savings of the product candidates to be used (lifetime database of the existing products). • The operation and maintenance strategy of the owner.
Information entitiess	• Quantitative OPRs • Digital twin model • Technical energy efficiency data of the selected products • Collaboration model to guarantee seamless real time editing and updating of the digital twin • Data for the operational phase
Information exchange	Creation of a **digital twin of the building** to be renovated. The team considers various existing product options to match the OPR e.g. the energy performance of the project's life cycle and cost of the future operation and maintenance and discusses the relevant design options with the owner to decide the most suitable one. The design team consults the operation and maintenance professionals to guarantee the energy performance and low maintenance costs during the lifetime of the selected products. The outcome of the phase is: • The products fulfilling the owners targets selected. • Technical energy efficiency specification and details for the selected products. • 'As designed' energy performance with products to be installed in the renovation in terms of KPIs. • The simulated indoor environment conditions. • Maintenance cost estimate of the selected technical products. • First version of the digital twin merging the domain specific designs.

Use Case	**3 Construction:** Implementation of the renovation measures ensuring their compliance with plans
Relevant requirements	• The strategic environmental and financial targets of the owner (OPR). • Targets for the Occupant's behavior and comfort by space type and purpose • Targets for the Indoor air quality by space type and purpose • OPR's compiled to more specific numerical energy efficency indicators (kWh/m²,a). • Detailed design of the digital twin • Plans for performance test • Operation and maintenance plan
Information entities	• Quantitative OPRs • The model for the digital twin data • Technical product data of the actual products selected for the renovation • Collaboration model of the work flow to guarantee seamless real time editing and updating of the digital twin from "As designed" to "As built" • Compliance and control data to guarantee the intented performance of the installed products • Handover data: performance tests, user guides, training material, service manuals
Information exchange	At the construction phase the design team continuously updates arisen fine tuning needs of the design into the digital twin of the renovated building and **checks and validates** the contractors **implementation** and its **compliance against the design** and against the OPR. The contractor has a direct connection to the BIM, of which 'as built' stage is continuously updated during the construction work by the contractor. The design team makes the automation system simulations with a digital twin to support tuning of the the control systems to be installed. The outcome of the phase is: • Updated energy performance with exact products installed in the renovation in terms of KPIs. • The simulated indoor environment conditions with installed products in terms of KPIs. • An updated operation, management and maintenance plan. • 'As built' digital twin to be used in the operational phase. • The product handover documents: performance tests, user guides, training material, service manual.

Use Case	**4 Building Use:** Evaluation of OPRs at the operational phase of the renovated building
Relevant requirements	• The strategic environmental and financial targets of the owner (OPR). • As-built operation and maintenance plan • As-built digital twin • The received handover materials of the installed products • User guides
Information entitiess	• Quantitative OPRs for operational phase • Digital twin • Process for seamless real-time updating of the digital twin from as-built to as-maintained • Compliance and control data to guarantee the intented performance of the replacement products during the operation • Handover data: performance tests, user guides, training material, service manuals
Information exchange	At this stage (during 20-30 years) the operation and maintenance personnel will operate the renovated building according to the maintenance plan. The performance is evaluated against the KPIs and energy services may be connected to the evaluations results. The construction contractors will handle and correct the performance claims arisen during the guarantee period of the renovation installations. The energy performance simulations are made with a digital twin to support predictable operation & maintenance and fine tuning of the controls during the operation. The smaller component changes during the lifecycle and their compliance against the agreed energy performance is checked every time when the replacement action is implemented. The outcome of the phase is: • The energy performance of the operation phase is inline with owners project requirements (OPR). • The satisfied end users related to the indoor environment conditions. • The 'As built and as operated and as maintained' -version of the digital twin.

Use Case	**5 End of Life:** Recycling of the products and materials of the renovated building
Relevant requirements	• Environmental strategy for recycling • As-maintained digital twin
Information entities	• The digital twin data of the building renovated • Process for seamless updating of the digital twin from as-maintained to the as-demolished • Recycling model
Information exchange	In case of the **no go decision** of the upcoming technical cycle of the building use, the digital twin of the building is delivered to the demolition contractor to support the separation of the building materials and products to be either recycled as a used product or to be decomposed for material re-use. The outcome of the phase is: • Zero waste to the landfills. • As-demolished digital twin. • Digital history of the recycled products and materials for the reuse (the technical energy efficiency details, possible detoriation and expected lifetime left).

Occupants behaviour and comfort use cases

Use Case	Early identification of behavioural & comfort parameters in building premises - Establishment of a comfort preserving framework for inhabitants
Relevant requirements	• Occupant is interested to control of indoor environmental conditions in premises • Occupant is interested to set a sustainable and comfort preserving framework in premises • Occupant’s agreement to share information concerning building operation, and comfort and behavioural preferences
Information entities	**Occupants configuration parameters:** Age, Gender Lifestyle, etc. **Occupancy profiling parameters:** presence, level of occupancy, etc. **Activity profiling parameters:** typical activities (drivers & actions) **Indoor environmental conditions**: temperature, humidity, luminance **Comfort preferences:** thermal comfort/visual comfort settings
Information exchange	Occupant is defined with a static profile (config parameters) and occupancy/activity profiling parameters Occupant directly set the preferences of indoor environmental conditions (comfort level) Occupant indirectly indicates the preferences in different environmental conditions in building premises.

Use Case	Establishment of a comfort preserving framework during the renovation process
Relevant requirements	• Occupant updates information about comfort conditions during renovation process • Occupant faces comfort or discomfort during the renovation process
Information entities	*As above but based on the specific conditions that apply during the renovation process*
Information exchange	*As above but based on the specific conditions that apply during the renovation process*

Use Case	Renovation tasks/processes to address occupant comfort conditions
Relevant requirements	• Designers of specific disciplines (structural, HVAC, MEP, waste, water) collates technical documentation from technical specialist consultants or from legislation about environmental and comfort conditions as required by applicable legislation. • Designers of specific disciplines (structural, HVAC, MEP, waste, water) produce architectural and technical plans and documentation describing the project to a level of detail as required for Planning or Building permit applications to address behavioural and comfort specific requirements
Information entities	**Comfort KPIs**: based on the applicable legislation with focus on environment and comfort criteria **Comfort KPI values:** definition of the bounds of indoor environment and comfort criteria
Information exchange	Expert or designer to specify the indoor environmental conditions standards

Use Case	Establishment of a comfort preserving framework post renovation process
Relevant requirements	• Occupant updates information about comfort conditions post renovation process • Occupant feels comfort/discomfort during post renovation process • Occupant uploads ad hoc information that may be useful for the renovation process
Information entities	**Indoor environmental conditions**: temperature, humidity, luminance **Comfort preferences:** thermal comfort/visual comfort settings **Comfort Related KPIs**: as identified in the national and international legislation with focus on environmental & comfort conditions in premises **Comfort KPI values:** as the boundaries for indoor environmental/comfort conditions in premises
Information exchange	Occupant directly set its preferences about indoor environmental conditions (comfort levels) Occupant indirectly indicates its preferences in different environmental conditions in building premises. Occupants behaviour and comfort profiling analysis based on IEC standards

Equipment and material use case

Use Case	To acquire preliminary material and equipment data from existing building
Relevant requirements	• Client develops an overall understanding of the status of the building. The existence of certain materials (e.g. asbestos) may have a serious impact on strategic decisions for the next steps • Client/Owner recruits an Adviser requesting an initial cost-benefit appraisal. Again, the non-availability of certain material or equipment will influence future strategic decisions
Information entities	**Material and Material Constituents:** Existence of certain material and link to main building elements **Equipment and Systems Inventory:** Existence of instances of certain classes without detailed analysis of descriptive attributes (e.g. Heating, Ventilation or AC-systems with capacity and year built)
Information exchange	Client/Owner or Archive Manager evaluates the quality of available BIM (CAD) models. Available information is filtered out and circulated to Client Adviser. Client Adviser executes an initial inspection and verifies/corrects information provided.

Use Case	Acquire detailed material and equipment data from existing building
Relevant requirements	SiteSurveyor (assisted by technical experts) executes different activities to identify, document, and model the status of the building, including it's core & shell systems and building services systems
Information entities	**Equipment and Systems Inventory:** Relevant building elements and DistributionElements are documented. Usually, building elements are documented in a BIM using their geometrical representation (depending on a previously agreed "Level of Detail". Building Services Systems are either documented through a schematic or (less frequently) using their geometrical representation. **Material and Material Constituents:** In case of existing buildings the specification of materials is not always possible. In case of systems to be replaced it is even not required. E.g. electrical supply or plumbing systems are often completely replaced in residential buildings. **Geometrical representation of building and distribution elements:** Data is acquired for load bearing building elements or for insulating components, such as facade elements. The geometrical representation of building services elements is usually only acquired from space consuming ductwork of air-based ventilation systems.
Information exchange	Client/owner will forward the existing documentation to the surveyor. Site Surveyour verifies (or acquires) the relevant model information during a site survey. The Level of Detail must be agreed in advance. The Site Survyor integrates the information and forwards the consolidated model to the Client/Owner.

Use Case	Define the kind of intervention, produce schedules, esteem cost based on materials and equipments
Relevant requirements	The Project Leader analyses the kind of intervention and the tasks to be developed. This is usually done by using different simulation packages.
Information entities	**Equipment and Systems Inventory:** Based on the documentation of the existing building major changes and additions can be added to the BIM. **Material and Material Constituents:** The documentation of materials used in the existing building helps to etimate demolition and recycling cost. **Geometrical representation of building and distribution elements:** Geometrical information helps to estimate the volume of material to be disposed.
Information exchange	Information Source: Surveyor, Project Leader Recipient of Information: Quantity Surveyor

Use Case	Agree to share information
Relevant requirements	The owner agrees to share information about the building
Information entities	**Representations of the building:** For information of tenants and negotiations with building authorities or the public
Information exchange	Owner (or Archtectural Designer on behalf of the owner) release the information (representations) to the public.

Use Case	To prepare the Concept Design (including its documentation)
Use Case	**To integrate work from discipline designers and produce merged results**
Relevant requirements	Architect creates proposals for concept design.
Information entities	**Equipment and Systems Inventory:** The core and shell and major system components are specified. **Material and Material Constituents:** Property sets can be included to an extent that general values from vendor-neutral databases are used. **Geometrical representation of building and distribution elements:** Appropriate geometrical representations are used.
Information exchange	Source of Information: Architectural Designer User of Information: Owner, Project Leader, Public Authorities.

Use Case	To prepare preliminary design documentation
Relevant requirements	Architectural Designer prepares graphic representations, produces preliminary drawings and collates information from technical specialists. Lead Designer integrates information into design documentation.
Information entities	**Equipment and Systems Inventory:** Major system components are specified. It is not expected that neither the number nor the type of terminal devices or DistributionFlowComponents of building services systems are specified. Usually, building services and automation systems are specified as schematics, i.e. using topological descriptions. **Material and Material Constituents:** Property sets can be included to an extent that general values from vendor-neutral databases are used. **Geometrical representation of building and distribution elements:** Appropriate geometrical representations are used.
Information exchange	Source of Information: Project Leader, Lead Designer User of Information: Owner, Local Authority

Use Case	Preparation of Developed Design Stage of BIM-Model
Relevant requirements	Project Leader prepares Design Drawings and produces specs. Architect gathers relevant technical documents.
Information entities	**Equipment and Systems Inventory:** Major system components are specified. BIM-elements are further detailed, e.g. the usage of "multi-layer wall elements" might be introduced or decisions about the number of glazing layers in windows is defined. **Material and Material Constituents:** Property sets to be included are defined on material level. However, the manufacturer of materials were not decided. **Geometrical representation of building and distribution elements:** Appropriate geometrical representations are available. Details about interfaces and connections might be still missing.
Information exchange	Source of Information: Architectural Designer User of Information: Project Leader, Clients

Use Case	Preparation of Detailed Design Stage of BIM-Model
Relevant requirements	Architect develops execution and detailed models (represented by drawings). Provided specifications intended to enable contractors to build the work.
Information entities	**Equipment and Systems Inventory:** All systems and components are specified **Material and Material Constituents:** Property sets to be included are defined for each component. This is required for proper procurement. **Geometrical representation of building and distribution elements:** Detailed geometrical representations are required.
Information exchange	Source of Information: Architectural Designer User of Information: Tenderer, Contractor, Owner

Use Case	To oversee construction and commissioning processes
Relevant requirements	The project leader oversees construction works, monitors errors and omissions management and manages the hand-over and commissioning process.
Information entities	**Equipment and Systems Inventory:** Documentation of major system components. The "as-built" is compared against the "detailed design". Deviations must be documented and – if required – approved. **Material and Material Constituents:** Property sets to be included must document the "as-built-status". **Geometrical representation of building and distribution elements:** All geometrical specifications captured from the "as-built" facility. This may deviate from the planned drawings, due to manufacturing and assembly tolerances.
Information exchange	Source of Information: Detailed Design Documentation, results from Progress Monitoring & Commissioning User of Information: Owner, Tenant

Use Case	To operate the building to the satisfaction of owner and tenants
Relevant requirements	To "balance" the expectations between user comfort (tenant) and benefits from operation (operator) and profits from ownership of the buildng
Information entities	See section 3.10
Information exchange	Source of information: Building Operator (FM, BAC, and BIM-systems) User of Information: Owner, Tenant

Use Case	Preparation of End-of-Live BIM-Model
Relevant requirements	The owner agrees to share information about the building. Energy suppliers and building operators must share all data of relevance for health and safety. Structural Engineering calculations migt be required, to prove that the building remains in secure and stable condition, even if parts of the building are already demolished.
Information entities	**Equipment and Systems Inventory:** All systems must be modeled in detail. It is essential to have a detailed documentation to make sound and save decisions when to disconnect what supply systems (gas, electricity, water, etc.). Uncontrolled release of any matter must be strictly avoided. The same holds for the documentation of the core & shell system. Even temporary storage of “non-load bearing components” may lead to unexpected load cases. Under no circumstances the design loads are not to be exceeded to avoid collaps of the superstructure. **Material and Material Constituents:** All material must be documented. Propoer disposal must be organised in strict compliance with environmental regulations (e.g. disposal of asbestos). **Geometrical representation of building and distribution elements:** All geometrical specifications are captured from the “as-built” facility before the demolition starts. Temporary stages need to be identified and planned, e.g. in case parts of a superstructure need to be supported by scaffolding during demolition.
Information exchange	Source of Information: Detailed Design Documentation, results from surveying before and during demolition User of Information: Contractors involved in the demolition process, Buildig authorities, Environmental consultants and authorities. If required owners of facilities in the neigborhoud to avoid negative impact on their propety.

Indoor air quality use cases

Use Case	Early identification of Indoor Air Quality model parameters
Relevant requirements	Occupant is interested to set a health environment framework in premises Occupant is interested for the management and control of indoor air quality conditions
Information entities	**Occupants config parameters:** Age, Gender Lifestyle, etc. **Occupancy profiling parameters:** presence, level of occupancy, etc. **Activity profiling parameters:** typical activities (drivers & actions) **Indoor air quality conditions**: VOC, CO, CO2 **Indoor air quality preferences:** IAQ settings
Information exchange	Occupant is defined with a static profile (config parameters) and occupancy/activity profiling parameters Occupant directly set its preferences about indoor environmental conditions (comfort level) Occupant indirectly indicates its preferences in different environmental conditions in building premises.

Use Case	Establishment of a health/IAQ environment during the renovation process - Renovation tasks/processes to address IAQ standards
Relevant requirements	Occupant interested about IAQ conditions during the renovation process Expert setting health/IAQ boundaries during the renovation process
Information entities	**Indoor air quality conditions**: VOC, CO, CO2 **Indoor air quality standards:** IAQ settings
Information exchange	Occupant indirectly indicates its preferences under different IAQ in building premises. Expert specifies the IAQ levels during the renovation process and monitors IAQ values to ensure that these are within boundaries

Use Case	Establishment of a IAQ preserving framework post renovation process
Relevant requirements	Occupant updates information about IAQ conditions post renovation process Continuous monitoring of IAQ conditions to ensure that the IAQ values are within thresholds of the operational environment in building premises
Information entities	**IAQ conditions**: VOC, CO, CO2 **IAQ boundaries**: as identified in the national and international legislation with focus on environmental & comfort conditions in premises **IAQ KPI values:** aligned with the boundaries as specified above
Information exchange	Occupant directly express its preferences and non-preferences about IAQ conditions IAQ analysis based on international and national standards and taking into account IAQ related values

Acoustics use cases

Use Case	Early identification of desired performance of acoustical properties for setting requirements (Concept design)
Relevant requirements	• Applicable standards/regulations currently on a national level. • International acoustic classification scheme for residential buildings is under development (ISO/TC43/SC2/WG29). • Ongoing work in CEN TC126/WG12 on BIM Acoustics (cooperation initiated). • If no relevant requirements exist and requirements corresponding to new buildings are not feasible, best practice is to require that the resulting acoustical performance should equal or exceed existing conditions.
Information entities	System level parameters: • **Sound Insulation *In Situ***: Apparent sound reduction index R' • **Impact Noise and Sound Insulation *In SItu***: Energy-average impact sound pressure level in a room L_i
Information exchange	To client: • Acoustic properties of existing building solutions (for requirement setting): Acoustic consultant • Requirements on National/European level: Corresponding standard organisations / regulatory bodies

Use Case	Early identification of acoustical parameters of structural/fluid model for planning, specifications and design of solutions (Preliminary/Developed/Detailed Design iterative process)
Relevant requirements	• Requirements on building elements from numerical models/simulation based on identified system requirement • Existing framework for acoustic prediction models in EN-12354.
Information entities	Component level parameters: • **Sound Insulation of Building Elements (walls, windows, doors etc):** Sound reduction index, R, Adaption terms etc. • **Flanking Sound Insulation:** Weighted normalized flanking level difference $D_{n,f,w}$ System level parameters: • **Façade Sound Insulation:** Weighted sound reduction index $D_{nT,w}$, Level differences etc. • **Impact Noise and Sound Insulation:** Impact sound pressure level L_i, Weighted reduction of impact sound pressure level ΔL_w by a floor covering, etc.
Information exchange	To 1) Structural engineer / Architect / Designer and (2) Client (iterative process): • Acoustic properties of building elements for use in renovation: Building element manufacturers • Acoustic system properties: Acoustic consultant (iterative process)

Use Case	Implementation of solutions of importance for the acoustic performance during the construction phases
Relevant requirements	• The requirement set from client (acoustic consultant)
Information entities	System level parameters: • **Façade Sound Insulation:** Weighted sound reduction index $D_{nT,w}$, Level differences etc. • **Impact Noise and Sound Insulation:** Impact sound pressure level L_i, Weighted reduction of impact sound pressure level ΔL_w by a floor covering, etc.
Information exchange	Parameters to Contractor/construction workers from Acoustic consultant /structural engineer

Use Case	Maintaining acceptable acoustical indoor conditions.
Relevant requirements	• National/international regulations for residential buildings • Agreed acoustic comfort level (if stated in contract) • General regulations for healthy environments (national health agencies)
Information entities	**Indoor sound level:** Sound level of noise from technical installations, external sources as traffic, neighbours etc. **Estimated/expected sound level:** Estimated values from prediction models
Information exchange	To Resident / Real estate manager: Regulatory requirements: National authorities/ Acoustic consultant

Building performance use cases

Use Case	Documentation of User Comfort
Relevant requirements	• National regulations and international standards for the definition of user comfort • Determine ServiceLevelAgreements and related KPI • Meassured comfort parameters ($\square_{air,int.;}$ RH, IAQ-CO_2, $E_{v, glob}$ /$E_{v, diff}$, /E_v) • Threshold values (e.g max$\square_{air,int.;}$ min$\square_{air,int.;}$) • Dimensional data (list of spaces, list of tenants, list of systems) • Time hierarchy data
Information entities	• Novel models to prepare monitoring data for analysis are required, integrating fact data and dimensional data. • Algorithms are required to specify how KPI can be calculated and frequently updated
Information exchange	To Owner, Operator, Tenant • Comfort parameters: compiled from sensors • Threshold values: compiled from national legislation • Dimensional data: compiled from BIM and ERP-systems • Time hierarchy data: generated by system administrator

Use Case	Documentation of System Usage
Relevant requirements	• National and international regulations for systems' operation and maintenance • Documented status of systems (on/off or used with x% intensity). • Documented relationship of controller and FlowControlDevice (e.g. switch for pump, valve for radiator) • Dimensional data (list of systems, list of operators) • Time hierarchy data
Information entities	• Novel models to prepare monitoring data for analysis are required, integrating fact data and dimensional data. • Algorithms are required to specify how KPI can be calculated and frequently updated
Information exchange	To Owner, Operator, Tenant : • Feedback signals from Controller: compiled from automation system • Dimensional data: compiled from BIM • Time hierarchy data: generated by system administrator

Use Case	Documentation Energy Consumption
Relevant requirements	• National and international standards to define the performance of meters (e.g. reading tolerances) • National and international standards to regulate privacy and ownership of data • Regulations to define the normalisation of metered date: against weather data, consumption/area, etc. • Metered consumption • Time hierarchy data • Weather data • Other data required for normalisation
Information entities	• Linking of recorded time series data to consumers (e.g. either tenant or apartments).
Information exchange	To Owner, Operator, Tenant: • Metered consumption: compiled from automation system • Time hierarchy data: generated by system administrator • Weather data: from local weather station • Dimensional data: compiled from BIM

Use Case	Documentation of occupation density
Relevant requirements	• Regulations and standards defining privacy rights of tenants and occupants • Presence detection data • Time hierarchy data • Relationship between presence detection and spatial model
Information entities	• Linking of recorded time series data to spaces or even zones in spaces
Information exchange	To Owner, Operator, Tenant: • Presence data: compiled from automation system • Time hierarchy data: generated by system administrator • Location hierarchy: compiled from BIM

Cost modelling use cases

Use Case	To estimate renovation cost
Relevant requirements	The landlord requests an initial cost estimation.
Information entities	**Relocation cost:** The landlord needs to estimate what construction management scenario is cost-efficient (and possible from a legal, health and safety perspective); (a) renovation under occupation, or (b) renovation when tenants stay away. **Demolition costs:** Demolition cost needs to be estimated considering the following cost components (i) labour cost, (ii) material disposal (incl. recycling), (iii) cost of support resources (e.g., scaffolding). **Renovation cost:** For the estimation of renovation cost the following components need to be considered (i) material cost, (ii) labour cost, (iii) cost of support material (scaffolding etc).
Information exchange	**Source of Information**: Approximate, estimated quantities and average market prices for unit cost (e.g. Euro per sqm or Euro per cubic metre), **User of Information**: Landlord, Investor

Use Case	To specify approximate renovation cost
Relevant requirements	The landlord requests an initial cost specification. Whenever possible, approximate quantities shall be calculated from the initial BIM-model.
Information entities	**Relocation cost:** Based on the estimated relocation cost the landlord (or investor) requests a specification of relocation cost based on market prices. **Demolition costs:** An initial BIM-model can be used for the specification of the masses of building elements to be demolished. However, with current BIM-technology it is not realistic to assume that BuildingDistributionElements (i.e. Building Services Elements) are included in the BIM-model. It should be noticed, that in certain renovation scenarios building services sub-systems (e.g. electrical, phone, etc.) are only de-activated but not demolished. **Renovation cost:**The quanities for new building elements can be now determined on the basis of BIM. This may also be applicable for certain parts of building services components, e.g. sockets per room-type.
Information exchange	**Source of Information**: Approximate quantities and average unit prices (e.g., €/m^2 or €/m^3), **User of Information**: Landlord, Investor

Use Case	To specify exact quantities and cost
Relevant requirements	The landlord requests precise quantity specifications. Before the completion of the procurement process prices are still average prices. After completion of the procurement process, quantities and prices can be exactly determined.
Information entities	**Relocation cost:** After the completion of the procurement process the construction company will present a precise schedule for the execution of construction works; i.e. the duration for the relocation can be more precisely determined. **Demolition costs:** Based on the precise BIM model and the agreed demolition technology (after procurement) a precise quantity take-off can be extracted from the BIM-model. The quantity take-off will also include a specification of construction resources. **Renovation cost:**A precise BIM-model with (i) exact geometrical specifications, (ii) precise material descriptions, and (iii) manufacturer specification (after procurement) can be used for the generation of the quantity take-off.
Information exchange	**Source of Information**: Approximate quantities and average market prices for unit cost (e.g. Euro per sqm or Euro per cubic metre), **User of Information**: Landlord, Investor

Use Case	To specify environmental cost and full life-cycle cost
Relevant requirements	In order to achieve an environmental certification (e.g. BREEM, LEED, DGNB, Energy Carbone) material specifications must be compatible to certified LCA-databases, such as EPD: Italy, Ireland, ITB: Poland, RTS-EPD: Finland or Ökobaudat: Germany
Information entities	Embodied Energy: This can be determined “from gradle to grave” for each certified building material. The energy required to manufacture, transport, and install a building component must be included in the “energy required” to operate a building. Environmental Impact: This must be determined in advance, to avoid the usage of unnecessary environmental impacts from manufacturing, transportation or operation of building components.
Information exchange	**Source of Information**: Quantities and material specification from BIM Environmental data from national catalogues, such as: ÖkoBauDat, EPD, etc. **User of Information**: Architect or environmental consultant on behalf of architect