SPHERE

BIM DIGITAL TWIN PLATFORM

White Paper

October 2022

ONTOLOGIES AND BUILDING DIGITAL TWINS

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement Nº 820805.

With the collaboration of:

BIMTWIN BIMERR BIM4EEB BIMprove COGITO

October 2022

DISCLAIMER

ACKNOWLEDGEMENT

SPHERE project has received funding from the European Union's Horizon 2020 research and innovation program under grant agreement Nº 820805.

These projects have received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement:

BIM2TWIN
https://bim2twin.eu/
No 958398

BIMERR
https://bimerr.eu/
No 820621

BIM4EEB
https://www.bim4eeb-project.eu/
No 820660

BIMprove
https://www.bimprove-h2020.eu/
No 958450

COGITO
https://cogito-project.eu/
No 958310

AUTHORS

- Mads H. Rasmussen (PhD), NIRAS
- Pieter Pauwels (PhD), Associate Professor at Eindhoven University of Technology
- María Poveda-Villalón (PhD), Universidad Politécnica de Madrid, BIMERR project
- Sanju Tiwari (PhD, Post-Doc), Universidad Autonoma de Tamaulipas, Mexico
- Raúl García-Castro (PhD), Universidad Politécnica de Madrid, COGITO Project
- Pierre Bourreau (PhD), IT PM, NOBATEK/INEF4
- Bruno Fies, CSTB Research Engineer, Bim2Twin project
- Jonas Sschlenger (PhD), Technische Universität München, Bim2Twin project
- Bruno Daniotti (PhD), Politecnico di Milano, BIMEEB project
- Seppo Törma, Metropolia Univ. | CEO of VisuaLynk, BIMEEB project
- Mark Thomas, CEO Nextspace
- Michel Böhms, Senior Linked Data Scientist, TNO
- Gabor Sziebig, Research Manager, SINTEF Manufacturing, BIMprove project
- Angel Font, BDTA senior researcher
- Eduard Loscos, BDTA President
- Pablo Vicente-Legazpi, BDTA PM
- Peter Imbrechts, Neanex

ABOUT SPHERE

SPHERE is a 4-yearHorizon 2020 project with 20 partners targeting the improvement and optimization of building energy design, construction, performance, and management, reducing construction costs and their environmental impacts.

SPHERE seeks to develop a building centered Digital Twin Environment, involving not only the design and construction of the building but also including the manufacturing and the operational phases.

https://sphere-project.eu/

ABSTRACT

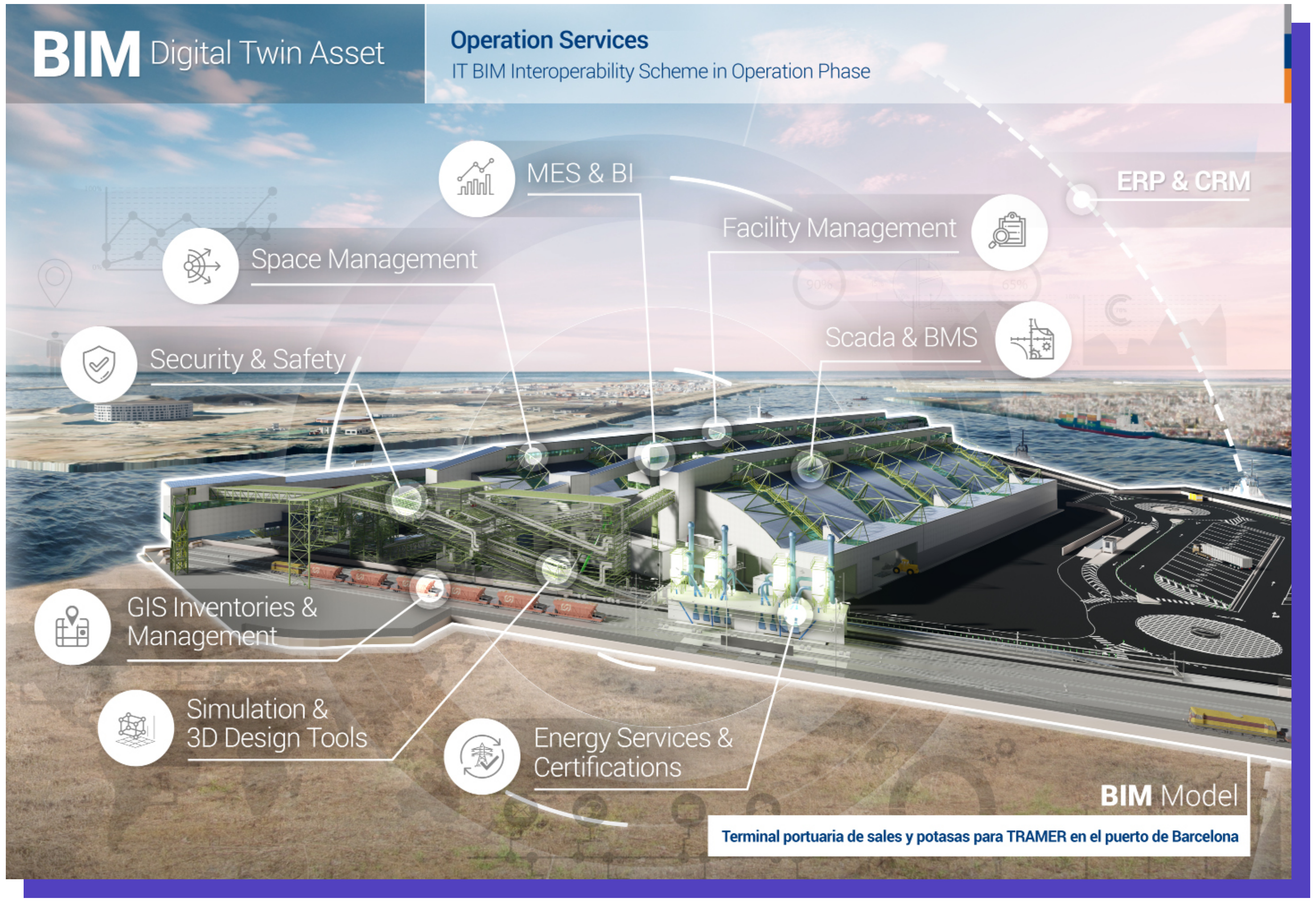
A review of existing ontologies in construction is presented (updated September 2022). Motivation, alignments and final applications with special interest on methodologies. Software Tools are presented in an Appendix. Some use cases are presented with different orientations.

Final conclusions are oriented to new paths to follow in future developments, considering other technologies and commercial solutions nowadays. More than new ontologies the use of existing ontologies as DICon[1] could help to standardize and make an ontologies map for construction possible. Progress with standards, as mentioned in a specific chapter about status of group CEN442 WG4, is essential to get a good level of interoperability, but it faces complex technical challenges. And the most sophisticated and perfect implementation sometimes means a non-practical approach, and too rigid to be used by the industry.

1 https://digitalconstruction.github.io/v/0.5/

KEYWORDS

- Building Digital Twin
- BDT
- Ontologies
- Building information modeling (BIM)
- IFC Semantic Web
- Linked data
- Semantic interoperability



Conceptualisation of a Building Digital Twin Instance Interoperability Services (sketch by: M.Borràs)

TABLE OF CONTENTS

1.-INTRODUCTION

Towards a unified reference architecture in AECOO industry

Nowadays, discussion across the construction industry about moving away from silos in different phases of the design and construction process have become a hot topic. Utilizing key project data through integrated systems and technologies in an effort to eliminate common problems such as re-entry of data or data redundancy is a matter of traceability and trustworthiness, hence controlling the risks. In order to achieve such an integrated information flow across the life cycle of an asset, data interoperability is extremely important in the AECOO (Architecture, Engineering and Construction, Owner and Operator) industry as the industry encounters a lot of complexity like many stakeholders, one-off projects and an ever-changing environment.

The (semantic) interoperability issue

The Issue of interoperability is not new within the AECO industry. Already in the nineties, efforts have been made tackling this issue, introducing exchange formats like STEP and IFC. Over the last two decades, the adoption and importance of open semantic standards from W3C[1] has increased significantly, resulting in a well-known set of ontologies/taxonomies developed at different TRL's[2] for different purposes or services across the building lifecycle. However, when introducing a more holistic approach like a Building Digital Twin, and thus taking in account the entire lifecycle and different domains of an asset, the connection between all the different existing ontologies is not properly unfolded.
Lacking a reference architecture for Building Digital Twins, ad hoc sets of relationships among ontologies are established for specific projects by mixing reference (open) with proprietary ones. This approach is time consuming and not practical since several non mature ontologies (Mid TRL[3]) are still evolving rapidly. A classic example is shown in figure 1[4].

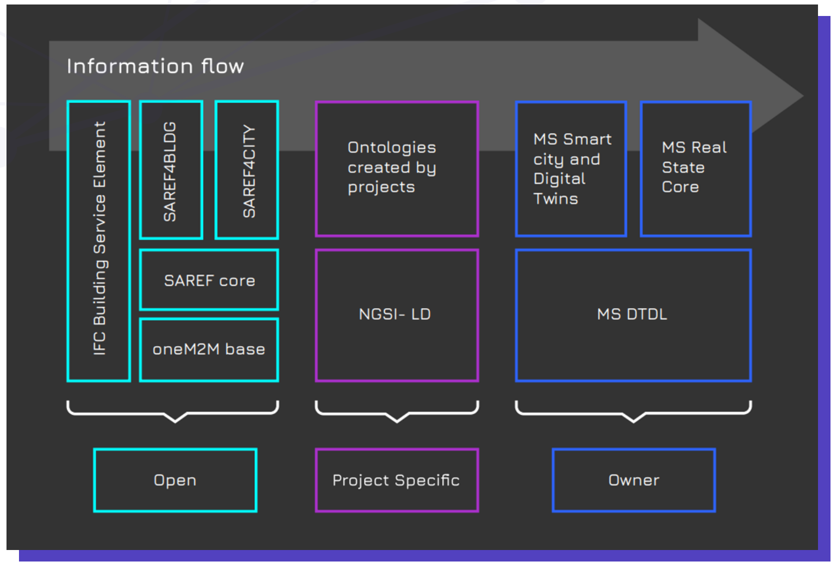


Figure 1: Example of Complexity map for project specific application

1 The World Wide Web Consortium is the main international standards organization for the World Wide Web. Founded in 1994 and currently led by Tim Berners-Lee, the consortium is made up of member organizations that maintain full-time staff working together in the development of standards for the World Wide Web.

2 TRL (Technology Readiness Level)

3 It's considered that SAREF4BLDG is TRL6

4 Adapted from The Digital Twin Hub

Building digital twin as a key cornerstone to enable interoperability across the building life cycle

The SPHERE project[5] consists of a robust framework for testing and implementing transcending data within a semantic interoperable Building Digital Twin ecosystem from both a technical and a business perspective. The cornerstone of the SPHERE BDT ecosystem is the definition of a Network of Ontologies where only OPEN reference ontologies are applied, extended, and empowered in a univocal architecture along the building life cycle.

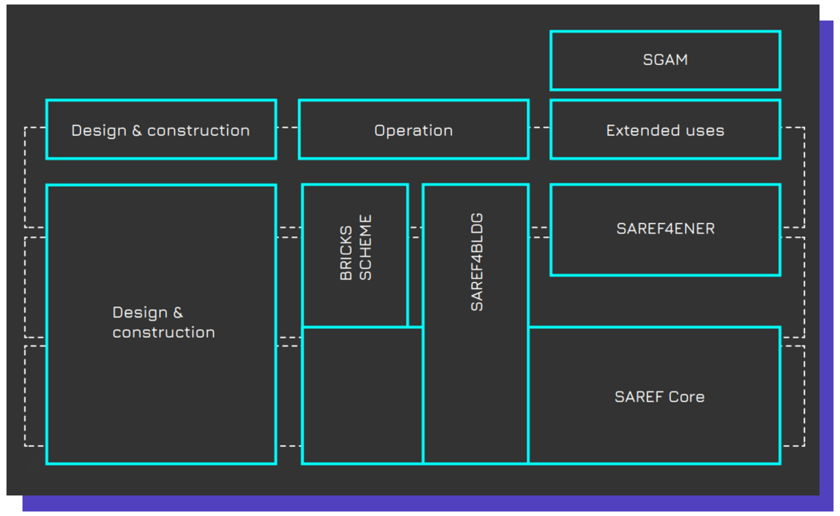


Figure 2: Digital Twin and map of main reference ontologies in AECO and extended sectors

Such a unified OPEN "Network of ontologies" approach enables use case-based extensions like a Distributed Energy Resource (DER) in the context of the smart grid development, or future smart energy networks and smart positive districts, and specifically connecting these emerging technologies to the building life cycle. And on the same hand, this approach allows the introduction of Use Cases within the circular economy like LCA and Material Passports, by extending the applied Network for the purpose of the selected Use Cases and without interfering with each other or having a need of one-size-fits-all ontology that reinvents what is already existing.

Future direction: towards a unified reference architecture for building life cycle management

The definition of a holistic Ontology Network for a Building Digital Twin is a necessity in the AECO industry. A next step is the reflection of a "Network of Ontologies" approach when applying a Building Digital Twin in a project context: A reference set of tools aimed to manage the entire building lifecycle under a PaaS or ecosystem business model. Within the SPHERE project the standardisation and definition of a Building Digital Twin Ontology Network will evolve in the development of an Open Digital Twin API. That API is a core part of the SPHERE Digital Twin Architecture enabling the ecosystem approach.

5 SPHERE is a 4-year, Horizon 2020 project (GA No. 820805) that aims to provide a BIM-based Digital Twin Platform to optimise the building lifecycle, reduce costs, and improve energy efficiency in residential buildings.

Alignment and interoperability

The use of formal ontologies in the AECOO industry has proved to have significant benefits to reduce the well-known issues of using data models in a fragmented industry: interoperability, 'siloization' of knowledge, discrepancies in the vocabulary… The use of knowledge graphs helps to ensure a high level of modularity and extensibility in the model through simple mechanisms, in particular **data alignment**, also mentioned as **data matching**, or **data mapping** in the literature. Alignment can be simply defined as the **task of defining a semantic relation between concepts**; this is particularly useful when the same concept is expressed in two different data models: structural engineers may refer to a wall in some way, and thermal engineers in a different one; when combining data models for structural engineering and for thermal engineering, one may mention that the two concepts are identical, or at least semantically related. Such a simple mechanism is powerful in helping data modelers reuse data models that were built up and consolidated by domain experts… and the diversity of domains in the AEC industry is wide[6].

But alignment can also be helpful to tackle **interoperability issues**; indeed, different software may work with different BIM models; for instance, a Revit native file used by architects for the design phase can be converted into IFC and then converted into a gbXML file format, that will be used by thermal engineers to compute energy consumption estimation. While IFC is a great step towards digital interoperability between software in the AEC industry, issues are still known. A way to work on interoperability between file formats is **to use alignment between models** (in the previous example, aligning concepts between the Revit model and the IFC model; and between the IFC model and the gbXML model); through alignment the conversion of one model to another model becomes straightforward[7], and this could be a major step towards BIM level 3: a central BIM model, shared on the web by all project stakeholders, where changes made by a partner are reflected to the rest of the project team, while each of them can still use their own domain language to query or change the BIM model.

6 As an example, see BOT ontology alignments modules at https://w3c-lbd-cg.github.io/bot/#AlignmentModules
7 as shown in https://research.aalto.fi/en/publications/dynamic-bim-format-conversion-as-inference-based-ontology-alignme

2.-SEMANTIC WEB TECHNOLOGIES

Sir Tim Berners-Lee, someone who changed the World

In 1989, Tim Berners-Lee invented the World Wide Web, an Internet-based hypermedia initiative for global information sharing while at CERN, the European Particle Physics Laboratory[8]. He wrote the first web client and server in 1990. His specifications of URIs, HTTP and HTML were refined as web technology spread.

He is Director of the World Wide Web Consortium (W3C), a Web standards organization founded in 1994 that develops interoperable technologies (specifications, guidelines, software, and tools) to lead the Web to its full potential. He is a founding Director of the Web Science Trust (WST) launched in 2009 to promote research and education in Web Science, the multidisciplinary study of humanity connected by technology. Berners-Lee is also a Director of the World Wide Web Foundation, launched in 2009 to coordinate efforts to further the potential of the Web to benefit humanity.

By Tim Berners-Lee[9]. In response to a request, a one page looking back on the development of the Web from my point of view. Written 1998/05/07

The World Wide Web: A very short personal history

There have always been things which people are good at, and things computers have been good at, and little overlap between the two. We were brought up to understand this distinction in the 50s and 60s and intuition and understanding were human characteristics, and that computers worked mechanically in tables and hierarchies.

One of the things computers have not done for an organization is to be able to store random associations between disparate things, although this is something the brain has always done relatively well. In 1980 I played with programs to store information with random links, and in 1989, while working at the European Particle Physics Laboratory, I proposed that a global hypertext space be created in which any network-accessible information could be refered to by a single "Universal Document Identifier". Given the go-ahead to experiment by my boss, Mike Sendall, I wrote in 1990 a program called "WorldWideWeb", a point and click hypertext editor which ran on the "NeXT" machine. This, together with the first Web server, I released to the High Energy Physics community at first, and to the hypertext and NeXT communities in the summer of 1991. Also available was a "line mode" browser by student Nicola Pellow, which could be run on almost any computer. The specifications of UDIs (now URIs), HyperText Markup Language (HTML) and HyperText Transfer Protocol (HTTP) published on the first server in order to promote wide adoption and discussion.

The dream behind the Web is of a common information space in which we communicate by sharing information. Its universality is essential: the fact that a hypertext link can point to anything, be it personal, local or global, be it draft or highly polished. There was a second part

8 https://www.internethalloffame.org/inductees/tim-berners-lee
9 https://www.w3.org/People/Berners-Lee/ShortHistory.html

of the dream, too, dependent on the Web being so generally used that it became a realistic mirror (or in fact the primary embodiment) of the ways in which we work and play and socialize. That was that once the state of our interactions was on line, we could then use computers to help us analyse it, make sense of what we are doing, where we individually fit in, and how we can better work together.

The first three years were a phase of persuasion, aided by my colleague and first convert Robert Cailliau, to get the Web adopted. We needed Web clients for other platforms (as the NeXT was not ubiquitous) and browsers Erwise, Viola, Cello and Mosaic eventually came on the scene. We needed seed servers to provide incentive and examples, and all over the world inspired people put up all kinds of things.

Between the summers of 1991 and 1994, the load on the first Web server ("info.cern.ch") rose steadily by a factor of 10 every year. In 1992 academia, and in 1993 industry, was taking notice. I was under pressure to define the future evolution. After much discussion I decided to form the World Wide Web Consortium in September 1994, with a base at MIT in the USA, INRIA in France, and now also at Keio University in Japan. The Consortium is a neutral open forum where companies and organizations to whom the future of the Web is important come to discuss and to agree on new common computer protocols. It has been a center for issue raising, design, and decision by consensus, and also a fascinating vantage point from which to view that evolution.

With the dramatic flood of rich material of all kinds onto the Web in the 1990s, the first part of the dream is largely realized, although still very few people in practice have access to intuitive hypertext creation tools. The second part has yet to happen, but there are signs and plans which make us confident. The great need for information about information, to help us categorize, sort, pay for, own information is driving the design of languages for the web designed for processing by machines, rather than people. The web of human-readable document is being merged with a web of machine-understandable data. **The potential of the mixture of humans and machines working together and communicating through the web could be immense**.

The Semantic Web

Although the initial idea of a machine-understandable Web appeared in 1989, it was first called the Semantic Web in 2001. Tim Berners-Lee defined the Semantic Web as: "extension of the World Wide Web (WWW) in which data are given meaning (semantics) to enable computers to" (…); "specifically a web of machine-readable information, whose meaning is well-defined by standards: it absolutely needs the interoperable infrastructure that only global standard protocols can provide". To achieve the Semantic Web, Web resources should be described in the way that makes their meaning explicit.

The vision of the Semantic Web is to present the web of data as a "thing" rather than documents on HTML pages. Semantic Web provides advancement of the data on the internet by allowing the queries along with retrieving and browsing the pages and deriving new knowledge from existing information to explore the inconsistencies. Like WWW uses HTTP and HTML for presenting documents, Semantic Web uses RDF and RDFs for presenting data. It has a layered architecture which has several components such as Unicode and URI, XML, Resource Description Framework (RDF), RDF Schema, Ontologies, Logic and Proof, Trust.

Berners-Lee outlined his vision for the Semantic Web as a layered architecture. In this architecture, the semantic languages are built upon URI (Uniform Resource Identifier) and Unicode, which are already present in the Web. URI became a W3C Recommendation in 1989, providing a means of identifying resources with NS (Namespaces). XML (Extensible Markup Language) and RDF (Resource Description Framework) were considered as two major technologies of the Semantic Web (Berners-Lee et al. 2001). XML became a W3C Recommendation in 1998, and RDF became a W3C Recommendation in 1999.

On the top of RDF there is **the Ontology layer**. Ontology means the **specification of a conceptualization**; which defines **terms and relationships** between terms, preferably in some machine-readable manner. The Ontology layer is in the form of the OWL (Web Ontology Language), which became a W3C Recommendation in 2004.

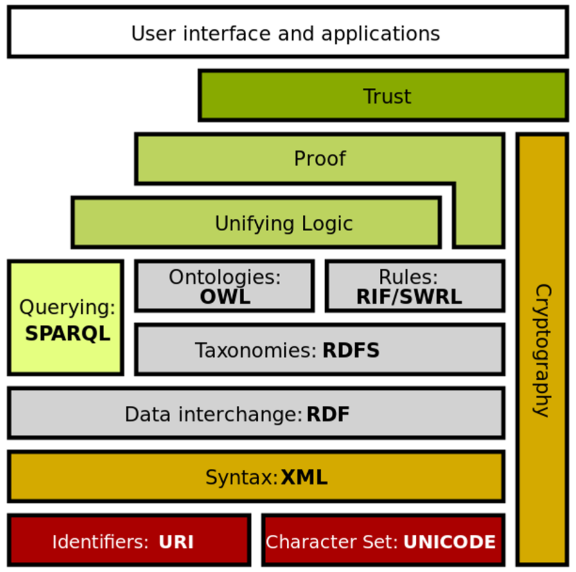


Figure 3: A Semantic Web Layer Cake, explains the complete architecture of the Semantic Web[10]

Brief overview of Internet and Semantic Web ontologies

Many people understand the Internet as documents which link to other documents[11]. But we can face a big problem of updating the links if information changes. What about if we could use the 'Web of Data' as a 'Content Management System'?. In that way we could use the community itself as a large 'content editor'.

Most of the time date are isolated, they are silos of information. Different databases without integration. Like documents available to download but no links between them.

We need an infrastructure for a real 'web of data': data is available on the Web and they are accessible via standard Web technologies, and data can be integrated over the Web, interlinked over the Web. This is where **Semantic Web technologies** come in.

In short: Map the various data onto an abstract data representation, making data independent of its internal representation. Merge the resulting representations and start **making queries on the whole directly on internet**.

10 Semantic Web - XML2000, slide 10. W3C. Retrieved 2008-05-13. https://en.wikipedia.org/wiki/Semantic_Web_Stack
11 Ivan Herman, W3C June 22nd, 2010 – Semantic Technology Conference, San Francisco

Semantic Web and construction

The construction industry is facing an increasing demand for information and communication between its global distributed partners. Current Web-based information management systems do not reach their full potential because the Internet is a "**web of links**", which is a place where **data can only be shared and processed by humans**. Therefore, it is inconvenient to resolve problems when project partners cannot meet together at the same time, or use different languages, or have different understandings of an issue. This leads to **low efficiency in communicating project information** and facilitating collaboration among partners, which are major hurdles for the success of projects.

One of the approaches that attempts to solve the above problems focuses on making the Web understandable by both machines and humans. The term, "Semantic Web", or "**Web of meaning**", is used to describe such a Web, in which information is given well-defined meaning. **Both computers and people can work in cooperation**. Since the information on the Semantic Web has a clearly defined meaning, it can be analysed and traced by computer programs. Although programs on the Semantic Web may be designed independently, they will be able to share and process data automatically.

It has been interesting to see the evolution of this technology over the last 10 years. An article in September 2004 about Semantic Web applications in construction[12] did not mean the IFC standard at all. Instead, it was declaring that "current information management systems cannot describe them" (speaking about construction items). This is a reality today and models can be represented and transferred as data. Procurement was described as an ideal situation where all stakeholders have their information on the web. And communications and change management, not considering some human factors, was seen as a promising technical achievement. We can see that there has been huge progress since 2004, but today we realize that the reality of product catalogues nowadays is far from ideal, and people management just going in and out a project may be destroying any good initial plan. In summary, only Semantic Web progress forgetting some human drawbacks leaves some room for improvement over the next years.

Even in the case that advanced knowledge systems would be retrieving the right documents thanks to the Semantic Web, or getting a high degree of automation, integration and reuse of data across various applications (URL 4), the human factor is still there, as an **intensive creation process** where many human decisions must be taken. Forgetting this fact results in useless technology as it was described and smartly pointed out by Mads H. Rasmussen as the SSoldac event in Cercedilla (Madrid, Spain) in June 2022.

Hence, we see today an imperative use of IFC, which is key in the design process, but with a huge **lack of design methodology and good practices**. There are new common data environments but many of them are reproducing the Ifc structure in the cloud, with many problems to integrate other data structures. This would be a divergent direction towards digital twins (supported with ontologies behind) and a real improvement in data management.

New ontologies are seen as an evolution of IFC to Semantic Web (IfcOwl, BOT). But again, lack of initial and well-structured information at IFC stage is frustrating for downstream operations.

12 Pan, J, Anumba, C J and Ren, Z (2004) Potential application of the semantic web in construction. In: Khosrowshahi, F (Ed.), 20th Annual ARCOM Conference, 1-3 September 2004, Heriot Watt University. Association of Researchers in Construction Management, Vol. 2, 923-9

We should **look backwards and help to improve the quality of IFC** data, as a methodology to get better data ontologies in the end. If we feed digital twin ontologies with IFC garbage we cannot expect good results in any way.

Semantic Web as checking platform

As ontologies can be collecting the information of a whole project in a web platform, it could be seen as a perfect platform for the implementation of checking rules. These rules could be complex and affecting many aspects of the project, not just limited to the IFC representation. They could be available at both sides of one permit or licensing interface and the reasoning would be open and available for anyone. And again, this is true in the paradise of perfect project information, perhaps not in today's reality. But a promising field of innovation and improvement over the next years.

Partial information and forgetting data

As important as collecting data may be how to detect that some information is outdated, changed or just that it must be completely deleted. Many small details in projects are totally irrelevant and they are hiding the real critical data needed to take decisions. Once again Semantic Web technologies can help to limit and encapsulate contents that will be used only if needed. And of course, 'bad information' can be detected as well. This inverse process of deconstructing information will be critical in big digital twins, and not only the graphical speed of the 3D viewers. Big digital twins will need to 'forget data', to clean data giving relevance or weight to important information when exploiting the digital twin.

3.-ONTOLOGIES REVIEW

An excellent source of information to know existing onlogies (and an independent evaluation of reliability) is found at DBpedia[13]. The DBpedia Association was founded in 2014 to support DBpedia and the DBpedia Community. DBpedia features around 20 language chapters. Members of the DBpedia community have been contributing to the expansion and stabilization of DBpedia since 2007. It started as a small project in Leipzig which grew into a large crowd-sourced project with scientists and researchers using the DBpedia Open Knowledge Graph worldwide.

At DBpedia Archivo we can find easily a list of available ontologies and their rating.

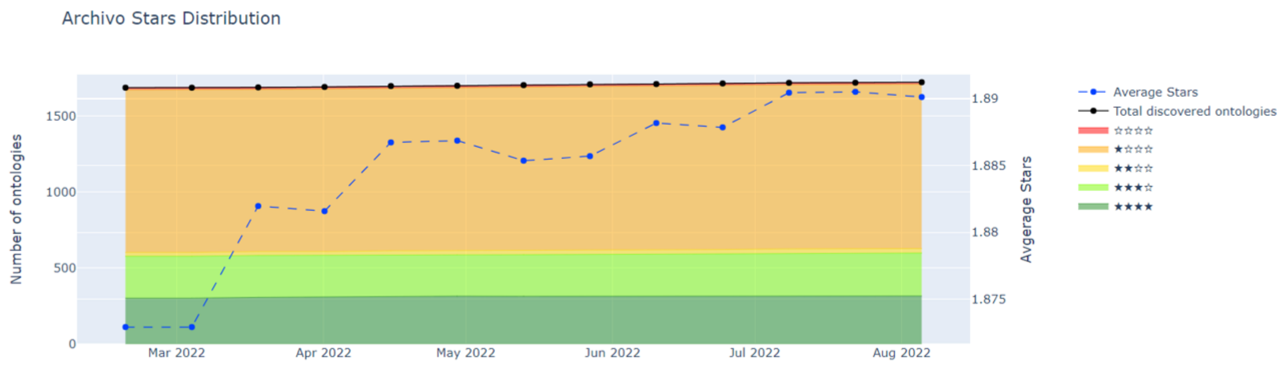


Figure 4: Total number on ontologies detected at Archivo DBpedia (extract from DBpedia webpage)

building

View Archived Ontology	Download Latest	Triples	Stars	Crawling Status
Type of building	owl, ttl, nt	27	★☆☆☆	✓
Buildings and Rooms Vocabulary	owl, ttl, nt	73	★☆☆☆	✗
Ontology of Building Accessibility	owl, ttl, nt	3187	★★★☆	✓
Building Element Ontology	owl, ttl, nt	1530	★★★★	✓
SAREF extension for building	owl, ttl, nt	3076	★★★★	✓
Building Concrete Monitoring Ontology (BCOM)	owl, ttl, nt	524	★★★☆	✓
Building Performance Ontology	owl, ttl, nt	587	★☆☆☆	✓
The Building Topology Ontology (BOT)	owl, ttl, nt	959	★★★★	✓
Building Product Ontology	owl, ttl, nt	427	★☆☆☆	✓
SAREF extension for building	owl, ttl, nt	3076	★★★★	✓
SBEO: Smart Building Evacuation Ontology	owl, ttl, nt	2253	★☆☆☆	✓
The SEAS Building Ontology	owl, ttl, nt	484	★★★★	✓
Building Performance Ontology	owl, ttl, nt	587	★☆☆☆	✓

Figure 5: Rating of ontologies related to “building”

13 https://www.dbpedia.org/

In the following chapters we will try to describe some of the main ontologies taking into account the work reported by Annex 81, BIMERR, BIM4Ren, SPHERE and COGITO EU projects. The final interest is to focus the problem of digital twin’s Semantic Web representation and to define future lines of development.

The Basic Formal Ontology, BFO[14]

The Basic Formal Ontology (BFO) is a small, upper-level ontology that is designed for use in supporting information retrieval, analysis and integration in scientific and other domains. BFO is a **genuine upper ontology**. Thus, it does not contain physical, chemical, biological or other terms which would properly fall within the coverage domains of the special sciences. BFO is used by more than 300 ontology-driven endeavours throughout the world.

The BFO project was initiated in 2002 under the auspices of the project Forms of Life sponsored by the Volkswagen Foundation. The theory behind BFO was developed first by Barry Smith and Pierre Grenon and presented in a series of publications[15].

Ontology in the life science is an example of success:

- 1990: Human Genome Project
- 1999: The Gene Ontology (GO)
- 2002: Open Biomedical Ontologies (OBO)
- 2004: Basic Formal Ontology
- 2004: OBO Foundry

First key for ontology success[16]: hub and spokes approach. That is why BFO is a good choice:

- It is very small
- Evolves very slowly
- Domain-neutral top-level ontology
- Active user forum
- Large user base
- Trained personnel with portable expertise

More than 300 ontologies today are re-using BFO. And it produces a virtuous snowball effect (more users, more mistakes detected, more things fixed, more other ontologies attracted).

Another key of ontology success is modularity. It makes possible ownership by experts (domain experts, not computer scientist). BFO is an ontology that enables other ontologies to plug into each other.

14 https://basic-formal-ontology.org/
15 https://basic-formal-ontology.org/publications.html
16 Barry Smith´s, Introduction to BFO, 2019. Youtube video:

3.1 BUILDING

The Annex 81

The International Energy Authority (IEA) coordinates a series of important initiatives and one of them has recently published the Annex81 ('Survey of metadata schemas for data-driven smart buildings', June 2022[17]). Main authors are Gabe Fierro[18] and Pieter Pauwels[19], and other contributors are listed in a footnote[20]. The mission of the IEA Energy in Buildings and Communities (IEA EBC) TCP[21] is 'to support the acceleration of the transformation of the built environment towards more energy efficient and sustainable buildings and communities, **by the development and dissemination of knowledge, technologies and processes and other solutions** through international collaborative research and open innovation'. Annex 81 is a survey of existing metadata schemas for data-driven buildings. It can be considered the most recent and updated evaluation at the time of publication of this white paper and gives interesting and practical information that point out future trends.

Figure 6: Buildings produce endless streams of sensor, meter, and IoT data

17 https://annex81.iea-ebc.org/

18 Gabe Fierro, Colorado School of Mines, Colorado, USA (gtfierro@mines.edu)

19 Pieter Pauwels, Eindhoven University of Technology, Netherlands (p.pauwels@tue.nl)

20 Aslak Johansen, University of Southern Denmark, Denmark (asjo@mmmi.sdu.dk) - Tianzhen Hong, Berkeley Lab - LBNL, California, USA (thong@lbl.gov) - Arne Hansen, Buildings Evolved, Australia (arne@buildingsevolved.com) - Dimitrios Rovas, University College London, UK (d.rovas@ucl.ac.uk) - Stephen White, CSIRO, Australia (stephen.d.white@csiro.au) - Chun Ping Gao, Building and Construction Authority, Singapore (gao_chun_ping@bca.gov.sg) - TK Wang, VBIS, Australia (tkwang@vbis.com.au) - Maggie Sullivan, Switch Automation, Colorado, USA (esullivan@switchautomation.com)

21 TCP, Technology Collaboration Program

Annex 81 lists the main ontologies available[22]. It makes a qualitative comparison between them taking into account key features, purpose of use and how they fit into a data driven smart building. The majority of the mentioned metadata schemas (6 of 7 in total) orient towards the use of RDF[23] and OWL[24]. Therefore, most instance data can simply be created and maintained using standard and generic RDF tooling[25].

A problem highlighted by the Annex 81 authors is the academic (non-commercial) point of view of some advanced metadata schemas. This opens two lines or resources which go in parallel. Proprietary data models of commercial software respond to other needs and times, but in any case, legacy software gives the possibility of integration and support with metadata schemas. And IFC, Bricks or LBD ontologies have some vendors using them. It is a diffuse line and the key value is the 'creation of an open and neutral format that can be used for external data exchange'[26].

Special reference is given to IFC (which is a metadata schema itself). IFC is an ISO standard and it is used extensively in design and construction phases of a project. It is the massive data entry of initial information and drawbacks can be complemented by the LBD alternative, expanded with Brick and Project Haystack graphs and data.

The final paragraph of Annex 81 finishes with this promising sentence:

'Despite the diversity of approaches and stakeholders for each metadata schema, there is a ***growing theme of unity and alignment*** *emerging from the various groups. We predict, hope, and recommend that future editions of most metadata schemas will focus more on complementing each other* ***through reductions in scope****, rather than expanding the modelling scope to compete on other perspectives of data-driven buildings. We also see* ***RDF-based metadata schemas emerging as the dominant modelling approach****. These demonstrate the highest degrees of interoperability and reusability compared to other proprietary models. New tools will emerge that raise the level of abstraction for interacting with RDF-based metadata schemas, ultimately democratising the use of rich metadata in data-driven smart buildings*[27].

The BIMERR ontology

The EU H2020 project BIMERR[28] -between some other objectives- has to develop an ontology for energy-driven renovation of existing residences. At one deliverable the team carry out a 'Survey of data models, ontologies and standards in the wider Energy Efficient Buildings domain'[29]. It is an interesting description and discussion of all the ontologies and data models available in 2019 which may have any relation with building renovation and energy analysis. In a later deliverable (BIMERR Ontology & Data Model) the final group of ontologies selected is described.

22 1. Project Haystack - 2. Brick Schema - 3. Real Estate Core (REC) - 4. BOT ontology and Linked Building Data (LBD) - 5. SAREF (SAREF4BLDG) - 6. SOSA / SSN - 7. Google Digital Buildings

23 RDF: Resource Description Framework

24 OWL: Web Ontology Language

25 https://www.w3.org/wiki/SemanticWebTools

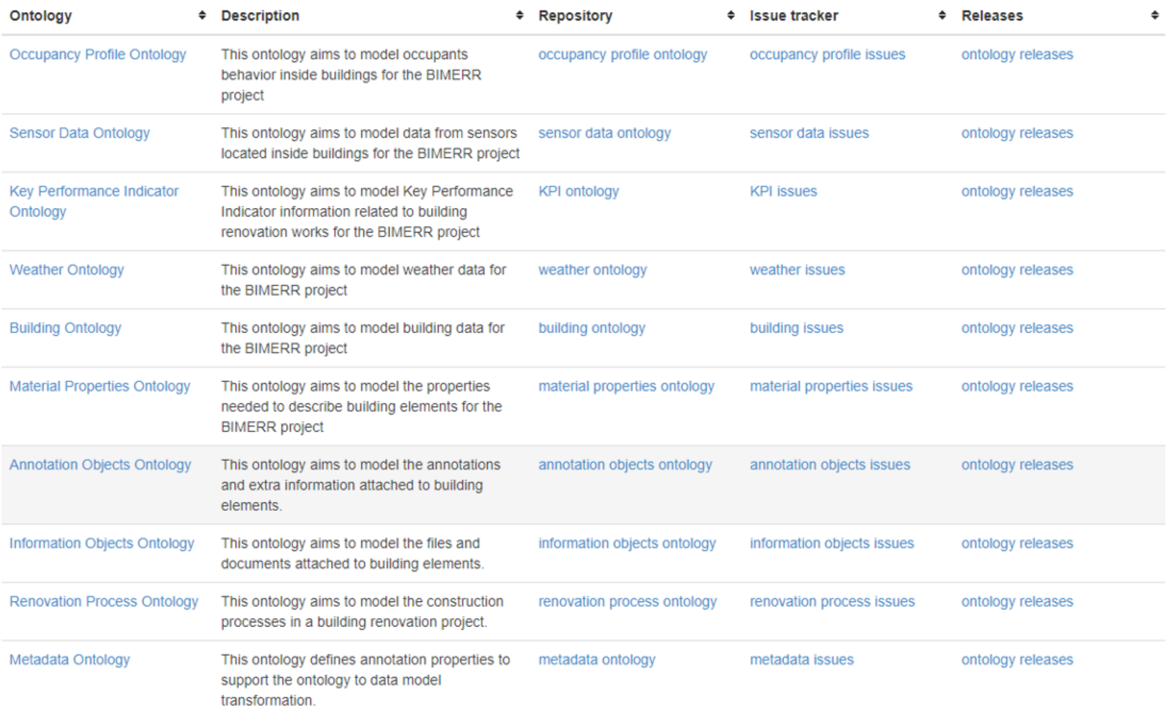
26 Annex 81, page 45

27 Annex 81, page 52, Summary

28 BIMERR: BIMERR BIM-based holistic tools for Energy-driven Renovation of existing Residences ID: 820621 - From: 1 January 2019 to: 30 September 2022

29 https://bimerr.eu/deliverables/

They were considering obvious domains such as building, materials, energy consumption of GIS data together with not so evident occupancy patterns, weather or 'reality capture'. The complete set of ontologies is published at https://bimerr.iot.linkeddata.es/ .



Ontology	Description	Repository	Issue tracker	Releases
Occupancy Profile Ontology	This ontology aims to model occupants behavior inside buildings for the BIMERR project	occupancy profile ontology	occupancy profile issues	ontology releases
Sensor Data Ontology	This ontology aims to model data from sensors located inside buildings for the BIMERR project	sensor data ontology	sensor data issues	ontology releases
Key Performance Indicator Ontology	This ontology aims to model Key Performance Indicator information related to building renovation works for the BIMERR project	KPI ontology	KPI issues	ontology releases
Weather Ontology	This ontology aims to model weather data for the BIMERR project	weather ontology	weather issues	ontology releases
Building Ontology	This ontology aims to model building data for the BIMERR project	building ontology	building issues	ontology releases
Material Properties Ontology	This ontology aims to model the properties needed to describe building elements for the BIMERR project	material properties ontology	material properties issues	ontology releases
Annotation Objects Ontology	This ontology aims to model the annotations and extra information attached to building elements.	annotation objects ontology	annotation objects issues	ontology releases
Information Objects Ontology	This ontology aims to model the files and documents attached to building elements.	information objects ontology	information objects issues	ontology releases
Renovation Process Ontology	This ontology aims to model the construction processes in a building renovation project.	renovation process ontology	renovation process issues	ontology releases
Metadata Ontology	This ontology defines annotation properties to support the ontology to data model transformation.	metadata ontology	metadata issues	ontology releases

Figure 7: BIMERR ontologies

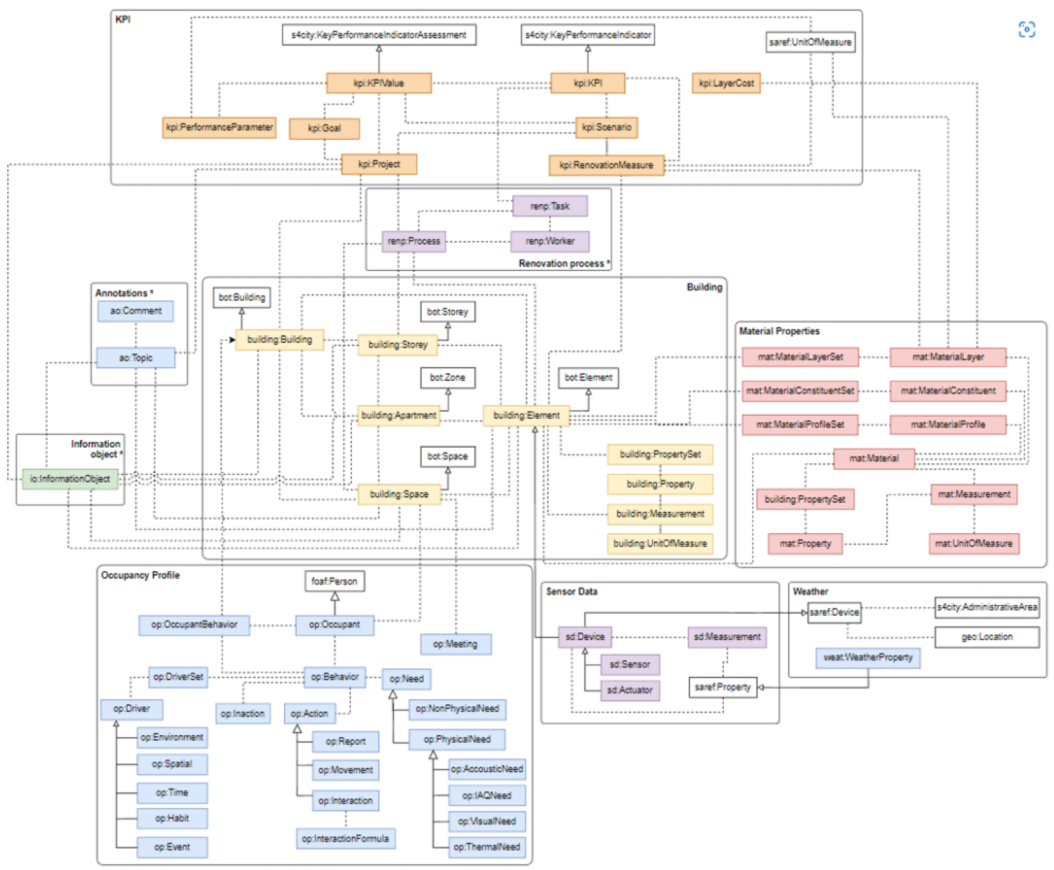


Figure 8: BIMERR project, graph Developed by Ontology Engineering Group

The following ontologies have been reused:

Prefix	Ontology namespace
foaf	http://xmlns.com/foaf/0.1/
geo	http://www.w3.org/2003/01/geo/wgs84_pos#
org	http://www.w3.org/ns/org#
xsd	http://www.w3.org/2001/XMLSchema#
owl	http://www.w3.org/2002/07/owl#
rdf	http://www.w3.org/1999/02/22-rdf-syntax-ns#
rdfs	http://www.w3.org/2000/01/rdf-schema#
saref	https://w3id.org/saref#
s4bldg	https://w3id.org/def/saref4bldg#
s4city	https://w3id.org/def/saref4city#
skos	http://www.w3.org/2004/02/skos/core#
time	http://www.w3.org/2006/time#

The ontologies follow the W3C best practices keeping metadata schemas modular and simple for easy maintenance (and non-functional requirements as 'reuse, modularity and best practices'). Some of the modules had to be developed, as the occupancy profile ontology or the KPI's. In case of the weather ontology some other ontologies have been reused (SAREF, SAREF for City (SAREF4CITY) and WGS84 Geo Positioning (WGS84_POS) ontologies).

Ontologies for Digital Twins: SPHERE[30] or COGITO[31]

These projects make use of ontologies available for their implementations. In case of BIMERR the ontology would be the base for energy analysis and renovation of existing building, whereas SPHERE and COGITO develop specific digital twin platforms. The scope and ambition of these DT implementations may have specific applications. So once again a survey of available ontologies and data models is a good starting point.
COGITO public deliverable 3.1[32] describes ontologies at domains

1. Construction site: Building;
2. Construction process;
3. Reality-capture: multi-source visual data, Internet of Things, Simulation;
4. Applications: Workflow management and smart contracts, Construction safety, Quality control.

30 SPHERE: Service Platform to Host and SharE REsidential data - ID: 820805 - From: 1 November 2018 to: 31 October 2022
31 COGITO: COGITO COnstruction-phase diGItal Twin mOdel - ID: 958310 - From: 1 November 2020 to: 31 October 2023
32 https://cogito-project.eu/library/public-deliverables/

A summary of data models and ontologies at each domain can be seen below:

Domain	Data model	Ontology
General	• IFC • cityGML	• DICO ontologies • ifcOWL • BIMERR ontologies
Building	-	• SAREF4BLDG • W3C BOT ontology • Brick schema • BPO
Process	• IFC Construction, Management, Process extension • Business Process Model and Notation	• W3C Time • BBO • DICO's Process Ontology
Multi-source visual data	• ES7 • Image data model by Grosky and Stanchev • Image data model by Clouard et al.	• Ontology model of the image object features -
Internet of Things	-	• SAREF ontology • SOSA/SSN ontology • Haystack model • W3C WoT Thing Description ontology
Simulation	-	• STATO
Workflow management	• ADOxx-based Modelling Tools data models • Novitech's I3D data model	-
Smart contracts	• bcBIM data model	• EthOn • OASIS and Smart contract extension
Construction safety	-	• SAREF4WEAR
Quality control	• GeometricQC tool • Observations data model	• BIM-based Automatic Selection Tool of Quality Control Specifications Ontology • CQIEOntology • Defect domain ontology • BrMontology

Figure 9: COGITO data models, ontologies and domains

The final overview COGITO ontology[33] network has five components:

1. COGITO Process
2. COGITO Facility
3. COGITO Resources
4. COGITO Quality
5. COGITO Safety

33 https://cogito.iot.linkeddata.es/

A graphical overview can be represented as follows:

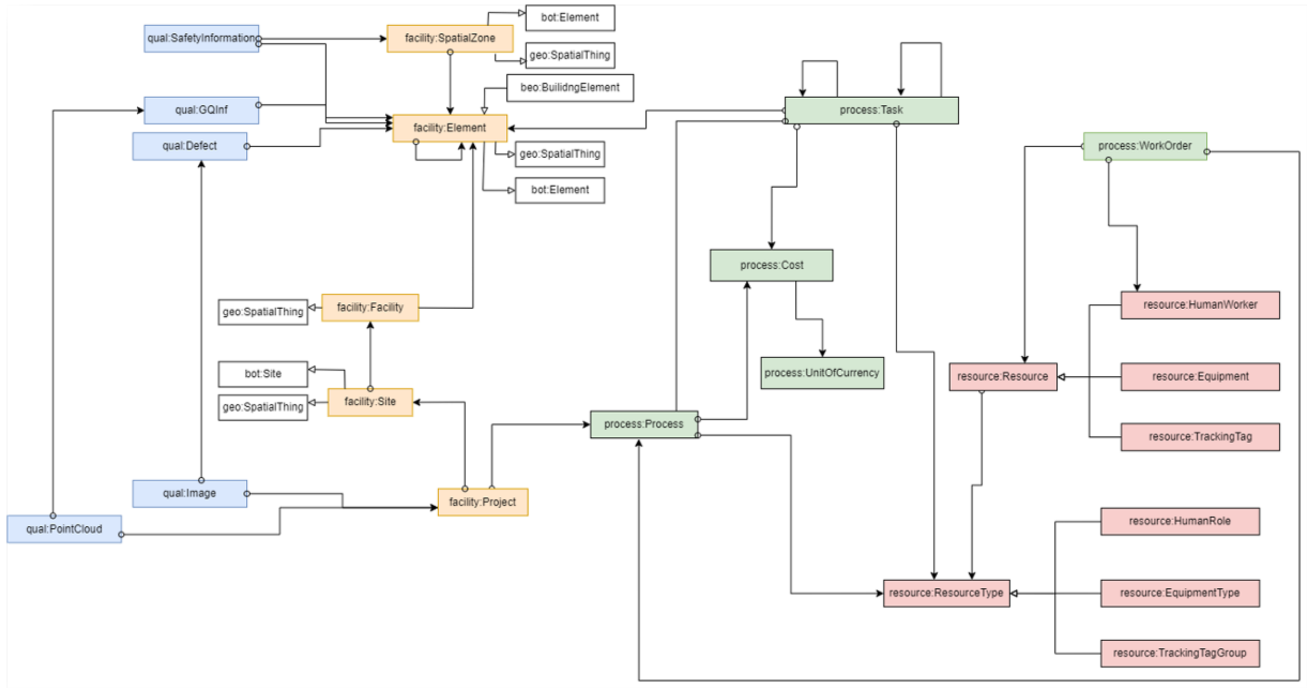


Figure 10: COGITO ontology network

SAREF and SAREF4BLDG ontologies

The **Smart Applications REFerence ontology** (SAREF)[34] is a standard ontology defined to model smart applications. This ontology has been developed and published by ETSI TC SmartM2M in 2015 and has been refined generating a new version published in March 2017, and its later version in February 2020[35].

SAREF4BLDG[36] is the SAREF extension for building devices, and aims for a more efficient interaction and integration of actors, methods and tools during the different phases of the building life cycle.

The current SAREF family of ontologies are published at the ETSI SAREF portal available at https://saref.etsi.org/. This family of ontologies includes SAREF core and the extensions for the following domains: energy (SAREF4ENER), environment (SAREF4), building (SAREF4BLDG), smart cities (SAREF4CITY), industry and manufacturing (SAREF4INMA), smart agriculture and food chain (SAREF4AGRI), automotive (SAREF4AUTO), eHealth and ageing-well (SAREF4EHAW), wearables (SAREF4WEAR), water (SAREF4WATR) and smart lifts (SAREF4LIFT). In addition, an extension modelling ontology patterns for systems is available (SAREF4SYST).

34 https://saref.etsi.org/
35 ETSI TS 103 264 V3.1.1 (2020-02)
36 https://saref.etsi.org/saref4bldg/

The SAREF core ontology aims to provide a **basic model for IoT** that could be extended and adapted in order to cover specific domains. As depicted in Figure 11, the core SAREF ontology[37] focuses on the **definition of smart applications**, therefore the main concept defined is saref:Device that include the subclasses saref:Sensor and saref:Actuator. SAREF allows the **representation of the functions** that a device can have by means of the concept saref:Function. A function could be linked to one or many **commands** (saref:Command) that represent the directive that a device must support to perform a certain function. Such commands might act upon a **state** (saref:State). SAREF also models the **concept of services** (saref:Service) that make the functions discoverable, registrable and remotely controllable. SAREF also allows the description of **tasks** (saref:Task) which a device might have been designed for. To represent the **observations** that a sensor might make the class saref:Measurement is included and linked to the saref:UnitOfMeasure in which the observation is measured and the saref:Property being measured. Finally, this ontology also represents the **consumption** (saref:Profile) of properties (e.g. energy) or commodities (e.g. water) saref:Commodity in a given period of **time** saref:Time and with an associated **price** saref:Price.

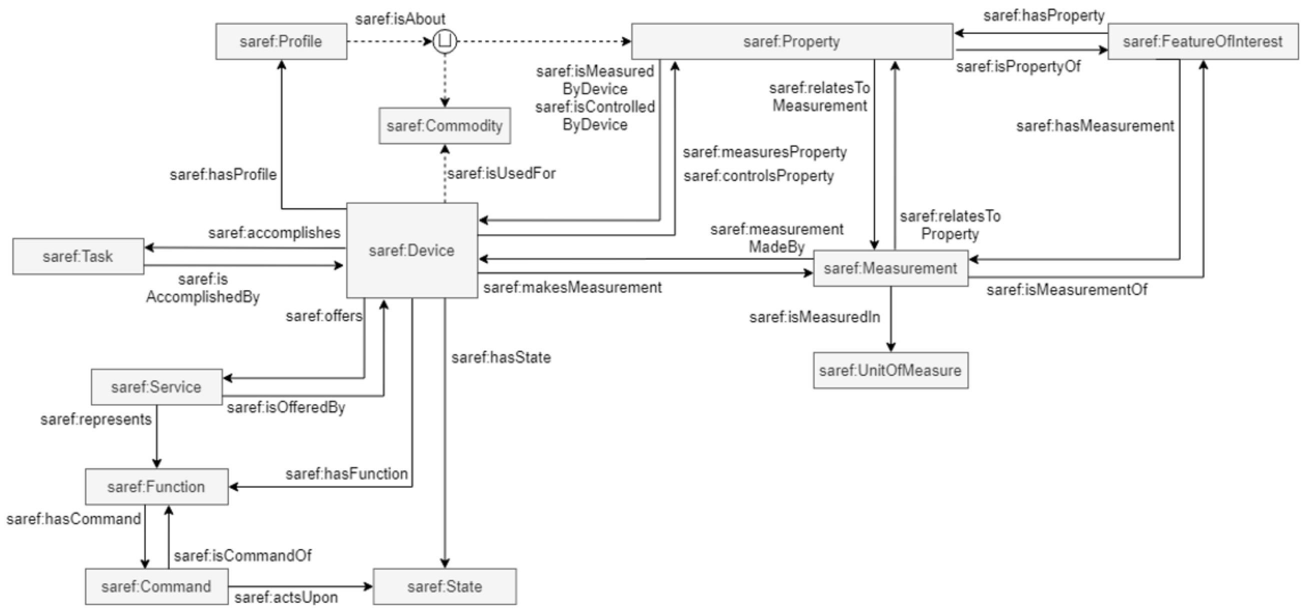


Figure 11: General overview of the SAREF ontology taken from https://saref.etsi.org/core/v3.1.1/.

37 In the following, "saref" will be used as a prefix for the namespace "https://w3id.org/saref#"

SSN/SOSA

The **Semantic Sensor Network ontology (SSN)**[38] is an ontology developed by the W3C Semantic Sensor Network Incubator Group (SSN-XG)[39]. This ontology is focused on the description of sensors, their capabilities and properties so that it "*allows the network, its sensors and the resulting data to be organized, installed and managed, queried, understood and controlled through high-level specifications*"[40]. This first version of SSN provided 4 perspectives: sensor, observation/data, system and feature and property, that can be adopted to many types of domains and applications. However, it didn't include descriptions for actuators.
In October 2017, the Spatial Data on the Web Working Group (joint effort between W3C and OGC[41]) released a new version of the SSN, that included the modelling of actuators and that differentiated between the SSN module[42] and the SOSA[43] (Sensor, Observation, Sampler and Actuator) module, among others.

This version of SSN/SOSA extends the SSO pattern (Stimulus Sensor Observation Pattern) by incorporating classes and properties for actuators and sampling. **The three major components of SOSA are "sensors and observations", "samplings and samples" and "actuators and actuations"**. As Figure 12 depicts, SOSA[44] provides the concept sosa:Sensor that make sosa:Observation about sosa:ObservableProperty to describe sensing acts. The observable property is a property of a sosa:FeatureOfInterest. In addition, to describe the results of sensing acts, SOSA present sosa:Sampler that make sosa:Sampling of some sosa:FeatureOfInterest to produce sosa:Sample but also the sosa:Result concept could be used. Finally, to describe the ability of using actuators to perform some actions, SOSA models the concept sosa:Actuator that is able to perform some sosa:Actuation over a sosa:ActuationProperty.

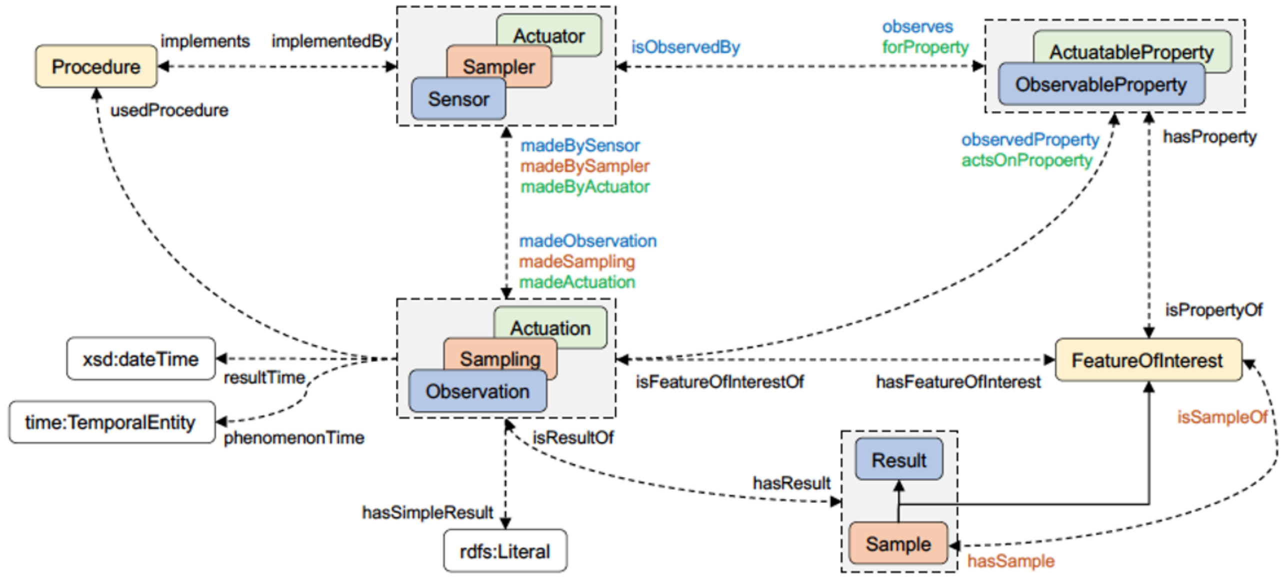


Figure 12: Overview of the core structure of the SSN ontology

38 https://www.w3.org/TR/vocab-ssn/
39 https://www.w3.org/2005/Incubator/ssn/
40 Compton et. al., 2012
41 Open Geospatial Consortium
42 http://www.w3.org/ns/ssn
43 http://www.w3.org/ns/sosa
44 In the following, "sosa" and "ssn" will be used as prefixes for the namespaces "http://www.w3.org/ns/sosa/" and "http://www.w3.org/ns/ssn/" respectively

BOT ontology[45]

BIM Levels of Maturity[46], with the web-based BIM Level 3 on the horizon, is the driving force of the Wide Web Consortium Linked Building Data Community Group (W3C LBD-CG) to introduce the Building Topology Ontology (BOT). It provides a **high-level description of the topology of buildings** including storeys and spaces, the building elements they contain, and their web-friendly 3D models.

Regardless of BIM level 3 requirements, **BOT has three great merits**: (1) BOT knows where it comes from and where it wants to go. (2) Some of the developers in the group have a deep knowledge of topological practical questions in buildings, opening the door for future queries and applications of automatic checking. And (3): it denotes a practical experience with real project management in big corporations. These are not minor observations as **BOT could be a real core solution** in the future digitalization paradigm in construction.

Initially BIM implementation and the IFC standard can be considered globally accepted and that could be considered the starting point. The W3C LBD-CG group aimed at creating a **lightweight BOT ontology** that would not have the same drawbacks found in IFC in terms of size and complexity. And that could open a Semantic Web space for construction, which would be the final objective of their implementation. Which is impossible using IFC but would be possible indeed integrating BOT and other aligned ontologies. BOT has a pragmatic and powerful implementation of topological problems which arise during practical work in a project. So, the community behind this ontology have a direct and deep knowledge of these questions. Doing an energy analysis connectivity of spaces and the interfaces (walls) are not always easy to extract from native programs to IFC, and additional entities (as bot:interface) can be a good help. Topology means problems of connectivity between entities, not just the description of the building.

And finally, the minimalist representation approach is far from being inefficient or incomplete, but a smart understanding of how the core of the building should be.

We could say that BOT offers a descriptive and topological tool for building representation, with potential alignments with other ontologies.

What is BOT?

BOT has three main classes: bot:Zone, bot:Element, and bot:Interface.

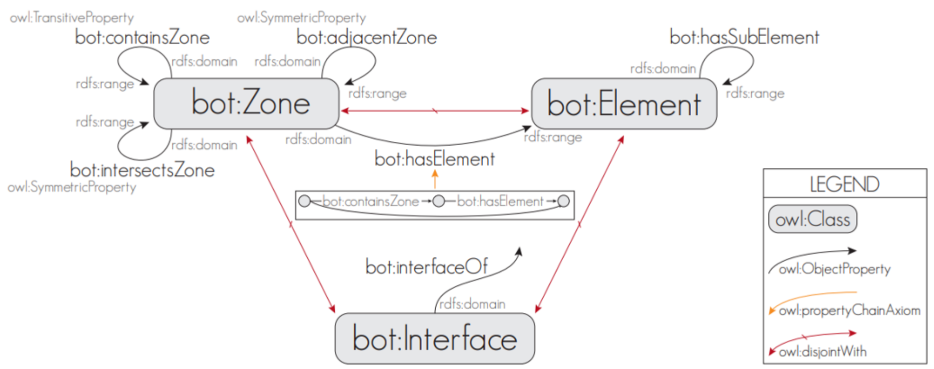


Figure 13: Illustration of the main three classes of BOT

45 https://w3c-lbd-cg.github.io/bot/

46 M. Bew and M. Richards, Bew-Richards BIM maturity model, in: BuildingSMART Construct IT Autumn Members Meeting, Brighton, UK, 2008.

A bot:Zone is defined as a part of the world that has a 3D spatial extent. Four sub-classes of bot:Zone are defined: bot:Site, bot:Building, bot:Storey and bot:Space.

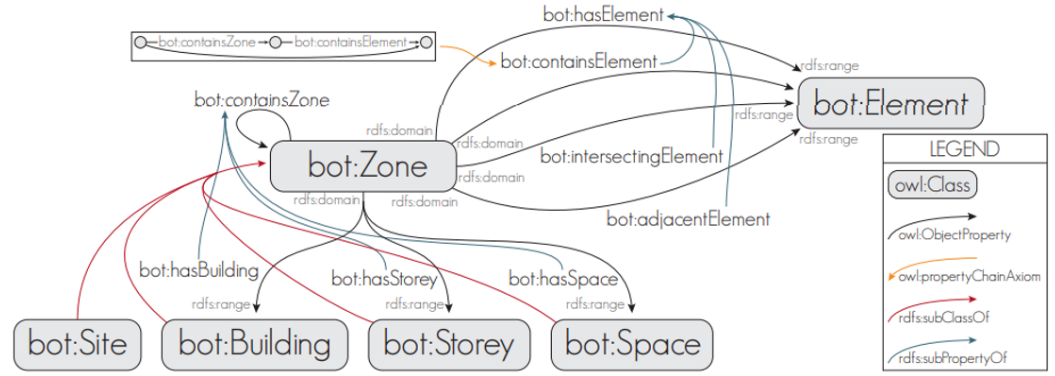


Figure 14: Four sub-classes of bot:Zone and the three sub-properties of bot:hasElement

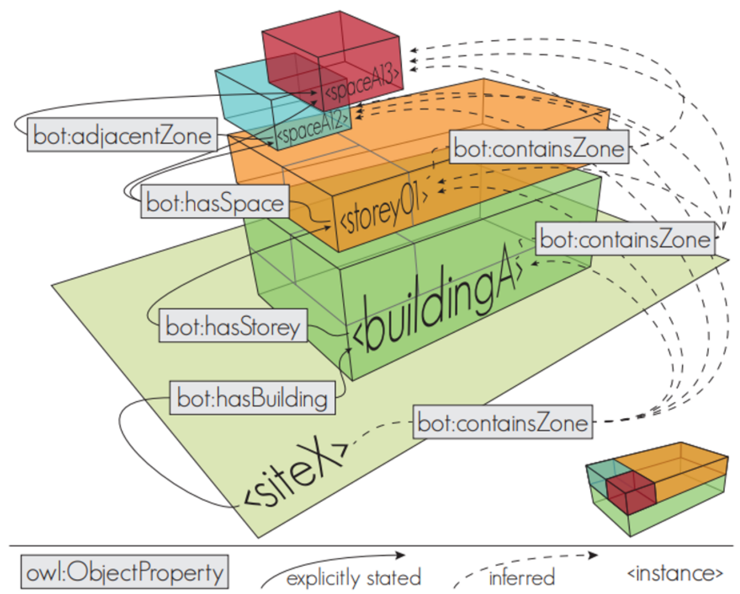


Figure 15: Zones in BOT follows a matrioska doll principle

A bot:Element is defined as a constituent of a construction entity with a characteristic technical function, form or position. Elements can host sub-elements, which is defined using the bot:hasSubElement property. Three main topological relationships between zones and elements are defined:

- bot:adjacentElement links a zone to an element that shares part of its boundary;
- bot:intersectingElement links a zone to an element whose 3D extents is partly shared;
- bot:containsElement links a zone to an element.

The class bot:Interface is used to describe the relationship between some specific zones and elements in detail. The concept of bot:Interface is useful in different situations:

a) the heat transmission area of the surface between a space and an adjacent wall can be used to determine the heat loss from that space through this wall;
b) the localisation of the intersection between a pipe and a wall can be used to specify where to apply fire sealing;
c) the type of access between two zones can be used to specify access restrictions for use in door navigation.

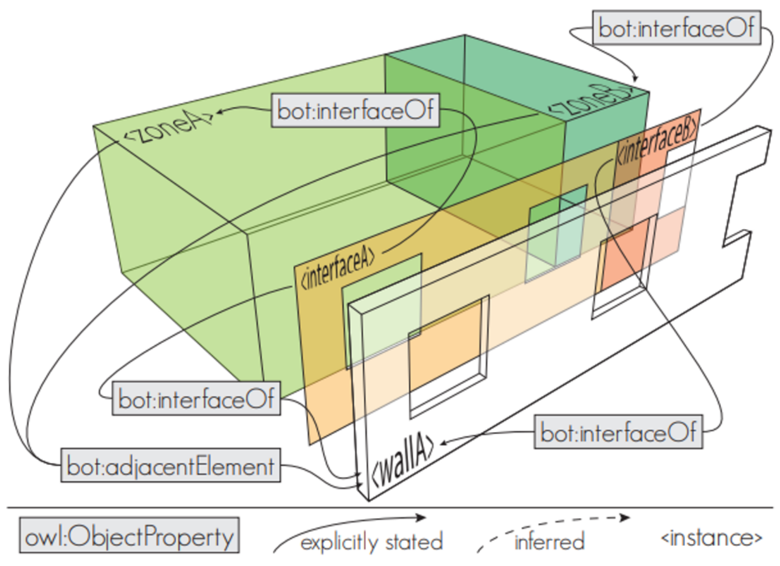


Figure 16: Interfaces between two zones and a wall. Inter faces can be used to qualify (i.e., attach additional information to topological relationships between zones, elements, or zones and elements.

Geometry and properties are outside the scope of BOT but there are ways to integrate them as well.

BOT is designed to function as a central element in the interdisciplinary communication. In addition, it aims at being the key entry point to connect the construction sector to adjacent domains.

BIM4EEB Digital Construction Ontologies, DiCon[47]

Digital Construction Ontologies (DiCon) were developed in the BIM4EEB based on the earlier ontology work carried out in the DiCtion project in Finland. DiCon is composed of a set of interrelated ontology modules that aim to capture the different aspects of construction and renovation projects. Since there are significant standardized data models and other ontology work in this domain – as reported in this white paper – the objectives of DiCon have been to link to and integrate with these other models.

47 https://digitalconstruction.github.io/v/0.3/index.html

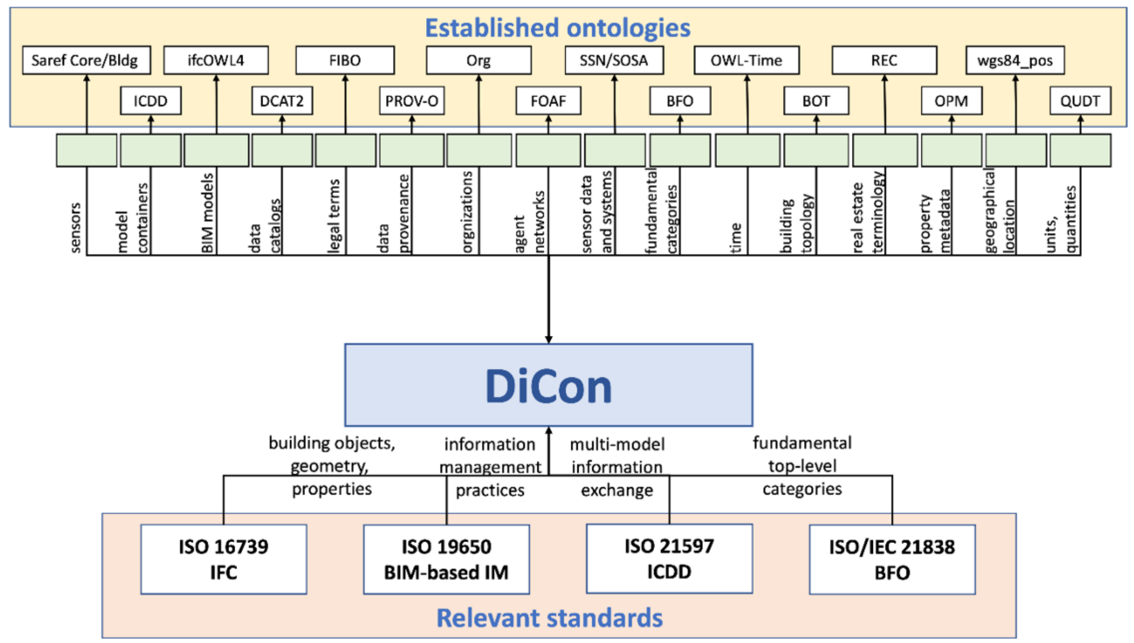


Figure 17: Relationships of DiCon with relevant standards and established ontologies

Figure 17 shows the relationship of DiCon ontologies with the relevant standards and established ontologies related to construction domain. As technical solutions for integration the DiCon uses the following:

- A top-level ontology for the uniform categorization of concepts in the different modules of DiCon.
- Alignment modules that establish the correspondences of DiCon terms with the terms of other ontologies. All the references to external concepts are placed in the alignment modules (except for instance vocabularies).

As shown in Figure 17, the standards considered most relevant for DiCon are:

- ISO 16739 Industry Foundation Classes (IFC): The universally supported BIM standard, according to which BIM models are used in BIM4EEB and DiCon.
- ISO 19650 BIM-based information management: A process standard whose relevant terminology is formalized in the DiCon.
- ISO 21597 Information Container for Linked Document Delivery (ICDD): A standard for the exchange of packages of interrelated documents or models.
- ISO/IEC 21838 Basic Formal Ontology (BFO): A standard top-level ontology used in DiCon.

DiCon is organized into several modules, shown in Figure 18 as blue boxes, where the arrows between the blue boxes represent the import relations (owl:imports) between the modules.

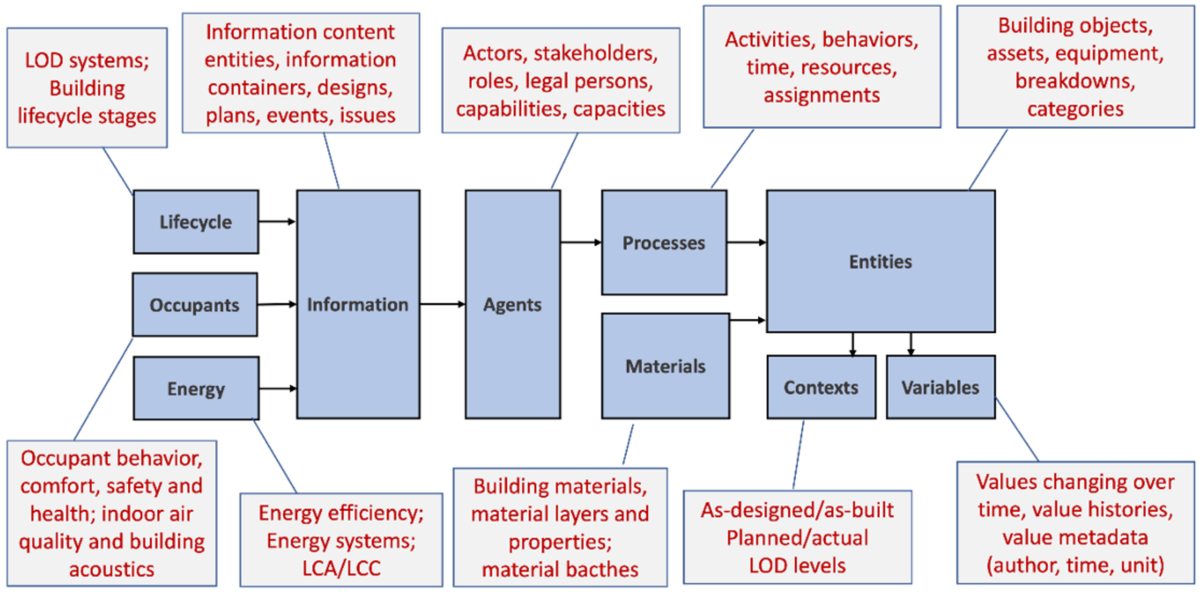


Figure 18: The modules of DiCon

The core modules of DiCon are the following

- ***Contexts***: The definitions to maintain different realms of information:
    - as-is, as-designed, as-built and as-maintained models,
    - planned versus actual execution of activities,
    - models containing information up to a specific LOD level,
    - plans for alternative construction methods or renovation scenarios.
- ***Variables***: A support for the objectification of properties to enable the attachment of richer data about them: values at different timepoints, quantity kinds, unit of measurement, and constraints on the value.
- ***Entities***: Based on BFO, the basic classes and properties needed for the representation of construction and renovation projects, such as building object, location, material batch, equipment, agent, and spatial and temporal regions.
- ***Processes***: Different kinds of processes, such as activities, services, and behaviour processes, including activity flows, capabilities, and resources.
- ***Agents***: Actors and stakeholders over the construction lifecycle, to support data sharing about social, organizational, legal, and contractual relations.
- ***Information***: Information content entities, information models, and information containers in construction and renovation, including models, plans, scenarios, messages, issues, videos and point clouds.

DiCon includes additional modules for energy efficiency and energy systems (*Energy*), occupant behaviour and profiles (*Occupancy*), building materials and material layers (*Materials*), and LOD levels and product lifecycle stages (*Lifecycle*). There are also some instance vocabularies relevant in renovation domain (*Units*, *Levels*, and *Stages*).

A part of the class hierarchy of DiCon is shown in the Figure 19. The prefixes are given in the left-up corner of the diagram; the prefixes are also indicted with a consistent colouring all through the diagram. The upmost layers of the ontology are based on the fundamental categories of BFO.

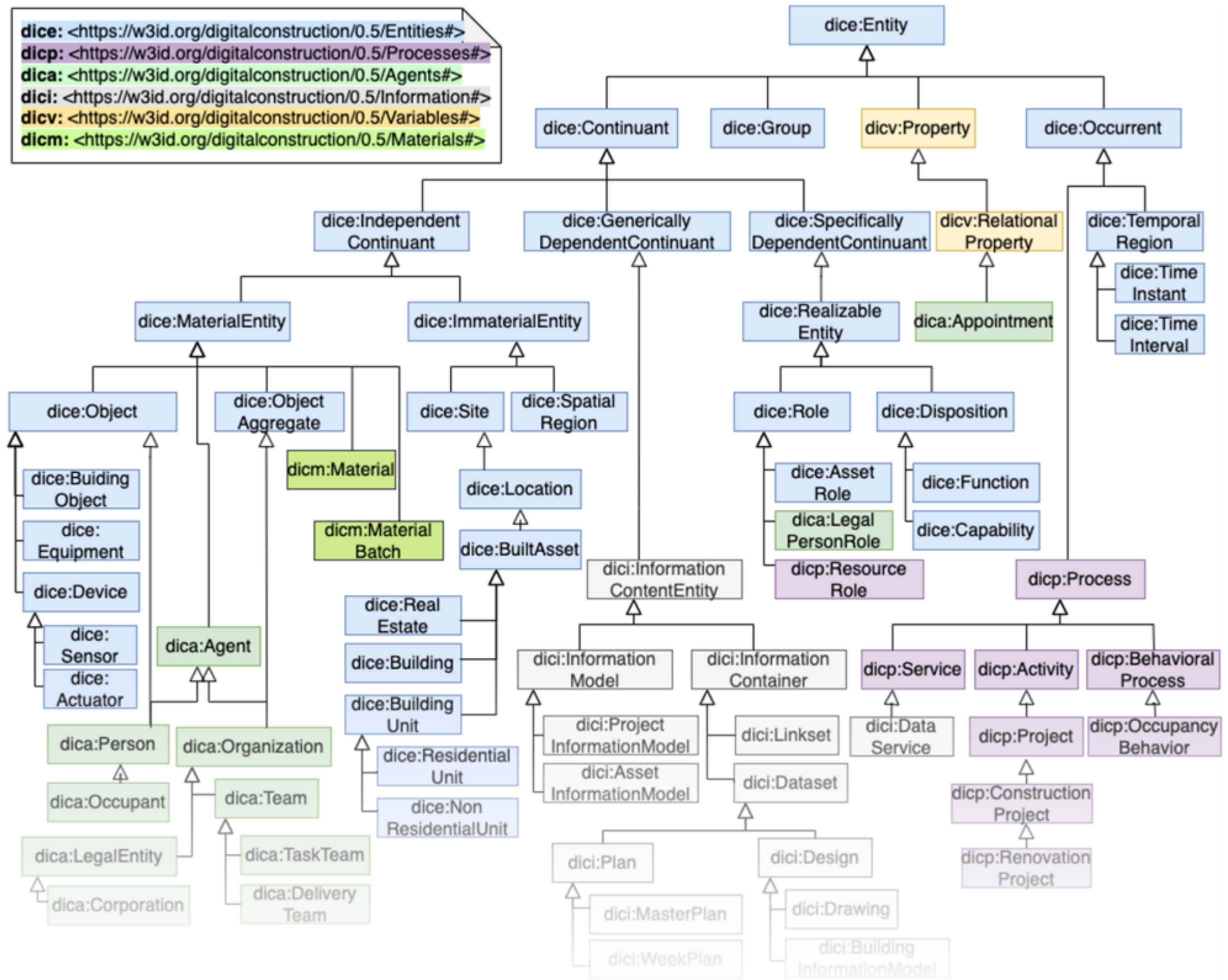


Figure 19: A part of the class hierarchy of DiCon

BRICK ontology[48]

Brick was originally developed by representatives of UC Los Angeles, UC Berkeley, UC San Diego, Carnegie Mellon University, University of Virginia, University of Southern Denmark and IBM Research. It is now developed and maintained by the Brick Consortium (non-profit organisation). It is supported by the Department of Energy of U.S., the European Commission, Innovation Fund Denmark, Intel, Johnson Controls, King Abdullah University of Science and Technology, NSF and TOPAs.

Brick is an ontology that captures the entities and relationships of buildings and their subsystems. Brick describes different operational, structural and functional facets of a building.

48 https://brickschema.org/

Core concepts of BRICKS

Core concepts of BRICKS are:

Entity: an entity is an abstraction of any physical, logical or virtual item; the actual “things” in a building.

- Physical entities are anything that has a physical presence in the world, like equipment, thermostat, or electric meters. Spatial elements as rooms or floors are physical entities as well.
- Virtual entities are anything whose representation is based in software. Examples are sensing status and setpoints.
- Logical entities are those entities or collections of entities that are defined by a set of rules. Examples are HVAC zones and Lighting zones. Concepts such as class names and tags (defined below) also fall into this category.

Tag: a tag is an atomic fact or attribute of an entity. Examples of tags are sensor, setpoint, air, water, discharge, leaving and vav[49]. Brick borrows the concept of tags from Project Haystack in order to preserve the flexibility and ease of use for annotation; however, Brick does not rely on tags alone to determine the type of an entity.

Class: a class is a named category used for grouping entities. Classes are organized into a hierarchy.

Relationship: a relationship defines the nature of a link between two related entities. Examples of relationships are *encapsulation* (one entity is contained within another), *sequence* (one entity takes effect before another in some process) and *instantiation* (one entity’s type is given by another entity).

Graph: an abstract organizational data structure representing a set of entities (nodes) and relationships (edges). Brick is represented by a directed, labelled graph.

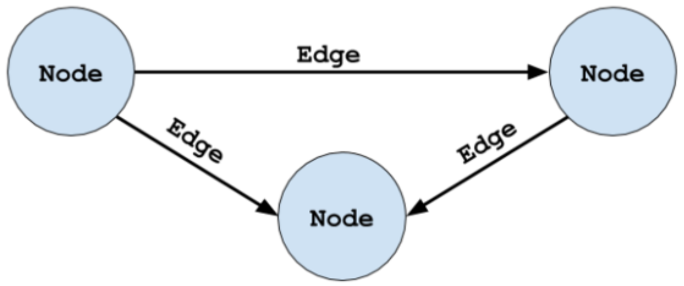


Figure 20: Brick´s graph

49 Variable Air Volume

Brick Model: a Brick model is a digital representation of a building that adheres to the Brick schema. Entities in a Brick model are classified according to the classes defined by Brick, and are connected using the relationships defined by Brick.

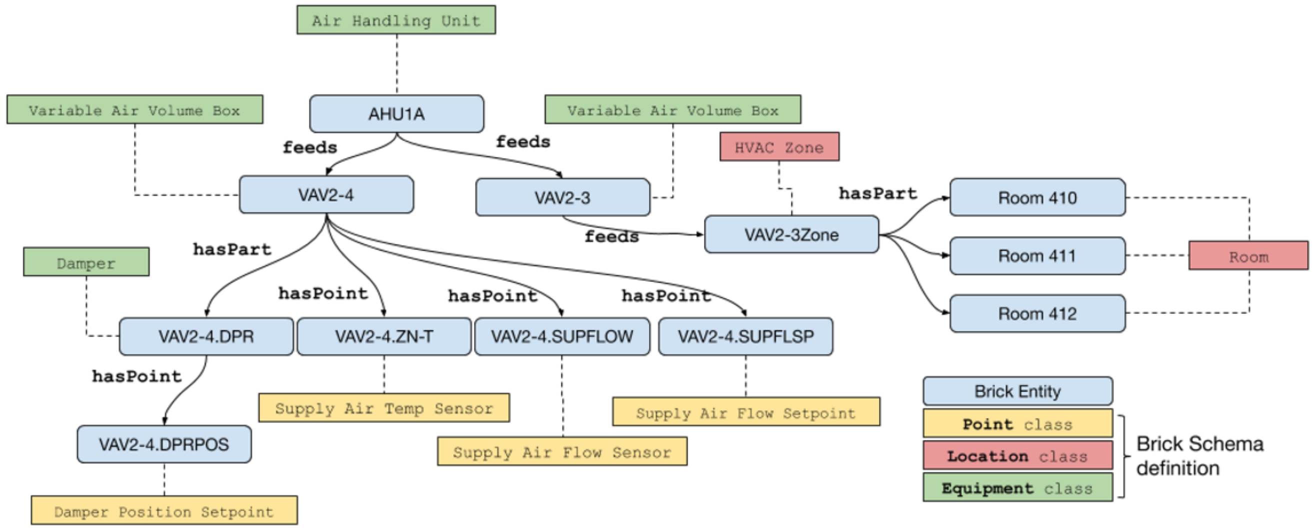


Figure 21: Brick model example

BPO: Building Product Ontology[50]

The Building Product Ontology defines concepts to describe (building) products in a schematic way. It provides methods to describe assembly structures and component interconnections, and attach properties to any component without restricting their types, as is often the case in template-driven product descriptions. To allow the description of complex properties, it also contains terms for unordered, two-dimensional lists.

ORKG, existing smart building domain ontologies comparison (August, 2022)

PhD Sanju Tiwari compiled at orkg.org[51] the main 15 ontologies we can find today in building domain. This detailed comparison describes details as number of classes, instances, serialization, URL or reused ontologies. It may be a good It may be a good starting point to study which options we have available today in construction.

50 https://w3id.org/bpo
51 https://orkg.org/comparison/R214164/

3.2. REAL ESTATE

Figure 22: Real Estate operations involve interactions within buildings and may share the knowledgein order to scale towards smartcities

The RealEstateCore (REC) ontology[52]

Property owners can use RealEstateCore to describe the data of interaction within the buildings that they operate – as well as the management, storage, and sharing of this data. RealEstateCore is a modular ontology, that is, a collection of data schemas that describe concepts and relations that can occur in data that is generated to model buildings and building systems, or that is sourced from such systems.

Having the shared language that these data schemas provide enables property owners to connect their buildings with new services on a large scale, and not have to worry about building- or technology-specific implementation details and formats.
RealEstateCore alignment table would be as follows:

STANDARD	Business administration	IoT	Building Management Systems (BMS)	Building Information Modeling (BIM/IFC)	RDF-based
BRICK Schema	No	Yes	Yes	Partial	Yes
BOT	No	No	No	Yes	Yes
SAREF4BLDG	No	Yes	Partial	Partial	Yes
NGSI-LD	No	Yes	No	No	Yes
BACnet/Ashrae 223P	No	Partial	Yes	No	Yes
Haystack 3	No	Partial	Yes	No	No
Haystack 4	No	Partial	Yes	No	Yes

52 https://www.realestatecore.io/

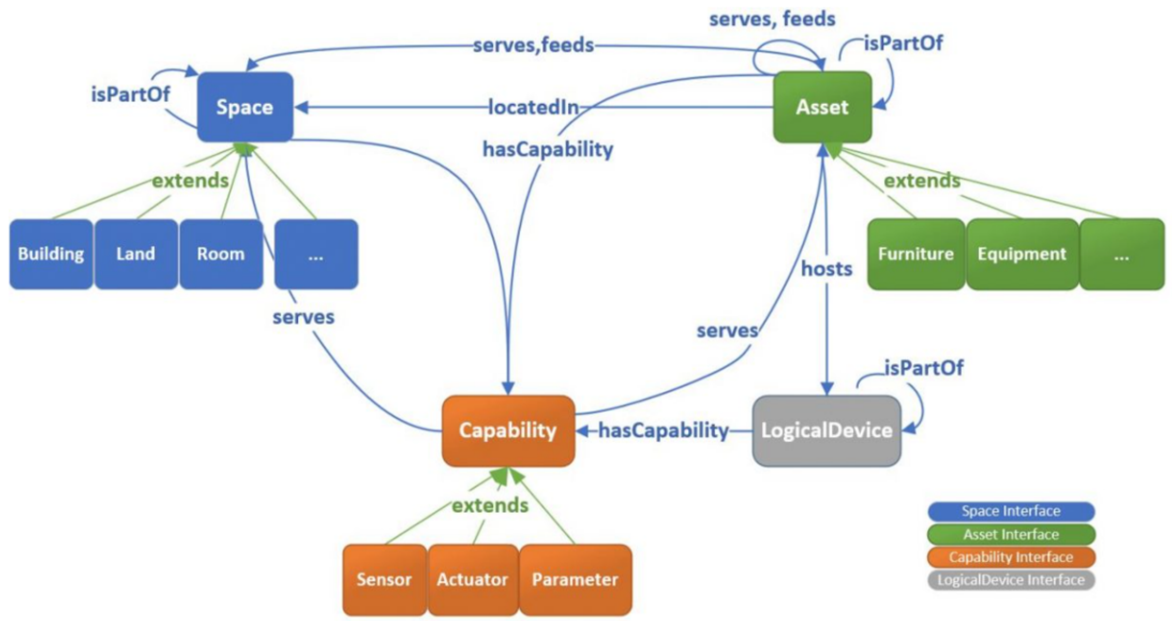


Figure 23: REC key classes subset (extracted from REC webpage)

A brief description of the main classes would be the following:

- **Spaces:** A core:Space is contiguous part of the physical world that has a 3D spatial extent and that contains or can contain sub-spaces. For example a Region can contain many pieces of Land, which in turn can contain many Buildings, which in turn can contain Levels and Rooms. This concept is comparable to a Zone in the BOT ontology.
- **Building components:** A core:BuildingComponent is a part that constitutes a piece of a building's structural makeup, for example Façade, Wall, Slab, RoofInner, etc
- **Assets:** A core:Asset is an object which is placed inside of a building, but is not an integral part of that building's structure. We provide a substantial hierarchy of assests, for example architectural, furniture, equipment, systems, etc.
- **Logical Devide:** core:LogicalDevice: A physical or logical object defined as an electronic equipment or software that communicates and interacts with a digital twin platform. A logical device could be an integrated circuit inside of a smart HVAC unit, or a virtual server running on a Kubernetes cluster. Logical devices can have Capability instances (through hasCapability) that describe their input/output capabilities. If Logical Devices are embedded within Asset entities (through the hostedBy property) such capabilities typically denote the capabilities of the asset.

Brick SchemaBRICK ontology and RealEstateCore have announced a major harmonization effort between these two smart building metadata standards.

'The Brick Schema and RealEstateCore will combine the best of both worlds. It will be a complete open-source standard based on Semantic Web technologies which the industry can use efficiently for a long time. We enable buildings to become good inhabitants of smart cities'[53].

53 Gabriel Fierro (Brick Consortium) and Erik Wallin (RealEstateCore Consortium)

The combination of Brick Schema and RealEstateCore comprehensively will cover the following domains:

- Building management systems (e.g., HVAC, Access control, Elevator)
- Business administration systems (e.g., CAFM, ERP, Sustainability reporting)
- IoT devices (e.g., Indoor climate, People counting)
- Blueprints (e.g., Building information modelling, DWG)

The harmonization of the two standards is also being done with an eye towards **compatibility with the upcoming ASHRAE 223 standard**. Brick Schema 1.3 and RealEstateCore 4.0 are both targeting a release on August 30, 2022. Full harmonization will occur in future releases of the two standards, which is expected to occur quickly after the August release.

3.3 METHODOLOGIES[54]

The development of methodologies, methods and techniques for building ontologies have produced numerous results since the 1990s. Some examples of methodologies are METHONTOLOGY[55], On-To-Knowledge[56], DILIGENT[57] and the NeOn Methodology[58]. Most of these approaches, generally considered “traditional” methodologies, were designed to build ontologies from scratch, except for the NeOn Methodology, which also contemplates development through the reuse of existing resources.

However, the use of these methodologies in some of the most recent ontology developments, especially those focused on the development of lightweight ontologies and vocabularies, has shown some of its difficulties and shortcomings.

In this sense, it is worth noting the rigidity of the traditional processes instead of following agile development practices, the non-orientation towards collaborative development and their orientation towards taking requirements based on experts, instead of being directed by data, for example. In addition, none of the above methodologies considers activities typical of linked data development projects, so they do not take into account, for example, their publication following the best practices for publishing vocabularies on the web.

For these reasons, in recent years work has begun on the development of **ontology development methodologies** that cover these shortcomings. One of these methodologies is the LOT Methodology[59] (Linked Open Terms Methodology), and has been proposed by the Ontological Engineering Group of the Universiddad Politécnica de Madrid, following good practices and lessons learned in the development of ontologies. This methodology follows the philosophy of agile ontology development and presents a series of simple steps based on the available data, and taking into account its particular characteristics. In addition, this methodology is oriented to the **reuse of vocabularies**, following an iterative approach that can be combined with agile software development methodologies, being able to organize the development phases in sprints. In addition, this methodology will be combined with external recommendations and guides known by the team members to carry out specific activities, such as: creation of **competence questions**, transformation of non-ontological resources, publication of ontologies, etc. These specific guidelines can come from methodologies such as the NeOn Methodology, from recommendations of the W3C web consortium, to name a few.
The phases proposed by the LOT Methodology, represented schematically in the following figure:

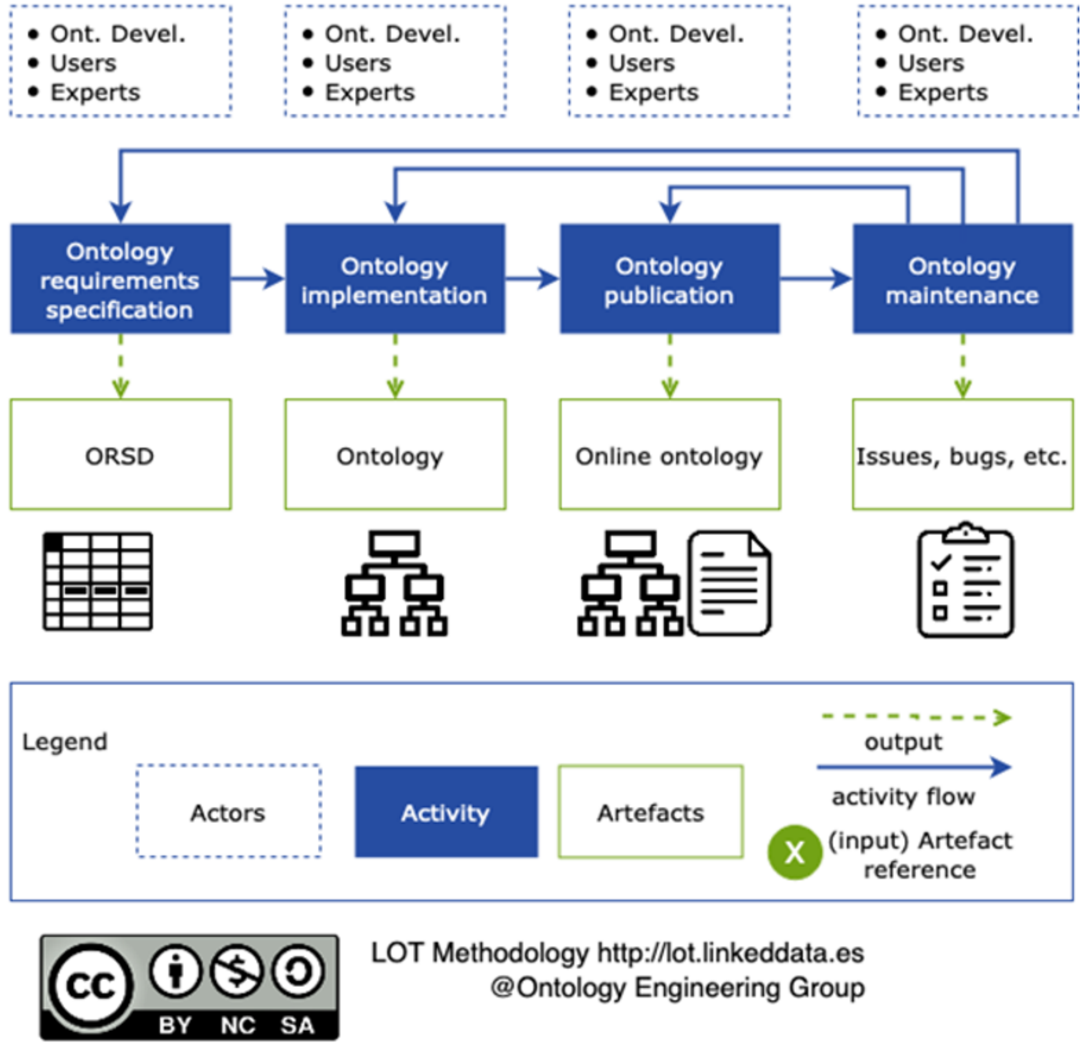


Figure 24: Methodologies, phases proposed

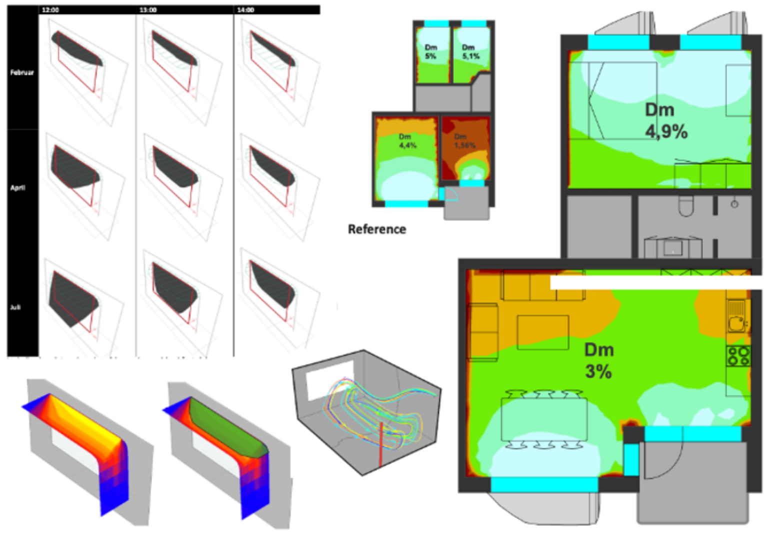
This methodology has already been applied in the creation of more than 20 ontologies. For example, it has been used in the development of some of the ontologies that were included in the “UNE 178301 Smart Cities Standard. Open Data”, the BIMERR (https://bimerr.iot.linkeddata.es/) and COGITO (https://cogito.iot.linkeddata.es/) ontologies and other European projects as VICINITY (http://vicinity.iot.linkeddata.es) or AURORAL (https://auroral.iot.linkeddata.es/), and some extensions of the SAREF standard ontology.

54 María Poveda Villalón, UPM
55 Ref: METHONTOLOGY: From Ontological Art Towards Ontological Engineering”
56 On-To-Knowledge Methodology (OTKM)” https://doi.org/10.1007/978-3-540-24750-0_6
57 Distributed Engineering of Ontologies (DILIGENT)” https://doi.org/10.1007/3-540-28347-1_16
58 Ontology Engineering in a Networked World” https://doi.org/10.1007/978-3-642-24794-1
59 Poveda-Villalón, M., Fernández-Izquierdo, A., Fernández-López, M., & García-Castro, R. (2022). LOT: An industrial oriented ontology engineering framework. Engineering Applications of Artificial Intelligence, 111, 104755. https://doi.org/10.1016/j.engappai.2022.104755

4.-USE CASES

NIRAS[60], HVAC deep checking

Niras is a global engineering company with HQ located at Copenhagen (Denmark). It is a value-driven, multi-disciplinary engineering consultancy committed to sustainable progress and service delivery. During the SSoLDAC (Summer School of Linked Data in Architecture and Construction), PhD Mads H. Rasmussen was describing with detail his professional trajectory from Architecture to HVAC engineer, and specific application they use today at NIRAS for data analysis and ontologies.



Radiance-derived window shade and the effect on daylight access

His long experience on design, integrated functional projects or HVAC installations makes him especially valuable as data analyst, as he may know why to use linked data for. As a consequence, the implementation of the BOT ontology. He is that ‘domain expert’ needed to make useful applications for ontologies.

Between other projects we can mention the Aarhus University Digital Twin, with a semantic data model of campus, buildings, spaces, sensors and mechanical system. The Bart project, with two lines of development towards building Solibri rule sets for checking Danish building regulations (track A), and a second track B for investigating new methods (including Linked Data). In May 2021 a new project is running in collaboration with Saint Gobain (Contech Pioneer Project 5), looking for automatic design of walls based on boundary conditions. Another interesting project called K4A was oriented to the formalization of design frames and checks that are continuously checked against. And last but not least, the project Almen Byggeportal, a digital project delivery for all subsidized residential buildings in Denmark, with examination declarations, document and model checking and final building element documentation (Material passports, O&M material etc.).

60 https://www.niras.com/


```
[?s, ex:segmentVolumeFlow, ?qTot]
:-
[?s, a, fso:Segment],
AGGREGATE(
   [?s, ex:indirectlySuppliesFluidTo, ?t],
   [?t, a, fso:Terminal],
   [?t, ex:terminalVolumeFlow, ?q]
   ON ?s
   BIND SUM(?q) AS ?qTot
) .
```

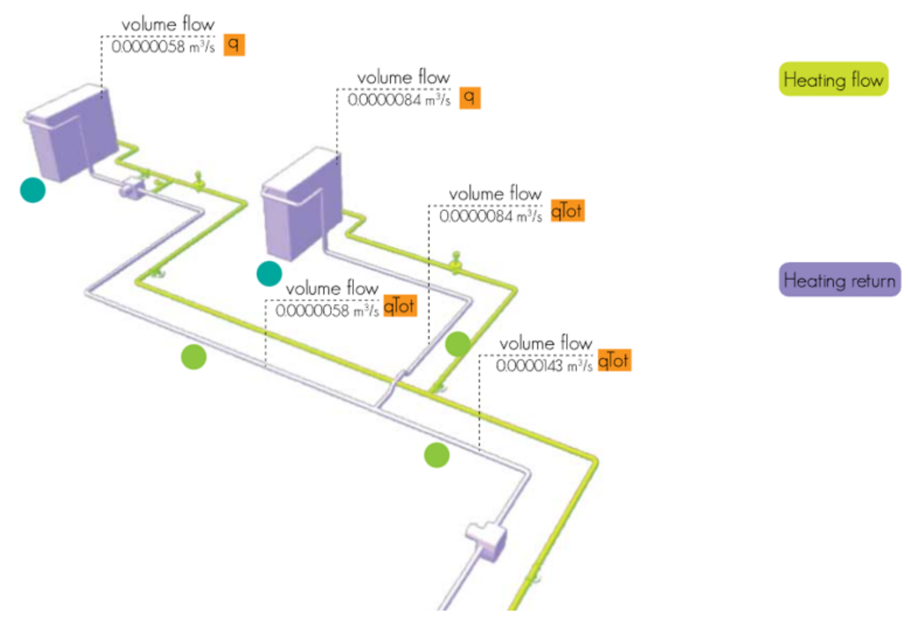


Figure 25: Semantic web technologies at NIRAS. Incremental reasoning

In all these projects NIRAS makes extensive use of Semantic Web technologies, including **incremental reasoning** for complex checking processes affecting MEP systems.

Even with this impressive demonstration of advanced technology Mr. Rasmussen points out the human factor, which means that in many cases due to the lack of time, change of project human resources, lack of coordination, … many processes are still (in the end) manual. This is the challenge and the opportunity we are facing, and with Rasmussen´s own words: “The technology is maturing - all the bits and pieces are already there. Go and create something meaningful with it!”

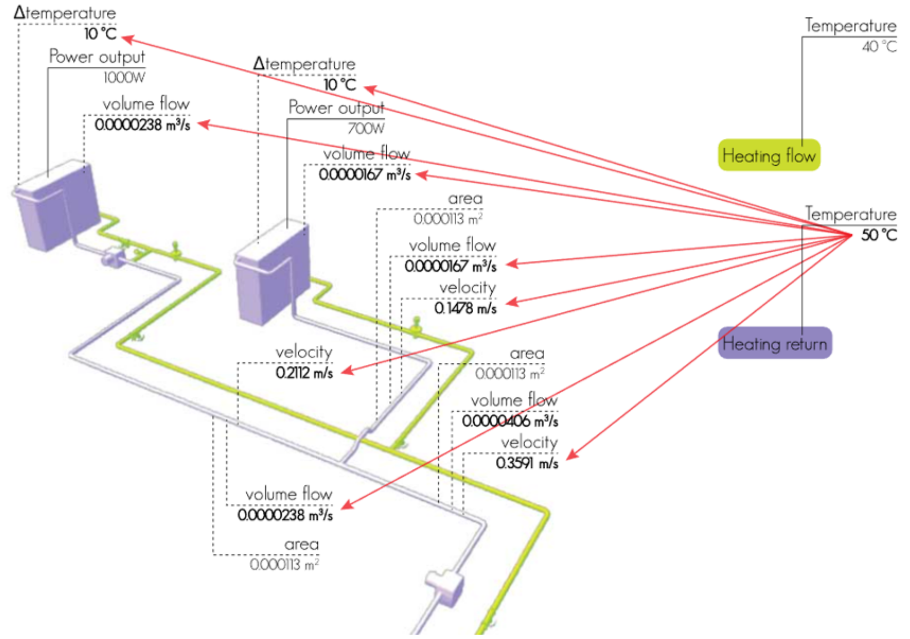


Figure 26: Incremental reasoning in MEP systems

The potential of incremental reasoning

Incremental reasoning would be the reasoning performed in response to updates by minimising the amount of re-computations needed. When an ontology is updated, only the reasoning task related to this update is performed and the whole semantic reasoning process does not need to be processed again. There are several techniques to improve the computation efficiency, such as windowing the scope of data or operator sharing between rules, and checking models to see if these techniques improve speed and quality.

BIMprove, safety-critical situation detection

The H2020 project BIMprove[61] harnesses the power of cutting-edge Digital Twin Technology to lead construction sites into the industry 4.0 revolution. Fire protection, fall prevention and scheduling and cost benefit are specific goals of the project.
One of the tasks in this project was to develop an automatic detection of safety measures, especially safety nets, from photographs taken at construction sites. From that inspection a data model is created using a specific ontology developed in this project.

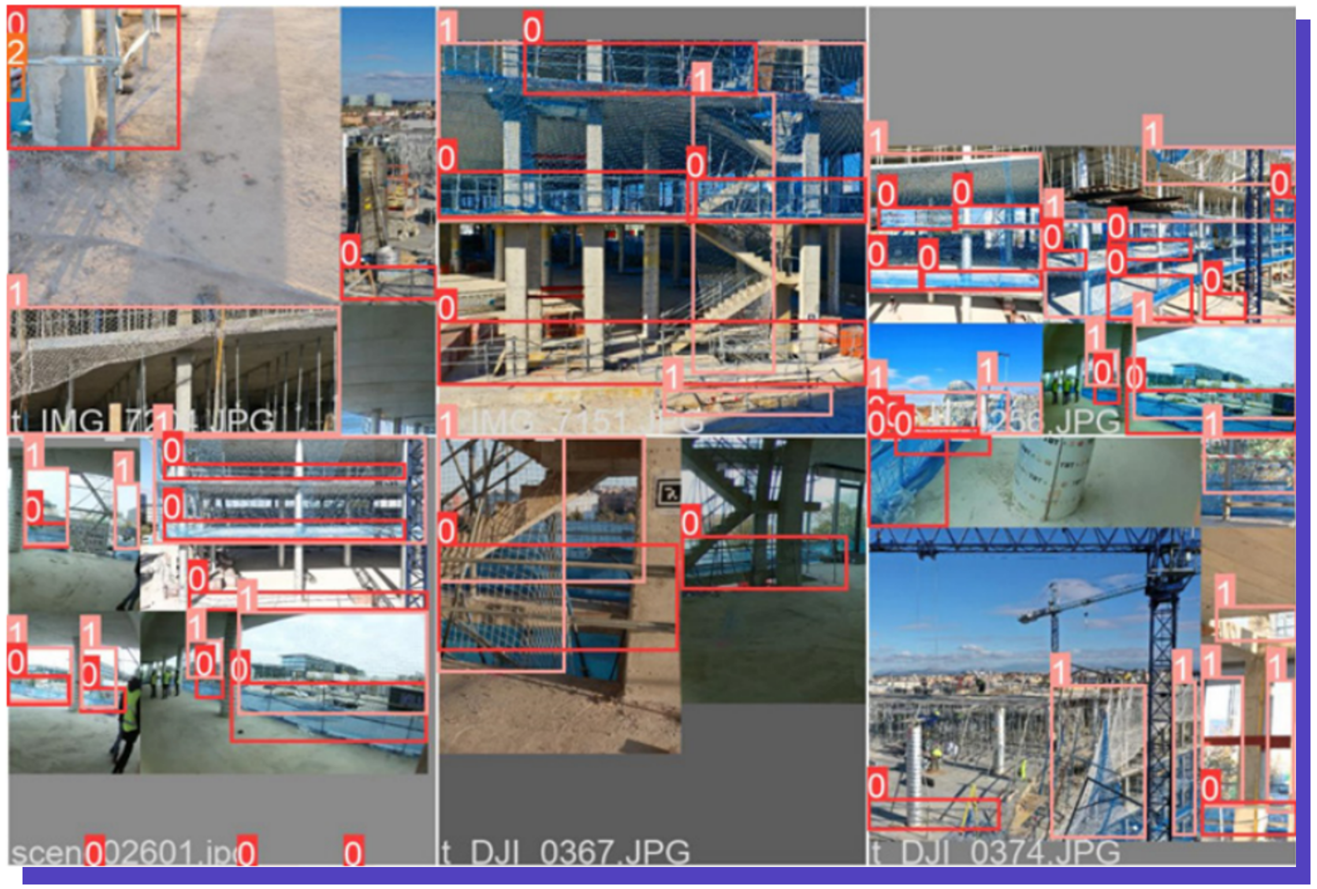


Figure 27: Tiled image samples of the augmented data with artificial scaling, shearing, flipping and mosaic of the original

61 https://www.bimprove-h2020.eu/

Detection results are returned as JSON text. The coordinates are counted as pixels of the original image dimensions.

The purpose of the BIMprove Risk and Image Data Management Service is to store and provide access to the data created by the Safety Related Visual Object Detection described above. The Risk and Image Data Management Service includes three main parts that are:

1. The database used to store the metadata provided by the Safety Related Visual Object Detection. The selected database solution is Apache Fuseki graph database that supports RDF (Resource Description Framework) format and SPARQL query language. The database structure follows the risk data ontology specifically developed to address the requirements of risk data management. A visualization of the ontology is shown in Figure 28.
2. REST[62] interface that provides an access to the database. The interface includes two basic access methods that are insert data and query data. The interface is implemented using Python language and Tornado web framework with its asynchronous networking library.
3. HTML based GUI that enables examining the content of the database (as well as the results of the Safety Related Visual Object Detection). The GUI is supported by the Flask micro web framework written in Python.

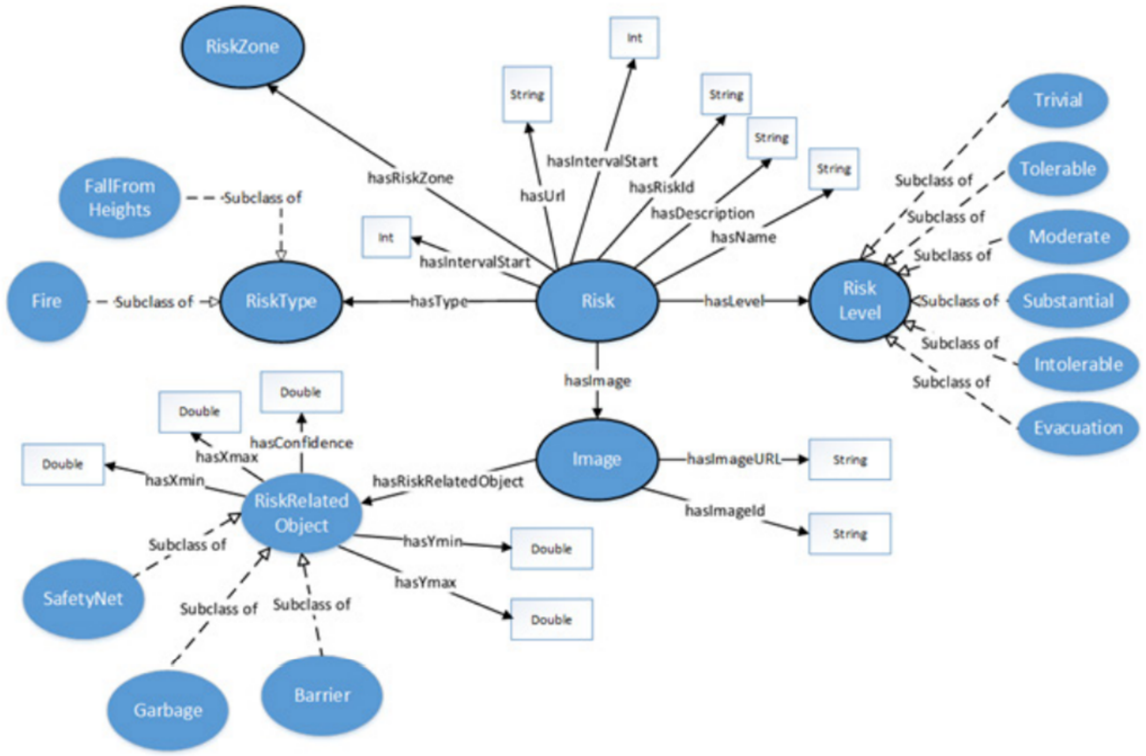


Figure 28: Risk ontology

The source code implementing the aforementioned modules is publicly available[63].

62 Representational State Transfer

63 https://github.com/superdupercodez/bimprove_rdc

IDP digital twin platform

The company IDP[64] located in Sabadell (Barcelona, SPAIN), develops a full platform of digital twin supporting several disciplines, from BIM and point clouds to real time data. The underlying data structure is a combination of resources and they allow 3D visualization together with sensor values or documents.

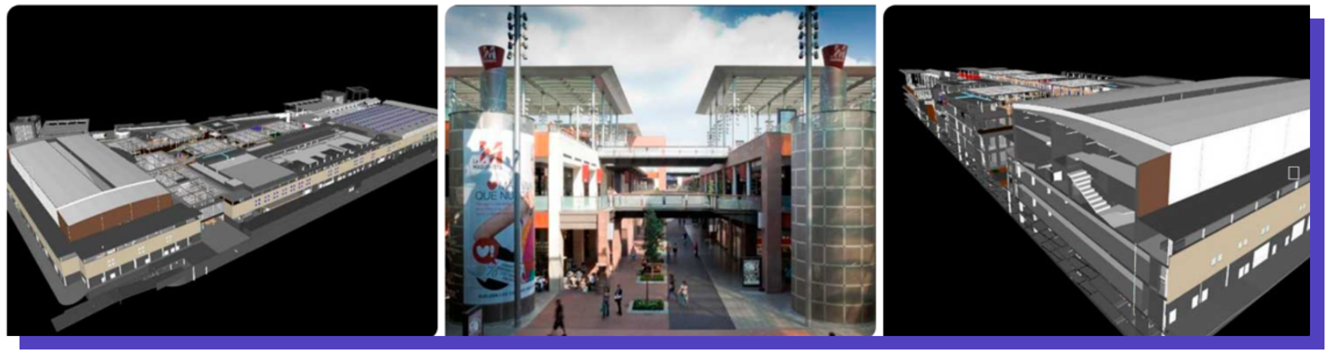


Figure 29: Digital Twin of commercial centre “La Maquinista”, Barcelona (SPAIN)

IDP uses this platform in their services and they are aiming to develop a technological framework ready for the evolution along the life cycle of assets whether a building, a process plant or a city. This vision is citizen centric and warranties ethics and privacy of data generated.

It is prepared to synchronize an updated view of smart buildings and connected with open information to the stakeholders. It allows a circular management of waste and residuals and an optimum energy management.

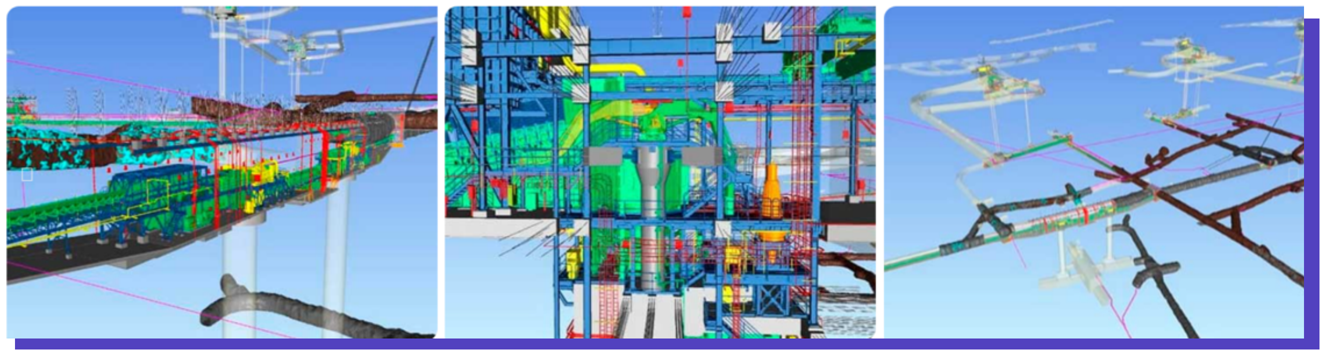


Figure 30: Digital Twin of an underground mine of CODELCO (Chile)

64 https://www.idp.es/idp-digital/

BIM4EEB

The goal of BIM4EEB was to create a toolkit for renovation projects to allow efficient mapping and modelling of a building, and to plan and coordinate the activities of the renovation project. The overall scenario is shown in the figure below:

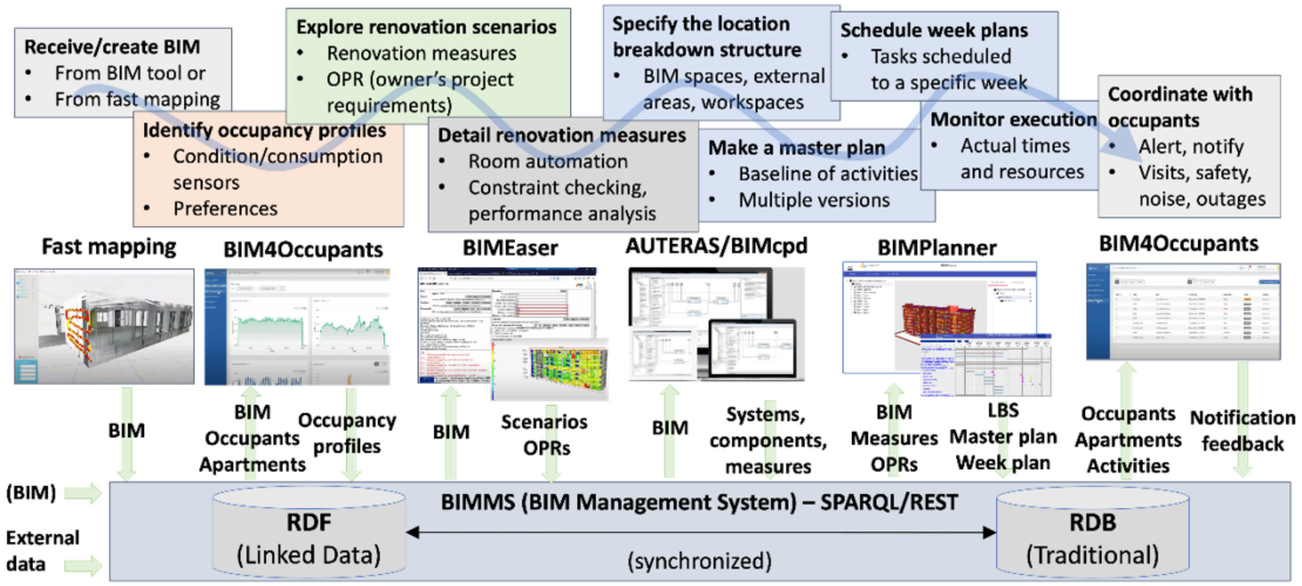


Figure 31: The motivating scenario of BIM4EEB toolkit

The figure presents the tools of BIM4EEB toolkit (in the middle): Fast mapping tool, BIM4Occupants, BIMEaser, AUTERAS, BIMcpd, and BIMPlanner. These tools share information with each other based on the BIMMS data sharing platform (shown at the bottom), that manages both linked data and traditional data (relational data and documents).

Shown at the top of the figure is a sequence of multi-part use cases of the BIM4EEB toolkit. Each of the toolkits produce data and can utilize data produced by the previous use cases. The **tools interact through the shared data**, not directly with each other. The use cases are as follows:

- ***Receive/create BIM***: A BIM model can be received from an authoring tool, or generated using the Fast mapping tool – or parts of an existing model can be refined with the Fast Mapping tool. Fast mapping is based on a combination of laser scanner and a sensor stick – a multisensor device for scanning walls – to detect the overall geometry and hidden structures of the building.
- ***Identify occupant profiles***: Based on sensors installed on the building and questionnaires to participants, the profiles and preferences of occupants can be found out with the BIM4Occupants tool, based on information about BIM models, apartments and occupants.
- ***Explore renovation scenarios***: Using the BIM and occupancy profiles, BIMeaser carries out energy simulations and allows the exploration of different kinds of renovations measures (e.g., insulation, heating alternatives, heat pumps), and combination of those, that is, renovation scenarios. For each renovation scenarios a set of parameters is computed, and they are stored as **Owner’s Project Requirements** (OPR).
- ***Detail renovation measures***: For a selected renovation scenario and OPR, the renovation measures can be detailed and analysed with AUTERAS and BIMcpd tools.

- ***Specify the location breakdown structure***: BIMPlanner allows the interactive definition of work areas to create a location breakdown structure for the project based on the BIM model.
- ***Make a master plan and week plans***: Create the project plans at different planning levels taking advantage of the location breakdown structure. The planning utilizes the activities specified in the selected renovation scenario.
- ***Monitor project execution***: Using BIMPlanner, the status and progress information about the execution is gathered, related to the BIM and operation plans, and compared with planned values.
- ***Coordinate the execution with occupants***: The BIM4Occupants tools is used to coordinate the execution – such as visits, power and water outages, etc. – with the occupants of the apartments. The information from BIM, plans, apartments, and occupants is used.

In the BIM4EEB the data sharing platform (BIMMS) was implemented as well as the different tools of the BIM4EEB toolkit. The crucial parts of the above presented encouraging use case scenarios which were validated in the three demonstration sites at Italy, Poland and Finland.

NEXTSPACE, a visual ontologies management tool

Despite its excellent graphic performance (supported by Nvidia and Unreal Engine partnership), the company Nextspace[65] with headquarters in New Zealand, declares its product as “a data-first approach, because it’s not a choice”. The long experience of CEO Mark Thomas and the team supporting the digital twin tool is a warranty of knowledge, finally producing a data centric tool based on ontology management. They know the way ahead that others are now beginning to take.

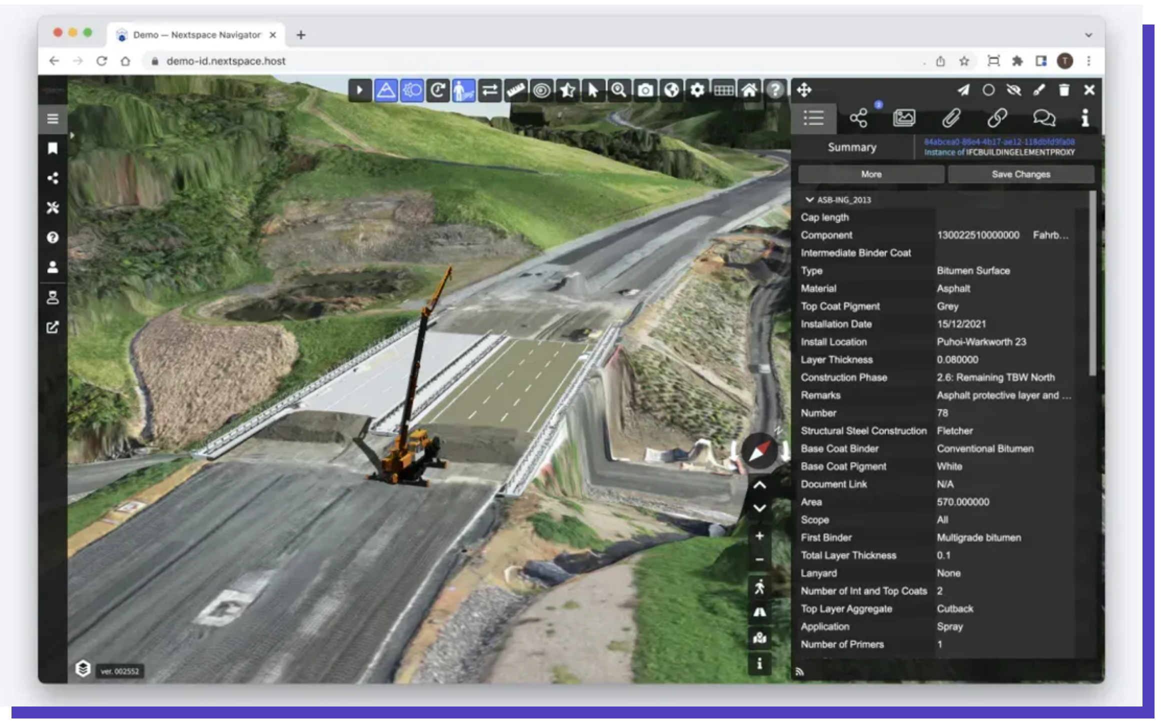


Figure 32: Nextspace interface of a highway digital twin

65 https://www.nextspace.com

Over the next decade, digital twins will become commonplace. This increased adoption is driven by a need to address increasingly complex challenges. Challenges that require humans to better leverage the power of data and computers. Data-first twins will easily cope with increasing complexity. The unique data architecture of Nextspace can connect a given digital twin to a network of others. The same architecture is capable of integrating emerging technologies, like AI, that will become commonplace too.

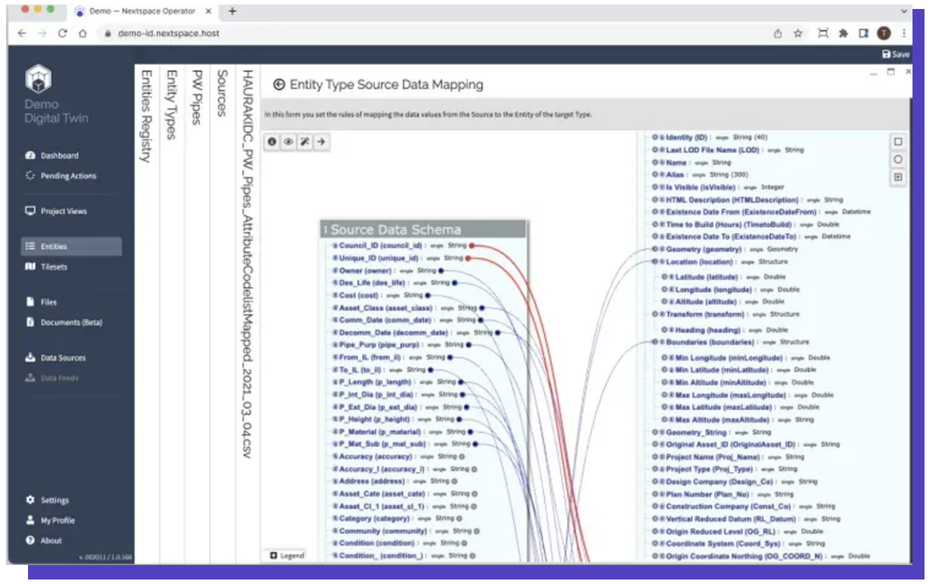


Figure 33: A graphic management of ontologies makes much easier complex data implementations

The real-world challenges you’re trying to solve are rarely simple enough to solve with just a single set of data. Nextspace treats all of your data equally—regardless of where it comes from, or the format it’s stored in. You can import graphics, relationships and data including IFC from ArchiCAD, Tekla, Bentley, Revit, and more. More formats include Revit, AutoCAD Plant 3D, KML and KMZ, JSON and GeoJSON, TIF and GEOTIFF, LAS and LAZ point clouds, DWG and DXF, and Navisworks. Dynamic source data can be connected as well: link sources from Esri, SAP, IBM Maximo, and SharePoint. Link IoT, SCADA, HMI, and sensor data. It is possible to import point cloud, LIDAR, and photogrammetry tile sets, documents and multimedia.

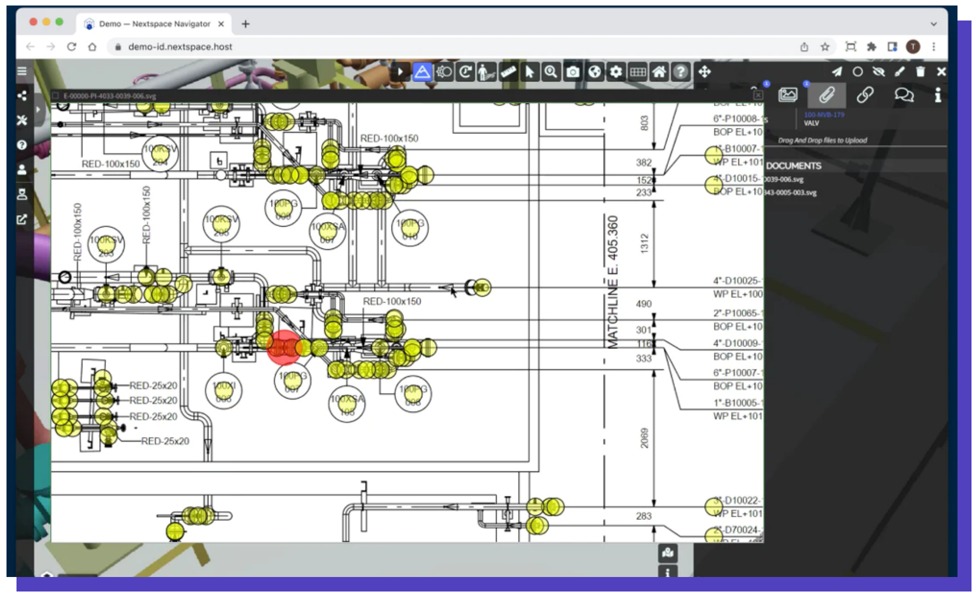


Figure 34: Hotspots can be inserted to link specific items to notes or documents

Solving any challenge starts by defining the right problem. Nextspace platform organizes data attributes based on the question of what matters most to your challenge. GUID-based ontology is the basis for data adaptability—and it is this ability to absorb change effortlessly which future-proofs your universal data schema, and allows you to connect with independent digital twins. There's no need to agree to any data standards before you start a project on our platform. Instead, you can adapt, evolve, and even abandon standards. You can start creating a digital twin on Nextspace without first having to "fix" any existing databases or data silos. Begin with incomplete data, and layer new data as it becomes available. There's no reason to wait, and many reasons to get started.

Within the Nextspace platform, every component is an entity. And every entity has its own unique ID. The Ontology describes the relationships that exist between each and every GUID, or component, in your digital twin. That relationship can be set up graphically.

Figure 35: 3D information can be combined with realtime data from sensors

COGITO

The COGITO H2020 project aims to describe a real-time digital representation (twin) of a construction project, **to monitor real construction** projects from their design to implementation stages. That fuses as-designed multidimensional Building Information Model information, with:

- as-is BIM information,
- live IoT and visual data,
- the (machine or human) project management decisions taken using that data.

One of the main objectives when developing a digital twin in the field of construction is to monitor all aspects related to the development of tasks in the work site. The constant digital monitoring supported by simulation gives the possibility to prevent any contingency that may occur in the construction site.

Specifically, the digital twin model for execution phase must consider the following aspects:

- Lifecycle: The digital twin model encompasses part or the entire life cycle of the construction project with the primary objective of optimising time, cost, quality, and workers' safety.
- Monitoring: The digital twin model is connected to the physical environment through an underlying sensor network, selecting and filtering the data required for daily operational management. These data are fed into decision-making by humans or Artificial Intelligence techniques.
- Simulation: The digital twin model supports simulating all the interactions among the different elements and actors of the construction site in a digital environment. Therefore, before on-site implementation, the digital twin is the virtual test bench for any construction process.
- Prediction: The digital twin model captures events from the physical construction site and enables an intelligent and adaptive construction process, predicting and issuing alerts related to the different physical assets.
- Optimisation: The digital twin model provides a holistic approach to all the construction project parameters and identifies the actions/decisions to achieve the best possible project outcomes.

Therefore, developing a digital twin in the construction domain helps streamline project delivery and achieve optimal control and intelligent management.

In this context, a bundle of novel services can facilitate the: (i) timely detection of health and safety hazards to humans, (ii) timely identification of quality defects, and (iii) provide means for real-time on-site workflow management.

Work must be done in different phases of the construction project lifecycle to mitigate health and safety related issues at construction sites. In the planning phase, potential construction site hazards and necessary protective equipment must be identified, considering the Building Information Model and its 3D geometric elements, by defining appropriate construction spatial zones in the 3D BIM model.

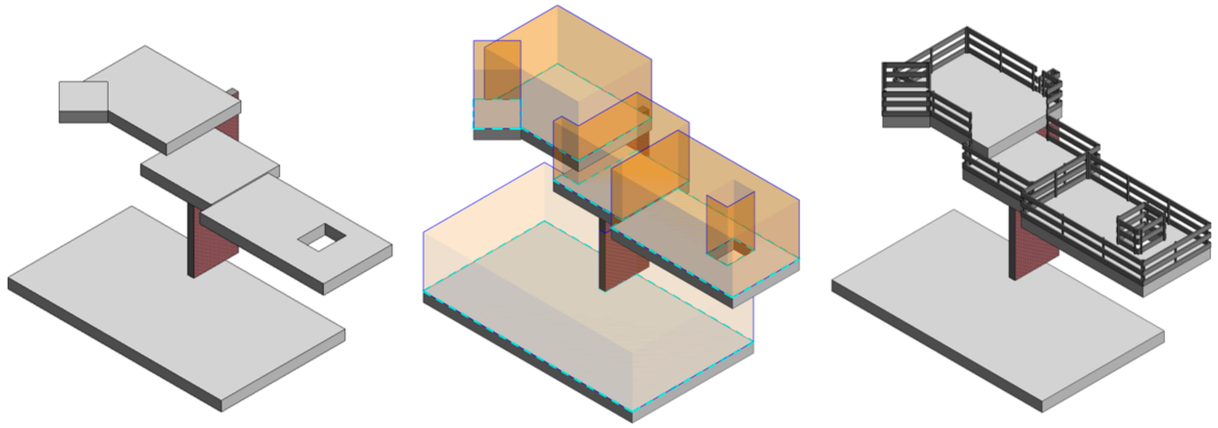


Figure 36: Geometric analysis of safety prevention

It is essential to detect rapidly unfolding risks on the construction site during the construction execution phase, using location tracking data to prevent potentially hazardous situations. This allows to predict potentially dangerous activities and issue personalised alerts.

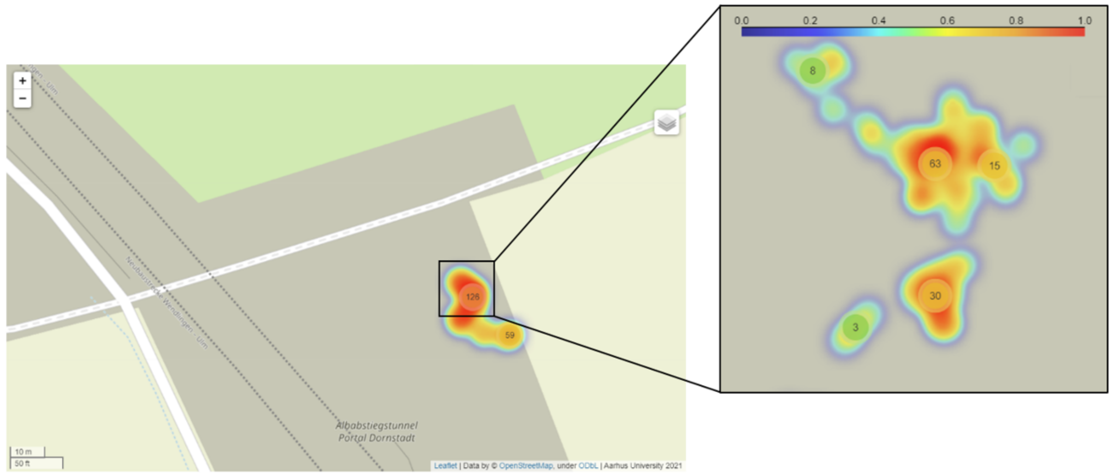


Figure 37. Identification of close proximity events between heavy machinery

Furthermore, during the whole life cycle, it is of interest to provide visual and interactive training materials (e.g., using virtual reality), taking into consideration worker-specific aspects and, if possible, the real construction environment.

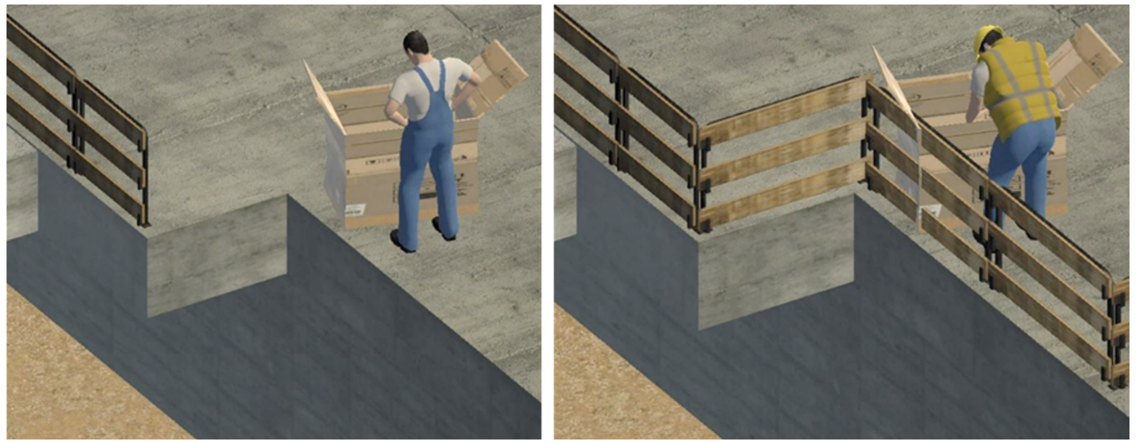


Figure 38: Safety training scenario

Quality control in construction sites can be approached from different perspectives. One of them is to ensure that geometric specifications are met during the construction process. To do so requires defining during the planning stage where geometric specifications apply in the project (by analysing the BIM model). Then, during construction execution, 3D data captured from the construction site is analysed and compared to information contained in the digital twin (more specifically the BIM model) to recognise the BIM components and then assess whether their attached geometric specifications are met.

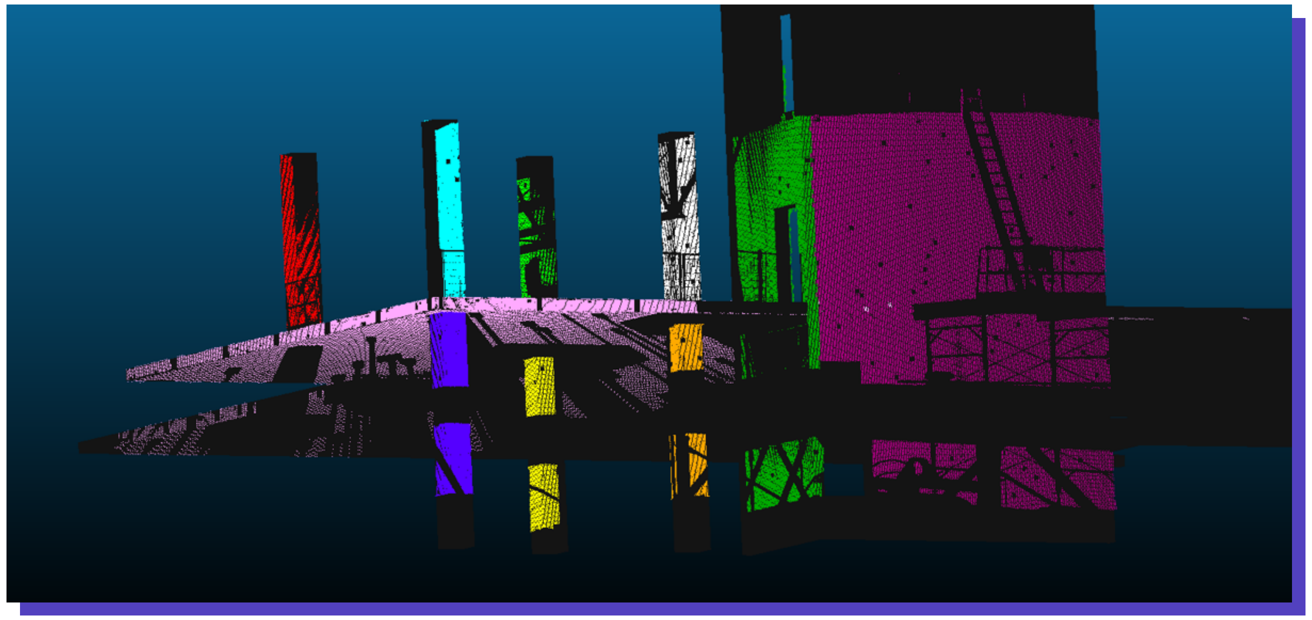


Figure 39: Segmented construction components after voxel space point matching

Another perspective is automatically detecting and classifying common visual construction defects in 2D images acquired on site (either manually by a worker or periodically). Artificial Intelligence techniques (e.g., deep learning) can automate detection with minimal need to manually confirm the detected defects. All this visual quality information is included in the digital twin to support other tasks in the construction process.

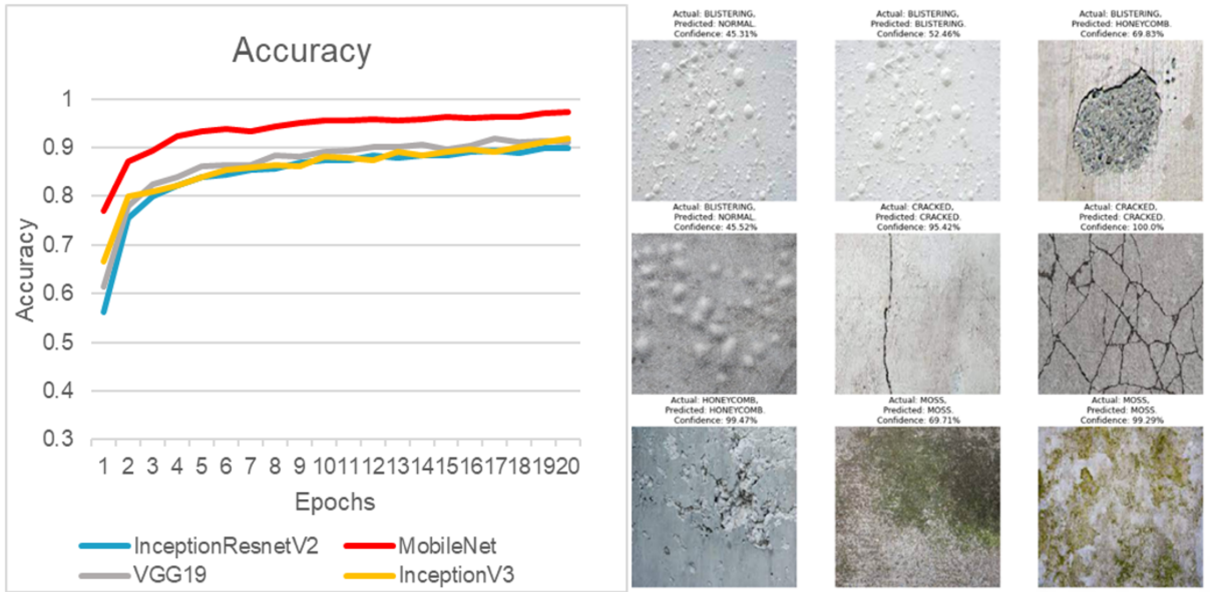


Figure 40: Sample defect predictions

On-site visual data capturing based on Augmented Reality enables the live visualisation of the geometric, visual, and safety issues. The relevant stakeholders will be able to visualise the detected issues to confirm them and propose remedial tasks in the workflow.

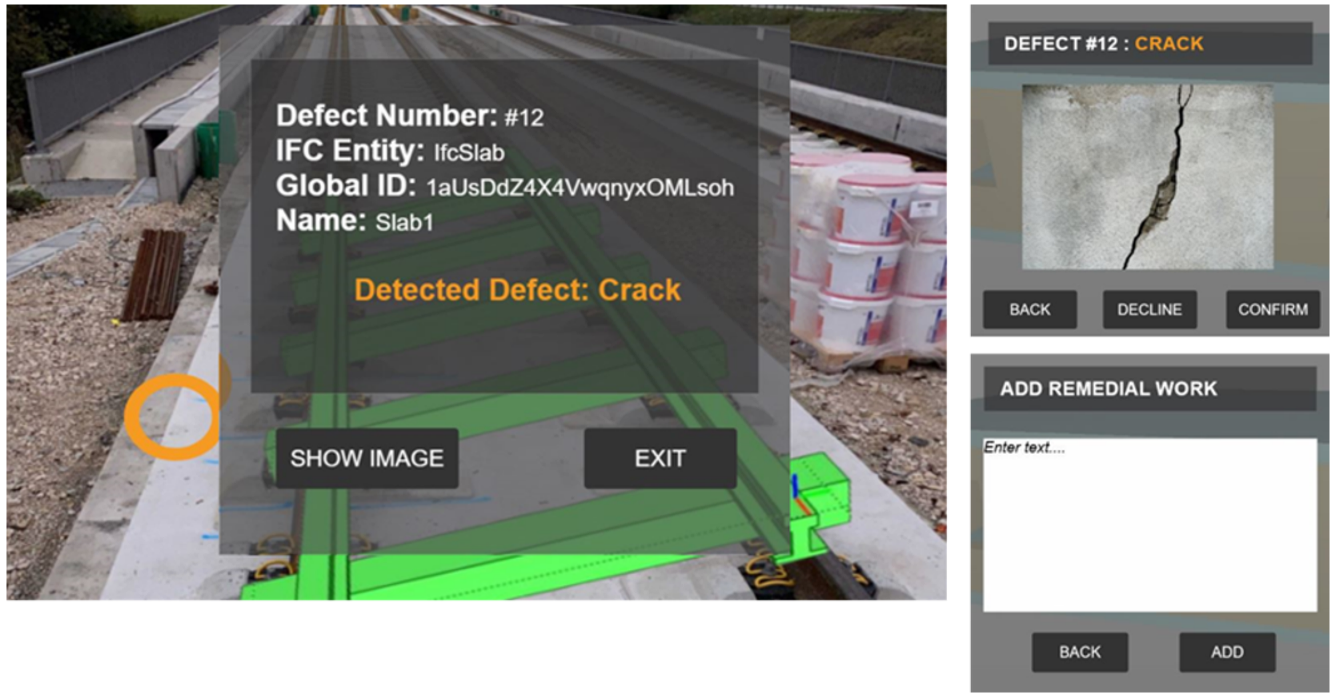


Figure 41: On-site defect visualisation, defect confirmation, and addition of remedial work

Finally, it is needed to define an environment where the construction process and workflow can be modelled and incrementally refined through simulation and optimisation mechanisms over-scheduled data and resource allocation.

Such an environment will allow managers or forepersons to keep track of the entire process and adapt the workflow in case of planned or unforeseen changes during the construction. In addition, it will provide on-site assistance to workers by providing them with the necessary information and instructions of the process, reporting work progress and issues, and receiving updates from the construction manager or foreperson.

NEANEX material passports and portal

The concept of material passports for constructions is gaining more and more traction in the AEC industry to reach evermore important circularity targets. At the same time, organizations are wondering how to efficiently create and exchange “digitally sustainable” material passports in practice.

Organizations are making circularity in the construction sector a top priority to reduce the pressure on natural resources, limit waste and prevent climate change by more efficient reuse of construction materials. Material passports for constructions are perceived as an important tool to reach this goal. A so-called material passport can be understood as a specific part of a digital asset register or construction logbook, i.e. a set of interlinked datasets to capture the value of the assets and materials across

construction lifecycles, projects, and time, allowing asset owners and their affiliated parties to meet governmental or self-imposed circularity targets. The contractors will also be required to deliver the starting point of the digital asset register to the asset owner, i.e. a dataset reflecting the construction they have built, compliant to the contractual information requirements of their client. The availability of good quality and up to date asset data to the asset owner’s team is a prerequisite for good decision making on both maintenance and reuse of construction assets and their materials.

To share and re-use asset and materials data in a “digitally sustainable” manner – i.e. independent from time, tools, organizations and software vendors – information exchange flows based on mature and widely applied standards is required. In this short Use Case, we dive deeper into the subject by demonstrating the application of standardized Linked Data (LD) technologies.

Application of Linked Data for the creation and handover of material passports

During the construction project, conceptual object types such as “Rail” (see Fig. 42) are instantiated as individual objects in the construction project data environment managed by the contractor.

Throughout the project, the contractor’s team collects and combines an entire array of distinct datasets, including data from manufacturers, as-built geometry from CAD/BIM authoring tools,
alphanumeric data delivered by the work planner, etc.

All these individual objects should be classified to one of the OTL object types shared by the client. In our example project (see upper part of Fig. 42), the individual “Rail 2” is classified to the “Rail” object type. As a result of the classification, a list of required OTL [66] aspects are presented to the user. By adding values for these aspects as an enrichment step during the course of the project, the information requirements of the client are fulfilled. As the original information requirements are structured in a standardized manner using Linked Data, a reliable and repeatable validation process over the “under construction” asset data register can be executed on the side of the contractor.

At the end of the construction process (see lower part of Fig. 42), the contractor does a handover to the client of the as-built datasets conforming to his/her initial requirements. By relying again on the Linked Data standards, this time to construct a dataset, the data is shared again in an entirely vendor-neutral manner.

66 An object type library (OTL) is a library with standardised object-types names (e.g. road, viaduct) and properties or specifications. An object is described with its object-type data, geometry data and metadata, Metadata are data (or information) about the data of objects. Metadata are needed because each object type has its own properties. How the object types are grouped is called an ontology. The OTL can be linked to a data dictionary, with the definitions of object-types.’
Within an OTL, assets are described with the standardised language, syntax and semantics required for a reliable information exchange.

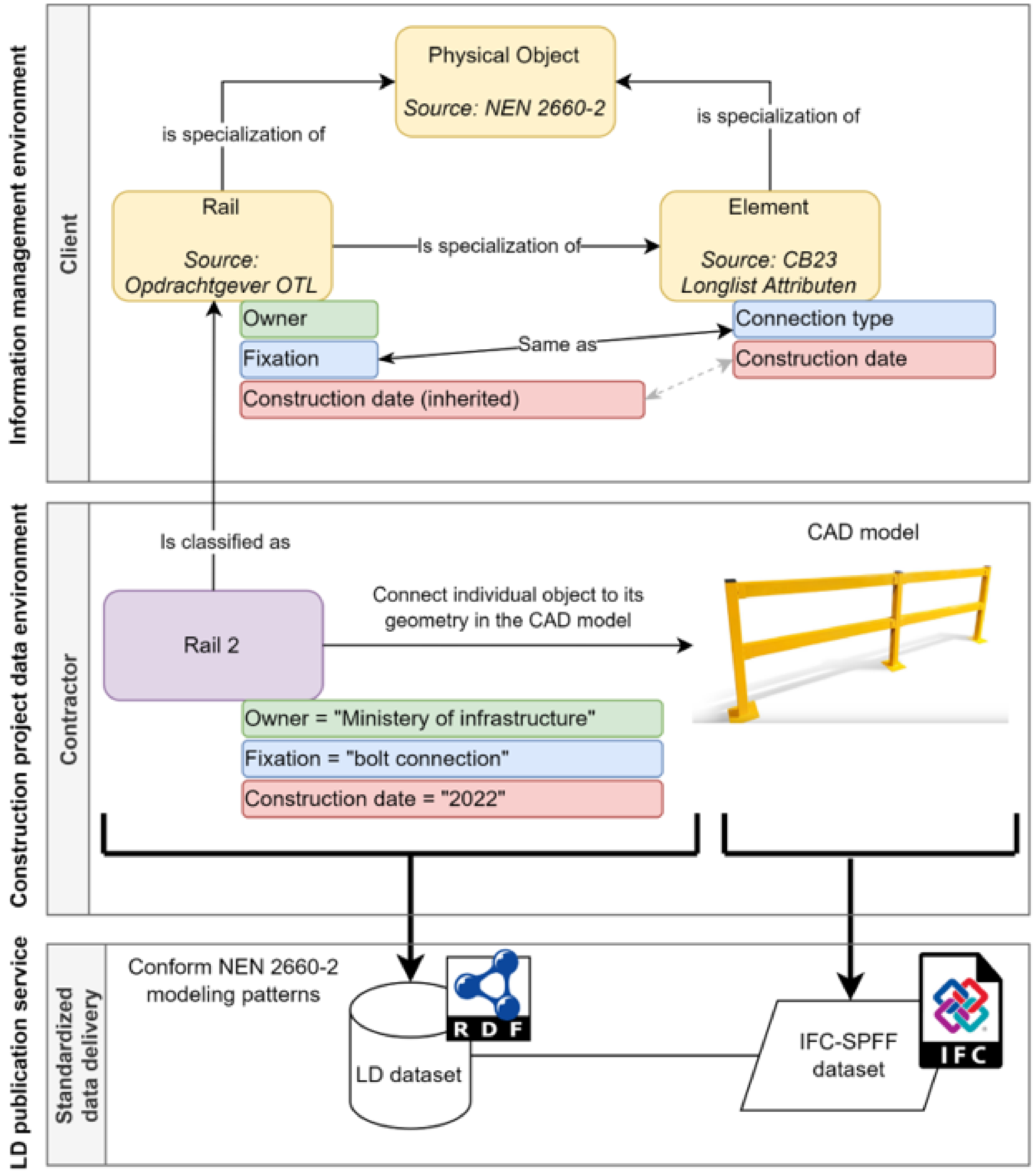


Figure 42: Producing an “as built” dataset and publishing it as Linked Data

The Neanex Portal, depicted in Fig. 43, is a web application suitable to apply the “combine-enrich-handover” principle. This construction project data environment offers among others a user-friendly plugin to collect data in a semi-automated workflow from various CAD/BIM standards to link documents to individual objects as well as features for the bulk import of alphanumeric data. The combined data, including links to geometry elements, can be further enriched and viewed with user provided aspects and relations between individual objects. As a result, it becomes a breeze to validate

datasets against the structured information requirements and to export data (including IFC) following the Linked Data standards, without the end-users needing to know the “ins and outs'' of the underlying Linked Data technologies.

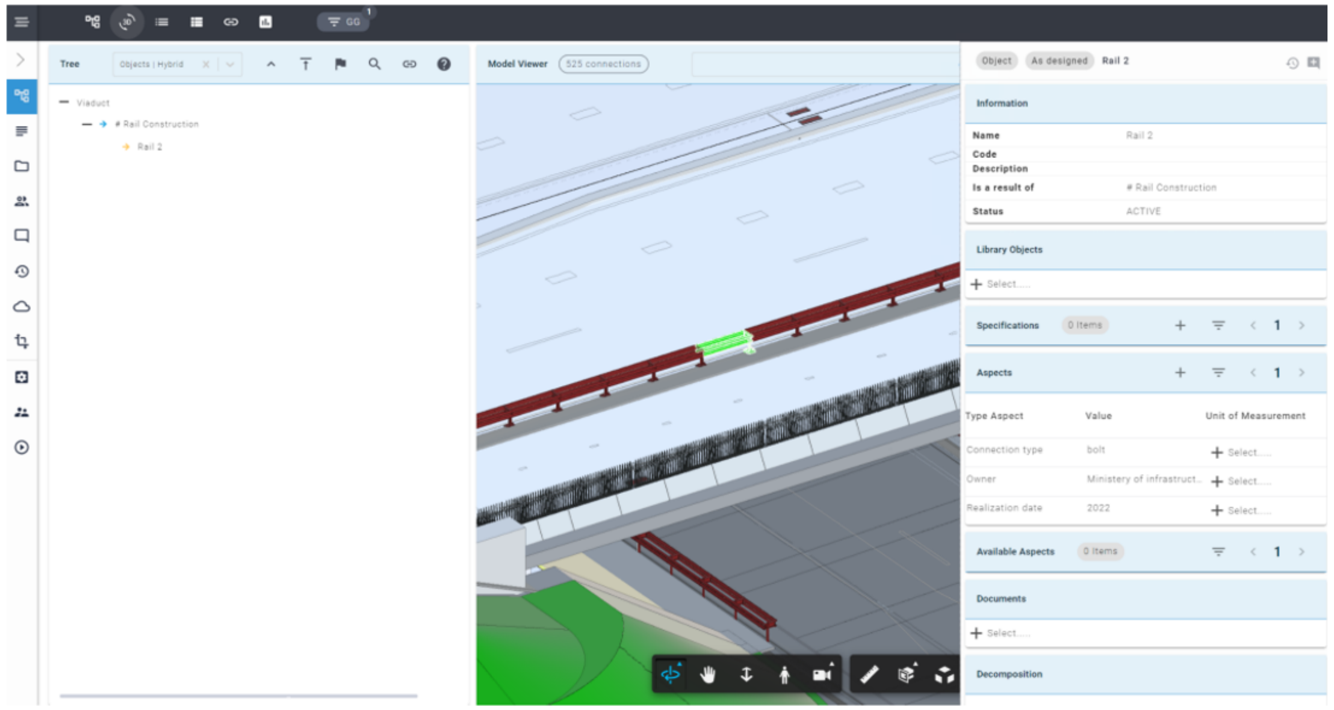


Figure 43: Producing an “as built” dataset and publishing it as Linked Data

Validating and applying the digital asset register

The final delivery of the “as built” digital asset register of the new bridge, including the necessary content of the material passport, can be compared by the client with generic tools against the initial information requirements that reference his/her own OTL and external information requirements from the Client. By doing so, all parties are assured that all necessary information is included in the data delivery to support current and future circularity targets.

From this point on, the asset register representing the bridge will have to be maintained together with the construction, to support tasks at a later point in time. Typically, the received datasets are assembled and loaded in one or more asset management systems. Since the dataset from the contractor was delivered using well-established Linked Data standards, the client (now the asset owner) can share the relevant parts of the asset register in a vendor-neutral manner with parties of his/her interest, e.g. to arrange maintenance tasks or to plan future projects.

5.-CONCLUSIONS. WHERE TO GO

We see today an **imperative use of IFC**, which is key in the design process, but with a huge **lack of design methodology and good practices**. There are new common data environments but many of them are reproducing the Ifc structure in the cloud, with many problems to integrate other data structures. Digital twins, supported with ontologies behind, will be a real improvement in data management.

The onto global map for building. DICO (0.3) and DICon (0.5) ontology.

While a digital solution of one of the subcontractors in a building project can mean great advances for individual tasks, the improvement of construction productivity depends also on what happens between the tasks: how the information produced in one task can and will be utilized by other tasks. This requires agreements and implementation, but before that can be done, there needs to be shared understanding of interoperability at technical, syntactic, and semantic levels. Large volumes of heterogeneous data will be produced at each lifecycle stage. It is a significant challenge to process and interrelate different pieces of data into a meaningful and accurate overall picture that is operationally useful and understandable to both human and automated agents. This has been already carried out by the DICO ontology.

The purpose of the **Digital Construction Ontology Suite**[67] is to address the semantic level of this challenge, by providing the essential concepts and properties of construction and renovation projects, thus paving the way to the ultimate integration of information from different decentralized sources over construction lifecycle.

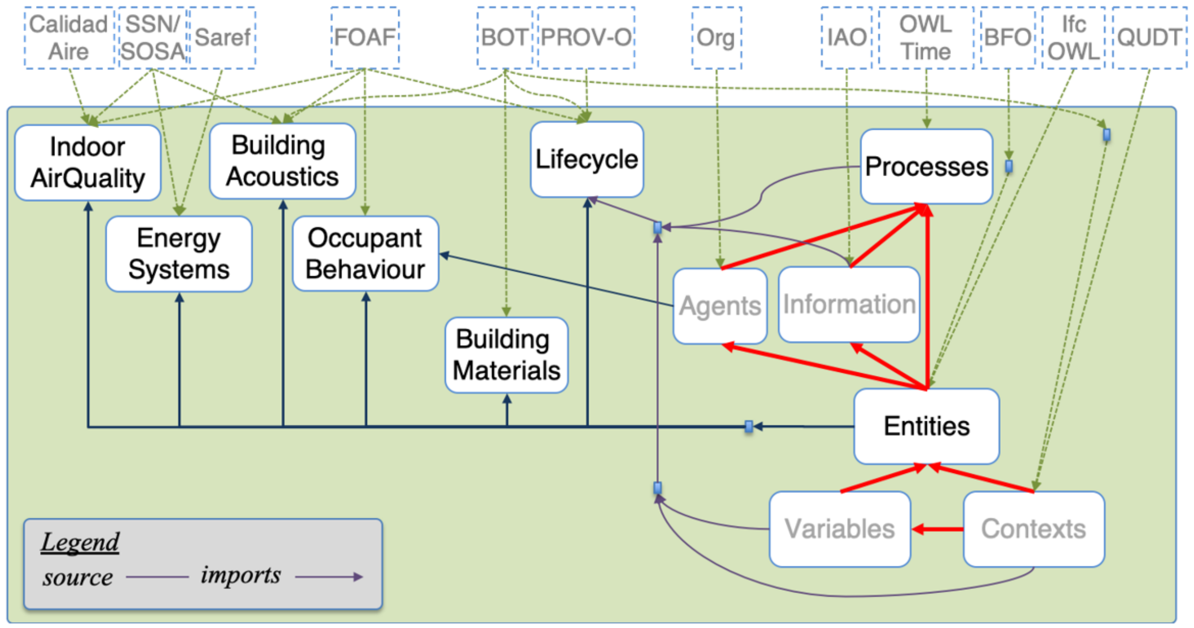


Figure 44: Digital construction ontology DICO 0.3, integration with the Semantic Web

67 https://digitalconstruction.github.io/v/0.3/index.html

DICO ontology may be **the most comprehensive effort** to include all the domains involved in construction making use of existing ontologies, considering BFO ISO/IEC 21838-2 standard as a base ontology. Digital construction ontologies aim to capture the relevant objects and properties (relationships and attributes) that can be referred to by people or systems during the management and execution of construction or renovation projects. This includes physical and spatial entities, temporal regions, information contents, agents, activities, and groupings of objects.

The top-level organization of the Digital Construction Ontologies are provided by BFO. It divides the entities into two classes, Occurrent (things taking place in time, such as processes) and Continuant (things taking place in space, such as physical entities or spacial regions). From the perspective of construction management, Activity - a subclass of Process - captures the intentional efforts of an Agent. An Agent can be a Person or an Organization, and can have Capabilities and assume Roles. The construction process is characterized by a set of Information-Content-Entities, such as Designs, Plans, Contracts and Issues.

DICon[68] (formerly the 0.5 version, updated in July 2022) is the continuation of DICO ontology.

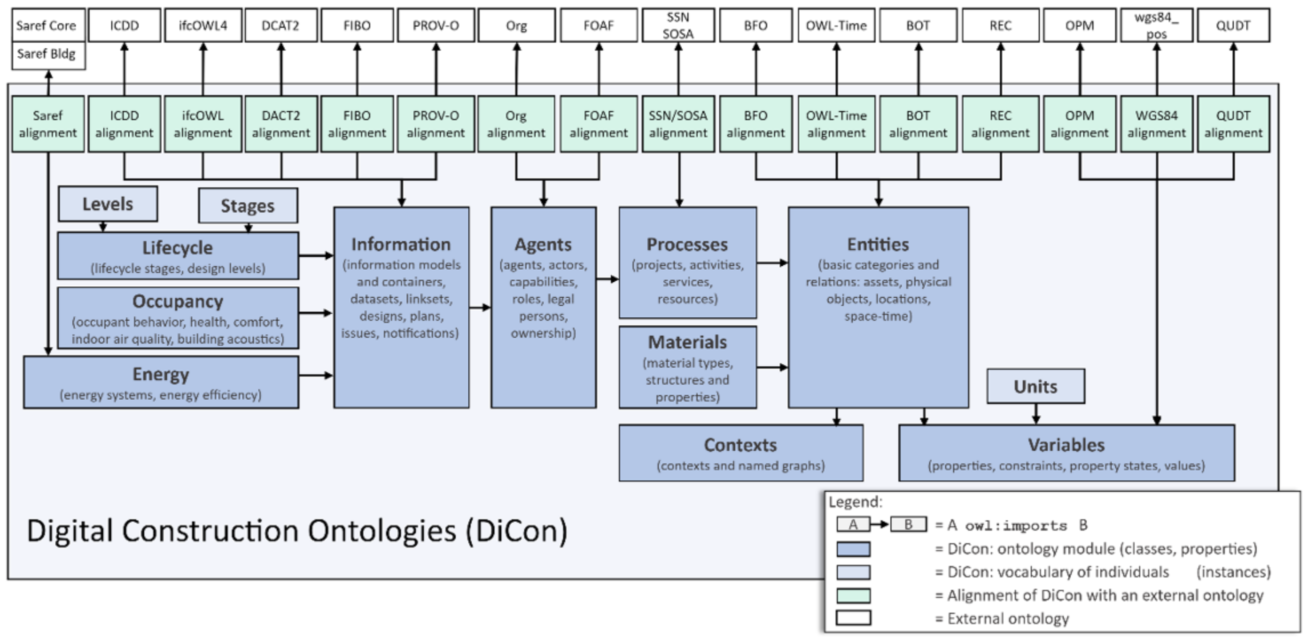


Figure 45: DICon ontology

DiCon is modularized using the vertical and horizontal segmentation approach of the Semantic Sensor Network Ontology. In the vertical dimension, a new module imports the previous one and deepens the representation of the underlying domain by defining additional subclasses, properties, restrictions, or alignments (therefore, an alignment is always means vertical segmentation). In the horizontal dimension, a new module broadens the domain by defining classes complementary to the previous ontology as well as properties to connect them to the previous concepts.

68 https://digitalconstruction.github.io/v/0.5/index.html

In the vertical dimension, the new module should support selected use cases better, having perhaps a narrower user base when compared to the previous module, while horizontal segmentation should extend the set of supported use cases, therefore broadening the potential user base. In the diagram below, the blue rectangles represent the modules of DiCon (the light-blue ones are vocabularies of individuals), green boxes are alignment modules, and the white boxes at top are external ontologies.

While anyone can implement new ontologies, **DICO efforts mark a route towards a common and broad use of ontologies in construction in the future**, and will inspire many new projects.

Normalization and standards, CEN442 WG4

Michel Böhms (TNO) shared in one of the LBD seminars[69] what the state of the art in normalization was at CEN442 WG4 Linked Data. TC442 is about BIM, and WG4 is about "Dictionaries" (but beyond, meaning terms/definitions: actually Data/Information Models or in Linked Data terminology: "Ontologies").

There are major social challenges in our built environment: Housing shortage, outdated infrastructures, energy transition or need for circularity. The right digitalization in project and asset management can help the solution of these challenges through an efficient and future-proof data and software landscape. **Semantic Web Technologies can contribute to deliver a future-proof data landscape**, as a guideline for uniform semantic modelling of assets (included products as 'assets-you-can-buy').

The scope would include all types of assets in the entire built environment, for the entire asset life cycles (and supply chains of these assets), and even for the entire information lifecycle (acquisition/creation, storage, transformation, derivation, integration, decision support/making). Typical use cases are processes of data exchange (or data transfer), data sharing or even existing Linked Data or Semantic Web applications:

1. **Data Exchange:** we can find a non-ideal situation where there are multiple copies (potentially outdated or without synchronization), or a real interface ('change of ownership') where these processes are needed (for example to integrate supply chain).
2. **Data Sharing:** Ideal situation if there is "one copy", and the best situation if there is 'no change of ownership' (for example life cycle management).
3. **LD/SW used for data exchange**: Data collection relies in 'centralized' repositories (in the future more distributed). There are federated queries (for reading and writing). An example would be the solution SOLID/PODS[70] (read-write web).

The guiding principles of this standardization group could be summarized as data fairness (gofair.org) and getting the right quality data depending on the context.

- **Fairness:** Data must be findable (typically in de cloud), accessible (securely accessible via the right identification, authentication & authorisation), interoperable via application of (sets of) open standards, reusable, and well-defined semantics via information models (like ontologies).
- **Right data quality depending on context: this** means data that is relevant, correct, timely, complete, consistent and precise.

69 STATUS OF CEN TC442 / WG4 / TG3 - Semantic Modelling and Linking (SML) standard, the 21st of September 2022

70 https://solidproject.org/

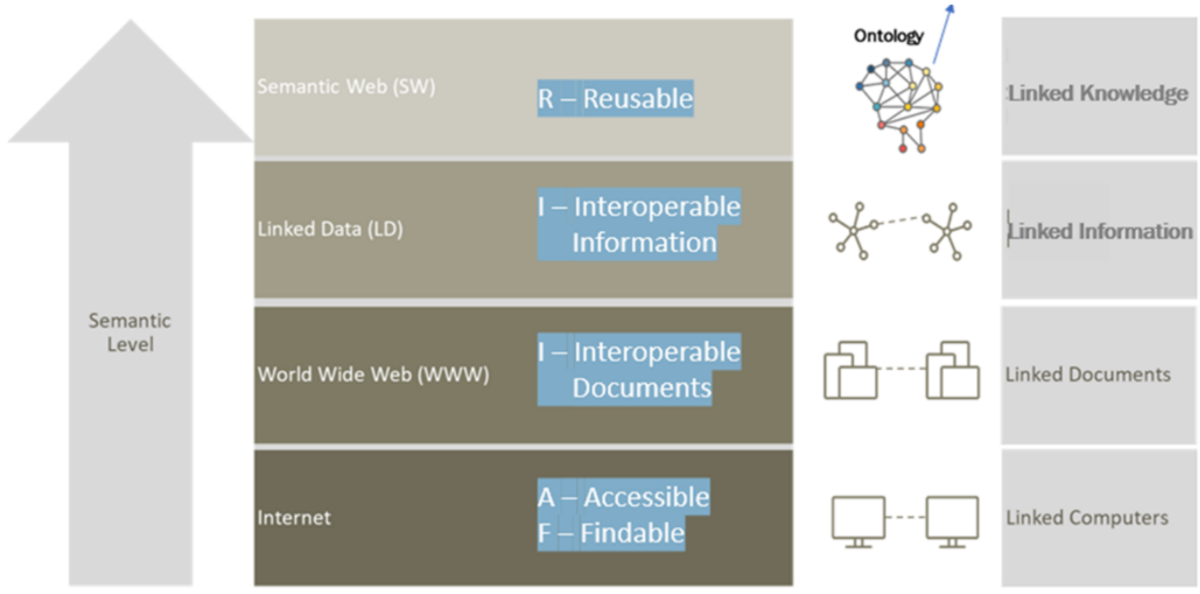


Figure 46: FAIRness powered by Linked data/Semantic Web

We can consider the implementation in three layers of complexity, as shown in the Figure 47.

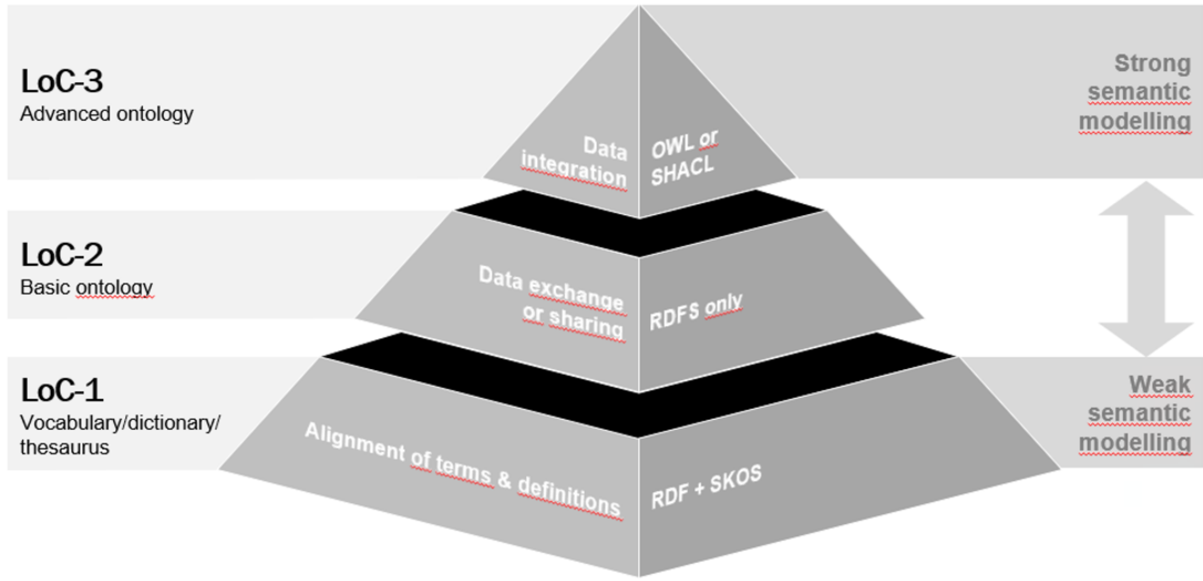


Figure 47: Layers of complexity in semantic modelling

A more complex implementation means a lack of generalized use, and the standardization is aware of this situation. Nevertheless, the technological solution responds and solves a complex problem. As an example, we can mention the **property modelling**. Modelling patterns and property modelling we should be consider as:

- **Simple** (Layer 1):
  - No objectification
  - Unit/Quantity kinds implicit, in name or in datatype
- **Complex** (Layer 2: subject/object)
  - Complex value: QuantityKind (QualityKind/RelationReference)
- **Complicated** (Layer 3)
  - Objectified predicate and value
  - W3C SOSA including extended QUDT usage
  - qudt:QuantityValue
  - qudt:value, qudt:numericValue
  - qudt:unit, qudt:hasQuantityKind

A simple approach could be much better understood and hence of massive use, but the "complicated" (Layer 3) could have a more difficult introduction. But the technical problem itself is not trivial and simple solutions in the long term demonstrate that they are useless at solving real problems. This standardization group has tried to use what is existing, to work with top ontologies and to study layers of complexity of implementation, thinking that the most sophisticated and perfect implementation sometimes means a non-practical approach and too rigid to be used by the industry.

Nevertheless, the implicit complexity of the 'data problem' is clear, and potential solutions and tools existing already available.

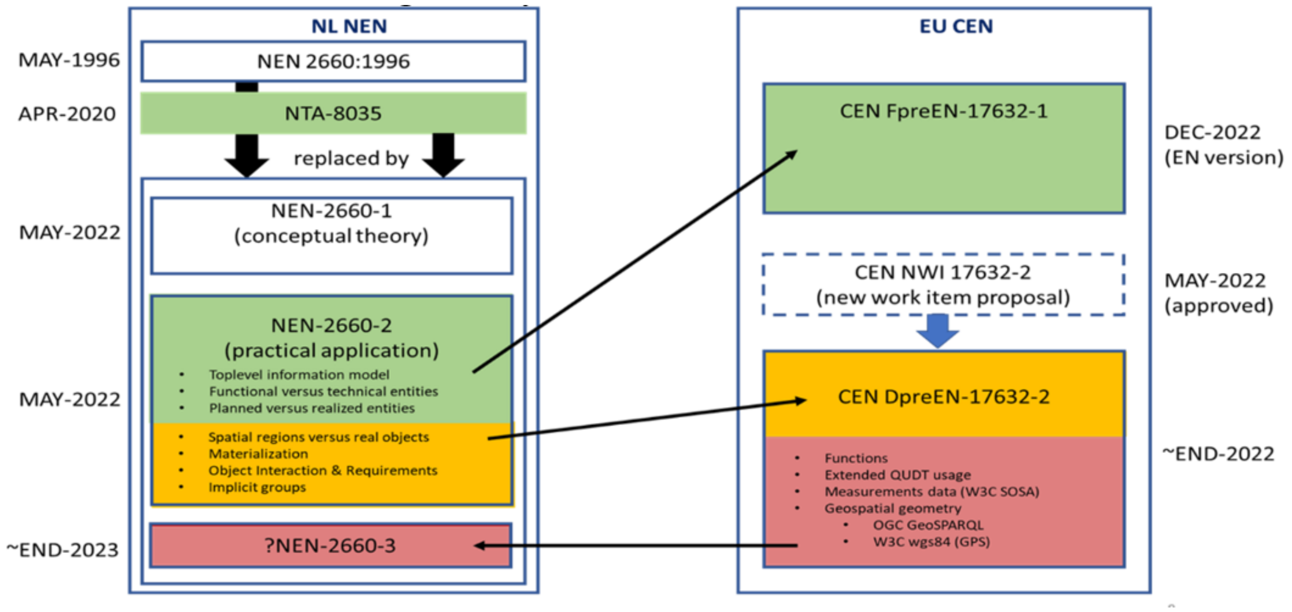


Figure 48: Expected progress of CEN standards for ontologies

The CEN FpreEN-17632-1 is expected at December 2022, and CEN DpreEN-17632-2 at the end of year 2022 (following the figure 48). Some progress in modelling patterns or property modelling should still be expected in the next few years as a result of this effort

6.-APPENDIX: SOFTWARE TOOLS

Many of the mentioned metadata schemas orient towards the use of RDF and OWL. Therefore, most instance data can simply be created and maintained using standard and generic RDF tooling. This typically includes the following types of tools:

- Ontology editors (e.g., Protégé): allow the creation or maintenance of a new metadata schema as an OWL ontology. Such tools are recommended for editing ontologies only, not the actual instance data.

- Triple stores (e.g., Stardog, OntoText GraphDB, Virtuoso): allow to store the instance data in what looks like a database management system (DBMS) for graph data. Such tools are recommended for storing actual instance data, not the OWL ontologies that should be published according to best practices.

- Dedicated software libraries (e.g., RDFLib, OWLAPI, Jena): allow to handle RDF data programmatically (parse, (de)serialise, query, write, create). These tools are crucial in working with the data in a data-driven smart building, as they enable the required level of automation.

- Additionally, for loading IFC to BOT ontology it is available this tool: https://github.com/NIRAS-MHRA/IFC2BOT

7.-ACRONYMS

AEC, Architecture, Engineering and Construction

AECCO, Architecture, Engineering, Construction, Owner Operator

AI, Artificial Intelligence

BDT, Building Digital Twin

BDTA, Building Digital Twin Association[71]

BDTCM, Building Digital Twin Construction Manager

BDTE, Building Digital Twin Environtment

BDTI, Building Digital Twin Instrumentation

BDTM, Building Digital Twin Manager

BDTSM, Building Digital Twin Simulation Manager

BEMS, Building Energy Management System

BFO, Basic Formal Ontology. See chapter 3, Ontologies review

BIF, BIMERR Interoperability Framework

BIM, Building information modeling

BOT, The Building Topology Ontology (BOT) is a minimal ontology for describing the core topological concepts of a building. See chapter 3.1

CERN, European Organization for Nuclear Research. The name CERN is derived from the acronym for the French **Conseil Européen pour la Recherche Nucléaire**

DBMS, Data Base Management System

DBpedia, https://www.dbpedia.org/

DiCon, Digital Construction Ontologies

DT, Digital Twin

DTC, Digital Twin Construction

FM, Facilities Management

gbXML, Green Building XML

GIS, Geographical Information System

GO, The Gene Ontology

GUI, Graphical User Interface

GUID, globally unique identifier

HTML, Hypertext Markup Language, a standardized system for tagging text files to achieve font, colour, graphic, and hyperlink effects on World Wide Web pages.

HTTP, Hypertext Transfer Protocol. It is an application protocol for distributed, collaborative, hypermedia information systems that allows users to communicate data on the World Wide Web

71 https://buildingdigitaltwin.org/

HVAC, Heating, Ventilating and Air Condition

IDD, Input Data Dictionary

IDF, Input Data File

IEA, International Energy Authority

IEA EBC, IEA Energy in Buildings and Communities

IFC, Industry Foundation Classes[72]

IfcOwl, provides a Web Ontology Language (OWL) representation of the Industry Foundation Classes (IFC) schema

INRIA, Institut national de recherche en sciences et technologies du numérique

IoT, Internet of Things

ISO, International Organization for Standardization

JSON, JavaScript Object Notation

LBD, Linked Building Data, https://w3c-lbd-cg.github.io/lbd/

LD, Linked Data

LOD, Level of Development

MEP, Mechanical, Electrical, and Plumbing

MIT, Massachusetts Institute of Technology

NS, Namespaces

OB, Occupant Behaviour

OBO, Open Biomedical Ontology

obXML, Occupant Behaviour XML

OGC, Open Geospatial Consortium

OPR, Owners Project Requirements

OTL, Object Type Library

OWL, The W3C Web Ontology Language (OWL) is a Semantic Web language designed to represent rich and complex knowledge about things, groups of things, and relations between things

PRUBS, Profiling Resident Usage of Building System

REST, stands for Representational State Transfer. REST APIs work by fielding requests for a resource and returning all relevant information about the resource, translated into a format that clients can easily interpret

RDF, stands for Resource Description Framework and is a standard for describing web resources and data interchange, developed and standardized with the World Wide

SAREF, Smart Appliance Reference

72 https://technical.buildingsmart.org/resources/ifcimplementationguidance/

simXML, SimModel XML

SML, Semantic Modelling and Linking standard

SPARQL, declarative programming language and protocol for graph database analytics

SSN, Semantic Sensor Network Ontology

TCP, Technology Collaboration Program

UDI, unique device identifier

URI, Uniform Resource Identifier. It is a character sequence that identifies a logical (abstract) or physical resource -- usually, but not always, connected to the internet.

VAV, Variable Air Volume

W3C, World Wide Web Consortium

W3C LBD-CG, W3C Linked Building Data Community Group

WST, Web Science Trust

WWW, World Wide Web

XML, eXtensible Markup Language

XSD, XML Schema Definition

SPHERE

BIM DIGITAL TWIN PLATFORM

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement Nº 820805.