2025 European Conference on Computing in Construction

CIB W78 Conference on IT in Construction

Porto, Portugal
July 14-17, 2025

DEVELOPMENT OF A BUILT ENVIRONMENT ONTOLOGY LOOKUP SERVICE (BE-OLS) WITH A NEW ONTOLOGY EVALUATION

Arghavan Akbarieh[1], Iryna Osadcha[2], Karim Farghaly[3], Frédéric Bosché[4], and Raido Puust[5]
[1]Eindhoven University of Technology, Information Systems in the Built Environment, Eindhoven, The Netherlands
[2]Faculty of Civil Engineering and Architecture, Kaunas University of Technology, Kaunas, Lithuania
[3]Bartlett School of Sustainable Construction, University College London, London, UK
[4]School of Engineering, University of Edinburgh, Edinburgh, UK
[5]Department of Civil Engineering and Architecture, Tallinn University of Technology, Tallinn, Estonia

Abstract

Recent years have seen a surge in development of formal ontologies within the domain of built environment. Yet, experts like novices face difficulties in locating relevant ontologies, hampering the use of ontologies. Besides, these ontologies demonstrate a range of development maturity. This highlights a gap: the need for a continuously updated repository to help users in discovering, evaluating and (re)using ontologies. This paper presents the Ontology Lookup Service for the Built Environment: BE-OLS, which allows to easily find and consult existing ontologies. It also enables researchers to find gaps or redundancies to help bring some order to the current state of the field.

Introduction, background and motivation

Semantic web technologies and linked data

Semantic Web technologies and Linked Data are a set of standards to enable machines to share and reuse data over the Web (W3C, 2023). They have gained significant traction in the construction industry over the past decade with official standards published on this subject, such as EN 17632- 1:2022 (CEN, 2022). Linked Data -based solutions offer several advantages, including improved semantic and technical interoperability, seamless information exchange between tools and platforms, the ability to link information across domains, and support for logical reasoning (Berners-Lee et al., 2001; Pauwels et al., 2017). These advantages address the challenges posed by the heterogeneous nature of the Built Environment (BE), where traditional monolithic data models have proven inefficient for managing and exchanging the diverse data associated with the construction products and processes (Borrmann et al., 2024).

Ontologies

As part of the Semantic Web stack, formal *ontologies* are explicit, structured, machine-readable, shareable, reusable, and web-compliant seta of vocabularies (classes and relationships) used to represent and transfer knowledge within a domain (Gruber, 1993). Modelled based on the Resource Description Framework (RDF) and Web Ontology Language (OWL), an ontology provides a set of logically interconnected concepts in a universe of discourse. This method also enables the scalable integration of smaller, specialised data models, each designed to focus on specific areas within the built environment. Formal ontologies modelled based on RDF and OWL facilitate Linked Open Data querying across diverse data sources.

Ontology development and recommendations

Like any artifact, ontology development follows a lifecycle. This typically includes the following steps: Defining requirements; Extracting terms and relationships; Conceptualising a basic model; Searching for ontologies and vocabularies to reuse; Selecting their suitable elements; Developing the ontology with appropriate tools and naming conventions; and evaluating it for logical consistency and usability (Radulovic et al., 2015).

An important recommendation by W3C is the alignment of new ontologies with existing ones by reuse or extension to avoid creating unnecessary new and redundant vocabularies (Villazón-Terrazas et al., 2011). Ontology alignment and reusing existing classes and properties reduces the disconnection of individually curated data models, enables Linked Data and minimises development time. Most of the frequently used ontology development methods also recommend reusing terms from existing ontologies (Alobaid et al., 2019; Poveda-Villalón et al., 2022). To promote reusability, they further recommend that ontologies include machine-readable metadata and human-readable online documentation.

Linked Data aims to provide interoperability over heterogenous datasets in a machine-understandable way. To ensure that ontologies and linked datasets over the Web are not only developed for their own sake, but are actually accessible and reuseable (Janowicz et al. 2014), several guidelines were developed and adjusted over the years.

The evolution of best practices for publishing and using Linked Data began with Berners-Lee's "*Linked Open Data 5 Star rating system*" in 2006. A star is granted if the logical sequence is followed: (1) Being available on the internet in any format, (2) being structured and ready for automated reuse, (3) being accessible in an online file format, (4) incorporating all of the above while also being available in RDF format (5) and finally, including metadata for enhanced discoverability and interoperability (Berners-Lee, 2006). Later in 2012, Bernard Vatant criticised that following the above rating system only make 5-star data, which are not as useful if the vocabularies are not well-designed. He proposed the "5-Star Linked Data Vocabulary" to emphasise that a semantically-valuable "good" linked data needs good vocabulary. To this end, Vatant (2012) proposed five principles: (1) Publishing the vocabulary on the Web at a stable Uniform Resource Identifiers (URI), (2) providing natural language documentation, (3) providing labels and descriptions for the documentation, (4) making the whole thing available in a dedicated URI using content negotiation, (5) and reusing concepts by establishing links between vocabularies.

With similar ambitions, Janowicz and colleagues (2014) proposed a "*5-star system for Linked Data vocabulary use*", with stars going from 0 to 5. (0) No web-accessible description of the used vocabularies, (1) Having a dereferenceable human- readable information about the used vocabulary. (2), information is available as machine-readable explicit axiomatization of the vocabulary, (3) having them linked to others, with (4) metadata about the vocabulary available in a dereferenceable and machine-readable form, (5) where these vocabularies are linked to by other vocabularies. In their definition, the 5 star is bestowed when it reflects the external usage and perceived usefulness. It is the only star on which the creator of the vocabulary has limited influence on.

In 2016, "*FAIR Guiding Principles for scientific data management and stewardship*" was published to advocate machine-actionability and support higher data computation (Wilkinson et al., 2016). Adopting the FAIR (Findable, Accessible, Interoperable and Reusable) principles for Linked Data is seen as necessary to improve the development and usage of ontologies for both humans and machines (Alobaid et al., 2019):

- *Findability:* is the first step in reusing data (Wilkinson et al., 2016). Data should be easy to locate by both humans and computers.
- *Accessibility:* Once found, users should be able to access data using standardised, open and free protocols.
- *Interoperability:* Data should integrate with other datasets and systems for analysis and exchange.
- *Reusability:* Data should be well-described and documented so it can be reused in different contexts.

As part of the FAIRsFAIR project, Hugo et al. (2020) published 17 recommendations and a list of 14 best practices for applying FAIR principles to enable semantic interoperability and improve the global FAIRness of semantic artefacts (e.g., ontologies, controlled vocabularies, thesauri, and taxonomies).

Most recently, the Onto4Reuse framework has been proposed by Moreira and colleagues (2024) to facilitate ontology reuse. It introduces a reference architecture for ontology reuse based on four key components: (1) ontology design patterns within an Ontology Pattern Language (OPL), (2) ontology matching based on the Unified Foundational Ontology (UFO), (3) reverse engineering for grounding operational ontologies in a foundational ontology, and (4) compliance with FAIR data principles. They proposed that enriching domain ontologies with foundational ontologies, i.e., high-level or upper ontologies, promotes their role as semantic bridges for aligning domain ontologies. This is consistent with an earlier work by Martins et al. (2023) who proposed the Ontology for Ontology Analysis (O4OA), a reference conceptual model anchored in UFO, to bridge gaps between (meta)ontological requirements and domain-specific conceptualisations to achieve conceptual clarity and terminological alignment.

Ontology discovery

Hugo et al. (2020) stressed the importance of enabling interoperability and making semantic artefacts discoverable through search engines. It must be highlighted that Findability (as promoted through FAIR) is not the same as Discoverability. Even though most ontologies are now published on the Web, they are often discovered through academic outlets, word of mouth, search engines, often with the help of luck. As a result, even high-quality data resources can remain unused, despite having identifiers and rich metadata, simply because their existence is unknown (GO FAIR, 2025). More explicit indexing is thus needed, ideally through repositories and web-based services.

Initial effort to collect existing ontologies were made by listing them in reports published as PDF files (StandICT.eu et al., 2023). Going a step further, several domain-specific ontology registries have been developed. For example, the Ontology Lookup Service (OLS) provides a centralised platform for accessing ontologies within the biomedical domain (Ontology Lookup Service (OLS), 2024). Finally, efforts were made to not only list but also evaluate ontologies, typically with regard to their level of FAIRness. An example is the DBpedia's Archivo with more than 1,800 registered ontologies. It evaluates the minimum viability of ontologies based on retrieval, license and consistency, awarding a "star" for each fulfilled parameter (Frey et al., 2020)

Automated ontology evaluation

In order to organise and automate the evaluation of ontologies, a series of studies developed web-based tools and platforms, each reflecting diverse objectives and evaluation criteria. The OntOlogy Pitfall Scanner! (OOPS!) was developed to evaluate ontologies for potential errors and their level of compliance with the FAIR principles (Poveda-Villalón et al., 2014). It was later extended into the OntOlogy Pitfall Scanner for

FAIR (FOOPS!) that also provides an overall FAIRness score, with detailed insight on compliance with each of the FAIR principles (Garijo et al., 2021).

Archivo's evaluation system assesses ontologies based on retrievability, licensing, and consistency. An ontology with zero stars is considered unusable, as it is neither retrievable nor parseable; reasoning and querying is not possible in such cases. A one- star ontology is automatically retrievable and parseable, but lacks a clear or usable license, negatively impacting its practical applicability. Two-star ontologies include a license statement, but it requires manual inspection or extra coding effort for verification. An additional star is granted if the license statement meets minimal interoperability standards, improving usability. Finally, the highest rating, adding two stars to the existing four, is awarded if the ontology successfully passes a consistency check by a reasoner, ensuring that loading it into an inference system is likely to succeed (Frey et al., 2020).

Another example is Open Biological and Biomedical Ontology Foundry, which employs the OBO principles with 20 distinct points (some similar to FAIR), an experimental OBO score, and metrics to evaluate the biomedical ontologies (Jackson et al., 2021).

Ontologies in the Built Environment

In the context of the built environment, a rapidly growing number of ontologies have recently been developed to meet specific objectives and for data integration and automation. They either define the new and necessary ontological classes and relationships, or extend existing ontologies (Farghaly et al., 2023). In section *Method* and *Results*, we present an initial reasonably exhaustive list of more than 100 ontologies for the built environment.

As highlighted earlier, reusability is an important principle in ontology development (Villazón-Terrazas et al., 2011). However, as our results show (see section Method and Results), ontology reuse within the existing landscape of BE ontologies remains under-practiced. Figure 1 summarises a root cause analysis of this issue. Ontologies in the BE domain have frequently been created in isolation, which makes their discovery and reuse difficult (Farghaly et al., 2023). Despite ontology developers recognising the importance of reusing other ontologies or making their ontologies reuse-friendly, they often opt not to pursue it or only reflect on it a posteriori. For instance, this was observed in the development of the FFDR ontology, where ontology alignment was thought about in the last step, only to realise that this would significantly restructure their data model (Guyo et al., 2023). This example highlights the issue of ontology silos and isolation, where ontology development is not the main goal, but a step towards a bigger goal that overshadows the aspect of usage of the ontology by others beyond the scope of the project.

Furthermore, ontologies are no longer used only by ontology engineers and experts, but also by less experienced users in various domains. However, since ontologies are scattered across the Web and often lack proper documentation, discovering, navigating, and assessing their quality and potential for reuse is not a trivial task.

There is, thus, a clear need to stop the current unmanaged development of ontologies in the BE domain, evaluate their quality, enhance their compliance with FAIR principles, rationalise and facilitate their reuse, and make them overall more discoverable.

As an earlier attempt to systematise ontologies related to the BE, the study of Poveda-Villalón et al. (2014) created SmartCity.LinkedData.es, an ontology catalogue for smart cities and related sub-domains (e.g., energy, climate, buildings), incorporating ontology evaluation features based on the earlier-described Linked Open Data 5-Star rating system. Further, some volunteer-driven initiatives aimed to collect and compare smart building-related ontologies (Tiwari, 2022). However, these initiatives offer snapshots and are often limited in scope, or are too general (e.g., Archivo) and/or are not systematically maintained - e.g., Smartcity.linkeddata.es, are discontinued.

This paper thus reports the initial results of a collective effort by members of the EC3 Modelling & Standards (M&S) Committee to develop an *Ontology Lookup Service for the Built Environment domain, BE-OLS*. The ultimate aim is to develop a repository and service enabling ontologies users and researchers to find and rapidly assess the suitability of existing ontologies for

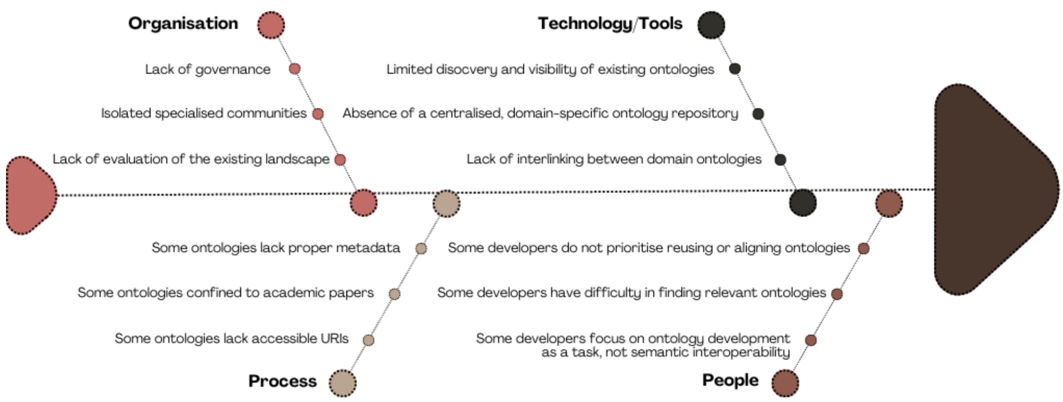

Figure 1: Root cause analysis of lack of reusability of ontologies in the BE domain.

Table 1: BE sub-domains considered for the ontology classification.

Acoustics	Facilities Management	Planning Permission
Air (quality)	Fire Safety	Production (Process)
BE Product (Building)	Geographic Information	Quality
BE Product (Infrastructure)	Geometry	Resources
Circular Economy	IoT Sensors/Actuators	Safety
Comfort	LCA	Structural (Performance)
Cost	Lighting	Water
Energy	Materials	Weather/Climate
Fabrication	Mobility	

their specific use cases in the BE domain. The service formally reports FAIR compliance issues in a more granular and practical-usage -oriented way, the service will hopefully also motivate ontology owners to review and improve compliance. Finally, the service may help identify opportunities for ontology rationalisation as well as ontology gaps.

Method and results

A BE-specific OLS would address many challenges in ontology discoverability, accessibility, interoperability, reusability, practical usability evaluation, and ultimately better use of ontologies in the field. To establish such a domain-specific repository, we followed a multi-step process detailed in the following sub-sections:

1. Ontology collection
2. Database development
3. BE sub-domain ontology classification
4. Ontology evaluation
5. OLS development

Ontology collection

We gathered ontologies through web and Scopus searches, building on the existing extensive knowledge from the authors. All authors contributed independently to that search, which yielded various sources, including academic papers, white papers, and domain- specific searches (e.g., circular economy ontologies for the built environment), whose systematic review sometimes revealed other relevant ontologies.

Database development

As an initial step in our overall project, the collected ontologies were compiled into a structured database table. Each entry was annotated with metadata including publication year and version, BE sub-domain classification, FAIR evaluation score through the FOOPS service, licensing, URI (when available) and connection to upper and domain- specific schemas. Some of this information was collected from ontology sources but BE sub-domain classification and the evaluation required expert and external evaluations done by the M&S committee members.

BE sub-domain classification

For the BE sub-domain classification, the sub-domains listed in Table 1 were established by the team from their knowledge of the BE domain alongside their review of the retrieved ontologies. Each ontology was then given a Primary sub-domain classification and a Secondary sub-domain classification from that list, in order to reflect the fact that some ontologies are often established to provide semantic bridging between two such sub-domains.

Ontology evaluation

We initiated the development of an evaluation framework that complements the FAIR and 5-star systems, with a specific emphasis on the data integrations needs of the Built Environment.

The framework aims to be data- and evidence-driven. During the initial data collection and database compilation, several shortcomings in the available metadata of existing ontologies became evident. Some failed to meet even basic metadata and data quality standards. All these shortcomings were flagged and documented, forming the foundation of our evaluation framework. In the next step, we scrutinised the available web-based ontology catalogues and the way they evaluate ontologies (cf. sections Ontology discovery and Automated ontology evaluation) and observed that their evaluations are based on the FAIR principles. And, while they provide an understanding of the metadata completeness of the ontologies, they do not offer sufficient insights into the practical applicability of ontologies.

Being computationally sound, syntactically valid and algorithmically compliant with FAIR and Linked Data principles ensures that an ontology is well-formed, logically consistent, and machine-processable. Not being fully FAIR compliant does not mean an ontology is irrelevant, and FAIR compliance does not account for alignment or practical usability in its intended context. As such, the evaluation presented in this study differs from FAIR in a few key areas. It does not conflict with FAIR but it (1) has a domain-specific evaluation angle (in our case, the BE domain); (2) and it accounts for ontologies whatever their level of FAIR compliance, even if they do not have a URI.

Our evaluation is conducted along five axes: Connectivity, Accessibility, Documentation & Reuse, Quantity and Robustness. Each axis consists of three evaluation criteria. The evaluation of each criterion relies on a simple binary assessment. If an ontology meets the criterion based on predefined guidelines, it is marked "YES" (1 point); otherwise, it is marked "NO" (0 point). Each axis can thus receive a score between 0 and 3.

- *Connectivity:* Assesses how well the ontology connects to upper ontologies, and other domain ontologies (here BE) and meta-schemas.

- *Accessibility:* Assesses whether the ontology is available as a URI, in a serialised format, and has, at least, a conceptual data model.
- *Documentation & Reuse:* Assesses documentation clarity, use of annotations, and whether the ontology has been reused or extended.
- *Quantity:* Assesses the ontology's size, attribute richness, structural complexity and density.
- *Robustness:* Assesses the validation and verification techniques used in the ontology development, and checks if real-world use cases were used to ensure practical reliability.

To date, we have only completed the development of the first three axes. For example, the three criteria for the Connectivity axis are: alignment with upper ontologies, alignment with other BE ontologies, and alignment with BE domain meta-schemas. Connectivity to upper schemas is just as important (if not more so) as connectivity to domain- specific ontologies, also noted by Moreira et al. (2024). For example, if an ontology incorporates upper ontologies such as FOAF or VANN, it is considered aligned with general Semantic Web standards. Note that we have excluded RDF, RDFS, OWL and XML from the evaluation of this criterion, because these are constituents of any ontology by default. For the BE domain meta-schemas, we have limited our list to the meta-schemas reported by Fierro and Pauwels (2022): Project Haystack (PH), Brick Schema (BRICK), Real Estate Core (REC), Building Topology Ontology (BOT), SAREF for Buildings (S4BLDG), SOSA / SSN, Google Digital Building Ontology (DBO).

OLS development

We built the BE-OLS web platform using HTML, CSS, and JavaScript, which is connected to our dataset. The BE-OLS, shown in Figure 4, is freely accessible at https://cyberbuildlab.github.io/BE-OLS/. The BE-OLS code is publicly available on the GitHub repository: https://github.com/CyberbuildLab/BE-OLS.

Results

Database

Our research led to the identification of 110 BE ontologies. As shown in Figure 2, a first observation is that 22 of the identified 26 sub-domains are the primary focus of at least two ontologies, confirming both the relevance of the selected sub-domains and the broad scope of development of ontologies in the BE domain. On the other hand, 5 of the sub-domains are the primary focus of more than 5 ontologies: BE Product (Building), IoT Sensors/Actuators, Production (process), Energy and Information Management. While, within each sub-domains, the ontologies may be complementary, this could also suggest "over-crowding" with the resulting need for some rationalisation.

Evaluation Framework

Observation and Evidence-based Criteria

During the database compilation step, we noticed that the metadata and level of details vary greatly from one ontology to another. These variations were recorded and discussed to see how impactful they are on usability and discoverability of the ontologies. When assessing publicly accessible URIs, we also noticed a pattern: ~25% of the collected ontologies provide only downloadable serialisations without a persistent URI where the ontology is hosted and explained in human-readable language. This distinction actually resulted in some refinement of our evaluation framework.

Another key observation was the varying levels of alignment between ontologies. Figure 3 shows a network graph presenting the alignment among BE ontologies, with the size of each node (ontology) representing the extent to which that ontology is reused/extended by other BE ontologies. BOT (15), SAREF Core (10), S4BLDG (3), SOSA/SSN (5), SEAS (5), DOT (5), IFC4-ADD2 (4) appear to be the most commonly reused ontologies. As expected (hoped), these include a few of the meta-schemas identified by Fierro and Pauwels (2022), although PH, BRICK, REC, LBD and DBO ontologies are not in that list. SEAS is the Smart Energy Aware Systems – Core ontology, DOT is the Damage Topology Ontology, and IFC4-ADD2 is ifcOWL ontology (IFC4_ADD2). Figure 3 also brings to light that 29 BE ontologies (> 25%) are not aligned with any other BE ontologies. While this could at times be somewhat justified (very niche sub-domains?), this generally suggests inadequate reuse practice in ontology development. These insights reinforced the need for an evaluation framework that not only considers technical completeness or expressiveness, but also explicitly assesses inter-ontology alignment within the broader domain as an aspect of practical adoption and usability.

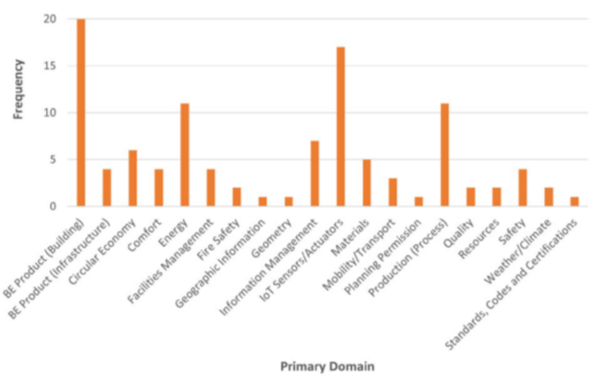

Figure 2: Distribution of the BE ontologies across their primary domains of application.

capture the extent to which an ontology is embedded in the broader network of practice, which makes it more likely to be adopted and extended. Overall, the way an ontology is reused, and how it reuses others, has a cascading impact on the structure and quality of the ontology that depends on it.

Our investigations showed that some ontologies are locked within scientific articles or exist solely as conceptual models, without ever being made available to users in a machine-readable, structured, reusable format. A conceptual model (e.g. diagrams in a PDF or paper) is not machine-readable; hence, not usable in applications. In other instances, ontologies may have a serialisation available, but not hosted at a persistent and dereferenceable URI, making them machine readable but preventing users from accessing human-readable documentation. While having a structured format (e.g. RDF, TTL files) is an improvement over a simple conceptual model, it is not FAIR-compliant without being also accessible via proper metadata or persistent URIs. The accessibility evaluation axis considers these observations to assess whether an ontology has progressed beyond a conceptual model into a reusable format, and minimally available at a dereferenceable URI. We observed two distinct cases of minimal accessibility: (1) Material Passport Ontology (MPO) introduced as a conceptual model, without a serialisation, a controlled case study, verification, or validation (three factors that form the foundation of Robustness in our evaluation framework) (2) In contrast, Building Design Ontology (BIMDO) has a complete, validated, and verified data model, yet, lacking a publicly available URI or serialisation.

We also observed that there were varying levels of machine- and human- readable descriptions across the ontologies. Most ontologies provided human-readable documentation. But there were instances where no human-readable descriptions were available or labels and annotations were entirely missing. In such cases, ontology reuse becomes significantly more challenging. Lack of proper "plain English" documentation makes it difficult for users to interpret, use and reuse an ontology effectively. For example, the Building Circularity Assessment Ontology (BCAO) lacks any accompanying human-readable documentation accessible through a URI. In addition, its serialisation contained no further annotations, such as rdfs:comment or dcterms:description.

BE-OLS Platform

The developed BE-OLS Platform is shown in Figure 4. On the main page (Figure 4a), all ontologies are displayed as cards, each containing key information, including: name, acronym, domain, sub-domain(s), FAIR score and a spider chart visualising the evaluation results across three axes: Connectivity, Accessibility, as well as Documentation & Reuse.

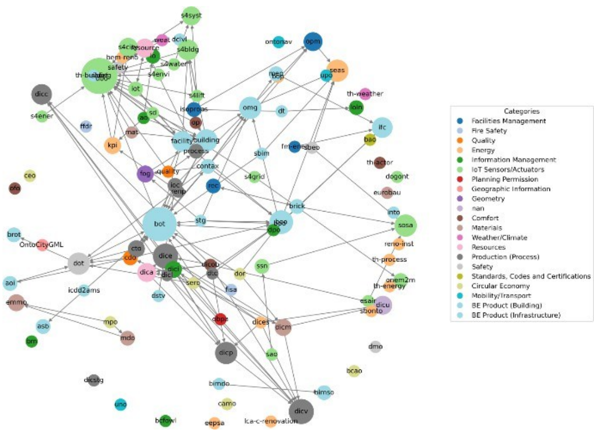

Figure 3: Network graph of the alignment among BE ontologies

Each card includes a button linking to the ontology's dereferenceable URI. Clicking on a card redirects users to a detailed ontology page (Figure 4b), where the user can view all the information contained in the database for that ontology, including a breakdown of how the ontology has been assessed against the evaluation criteria, accompanied by the spider chart. The platform offers two search functions: (1) Search by domain, filtering ontologies based on their primary domain as demonstrated in Figure 2; and (2) General search, enabling users to enter any term or digit in the search box to query the entire database.

Note that we have so far identified five ontology clusters: COGITO, BIMERR, Digital Construction, Brick and SAREF. If an ontology is part of a cluster, this is reported on the individual ontology pages. This part aims to guide the user to see the rest of the ontologies of that cluster/family.

Discussions

The initial purpose of this work was to enhance discoverability and provide BE user -centric insights into the relevance and reliability of an ontology. As a result, we developed a BE-customised lookup service equipped with a customised evaluation system that aims to help stakeholders assess the relevance and reliability of an ontology within its broader domain, in order to determine if it meets their needs. It must be noted that, while the evaluation contributed is shown with focus on the BE domain, it is not domain-specific and can be used in other domains.

The BE-OLS presented in this work is separate from other platforms. For example, Linked Open Vocabularies (LOV) (Dumontier et al., 2017) is a general collection of well-documented vocabularies that partially supports automatic discovery, whereas BE-OLS is a collection of domain-specific ontologies, encompassing ontologies at various levels of development maturity, including those described only in academic articles (which consequently affects their evaluation scores), with the aim of providing a more comprehensive understanding of the ontology landscape in the BE domain. Our evaluation framework

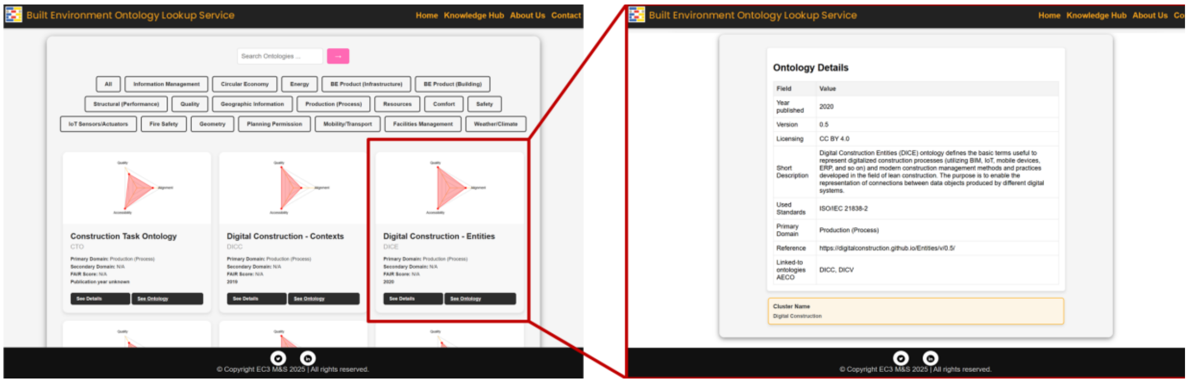

Figure 4: The BE-OLS platform. (a) A snapshot of the main page of the BE OLS cards, filter and search function; (b) A snapshot of the Individual ontology details display

adds to FAIR evaluation tools (e.g. FOOPS) by focusing not only on its technical soundness, but also on the "knowledge" it encodes, its conceptual clarity, and its potential to be understood and reused within the broader domain. Because this knowledge is embedded within the ontology's structure and constrained by its formal constructs, certain aspects not specifically addressed in FAIR, become essential such as evaluating alignment with upper ontologies and other domain ontologies. Indeed, understanding how and to what extent an ontology is aligned with other vocabularies impacts its appeal for reuse. If an ontology developer does not know how an existing ontology works and how it fits in the broader domain, then they cannot reuse it correctly. When users compare ontologies, those with stronger and clearer semantic connections will be preferred over isolated or poorly aligned alternatives. Adequately documented vocabularies (including through plain language English) is naturally also important to ensure that it is understandable and reliable.

Limitations and future steps

Our search for ontologies in the BE domain remains a work in progress, with more ontologies to be added to the dataset. Since BE-OLS is an open project (see GitHub page above), the whole community can in fact (and is warmly invited to) contribute to ensure our collection is complete and our evaluation is robust.

The evaluation framework is also still in the process of refinement with two axes of the evaluation framework also remain to be completed: Quantity and Robustness. Furthermore, the overall ontology evaluation framework and BE-OLS service need to be formally evaluated.

Finally, we plan to further automate metadata extraction, streamline the evaluation process.

Future work could also look into the integration of AI (i.e. LLM) tools to ease ontology searching and analysis.

Conclusion

There is a growing demand for understanding and using ontologies to achieve higher data integration in the built environment. Despite efforts to promote open ontologies through initiatives such as FAIR and Linked Data principles, there is limited insight into the depth and breadth of existing formal ontologies across the heterogeneous domains of the built environment, making their use, reuse and extension challenging. Ontologies' existence, semantic interactions within and beyond the domain, their sub-domain classification, conceptual structure, practical usability and availability are not systematically recorded or reported.

The Built Environment Ontology Lookup Service (BE-OLS) addresses these challenges by providing a structured repository for ontology discovery, alongside a customised evaluation framework that benchmarks ontologies against our Alignment, Accessibility and Quality criteria. By systematically cataloguing and evaluating over 100 ontologies, BE-OLS aims to help practitioners to discover, compare and reuse relevant ontologies on one hand, and to encourage developing reuseable ontologies on the other.

References

Alobaid, A., Garijo, D., Poveda-Villalón, M., Santana-Perez, I., Fernández-Izquierdo, A., & Corcho, O. (2019). Automating ontology engineering support activities with OnToology. Journal of Web Semantics, 57, 100472.

Berners-Lee, T. (2006). Linked Data-Design Issues. https://www.w3.org/DesignIssues/LinkedData.html

Berners-Lee, T., Hendler, J., & Lassila, O. (2001). The Semantic Web. Scientific American, 284(5), 35-43.

Borrmann, A., Schlenger, J., Bus, N., & Sacks, R. (2024). AEC Digital Twin Data—Why Structure Matters. In S. Skatulla & H. Beushausen (Eds.), Advances in Information Technology in Civil and Building Engineering (pp. 651–669). Springer International Publishing.

CEN. (2022). EN 17632-1:2022—Building information modelling (BIM)—Semantic modelling and linking (SML)—Part 1: Generic modelling patterns.

Dumontier, M., Vandenbussche, P.-Y., Atemezing, G. A., Poveda-Villalón, M., & Vatant, B. (2017). Linked Open Vocabularies (LOV): A gateway to reusable semantic vocabularies on the Web. Semant. Web, 8(3), 437–452.

Farghaly, K., Soman, R. K., & Zhou, S. A. (2023). The evolution of ontology in AEC: A two-decade synthesis, application domains, and future directions. Journal of Industrial Information Integration, 36(100519).

Frey, J., Streitmatter, D., Götz, F., Hellmann, S., & Arndt, N. (2020). DBpedia Archivo: A Web-Scale Interface for Ontology Archiving Under Consumer-Oriented Aspects. In Semantic Systems. In the Era of Knowledge Graphs (pp. 19–35). Springer International Publishing.

Garijo, D., Corcho, O. & Poveda-Villalón, M. (2021) FOOPS!: An Ontology Pitfall Scanner for the FAIR principles. In International Semantic Web Conference (ISWC) 2021. CEUR-WS.org.

GO FAIR. (2025). F4: (Meta)data are registered or indexed in a searchable resource. GO FAIR. https://www.go-fair.org/fair-principles/f4-metadata-registered-indexed-searchable-resource/

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

Guyo, E. D., Hartmann, T., & Snyders, S. (2023). An ontology to represent firefighters data requirements during building fire emergencies. Advanced Engineering Informatics, 56, 101992.

Hugo, W., Le Franc, Y., Coen, G., Parland-von Essen, J., & Bonino, L. (2020). D2.5 FAIR Semantics Recommendations Second Iteration (1.0). Zenodo.

Jackson, R., Matentzoglu, N., Overton, J. A., Vita, R., Balhoff, J. P., Buttigieg, P. L., Carbon, S., Courtot, M., Diehl, A. D., Dooley, D. M., Duncan, W. D., Harris, N. L., Haendel, M. A., Lewis, S. E., Natale, D. A., Osumi-Sutherland, D., Ruttenberg, A., Schriml, L. M., Smith, B., … Peters, B. (2021). OBO Foundry in 2021: Operationalizing open data principles to evaluate ontologies. Database.

Janowicz, K., Hitzler, P., Adams, B., Kolas, D., & Vardeman II, C. (2014). Five stars of Linked Data

Martins, B. F., Souza, R. G.-S., Román, J. F. R., Hadad, M., & Pastor, O. (2023). The Ontology for Conceptual Characterization of Ontologies. In Conceptual Modeling. ER 2023, 105–124.

Moreira, J., Donkers, A. J. A., Pauwels, P., Bektas, E., & van Ee, T. (2024). Onto4Reuse: Towards an Ontology Reuse Framework for Knowledge-intensive Software Engineering. In 13th International Workshop on Formal Ontologies Meet Industry. JOWO 2024. Enschede, Netherlands.

Ontology Lookup Service (OLS). (2024). https://www.ebi.ac.uk/ols4

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

Poveda-Villalón, M., Asunción, G.-P., & Suárez-Figueroa, M. C. (2014). OOPS!(Ontology Pitfall Scanner!): An on-line tool for ontology evaluation. International Journal on Semantic Web and Information Systems (IJSWIS), 10(2), 7–34.

Poveda-Villalón, M., Espinoza-Arias, P., Garijo, D., & Corcho, O. (2020). Coming to Terms with FAIR Ontologies. In Knowledge Engineering and Knowledge Management (pp. 255–270). Springer International Publishing.

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.

Poveda-Villalón, M., García-Castro, R., & Gómez-Pérez, A. (2014). Building an ontology catalogue for smart cities (A. Mahdavi, B. Martens, & R. Scherer, Eds.; pp. 829–836). CRC Press.

Radulovic, F., Poveda-Villalón, M., Vila-Suero, D., Rodríguez-Doncel, V., García-Castro, R., & Gómez-Pérez, A. (2015). Guidelines for Linked Data generation and publication: An example in building energy consumption. Automation in Construction, 57, 178–187.

StandICT.eu, Sarkar, A., Frost, L., Walshe, R. & Muscella, S. (2023) Report of TWG Ontologies: Landscape of Ontologies Standards. Zenodo.

Tiwari, S. (2022, August). Existing smart building domain ontologies comparison. https://doi.org/10.48366/r214164

Vatant, B. (2012) in other words: Is your linked data vocabulary 5-star? in other words. Available from: <https://bvatant.blogspot.com/2012/02/is-your-linked-data-vocabulary-5-star_9588.html> [Accessed 31 January 2025].

Villazón-Terrazas, B., Vilches-Blázquez, Luis. M., Corcho, O., & Gómez-Pérez, A. (2011). Methodological Guidelines for Publishing Government Linked Data. In Linking Government Data (pp. 27–49). Springer New York.

Wilkinson, M. D., Dumontier, M., Aalbersberg, Ij. J., Appleton, G., Axton, M., Baak, A., Blomberg, N., Boiten, J.-W., da Silva Santos, L. B., Bourne, P. E., Bouwman, J., Brookes, A. J., Clark, T., Crosas, M., Dillo, I., Dumon, O., Edmunds, S., Evelo, C. T., Finkers, R., … Mons, B. (2016). The FAIR Guiding Principles for scientific data management and stewardship. Scientific Data, 3(1), 160018.

World Wide Web Consortium (W3C). (2023). Linked data. W3C Wiki. https://www.w3.org/wiki/LinkedData