Linked data for the life cycle assessment of built assets

Calin Boje[1], Tomas Navarrete[1], Sylvain Kubicki[1] and Thomas Beach[2]

[1] *Luxembourg Institute of Science and Technology, Esch/Alzette, Grand Duchy of Luxembourg*
[2] *School of Engineering, Cardiff University, Cardiff, United Kingdom*

Abstract

Life Cycle Assessment (LCA) is a scientific method for the quantification of environmental impacts on a product system, which is important for sustainable design and management of our built environment. Conducting LCA on buildings requires access to highly contextualized information which can be sourced from the Building Information Model (BIM) or monitoring systems in place. The interoperability between LCA domain tools and BIM tools is lacking. Our motivation lies in semantically bridging LCA and built environment domains by adopting a Semantic Web (SW) technologies. This would result in increased interoperability on the web, increased automation of information pipelines and more explainable impacts of complex contexts. In this paper we introduce the work in progress under the SemanticLCA ontology where we modelled several use cases for LCA of built assets. To demonstrate this, we showcase one case study at the building level, highlighting the semantic alignments between BIM models, LCA data and sensing devices. The paper discusses the implementation challenges and offers suggestions on how such an ontology can be used in the future.

Keywords

ontology, lca, bim, alignment, linked data, sustainability, built environment

1. Introduction

For the built environment sector to transition from a linear supply chain to a sustainable and circular economy, it needs tools to quantify and measure its impacts, as well as convenient ways to integrate data from multiple application domains. The motivation behind this work lies in bridging the LCA and the built environment domains semantically, using SW technologies across the LCA process. The direct benefits are threefold: (1) increased interoperability between sustainability tools and built environment tools, (2) increased automation for sustainability assessment within the scope of buildings and infrastructure and (3) more explainable impacts assessment given complex contexts. The indirect benefits result in smarter tools and more informed decisions for the building design and operation stages, and potentially reduced emissions because of smart built asset management and design.

In this article we will rely on a standard knowledge engineering approach wherein we formulated several case studies as part of the SemanticLCA[2] project. Additional non-functional requirements on BIM integration with real-time data monitoring systems was also investigated. The formulated ontology models are made available[3], currently still a work in progress. The aim of this paper is to showcase the ontology design process and utility of linking LCA and built environment hybrid data, with a focus on buildings.

The article structure provides background on using LCA and BIM, existing LCA databases and relevant semantic models in section 2. The ontology engineering methodology adopted is outlined in section 3. Section 4 will provide a description of the outputs of the ontology engineering methodology along with a discussion of the potential alignments with existing established ontologies. To demonstrate the use of the modelled ontologies, we simulate its use on several use cases in section 5. Finally, in

---

Proceedings LDAC2023 – 11th Linked Data in Architecture and Construction, June 15–16, 2023, Matera, Italy
EMAIL: calin.boje@list.lu (A. 1); tomas.navarrete@list.lu (A. 2); sylvain.kubicki@list.lu (A. 3); beachth@cardiff.ac.uk (A. 4)
ORCID: 0000-0002-5150-9355 (A. 1);

CEUR Workshop Proceedings (CEUR-WS.org)

[2] https://www.list.lu/fr/recherche/environment/projet/semanticlca/
[3] https://github.com/bojecp/slcao/

Section 6, the ontology is discussed through the prism of available tools, its completeness, and limitations.

2. Related research

The sustainability assessment of buildings and infrastructure is done as part of the design process, under frameworks such as BREAM[4] and LEVEL(s)[5]. This requires specialized expertise and including the aggregation of multiple aspects, covering calculation methods on the use of materials, water, and energy, but also the impacts on human wellbeing. Both LEVEL(s) and BREEAM include LCA as part of the process, while BIM tools and platforms already facilitate some degree of integration with LCA tools [1]. The maturity of the use of these methods and tools highly varies depending on region, design practices and real-estate development habits. Within this article we tackle a multi-disciplinary problem, covering buildings, materials, sensors, human activities, and software systems. Considering the complexity of things involved, a SW approach is a first step towards achieving interoperability between the built environment and sustainability assessment communities. The use of SW is already preponderant across the architectural, engineering and construction industry[2], [3], covering many adjacent application domains.

LCA requires full life cycle data for complete analyses. BIM model-sourced data can be used for design stage LCAs, but operation stage data is usually estimated, and not measured. To compensate for a lack of data post design and construction, we look at various building monitoring paradigms using sensors. This enables dynamic data gathering on energy and water consumption and occupancy use [4], which is useful in gathering operation stage data and monitoring if sustainability practices are followed.

2.1. LCA and BIM integration

LCA is an established scientific methodology, based on several ISO standards (such as ISO 14040[6] and ISO 14044[7]), which is applied on product systems, allowing the quantification of environmental impacts such as emissions of gasses, pollution to water and land, but also effects on human health, land, and resources usage. The built environment encompasses a large range of product systems, from simple construction components to sophisticated buildings, which are considered for sustainability assessment across several standards, such as the ISO 21931[8] series. The assessment of built assets includes natural resource depletion, considers the operation stage energy and water usage, but also how the embedded materials are treated at the end of the life cycle (for demolition waste, or alternatively deconstruction and reuse).

There are numerous LCA tools which are BIM compatible, and the process overall is more automatic and streamlined than in the past [5]. Looking at several industry case studies, [6] noticed 5 distinct integration methods between BIM model information and LCA tools, with the most advanced in terms of data structure being characterized at BIM object level. The process so far is highly fragmented [7], with no common data models [8]. The use of standards such as Industry Foundation Classes (IFC) is also limited, while the BIM model data is lacking for complete LCA, usually resulting in simplified calculations [9].

2.2. LCA domain tools and models

LCA tools which integrate with BIM range from conventional to static and dynamic calculation processes as analyzed by [10]. Across these processes, LCA tools define the assessment process, but can be configured to work with various LCA databases. There are many LCA databases by domain and region, which must be adapted on a by-project basis to characterize each case, and the work by [5] lists

---

[4] https://bregroup.com/products/breeam/

[5] https://environment.ec.europa.eu/topics/circular-economy/levels_en

[6] https://www.iso.org/standard/37456.html

[7] https://www.iso.org/standard/38498.html

[8] https://www.iso.org/standard/71183.html

several construction specific databases per country. Several open-source tools for creating LCA inventories and running the calculations are available, such as the Brightway2[9] suite of libraries in python, or the OpenLCA[10] software. These were designed generically for LCA processes, but not specialized for built asset design, construction of operation, and therefore the definition of the LCA within these contexts has to be defined. While the building and infrastructure domains already benefit from some degree of standardization thanks to formats such as the IFC, LCA domain tools and databases lack a common data standard.

2.3. Existing LCA-related ontologies

The development of the LCA domain into a semantic web ontology has been proposed by [11] for a specific LCA database. A minimal ontology was proposed by [12]. These are preceded by the newer BONSAI[11] open database for carbon foot printing, which is accompanied by an ontology definition. There are several known schema models and ontologies within the built environment which are found useful for LCA data acquisition and integration, which we list in Table 1 as part of the ontology development (detailed in section 4). Common patterns can be discerned from these models, with a very restricted LCA vocabulary around activities, inventories of such activities, flows, emissions, and processes, in line with several LCA databases.

Whilst a focus on LCA concept is highlighted, the process of carrying out an LCA (for a building for example) is not modelled explicitly, where additional concepts are needed to create the context, and link hybrid data. Thus, the scope of the aforementioned ontologies is not adequate for the built environment case. To account for this limitation and use LCA concepts linked to the context of each product system under evaluation (e.g. component LCA, building LCA, city district energy LCA, etc.), we adopted a standard ontology design approach on several applied use cases.

3. Ontology development methodology

The ontology development process of the SemanticLCA project undertook an adapted NeON methodology[13], outlined in several steps in Figure 1. We created a glossary of terms within the fields of LCA (step 1), which leads to the identification of a fundamental LCA vocabulary (step 2). Important vocabularies, schemas, models, and ontologies were also analyzed (step 3), from the LCA field and adjacent domains, as needed by the applied use cases (step 4). The definition of the distinct applied use cases allowed a narrower scope of the application domains. This in effect has caused our ontology to develop in the direction of a network of smaller interconnected ontologies [13]. The use cases from building to city districts, attempt to cover several aspects of the built environment. They include domain specific tools and systems, with different scopes and objectives. For each use case we had several workshops with experts and a series of competency questions (CQ) were formulated (step 5). The ontology itself (step 6) had to accommodate each case study. The focus of the ontology was to replicate a common data schema which could answer CQ from an ontology-supported system. Step 7 looks at iterative testing and validation cycles, which is demonstrated in section 5 of this article.

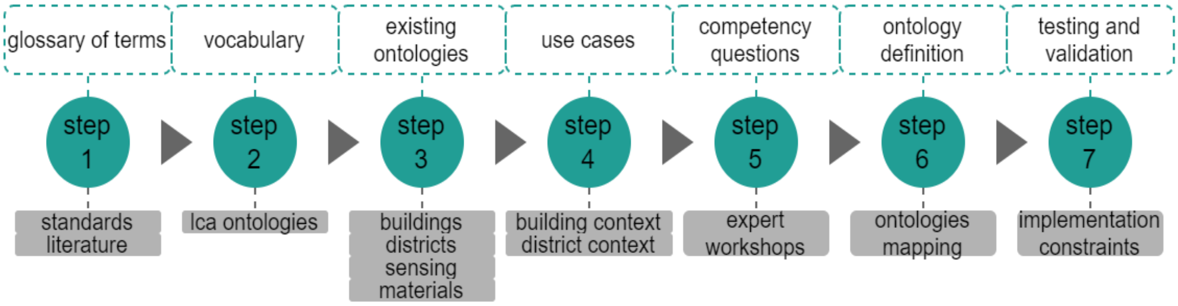


Figure 1: SemanticLCA ontology development methodology steps

---

[9] https://documentation.brightway.dev/en/latest/index.html
[10] https://www.openlca.org/
[11] https://github.com/bonsamurais

4. SemanticLCA Ontology use cases and evolution

Following an initial analysis of LCA tools and databases, fundamental LCA vocabularies and existing ontologies, the overall scope of the SemanticLCA ontology was elicited through an analysis of a series of use cases. All use cases are meant to employ Life Cycle Impact Assessment (LCIA), meaning that certain product systems are analyzed through the lens of LCA. These were proposed by the researchers and developers seeking to build software tools to meet the needs of those use cases, and they are: (1) building materials analysis from BIM models, (2) building operational energy usage, (3) indoor air quality impacts on human health, (4) building energy performance optimization considering human health impacts, (5) energy consumption prediction of city districts (6) extension of district buildings, (7) weighting methodologies for impact results.

The teams working on the use cases were supported to develop a set of CQs, which were reviewed (in isolation for each use case). Once finalized, the CQ[12] were categorized into the ontology domains and any duplicates removed. The initial structure of the SemanticLCA ontology and its overarching domains are shown schematically in Figure 2. The remainder of this section explains the domains and other nearby ontologies which can be mapped to or integrated into the process, with Table 1 presenting the related ontologies analyzed for each domain. Class concepts used in section 5 are presented in more detail here for the building, LCA and sensor domains.

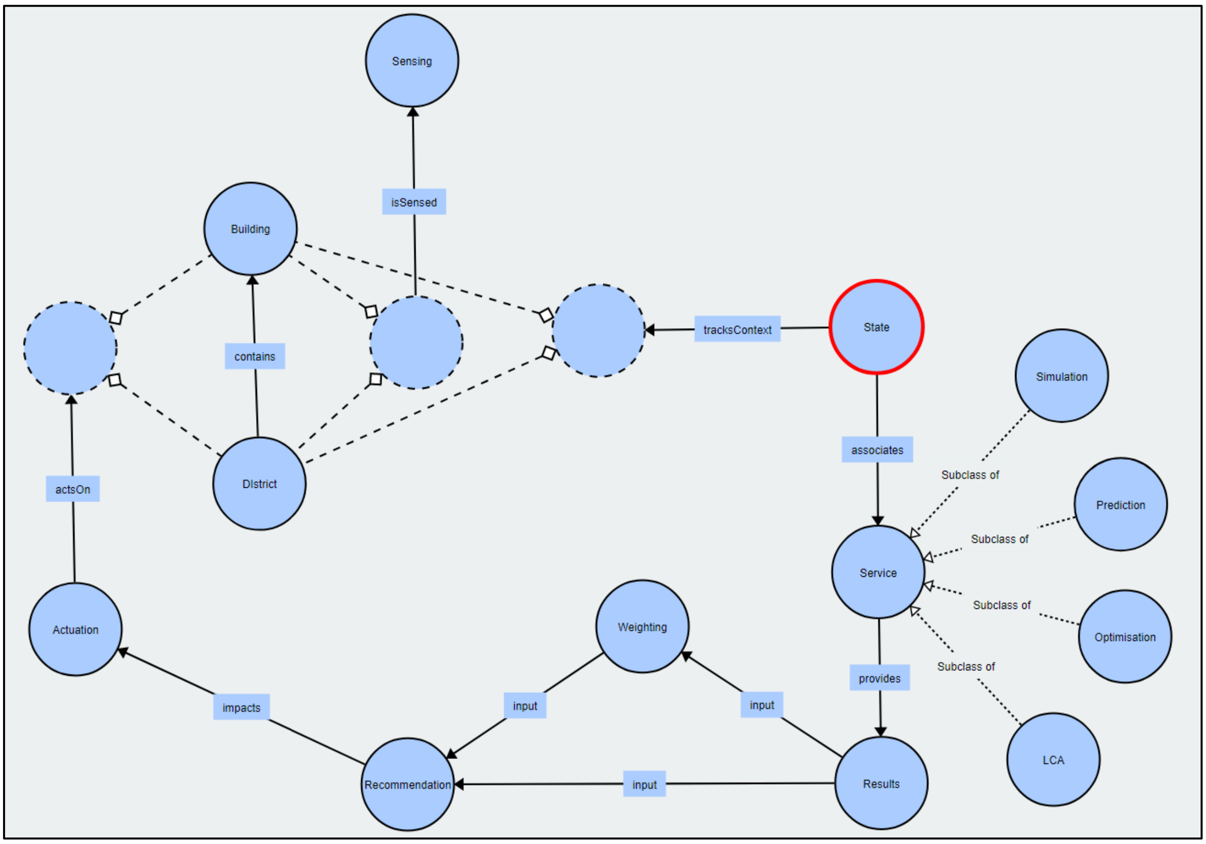


Figure 2: Overview of SemanticLCA concepts from key domains

LCA and weighting concepts: This domain of the ontology provides a generic view on how LCA calculates impacts for different collections of physical assets or other activities, in relation with weighting methodologies. More specifically it consists of key concepts such as LCA *ImpactCategories*, LCA *Results*, *Methodologies* (or collections) of *Activities* and *Processes*. It also incorporates the key concepts of weighting an LCA result (i.e., allowing for a method where each lifecycle cost is apportioned appropriately for the problem being considered), this introduces concepts such as weighting

---

[12] https://github.com/bojecp/slca/

methods and representation of scores. This domain also focuses on integration with relevant LCA tools and services, such as including LCA databases (i.e., Ecoinvent).

Building concepts: This domain incorporates the integration with building data, overlapping with the BIM model data. This includes the representation of key building topological concepts, building *Elements* along with their *Materials*. The *Element* class is the equivalent to *bot:Element*, and the more generic *ifcowl:IfcElement* in the IFC schema. Useful generic properties such as identifiers are represented as data properties, while more specific object characteristics are defined at the *ElementType* class level, an equivalent of *ifcowl:IfcObjectType*. Building materials are defined similarly, with the *Material* class defining instances of material quantities linked to a specific *MaterialType*. This is more convenient for LCA datasets, but different from IFC modelling. The domain also models some key properties of the building (and their elements) which are important to LCA. This includes building types, building system types and configuration, and occupancy profiles. This element of the ontology provides a high-level semantic conceptualization that will then be related to other broader representations i.e., BOT or the more detailed and heavyweight ifcOWL.

District concepts: The domain represents city district representations, where building (and other built assets) are grouped by zones, streets or other containers. For the considered use cases, this domain is less concerned with geometry, but more concerned on general characteristics of buildings, such as archetypes, their age, and energy profiles, occupancies and relationships between built assets. This is set to allow the scaling up of LCA application from the building domain to a district level on both embedded materials, but also operational water and energy usage. We adopt a bottom-up approach when considering the city district (i.e. a district is a collection of buildings).

Sensing concepts: This domain models a high-level view of sensors applied to a built asset (inside a building, or within a city district location). It defines the *Sensor* class which represents physical or virtual sensors, adopting an IoT modelling perspective. This is linked to a *SensorType* class which describes device metadata, and a Reading class which is an instance of a measure, similar to a sosa:Observation. The *MeasurementType* class describes data about read values, such as labels and units, similar to sosa:ObservableProperty. Furthermore, it provides a lightweight set of concepts to enable the modelling of a timeseries set of sensor data. Critically, this domain also includes the concept of a sensor's history, enabling the modeling of portable sensors that may be moved during their lifetime. This domain is modelled specifically to define building and district level sensing devices, more specific than the SNN/SOSA's sensing network. They are important to integrate timeseries databases (i.e., InfluxDB), but easily compatible with more comprehensive sensing ontologies such as SOSA/SSN or Saref4Buildings.

Scenario and state concepts: This domain models the concepts of scenarios and states of built assets (at building or district levels), thus enabling the exploration of several configuration scenarios. This is done through modelling the building moving through a series of linear or branching states. The states can be physical to represent modifications to the building that have occurred, or virtual to represent states that do not current represent the physical reality of the building. This is critical to enable the modeling of the exploration of refurbishment scenarios.

Metadata concepts: Several support and linking concepts were needed to map the diverse domains for a complete network of ontologies. Dublin Core ontologies and SKOS are considered for annotations and inferred mappings respectively, for example.

The development has followed a pragmatic approach, based on a set of use cases, and will be further iterated as part of ongoing and future work. A standard ontology engineering approach was followed, with the identification of use cases, derivation of competency questions, analysis of existing relevant semantic resources and then a "from-scratch" modelling exercise conducted. An interesting point to note is that, when analyzing the relevant existing semantic resources (as shown in Table 2), unsurprisingly many BIM related resources were identified, but very few updated LCA ontologies. The main candidate in this field is the existing BONSAI ontology, however, while this is a related ontology, it does not provide any concepts necessary to link with BIM, sensing or any of the other domains considered by this paper; nor does it model concepts related to impacts for specific types of product systems.

Table 1
Relevant tools with concepts related to Semantic LCA domains

Ontology/Schema	Type	Domain	Concepts of interest
BONSAI	ontology	LCA	activities, flows, properties
LCA Commons[13]	library	LCA	flows, mappings
OpenLCA Schema[14]	schema		methods, processes, results, impact categories
Brightway2[15]	Library	LCA	toolkits for calculation
IfcOwl[16]	ontology	Building	elements, materials, properties, aggregations, hierarchies, types, spatial constructs, sensors
BOT[17]	ontology	Building	spatial constructs, elements
OPM[18]	ontology	Building	properties
RealEstateCore[19]	ontology	Building	building systems, elements, spatial constructs, sensors and controls
Building Product Ontology[20]	ontology	Building	elements, aggregations, properties
SSN-SOSA[21]	ontology	Sensing	sensors, observations, results
SAREF[22]	ontology	Sensing Integration	devices, appliances, energy, gas
Freeclass ontology[23]	ontology & classification	Building	hierarchies, building products, processes, properties
DCMI[24]	ontology	Metadata	annotations
SKOS[25]	ontology	Metadata	semantic associations
XKOS[26]	ontology	Metadata	associations of classifications
OWL-time[27]	ontology	Scenario & State	time spans, start, end times

5. Case study on building data

For this study we use a system under development to integrate concepts related to buildings and LCA from the SemanticLCA ontology with IFC model data for real buildings. The case study showcases different scenarios on working with various datasets, and several integrated ontologies, which can be used as-is on a dedicated triple store, or to act as a knowledge base for a dedicated LCA calculation service or tool.

Several SWRL rules and SPARQL queries were formulated from the competency questions to show how data (Abox) can be mapped and retrieved. The ontology has to provide matches between several concepts coming from different domains. Although these relationships exist at the Tbox level, according to the ontology model, the Abox assertions need to be created and their mappings found. In

---

[13] https://github.com/USEPA/Federal-LCA-Commons-Elementary-Flow-List
[14] http://greendelta.github.io/olca-schema/
[15] https://github.com/brightway-lca/brightway2
[16] https://technical.buildingsmart.org/standards/ifc/ifc-schema-specifications/
[17] https://w3c-lbd-cg.github.io/bot/
[18] https://w3id.org/opm
[19] https://doc.realestatecore.io/3.2/full.html
[20] http://w3id.org/bpo/1-2/
[21] https://www.w3.org/TR/vocab-ssn/
[22] https://ontology.tno.nl/saref/
[23] https://www.freeclass.eu/semanticSearch_Information
[24] https://www.dublincore.org/specifications/dublin-core/dcmi-terms/
[25] https://www.w3.org/2004/02/skos/
[26] https://ddialliance.org/Specification/RDF/XKOS
[27] https://www.w3.org/TR/owl-time/

practice, each BIM model will likely have different types of objects and materials, while the LCA data can come from different providers.

Within this case study we address several problems as part of the building LCA, and aggregate data according to Figure 3. Firstly, we gather BIM, LCA and sensed data from different domains and construct a series of triple graphs for the individuals of each domain. The SemanticLCA ontology constructs the context on a triple store database (Jena Fuseki), and connects several existing ontologies (some mentioned in Table 2); within this case study we use the SKOS, Bonsai and SOSA ontologies. Mapping the *Element* class to BOT is logical, but BOT alone cannot provide the constituent materials and their types. Secondly, we define several SWRL rules to help associate Abox assertions from the LCA domain with building materials from the BIM. We use several SKOS semantic relationships to implement this. Thirdly, we query the instance data using SPARQL to test several competency questions on identifying the LCA impacts of building materials and sensor readings.

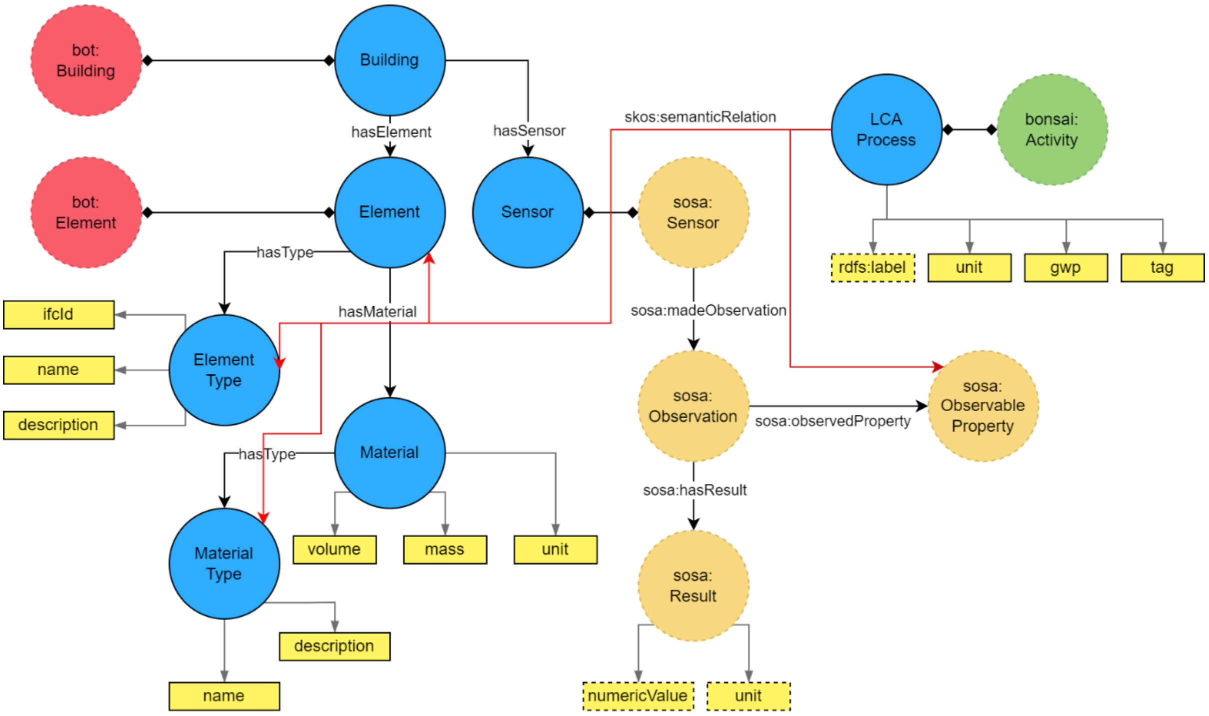


Figure 3: Schema level representation of Tbox assertions, with imported concepts from nearby domain ontologies.

5.1. Matching materials using SKOS

We can use SKOS to infer some relationships between the building element materials and their equivalent materials in the LCA domain. The SKOS ontology provides means to map concepts across different domains and supports two main types of semantic relationships: hierarchical and associative. Hierarchical relationships can be used to classify concepts from “broader” (more generic) to “narrower” (more specific), whilst the associative relationships are specified from “related” to “close” and “exact” matches. Within this case study, we use rules to infer associative relationships between IFC elements’ materials and LCA specific processes.

For the implementation of dynamic mapping of instance data from multiple domains, we developed a multi-tier process for rule sets, which use SKOS associations. We defined 3 distinct tiers (or typologies) of rule sets, based on the relationships possible from SKOS reasoning:

- Tier III – related data; a broad match, with many potential results;
- Tier II – relevant data; a close match; high possibility of exact concepts with several results;
- Tier I – exact data; an identical match; 1 specific result.

Within Table 3 we shown several examples of rules expressed in SWRL, which match related building element materials to LCA processes. These rules use string matching functions to find semantic relations between *tags* for LCA processes, which are matched against building element names, and the related material type names. *Tags* are simple strings used to describe each LCA concept with key words, in order to parse and find them more conveniently. Rules 1 and 2 in Table 3 are considered Tier III, where simple matches are found. To increase the accuracy, we defined a Tier II rule, whereby if the LCA process already has inferred a *relatedMatch* for both the building element and the material type individuals, then this considered a *closeMatch* in SKOS for the materials. A tier I rule for exact matches only makes sense if certain symbols are used, such as classification codes, which would be better suited for the XKOS ontology instead. Another example of a Tier II rule is the correct identification of the location of the products. Rules to match the location of the LCAProcess, with the location of the building from the BIM, assuming that the closest materials are bought at the construction site can be defined. This would further filter results to the defined LCA process scope.

Table 2

Example rules for matching LCA processes with building material types using SKOS relationships

no	Tier	Body	Head
1	III	lca:tag(?p, ?tag) ^ building:elementName(?e, ?name) ^ swrlb:matches(?name, ?tag)	-> skos:relatedMatch(?e, ?p)
2	III	lca:tag(?p, ?tag) ^ building:elementTypeName(?et, ?name) ^ swrlb:matches(?name, ?tag)	-> skos:relatedMatch(?et, ?p)
3	III	lca:tag(?p, ?tag) ^ building:materialTypeName(?m, ?name) ^ swrlb:matches(?name, ?tag)	-> skos:relatedMatch(?m, ?p)
4	II	building:Element(?e) ^ skos:relatedMatch(?e, ?p) ^ building:MaterialType(?m) ^ skos:relatedMatch(?m, ?p)	-> skos:closeMatch(?m, ?p)
5	II	building:ElementType(?et) ^ skos:relatedMatch(?et, ?p) ^ building:MaterialType(?m) ^ skos:relatedMatch(?m, ?p)	-> skos:closeMatch(?m, ?p)

We used several SPARQL queries to find materials, with both the related and close matches for SKOS, and compare the results with the building materials. These are summarized in table 3 below, after testing on two different IFC models: (A) open building model[28] with 218 elements (E), 37 types (T) and 18 distinct material types (M); (B) private model with 1661 elements, 187 types and 43 distinct material types. These were associated in turn with two distinct LCA datasets, one with 40 and another with 600 processes. Model (A) is smaller in size, with rather vague material descriptions (as placeholder materials, whereas model (B) is larger, each element has an associated type (and IfcType), with material descriptions from the NL-SfB[29] classification system.

Table 3

SPARQL results on BIM models in combination with LCA datasets from testing Table 2 rules

no	Tier	BIM	LCA ind.	BIM ind.	matches	correct	comments
1	III	A	40	E - 93/218	146	86/146	openings not matched, furniture
		A	600	E - 140/218	1522	596/1522	out of scope for the LCA dataset
		B	40	E - 1251/1661	1865	n/a	
		B	600	E - 1476/1661	23703	n/a	
2	III	A	40	T - 4/37	4	4/4	Furniture not matched,
		A	600	T - 9/37	34	32/34	very few types in model
		B	40	T - 22/187	27	26/27	
		B	600	T - 116/187	596	207/596	
3	III	A	40	M - 14/18	28	23/28	
		A	600	M - 15/18	302	221/302	
		B	40	M - 26/43	52	30/52	
		B	600	M - 31/43	625	328/625	
4	II	A	40	M – 11/18	16	9/16	Lots of types of concrete on LCA
		A	600	M – 11/18	197	147/197	data, BIM materials are vague

---

[28] http://openifcmodel.cs.auckland.ac.nz/Model/Details/274

[29] http://nl-sfb.bk.tudelft.nl/eng.htm

		B	40	M – 20/43	38	24/38	
		B	600	M – 25/43	415	256/415	
5	II	A	40	M – 0/18	0	n/a	poor BIM quality
		A	600	M – 0/18	0	n/a	poor BIM quality
		B	40	M – 3/43	3	1/3	
		B	600	M – 6/43	26	13/26	

The tests show slightly better results when the BIM data quality improves (for model B), and also more materials are matched when the LCA dataset increases. However, the possible correct combinations (column 7 in Table 3) are many because BIM materials are vague compared to LCA processes, and this can lead to very different results. The initial matches are useful to filter initial LCA processes, but the creation of an LCA scenario requires more advanced interpretation. Alternatively, more precise specification of BIM model materials using codes, and having them pre-matched to specific LCA processes.

5.2. Calculating the global warming potential

To demonstrate the utility of connecting LCA processes with the BIM model data, we asserted the correct mappings between materials for model (A) previously described, using *skos:closeMatch*. In this scenario we can then directly use the SemanticLCA ontology to calculate the impacts of each building component, as each *lca:LcaProcess* (or *bonsai:Activity*) individual has a precalculated global warming potential (GWP) property, denoted by *bw2ont:gwp*, the units of which are specified by *bw2ont:unit*., which here are in kg of $CO_{2equivalent(eq)}$. Query (SQ1) in Figure 4 shows the GWP in kg of CO_{2eq} of one building element with several materials. The element in question is an external wall with several material layers. Working in reverse order, we can query the material type in question to calculate the total impacts (on all building elements combined), as shown in Figure 5.

```
PREFIX lca: <http://slca.org/slca/lca/>
PREFIX building: <http://slca.org/slca/building/>
PREFIX skos: <http://www.w3.org/2008/05/skos#>
PREFIX rdfs: <http://www.w3.org/2000/01/rdf-schema#>
PREFIX bw2ont:<http://ontology.brightway.dev/>

SELECT DISTINCT ?materialName ?lcaName ?mass ?gwp (ABS(?gwp * ?mass) as ?calculatedGwp)
WHERE {
  ?element building:hasMaterial ?material .
  ?material building:materialMass ?mass .
  ?material building:hasMaterialType ?materialType .
  ?materialType building:materialTypeName ?materialName .
  ?materialType skos:closeMatch ?lcaProcess .
  ?lcaProcess rdfs:label ?lcaName .
  ?lcaProcess bw2ont:gwp ?gwp .
}
```

	materialName	lcaName	mass	gwp	calculatedGwp
1	"Masonry - Brick"	"market for clay brick"	"2504"^^xsd:double	"0.318174743118504"^^xsd:decimal	"796.709556768734e0"^^xsd:double
2	"Plasterboard"	"market for gypsum, mineral"	"174.16"^^xsd:double	"0.06483115139577918"^^xsd:decimal	"11.290993327088902e0"^^xsd:double
3	"Metal - Stud Layer"	"market for steel, low-alloyed"	"30"^^xsd:double	"1.9957075713862387"^^xsd:decimal	"59.87122714158716e0"^^xsd:double
4	"Insulation / Thermal Barriers - Rigid insulation"	"market for stone wool"	"24.5"^^xsd:double	"1.3369904608859726"^^xsd:decimal	"32.75626629170633e0"^^xsd:double
5	"Masonry - Concrete Block"	"market for concrete block"	"5135"^^xsd:double	"0.07056324260584888"^^xsd:decimal	"362.342250781034e0"^^xsd:double

Figure 4: SPARQL (SQ1) calculating an element's GWP with results in kg CO_{2eq}

The results of query SQ1 use kilograms of the materials embedded which needs to rely on explicit and accurate BIM object properties. Query SQ2 uses the volumes of materials, as modelled in BIM and calculated by BIM platforms, but this can also be problematic. To correctly identify the impacts of elements or materials, the correct units must be associated with LCA processes. Imagine a scenario where the material type named “*Metal - Stud layer*” within a composite wall, with a volume of 0.4 m^3, which denotes the volume of the entire width that represents the metal studs positioning within the wall. If we consider this volume for the GWP calculation, the result will not be representative, as within the containing volume we have much less material, as there are several hollow metal struts at a distance from each other, which for our example equated to 30 kg along the entire length, as opposed to 1000kg equivalent from the volume. To limit this, the units of LCA processes are checked to match the units of each material where possible, whereas the values are denoted separately here by data properties *building:materialVolume* and *building:materialMass*.

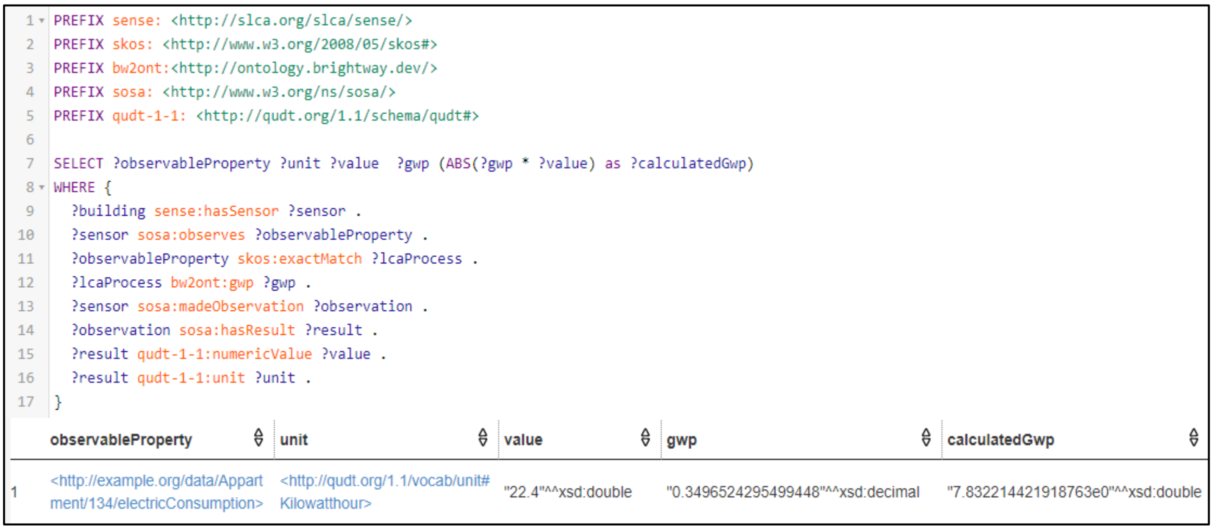
```
PREFIX building: <http://slca.org/slca/building/>
PREFIX skos: <http://www.w3.org/2008/05/skos#>
PREFIX bw2ont:<http://ontology.brightway.dev/>

SELECT ?materialName  ?totalVolume (ABS(?gwp * ?totalVolume) as ?totalGwp)
WHERE {
  ?materialType building:materialTypeName ?materialName .
  ?materialType skos:closeMatch ?lcaProcess .
  ?lcaProcess bw2ont:gwp ?gwp .
  {
    SELECT ?materialType (SUM(?vol) as ?totalVolume)
    WHERE {
      ?material building:materialVolume ?vol .
      ?material building:hasMaterialType ?materialType .
    } GROUP BY ?materialType
  }
  FILTER STRSTARTS(?materialName, "Concrete")
}
```

	materialName	totalVolume	totalGwp
1	"Concrete"	"48.761477286999956e0"^^xsd:double	"14301.174376684177e0"^^xsd:double

Figure 5: SPARQL (SQ2) calculating a material type’s GWP across the BIM model

Lca processes relate to various product systems, some of which can be defined as the production of materials (shown previously), but this can also define the production of energy. To highlight the utility of smart devices (sensors/actuators) within the SemanticLCA ontology, we linked it with a sample SNN/SOSA dataset[30], where a sample sensor registers the energy use within a building; we asserted that the sensor is connected to our building model. In a similar fashion to mapping materials to elements, we mapped a *sosa:ObservableProperty* to an electricity production LCA *bonsai:Activity*. SPARQL query SQ3 (in Figure 6) shows a way of calculating the GWP from the electricity consumption from the last *sosa:Observation*.



```
PREFIX sense: <http://slca.org/slca/sense/>
PREFIX skos: <http://www.w3.org/2008/05/skos#>
PREFIX bw2ont:<http://ontology.brightway.dev/>
PREFIX sosa: <http://www.w3.org/ns/sosa/>
PREFIX qudt-1-1: <http://qudt.org/1.1/schema/qudt#>

SELECT ?observableProperty ?unit ?value  ?gwp (ABS(?gwp * ?value) as ?calculatedGwp)
WHERE {
  ?building sense:hasSensor ?sensor .
  ?sensor sosa:observes ?observableProperty .
  ?observableProperty skos:exactMatch ?lcaProcess .
  ?lcaProcess bw2ont:gwp ?gwp .
  ?sensor sosa:madeObservation ?observation .
  ?observation sosa:hasResult ?result .
  ?result qudt-1-1:numericValue ?value .
  ?result qudt-1-1:unit ?unit .
}
```

	observableProperty	unit	value	gwp	calculatedGwp
1	<http://example.org/data/Appartment/134/electricConsumption>	<http://qudt.org/1.1/vocab/unit#Kilowatthour>	"22.4"^^xsd:double	"0.3496524295499448"^^xsd:decimal	"7.832214421918763e0"^^xsd:double

Figure 6: SPARQL (SQ3) calculating the impact in GWP of measured electricity use from a sensor

---

[30] https://github.com/w3c/sdw/blob/gh-pages/ssn/integrated/examples/apartment-134.ttl

6. Discussion and conclusion

Within this paper we showcased the SemanticLca ontology which is focused on practical use cases, but still a work in progress. The scope of the full ontology is quite broad, but it was designed to map and incorporate several nearby domain ontologies. The literature review shows several domain ontologies related to buildings and LCA (listed in Table 2), but none deal with more precise context definitions which are required for specific built environment product systems in LCA. We showed several mappings between building components with BONSAI, SSN/SOSA and BOT (in Figure 3), and the utility of matching materials to LCA processes dynamically, using rules and existing datasets on a knowledge base in Jena Fuseki. The presented ontology development method specified several steps, and it showcased the implementation on the building use case, where building component materials and monitored building data is used for LCA. The ontology is developed and tested as a support schema for software tools to collect and integrate information across multiple built environment domains. In this sense, it is a simple ontology for generic concepts which are complemented by several more specialized ontologies (several of which we demonstrated in the case study). We do not directly import smaller specialized ontologies at this stage, but this is foreseen for future versions. The SemanticLCA ontology was not designed specifically for reasoning, but more for mapping generic concepts across domains, allowing more complex multi-domain data integration, such as the LCA for buildings use case shown here. We do however demonstrate reasoning using SKOS here.

The implementation of SWRL rules can be replaced by systematic SPARQL queries which achieve similar results, and inferred triples can be asserted back into the graph. The SKOS ontology allows a convenient way to classify concepts on schema (Tbox) and data (Abox) levels, but it was not meant to infer “facts”. The limitations of the dynamic mapping is the appearance of false positives and mismatches due to a lack of clear information from the two domains. Although correct matching is difficult to achieve with sematic tools alone, this is a first step in filtering related concepts. This can be improved with additional rules, external algorithms or human validation, which could be used to assert close and exact matches as well. Configuring rules for exact matches proves problematic due to the vagueness of the datasets on both sides. This could partly be solved by relying on explicit annotations or symbols, such as classification codes, and implemented using the XKOS ontology.

The implemented SPARQL allow us to find matches of materials at different levels, and calculate their impacts, as shown in section 5.2 for the GWP of an element or an entire material within the building model for static building data. The dynamic sensed data can be considered in very similar ways, as shown in calculating the GWP impacts of electricity consumption from smart meter readings. This can be further developed by external applications to create various scenarios and compare results.

Future works is set around integrating other relevant domains, such as city districts, sensors and the use of weighting methodologies for LCA results, which will further increase the completeness of the SemanticLCA ontology for built environment use cases.

7. Acknowledgements

The authors would like to acknowledge the financial support of the Fonds National de la Recherche (FNR) in Luxembourg and the Engineering and Physical Sciences Research Council (EPSRC) in the UK under grant agreement (INTER/UKRI/19/14106247; EP/T019514/1) for the SemanticLCA research project.

8. References

[1] M. Röck, A. Hollberg, G. Habert, and A. Passer, “LCA and BIM: Visualization of environmental potentials in building construction at early design stages,” *Build Environ*, 2018, doi: 10.1016/j.buildenv.2018.05.006.

[2] P. Pauwels, S. Zhang, and Y.-C. Lee, “Semantic web technologies in AEC industry: A literature overview,” *Autom Constr*, vol. 73, pp. 145–165, Jan. 2017, doi: 10.1016/j.autcon.2016.10.003.

[3] F. H. Abanda, J. H. M. Tah, and R. Keivani, “Trends in built environment semantic Web applications: Where are we today?,” *Expert Syst Appl*, vol. 40, no. 14, pp. 5563–5577, Oct. 2013, doi: 10.1016/j.eswa.2013.04.027.

[4] C. Boje, A. Marvuglia, J. Hermen, A. Guerriero, T. Navarrete, and E. Benetto, “A pilot using a Building Digital Twin for LCA-based human health monitoring,” *Proceedings of the CIB W78 Conference*, pp. 568–577, 2021, [Online]. Available: https://itc.scix.net/pdfs/w78-2021-paper-057.pdf

[5] C. Cavalliere, “BIM-led LCA: Feasibility of improving Life Cycle Assessment through Building Information Modelling during the building design process,” *Politecnico di Bari*, 2018.

[6] L. Wastiels and R. Decuypere, “Identification and comparison of LCA-BIM integration strategies,” *IOP Conf Ser Earth Environ Sci*, vol. 323, no. 1, 2019, doi: 10.1088/1755-1315/323/1/012101.

[7] A. Hollberg, G. Genova, and G. Habert, “Evaluation of BIM-based LCA results for building design,” *Autom Constr*, vol. 109, no. October 2019, p. 102972, 2020, doi: 10.1016/j.autcon.2019.102972.

[8] W. W. Ingwersen *et al.*, “A new data architecture for advancing life cycle assessment,” *International Journal of Life Cycle Assessment*, vol. 20, no. 4, pp. 520–526, Apr. 2015, doi: 10.1007/s11367-015-0850-6.

[9] E. Meex, A. Hollberg, E. Knapen, L. Hildebrand, and G. Verbeeck, “Requirements for applying LCA-based environmental impact assessment tools in the early stages of building design,” *Build Environ*, vol. 133, no. February, pp. 228–236, 2018, doi: 10.1016/j.buildenv.2018.02.016.

[10] K. Safari and H. AzariJafari, “Challenges and opportunities for integrating BIM and LCA: Methodological choices and framework development,” *Sustain Cities Soc*, vol. 67, no. December 2020, p. 102728, 2021, doi: 10.1016/j.scs.2021.102728.

[11] B. Bertin, V.-M. Scuturici, J.-M. Pinon, and E. Risler, “CarbonDB: A Semantic Life Cycle Inventory Database,” in *Proceedings of the 21st ACM International Conference on Information and Knowledge Management*, in CIKM ’12. New York, NY, USA: Association for Computing Machinery, 2012, pp. 2683–2685. doi: 10.1145/2396761.2398725.

[12] K. Janowicz *et al.*, “A minimal ontology pattern for life cycle assessment data,” in *Proceedings of the 6th Workshop on Ontology and Semantic Web Patterns (WOP 2015) co-located with the 14th International Semantic Web Conference (ISWC 2015)*, E. Blomqvist, P. Hitzler, A. Krisnadhi, T. Narock, and M. Solanki, Eds., in CEUR Workshop Proceedings, vol. 1461. CEUR Workshop Proceedings, 2015.

[13] M. C. Suárez-Figueroa, A. Gómez-Pérez, and M. Fernández-López, “The NeOn Methodology framework: Ascenario-based methodology for ontologydevelopment,” *Appl Ontol*, vol. 10, no. 2, pp. 107–145, Sep. 2015, doi: 10.3233/AO-150145.