Costin, A. and Eastman, C. (2017). "Requirements for Ontology Development in the AECO Industry". In: *LC3 2017: Volume I – Proceedings of the Joint Conference on Computing in Construction (JC3)*, July 4-7, 2017, Heraklion, Greece, pp. 533-540. DOI: https://doi.org/10.24928/JC3-2017/0149.

REQUIREMENTS FOR ONTOLOGY DEVELOPMENT IN THE AECO INDUSTRY

Aaron Costin[1] and Charles Eastman[2]

Abstract: This paper presents and discusses the requirements needed for the development of ontologies in the Architecture, Engineering, Construction, and Operation (AECO) Industry. With the increase of information modeling for all aspects of a construction project with a variety of software tools and technologies, there has been a major need of communication and exchange of information. An approach to improve seamless information exchanges is the use of ontologies. One major benefit of using ontologies is that the information and knowledge defined in the ontologies can be shared across domains. However, to do so requires standardized rules and requirements in order to share and promote reuse at the domain level. Significantly, with the increased demand of ontologies in the AECO industry, there needs to be standardization and consensus in the development and use of the ontologies to ensure the seamless transfer of information as well as realizing the full benefits of ontologies.

Keywords: Ontology, Information Exchange, Semantics, Logic, Taxonomy, AECO industry.

1 INTRODUCTION

Over the last few decades, the Architecture, Engineering, Construction, and Operation (AECO) Industry has seen an increase in the amount of computer software, technologies, and automation to help improve all facets of the industry. With the increase of such innovations, interoperability between the different domain facets is a major need for the seamless transfer of information. Interoperability is defined as the ability of one system (e.g. operating system, software application, workflow, process, etc.) to work with other systems without any effort on the part of the end user of the system (Costin, 2016). Non-interoperable systems require a significant manual effort to transfer information from one application to the other, which could result in errors, omissions, or data loss.

One method to increase interoperability in computer systems is the use of ontologies (Costin 2016). An ontology is a formal representation of an abstract view of a domain that describes the objects, concepts and relationships between them that hold in that domain (Gruber 1993). Ontologies are used to define and standardize machine-readable information that describes domain knowledge to be utilized for a broad range of applications. Reasons for using ontologies include sharing common understanding of the structure of information among people or software agents, enabling the reuse of domain knowledge, making domain assumptions explicit, separating domain knowledge from the operational knowledge, and analyzing domain knowledge (Noy and McGuinness 2001). With the increasing demands for domain-wide integrated construction and infrastructure development, there is a growing need for the development of an ontology for the

1 Assistant Professor, M. E. Rinker, Sr. School of Construction Management, University of Florida, Gainesville, Florida, USA, aaron.costin@ufl.edu

2 Professor, School of Architecture, Georgia Institute of Technology, Atlanta, Georgia, USA, eastman@design,gatech.edu

construction domain that supports the multi-taskholder project development process (El-Gohary et al. 2010). Ontologies for the AECO industry have the potential to provide seamless information exchanges across the domain, however, various rules and requirements must be met in order to do so. Therefore, the purpose of this paper is to provide an overview of relevant topics needed for implementing an ontological framework. First, semantics and logic are two important aspects needed to be captured to promote information exchanges. Next, understanding the various categories and relations between ontologies are important for ontology development. Specific programming languages are needed to develop and use ontologies. Finally, ontologies in the AECO industry are reviewed followed by the requirements for effective implementation.

2 SEMANTICS AND LOGIC

Ontologies need to capture two important elements: semantics and logic. Semantics are the meanings and interpretations of a word or phrase in a specific context. At the core level of computer science, data are essentially bits and bytes that the computer uses in processes in which are essentially useless to both human and computer function without any context (Costin, 2016). Therefore, it is important that the data be given the semantic information needed to represent what human function represents. For example, when data are exchanged between BIM software, it is insufficient to solely rely on 3D visual properties of the objects (Eastman et al. 2010). Although the geometries of the objects in the 3D model are important, they alone are not sufficient to describe the needed meaning of the modeled objects. At the exchange level (i.e. passing information), semantics may cause issues for humans and computers that are interpreting the context. The goal is to have semantic consistency in an information exchange in which the human-based knowledge and computer based interpretation of the information are equivalent, i.e. the computer understands what the user intends.

Logic is the reasoning behind the development of the ontology of how the domain works. Logic is represented by Description Logic (DL), which is the basis of most ontology languages. DL is the formal knowledge representation used to express the conceptualization of domains in an organized and formally well-understood manner (Taye 2011). The logic structure contains the additional axioms (logic assertions) provided by the ontology language in a common form (structure). In this research, an axiom is a "stated rule or principle that helps govern the ontology" (Costin 2016). Axioms are similar to postulates (e.g. math or geometry postulates), in which they are assertions without any formal proofs (although formal definitions are usually provided), and are used for deducing other truths. Axioms are typically written out in first order logic, such as in Uschold and Gruninger (1996). As part of mathematical logic, these types of rules are associated with type theory. Table 1 displays examples of axioms used in the BrIM ontology for bridges described in (Costin 2016).

3 ONTOLOGY CATEGORIES

Ontologies can be grouped in different categories, or levels, based on the purpose and information contained within. Borst et al. (1997) defined three categories of ontologies: 1) highly generic, or "super theory," ontologies, 2) base ontologies representing natural categories or viewpoints within a broad field; and 3) domain ontologies which are specializations of base ontologies to a specific domain. Understanding these categories and their relations are important for ontology development (see Figure 1). There is no "one" ontology, but rather multiple ontologies to define needed information. In order to define or reuse ontologies for the AECO domain, all three categories must be addressed.

Table 1: Examples of defining axioms (Costin 2016).

Relations	First Order Logic	Definition	Example
ComposeOf	Composed-of (A,B) $\leftrightarrow$ (B⊆A) ∧ (A⊈ B)	B is composed of A, if A is a subset of B, and B is not a subset of A	Bridge is composed of smaller parts, e.g. columns, beams, etc.
InverseTo	∀A,B**f**(A,B)↔**g**(B,A)	For all A and B, relation **g** is the inverse of relation **f** if A maps to B and B maps to A.	If beam is partOf bridge, then bridge hasPart beam.

3.1 Super Theories

Super theories can be seen as the "ontologies of ontologies" since they describe the main principles and rules of ontologies. Three main super theories include mereology, topology, and systems theory. First, mereology is that of parenthood relations, such as the relations of part-to-whole or the relations of part-to-part within a whole. Mereology also includes the relation of composition (combining parts to make an object) and decomposition (breaking an object into parts), which essentially describes how an ontology is represented. For example, a major feature of an ontology is that it is organized in a hierarchical fashion, called a taxonomy. This feature allows for the assertion of mereology relationships, among other factors. For example, a building (whole) is composed of various components such as structure, floors, rooms (parts) and the relations between each can be represented by mereology relations. Additionally, each of these parts, themselves, can be composed of additional parts. The taxonomy is important since it can visually and programmatically display the mereological relations, such as showing how the parts of a building are related.

Topology is the study of shapes and how the properties are preserved due to deformations, in which topology can be used to abstract the inherent connectivity of objects. The topological ontology is a projection (extension) of mereology because it defines a relation to express the fact that mereological individuals are connected (Borst et al. 1997). Shape representation and relations are an important aspect of modeling infrastructure, and topology is needed to determine and infer the relations between objects. Lines, cubes and spheres are all geometry primitives that represent certain types of objects that can be used to represent topology or spatial volumes. For example, various software CAD tools utilize these primitive objects to design components and features. Knowing how these shapes relate to one another can result in compositions of more complex shapes, such as a nut fastening to a bolt.

Finally, systems theory is the study of a system, and how system components interact within the system and the outside environment. This extension of topology defines the system properties and system boundaries. For example, systems theory provides the abilities to study how a structure and its components act with other environments, while potentially providing the means for enabling the realization and deployment of a structure as a system.

3.2 Base Ontologies

Base ontologies are specializations of super theories that formalize a conceptual category of concepts for a domain. Each domain will have their own base ontologies, in which for science and engineering fields include the configuration of components, physical processes underlying behavior, and the engineering mathematics that describe the processes (Borst et al. 1997). For example, a configuration of components would include the structural components (beam, wall, etc.) and how they interact with the other components. Physical processes describe the mechanisms of how things flow or behave, such as dynamic behavior (hydraulics, acoustics, electricity, etc.). The engineering mathematics describe the physical processes, such as time, units, and equations.

3.3 Domain Ontologies

Domain ontologies form an integral and coherent conceptualization (view of) of a domain (Borst et al. 1997). Since the domain ontology is a specialization of the base ontologies, the information contained in it can be extended and reused by other domains. This is important since many of the domains in the AECO industry overlap, such as construction, safety, transportation, buildings, bridges, etc.

Figure 1 describes how the three categories of ontologies relate. Clearly, without addressing each category above would result in incomplete domain ontologies for the AECO industry by lacking the semantics or logic needed to express the full intent. The ability for domains to share knowledge is a significant aspect since it can help promote the goal of interoperability. Thus, this is why it is extremely important that ontologies follow standardized rules and guidelines in their development to guarantee consistency and reliability. Moreover, interoperability will allow the base and super theory ontologies to be shared with other fields. This is import because each domain will not need to recreate the ontologies since they can link to information captured in other ontologies.

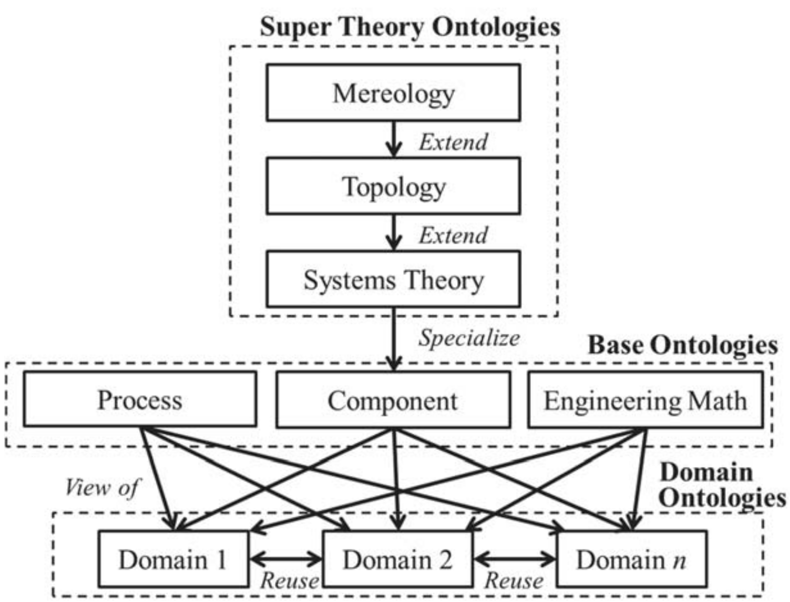

Figure 1: Integration of the three levels of ontologies: super theory, base, and domain. (Adopted and modified from Borst et al.1997).

4 Ontology Languages

Just as there are computer languages to code software, ontology languages are needed to describe and code ontologies in a machine-readable format. With the development of computer systems, there have been numerous research efforts to design ontology languages. The Knowledge Interchange Format (KIF), now obsolete, was one of the first to be developed in the early 1990s (Genesereth and Fikes, 1992). The Resource Description Framework (RDF) is a framework for conceptual description and the modeling of information that is implemented in web resources. RDF is composed of three components, known as RDF triples: subject, predicate, and object RDF triples state a single fact about a resource, in which the subject is the subject being described, the predicate is the relationship of the subject, and the object represents what is related to the subject by the predicate. "Although RDF is a good basic language for building many other languages, it is not very expressive and has limitations in describing resources, including descriptions of existence, cardinality, localised range and domain constraints designating transitive, inverse or symmetrical properties" (Taye 2011). However, RFD provides basic concepts that are used as part of other languages.

The Web Ontology Language (OWL), now in its second release (OWL2), has recently become a go-to language for ontologies. Based on previous languages, OWL 2 is designed to facilitate machine-readable ontology development and sharing over the internet. There are three sublanguages of OWL2 that offer advantages in particular scenarios: OWL2 EL, OWL2 QL, and OWL2 RL (W3C OWL Working Group 2012). OWL2 EL enables polynomial time algorithms for all the standard reasoning tasks that is suited for large and complex ontologies. OWL2 QL enables conjunctive querying that can be used to query information from databases (e.g. SQL). Lastly, OWL2 RL enables the implementation of the polynomial time algorithms since it uses rule-base implementation that operates directly on RDF triples.

5 Ontologies in the AECO Industry

Only a few projects have been undertaken formalizing an ontology specifically for the AECO domain. Issa and Mutis (2015) present a comprehensive overview of ontology research which focus on enabling information exchange, sharing, and integration during construction project activity. These include ontologies for semantic interoperability for construction projects, healthcare facility management, retrieval of construction documents, and construction contracts. Additionally, Issa and Mutis (2015) include research efforts that discuss valuation methods and metrics of AEC domain ontologies.

As part of a construction project called e-COGNOS ("COnsistent knowledGe management across prOjects and between enterpriSes in the construction domain"), El-Diraby et al. (2005) present an ontology to support semantic knowledge management for applications, including semantic indexing and collaborative project development. Formal ontologies have been developed to support corporate memory management in building construction (El-Diraby and Zhang 2006) to represent knowledge related to various actors and actor roles in the construction industry (Actor-Onto) (Zhang and El-Diraby, 2009), to support knowledge-enabled process management and coordination across the construction project process (IC-Pro-Onto) (El-Gohary et al. 2010), to represents infrastructure product knowledge (IPD-Onto) (El-Diraby and Osman 2011), and to represent construction knowledge (DOCK) (El-Diraby 2013). Recently, ontological approaches have been proposed to support building information modeling exchanges

(Venugopal et al. 2012), to produce construction claim documentation (Niu and Issa 2013), to automate safety planning for job hazard analysis (Zhang et al. 2015), and to promote interoperability of heterogeneous bridge information models (Costin 2016).

With the benefits of ontologies for interoperability becoming more popular over the internet (i.e. Semantic Web), there have been research efforts to map various non-web based languages to web based ontologies, such as the Web Ontology Language (OWL). Schevers and Drogemuller (2005) was the first to propose that the building and construction industry make use of the Semantic Web by proposing that the industry foundation classes (IFC), as part of ISO-STEP, should be available as an OWL ontology. The authors propose various ways of converting (mapping) an IFC file to an OWL file, which sets the foundation for more developed mapping techniques. Beetz et al. (2009) introduces the development of the first ontology for the building and construction sector. Coined "ifcOWL" (properly named for the IFC naming convention), the ontology is transformed from the EXPRESS schema (on which IFC is based) by a semiautomatic approach using various methods, including generic query, RDF, and reasoning algorithms.

Recently, buildingSMART International has adopted ifcOWL and is currently managing the development, as well as hosting the current and live versions of ifcOWL. To help users and developers, buildingSMART has published a recommended usage guide for using ifcOWL (Pauwels et al., 2015). Additionally, the buildingSMART Linked Data Working Group is currently collaborating on developments to use linked data and semantic web technologies to support other buildingSMART International development efforts, such as model view definitions (MVD), the buildingSMART Data Dictionary (bSDD), regulations, and BIM Guides.

6 Ontology Requirements

Below is a discussion of the requirements needed for ontology development for the AECO industry (Costin 2016). These requirements are currently being externally validated by an ongoing case study (Costin et al. 2017).

6.1 Clear and Concise Terminology

At the development level, it is important for the ontology to be based on semantically clear terminology. The terminology must portray domain knowledge in a clear and logical fashion, so that computer systems can interpret human intent. If all of the terminology were defined upfront, along with applicable use cases, the ontology development would be more efficient and accurate (versus ad-hoc development). However, in reality, no one ontologist or computer programmer knows everything about the domain. Therefore, it is imperative that the industry domain work together in order to produce the terminology and knowledge base needed to be modeled into an ontology. Costin (2016) proposed an industry workflow driven ontology development, in which domain knowledge was produced by industry and captured into a taxonomy and then converted into an ontology.

6.2 Sufficiency and Consistency

The ontology needs to meet the domain requirements required to exchange information. To sufficiently address the needs of the scientific and engineering oriented AECO industry, the super theories and base ontologies discussed in Figure 1 need to be addressed. Additionally, there should be no inconsistencies, duplications, or over constraining of definitions (via axioms) in the ontology. The use of reasoners or rule engines for consistency checking is recommended for the ontology.

6.3 Reusable and Extensible

Since one of the benefits of ontologies is sharing knowledge, it is imperative that the ontology be developed in a way to promote reuse and extensibility. To do so requires standard practices for each of the following: methodology for knowledge capture, ontology language, information databases, and reuse of existing defined ontologies. There are existing standards that the AECO industry can utilize (e.g. National BIM Standard, ifcOWL, OWL2, etc.), and following suit of major ontological research will enable the ease of use of the super theories ontologies. This means that if the AECO industry follows the standard practices of the major ontology groups (e.g. W3C), the industry will have more defined ontologies to utilize, and thus can focus more on the domain ontologies.

6.4 Reliability and Security

One major benefit of open ontology languages is the fact that anybody can develop and share ontologies. On the flip side, there is a lack of oversight on the reliability and security of those ontologies. Therefore, it is imperative to establish security measures and safeguards to assure reliability and the prevention of unauthorized altering of the ontology. One solution is to have an organizational body validate, approve, and safeguard the ontology, but also allows for authorized modifications or expansion if needed.

7 Conclusion and discussion

This paper presents an overview of requirements needed to implements ontologies in the AECO industry. Compared to other industry sectors (e.g. healthcare), ontologies usage and development are fairly new to the AECO industry. There have been great research efforts in this industry, and this paper only scratches the surface. The industry can learn a lot from other industry's best practices and standards. In conclusion, this paper ends with questions to induce thought provoking discussions to promote alignment and consensus in the industry.

- What are the specific information exchanges requirements needed to represent the activities and processes in the AECO industry (geometry, human activities, environmental properties, energy models, etc.)?
- What super theories or base ontologies exist that the AECO industry utilize?
- What established domain ontologies in the AECO industry (whether active or inactive) can be utilized and improved upon?
- What are the methods and benchmarks to validate and authenticate currently established ontologies?
- What organization bodies can steward and oversee the development of AECO domain ontologies to promote consistency and alignment?

8 References

Bechhofer, S., van Harmelen, F., Hendler, J., Horrocks, I., McGuinness, D., L., Patel-Schneider, P., F., and Stein, L., A. (2004). *OWL Web Ontology Language Reference,* <https://www.w3.org/TR/owl-ref/> [31 Jan., 2016].

Beetz, J., Van Leeuwen, J., and De Vries, B. (2009). IfcOWL: A Case of Transforming EXPRESS Schemas into Ontologies. *Artificial Intelligence for Engineering Design, Analysis and Manufacturing*, 23(01), 89-101.

Borst, P., Akkermans, H., and Top, J. (1997). Engineering Ontologies. *International Journal of Human-Computer Studies*, 46(2), 365-406.

Costin, A. (2016*). A New Methodology for Interoperability of Heterogeneous Bridge Information Models*. PhD Diss., CEE, Georgia Institute of Technology, Atlanta, GA.

Costin, A., Eastman, C., and Issa, R. (2017). The Need for Taxonomies in the Ontological Approach for Interoperability of Heterogeneous Information Models. *IWCCE*, Seattle, June 25-27.

Eastman, C. M., Jeong, Y. S., Sacks, R. & Kaner, I. (2010). Exchange Model and Exchange Object Concepts for Implementation of National BIM Standard. *Journal of Computing in Civil Engineering*, 24(1), 25–34.

El-Diraby, T. E., and Osman, H. (2011). A Domain Ontology for Construction Concepts in Urban Infrastructure Products. *Automation in Construction,* 20(8), 1120-1132.

El-Diraby, T. E., Lima, C., and Feis, B. (2005). Domain Taxonomy for Construction Concepts: Toward a Formal Ontology for Construction Knowledge. *J. Comput. Civ. Eng.*, 19(4), 394–406.

El-Diraby, T. E., and Zhang, J. (2006). A Semantic Framework to Support Corporate Memory Management in Building Construction. *Automation in Construction*, 15(4), 504-521.

El-Diraby, T. E. (2013). Domain Ontology for Construction Knowledge. *J. Constr. Eng. Manage.*, 139(7), 768–784.

El-Gohary, Nora M., and El-Diraby, T.E. (2010). Domain Ontology for Processes in Infrastructure and Construction. *J. of Constr. Eng. Manage.*, 136(70), 730-744.

Genesereth, M. and Fikes, R. (1992). *Knowledge Interchange Format Version 3.0 Reference Manual*. Technical Report Logic-92-1, Stanford University.

Gruber, T. (1993). A Translation Approach to Portable Ontology Specifications. *Knowledge Acquisition,* 5(2), 199–220.

Issa, R. R. and Mutis, I. (2015). *Ontology in the AEC Industry*. ASCE Press.

Niu, J, and Issa, R. R. (2013). Conceptualizing methodology for building an ontology for construction claim knowledge." *ASCE IWCCE*, June 23-25, L.A., CA, 492-499.

Noy, N. F. and McGuiness, D. L. (2001). *Ontology Development 101: A Guide to Creating Your First Ontology*. Stanford University, Stanford, CA.

Pauwels, P., Terkaj, W., and Beetz, J. (2015). *EXPRESS-to-OWL Conversion Routine*. Proposed Recommendation, Final. buildingSMART International.

Schevers, H. and Drogemuller, R. (2005). Converting the Industry Foundation Classes to the Web Ontology Language. *Semantics, Knowledge and Grid*, Nov. 27-29.

Taye, M. (2011). Web-Based Ontology Languages and its Based Description Logics. *International Journal of ACM Jordan*, 2(1), 1-9.

Uschold, M., and Gruninger, M. (1996). Ontologies: Principles, Methods and Applications. *Knowl. Eng. Rev.*, 11(02), 93-155.

Venugopal, M., Eastman, C., and Teizer, J. (2012). An Ontological Approach to Building Information Model Exchanges in the Precast/Pre-Stressed Concrete Industry. *Construction Research Congress*.

W3C OWL Working Group. (2012). *OWL 2 Web Ontology Language Document Overview* (2d Edition). < https://www.w3.org/TR/owl2-overview/> [8, Nov. 2016].

Zhang, J. and El-Diraby, T. E. (2009). SSWP: A Social Semantic Web Portal for Effective Communication in Construction. *Journal of Computers*, 4(4), 330-337.

Zhang, S., Boukamp, F., and Teizer, J. (2015). Ontology-Based Semantic Modeling of Construction Safety Knowledge: Towards Automated Safety Planning for Job Hazard Analysis (JHA)." *Automation in Construction,* 52, 29-41.