ENGAGED ONTOLOGY DEVELOPMENT TO BRIDGE FRAGMENTED DIGITAL REALITIES

Ramon ter Huurne[1], Léon Olde Scholtenhuis and André Dorée

Department of Construction Management and Engineering, University of Twente, PO BOX 217, 7500AE, Enschede, The Netherlands

Fragmented knowledge bases in construction are the result of varying views on how to capture and represent knowledge. For asset management of utility infrastructure, knowledge is often either stored implicitly or in heterogeneously structured data models. Consequently, it becomes difficult to connect and implement the different data models coming from numerous individual asset owners. To unify these various knowledge representations of utility infrastructure data, we adopt the computer science concept of the 'ontology'. By using a design science-inspired research approach, we demonstrate the development of an ontology with the intent to cope better with the fragmentation in the utility infrastructure sector. We further demonstrate that the co-development of an ontology with domain professionals may emerge into a shared conceptualisation of the domain. Based on this process, we claim that engaged ontology development can play an important role in bridging the fragmentation between digital realities, in turn making digital modelling concepts such as digital twins more likely to become adopted by the utility construction sector. Future work is required to assess the impact of the ontology once applied on a larger scale.

Keywords: digital twin; fragmentation; representation; ontology; utility infrastructure

INTRODUCTION

The construction industry increasingly digitizes the life cycle of construction assets by defining concepts, attributes, and their relations (El-Diraby and Osman 2011). Digitisation is typically done through the creation of virtual representations of physical counterparts in, for example, Building Information Modelling (BIM) and Geospatial Information Systems (GIS) environments. Digitisation is further supported by the rapid development of technologies like artificial intelligence (AI), big data, the internet of things (IoT), cloud computing, wireless sensor networks, and the fifth-generation cellular network (5G) (Lu et al., 2015; Syafrudin et al., 2018). Altogether, these digital advancements nowadays drive state-of-the-art engineering and problem-solving in the construction industry.

A yet to explore issue in the construction context is the fragmentation of knowledge bases. Prior research shows that nations, organisations and even individuals may portray varying views on how to capture and represent the knowledge relevant to a construction asset's life cycle (Azhar 2011; Ter Huurne and Olde Scholtenhuis 2018;

[1] r.b.a.terhuurne@utwente.nl

ter Huurne, R, Olde Scholtenhuis, L and Dorée, A (2022) Engaged Ontology Development to Bridge Fragmented Digital Realities *In:* Tutesigensi, A and Neilson, C J (Eds) *Proceedings of the 38th Annual ARCOM Conference,* 5-7 September 2022, Glasgow, UK, Association of Researchers in Construction Management, 328-337.

Voordijk et al., 2022). Such knowledge is namely either implicit or stored in heterogeneously structured data models (Figure 1). Consequently, it becomes difficult to align and connect the different data models coming from numerous individual asset owners. Such distinct and self-centred knowledge bases confuse, fragment, and ultimately delimit collaborative asset management practices (Ter Huurne and Olde Scholtenhuis 2018). Accordingly, the literature argues that uniformity of knowledge bases is necessary before the alignment of digital practices and their data models can be achieved (Turk 2001). This insight stresses the relevance of defining a shared domain understanding in the current - and most likely expanding - realm of digital construction practices.

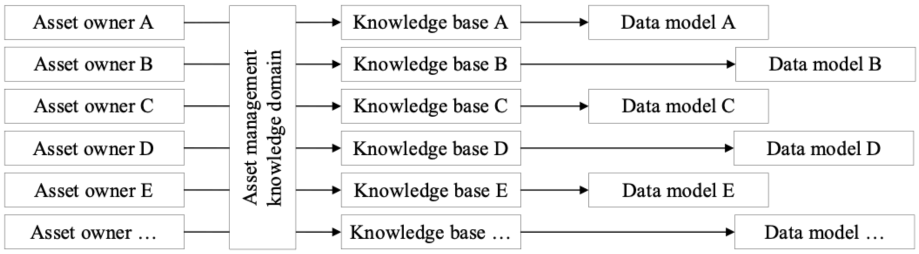

Figure 1: Heterogenous data models in construction describing the same knowledge domain

This study explores the concept of an 'ontology'. Ontologies provide the metadata that describes domain knowledge, bridging any varieties that may exist between distinct knowledge bases and their subsequent data models. This allows practitioners to understand both the content and structure of a knowledge domain (El-Diraby and Osman 2011). One sector in the construction industry that exhibits many varying data models - yet of which no one intends to be an ontology - is the utility sector. Utilities are the (typically below-surface) cables and pipes that transport services like water, gas, electricity, and telecommunication. Data models in this sector differ greatly between application domains, utility disciplines, geographical districts, and organisations.

The objective of this study is twofold. First, we demonstrate the co-development of an ontology with domain professionals with the intent to cope better with the fragmentation in the utility infrastructure sector. Specifically, we explain our efforts in attempting to conceptualize a shared domain understanding by adopting a design science-inspired research approach. Second, we reflect upon the ontology's implications regarding the future digitisation of construction environments and discuss its potential impact on the design context. Therewith, this study contributes to the construction management literature with an exploratory study of an ontology development process and expands our understanding of how ontologies may play a role in unifying fragmented data models in future digitisation efforts.

Fragmented digital practices complicate the forming of shared conceptualisations. Prior research provides evidence for the co-existence of multiple distinctive knowledge bases as the result of varying views on how to capture and represent domain knowledge (Azhar 2011; Ter Huurne and Olde Scholtenhuis 2018; Voordijk et al., 2022). This is likely to result in many heterogeneous structured data models. In turn, asset owners use these data models in their adopted information technology. Consequently, little uniformity in knowledge representation exists. This hampers the sharing of knowledge, information and data, whereas the concepts modelled are more prone to misunderstanding and misinterpretation (Turk 2001). Therefore, alignment and connection between the distinct knowledge bases and their data models become difficult. Altogether, this hampers collaborative engineering practices and complicates software interoperability (Lu et al., 2015).

This notion of 'varied knowledge representation' is expected to further increase in the modernizing and increasingly digital and virtual construction environment. In response to a greater focus on the operations and maintenance of construction assets, asset owners now increasingly enrich their existing knowledge bases and data models. This is illustrated by the recent introduction of the 'digital twin' in the construction domain (Opoku et al., 2021). Digital twins "facilitate the means to monitor, understand, and optimize the functions of physical entities, living as well as non-living, by enabling a seamless transmission of data between the physical and the virtual world" (El Saddik 2018). Digital twins are used to describe and monitor an asset's entire lifecycle via the use of sensor data, analytical and predictive models, and visualisations. The insights derived from the digital twin are then used in real-life to make decisions about the physical asset. Although still in their nascent phase, digital twins are considered amongst the most promising advancements to further modernize digital and virtual construction environments and their processes (Khajavi et al., 2019).

Based on these insights, knowledge needs to be unified to achieve the collaboration benefits of data models before 'going digital' (Gustavsson et al., 2012). This can provide a solid basis for the sharing and exchange of data models like digital twins. Literature on computer sciences advocates the use of ontologies to represent knowledge domains. Ontologies describe the world as seen by a group of people at a certain time according to a school of thought that is based on a set of fundamental propositions or world views (El-Diraby and Osman 2011). An ontology can be defined as formal and explicit specifications of shared conceptualisations (Sure, Staab, and Studer 2009). Conceptualisation refers to the universe of discourse. Shared refers to the multiple views an ontology should be able to represent. Formal and explicit refers to the fact that the concepts within the ontology should be described in a clear computer-interpretable format. Once adopted and shared amongst practitioners, ontologies are thus used to represent knowledge in a unified, simplified, and consistent way.

Brachman and Levesque (2004) describe the logic behind such knowledge representation as "the field of study concerned with using formal symbols to represent a collection of propositions believed by some putative agent". Knowledge representations are thus the internal representations of such an agent (Jakus et al., 2013). This means that representing knowledge through ontologies requires thought about phenomenology - a branch of philosophy that deals with how to take things for what they are and what it means 'to be' - and hermeneutics - a branch of philosophy focussing on interpretation. Intention and interpretation are relevant when capturing 'realities' because their meanings can be shaped both by the authors and users of ontologies (Turk 2001). Ontologies thus rely on consensus amongst the domain professionals using it to enable shared conceptualisations of the knowledge it captures and represents.

To date, ontology development efforts to bridge fragmented realities have received limited attention in construction management literature. Considering this research gap, this study showcases the development of an ontology with the intent to cope better with the fragmentation in one construction sector in particular: the utility sector. In this development process, we explain our efforts in attempting to create a shared domain understanding by adopting a design science-inspired research approach. We then zoom out and discuss the ontology's impact on its design context, while providing an outlook on the role of ontologies in the future digitisation of construction environments.

METHOD

To develop the ontology and assess its impact on its design context, we adopted Hevner's (2007) three-cycle view of design science (Figure 2). We chose design science as our research approach since it explicitly focuses on creating knowledge and understanding of a problem domain by building and designing an artefact. Furthermore, this approach supports both the development (combination of the rigor and design cycle) and the assessment of the impact of the ontology on its design context and application domain (the relevance cycle).

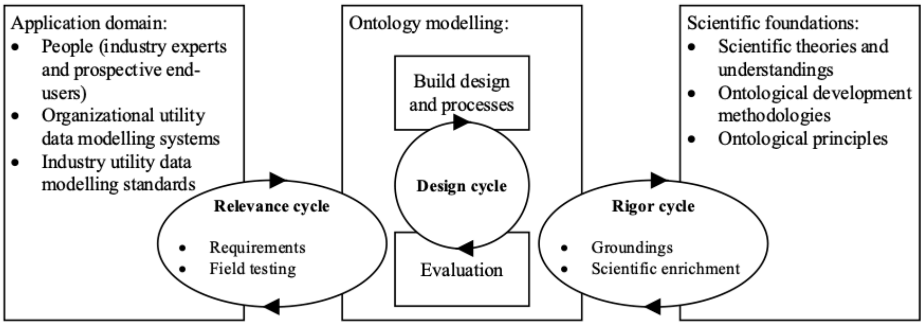

Figure 2: Adopted design science research approach (adapted from Hevner, 2007)

To build the ontology, we adopted a hybrid approach of existing ontology development principles and methodologies as scientific foundations. In specific, elements of the studies by Corcho *et al.* (2003), Noy and McGuinnes (2001), López *et al.* (1999), Sure *et al.* (2004), Gasevic *et al*., (2009), Jakus *et al.* (2013), and Pinto *et al.* (2009) were used and integrated into the design science cycle. We investigated the application domain of the ontology, defined the ontology's requirements, designed, and build the ontology, evaluated the ontology, and performed a field test. This entire process was conducted iteratively over a two-year timespan.

To align the potential varying intentions and interpretations of domain professionals, the ontology was co-developed in close collaboration with industry experts and prospective end-users. They were mainly involved in the requirements engineering and evaluation phase of the ontology. Specifically, over twenty industry meetings were held, either with groups or individuals. Further, utility data models (including models such as CityGML Utility Network and INSPIRE), existing utility design guidelines and real-life sources of domain data (obtained through observational case studies of the digital modelling practices of twelve major utility owners) were inspected to get an understanding of the knowledge represented in the distinctive knowledge bases.

The ontology was modelled by using a hybrid approach of top-down and bottom-up modelling. A top-down approach starts building with the most generic concepts. A bottom-up approach starts building with the most specific concepts. The proposed hybrid approach has the benefit of grasping the generic concepts of the domain, while at the same time being able to connect these with the detailed aspects of real-life practices. The concepts of the ontology were modelled in the Unified Modelling

Language (UML) which is an often-applied language in class modelling due to its graphical notations. We further adopted the typology of Gasevic *et al. (*2009) to communicate and describe the ontology design. This topology comprises the three main elements of an ontology, being taxonomy and hierarchy, vocabulary and terms, and semantics. The taxonomy and hierarchy refer to the hierarchical categorisation of the concepts within the ontology. The vocabulary refers to the set of terms and names that are used in the subject area captured by the ontology. Semantics refers to the linguistic meaning of these applied terms and names.

Given the iterative nature of the design cycle, simultaneous to its development evaluation of the ontology took place. Four evaluation techniques were applied: assessment against sources of domain data, assessment against competency questions, assessment against modelling rules, and assessment against end-user and expert input. The ontology was also implemented as a field test in a simulated utility operations and maintenance case. During this field test, we assessed once more whether the ontology could satisfy the 'competency questions' (questions the ontology should be able to provide the knowledge for to answer). Altogether, the various evaluation measures helped to assess the ontology against eight ontology evaluation criteria: accuracy, adaptability, clarity, completeness, computational efficiency, conciseness, consistency, and organisational fitness. Ultimately, eight versions were developed before the ontology was considered capable of satisfying all evaluation measures, and therewith, creating a shared domain understanding.

To assess the ontology's impact on its design context, a semi-structured plenary session was held with the prospective-end users and industry experts. The ontology and its intended aim were presented in combination with a demonstration of the field test. Specifically, the participants were asked to express their thoughts about how the ontology would implicate their current digital modelling efforts, as well as their current asset management practices.

An Ontology for Utility Asset Modelling

The ontological model presented in this study is an empirically grounded ontology whose targeted use lies in the management of the operations and maintenance of utilities during their lifespan. The ontology applies to the following utility disciplines: electricity, oil, gas, chemicals, sewage, water, thermal and telecommunication. The use of the ontology has no geographic boundary but is based on utility networks of developed countries. The ontology models utility networks, their subnetworks, and their superordinate networks. Utilities can be modelled both in two and three dimensions. The next sections provide an overview of the developed ontology's design. The complete ontological model and its accompanying documentation are available from the corresponding first author upon reasonable request.

The ontology describes the entire chain of knowledge represented in a utility asset management context. The model describes the interplay between actors, projects and processes, physical objects and knowledge items, and how these are described through their spatial characteristics, functions, attributes and performances. This is visualized in Figure 3.

Concepts and their attributes in the model are related to one another through different types of relations. We used the relation types of UML to represent whole-part, parent-child, and one-to-one relationships. For each of the concepts as presented in Figure 3, the ontological model allows the representation of a variety of instances. We illustrate some of its instances in Table 1.

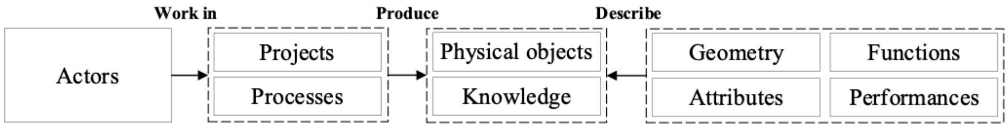

Figure 3: Concepts modelled in the ontology

Table 1: An example of instances from the ontology

	Example instances from the ontology
Actors	Owner, manufacturer, contractor
Projects	Rehabilitation of utilities, maintenance of utilities, replacement of utilities
Processes	Inspection, surveillance, corrective maintenance
Physical objects	Distribution component, controller component, protective shell
Knowledge items	Maintenance report, localization guidelines
Spatial characteristics	Geometry, construction area, project boundary
Functions	Network function, network usage, object roles
Attributes	Object material, electricity operating voltage, network status
Performances	Safety performance, engineering performance, financial performance

Table 2 illustrates how these relations were applied.

Table 2: An example of relations from the ontological model

	Instance 1	Relationship description	Instance 2
Whole-part	SubNetwork	is part of	Network
Parent-child	ElectricityCable	is a	DistributionComponent
One-to-one	PerformanceProperties	is associated to	NetworkFeature

The vocabulary and terms of the concepts within the ontology were carefully chosen in close collaboration with end-users and domain experts. An additional catalogue of all used terms was also developed to prevent semantic issues such as differing interpretations during the sharing and exchange of information.

Additionality, the ontology was implemented in a simulated utility operations and maintenance field test. Specifically, two utility types were modelled in a renovation project of street works on a university campus. We posed competency questions to verify whether the knowledge captured and represented by the ontology was deemed sufficient for its design context and application domain. Examples of asked questions are: "What is the state of operation of a utility network?", "What is the nominal flow of a commodity through a distribution line?", and "When was the last maintenance activity performed?". The ontology was able to satisfy the requirements of the engineering tasks required and assessed as relevant by the involved domain professionals.

The developed ontology was positioned as an intermediate that provides the metadata for all those concepts modelled in the current distinctive knowledge bases of utility asset owners (Figure 4). Compared to the the-before-situation (Figure 1), the introduction of an ontology does not necessarily mean knowledge bases or data models are integrated into one singular model. Instead, the developed ontology provides a basis for the sharing and exchange of, in this regard, utility asset information. This allows heterogeneous data models to share and exchange their information with one another, as to how they capture and represent their knowledge is unified through the ontology.

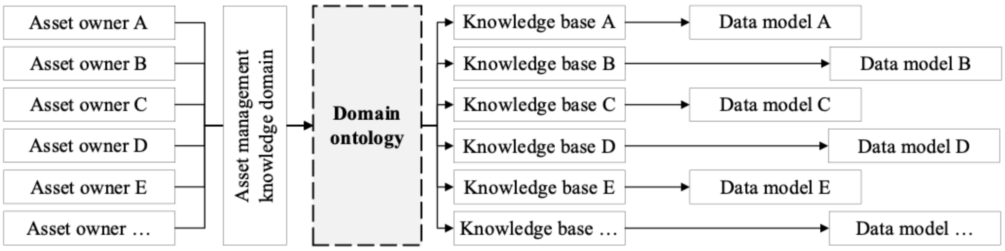

Figure 4: The ontology as an intermediate to heterogeneous data models

When assessing the impact of the ontology on its design context, the industry experts and prospective end-users first acknowledged that the ontology could stimulate the sharing and exchange of information in digital and virtual construction environments. Second, the ontology was considered helpful in the future co-development of digital models by multiple asset owners. Third and last, the richness of the knowledge represented by the ontology was considered supportive for data-driven engineering and smart reasoning. Especially for those asset owners currently working with less rich and digitized asset data models, the ontology was deemed helpful in further digitisation efforts. In the next section, we reflect upon the role ontologies may play in future digitisation efforts of construction environments.

FINDINGS

This study demonstrated the development of an ontology with the intent to cope better with the fragmentation of digital practices in the utility infrastructure sector. We demonstrated, by adopting a design science-inspired approach (Hevner 2007), the efforts undertaken to establish a shared understanding of the knowledge domain. Subsequently, we reflected with domain professionals on the impact of the ontology on the design context. Based on this process, the contributions of this study to the construction management literature are twofold.

First, this study demonstrates that an engaged ontology development process can emerge into shared conceptualisations of a knowledge domain. The perceptions of human beings - including the modellers of ontologies - are incomplete and bounded by rationality, resulting in various viewpoints on the to-be-modelled reality (Olde Scholtenhuis and Hartmann 2015). Therefore, developing an ontology that meets everyone's perception of this reality is considered a highly complex task (Turk 2001). In this study, we experienced that co-developing the ontology with its prospective end-users and industry experts stimulates a consensus-seeking behaviour that helps to align different realities. Based on this, we claim that the co-development of ontologies plays an important role in establishing shared conceptualisations, and in turn, in bridging distinct digital realities. However, future work is required to assess the ontology's role once applied on a larger scale.

Second, this study argues that the co-development of an ontology with domain professionals may form a solid basis for the future co-development and adoption of exchangeable data models. New data models continuously enter the construction environment in, for example, the form of the recently emerging digital twin (Opoku *et al.,* 2021). In this context of a modernizing and increasingly digital and virtual construction environment, where fragmentation is expected to only further increase, we claim that co-developed ontologies and data models help in bridging fragmented data realities. In turn, such co-developed models may improve inter-organisational communication, cooperation, and coordination (Adriaanse *et al.,* 2010; Peansupap and

Walker 2005). However, also here, future work is required to assess the ontology's impact on the development and adoption of co-developed data models once applied on a larger scale.

This study further provides opportunities and recommendations for future work. First, we argue that an ontology development process itself is not sufficient in a continuously evolving industry such as construction. Knowledge domains and domain professionals and their preferences and perceptions may change over time. This implies maintenance and alignment to these altering knowledge domains are needed to ensure developed ontologies are still capable of establishing a shared conceptualisation of the domain. Based on this, we emphasize that standardisation of knowledge representation in the form of ontologies is most likely an ongoing effort. We urge scholars and practitioners to take this notion into account in their future adoption or studies of ontologies.

Second, the ontology in this study was implemented as a field test in a single simulated case study. Although this test did provide evidence of the ontology's capability of generalizing the necessary knowledge of two utility disciplines in a specific context of use, further work is required to investigate whether the ontology can generalise the entire knowledge domain it was intended for to capture. As explained by Gruber (1995), the generic nature of ontologies is a key requirement of ontologies to enable the representation of entire knowledge domains. Yet, the dynamics of real-life environments between utility disciplines may significantly differ from simulated environments. This could display deficits in the current design of the ontology, requiring additional design cycles of most likely both the design and relevance cycle (Hevner 2007).

CONCLUSIONS

This study demonstrated the development of an ontology with the intent to cope better with the fragmentation of digital realities in the utility infrastructure sector. By adopting a design science-inspired research approach, we explained our efforts to establish a shared understanding of the knowledge domain. Via an engaged ontology development process in close collaboration with prospective end-users and industry experts, we abstracted concepts from distinctive knowledge bases and industry standards. Through a combination of multiple evaluation measures and a partial field test of the ontology in a utility asset management case, the ontology was considered complete and suited to its design context and application domain.

Contributions of this study are twofold. First, we provide to the construction management literature an exploratory study of an engaged development process. We demonstrate that such co-development of an ontology with domain professionals is more likely to emerge into a shared conceptualisation of the domain. This can form the solid basis of exchangeable digital models of public space. Second, we contribute to construction management literature the notion that engaged ontology development may play an important role in bridging fragmented realities in digital and virtual construction environments. This, in turn, makes asset management concepts more likely to become shared, exchanged, and adopted by the utility construction sector, engaging collaborative asset management practices. This insight stresses the relevance of an ontology in the modernizing construction sector, where the development and application of digital models like digital twins will only increase over time. However, future work is required to explore the impact of the ontology once applied on a larger scale.

Data Availability

Some or all data that support the findings of this study are available from the corresponding first author upon reasonable request. This includes the complete ontological model and parts of the supporting documents of the case studies and industry meetings.

REFERENCES

Azhar, S (2011) Building information modeling (BIM) Trends, benefits, risks and challenges for the AEC industry, *Leadership and Management in Engineering*, **11**(3) 241-252.

Brachman, R J and Levesque, H J (2004) *Knowledge Representation and Reasoning*, Amsterdam: Morgan Kaufmann.

Corcho, O, Fernández-López, M and Gómez-Pérez, A (2003) Methodologies, tools and languages for building ontologies Where is their meeting point? *Data and Knowledge Engineering*, **46**(1), 41-64.

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.

Gasevic, D, Djuric, D and Devedzic, V (2009) *Model Driven Engineering and Ontology Development*, New York: Springer.

Gruber, T R (1995) Toward principles for the design of ontologies used for knowledge sharing? *International Journal of Human-Computer Studies*, **43**(5-6), 907-928.

Gustavsson, T K, Samuelson, O and Wikforss, Ö (2012) Organizing IT in construction: Present state and future challenges in Sweden, *Journal of Information Technology in Construction (ITcon)*, **17**(33), 520-534.

Hevner, A R (2007) A three cycle view of design science research, *Scandinavian Journal of Information Systems*, **19**(2), 87-92.

Ter Huurne, R B A and Olde Scholtenhuis, L (2018) *Digitisation for Integration: Fragmented Realities in the Utility Sector*, Working Paper, Association of Researchers in Construction Management.

Jakus, G, Milutinović, V, Omerović, S and Tomažič, S (2013) *Concepts, Ontologies, and Knowledge Representation*, Springer Briefs in Computer Science, i-vi.

Khajavi, S H, Motlagh, N H, Jaribion, A, Werner, LC and Holmström, J, (2019) Digital twin: vision, benefits, boundaries, and creation for buildings, *IEEE Access*, **7**, 147406-147419.

López, M F, Gómez-Pérez, A, Sierra J P and Sierra, A P (1999) Building a chemical ontology using methontology and the ontology design environment, *IEEE Intelligent Systems and Their Applications*, **14**(1), 37-46.

Lu, Y, Li, Y, Skibniewski, M, Wu, Z Wang, R and Le, Y (2015) Information and communication technology applications in architecture, engineering and construction organisations: A 15-year review, *Journal of Management in Engineering*, **31**(1), A4014010.

Noy, N and McGuinnes, D (2001) *Ontology Development 101: A Guide to Creating Your First Ontology*, Stanford, CA: Knowledge Systems Laboratory.

Olde Scholtenhuis, L L and Hartmann, T (2015) Fieldwork-based method for end-user engagement in domain ontology development, In: Raja R A Issa and Ivan Mutis (Eds) *Ontology in the AEC Industry*, American Society of Civil Engineers,149-167.

Opoku, D G J, Perera, S, Osei-Kyei, R and Rashidi, M (2021) Digital twin application in the construction industry: A literature review, *Journal of Building Engineering*, **40**, 102726.

Pinto, H, Tempich, C and Staab, S (2009) Ontology engineering and evolution in a distributing world using DILIGENT, In: S Staab and R Studer (Eds) *Handbook on Ontologies*, Berlin, Heidelberg: Springer.

El Saddik, A (2018) Digital Twins: The convergence of multimedia technologies, *IEEE Multimedia*, **25**(2), 87-92.

Sure, Y, Staab, S and Studer, R (2009) Handbook on ontologies, *Handbook on Ontologies*. Cham: Springer, 135-152.

Sure, Y., Staab, S and Studer, R (2004) On-to-knowledge methodology (OTKM), In: *Handbook on Ontologies*, Springer, Berlin, Heidelberg, 117-132

Syafrudin, M, Alfian, G, Fitriyani, N L and Rhee, J (2018) Performance analysis of IoT-based sensor, big data processing and machine learning model for real-time monitoring system in automotive manufacturing, *Sensors*, **18**(9), 2946.

Turk, Ž (2001) Phenomenologial foundations of conceptual product modelling in architecture, engineering and construction, *Artificial Intelligence in Engineering*, **15**(2), 83-92.

Voordijk, H, ter Huurne, R and Olde Scholtenhuis, L L (2022) Adoption and use of ontologies in the utility sector: A technological mediation perspective, *Journal of Management in Engineering*, **38**(1), 05021016.