Semantic enrichment of Building and City Information Models: a ten-year review

Fan Xue, Liupengfei Wu*, and Weisheng Lu

Department of Real Estate and Construction Management, The University of Hong Kong, Hong Kong, China

Abstract:

Building Information Models (BIMs) and City Information Models (CIMs) have flourished in building and urban studies independently over the past decade. Semantic enrichment is an indispensable process that adds new semantics such as geometric, non-geometric, and topological information into existing BIMs or CIMs to enable multidisciplinary applications in fields such as construction management, geoinformatics, and urban planning. These two paths are now coming to a juncture for integration and juxtaposition. However, a critical review of the semantic enrichment of BIM and CIM is missing in the literature. This research aims to probe into semantic enrichment by comparing its similarities and differences between BIM and CIM over a ten-year time span. The research methods include establishing a uniform conceptual model, and sourcing and analyzing 44 pertinent cases in the literature. The findings plot the terminologies, methods, scopes, and trends for the semantic enrichment approaches in the two domains. With the increasing availability of data sources, algorithms, and computing power, they cross the border to enter each other's domain. Future research will likely gain new momentums from the demands of value-added applications, development of remote sensing devices, intelligent data processing algorithms, interoperability between BIM and CIM software platforms, and emerging technologies such as big data analytics.

Keywords: Building Information Model; City Information Model; Geographic Information System; Semantic enrichment; Smart city; Intelligent building

1. Background

In the architecture, engineering, and construction (AEC) industry, information is the common keyword of the Building Information Model (BIM) and City Information Model (CIM). The concept of BIM is coined to refer to a digital representation of a facility's physical and functional characteristics; it is a shared knowledge resource for information about a facility, forming a reliable information foundation for decisions throughout its life cycle (NIBS, 2015).

Building information modeling is a nomenclatural term used to refer to a family of technologies and related practices used to represent and manage information used and created for the process of designing, constructing, and operating buildings (Davies and Harty, 2011). BIM has roots that can be traced back to the production and manufacturing industry, where designers tend to develop a digital model of the product for optimization and prototyping before it is mass-produced (Eastman et al., 2011).

The concept of CIM has emerged in parallel to refer to a system of urban elements and environments represented in 2D and 3D symbols (Stojanovski, 2013). The gerund, city information modeling, is used to represent the technologies and practices used to develop a city information model and harness its power for various smart applications, e.g., navigations, transportation, urban climate, and urban morphology (Xu et al., 2014), which are normally placed under the umbrella of a smart city. 3D city models as the baseline of CIM has a much longer history if one considers its roots in Geographic Information System (GIS), a computer system for capturing, storing, checking, and displaying data related to positions on the Earth's surface (NGS, 2012). According to Julin et al. (2018), such 3D city models are typically built by merging photogrammetry and laser scanning data with GIS data.

The essential difference between CIM and GIS is the scope and form of the "I," i.e., information. The information in GIS is explicitly referenced according to the scope of the Earth's surface and often managed in forms of 'layers' (Goodchild, 1991; Lu et al., 2018). In contrast, a CIM's scope focuses on urban areas but includes more types of non-GIS information, e.g., BIM, Light Detection and Ranging (LiDAR), inhabitant behaviors, and city energy simulation, and managed in the form of cross-reference relation 'networks'—like a BIM (Xu et al., 2014; Liu et al., 2017). Thus, in this paper, at the risk of oversimplification, a 3D city model in GIS is treated as a specific type of CIM. In recent years, other buzzwords such as the 'Cyber-Physical System' and 'digital twin' are emerging (Xue et al., 2020) from agriculture, defense, energy, healthcare, transportation, and manufacturing systems to represent the virtual, digital models of these systems. In non-technical language, one can perceive the 'real-time' BIM and CIM as the cyber replica or digital twin of a building and a city, respectively.

Characteristics of BIM and CIM information are essential for differentiating and understanding these buzzwords. Information is the shared key to enable their designated applications, though the scope and organization are different. As listed in Table 1, BIM represents building elements (e.g., beams and columns) with building-level information (e.g., geometric details, energy usage data, and life cycle costing information) to meet AEC industrial needs (Karan et al., 2016). The taxonomy of BIM information distinguishes geometric, semantic, and topological types (Pratt-Hartmann, 2004; Schlueter and Thesseling, 2009). Geometric information directly relates to the shapes and forms of facilities, whereas semantic information captures their intrinsic properties (e.g., functionality), and topological information gives the relationships among these objects. In comparison, CIM involves a three-dimensional model that connects buildings and other urban information sources (e.g., roads, public spaces, street lights, and even people on the street) in a city (Xu et al., 2014).

Table 1. Scope and organization of information in BIM, and CIM

	Dimension of geometry	Scope of information	Organization of relational information
BIM	3D	Up to building level (e.g., rooms, buildings, and bridges)	Explicit via relations between 3D components (e.g., joints)
CIM	3D	Up to city level (e.g., metro shop, park bench, city hall, street greenery, and landscape)	Geo-referenced 3D City elements

The term 'semantic enrichment' has been defined in several ways. For example, Clarke and Harley (2014) defined semantic enrichment as the process of adding a layer of topic metadata to content so that the machine can understand and connect with it. Later, Sacks et al. (2017) pointed out that semantic enrichment aims to add absent or new information, compile as-is or as-built conditions, and extend the schemas to existing BIMs/CIMs. Bloch and Sacks (2018) further define semantic enrichment as an indispensable process that encompasses the classification of objects, aggregation and grouping, unique identification, completion of missing objects, and reconstruction of occluded objects in the case of application to models. This study adopts Bloch and Sacks' (2018) definition given its comprehensiveness. The exterior geometry of a building or a city might be fixed from the outset; semantic information, however, will update along the way when they are developed, utilized, and updated. These updated semantics, e.g., as-built or as-is conditions, should be properly captured in BIM or CIM. Over the past decade, semantics enrichment has been developed in BIM and CIM independently, and it remains so. Nevertheless, scholars are increasingly interested in seeing more details of buildings in a CIM. Many review papers have reflected the up-to-date processing of building and city data to create new BIMs/CIMs, e.g., for buildings (Volk et al., 2014), indoor environments (Kang et al., 2020), roads (Wang et al., 2016), and applications such as building energy (Reinhart and Davila, 2016) and urban ecology (Berling-Wolff and Wu, 2004). However, a study to juxtapose the research niche of enriching existing models is absent and thus much desired. This study is particularly opportune because BIM and CIM, previously dominated by giant software vendors, are starting to join hands with each other (Esri, 2019). Therefore, it would be intriguing to review what has been researched in BIM and CIM semantic enrichment and to extrapolate what will likely happen under the general trend of BIM and CIM integration.

The primary aim of this research is to compare the semantic enrichment studies in BIM and CIM in the past 10 years for outlining the research trends and prospects. The goal is achieved by deploying an analytical framework that includes a conceptual model, a set of comparative studies, and future direction predictions. The rest of this article consists of four sections.

Section 2 describes the research methodology. Section 3 compares the semantic enrichment of BIM and CIM. Section 4 gives the trends and prospects of semantic enrichment of BIM and CIM. Findings are drawn in Section 5.

2. Research methods

The primary research method adopted in this paper is a comparative archival study. It follows a four-step methodology to conduct the study: (1) Developing a guiding conceptual model, (2) Data collection by published case selection and decoding, (3) Comparative study of semantics enrichment for BIM and CIM, and (4) Extrapolation of the future development.

2.1 Guiding conceptual model

A guiding conceptual model was developed to facilitate data collection and analysis. As shown in Figure 1, the conceptual model includes six interrelated components, namely, (1) The subject facility, (2) Baseline model, (3) Sources of new semantics, (4) Semantic enrichment method, (5) Enriched model, and (6) Application. For a given facility, semantic enrichment is the process that extracts new semantics from the new semantics sources and adds them to the baseline BIM or CIM so that the updated model has a richer digital representation of the facility.

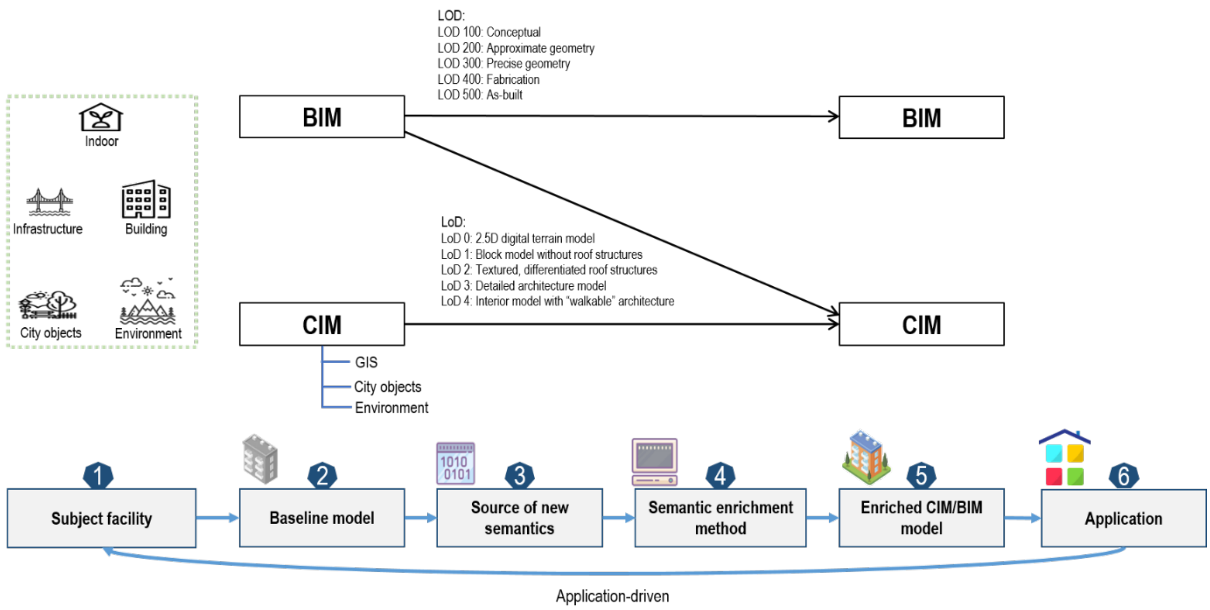

Figure 1. A guiding conceptual model of semantic enrichment in BIM and CIM

The matrices for measuring the new semantic richness of BIM and CIM are the Level of Development (LOD) and the Level of Detail (LoD), respectively, as shown in Figure 1. LOD for BIM is defined on the temporal development of buildings, where LOD 100 is the conceptual design without shape information or geometric representation and LOD 500 is the as-is BIM with verified representation in terms of size, shape, location, quantity, and orientation (Boton et al., 2015; Lu et al., 2018). There are intermediate levels, such as LOD 350 and LOD 450 (Chen et al., 2015; Abualdenien and Borrmann, 2019). So, the semantics in a BIM can be enriched by designs, plans, surveying, documents, inspections, and operations along the building life cycle. In this view, BIM with replenishing semantics can be a potential baseline model for enriching CIM. In contrast, LoD, which originated in the CityGML (City Geography Markup Language) standard (Chen et al., 2020), emphasizes the spatial detailing in CIM. For buildings, LoD 0 denotes 2.5D block models, and LoD 4 means a 'walkable' model with building interiors. For example, a well-designed BIM in the tendering phase can be LoD 4.3 for CityGML, but only LOD 200 for BIM. In non-building facilities such as vegetation, LoD 0 represents crown projection, whereas LOD 4 gives third-level branches (Ortega-Córdova, 2018). There also can be intermediate LoD levels, such as LoD 1.1 and LoD 3.4 in Biljecki et al. (2016b).

The six components in the conceptual model can be further described with categorical values, as listed in Table 2. The subject facility can be indoor space and furniture, buildings, infrastructure, city objects, or even the environment. Traditionally, there are two silos, i.e., building level and city level, which do not talk to each other. A baseline model is a given, normally preliminary, digital representation of the facility. It can be a BIM or CIM. Here, a CIM is a fully 3D city model or a 2D GIS model with certain city objects linked to their 3D models. Semantic enrichment in the BIM and CIM integration scenario focuses on the interoperability problems between the two. The minor descriptor in Table 2 shows that the baseline model could be in an open or commercially protected format.

Table 2. Descriptors of the six components in the conceptual model of semantic enrichment

Component	Descriptor			Examples
	Major	Minor	Code	
Subject facility	Indoor		IN	Furniture, room space
	Building		BLD	Facades, structural elements
	Infrastructure		INFR	Bridge, tunnel
	City objects		CO	Lamppost, park
	Environment		RE	Vegetation
Baseline model	BIM	Open	BIM-O	IFC
		Commercial	BIM-C	Autodesk Revit
	CIM	Open	CIM-O	CityGML
		Commercial	CIM-C	Esri ArcGIS
Source of new semantics	Data	Raster (2D/3D)	D-R	Digital pictures, radar images, voxels
		Vector (2D/3D)	D-V	LiDAR points, lines, polygons
		Tabular	D-T	Descriptive data stored in rows and columns
	Existing components	BIM	C-BIM	BIM components
		CIM	C-CIM	3D city objects
	Ontology	Formal	O-F	Complete ontology definition
		Ad hoc rules	O-R	Case-specific rules
		Spatial relations	O-S	Geological relationship
Automated method	Semantic reasoning		Rsn	Rule-based reasoning
	Semantic registration		Rgst	ICP algorithm
	Semantic segmentation		Sgmt	Supervised
Enriched CIM/BIM	Format	BIM		
		CIM		

	New semantics	Geometric	NS-G	New openings on a wall
		Non-geometric	NS-N	A group of a furniture
		Both	NS-B	
Application	Pre-operation		PO	Construction progress control
	Operation & Maintenance		OM	Facility maintenance
	Urban environment		UB	Urban planning
	Energy		EN	Energy performance simulation
	Unspecified case		UNS	No application mentioned

The source of new semantic can be classified into data, existing components from BIM/CIM, and ontology. The data can be further divided into raster data, vector data, and tabular data. Raster data (GISGeography, 2015) could be 2D or 3D, as represented in satellite photos, radar images, and voxels. Vector data consists of LiDAR points, lines, and polygons (Koch and Heipke, 2006). Tabular data is descriptive data stored in rows and columns in the database and can be linked to spatial data. The second category of new semantics sources is the components stored in various BIM or CIM libraries. They can be produced by manufacturers, software vendors, or other interested modelers for private or open access. The ontology sources contain formal ontology, *ad hoc* rules, and geological relationships. Unlike data and existing BIM/CIM components, ontology is an often-neglected source of new semantics, particularly position and typology information.

In general, there are three types of semantic enrichment methods:

(i) Semantic reasoning based on pre-determined rules or ontology,
(ii) Semantic registration based on many-round trial-and-error searches, and
(iii) Semantic segmentation based on an annotated training data set.

It needs to be pointed out that we excluded manual semantic annotation methods in this study in that they are labor-intensive, time-consuming, and error-prone (Tang et al., 2010; Brilakis, 2010; Xiong et al., 2013). This is particularly true when it comes to larger sites or city level. The literature search later in this paper also suggests that pure manual semantic methods are rarely reported, although they are indispensable to fine-tune or correct the models generated from automated methods. Some research created new geometric and non-geometric semantics with a higher LOD/LoD. The output-enriched models could appear as BIM or CIM. Various applications can be realized through semantic enrichment. As shown in Figure 1, in a sense, semantic enrichment is the process of inferring and enriching the semantics to a given model based on domain knowledge (Bloch and Sacks, 2018).

2.2 Data collection

The data collection process encompassed collecting, identifying, scoping, synthesizing, screening, and reducing following the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) standard. The first step encompasses the collection of related publication by applying the combinations of query strings of “"Semantic enrichment" ("building information model*" OR BIM OR "geographical information system*" OR GIS OR "city information model*" OR CIM)" and "((BIM GIS) OR CIM) (semantic* OR LOD OR interoperability)" in the full-text search engines of Google Scholar. The first query string focused on the papers explicitly declares the keyword semantic enrichment in related fields, while the second one aims to include more papers on semantic enrichment in a BIM and CIM integration scenario. Similar query terms were applied with the 'snowball' method in Scopus and the university library search engine to ensure the completeness of collected data.

Then, in collecting, identifying, and scoping the literature, a quick screening was conducted to identify relevant articles regarding the titles, abstracts, and keywords. The priority was given to the journals in the building and urban domains. The reason is that journal papers offer more reliable and accurate materials due to rigorous review procedures (Akram et al., 2019). Based on the proposed conceptual model, non-empirical studies such as commentary and position papers were excluded. Empirical research can provide 'verifiable' evidence, which is rigorously resultant from specific empirical studies. A snowballing technique (Oraee et al., 2017) was employed to ensure the scope of the review is complete. The full-texts of the screened publications were reviewed by all of the authors to confirm the coverage's completeness. Finally, the process resulted in 44 publications, including 36 journal papers, 4 conference papers, 2 reports, 1 Ph.D. thesis, and 1 book chapter. We extended the scope to include the early years, but most of the pertinent publications fell between 2010 and 2019 (both inclusive). The reason is plausible because the National BIM Standard was published in the United States in 2007, and Open Geospatial Consortium (OGC) members adopted CityGML as the official standard almost simultaneously in 2008. Thus, we crossed out a few cases before 2010 to reflect the most recent advancements in the fields.

In the third step of synthesizing, screening, and reducing, we examined the semantic enrichment methodologies in the 44 publications guided by the conceptual model, as shown in Figure 1. With the descriptors defined in Table 2, the basic information of the selected empirical studies is extracted and tabulated in Table 3. It can be seen from Table 3 that the studies are grouped into clusters. The first 22 publications were enrichment cases for BIM, while the remaining 22 were for CIM. The publication year, subject facility, baseline model, source of new semantics, semantic enrichment methods, new semantics, and output enriched model were coded for each paper. For example, the fourth item Kim et al. (2013) started from a commercial BIM baseline model (BIM-C), added a new source of vector data (D-V), employed the Iterative Closest Point (ICP) – a registration method (Rgst), enriched building's BIM format (BLD) with new geometric and non-geometric semantics (NS-B), and implemented an application in the pre-operation phase (PO).

Table 3. Basic information of the empirical studies on semantic enrichment for BIM and CIM

No.	Reference	Subject facility	Baseline model	Source	Method	New semantics	Enriched model	Application

BIM semantic enrichment								
1	Eastman et al. (2010)	BLD	BIM-O	D-T	Rsn	NS-G	BIM	PO
2	Golparvar-Fard (2011)	BLD	BIM-O	D-R	Rgst	NS-G	BIM	PO
3	Bosché (2012)	BLD	BIM-C	D-R	Rgst	NS-G	BIM	UNS
4	Kim et al. (2013)	BLD	BIM-C	D-V	Rgst	NS-B	BIM	PO
5	Xiong et al. (2013)	BLD	BIM-C	D-V	Rgst	NS-G	BIM	UNS
6	Irizarry et al. (2013)	BLD	BIM-C	C-CIM	Rsn	NS-N	BIM	PO
7	Zhang et al. (2014)	BLD	BIM-O	D-T	Rsn	NS-B	BIM	PO
8	Lee et al. (2014)	INFR	BIM-O	O-R	Rsn	NS-B	BIM	OM
9	Hong et al. (2015)	IN	BIM-C	D-R	Sgmt	NS-G	BIM	PO
10	Wang et al. (2015)	BLD	BIM-C	D-R	Sgmt	NS-G	BIM	EN
11	Belsky et al. (2016)	BLD	BIM-O	O-R	Rsn	NS-G	BIM	PO
12	Zeibak-Shini et al. (2016)	BLD	BIM-C	D-V	Rsn	NS-B	BIM	OM
13	Bassier et al. (2016)	BLD	BIM-O	D-V	Sgmt, Rsn	NS-G	BIM	PO
14	Sacks et al. (2017)	INFR	BIM-O	D-T	Rsn	NS-N	BIM	PO
15	Bloch et al. (2018)	IN	BIM-C	O-R, D-T	Sgmt, Rsn	NS-B	BIM	PO
16	Xue et al. (2018)	BLD, IN	BIM-C	D-R	Rgst	NS-B	BIM	PO
17	Hamid et al. (2018)	BLD, IN	BIM-C	D-T	Rsn	NS-B	BIM	PO
18	Bienvenido-Huertas et al. (2019)	BLD	BIM-C	D-R	Sgmt	NS-B	BIM	OM
19	Koo et al. (2019)	BLD	BIM-O	C-BIM	Sgmt	NS-B	BIM	PO
20	Simeone et al. (2019)	BLD	BIM-C	O-R	Rsn	NS-B	BIM	OM
21	Xue et al. (2019a)	IN	BIM-C	D-V	Rgst	NS-B	BIM	PO
22	Xue et al. (2019b)	IN	BIM-C	D-V	Rgst	NS-B	BIM	PO
CIM semantic enrichment								
23	Wittner (2010)	BLD, IN	CIM-C	D-V, D-R	Rsn	NS-B	CIM	UB
24	El-Mekawy et al. (2011)	BLD	CIM-O	C-BIM, O-F	Rsn	NS-G	CIM	PO
25	He et al. (2012)	CO	CIM-C	O-R, D-R	Rsn	NS-G	CIM	UB
26	Irizarry et al. (2012)	BLD	BIM-C	C-CIM	Rgst	NS-N	CIM	PO
27	Löwner et al. (2013)	CO	CIM-O	O-R	Rsn	NS-G	CIM	PO
28	Xu et al. (2014)	CO	CIM-O	C-BIM	Rsn	NS-N	CIM	UB
29	Mignard et al. (2014)	CO, BLD	CIM-O	O-F	Rsn	NS-N	CIM	UB
30	Borrmann et al. (2015)	INFR	BIM-O	C-CIM	Rsn	NS-B	CIM	PO
31	Amirebrahimi et al. (2015)	BLD	BIM-O	D-V	Rsn	NS-B	CIM	UB
32	Kang et al. (2015)	BLD	CIM-O	C-BIM	Rsn	NS-B	CIM	OM
33	Biljecki et al. (2016a)	BLD	CIM-O	O-R	Rsn	NS-G	CIM	PO
34	Biljecki et al. (2016b)	BLD	CIM-O	O-F	Rsn	NS-G	CIM	EN
35	Deng et al. (2016)	CO	CIM-O	O-R	Rsn	NS-G	CIM	UNS
36	Hor et al. (2016)	CO, BLD	CIM-C	O-F	Rsn	NS-B	CIM	UB
37	Karan et al. (2016)	BLD, RE	CIM-O	O-F	Rsn	NS-B	CIM	PO
38	Peng et al. (2016)	BLD	CIM-C	C-BIM, D-V, D-T	Rsn	NS-B	CIM	PO
39	Howell et al. (2017)	RE	BIM-O	D-T	Rsn	NS-G	CIM	OM
40	He et al. (2018)	RE	CIM-C	O-S	Rsn	NS-B	CIM	UB
41	Lee et al. (2018)	INFR	CIM-C	C-BIM, D-R, D-T	Rsn	NS-B	CIM	OM
42	Wang et al. (2019)	CO	BIM-O	C-CIM	Rsn	NS-B	CIM	OM
43	Zhao et al. (2019)	INFR	CIM-C	C-BIM	Rsn	NS-G	CIM	UB
44	Aleksandrov (2019)	BLD, RE	CIM-O	D-V, D-R	Rsn	NS-G	CIM	UB

3. A review of the semantic enrichment of BIM and CIM

An overview of the 44 semantic enrichment studies is visualized as a Sankey chart in Figure 2. In Figure 2, the six major components in the conceptual model were shown as columns from left to right, while two columns, 'model format' and 'new semantics,' were added to explain the formats of baseline models and the results of methods, respectively. The size of a rectangular component 'stock' indicates the number of cases associated with the label in Table 3. A curved 'flow' between two components stands for the frequency of co-occurrence in the 44 cases.

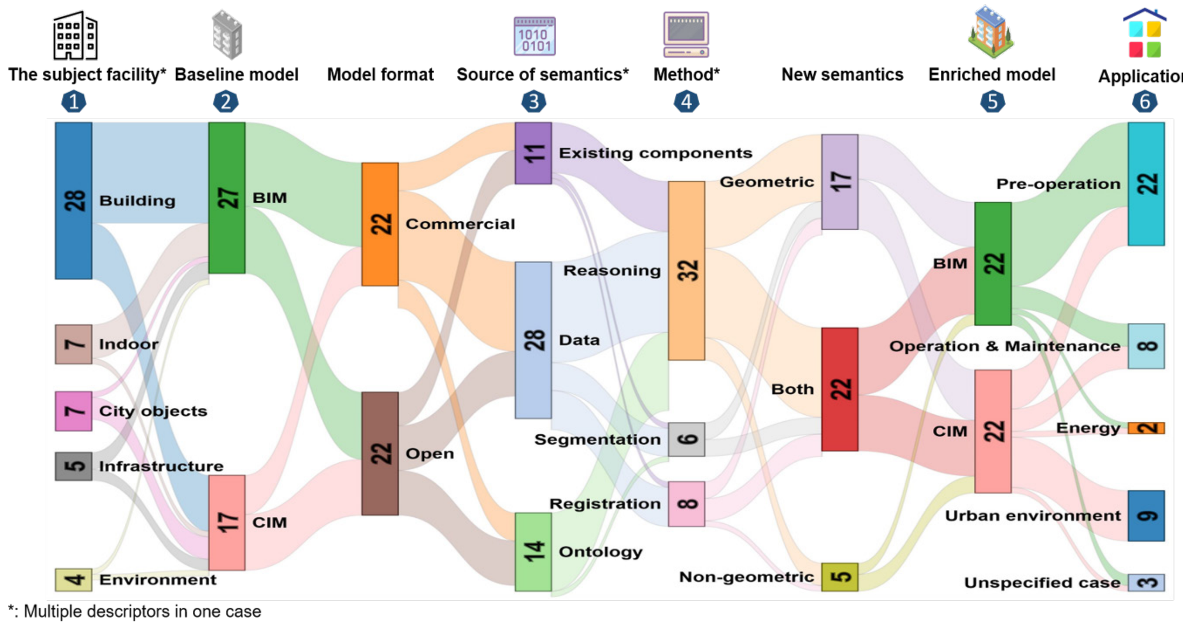

Figure 2. Sankey chart of the semantic enrichment studies for BIM and CIM (2010 – 2019)

3.1 The subject facilities

As shown in Figure 2, the distributions of subject facilities were considerably uneven, where 28 out of 44 studies focused on buildings. In BIM semantic enrichment, seventeen studies targeted buildings, six on indoor areas, and two on infrastructures. Note that some papers, e.g., Xue et al. (2018) and Hamid et al. (2018), involved both building and indoor furniture in their validations. In contrast, the subjects were more balanced for CIM. Twelve CIM cases targeted buildings, one on indoor areas, seven on city objects, four on environments, and three on infrastructure. The building, which is vital to the daily work and life of humankind in cities, is a critical meeting point between BIM and CIM.

3.2 The baseline models and formats

The baseline models, by definition, can be divided into two types: BIM and CIM. Each type had two formats, i.e., commercial and open, as shown in Figure 2. Overall, 27 cases were based on BIMs, while 17 on CIMs. For BIM semantic enrichment, all 22 cases used BIMs as the baseline models, 14 of which were commercial BIM formats and solutions (e.g., Autodesk

Revit). In comparison, open BIM formats and solutions, such as IFC and ISO 16739-1:2018, were adopted in 8 studies, slightly less than the commercial ones. Powerful and open—to a certain extent—APIs (Application Programming Interfaces) and online free development documents offered by commercial BIM software vendors can be good explanations (Xue et al., 2019a).

Unlike the BIM ones, only 5 out of 22 CIM cases were enriched above baseline BIMs (Irizarry et al. 2012; Borrmann et al., 2015; Amirebrahimi et al., 2015; Howell et al., 2017; Wang et al., 2019), while the majority was based on CIMs. Over half (14 out of 22) of baseline models had open formats such as CityGML, while the remaining employed commercial formats and platforms such as ESRI ArcGIS. The open formats, such as CityGML or IFC, are known to have better transparency, operability, or scalability. For example, Deng et al. (2016) applied an instance-based method to generate mapping rules between CityGML and IFC to enable bidirectional semantic conversions. The commercial solutions to CIM, similar to those for BIM, were good at handling large-scale models and data. For instance, He et al. (2018) enriched geological semantics on top of a fully 3D city model in ArcGIS.

3.3 The sources of new semantics

The sources of new semantics consist of three major types: data, ontology, and existing BIM/CIM components. As shown in Table 3 and Figure 2, 17 out of 22 semantic enrichment cases for BIM extracted new semantics from various types of data. Among these, 6 studies involved vector data. For instance, Kim et al. (2012) and Xue et al. (2019a; 2019b) processed LiDAR points to produce semantically enriched BIMs. Also, 6 studies employed raster data. For example, Xue et al. (2018) used photos. Eastman et al. (2010), Zhang et al. (2014), Sacks et al. (2017), and Hamid et al. (2018) exploited tabular data. A few studies looked into the ontology sources (Lee et al., 2014; Belsky et al., 2016). Simeone et al. (2019) deployed an ontology-based system together with *ad hoc* rules as the source for enriching BIM; one study used both tabular data and *ad hoc* rules in semantic enrichment (Bloch et al., 2013). For the existing components, Irizarry et al. (2013) and Koo et al. (2019) reused BIM and CIM components.

In 11 out of 22 CIM studies, data was the semantic source. For instance, Howell et al. (2017) applied tabular data, including water metering as the source; Wittner (2010) and Aleksandrov (2019) triangulated both raster and vector data. Ten studies applied ontologies to inferring new semantics (e.g., geological relationships utilized in He et al. [2018]). Various rules, as implicit or incomplete ontologies, were also popular as the semantic sources. For example, Löwner et al. (2013) studied 44 combinations of Geometric and Semantics Level of Detail (GLOD/SLOD). BIM components were adopted as the semantic source in six cases, where the most frequent research design trend was the open BIM to CIM conversion, such as from IFC to CityGML. For instance, Xu et al. (2014) and Kang et al. (2015) tested different ways for importing IFC components into CIMs. A few cases channeled semantic elements from reference CIMs to generate new CIMs (Irizarry et al., 2012; Borrmann et al., 2015; Wang et al., 2019). Some studies integrated multiple types of semantic resources. For example, He et al. (2012) exploited both raster data (e.g., 2D cadastral maps) and *ad hoc* rules; EI-Mekawy (2011) used BIM components and formal ontology; Lee et al. (2018) involved raster data, tabular data, and BIM components simultaneously.

3.4 The enrichment methods

3.4.1 Building information model

The 22 BIM cases covered all three types of enrichment methods, namely semantic reasoning, semantic registration, and semantic segmentation. A semantic reasoning method holds a set of fixed, usually predetermined processes without computational evolution or adaptations. More specifically, semantic reasoning methods included rule-based, formal ontology-based, and simulation ones. As exhibited in Table 3, semantic reasoning was the method for 11 of the BIM cases. An example was Belsky et al.'s (2016) *SeeBIM* engine which reasoned IFC models with five parts: parser, database, three-tiered structure inference rules, IFC writer, and rule-processing engine; through the reasoning engine, missing precast joints, connections, and slab aggregations of a concrete parking garage were enriched automatically.

The semantic segmentation consists of supervised machine-learning methods, which concluded hyper-models of classification from annotated training data sets. Six cases relied on semantic segmentation. For example, Bienvenido-Huertas et al. (2019) applied the J48 algorithm; Koo et al. (2019) applied support vector machines for semantic enrichment. Some segmentation studies involved multiple methods. For instance, Bloch (2018) compared supervised machine learning to reasoning methods in terms of feature collection and datasets; Bassier et al. (2016) applied both segmentation and reasoning methods.

In total, 7 out of 22 BIM cases adopted semantic registration methods, which were evolutionary computations through multiple rounds (iterations) of searching in a trial-and-error fashion. Kim et al. (2012) employed the ICP method for matching 3D measurement data with existing BIM. Advanced evolutionary computation (EC) methods, such as the covariance matrix adaptation evolution strategy (CMA-ES) and Niching Migratory Multi-Swarm Optimization (NMMSO), were applied to process 2D and 3D raster data to semantically enriched as-built BIMs for complex indoor and outdoor scenes (Xue et al., 2018; 2019b). The key to semantic registration is a reward (or penalty) function, which can be the photo's structural similarity (SSIM) metrics or root-mean-square errors (RMSE) between 3D data and models (Xue et al. 2019a).

3.4.2 City Information Model

The variety of enrichment methods for CIM was much lower. Semantic reasoning dominated the CIM studies. Many reasoning approaches were applied for mapping and converting IFC, an open BIM standard, to CityGML. Example methods include semantic web technology (Zhao et al., 2019) and data interface modeling (Amirebrahimi et al., 2015). Enrichment to CityGML's LoD standard was a topical research direction in CIM. Examples included Biljecki et al.'s (2016a; 2016b) 16 levels of LoDs and Löwner et al.'s (2013) significant extensions like GLOD, SLOD, and ALOD (appearance LoD). With the help of proprietary open APIs and free development documents available online, commercial formats and models were also used for semantic reasoning. For instance, He et al. (2018) enriched the expressions in geological relationships via a Diagrammer Geodata baseline model. Whether open or commercial, the key to enabling semantic reasoning was the interoperability issues between different formats and standards. Once the interoperability issues were solved, e.g., by mapping the ontologies or functional APIs, semantic reasoning would become a shortcut to reuse existing BIM and CIM resources to create CIMs at high LoD levels. Only one study (Irizarry et al., 2012) adopted the registration method.

3.5 The enriched models and semantics

In the 22 cases of enriched BIMs, twelve received both geometric and non-geometric new semantics; eight had new geometric semantics only, and two gained non-geometric semantics only. In contrast, ten out of 22 CIM cases had both types of new semantics; nine cases aimed at geometric semantics only; while only three cases, Irizarry et al. (2012), Mignard et al. (2014), and Xu et al. (2014) enriched non-geometric semantics to its CIM. To sum up, semantic enrichment studies for BIM weighed the non-geometric semantics more (14 out of 22) apart from the geometry (20 out of 22); CIM focused on the geometry (19 out of 22). The result was consistent with the guiding scales, i.e., LOD and LoD, to BIM and CIM.

More interesting was about the subtypes of the physical facilities and their new semantics. As listed in Table 4, there were over 20 types of objects (e.g., furniture and room space) under the five major types of subject facilities. Their geometric attributes (e.g., area), non-geometric types (e.g., room type), identifiers (e.g., manufacturer), and relations (e.g., spatial relationship) were concluded from the literature. First, the objects listed under the building and city objects were the most diversified; meanwhile, these objects also had the most diversified geometric attributes, e.g., a building object can have a long list of attributes including height, distance, size, area, material, shape, dimension, volume, thickness, length, and proximity. The reason for the most listed objects under the building and city objects may be the diversification of related University majors and faculties. For example, the faculty of Architecture, the faculty of Civil Engineering, and the faculty of Geographic Information Technology have various studies on the semantics of buildings and cities in BIM and CIM.

Table 4. List of geometric and non-geometric semantics for BIM and CIM in the literature

Enriched model	Subject facility	Geometric	Non-geometric

BIM	Indoor facility	**Objects**: furniture, room space, kitchen cabinetry **Attribute**: size, area, material, height	**Type**: Room type, OmniClass number, model, style number, identity data **Identifier**: manufacturer, material product line, phasing, release date, cost, assembly code, assembly description
BIM & CIM	Building	**Objects**: interior entities (e.g., door, wall, window, column, beam, covering, ceiling, footing, pile, plate, roof, slab, railing, ramp, stair, floor, precast concrete component are BIM objects defined inside building or as building parts), balconies, exterior entities (e.g., door, wall, window, roof, facade are CIM objects because they are part of the external environment of the building), building, parking garage **Attributes**: height, distance, size, area, material, shape, dimension, volume, thickness, length, proximity	**Type**: functional classification **Identifier**: construction progress, degree of intervention, knowledge representation and management, cost **Relation:** spatial relation, logical relation
	Infra-structure	**Objects**: bridge, highway, tunnel **Attribute**: quantity, shape	**Type**: design options **Identifier**: cost, environmental impact **Relation:** spatial relation
CIM	City objects	**Objects**: exterior entities (e.g., in CIMs, the outer surfaces of windows, doors, roofs, and facades are regarded as part of the exterior of buildings), building, outbuilding shell, road, public garden, drainage system, structure elements, ground, streetlight, pedestrian path **Attribute**: length, shape, geometrical resolution, semantic depth, texture, distance	**Type**: place type, component type, land title number **Identifier**: land use, land cover, building facilities information, campus water supply pipeline information, component information, coordinate **Relation:** spatial relation
	Environ-ment	**Objects**: water utility, vegetation, geological object **Attribute**: size, shape, texture	**Type**: telemetry data, valve and pump states, smart metering data **Identifier**: pipe description, household description **Relation:** spatial relation, geological structural relation, geological phenomenon relation, geological object relation

The most diversified identifiers were found in indoor facilities and city objects. Most subtypes of relations were found from the enrichment cases on the environment semantics. For example, the spatial relation in the geological environment gives the position of the rock in space relative to a reference object (e.g., the position of a tree). As another example, the logical relation in the geological environment gives the dependence between two rock layers. The items in Table 4 reveal each subject facility's potential in semantic enrichment for BIM and CIM, which can partially explain why buildings were the most targeted facilities in literature. In the future, the indoor objects' identification can become an emerging trend for BIM, while city objects' potentials on the object types and the environment's potential on the relations can lead to new explorations in the semantic enrichment for CIM.

3.6 The applications

Applications and the values embedded specify the target semantics and initiate enrichment processes. In BIM semantic enrichment studies, 15 were applied in pre-operation phases (e.g., planning and construction), four in operation and maintenance, one in energy, while two were unspecified. Most novel value-added applications of BIMs lie in the pre-operation phases of buildings. The distribution reflects the scope of BIM studies and the interest of BIM researchers and practitioners. In comparison, applications of CIM semantic enrichment showed more balanced and diverse fields. Nine CIM cases were applied in the urban environment (e.g., urban planning), seven in pre-operation, four in operation and maintenance, one in energy, while one paper did not specify its application. The diverse applications echo the diverse sources of CIM semantics, as shown in Table 3.

4. The trends and prospects of semantic enrichment of BIM and CIM

Figure 3 shows the stacked bar charts of the yearly semantic enrichment cases for BIM and CIM from different perspectives. First, Figure 3a shows that in BIM studies, there has been a surge of indoor facility studies since 2018, while the study of buildings has increased relatively steadily since 2010. In comparison, the CIM cases had more diversified and balanced subject facilities. For example, in 2019, the four cases covered four different types of subject facilities, while the environment was the latest emerged subject (in 2016). Virtual reality and augmented reality (VR/AR), which had recently become a household term, was one enabler of indoor facility studies (Chen and Xue, 2011). For instance, Google's Project Tango was a pioneer project for household VR/AR, from which two smartphone models were mass-produced. One was *Lenovo Phab 2 Pro*, released in December 2016, and the other was *Asus Zenfone AR*, announced in 2017. Although Tango was not a successful project in terms of business, the new VR/AR sensor data directly inspired a few indoor facility studies as semantic sources: e.g., Xue (2019a). In 2020, new mainstream smartphone models have Apple's AR-Kit or Google's ARCore installed; furthermore, VR/AR is growing extensively. For example, IDC (2019) predicted that global VR/AR spending would have a five-year compound annual growth rate (CAGR) of 77.0%. Thus, semantic enrichment of indoor facilities should gain new momentum in the future.

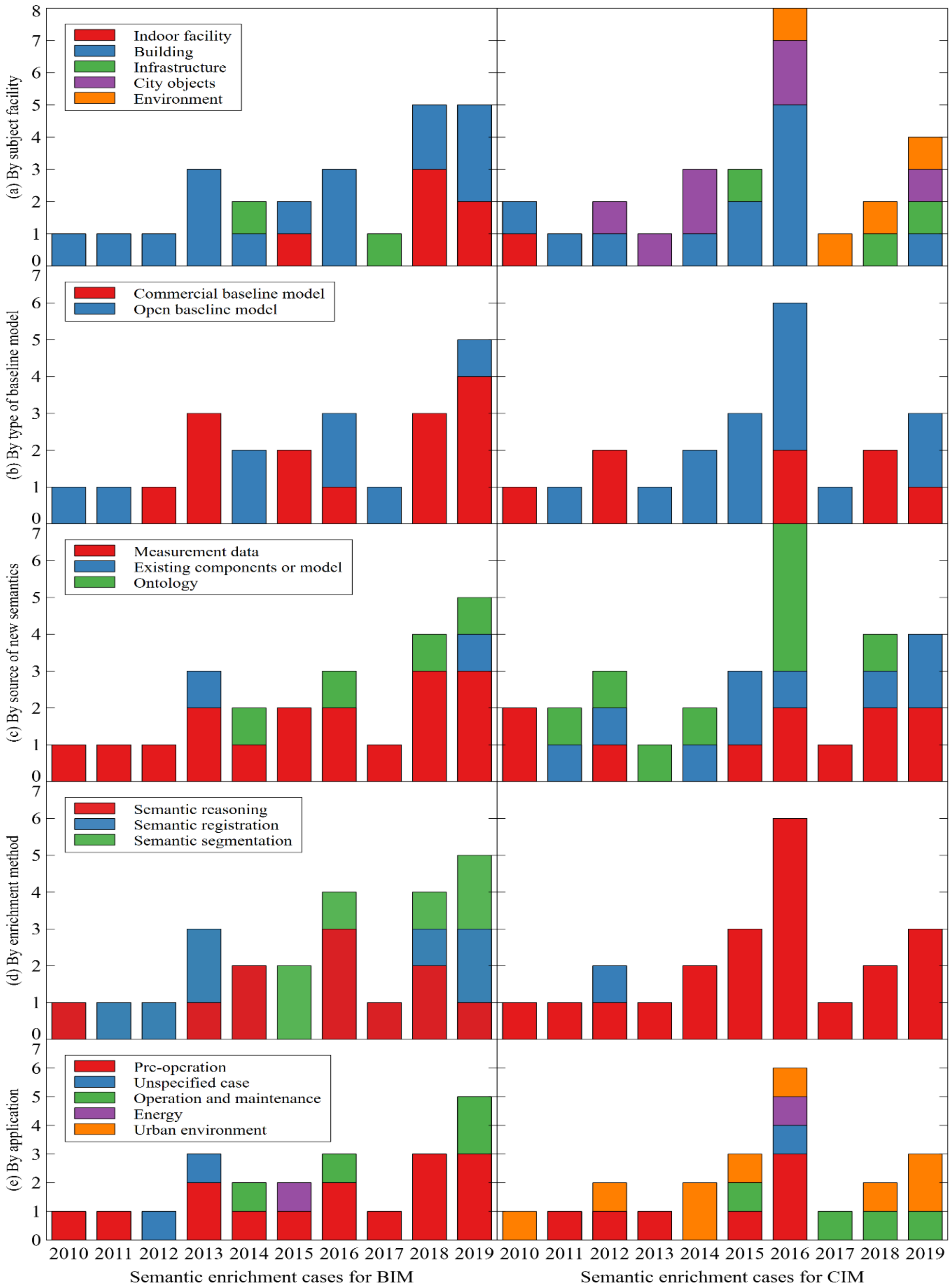

Figure 3. Stacked bar charts of yearly published semantic enrichment cases for BIM and CIM

Figure 3b shows that since 2018, commercial baseline models have been increasingly used in BIM cases. During the same period, only a few CIM cases adopted commercial baseline models. For the BIM cases, one reason was the poor efficacy of the open BIM standard— IFC. For example, Zhiliang et al. (2020) pointed out that the information corresponding to the established property sets can only be read, but its semantics cannot be automatically identified in the current IFC. De Gaetani (2020) recognized faults when exchanging sample IFC BIM models and related Gantt charts. Besides, information redundancy, such as the randomized sequential identifiers (#-Ids) and complicated cross-referencing relation systems, makes it difficult to extract and, particularly, to enrich BIM information efficiently using IFC (Borrmann et al. 2018; Xue and Lu, 2020). In comparison, CIM semantic enrichment studies used to be benefited from OGC's long list of open standards covering a wide spectrum from data to model to tool. However, a few recent semantic enrichment studies for CIM were benefited from commercial but powerful APIs, e.g., He et al. (2018), or the CIM-BIM integration, e.g., Lee et al. (2018) and Zhao et al. (2019). Considering the development of IFC and recent collaborations between BIM and CIM software giants, we would likely see more semantic enrichment studies utilizing commercial standards and solutions in the next five years for BIM and CIM. In the longer run, open standards have the ground to continue its prevalence for CIM, as well as for BIM in case the APIs to IFC such as the *IfcOpenShell* project (Krijnen, 2011) are competitive against commercial platforms.

For both BIM and CIM, all three types, i.e., data, model, and ontology, have been involved as sources of new semantics, as shown in Figure 3c. In the BIM cases, data was the most frequently used, while existing components were the least. In contrast, situations were the opposite for CIM – data was the least frequent while existing components and ontologies were the most. Availability of city-scale measurement data should not be among the reasons because accurate data sets such as satellite photos, radar images, and 3D LiDAR point clouds became much available and affordable in the recent decade. However, the availability of effective algorithms and methods to pre-process and process the city-scale data sets can be a cause. For example, the 2015 Dublin LiDAR was an open-source data set consisting of 1.4 billion 3D points; it wasn't until 2019 that Zolanvari et al. published a pre-processed data set with a sheer 18.6% point cloud with semantic annotations. Nevertheless, researchers on CIM can refer to up-to-date data processing methods for BIM (e.g., deep learning, statistical modeling, model-based down-sampling, and optimization algorithms) to handle city-scale data. Therefore, we are open to seeing more data-driven semantic enrichment studies for CIM in the future. For BIM, components and model is a promising source of new semantics to take advantage of the CIM-BIM integration.

The three types of methods were roughly equally employed for BIM semantic enrichment in terms of frequency, while semantic segmentation methods appeared in the second half of the review period. The situation for CIM was different, where most cases were conducted through semantic reasoning methods. The major reason should be that the rule of ontology-based reasoning is the most efficient (i.e., least time cost) when data grows to the city scale. In comparison, semantic registration methods for BIM applied a considerable number of trial-and-error tests to match the indoor facility to correct 3D models; and semantic segmentation methods require a large number of training examples (e.g., Zolanvari et al. 2019). However, we are optimistic to see that the research communities for BIM and CIM can learn effective and efficient technologies from each other, along with the emerging trend of crossing the disciplinary boundary. Besides, new technologies, such as big data analytics, which can discover new relationships from complex urban big data, are also potential game-changers to process city-scale data sets (Lu et al., 2018). Therefore, the semantic enrichment methods for CIM are very likely to be more efficient and diversified.

A shared value-added application can be a strong driver for promoting and integrating the complementary semantics in BIM and CIM. For example, Figure 3e shows that both started to impact the operation and maintenance phase in 2014. However, there still are some obstacles:

- Different standard formats (Liu et al., 2017), for example, the IFC schemas are not fully compatible with the definitions in CityGML;
- Lack of legal protection for intellectual property rights (Ku and Taiebat, 2011), for example, authorship, readership, and ownership of semantics within the model;
- Funds and time costs required for integration (Suermann and Issa, 2009), for example, employing modelers and consultants;
- Unsuitable external environmental conditions (Azhar et al., 2012), for example, lack of professional interactivity and external motivations;
- Insufficient management (Eastman et al., 2011), for example, lack of a well-established workflow and immature dispute resolution;
- Different spatial scale (Song et al., 2017), for example, data and information with multiple spatial scales; and
- Different user focus (Liu et al., 2017), for example, BIM users used to focus on construction projects, while CIM users may focus on spatial analysis of outdoor environments.

In the past, much focus was on resolving technical difficulties, while management, environmental and legal obstacles received less attention. Therefore, researchers should realize that openness and collaboration are the keys to favorably integrating BIM and CIM. Application-driven semantic enrichment, frequent communication, and government policies are also essential for successful integrations. For example, recent smart buildings and city initiatives require people in these two fields to collaborate to develop more semantic enrichment applications (e.g., energy). It would be a misconception to assume that enriched models will be automatically implemented in real life. Instead, realizing semantic enrichment applications requires positive attitudes from people in different sectors. This is the only way the nexus between the research and practice of semantic enrichment will spiral upward.

In short, this research developed a conceptual model consisting of six components to revisit the existing semantic enrichment studies in BIM and CIM for outlining the trends and prospects. Value-added applications and research gaps brought opportunities for future investigations. Firstly, BIM semantic enrichment starting from CIM was rare in the past, and future research breakthroughs can be extended from this perspective. Secondly, compared with other subject facilities, researchers can also pay attention to indoor facilities. In the era of big data, people can generate and access more and more indoor data, which can involve humans in the loop through the semantics of indoor behaviors and facility usages. Thirdly, with the rapid development of data processing algorithms, future research should attempt to use city-scale data sets as a source of new semantics to help develop smart buildings and cities and manage the life cycle of infrastructures. Last, but not least, research using semantic reasoning is relatively mature and saturated, and future research can tackle more semantic registration or segmentation methods.

The findings of this study also have practical implications. Although semantic enrichment is relatively successful in the pre-operation phase, there is still a lack of systematic considerations to consolidate isolated applications in different phases and departments. One reason is that user interfaces are isolated at different phases of the construction life cycle (e.g., design, construction, and post-maintenance). Another possible reason is that different government departments hold CIM datasets. For example, in Hong Kong, the Land Department holds 3D digital maps, while building semantics are separately managed by the Building Department, Housing Department, Planning Department, and Home Affairs Department. Therefore, in addition to technical obstacles, there are also institutional barriers. Furthermore, neither private companies nor the public can access the datasets. Therefore, common urban semantic platforms, such as Queriocity (Lopez et al., 2012) and Smart City Platform (ASTRI, 2019), should aim to solve transparency and accessibility issues. Researchers interested in semantic enrichment can also pay attention to emerging theoretical challenges from common urban semantic platforms.

5. Conclusions

Semantic enrichment is a process that adds semantic information to a building information model (BIM) or a city information model (CIM) to enable applications pertinent to construction management, geoinformatics, and urban planning. However, the process of harnessing semantic enrichment to improve the existing BIM and CIM seems a conundrum. This paper proposes a conceptual model representing the series of generic semantic enrichment processes for both BIM and CIM. In the model, a semantic enrichment study first targets a subject facility, then adopts a baseline BIM or CIM, locates sources of new semantics, distills the new semantics using various methods, which results in an enriched BIM or CIM by appending the newly distilled semantics, and finally realizes applications. The proposed model was applied to 44 empirical studies in the literature.

A comparative analysis of semantic enrichment in BIM and CIM showed considerable differences in terms of scopes, terminologies, methodologies, and scales. For example, a recent study on BIM semantics focused more on buildings and indoors, while CIM research covered buildings, infrastructure, urban objects, and environments. The future research will probably gain opportunities from the penetrations of remote sensing devices (such as VR/AR smartphone), intelligent data processing algorithms, BIM and CIM software platforms, and other emerging technologies such as big data analytics. Since 2018, commercial platforms that serve as baseline models for semantic enrichment in BIM and CIM have become common in literature. There are BIM cases since open BIM standard IFC is not effective. CIM semantic enrichment studies used to benefit from OGC's long list of open standards; however, a few recent CIM studies benefited from commercial but powerful APIs. When it comes to the source of new semantics for BIM cases, data was the most frequently used, while existing components were the least. In contrast, situations were the opposite for CIM. Notably, the enrichment methods were rather different, too. Semantic reasoning, semantic registration, and semantic segmentation were roughly equally employed for BIM semantic enrichment in terms of frequency, while most CIM cases were conducted through semantic reasoning methods. BIM semantic enrichment was mainly applied in pre-operation phases, while CIM semantic enrichment was mainly applied in urban environments. However, recent trends indicated that BIM and CIM have similar applications in the operation and maintenance phase.

The study is the first-ever one of its kind to examine the niche area of semantic enrichment. Moreover, the study combs the development in fields and helps readers understand the changes in semantic enrichment in the past ten years and its prospects or future research. Another contribution of this research is the conceptual model of semantic enrichment comprising six components that explain the rationale of semantic enrichment from a data processing perspective. By triangulating the conceptual model with the historical data from the niche of semantic enrichment research, this study summarized the research findings, based on which future development could proceed on a more streamlined footing.

Although the study in this paper has many merits, it also has the following limitations. Firstly, manual screening and coding work could inevitably present subjectivity. The accuracy can be improved by discovering more related publications in future work. Secondly, the selection and discussion of the semantic enrichment studies were specified to the empirical city and building cases in the scholarly publications in the urban-related fields. Practitioners might consider real-life demands and conditions as well. A systematic recommendation of semantic enrichment methods distilled from the best practices in the literature should be included in future work. Amidst the general trend of enhancing semantic enrichment, researchers and practitioners alike are encouraged to think out of the box, e.g., where are the optimal 'sweet points' of BIM and CIM between the high LOD/LoD semantics and various costs (e.g., funds, privacy, social inequality, and human-machine conflict), as it is certainly not the higher the LOD/LoD, the better.

Acknowledgements

The work described in this paper was fully supported by a grant from the Research Grants Council of the Hong Kong SAR (Project No. 27200520). We would like to express our gratitude to the anonymous referees for the constructive comments on the clarity and quality of this paper.

References

* Indicates *references for empirical studies analyzed by the author*s

Akram, R., Thaheem, M. J., Nasir, A. R., Ali, T. H. & Khan, S., 2019. Exploring the role of building information modeling in construction safety through science mapping. *Safety Science*, 120, 456-470. https://doi.org/10.1016/j.ssci.2019.07.036

*Aleksandrov, M., Diakité, A., Yan, J., Li, W., Barton, J., Doig, J. & Zlatanova, Z., 2019. Progressing precinct modelling on the UNSW campus and beyond: BIM/PIM and 3DGIS. Project Report. Australia Government. Sydney: CRC for Low Carbon Living. https://www.be.unsw.edu.au/sites/default/files/upload/grid/LCL_CRC_3DPIM2_Data_Modelling.pdf.

*Amirebrahimi, S., Rajabifard, A., Mendis, P. & Ngo, T., 2015. A Data Model for Integrating GIS and BIM for Assessment and 3D Visualisation of Flood Damage to Building. *CEUR Workshop Proceedings. 1323*, pp. 78-79. Brisbane: Locate. http://ceur-ws.org/Vol-1323/paper27.pdf.

Abualdenien, J. & Borrmann, A., 2019. A meta-model approach for formal specification and consistent management of multi-LOD building models. *Advanced Engineering Informatics*, *40*, 135-153. https://doi.org/10.1016/j.aei.2019.04.003

Azhar, S., Khalfan, M. & Maqsood, T., 2012. Building information modelling (BIM): now and beyond. *Construction Economics and Building*, *12*(4), 15-28. https://doi.org/10.5130/AJCEB.v12i4.3032

Borrmann, A., Beetz, J., Koch, C., Liebich, T. & Muhic, S., 2018. Industry Foundation Classes: A standardized data model for the vendor-neutral exchange of digital building models. In B. A., K. M., K. C. & B. J., *Building Information Modeling* (pp. 81-126). Cham, Switzerland: Springer. https://doi.org/10.1007/978-3-319-92862-3_5

Belsky, M., Sacks, R. & Brilakis, I., 2015. A framework for semantic enrichment of IFC building models. Doctoral dissertation, Technion-Israel Institute of Technology, Israel, viewed 15 December 2019, https://bit.ly/2RRPyg6

Berling-Wolff, S. & Wu, J., 2004. Modeling urban landscape dynamics: A review. *Ecological Research*, *19*(1), 119-129. https://doi.org/10.1111/j.1440-1703.2003.00611.x

*Belsky, M., Sacks, R. & Brilakis, I., 2016. Semantic enrichment for building information modeling. *Computer-Aided Civil and Infrastructure Engineering*, 31(4), 261–274. https://doi.org/10.1111/mice.12128

*Biljecki, F., Ledoux, H. & Stoter, J., 2016a. An improved LOD specification for 3D building models. *Computers, Environment and Urban Systems*, 59, 25–37. https://doi.org/10.1016/j.compenvurbsys.2016.04.005

*Biljecki, F., Ledoux, H., Stoter, J. & Vosselman, G., 2016b. The variants of an LOD of a 3D building model and their influence on spatial analyses. *ISPRS Journal of Photogrammetry and Remote Sensing*, 116(C), 42–54. https://doi.org/10.1016/j.isprsjprs.2016.03.003

*Bloch, T. & Sacks, R., 2018. Comparing machine learning and rule-based inferencing for semantic enrichment of BIM models. *Automation in Construction*, 91, 256–272. https://doi.org/10.1016/j.autcon.2018.03.018

*Bienvenido-Huertas, D. E., Nieto-Julián, J. J., Moyano, J. M., Macías-Bernal, J. & Castro, J., 2019. Implementing Artificial Intelligence in H-BIM Using the J48 Algorithm to Manage Historic Buildings. *International Journal of Architectural Heritage*, 1-13. https://doi.org/10.1080/15583058.2019.1589602

*Bassier, M., Vergauwen, M. & Van Genechten, B., 2016. Automated semantic labelling of 3D vector models for scan-to-BIM. Paper presented at the 4th Annual International Conference on Architecture and Civil Engineering, 25–26 April, België, viewed 1 May 2020. https://doi.org/10.5176/2301-394X_ACE16.83.

*Bosché, F., 2012. Plane-based registration of construction laser scans with 3D/4D building models. *Advanced Engineering Informatics*, 26(1), 90-102. https://doi.org/10.1016/j.aei.2011.08.009

*Borrmann, A., Kolbe, T. H., Donaubauer, A., Steuer, H., Jubierre, J. R. & Flurl, M., 2015. Multi-scale geometric-semantic modeling of shield tunnels for GIS and BIM applications. *Computer-Aided Civil and Infrastructure Engineering*, 30(4), 263-28. https://doi.org/10.1111/mice.12090

Brilakis, I., Lourakis, M., Sacks, R., Savarese, S., Christodoulou, S., Teizer, J. & Makhmalbaf, A., 2010. Toward automated generation of parametric BIMs based on hybrid video and laser scanning data. *Advanced Engineering Informatics*, 24(4), 456-465. https://doi.org/10.1016/j.aei.2010.06.006

Boton, C., Kubicki, S. & Halin, G., 2015. The challenge of level of development in 4D/BIM simulation across AEC project lifecycle. A case study. *Procedia Engineering*, *123*, 59-67. https://doi.org/10.1016/j.proeng.2015.10.058

Clarke, M., & Harley, P. 2014. How smart is your content? Using semantic enrichment to improve your user experience and your bottom line. *Science*, *37*(2), 41. http://www.councilscienceeditors.org/wp-content/uploads/v37n2p40-44.pdf

Chen, K., Johan, H., & Erdt, M., 2020. Appearance-driven conversion of polygon soup building models with level of detail control for 3D geospatial applications. *Advanced Engineering Informatics*, 44, 101049. https://doi.org/10.1016/j.aei.2020.101049

Chen, K., & Xue, F. 2021. The renaissance of Augmented Reality in construction: history, present status and future directions. *Smart and Sustainable Built Environment*, in press. Doi: https://doi.org/10.1108/SASBE-08-2020-0124

Davies, R. & Harty, C., 2011. Building Information Modelling as innovation journey: BIM experiences on a major UK healthcare infrastructure project. Paper presented at the 6th Nordic conference on construction economics and organization–shaping the construction/society nexus, 13–15 April, Copenhagen, viewed 1 January 2020. https://www.semanticscholar.org/paper/Building-Information-Modelling-as-innovation-BIM-on-Davies-Harty/1c26f66825994ca920f452151a30093b234af677

*Deng, Y., Cheng, J.C.P. & Anumba, C., 2016. Mapping between BIM and 3D GIS in different levels of detail using schema mediation and instance comparison. *Automation in Construction*, 67, 1–21. https://doi.org/10.1016/j.autcon.2016.03.006

De Gaetani, C. I., Mert, M., & Migliaccio, F. (2020). Interoperability Analyses of BIM Platforms for Construction Management. *Applied Sciences*, 10(13), 4437. https://doi.org/10.3390/app10134437

Eastman, C., Teicholz, P., Sacks, R. & Liston, K., 2011. A guide to building information modeling for owners, managers, designers, engineers and contractors: Handbook, B. I. M, viewed 1 January 2020. 10.1002/9781119287568.

*Eastman, C., Sacks, R., Panushev, I., Venugopal, M. & Aram, S., 2010. *Precast/Prestressed Concrete National BIM Standard Documents: Model View Definitions for Precast Concrete*. Technical Report, BIM Advisory Committee, Precast BIM Standard Project, Precast/Prestressed Concrete Institute. https://itc.scix.net/pdfs/w78-2010-18.pdf

*El-Mekawy, M., Östman, A. & Shahzad, K., 2011. Towards interoperating CityGML and IFC building models: a unified model based approach. In *Advances in 3D geo-information sciences* (pp. 73-93). Berlin, Heidelberg: Springer. https://doi.org/10.1007/978-3-642-12670-3_5

ESRI, 2019. *ESRI-Autodesk Partnership*. Retrieved December 8, 2019, from https://bit.ly/38hz7Q7

*Golparvar-Fard, M., Peña-Mora, F. & Savarese, S., 2011. Integrated sequential as-built and as-planned representation with D 4 AR tools in support of decision-making tasks in the AEC/FM industry. *Journal of Construction Engineering and Management*, 137(12), 1099-1116. https://doi.org/10.1061/(ASCE)CO.1943-7862.0000371

GISGeography., 2015. *Vector vs Raster: What's the Difference Between GIS Spatial Data Types?* Retrieved May 8, 2020, from GISGeography: https://gisgeography.com/spatial-data-types-vector-raster/

Goodchild, M. F., 1991. Geographic information systems. *Progress in Human geography*, 15(2), 194-200. https://doi.org/10.1177%2F030913259101500205

*He, H., Hu, D. & Lü, G., 2018. GIS Application to Regional Geological Structure Relationship Modelling Considering Semantics. *ISPRS International Journal of Geo-Information*, 7, 267. https://doi.org/10.3390/ijgi7070267

*He, S., Besuievsky, G., Tourre, V., Patow, G. & Moreau, G., 2012. All range and heterogeneous multi-scale 3D city models. *Usage, Usability, and Utility of 3D City Models–European COST Action TU0801*, 02006. https://doi.org/10.1051/3u3d/201202006

*Howell, S., Rezgui, Y. & Beach, T., 2017. Integrating building and urban semantics to empower smart water solutions. *Automation in Construction*, 81, 434. https://doi.org/10.1016/j.autcon.2017.02.004

*Hong, S., Jung, J., Kim, S., Cho, H., Lee, J. & Heo, J., 2015. Semi-automated approach to indoor mapping for 3D as-built building information modeling. *Computers, Environment and Urban Systems*, 51, 34-46. https://doi.org/10.1016/j.compenvurbsys.2015.01.005

*Hamid, M., Tolba, O. & El Antably, A., 2018. BIM semantics for digital fabrication: A knowledge-based approach. *Automation in Construction*, 91, 62-82. https://doi.org/10.1016/j.autcon.2018.02.031

*Hor, A. H., Jadidi, A. & Sohn, G., 2016. BIM-GIS integrated geospatial information model using semantic web and RDF graphs. *ISPRS Annals of Photogrammetry, Remote Sensing and Spatial Information Sciences*, 3(4), 73-79. 10.5194/isprs-annals-III-4-73-2016

Hong Kong Applied Science and Technology Research Institute (ASTRI). 2019. *Smart City Platform*, Retrieved October 27, 2020, from https://www.astri.org/tdprojects/smart-city-platform/.

IDC. 2019. *Worldwide Augmented and Virtual Reality Spending Guide.* Framingham, MA, USA: International Data Corporation. https://www.idc.com/getdoc.jsp?containerId=IDC_P34919

*Irizarry, J., Karan, E. P. & Jalaei, F., 2013. Integrating BIM and GIS to improve the visual monitoring of construction supply chain management. *Automation in construction*, 31, 241-254. https://doi.org/10.1016/j.autcon.2012.12.005

*Irizarry, J. & Karan, E. P., 2012. Optimizing location of tower cranes on construction sites through GIS and BIM integration. *Journal of information technology in construction*, 17(23), 351-366. https://www.itcon.org/2012/23

Julin, A., Jaalama, K., Virtanen, J. P., Pouke, M., Ylipulli, J., Vaaja, M. ... & Hyyppä, H., 2018. Characterizing 3D city modeling projects: towards a harmonized interoperable system. *ISPRS International Journal of Geo-Information*, *7*(2), 55. https://doi.org/10.3390/ijgi7020055

*Kim, C., Son, H. & Kim, C., 2013. Automated construction progress measurement using a 4D building information model and 3D data. *Automation in Construction*, 31, 75–82. https://doi.org/10.1016/j.autcon.2012.11.041

*Koo, B., La, S., Cho, N. & Yu, Y., 2019. Using support vector machines to classify building elements for checking the semantic integrity of building information models. *Automation in Construction*, 98, 183-194. https://doi.org/10.1016/j.autcon.2018.11.015

*Karan, E., Irizarry, J. & Haymaker, J., 2016. BIM and GIS Integration and Interoperability Based on Semantic Web Technology. *Journal Of Computing In Civil Engineering*, 30(3), 04015043. https://doi.org/10.1061/(ASCE)CP.1943-5487.0000519

Koch, A. & Heipke, C., 2006. Semantically correct 2.5D GIS data — The integration of a DTM and topographic vector data. ISPRS Journal of Photogrammetry and Remote Sensing, 61(1), 23-32. https://doi.org/10.1007/3-540-26772-7_39

*Kang, T. W. & Hong, C. H., 2015. A study on software architecture for effective BIM/GIS-based facility management data integration. *Automation in construction*, *54*, 25-38. https://doi.org/10.1016/j.autcon.2015.03.019

Kang, Z., Yang, J., Yang, Z. & Cheng, S., 2020. A review of techniques for 3D reconstruction of indoor environments. *ISPRS International Journal of Geo-Information*, 9(5), 330. https://doi.org/10.3390/ijgi9050330

Krijnen, T., 2011. IfcOpenShell. http://ifcopenshell. org

Ku, K. & Taiebat, M., 2011. BIM experiences and expectations: the constructors' perspective. *International Journal of Construction Education and Research*, *7*(3), 175-197. https://doi.org/10.1080/15578771.2010.544155

*Löwner, M., Benner, J., Gröger, G. & Häfele, K.H., 2013. New concepts for structuring 3D city models–an extended level of detail concept for CityGML buildings, Paper presented at International Conference on Computational Science and Its Applications, 24-27 June, Berlin, viewed 27 October 2020. https://doi.org/10.1007/978-3-642-39646-5_34

*Lee, S., Kim, K. & Yu, J., 2014. BIM and ontology-based approach for building cost estimation. *Automation in Construction*, 41, 96-105. https://doi.org/10.1016/j.autcon.2013.10.020

Lu, W., Lai, C. & Tse, T., 2018. *BIM and Big Data for Construction Cost Management.* HongKong, China: Routledge. 10.1201/9781351172325

Lu, W., Peng, Y., Xue, F., Chen, K., Niu, Y. & Chen, X., 2018. The fusion of GIS and Building Information Modeling for big data analytics in managing development sites. *Reference Module in Earth Systems and Environmental Sciences*. https://doi.org/10.1016/B978-0-12-409548-9.09677-9

*Lee, P., Wang, Y., Lo, T. & Long, D., 2018. An integrated system framework of building information modelling and geographical information system for utility tunnel maintenance management. *Tunnelling and Underground Space Technology Incorporating Trenchless Technology Research, 79*, 263-273. https://doi.org/10.1016/j.tust.2018.05.010

Lopez, V., Kotoulas, S., Sbodio, M. L., Stephenson, M., Gkoulalas-Divanis, A. & Mac Aonghusa, P. 2012. Queriocity: A linked data platform for urban information management. Paper presented at the International Semantic Web Conference, 11-15 Novermber, Berlin, viewed 27 October 2020. https://doi.org/10.1007/978-3-642-35173-0_10

Liu, X., Wang, X., Wright, G., Cheng, J. C., Li, X. & Liu, R., 2017. A state-of-the-art review on the integration of Building Information Modeling (BIM) and Geographic Information

System (GIS). *ISPRS International Journal of Geo-Information*, *6*(2), 53. https://doi.org/10.3390/ijgi6020053

*Mignard, C. & Nicolle, C., 2014. Merging BIM and GIS using ontologies application to urban facility management in ACTIVe3D. *Computers in Industry*, 65(9), 1276-1290. https://doi.org/10.1016/j.compind.2014.07.008

National Institute of Building Sciences, 2015. National Building Information Modeling Standard. Version 3, https://www.nationalbimstandard.org

National Geographic Society, 2012. GIS (geographic information system). http://www.nationalgeographic.org/encyclopedia/geographic-information-system-gis/

Oraee, M., Hosseini, M. R., Papadonikolaki, E., Palliyaguru, R. & Arashpour, M., 2017. Collaboration in BIM-based construction networks: A bibliometric-qualitative literature review. *International Journal of Project Management*, 35(7), 1288-1301. https://doi.org/10.1016/j.ijproman.2017.07.001

Ortega-Córdova, L., 2018. Urban Vegetation Modeling 3D Levels of Detail. http://resolver.tudelft.nl/uuid:8b8967a8-0a0f-498f-9d37-71c6c3e532af

Pratt-Hartmann, I., 2004. Fragments of lLanguage. *Journal of Logic, Language and Information*, 13(1), 207–223. https://doi.org/10.1023/B:JLLI.0000024735.97006.5a

Ping, J., Changxiu, C., Jiong, X. & Rongguo, C., 2013. Design and Implementation of Manage Model for Raster Metadata in Spatial Database. *Remote Sensing Technology and Application*, *27*(2), 315-321. CNKI:SUN:YGJS.0.2012-02-024

*Peng, L., Tang, S., Liu, M., Li, P., Zhu, Q. & Guo, H., 2016. Building Interval Planning Approval Approach of BIM and GIS Integration. *Geomatics World, 23*(2), 32-37. http://en.cnki.com.cn/Article_en/CJFDTotal-CHRK201602007.htm

Reinhart, C. F. & Davila, C. C., 2016. Urban building energy modeling–A review of a nascent field. *Building and Environment*, *97*, 196-202. https://doi.org/10.1016/j.buildenv.2015.12.001

*Sacks, R., Ma, L., Yosef, R. & Borrmann, A., 2017. Semantic enrichment for Building Information Modeling: Procedure for compiling inference rules and operators for complex geometry. *Journal of Computing in Civil Engineering*, 31(6), 04017062. https://doi.org/10.1061/(ASCE)CP.1943-5487.0000705

Schlueter, A. & Thesseling, F., 2009. Building information model based energy/exergy performance assessment in early design stages. *Automation in Construction*, 18(2), 153–163. https://doi.org/10.1016/j.autcon.2008.07.003

*Simeone, D., Cursi, S. & Acierno, M., 2019. BIM semantic-enrichment for built heritage representation. Automation in Construction, 97, 122–137. https://doi.org/10.1016/j.autcon.2018.11.004

Stojanovski, T., 2013. City Information Modeling (CIM) and urbanism: blocks, connections, territories, people and situations. Paper presented at the Symposium on Simulation for Architecture & Urban Design, 7–10 April, California, viewed 1 January 2020, http://orcid.org/0000-0002-1790-0254

Suermann, P. C. & Issa, R. R., 2009. Evaluating industry perceptions of building information modelling (BIM) impact on construction. *Journal of Information Technology in Construction*, *14*(37), 574-594. http://www.itcon.org/2009/37

Song, Y., Wang, X., Tan, Y., Wu, P., Sutrisna, M., Cheng, J. C. & Hampson, K., 2017. Trends and opportunities of BIM-GIS integration in the architecture, engineering and construction industry: a review from a spatio-temporal statistical perspective. *ISPRS International Journal of Geo-Information*, *6*(12), 397. https://doi.org/10.3390/ijgi6120397

Tang, P., Huber, D., Akinci, B., Lipman, R. & Lytle, A., 2010. Automatic reconstruction of as-built building information models from laser-scanned point clouds: A review of related techniques. *Automation in construction*, 19(7), 829-843. https://doi.org/10.1016/j.autcon.2010.06.007

Volk, R., Stengel, J. & Schultmann, F. 2014. Building Information Modeling (BIM) for existing buildings—Literature review and future needs. *Automation in construction*, *38*, 109-127. https://doi.org/10.1016/j.autcon.2013.10.023

*Wang, C., Cho, Y. K. & Kim, C. 2015. Automatic BIM component extraction from point clouds of existing buildings for sustainability applications. *Automation in Construction*, 56, 1-13. https://doi.org/10.1016/j.autcon.2015.04.001

*Wittner, Y., 2010. A GIS-Based Building Data Model Storing, Visualizing, and Analyzing Building Information for United States Coast Guard and Environment Systems Research Institute. PhD Thesis, University of Redlands, University of Redlands, Redlands. 10.26716/redlands/master/2010.23

*Wang, M., Deng, Y., Won, J. & Cheng, J. C., 2019. An integrated underground utility management and decision support based on BIM and GIS. *Automation in Construction*, *107*, 102931. https://doi.org/10.1016/j.autcon.2019.102931

Wang, W., Yang, N., Zhang, Y., Wang, F., Cao, T. & Eklund, P., 2016. A review of road extraction from remote sensing images. *Journal of traffic and transportation engineering (english edition)*, *3*(3), 271-282. 10.1016/j.jtte.2016.05.005

*Xu, X., Ding, L., Luo, H. & Ma, L., 2014. From building information modeling to city information modelling. *Journal of information technology in construction*, 19, 292–307. https://www.itcon.org/2014/17

*Xue, F., Chen, K., Lu, W. & Zetkulic, A., 2019a. From Semantic Segmentation to Semantic Registration: Derivative-Free Optimization–Based Approach for Automatic Generation of Semantically Rich As-Built Building Information Models from 3D Point Clouds. *Journal of Computing in Civil Engineering*, 33(4), 04019024. https://doi.org/10.1061/(ASCE)CP.1943-5487.0000839

*Xue, F., Lu, W. & Chen, K., 2018. Automatic generation of semantically rich as-built Building Information Models using 2D images: A Derivative-Free Optimization approach. *Computer-Aided Civil and Infrastructure Engineering*, 33(11), 926–942. https://doi.org/10.1111/mice.12378

*Xue, F., Lu, W., Chen, K. & Webster, C.J., 2019b. BIM reconstruction from 3D point clouds: A semantic registration approach based on multimodal optimization and architectural design knowledge. *Advanced Engineering Informatics*, 42, Advanced Engineering Informatics, 42, 100965. https://doi.org/10.1016/j.aei.2019.100965

Xue, F. & Lu, W., 2020. A semantic differential transaction approach to minimizing information redundancy for BIM and blockchain integration, *Automation in Construction*, in press. https://doi.org/10.1016/j.autcon.2020.103270

*Xiong, X., Adan, A., Akinci, B. & Huber, D., 2013. Automatic creation of semantically rich 3D building models from laser scanner data. *Automation in construction*, 31, 325-337. https://doi.org/10.1016/j.autcon.2012.10.006

Xue, F., Lu, W., Chen, Z. & Webster, C. J., 2020. From LiDAR point cloud towards digital twin city: Clustering city objects based on Gestalt principles. *ISPRS Journal of Photogrammetry and Remote Sensing*, *167*, 418-431. https://doi.org/10.1016/j.isprsjprs.2020.07.020

*Zeibak-Shini, R., Sacks, R., Ma, L. & Filin, S., 2016. Towards generation of as-damaged BIM models using laser-scanning and as-built BIM: First estimate of as-damaged locations of reinforced concrete frame members in masonry infill structures. *Advanced Engineering Informatics*, 30(3), 312-326. https://doi.org/10.1016/j.aei.2016.04.001

*Zhao, L., Liu, Z. & Mbachu, J., 2019. Highway alignment optimization: An integrated BIM and GIS approach. *ISPRS International Journal of Geo-Information*, 8(4), 172. https://doi.org/10.3390/ijgi8040172

*Zhang, J., Yu, F., Li, D. & Hu, Z., 2014. Development and Implementation of an Industry Foundation Classes-Based Graphic Information Model for Virtual Construction. *Computer-Aided Civil and Infrastructure Engineering*, 29(1), 60-74. https://doi.org/10.1111/j.1467-8667.2012.00800.x

Zolanvari, S. M., Ruano, S., Rana, A., Cummins, A., da Silva, R. E., Rahbar, M. & Smolic, A., 2019. DublinCity: Annotated LiDAR Point Cloud and its Applications. *arXiv preprint arXiv,* 1909.03613. https://arxiv.org/abs/1909.03613

Zhiliang, M., Zhenhua, W., Wu, S. & Zhe, L., 2011. Application and extension of the IFC standard in construction cost estimating for tendering in China. *Automation in Construction*, 20(2), 196-204. https://doi.org/10.1016/j.autcon.2010.09.017