



Knowledge Extraction and Discovery Based on BIM: A Critical Review and Future Directions

Zhen-Zhong Hu^{1,2} · Shuo Leng² · Jia-Rui Lin² · Sun-Wei Li¹ · Ya-Qi Xiao²

Received: 19 January 2020 / Accepted: 20 March 2021
© The Author(s) 2021

Abstract

In the past, knowledge in the fields of Architecture, Engineering and Construction (AEC) industries mainly come from experiences and are documented in hard copies or specific electronic databases. In order to make use of this knowledge, a lot of studies have focused on retrieving and storing this knowledge in a systematic and accessible way. The Building Information Modeling (BIM) technology proves to be a valuable media in extracting data because it provides physical and functional digital models for all the facilities within the life-cycle of the project. Therefore, the combination of the knowledge science with BIM shows great potential in constructing the knowledge map in the field of the AEC industry. Based on literature reviews, this article summarizes the latest achievements in the fields of knowledge science and BIM, in the aspects of (1) knowledge description, (2) knowledge discovery, (3) knowledge storage and management, (4) knowledge inference and (5) knowledge application, to show the state-of-arts and suggests the future directions in the application of knowledge science and BIM technology in the fields of AEC industries. The review indicates that BIM is capable of providing information for knowledge extraction and discovery, by adopting semantic network, knowledge graph and some other related methods. It also illustrates that the knowledge is helpful in the design, construction, operation and maintenance periods of the AEC industry, but now it is only at the beginning stage.

1 Introduction

In the past decades, knowledge in the fields of Architecture, Engineering and Construction (AEC) industries mainly comes from experiences and is commonly documented in hard copies or specific electronic databases. Such documents include regulations, standards, manuals and text books [1, 2]. Retrieving information from documents and specific electronic media is, however, difficult and error-prone [3], which brings troubles to extract and reuse the knowledge [4]. With the developments in information technologies such as Building Information Modeling (BIM), Geographic Information System (GIS), Internet of Things (IoT), cloud computing and so on, more and more data are collected in the AEC industry. However, data are not information when they are not useful, regardless of providing knowledges.

Therefore, a lot of studies focused on extracting knowledge from the data via various methods, including ontology, semantic network, data mining algorithms, etc. For example, ontology is adopted to define and to represent the categories, properties and relationships among the concepts in the building industry. The concept of semantic network was proposed in the early 1990s when Berners-Lee et al. [5] extended the concept of World Wide Web (WWW) to handle the data accumulated in network communications without manual works. More specifically, ontology helps transforming the natural-language-based network data to information recognized by computers, which facilitates the process of inquiries or inferences conducted by the computers as it “knows” the concepts and the relationships among them [6]. The semantic network, on the other hand, extracts new knowledge in the aspect of “facts” from the internet via different approaches. These facts are related to each other, and thus can be organized in a series of relationship graphs which is known as Knowledge Graph (KG). The KG combines data and discovers knowledge from different sources by analyzing the grammar, vocabulary and structure characteristics of the texts. Meanwhile, data mining algorithms such as clustering and pattern recognition are used to retrieve the hidden information from big data. All these

✉ Zhen-Zhong Hu
huzhenzhong@tsinghua.edu.cn

¹ Tsinghua Shenzhen International Graduate School, Shenzhen 518055, China

² Department of Civil Engineering, Tsinghua University, Beijing 100084, China

methods are useful in terms of providing new knowledge for the AEC industry.

The concept of BIM technology was proposed in the 1970s [7]. In the recent 2 decades, it has brought great positive impacts to the AEC industry [8]. A BIM model involves physical and functional digital models for all the facilities of the building by applying digital modeling and associated technologies (DMAT) to collect data [9] and adopting unique, readable data standard to support the share of data among participants in different phases within the life-cycle of the project. It should be noted that it improves the collaborations among various participants from different aspects of the project as well. BIM can also integrate domain knowledge and expert methodologies for automated and intelligent applications [10]. Since BIM grows into a mature stage nowadays, a considerable number of BIM systems and corresponding data platforms are applied to the AEC industry [11]. With the growth in data amount, the participants have gradually realized the potential of discovering new knowledge from the data accumulated in the BIM. For example, researchers found that the value of BIM can be improved by adopting semantic network to support the data integration and complicated search requirements among different data sources [12]. Such trend is observed by the very influential organization BuildingSMART who revised its technical roadmap into three long-term levels and introduce the fourth layer to encompass the “semantic search” and “cloud database” domains. Moreover, studies concerning the application of BIM in the knowledge generation/extraction are continuously reported, including ontology-based data management and sharing, knowledge fusion among different domains and automatic logic inference and knowledge retrieve and so on.

Nevertheless, there are still a lot of works remained to be studied in the BIM-based knowledge related domain to improve the generality and efficiency of knowledge management. In order to provide an overview of the state-of-the-art in the knowledge extraction and discovery based on the data accumulated in BIM, to identify possible challenges and to indicate future directions, this paper conducts a systematic review on relevant literature. The following sections are arranged as follows. Section 2 illustrates the methodology and research framework. Sections 3 to 7 discuss the current up-to-date research topics in the aspects of knowledge description, knowledge discovery, knowledge storage and management, knowledge inference and knowledge application, respectively. The last section discusses the challenges and propose the future directions accordingly.

2 Methodology and Framework

The relevant literature is reviewed in a systematic and organized way in 5 steps. First, the scope of the review is determined as knowledge extraction and discovery based

on BIM. Consequently, searches on the existing literature in the scholar database are conducted by combining keywords as TITLE-ABS-KEY (“construction industry”) OR (“BIM”) OR (“AEC”) OR (“construction management”) AND TITLE-ABS-KEY (“ontology”) OR (“semantic web”) OR (“knowledge”) OR (“linked data”). Given the raw database, the less relevant studies are filtered out via examining the title and the abstract manually. Then summarize the high cited authors, journals and keywords, followed by the in-depth study of the main stream literatures to sum up the state-of-art research topics and future directions.

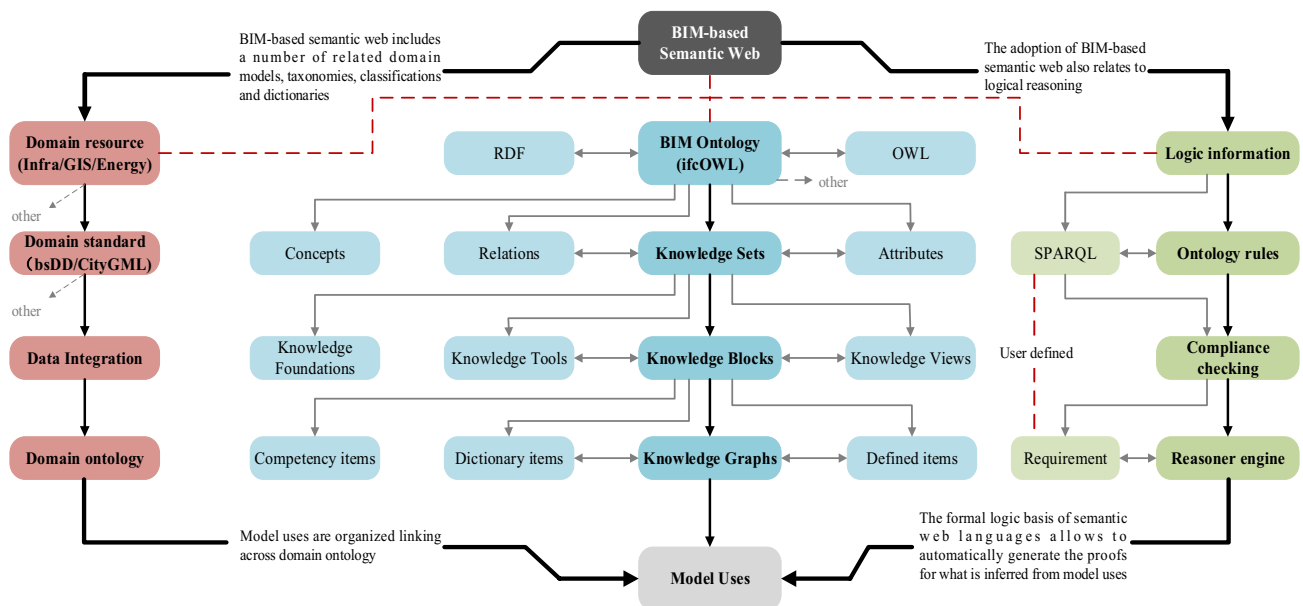
2.1 Research Scope

In order to focus on the BIM-based knowledge extraction and discovery, this review mainly covers the studies in a close relationship with BIM, semantic web, ontology, resource description framework, web ontology language, knowledge management, etc. Particularly, the combination of knowledge domain and BIM has brought great changes to the AEC industries [13]. In the early twenty-first century, domain-knowledge-based semantic technologies were introduced to the AEC industries [14]. For example, Pan et al. [15] and Elghamrawy and Boukamp [16] discussed the extra values that the semantic web would benefit the construction projects. At the same time, some recent reviews discussed the knowledge of BIM [17] and its application to architecture design, energy simulation, intelligent optimization, safety management [18], design and analysis of city spaces, integration of BIM and geographic information system, design codes compliance [19], facility management [20], etc. These topics were summarized according to the Latent Semantic Analysis (LSA) of the themes. Because BIM itself emphasizes the information within a digital model, the reviewed studies mostly focused on how to provide essential information to meet the requirements of specific domains or applications. In order to be organized, the studies are divided into 5 main stream topics, i.e., knowledge description, knowledge discovery, knowledge storage and management, knowledge inference and knowledge application as listed in Table 1.

Each research topic listed above relies on the building information in particular form and some of them even rely on information from different data source. BIM technology is capable to provide the information framework including data standard, data management and data platform to integrate these essential data. Meanwhile, semantic web provides technical framework for knowledge description, query language and inference engine for knowledge extraction and discovery from BIM-based data platform. Therefore, the integration of these two technologies is of significant potentiality as summarized in Fig. 1 for knowledge engineering in AEC industries.

Table 1 Research topic of knowledge extraction and discovery on building information modeling

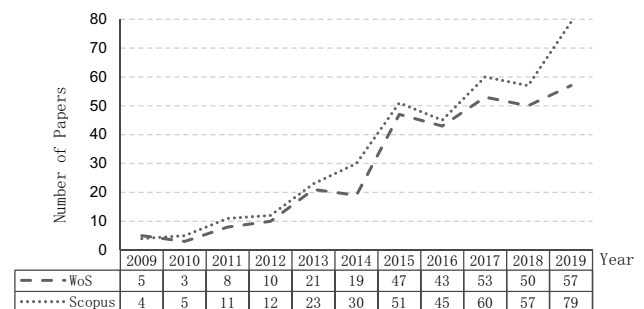
Research topics	Contents
Knowledge description	BIM data standard Semantic framework Transformation and extension of ifcXML and ifcOWL
Knowledge discovery	Knowledge item and knowledge graph AEC entities extraction Relationship extraction
Knowledge storage and management	Storage medium for data and knowledge The framework of knowledge storage Linking knowledge from different sources and domains Data fusion for building energy and facility maintenance Multi-scale information model and multi-layer building knowledge
Knowledge inference	Query language for BIM model and domain knowledge Knowledge inference method including knowledge description, knowledge retrieval and ontology reasoning machine
Knowledge application	For design period For construction period For operation and maintenance period

**Fig. 1** Framework of knowledge extraction and discovery on building information modeling

2.2 Statistics of the Literature

This research mainly examines the literature published from 2009 to 2019 in the Web of Science (WoS) [21] and Scopus [22] database. The numbers of the reviewed papers are illustrated in Fig. 2. Generally, the number of papers in the field is growing year by year especially after 2014, implying that BIM-based knowledge management becomes an important trend attracting scholars concerning information technologies in AEC industries.

According to the filtered literature, the top cited and hence the most influential scholars such as C. Eastman, J.

**Fig. 2** Literature numbers from WoS and Scopus database from 2009 to 2019

Beetz and P. Pauwels, and international academic journals such as Automation in Construction and Advanced Engineering Informatics are summarized in Fig. 3. Figure 4 shows the most active keywords include BIM, ontology and semantic web, etc.

3 Research Topic 1: Knowledge Description

Hannus et al. [23] used a term “island” to depict the gaps between information. When the islands can exchange information, it means that a public data model, instead of specific models, should be established for describing the objects in a common form, regardless of professions, processes and systems. Industry Foundation Classes (IFC) related standards

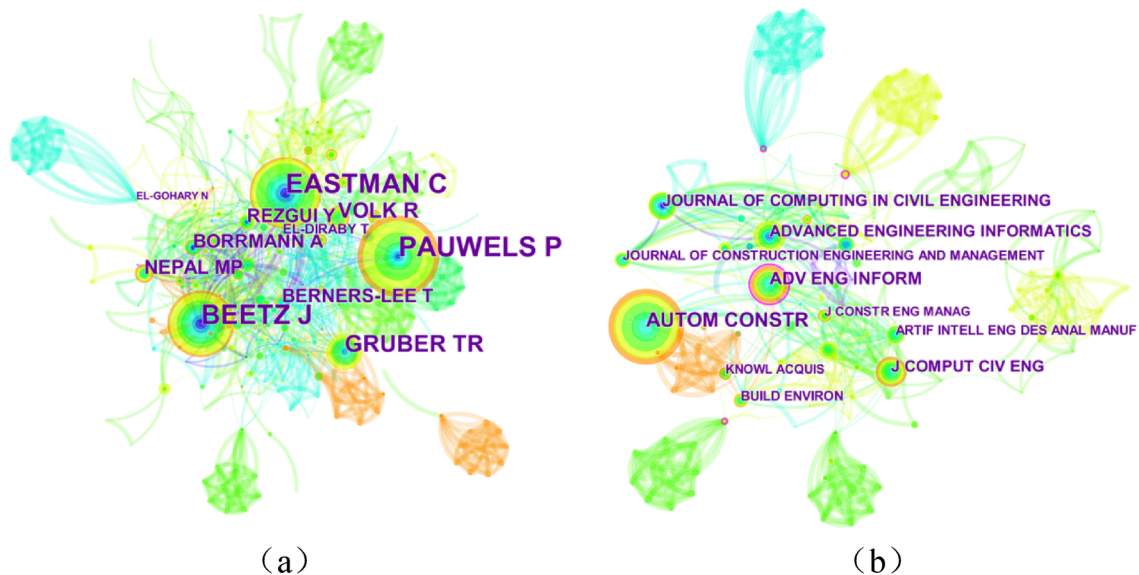


Fig. 3 **a** Top10 authors and **b** Top10 publishers related to knowledge extraction and discovery on building information modeling

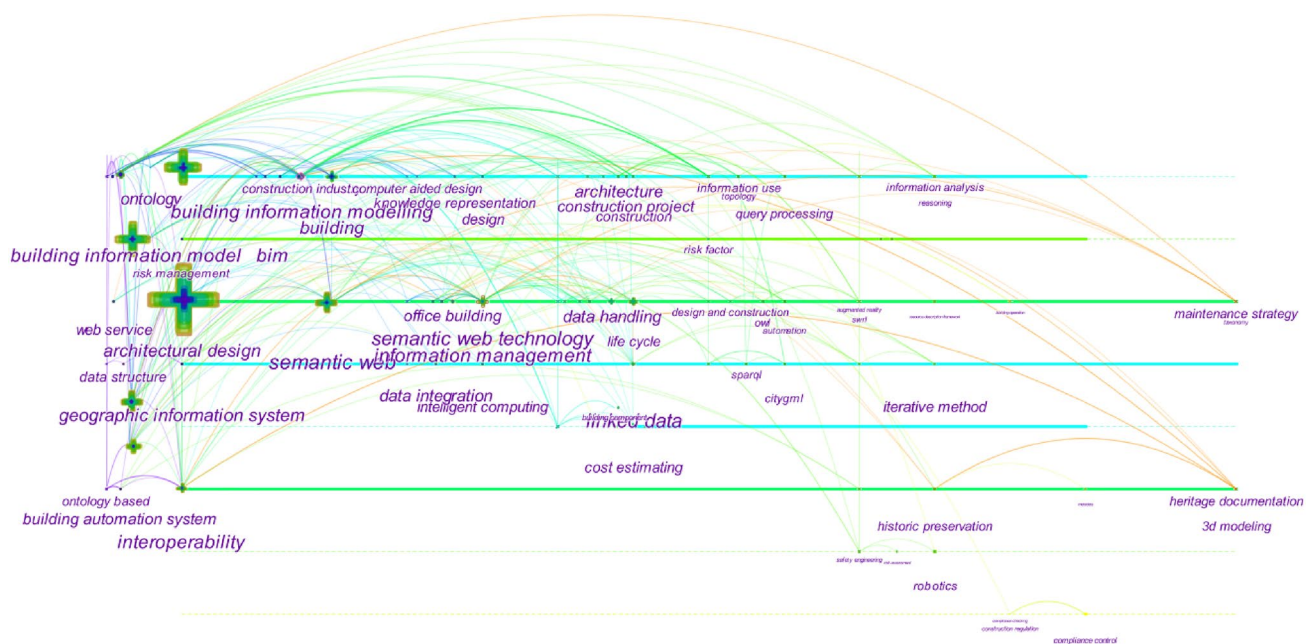


Fig. 4 Key words related to knowledge extraction and discovery on building information modeling

are the model of the kind for building information description and Resource Description Framework (RDF) [24] and Web Ontology Language (OWL) [25] are the models for semantic description.

3.1 IFC Standard Framework for Building Information Description

The IFC standard framework which was proposed by BuildingSMART as an open standard framework for the delivery and support of assets in the built environment, consists of IFC, Data Dictionary (bSDD), Information Delivery Manual (IDM) and Model View Definition (MVD) and BIM collaboration format (BCF) [26]. The IFC is an open and neutral data format for describing BIM and all the elements inside the model, to provide a unique data exchange standard. Nowadays, a lot of BIM compatible software adopt IFC as one of the data exchange formats [27]. The bSDD is developed from International Framework for Dictionaries (IFD), constitute a library of objects and their attributes to identify objects in the built environment and their specific properties regardless of language. The IDM presents a method to define the information exchange requirements through process modeling, so that all participants in the organization know which and when different kind of information has to communicated. The MVD is an IFC view definition which defines a subset of the IFC schema to meet the needs of one or many exchange requirements within the IDM. The MVDs are encoded in a format “mvdXML”, and define allowable values at particular attributes of particular data types. The BCF is an open file XML format “bcfXML” that supports the workflow communication in BIM processes and provides webservice “bcfAPI” for software development to exchange the BCF data. All these standards are nowadays formulating the most important and acceptable BIM standard framework to store the information and knowledge related to the AEC industries.

There are certainly studies that have developed an object-oriented information model [28] or common applications by adopting the IFC framework regardless of platform, machine or data source [29]. Commonly in these studies, the researchers adopted the data modeling language EXPRESS to describe an Entity-Relationship (E-R) model, including several hundreds of object definitions through a tree structure. In current released IFC schema, there are 4 layers, i.e., resource layer, kernel layer, shared layer and domain layer. Each layer has several modules. The resource layer consists of definitions for basic information resources such as materials, geometries and topologies. But these definitions can not be used alone without linking to entities from other layers because they do not contain the Globally Unique Identifier (GUID). While the entities in other 3 layers have GUID. The kernel layer defines the core data models including object

definition (such as the position and geometric appearance of an engineering object and the relationships among these objects). The shared layer defines common entities that can be utilized by various professional domains or processes (such as walls, beams, doors and windows). The domain layer defines professional domain dependent entities for information exchange within the domain (such as steam boilers, fans and dampers in the Heating, Ventilation and Air Conditioning (HVAC) domain).

3.2 RDF and OWL for Semantic Description

The RDF is proposed by the World Wide Web Consortium (W3C). By using a variety of syntax notations and data serialization formats, it aims to describe the metadata data model to serve as the semantic network descriptions. Specifically, the RDF adopts expressions of subject–predicate–object, known as “triples”, to describe resources. It is capable and widely adopted to describe the semantic-based knowledge because the triples can be presented in RDF graph which is a kind of directed label graph. In RDF graph, each node refers to a concept or an object of the real world and the node is identified by a Uniform Resource Identifier (URI). Through the directed linkages between these nodes, the information become readable and reusable by computers.

Being built on the RDF, the OWL [30] is also a family member of knowledge representation languages for authoring ontologies. The RDF graph that built based on OWL concept is called OWL ontologies. The current version of OWL is OWL2. Its profile is summarized in Fig. 5. The ovals upside represents the abstract concepts in ontology and therefore can be considered as an abstract structure or an RDF graph. In fact, any OWL ontology can also be regarded as an RDF graph. The relationship between ontology

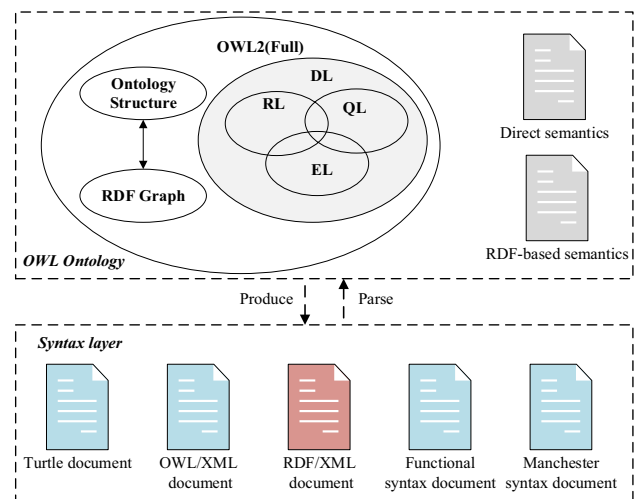


Fig. 5 The OWL2 profiles

structure and RDF graph is prescribed in the RDF graph documents (OWL2/RDF mapping). OWL Lite, OWL DL and OWL Full are the three sub languages for OWL while OWL DL can be further divided into three languages OWL EL, OWL QL and OWL RL. Different language is suitable for different complexity of the inference based on ontologies.

3.3 ifcXML and ifcOWL

Because EXPRESS language lacks semantic information, logic-based languages such as OWL are believed to have advantages in the aspects of knowledge description, semantic data sharing, reusing the existing ontologies and cooperativity among software [31]. As a result, in order to improve the universality and accessibility, IFC also provides another two formats for building information, i.e., IFC-XML and IFC-OWL besides EXPRESS format. IFC-XML files, with a “.ifcXML” suffix, follows the rules of Extensible Markup Language (XML) and the constraints of IFC, ifcXML schema and XML schema. Based on the ifcXML, the ifcOWL language extends the OWL standards to provide BIM-based ontology language and thus improves the scalability problems caused by EXPRESS. Furthermore, in order to integrate the descriptions of both building information and semantic, some scholars proposed platform independent frameworks to transform the IFC data into semantic representations [32]. Such platform provides semantic rich and human readable information by exchanging data from different product information models. Schevers and Drogemuller [33] presented a transition diagram, and Beetz et al. [31] suggested a semi-automatic method to convert the EXPRESS-based IFC data to OWL ontologies. Barbau et al. also proposed regulations for such transition and developed an OWL plugin based on Protégé [34]. Zhang and Issa [35] asserted that by converting IFC to OWL, the information model can be restructured by adopting information technologies, and retrieving information from IFC can be more efficient. Pauwels et al. [36] demonstrated the use of Semantic Web Rule Language (SWRL) to enrich the OWL version for IFC and create the semantic rule checking environment. They also suggested that specific rule ontologies should be developed based on SWRL and connected to the kernel existing ontologies in the AEC industries.

4 Research Topic 2: Knowledge Discovery

It is easier for people, not computers, to understand the building information and the contents behind the big data, thus it is always a prospect to achieve the machine readable and exchangeable information so that the knowledge can be discovered by the computers themselves [37].

4.1 Knowledge Item and KG

Knowledge can be divided into 2 kinds [38], explicit knowledge and implicit knowledge. The explicit knowledge is usually recorded in natural language or readable symbols for communication while the implicit knowledge is gained through incidental activities, or without awareness [39]. In a project, solutions and technical frameworks often rely on experienced participants and such kind of knowledge is also considered as implicit [40]. Most information-based knowledge management studies focus on capturing implicit knowledge in construction projects [41, 42]. Ugwu et al. [43] introduced ontology idea to support the mining, representation and reusing of knowledge for constructability assessment of steel structures, demonstrating that ontology is capable of obtaining domain knowledge, as well as turning the implicit knowledge into explicit. However, capturing implicit knowledge requires huge amount of manual works and lacks of a common way to establish the knowledge from bottom to top. Then the KG and Artificial Intelligence (AI) including deep learning model and Natural Language Processing (NLP) tools are the possible means to make discovering knowledge easier.

The KG was first named for the search engine of Google in 2012. Ever since then it has been promoted to a concept that refers to the network-based semantic database [44]. A KG which contains a series of entities, as well as the attributes and relationships between these entities can be extracted from various data sources (such as websites) through analyzing the text syntax, words and phrases, and the structure characteristics of the text. Compared to semantic network, the KG requires fewer manual intervention because it integrates the algorithms for information acquisition and handling, and is extendable via automatically extracting knowledge from the internet. Besides Google KG, there are now some generic KG such as DBpedia [45] and YAGO [46] and geoscience KG as AEC domain KG [47]. The KG has been widely applied affecting the people's daily lives, particularly in the area of information retrieval [48] and knowledge inference [49]. In AEC industries, researches have been carried out to construct KGs for managing interrelated project information from multiple participants [50] and identifying hazards on construction sites [51]. However, current KG is far from sufficient for the AEC industries. One challenge of generating a domain-specific KG is to integrate different information sources by a universal method.

4.2 Extraction of AEC Entities

The objective of entity extraction is to recognize those AEC domain related entities automatically from text contents. Here the entities refer to those AEC terminologies and will be considered as network nodes of a KG.

In non-structured building information, the process of entity extraction belongs to Named Entity Recognition (NER) in NLP, which has been presented with several technical algorithms. The early methods for NER are rule-based [52] or dictionary-based [53]. Although simple, these methods have found their applications in specific fields of AEC industry, such as recognizing the affected infrastructure and contractor entities in failure reports [54]. Machine-learning methods are then employed to replace the entity recognition to sequence labeling tasks, in which the labels are determined by their probabilities. Some typical machine-learning models for NER include Hidden Markov Model (HMM) [55] and Conditional Random Fields (CRF) [56]. Recently, the accuracy of NER is improved by adopting the deep learning technology, which also labels the sequence but the model is more complex. Recurrent Neural Network (RNN) is one successful example of such technology [57]. The neural in the same layer of an RNN is directly connected to form a circulation so that the dependencies between the words within a sentence can be analyzed easier. Moreover, the RNN can be combined with traditional machine-learning models such as CRF to further improve the accuracy [58]. Attention mechanism can also be compounded with neural networks to improve the prediction performance [59]. Evidence shows that as a branch of RNN, long- and short-term memory bi-directional neural network (Bi-LSTM) solves the problem that a normal RNN cannot “remember parameters” for a long time, and shows the highest accuracy and performance [60], thus it is expected to be potential in entity extraction in the AEC domain as well [61].

In structured BIM models, building entities are defined and represented object-orientally already. The BIM software provides geometric (such as length, width and depth) and non-geometric (such as color, fireproofing grade) parameters for various basic building objects. However, these default parameters only provide the information of an object, not the direct knowledge. Thus, some researches encourage BIM users to define their own parameters to attach knowledge to building objects or projects in the BIM model [62, 63]. For example, in Autodesk REVIT, users can define 2 types of parameters, i.e., project parameters and shared parameters but only the latter ones can be shared between families or projects through an Open DataBase Connection (ODBC). Peng et al. [64] utilized this feature and generated evacuation entities by defining evacuation parameters in REVIT models for public building safety management. Deshpande, Azhar and Amireddy [65] developed a BIM-based knowledge system for users to define parameters representing the human experiences so that the knowledge can be extracted.

4.3 Extraction of Relationships Between AEC Entities

The objective of relationship extraction is to recognize different kinds of relationships between the entities. These relationships become the directed edges to connect the nodes in a KG, forming a net-shaped structure and depicting how the entities work together in logic and physical manners.

In non-structured texts, researchers also proposed similar algorithms as extraction of entities including adopting rule-based and deep-learning models. Particularly, Residual Convolutional Neural Networks (ResCNN) is widely adopted because of its high accuracy and low requirements of manual labeling work [66]. The structure of ResCNN is illustrated in Fig. 6. It consists of convolutional layer, pooling layer, full connection layer and Softmax process. The ResCNN firstly transforms the word sequences into vectors. Then it has shortcuts between several convolutional layer to form residual blocks so that the input can be directly involved in the calculations. The neural network is hence more stable because the learning objective becomes the residual instead of predicted results [67]. The Res-CNN finally predicts the probabilities of each relationships and the one with the highest probability can be regarded as the extracted relationships.

Besides constructing pipeline systems that extract entities and relationships successively, the concept of end-to-end knowledge extraction which achieves entity recognition and relationship discovery in one neural network has emerged in recent years. End-to-end tasks also adopt deep neural models such as Bi-LSTM-CRF networks, while the encoding and decoding process is carefully designed to consider both entity and relationship information [68]. The end-to-end model solves the problem of error propagation that is common in pipeline systems, and is proved to get state-of-the-art performance.

Besides the semantic logic, the extraction of spatial relationships is also important in the building management [69]. Current BIM supported knowledge management gradually concerns extracting the spatial relationships by geometric information within the BIM model [70]. Most BIM models can be converted to IFC structure represented data and IFC describes geometric information by taking Curve2D, GeometricSet and GeometricCutterSet as basic elements. IFC also adopts SurfaceModel and SolidModel to describe the 3D models in surface and solid modes. The SolidModel can be further divided into types such as SweptSolid, B-rep and CSG, etc. According to such decomposition of geometric information, Borrmann discussed the spatial data structures and the definitions of analytic operators [71] and summarized the topologies among points, lines, surfaces and volumes as boundaries, interiors, exteriors and close-loop. Table 2 compares different kinds of operators, their theoretical basis and relationships. The logic chain can also

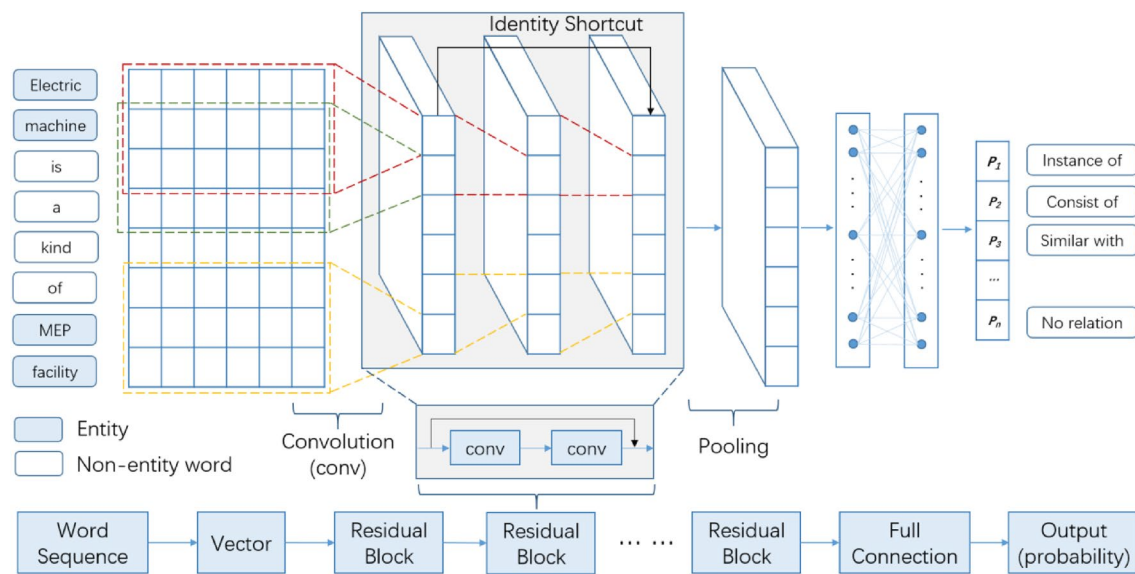


Fig. 6 The structure of the ResCNN model

be automatically generated according to the spatial information within a BIM model and pre-defined identification rules [72].

5 Research Topic 3: Knowledge Storage and Management

Knowledge should be shared between members in an organization and between organizations. Given the goal, knowledge is better stored in a computer readable form. The storage strategy has great impacts on the efficiency of knowledge retrieval and reuse. Only the knowledge is appropriately represented can it be effectively stored.

5.1 Storage Medium for Knowledge

Wang and Meng [75] gave a review on BIM-supported knowledge bases and concluded that BIM have multiple benefits including 3D visualization and collaborative work compared with other IT-based knowledge bases. However, the data and even some knowledge stored in the BIM

database are mostly useful for a specific project, but not for common purpose. In order to reuse such knowledge, some studies tried to establish independent knowledge bases that were linked to BIM model through information standards (such as IFC). For example, Motawa and Almarshad [63] developed a system that separated the knowledge base and BIM model. Ding et al. [76] also presented a system to achieve the idea that knowledge could be collected and integrated in a platform and reused in other relevant projects.

Ontology uses sharing formats to conceptualize domain knowledge but normally it does not support knowledge exchange between domains [40]. Combining ontology and semantic technology, domain knowledge is meaningful to other domains, given that the data is properly interpreted. For example, the knowledge management model developed by Lee and Jeong [77] can transform a specific domain format to a neural one, which is an opposite direction by the semantic mechanism. Even in the same knowledge domain, ontology regulations are heterogeneous [78]. Beside ontology, linked data [79] and fuzzy multi-criteria model [80] can also be a media to link cross-domain knowledge.

Table 2 3D spatial operators

Operators	Theoretical basis	Relationships
Directional operators [73]	Point-set theory notation	<i>NorthOf, southOf, westOf, eastOf,</i>
Topological operators [71]	9-intersection model	<i>Above, below, within, contain, touch,</i>
Metric operators [74]	Point-set theory notation	<i>Overlap, disjoint, equal, mindist, maxdist, isCloser, isFarther, union, intersection, difference</i>
Boolean operators		

The development of ontology framework brings new ways to knowledge retrieval in BIM environments. One example is the storage and management of the knowledge concerning construction risk based on both ontology and BIM [81]. But usually, ontology is responsible for reflecting a set of concepts and the relationships among them for specific domain knowledge. Therefore, some efforts were made on developing a shared ontology to embed various common domain concepts in the BIM environment [82]. The shared ontology can be considered as the semantic agency for the alignment of domain knowledge, so that users can create their own ontology based on the shared ontology.

To achieve the semantic collaboration among various domain knowledge, a common semantic mechanism should be established by the integration of different ontology knowledge in BIM environment [83]. As a result, the integration of BIM and knowledge-based system has become a new technical trend. Deshpande et al. [65] proposed a method to acquire, retrieve and store information and knowledge within BIM, as well as a knowledge classification and propagation framework. Ho, Tserng and Jan [84] developed a BIM-based knowledge sharing management system for both managers

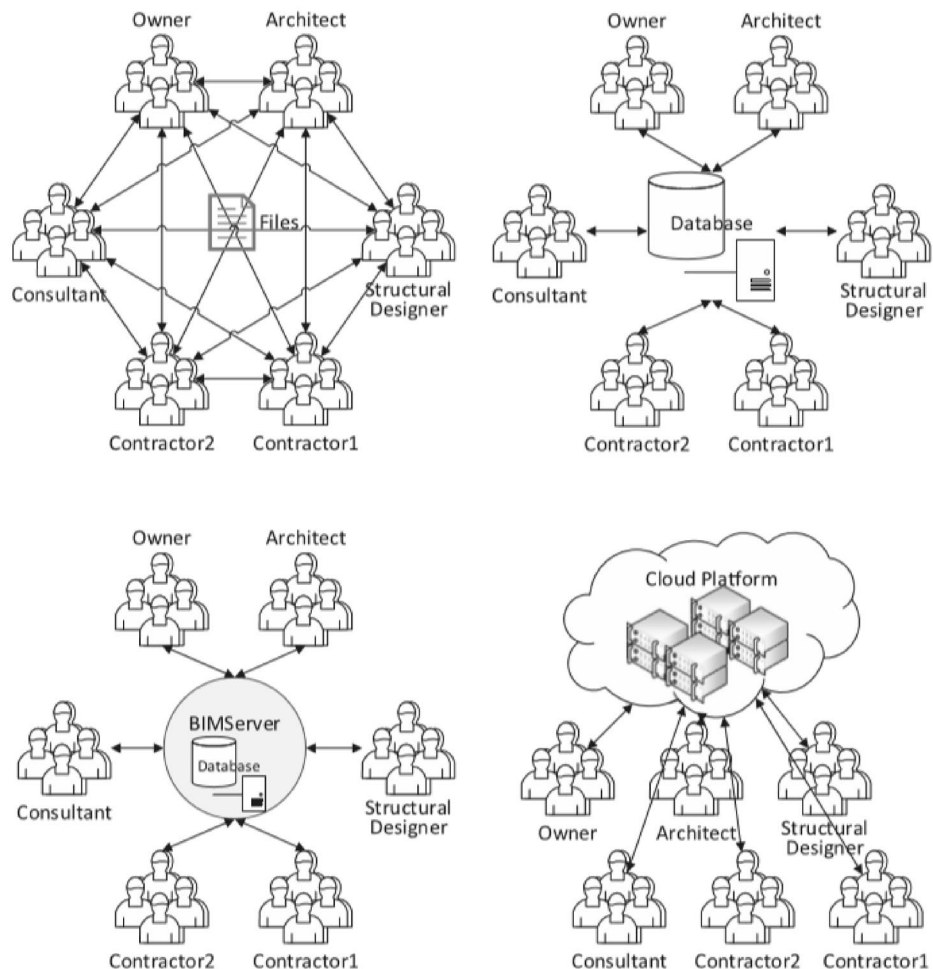
and engineers to share knowledge and experience within BIM environment. Efforts have also been made to link heterogeneous BIM data by constructing an ontology-based vaults database and prompt data sharing among different domains include architecture, engineering, construction and facility management [85].

5.2 Framework of Knowledge Storage

As summarized in Fig. 7, there are 4 typical BIM knowledge storage frameworks, i.e., file-based, central database, single server and cloud server [86].

Most of the file-based frameworks are designed on top of IFC standard and its file formats are frequently IFC-based, such as IFC-SPF with a .ifc suffix [87]. In the central database framework, researchers adopted all kinds of databases including relational database, object-oriented database, key-value database and relational-object database as their storage medias. However, neither file-based nor central database framework contains a generic layer for the use of model and knowledge. Instead, users should develop their own applications according to the functional requirements.

Fig. 7 BIM-based knowledge storage **a** file transfer, **b** central database, **c** single server, and **d** cloud server



A way to improve the data share among participants is to set up a BIM server for users and their applications. This server should provide extra functions based on the BIM data such as model review, 3D visualization, version control and collision detection, etc. Three representative BIM servers are the IFC model server developed by VTT Building and SECOM Co. Ltd. [88], EDM model server by Jotne EPM [89] and Bimserver.org by TNO Netherlands and Eindhoven TU [90]. Since cloud computing technology has the potential to reform the information management for the AEC industry, the cloud computing platform for BIM model management, which is helpful to reduce the cost while providing higher performance [91], has been proposed in the past a few years and become more and more mature nowadays. This idea can be traced back a decade, when Zhang proposed a BIM-based construction Information Integration Platform (BIMIIP) [92]. Based on further researches on data collaboration within BIM systems [93], Zhang et al. developed a prototype system (BIMDISP) to establish a multi-server data-sharing environment [86]. Some typical commercial cloud computing platforms for BIM are Autodesk BIM 360, Cadd Force, BIM9, BIMServer, BIMx and STRATUS, etc. Linking domain knowledge.

5.3 Linking Domain Knowledge

How to manage the domain knowledge that depends on semantic ontology is an emerging research area because such technologies provide collaborative representation for domain knowledge such as data from BIM, GIS, sensors and building automation systems (BAS), etc. [94]. Besides, they also provide linkages between data [95]. These technologies often adopt ifcOWL to link building data from a number of data sources because ifcOWL brings the advantages of (1) providing RDF to present any data type; (2) allowing extensions of logic inference by adopting OWL and (3) linking information graph in a network. The framework of building data linkage across domain knowledge based on ifcOWL is illustrated in Fig. 8.

To integrate the data in BIM, GIS and CAD platforms is also an ongoing research topic, which in some aspects leads the integration of knowledge management [96]. GIS is widely applied in infrastructure projects within their life-cycle. The data standard for GIS such as CityGML is organized by the Open Geospatial Consortium (OGC). The GML model, having close relationship with RDF, is appropriate to link data. Taking advantage of this characteristic, a large number of software and products support effective storage and spatial inquiry of large GIS data set, by introducing the geometric description WKT [97] (Well Known Text) and a query language GeoSPARQL [98].

Particularly in large infrastructure projects, the integration of building data and geographic data is essential to support

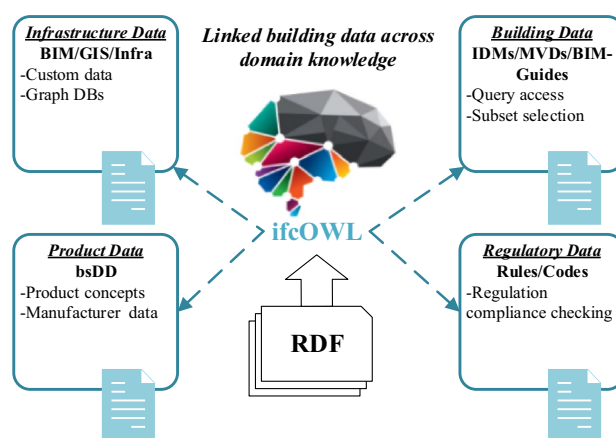


Fig. 8 Linking building data across domain knowledge

project knowledge management and exchange among all participants from multi-disciplines [99, 100]. Mignard et al. [101] developed a Siga3D system to integrate BIM data and GIS data and achieve the city-scale infrastructure management. Beetz and Borrmann [102] introduced linked data and discussed the spatial semantic data exchange, management and related applications to the design, construction and operation management of road projects, and developed a system to integrate various data sources. Zhao, Liu and Mbachu [103] linked similar concepts between BIM and GIS ontologies by ontology mapping and used integrated BIM-GIS data to support highway planning process.

5.4 Building Energy Consumption and Facility Maintenance Knowledge

Knowledge integration management also focuses on information integration across different phases in the building life-cycle [104]. A typical example is constructing a domain ontology schema for the building information and Mechanical, Electrical and Plumbing (MEP) data to provide knowledge for building energy analysis and optimization, especially for applications during design, and operation and maintenance periods.

In the design period, Korman, Fischer and Tatum [105] established an MEP knowledge base to represent the complex compositions of MEP systems by retrieving, analyzing and summarizing related knowledge, based on which they developed a knowledge inference method as well as a prototype system to utilize such knowledge for MEP design and conflict analysis. Olofsson et al. [81] through a case study, discussed the integrated application of BIM and Virtual Design and Construction (VDC) to the collaborative design of MEP system, and proposed an implementation routine for the design process and installation according to the roles of contractor and sub-contractors.

During the operation and maintenance period, the researches focus on the information integration for building performance evaluation and optimization. For example, some integrated delivery technologies for intelligent MEP management in the operation and maintenance phase were proposed based on BIM models [106]; S. Jiang, Wang and Wu [107] achieved the building energy integrated management based on semantic network and Wu [108] proposed an intelligent evaluation system for green building also based on ontology. Dibley et al. [109] developed an OntoFM system to monitor the building in real-time based on multi-agents. In the study, the domain ontology is built up based on ifcOWL and Ontosensor, which is an ontology for sensors themselves. Zhang et al. [110] proposed a comprehensive data model for building performance monitoring by reusing ifcOWL, semantic sensor network (SSN) and building topology ontology (BOT) as references. The model is designed to integrate static facility information and dynamic monitoring data to support the performance management platform. A similar sensing-ontology-based analysis conducted by Marroquin, Dubois and Nicolle [111] was applied to find out the relationships between indoor occupants and building energy consumption. One thing to emphasize in Marroquin's research is that the knowledge learning and inference were employed to provide intelligent analyses. Besides sensing data, construction materials are also capable of supporting the performance analysis of a building by accurately enquiring materials for each element [112]. Knowledge can always be regarded as the kernel to support information retrieval algorithms for energy consumption calculations [113] and most of the energy analyses were based on the kernel building data such as IFC described models, and the energy consumption data [114]. Here, a typical energy data model is SimModel which is widely adopted in energy simulation software for data exchange and share. The SimModel now is included in the building domain ontology [115] and can be transformed to other data model such as RDF graph [116]. Transforming the energy data to RDF model has the advantages in the information acquisition and analysis but further researches are needed in the data exchange standards between IFC and SimModel.

5.5 Multi-layer Building Knowledge Integration and Management

In a large-scale building, due to the huge workload of modeling all the details, some researchers prefer to define multi-scale information models to provide knowledge in various details. In these studies, the term LOD may refer to Level of Details or Level of Development. The former focuses on the details of the model elements, especial the geometric information, while the latter focuses on the details of additional information attached to the model. The lower the LOD is,

the fewer details provided by the model. At the same time, Higher LOD models mostly builds up on top of Lower LOD ones, but any LOD model should be established gradually. In the GIS systems, LOD is also adopted for rapid and multi-scaled visualizing the city-level models [117]. According to the differences between GIS and BIM models, how to map and reuse the models within these two kinds of systems are essential for large-scale projects. For example, Kang and Hong [117] proposed a multi-scale mapping method for BIM and GIS based on semantic and multiprocessing-based screen-buffer scanning including mapping rules. Hu et al. [118] also presented a multi-scale management framework based on a multi-scale model for both construction and facility management of large public buildings. The proposed multi-scale model consists of several macro-, micro- and schematic-scale models. The management framework makes use of the multi-scale model and embeds a data management mechanism, as well as algorithms to transform BIM model to GIS map model, according to multi-scale management requirements.

6 Research Topic 4: Knowledge Inference

6.1 Query Language for Domain Knowledge

Because IFC is built on the base of EXPRESS, which can provide limited query and analysis support for large data set, few query languages are established for IFC [119]. In order to solve this problem, studies have been conducted focusing on the framework of XML or RDF, which provide better supports for query language. Within such frameworks, Structured Query Language (SQL) can be adopted for queries in relational database, XQuery language for XML format, and SPARQL for RDF format.

XML is one of the standard representations of structured knowledge. In most situation, XML is an object-oriented mode and can be adopted in the instance files for data exchange. Nowadays, XML and XML Schema are widely adopted in the BIM environment to record AEC knowledge. Scholars have even adopted XML mapping (ifcXML) for IFC models [120]. Furthermore, in order to analyze and filter XML data, XQuery language is proposed and accepted in the set of W3C standard [121].

SPARQL, a kind of RDF query language, is based on graph structured resource descriptions. It is developed based on semantic network. SPARQL adopts "Select-Where" expressions which is similar to relational database queries except that it is combined with ifcOWL for use. Zhang et al. presented a more BIM-compliant query language based on SPARQL, named BimSPARQL [122]. As shown in Fig. 9, SPARQL as a public interface language, with extensive functions designed for querying data from outside the data

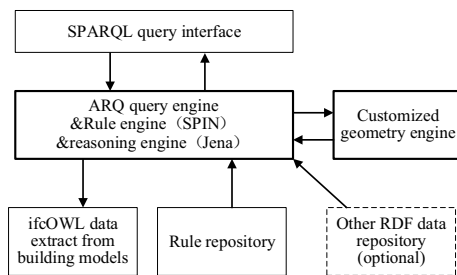


Fig. 9 SPARQL-supported domain ontology retrieval extended flow chart

source, is adopted by W3C [123] and has a series of Application Programming Interface (API) for RDF applications.

OWL, is another kind of ontology language that based of RDF and providing the ontology information framework. Generally, ontology makes data readable in computer and therefore can be deduced by computers according to known knowledge and defined regulations. However, RDF and OWL only support low level inferences. When complicated inferences are needed, more professional languages are required for rule definitions. Such language can customize rules to illustrate complicated logics and support the inference process by using these customize rules. Currently, SWRL, Rule Interchange Format (RIF) and N3logic are such kinds of the rule languages. Here SWRL is a common semantic network rule language based on OWL and Rule ML [124] and can link different data models [125]. The SWRL language has been applied in BIM-based knowledge systems for complicated analysis tasks such as checking whether masonries belong to the same wall by comparing their laying sequences and topologies [126].

It should be emphasized that the formalisms for retrieving information should be established before performing the query. Liu et al. [127] developed the lexicon and syntax of the domain-specific query language, as well as a set of mechanisms to facilitate users formulating query statements, parse the query, and retrieve the information. Their work focused on HVAC systems but seemed to be generic.

6.2 Methods for Domain Knowledge Inference

6.2.1 CWA and OWA for Knowledge Description

When describing the real world, people use two kinds of knowledge descriptions, i.e., Closed World Assumption (CWA) and Open World Assumption (OWA). According to the CWA, any unknown fact is considered as wrong. When the knowledge is complete in the knowledge base, or users have to make decisions according to incomplete knowledge with no risk, CWA is a good assumption for knowledge inference. In contrast, OWA is more open to handle the

incomplete knowledge, and it will return unknown to those undetermined results.

Most existing AEC applications, including BIM systems and public databases, adopt CWA to make decisions according to existing knowledge. But semantic web technologies depend on OWA, because the semantic network is a system with incomplete knowledge. For example, in the semantic network, a common knowledge may imply that A is a sub system of B in the MEP engineering. Then CWA may infer that A is the only one sub system of B because the knowledge base does not show any other sub systems of B. This infer may not be true in many cases. Thus, some researches discussed how to map the information in CWA to those in OWA and extend OWL with integrity constraints [128], ensuring that both CWA and OWA were valuable to AEC software [129].

6.2.2 CBR, RBR and KG for Knowledge Retrieval

Case-Based Reasoning (CBR) and Rule-Based Reasoning (RBR) are adopted in BIM-based knowledge retrieval [130, 131]. For example, knowledge manage systems were developed to carry out schedule evaluation [132] and building maintenance based on CBR in a BIM environment [62]. In these systems, BIM model provides the parameters of elements through IFC standard to CBR. Zhang et al. [131] developed an RBR-based system for safety check, also within the BIM environment. This system was responsible of checking the model contents according to pre-defined rules, and then provided the construability and safety optimization reports automatically. GhaffarianHoseini et al. [133] combined CBR and RBR to support BIM-integrated facility failure management where CBR is used for capturing experiences from past problems and RBR is applied to give explicit solutions. The visualization of BIM was proved to be essential in these CBR or RBR systems. It was also demonstrated that when combining CBR-BIM and RBR-BIM, the BIM-based knowledge management system could also be helpful in safety and security recognition within the lifecycle of an AEC project [134].

KG shows great potential in knowledge retrieval in common areas. Its main process includes the link prediction [135] and entity resolution [136]. The former one refers to the prediction of possible relationships between entities in the KG while the latter refers to the recognition and fusion of entities in case that different entities' names represent a unique object or a single entity's name represents several different objects. Thus, Socher, Chen, Manning and Ng [137] proposed a machine learning and knowledge retrieving mechanism based on Neural Tensor Networks (NTN). Wang, Wang and Guo [138] embedded the KG in the low-dimensional vector space and extended both the physical and

logical rules as constraints, to improve the learning behaviors of the KG and the accuracy of the knowledge retrieval.

6.2.3 Ontology Reasoning Machine for Knowledge Inference

Ontology reasoning is a key technique to realize the semantic network for knowledge inference and thus a number of ontology reasoning machines, such as testing machine by W3C and the one embedded in the semantic web framework by HP labs [139], were developed as the basic, supportive tool for ontology creation and application. These ontology reasoning machines, providing the following two main functions, can retrieve the hidden information in the ontology and then generate new knowledge. (1) Consistency test for the ontology. The consistency test ensures that all the logics between the entities and instances, and the entities themselves are consistent, as well as no conflicts within axiom constraints of the entities, the properties and instances. (2) Inference of implicit knowledge. The creation of ontology usually follows a principle that the entities and their relationships should be as simple as possible, with sufficient information. Thus, when adopting the ontology, inference of implicit knowledge is important to provide knowledge.

Some typical reasoning machines are listed in Table 3. All of these machines maintain the above-mentioned functions and have their own advantages and disadvantages. Among them, Racer [140], Pellet [141] and FaCT++ [142] are all specialized reasoning machines for ontology. Their inference methods rely on traditional description of logics and are optimized by adopting Tableau algorithm. As a result, they show advantages on the inference efficiency but are limited in specific ontology languages and the ability of extension and customization. Specifically, Racer doesn't provide inference for enumeration or user-defined data types, Pellet only supports a few ontology query languages and Jena can only support simple inference rules and can't support OWL inference. Jess [143] is an open source engine which is easier to connect with other applications, adopts generic inference engine to support cross-domain inference. But the efficiency of inference is low.

7 Research Topic 5: Knowledge Application

There are a lot of knowledge applications to the AEC industry within the project lifecycle. Rivera et al. [144] proposed a methodological-technological framework for the emerging concept Construction 4.0, which took BIM-based knowledge application as a core process to achieve automation and digitization. In this section, some typical and effective applications are studied and summarized according to the design, construction and maintenance phases.

7.1 Design Phase

Intelligent design of buildings based on domain knowledge is a fast-growing field in the aspect of application knowledge management to the AEC industry. Computer aided design (CAD) and engineering (CAE) are combined with computational intelligence (CI) to provide integrated application in the design phase. MacCallum once raised a question: "Does intelligent CAD exist?" in 1990 [145] which showed that applying algorithms such as machine learning to the building design was already in place in the early 1990s [146]. The rise of BIM provides an interactive visualization platform and powerful knowledge management system for building design [147], and knowledge-based AI technologies were applied to optimization and configuration [148–150], patterns and philosophies [151–153], intelligent design [154, 155] and interactive design [156], etc. to achieve the goals of automation of manual tasks, personality supports for domain specialist and professional guidance to the amateurs. Moreover, Merrell et al. [157] applied the Bayesian Network based on domain knowledge and real data to stochastically generate a series of layouts and even a complete 3D residential building. Also taking Bayesian model as a kernel element, Fisher et al. [158] presented an approach to arrange 3D furniture objects within a building base on existing examples. These researches explored the possibilities of intelligent building design based on domain knowledge.

Automatic rule-based checking in the AEC industry means assessing the building designs according to various

Table 3 Comparative analysis of various reasoning engines

Name	Reasoner				
	Jena	Jess	Racer	Pellet	FaCT++
Research Institutions	HP Labs	Sandia National Laboratories	Racer Systems GmbH & Co. KG	University of Maryland	University of Manchester
Open Source	✓	✓		✓	✓
Language	Java	Java	Lisp	Java	C++
Algorithm	Rete	Rete	Tableau	Tableau	Tableau
Efficiency	Low	Low	High	High	High

criteria by computer programs [159]. The objective is to automatically check the designed model by interpreting the rules and coding the standards to give the results such as “pass”, “failed”, “warning” or “unknown” [160]. The rules and standards considered in the rule-based checking are mostly the minimum criteria that a building or an infrastructure should follow to ensure the safety and proper functions. At least two steps are needed in such checking process: (1) formalize the building regulations and BIM models to rule models and representation models, respectively, and (2) make sure that the computer program can parse these two models and then perform the rule-based checking according to the rules and the designed object [161]. Recently, a lot of achievements were observed. For instance, efforts were made to propose a regulation compliance checking for buildings [162, 163]. Solihin [164] developed an automated code checking system by adopting IFC model and Express Data Manager (EDM) to evaluate the code compliance in Singapore. Autocodes, developed by Fiatch [165], is capable of code compliance checking for American building according to Existing BIM Standards and Guidelines. Besides, some special checks have also been considered including curriculum problems, spatial requirements and special site needs. Han, Kunz and Law [166] proposed a hybrid approach to compliance analysis for disabled access, combining the rule coding and performance-based method. Lee et al. [167] built a CBR method to combine with rule checking process and give recommendations after the checking process. Kadolsky et al. [168] showed a case to illustrate how to define the building information management rules based on ontology and then Hu [169] presented an automatic fire safety checking in building design based on BIM and ontology. Baumgärtel et al. [170] discussed how to preserve heat in building design by using rules.

7.2 Construction Phase

Knowledge provides an innovative way to better support the construction management in the sub divisions of cost estimation [171], safety management [134, 172], and computer-aided construction [173, 174] and so on.

Specifically, when AI technologies are introduced to the construction management, the safety, cost and schedule can be optimized. Sigalov and König [175] presented a process pattern recognition method for BIM-based construction schedules, to solve the problem of manual definition of proper and application-specific process templates, by estimating the similarities in construction schedules. Genetic algorithm (GA) [176] and resource constraints [177–179], spatial constraints [175, 180] can also be adopted to generate and optimize the construction schedule based on the knowledge provided by the BIM model and environment. The prediction of schedule is also possible for specific construction

activities [181]. A systematic review on BIM-enabled risk management applications was carried out by Ganbat et al. [182], and their conclusions shows that timely uploading, recording and checking of data is the key to reduce potential risk.

Construction cost is considered the fifth dimension (5D) of BIM [183], and knowledge also substantiates the results of model-based cost estimation and process optimization. For example, de Soto and Adey [171] presented a resource requirement prediction model to estimate the construction cost according to the detailed material prices. In some researches, BIM and knowledge were combined to select appropriate tower cranes [173] and make a reasonable layout planning for the tower cranes [174] during construction. Equipment travel path can also be optimized to avoid obstacle and reduce construction cost based on BIM and construction knowledge [184]. Most of these researches adopted existing AI algorithms to generate the knowledge hidden in the data or regular activities.

Moreover, in the area of safety and risk management, based on the semantic regulation checking, Zhang et al. [185] proposed a construction safety knowledge for job hazard analysis, which determines the safety issues in the aspects of tasks, activities and resources by transforming the Tekla structural model to RDF graph and combining the ontology and SWRL regulations. Ding et al. [76] constructed a semantic network based on an ontology-described BIM-supported management framework for construction safety knowledge, so as to generate the safety mappings and their relationships. Then by searching the semantics, the application knowledge is combined to a BIM object. With these ideas, knowledge-based BIM system can be developed to capture and to store various types of information and knowledge from different participants during construction [62].

7.3 Operation and Maintenance Phase

Knowledge is popular and of high expectations during the operation and maintenance of an AEC project, particularly in the aspects of safety management [186], automatic control [187], energy consumption management [188, 189] and decision-making supports [62, 190]. In the following three aspects of applications are discussed in detail.

Intelligent answering system is a common knowledge-based application in the operation and maintenance period. A typical intelligent answering system have at least 3 modules, i.e., information handling, question indexing and answer recommending. The system should at first carry out semantic analysis on the queries sent by the users. Specifically, based on the semantic knowledge base, the sentences are preprocessed by analyzing the lexical, syntactic and grammar to extract the semantic in the form of statements and text collections. A BIM-based dialogue system for users

to seek answers for building maintenance problems has been constructed by Motawa [191]. In such system, the domain KG is essential to provide comprehensive and deeply related summarization for the answers. Corry et al. [190] proposed an approach to use semantic web technologies to access soft AEC data from social medias, personal communications, mobile networks, indoor locations and financial reports etc., to build up a knowledge base. Through the integration of BIM and semantic network, the building information can be accessed together with other open data sources. This building information includes material, system, occupation information and even the weather. Such information is helpful to predict the energy consumption, future material, and the relationships between environment, people and cost [189]. Particularly in the energy consumption domain, more researches have been carried out in the aspects of human behavior modeling for energy saving in residential buildings [192], and semantic analysis [193] and decision support models [194] for energy efficiency in smart house environments.

Fault detection and diagnostics (FDD), usually consists the processes of fault detection, fault diagnostics and fault recognition, is another knowledge-based application in this period. The methods for FDD can be divided into two categories, i.e., model-based FDD and data-driven FDD [195]. The model-based FDD calculated the deviations between the predictions by the proposed model and the measured values, followed by the comparison of such deviations and pre-defined thresholds to assert the faults. The data-driven FDD, however, extracted the characteristics of the historical accumulated data and then considered these characteristics as the prior knowledge to detect the faults. Nevertheless, with the development of the advanced MEP facilities, the complexity is an obstacle to detect faults and requires skilled professionals to deal with the data [196]. Recently, some researches took advantages of the information integration in BIM to provide the FDD model with information stored in the BIM models, so as to reduce the dependence on professionals. For example, Liu et al. [197] proposed an integrated information framework for automated performance analysis of HVAC systems and identified the requirements for the framework of self-managing HVAC systems [198]. The framework was presented based on the existing standards such as IFC, sensor modeling language (SensorML) and BACnet and so on to establish the automatic retrieval and integration mechanism of related information. Then combined with existing algorithms, the automatic FDD could be achieved. Besides integrated framework, some information knowledge bases for HVAC systems based on BIM were also proposed to improve the efficiency and accuracy of the FDD process [199, 200]. Dong et al. [201] focused on the data modeling process of the BIM-based quantified FDD method, and built up an information model for FDD by

retrieving and integrating the static information of facilities, topologies, rooms and spaces from IFC and green building XML (gbXML), and dynamic monitoring information from sensors from BACnet protocol. A standalone FDD system was developed to locate the defective facilities in the BIM environment according to the monitoring temperatures of inside and outside the rooms, and the minimum message length principle [202]. These researches proved that knowledge, particularly integrated information is useful for facility management and FDD.

Building energy management is another area that relies on knowledge. The Building Energy Management System (BEMS) is a kind of system that developed to improve the building energy efficiency by collecting and analyzing the current energy consumption data. Thus technically, a BEMS should integrate the BAS and related operating and energy parameters, as well as the human behaviors and environment parameters within the building, to provide knowledge for making decisions. A BEMS has at least three physical layers, i.e., sensor layer, computing layer and application layer [203] and four modules, i.e., sensor and driver module, middleware, handling engine and user interface [204]. The sensor and driver exchange information between digital infrastructure and the physical environment; the middleware integrates generic data interface; the handling engine is responsible to retrieve and analysis the collected data, which include the environment data (such as temperature, humidity and carbon dioxide) and the human behavior (such as working, having a meeting and taking a break) [205]. The user interface interacts with the user after transforming the analysis result to a kind of knowledge. BIM usually integrates domain knowledge of energy management and can be used as knowledge sources to construct BEMS. A transformation workflow from IFC-based BIM to BEMS has been proposed [206].

8 Discussions and Future Directions

Even though some of the reviewed publications do not accurately use the terms of “knowledge” or “BIM”, the ideas of data integration, data-driven approach, semantic analysis, knowledge discovery, etc. have been widely accepted and show that there’s a growing interest to combine the knowledge science with the BIM technology to provide better support for the AEC industry. But the development seems to be at the preliminary stage and lacks of the integral, systematic and generic achievements. The review reveals that we are still facing at least the following four challenges.

1. Lack of fundamental innovation. Fundamental innovation is a boost for the revolution of the technology to level up an industry. However, knowledge science is a generic term in most of the disciplines and the funda-

mental principles mainly comes from the biology and computer science domains. That is one of reasons that AI, KG etc. show more powerful influence on the information and biological fields. BIM is for building and civil engineering but the idea comes from the machinery industry. Compared to producing an industrial equipment or components, the lifecycle management of buildings and infrastructures is far more complicated. As a result, it is a common view that BIM does not go beyond the product information model (PIM) in the machinery industry. The AEC industry adopts a lot of newly technologies but which one is this industry desperately struggling for remains unclear.

2. Lack of accurate on-site data. The BIM platforms or systems provide an integrated media to store and manage the big data collected within the lifecycle of a project. They also arouse people's attitude on the AEC data. However, because of the low level of informationization and the inaccuracy of collected data, the AEC industry is in a situation of lacking accurate on-site data, regardless of data-driven project management. This drawback is gradually changing since more and more advanced information technologies are applied, but without running in a right track for a period of time. The data can not be fully dependable, neither are the ways to manage the data or discover knowledge from the data.
3. Lack of consolidated knowledge. Ontology, semantic network and KG provide managers with knowledge, but decisions in the AEC industry nowadays are mostly based on intuition or experience because the knowledge is non-unique, incomplete and hard to use. The experience can be considered as a kind of knowledge but with no clear expressions. How to accumulate and then to represent knowledge in a computer-readable and human-reusable way in order to aid project decisions, is still out of good solution. Accumulated, well-organized, accurate represented, retrieval-supported knowledge is not yet ready to revolve the activities within the project lifecycle.
4. Insufficient level of application. It is of potential to combine knowledge with BIM because BIM provides the environment and tools to integrate and manage data while knowledge science can not only discover knowledge from the BIM model and its data but also promoted ways to make use of the knowledge. A lot of trials have been run in diverse phases and multiple aspects but the conclusion can still be drawn that partial application to projects, without a clear and systematic roadmap or solution, does not fully show the bright future. In fact, most of the applications are kind of proving the effectiveness of the information technology, instead of promoting the building or the infrastructure to a more intelligent level.

Facing these challenges, there are a lot of future works to be carried out. Some of the future directions are summarized below.

1. Expand the semantic network and KG. The larger the semantic network and the KG are, the more intelligence people can be benefited from. Thus, one essential way to improve the intelligence level of the AEC industry is to expand the semantic network and KG by creating and managing various domain ontologies, regulation sets and open them to the public. However, manual generation of knowledge is unacceptable. Therefore, new methods on automatic generation and update of semantic network and KG are of great importance. Furthermore, more accurate NLP and more efficient knowledge enquiries should be taken into account considering the explosion of the knowledge.
2. Connected with other advanced technologies. Knowledge science and BIM are considered sub-divisions of information technology. They can be further integrated with other advanced technologies. For example, AI algorithms such as deep learning and data mining are helpful to generate new knowledge from existing data. Cloud computing can enhance the efficiency of knowledge storage, as well as the computational ability of knowledge retrieval and analysis. IoT supports data acquisition and immediate feedback to the environment, VR/AR/MR technologies provide more immersive environment to sense the new world, etc. Only with the combination of these technologies, the knowledge science has the potential to explode into a new generation.
3. Improve the information platform. Another way to make progress is to improve the knowledge management systems and BIM systems. Most current knowledge management systems focus on the storage and retrieval of knowledge, lacking of the ability to generate new knowledge according to accumulated big data with reliable rules. Moreover, knowledge management systems are still yet popular in real project management. Nowadays there are many BIM systems provided by a lot of software companies and research institutes., however, because of the diversity of AEC projects, these systems are either generic but failed in deep penetrations in management or vice versa. On the other hand, BIM systems are still developing in the aspects of distribution of data, integration of information, and centralization of management activities, etc.
4. Towards the high intelligent buildings and infrastructures. The BIM technology integrates, shares and manages the data of buildings and infrastructures, at the same time knowledge science widens the road of applying these data. The combination of these two technologies can also predicts the development trends of

intelligence in the AEC industry. The intelligence here not only includes providing a more intelligent way for making use of the building or the infrastructures, but also enhancing the perceptible and controllable environment, and providing a more knowledge-driven methods in managing the lifecycle of the project. With these development, automatic design, construction robotics, and smart operations are reachable in the coming years to bring great improvements to the industry and peoples' daily lives.

5. Vitalize the potential of the accumulated knowledge. It is a common sense that knowledge is not playing a crucial role in the AEC industry even though the information technologies have been so advanced today. With the development of BIM in the past decade, the value of information has been realized by all the participants and thus data are accumulated, information is managed and knowledges are generated, gradually. Just as the accumulated data provides new knowledge to the industry, accumulated knowledges are capable to reform the cooperation, activity and more aspects of the industry. A knowledge-driven intelligent world is coming.

9 Conclusions

In the past decades, knowledges for AEC industries mainly come from experiences and are mostly recorded in documents in paper or electronic forms. In order to make use of these knowledge, a lot of researches focused on retrieving the knowledges by applying various of methods including ontology, semantic network, data mining algorithms, etc. These methods rely on valuable data. BIM, seems to be a valuable media to provide information because it provides physical and functional digital models for all the facilities within the lifecycle of the project by adopting unique, readable data standard. Therefore, the combination of the knowledge science with BIM shows great potential. Based on the review of existing publications, this research summarizes the latest research achievements of these two technologies, in the aspects of knowledge description, knowledge discovery, knowledge storage and management, knowledge inference and knowledge application.

By thoroughly studying of these publications, it shows that the ideas of data integration, data-driven approach, semantic analysis, knowledge discovery, etc. have been widely accepted to provide better support for the AEC industry. But the development seems to be at the preliminary stage and lacks of integral, systematic and generic achievements. This study identifies 4 major challenges for current situation, i.e., lack of fundamental innovation, lack of accurate on-site data, lack of consolidated knowledge, and insufficient level of applications.

Finally, this study predicts the future directions for the development of knowledge-driven intelligent AEC industry, including the aspects of expanding the semantic network and KG, being connected with other advanced technologies, improving the information platform, moving towards the highly intelligent buildings and infrastructures, and vitalizing the potential of the accumulated knowledge.

According to the review and discussions above, it can be considered that the research of combing knowledge science with BIM, particular in the area of knowledge extraction and discovery on BIM, is still in the very beginning stage with a lot of challenges to overcome. However, the further research and development show great potential to reform the AEC industry to provide a more intelligent environment for people.

Acknowledgements This research was supported by the National Natural Science Foundation of China (No. 51778336) and the Tsinghua University-Glodon Joint Research Centre for Building Information Model (RCBIM).

Funding This study was funded by the National Natural Science Foundation of China (No. 51778336) and the Tsinghua University-Glodon Joint Research Centre for Building Information Model (RCBIM).

Declarations

Conflict of interest On behalf of all authors, the corresponding author states that there is no conflict of interest.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

References

1. Kasvi JJJ, Vartiainen M, Hailikari M (2003) Managing knowledge and knowledge competences in projects and project organisations. *Int J Proj Manag* 21(8):571–582
2. Rezgui Y (2001) Review of information and the state of the art of knowledge management practices in the construction industry. *Knowl Eng Rev* 16(3):241–254
3. Hari S, Egbu C, Kumar B (2005) A knowledge capture awareness tool: an empirical study on small and medium enterprises in the construction industry. *Eng Constr Archit Manag* 12(6):533–567
4. Wang H, Meng X (2018) Transformation from IT-based knowledge management into BIM-supported knowledge management: a literature review. *Expert Syst Appl* 121:170–187

5. Berners-Lee T, Hendler J, Lassila O (2001) The semantic web. *Sci Am* 284(5):28–37
6. Kim K, Kim H, Kim W et al (2018) Integration of ifc objects and facility management work information using Semantic Web. *Autom Constr* 87:173–187
7. Eastman C, Teicholz P, Sacks R et al (2011) BIM handbook: a guide to building information modeling for owners, managers, designers, engineers and contractors. Wiley
8. Shou W, Wang J, Wang X et al (2015) A comparative review of building information modelling implementation in building and infrastructure industries. *Arch Comput Methods Eng* 22(2):291–308
9. Lee C-Y, Chong H-Y, Wang X (2018) Streamlining digital modeling and building information modelling (BIM) uses for the oil and gas projects. *Arch Comput Methods Eng* 25(2):349–396
10. Pezeshki Z, Ivani SAS (2018) Applications of BIM: a brief review and future outline. *Arch Comput Methods Eng* 25(2):273–312
11. Botton C, Rivest L, Ghnaya O et al (2020) What is at the Root of Construction 4.0: a systematic review of the recent research effort. *Arch Comput Methods Eng* 1–20
12. Shen L, Chua DKH (2011) Application of building information modeling (BIM) and information technology (IT) for project collaboration. In: Proceedings of the international conference on engineering, project, and production management
13. Pauwels P, Törmä S, Beetz J et al (2015) Linked data in architecture and construction. *Autom Constr* 2015(57):175–177
14. Pauwels P, Zhang S, Lee Y-C (2017) Semantic web technologies in AEC industry: a literature overview. *Autom Constr* 73:145–165
15. Pan J, Anumba CJ, Ren Z (2004) Potential application of the semantic web in construction. In: Proceedings of the 20th annual conference of the association of researchers in construction management (ARCOM). Heriot-Watt University Edinburgh, UK, pp 923–929
16. Elghamrawy T, Boukamp F (2008) A vision for a framework to support management of and learning from construction problems. In: Proceedings of the 25th international conference on formation technology in construction: improving the management of construction projects through IT adoption, Santiago, Chile
17. Yalcinkaya M, Singh V (2015) Patterns and trends in building information modeling (BIM) research: a latent semantic analysis. *Autom Constr* 59:68–80
18. Wu S (2018) Integrating Ontology and NLP for Automated Construction Process Safety Rule Checking in 4D BIM. South China University of Technology
19. Zhou H (2017) Semantic method research supporting BIM model-based compliance checking. Dalian University of Technology
20. Chen G (2017) A BIM and ontology-based approach for the management of building operation and maintenance. *J Inf Technol Civ Eng Archit* 9(4):67–73
21. Clarivate. Web of Science [EB/OL] (2018) [2018–02–25]. <https://www.webofknowledge.com>
22. Elsevier B.V. Scopus [EB/OL] (2018) [2018–03–23]. <https://www.scopus.com/>
23. Hannus M, Penttilä H, Silén P (1996) Islands of automation in construction. *Constr Inf Highw* 198:20
24. Schreiber G, Raimond Y RDF 1.1 Primer—W3C Working Group Note 24 June 2014[EB/OL](2014). <http://www.w3.org/TR/2014/NOTE-rdf11-primer-20140624/>
25. Hitzler P, Krötzsch M, Parsia B et al OWL 2 web ontology language primer (second edition)—W3C recommendation 11 December 2012[EB/OL](2012). <http://www.w3.org/TR/2012/REC-owl2-primer-20121211/>
26. BuildingSMART. Home—Welcome to buildingSMART-Tech.org[EB/OL](2016)[2016–05–06]. <http://www.buildingsmart-tech.org/>
27. Lee Y-C, Eastman CM, Solihin W et al (2016) Modularized rule-based validation of a BIM model pertaining to model views. *Autom Constr* 63:1–11
28. Lucas J, Bulbul T, Thabet W (2013) An object-oriented model to support healthcare facility information management. *Autom Constr* 31:281–291
29. Vanlande R, Nicolle C, Cruz C (2008) IFC and building life-cycle management. *Autom Constr* 18(1):70–78
30. Motik B, Grau BC, Horrocks I et al OWL2 web ontology language reference profiles (second edition)—W3C recommendation 11 December 2012[EB/OL](2012). <http://www.w3.org/TR/owl2-profiles/>
31. Beetz J, Van Leeuwen J, De Vries B (2009) IfcOWL: a case of transforming EXPRESS schemas into ontologies. *Artif Intell Eng Des Anal Manuf* 23(1):89–101
32. Barbau R, Krima S, Rachuri S et al (2012) OntoSTEP: enriching product model data using ontologies. *Comput Aided Des* 44(6):575–590
33. Schevers H, Drogemuller R (2005) Converting the industry foundation classes to the web ontology language. In: 2005 First international conference on semantics, knowledge and grid. IEEE, p 73
34. Knublauch H, Fergerson RW, Noy NF et al (2004) The Protégé OWL plugin: an open development environment for semantic web applications. In: International semantic web conference. Springer, pp 229–243
35. Zhang L, Issa RR (2011) Development of IFC-based construction industry ontology for information retrieval from IFC models. In: Proceedings of the 2011 Eg-Ice workshop, University of Twente, The Netherlands
36. Pauwels P, Van Deursen D, Verstraeten R et al (2011) A semantic rule checking environment for building performance checking. *Autom Constr* 20(5):506–518
37. Veltman KH (2001) Syntactic and semantic interoperability: new approaches to knowledge and the semantic web. *New Rev Inf Netw* 7(1):159–183
38. Smith EA (2001) The role of tacit and explicit knowledge in the workplace. *J Knowl Manag* 5(4):311–321
39. Alavi M, Leidner DE (2001) Knowledge management and knowledge management systems: conceptual foundations and research issues. *MIS Q* 25:107–136
40. Kazi AS (2005) Knowledge management in the construction industry: a socio-technical perspective. Igi Global
41. Kivrak S, Arslan G, Dikmen I et al (2008) Capturing knowledge in construction projects: knowledge platform for contractors. *J Manag Eng* 24(2):87–95
42. Lin Y-C, Lee H-Y (2012) Developing project communities of practice-based knowledge management system in construction. *Autom Constr* 22:422–432
43. Ugwu OO, Anumba CJ, Thorpe A (2005) Ontological foundations for agent support in constructability assessment of steel structures—a case study. *Autom Constr* 14(1):99–114
44. Paulheim H (2017) Knowledge graph refinement: a survey of approaches and evaluation methods. *Semantic web* 8(3):489–508
45. Lehmann J, Isele R, Jakob M et al (2015) DBpedia—a large-scale, multilingual knowledge base extracted from Wikipedia. *Semantic Web* 6(2):167–195
46. Fabian MS, Gjergji K, Gerhard W Yago: a core of semantic knowledge unifying wordnet and Wikipedia. In: 16th International world wide web conference, WWW

47. Wang C, Ma X, Chen J et al (2018) Information extraction and knowledge graph construction from geoscience literature. *Comput Geosci* 112:112–120
48. Eder JS Knowledge graph based search system. Google Patents, 2012(2012–06–21)
49. Xiong W, Hoang T, Wang WY (2017) DeepPath: a reinforcement learning method for knowledge graph reasoning. *arXiv preprint <http://arxiv.org/abs/1707.06690>*
50. Rasmussen MH, Lefrançois M, Pauwels P et al (2019) Managing interrelated project information in AEC Knowledge Graphs. *Autom Constr* 108:102956
51. Fang W, Ma L, Love PED et al (2020) Knowledge graph for identifying hazards on construction sites: integrating computer vision with ontology. *Autom Constr* 119:103310
52. Rau LF (1991) Extracting company names from text. In: [1991] Proceedings. The seventh IEEE conference on artificial intelligence application. IEEE, pp 29–32
53. Rindfleisch TC, Tanabe L, Weinstein JN et al (1999) EDGAR: extraction of drugs, genes and relations from the biomedical literature. In: *Biocomputing 2000*. World Scientific, pp. 517–528
54. Zhou S, Ng ST, Yang Y et al (2020) Delineating infrastructure failure interdependencies and associated stakeholders through news mining: the case of Hong Kong's water pipe bursts. *J Manag Eng* 36(5):4020060
55. Zhou G, Su J (2002) Named entity recognition using an HMM-based chunk tagger. In: *Proceedings of the 40th annual meeting on association for computational linguistics*. Association for Computational Linguistics, pp 473–480
56. McCallum A, Li W (2003) Early results for named entity recognition with conditional random fields, feature induction and web-enhanced lexicons. In: *Proceedings of the seventh conference on Natural language learning at HLT-NAACL 2003-Volume 4*. Association for Computational Linguistics, pp 188–191
57. Khashabi D On the recursive neural networks for relation extraction and entity recognition[EB/OL](2013)[2019–11–06]. <http://hdl.handle.net/2142/46992>
58. Li L, Jin L, Jiang Z et al (2015) Biomedical named entity recognition based on extended recurrent neural networks. In: *2015 IEEE international conference on bioinformatics and biomedicine (BIBM)*. IEEE, pp 649–652
59. Jin Y, Xie J, Guo W et al (2019) LSTM-CRF neural network with gated self attention for Chinese NER. *IEEE Access* 7:136694–136703
60. Huang Z, Xu W, Yu K (2015) Bidirectional LSTM-CRF models for sequence tagging. *arXiv preprint <http://arxiv.org/abs/1508.01991>*
61. Leng S, Hu Z-Z, Luo Z et al (2019) Automatic MEP knowledge acquisition based on documents and natural language processing. In: *Proceedings of the 36rd CIB W78 conference*
62. Meadati P, Irizarry J (2010) BIM—a knowledge repository. In: *Proceedings of the 46th annual international conference of the associated schools of construction*
63. Motawa I, Almarshad A (2013) A knowledge-based BIM system for building maintenance. *Autom Constr* 29:173–182
64. Peng Y, Li S-W, Hu Z-Z (2019) A self-learning dynamic path planning method for evacuation in large public buildings based on neural networks. *Neurocomputing* 365:71–85
65. Deshpande A, Azhar S, Amireddy S (2014) A framework for a BIM-based knowledge management system. *Procedia Eng* 85:113–122
66. Huang YY, Wang WY (2017) Deep residual learning for weakly-supervised relation extraction. *arXiv preprint <http://arxiv.org/abs/1707.08866>*
67. He K, Zhang X, Ren S et al (2016) Deep residual learning for image recognition. In: *Proceedings of the IEEE conference on computer vision and pattern recognition*
68. Dai D, Xiao X, Lyu Y et al (2019) Joint extraction of entities and overlapping relations using position-attentive sequence labeling. In: *Proceedings of the AAAI conference on artificial intelligence*
69. Zhou Y, Hu Z, Lin J et al (2019) A review on 3D spatial data analytics for building information models. *Arch Comput Methods Eng* 27:1–15
70. Borrmann A, Van Treeck C, Rank E (2006) Towards a 3D spatial query language for building information models. In: *Proceedings of joint international conference of computing and decision making in civil and building engineering (ICCCBE-XI)*
71. Borrmann A, Rank E (2009) Topological analysis of 3D building models using a spatial query language. *Adv Eng Inform* 23(4):370–385
72. Xiao Y-Q, Li S-W, Hu Z-Z (2019) Automatically generating a MEP logic chain from building information models with identification rules. *Appl Sci* 9(11):2204
73. Borrmann A, Rank E (2009) Specification and implementation of directional operators in a 3D spatial query language for building information models. *Adv Eng Inform* 23(1):32–44
74. Borrmann A, Schraufstetter S, Rank E (2009) Implementing metric operators of a spatial query language for 3D building models: octree and B-Rep approaches. *J Comput Civ Eng* 23(1):34–46
75. Wang H, Meng X (2019) Transformation from IT-based knowledge management into BIM-supported knowledge management: a literature review. *Expert Syst Appl* 121:170–187
76. Ding LY, Zhong BT, Wu S et al (2016) Construction risk knowledge management in BIM using ontology and semantic web technology. *Saf Sci* 87:202–213
77. Lee J, Jeong Y (2012) User-centric knowledge representations based on ontology for AEC design collaboration. *Comput Aided Des* 44(8):735–748
78. Abanda FH, Tah JHM, Keivani R (2013) Trends in built environment semantic Web applications: Where are we today? *Expert Syst Appl* 40(14):5563–5577
79. O'Donnell J, Corry E, Hasan S et al (2013) Building performance optimization using cross-domain scenario modeling, linked data, and complex event processing. *Build Environ* 62:102–111
80. Martínez-Rojas M, Marín N, Miranda MAV (2016) An intelligent system for the acquisition and management of information from bill of quantities in building projects. *Expert Syst Appl* 63:284–294
81. Olofsson T, Lee G, Eastman C et al (2007) Benefits and lessons learned of implementing building virtual design and construction (VDC) technologies for coordination of mechanical, electrical, and plumbing. Citeseer, 2007
82. Niknam M, Karshenas S (2017) A shared ontology approach to semantic representation of BIM data. *Autom Constr* 80:22–36
83. Curry E, O'Donnell J, Corry E et al (2013) Linking building data in the cloud: integrating cross-domain building data using linked data. *Adv Eng Inform* 27(2):206–219
84. Ho S-P, Tserng H-P, Jan S-H (2013) Enhancing knowledge sharing management using BIM technology in construction. *Sci World J* 2013, Article ID 170498
85. Previtali M, Brumana R, Stanga C et al (2020) An ontology-based representation of vaulted system for HBIM. *Appl Sci* 10(4):1377
86. Zhang J, Liu Q, Hu Z et al (2017) A multi-server information-sharing environment for cross-party collaboration on a private cloud. *Autom Constr* 81:180–195
87. Venugopal M, Eastman CM, Sacks R et al (2012) Semantics of model views for information exchanges using the industry foundation class schema. *Adv Eng Inform* 26(2):411–428

88. Kiviniemi A, Fischer M, Bazjanac V (2005) Integration of multiple product models: Ifc model servers as a potential solution. In: Proceedings of the 22nd CIB-W78 conference on information technology in construction
89. Jørgensen K, Skaug J, Christiansson P et al (2008) Use of IFC model servers: modelling collaboration possibilities in practice. Aalborg Universitet
90. Beetz J, Van Berlo L, De Laat R et al BIMserver. org—An open source IFC model server. In: Proceedings of the CIP W78 conference
91. Beach TH, Rana OF, Rezgui Y et al (2013) Cloud computing for the architecture, engineering & construction sector: requirements, prototype & experience. *J Cloud Comput Adv Syst Appl* 2(1):8
92. Zhang Y (2009) BIM-based construction information integration and management. Tsinghua University
93. Yu F (2014) Research on a modeling and application technology of BIM oriented to building lifecycle. Tsinghua University
94. El-Diraby TE (2012) Domain ontology for construction knowledge. *J Constr Eng Manag* 139(7):768–784
95. Radulovic F, Poveda-Villalón M, Vila-Suero D et al (2015) Guidelines for Linked Data generation and publication: an example in building energy consumption. *Autom Constr* 57:178–187
96. Zhu J, Wright G, Wang J et al (2018) A critical review of the integration of geographic information system and building information modelling at the data level. *ISPRS Int J Geo Inf* 7(2):66
97. Open Geospatial Consortium (2015) Geographic information-Well-known text representation of coordinate reference systems. OGC document, 2015
98. Open Geospatial Consortium (2010) GeoSPARQL-A geographic query language for RDF data. November, 2010
99. Jusuf SK, Mousseau B, Godfroid G et al (2017) Integrated modeling of CityGML and IFC for city/neighborhood development for urban microclimates analysis. *Energy Procedia* 122:145–150
100. Hor AH, Jadidi A, Sohn G (2016) BIM-GIS integrated geospatial information model using semantic web and RDF graphs. *ISPRS Ann Photogramm Remote Sens Spatial Inf Sci* 3(4):73–79
101. Mignard C, Gesquiere G, Nicolle C SIGA3D: a semantic bim extension to represent urban environment. In: Proceedings of the 5th international conference on advances semantic processing, Lisbon, Portugal
102. Beetz J, Borrmann A (2018) Benefits and limitations of linked data approaches for road modeling and data exchange. In: Workshop of the European group for intelligent computing in engineering. Springer, pp 245–261
103. Zhao L, Liu Z, Mbachu J (2019) Highway alignment optimization: an integrated BIM and GIS approach. *ISPRS Int J Geo Inf* 8(4):172
104. Ali M, Mohamed Y (2018) A framework for visualizing heterogeneous construction data using semantic web standards. *Adv Civ Eng* 2018, Article ID 8370931
105. Korman TM, Fischer MA, Tatum CB (2003) Knowledge and reasoning for MEP coordination. *J Constr Eng Manag* 129(6):627–634
106. Hu Z, Tian P, Li S et al (2018) BIM-based integrated delivery technologies for intelligent MEP management in the operation and maintenance phase. *Adv Eng Softw* 115:1–16
107. Jiang S, Wang N, Wu J (2018) Combining BIM and ontology to facilitate intelligent green building evaluation. *J Comput Civ Eng* 32(5):4018039
108. Wu J (2017) Research on intelligent assistant green building evaluation based on ontology and BIM. Dalian University of Technology
109. Dibley M, Li H, Rezgui Y et al (2012) An ontology framework for intelligent sensor-based building monitoring. *Autom Constr* 28:1–14
110. Zhang Y-Y, Kang K, Lin J-R et al (2020) Building information modelling-based cyber-physical platform for building performance monitoring. *Int J Distrib Sens Netw* 16(2):1550147720908170
111. Marroquin R, Dubois J, Nicolle C (2018) Ontology for a Panoptes building: exploiting contextual information and a smart camera network. *Semantic Web (Preprint)*, pp 1–26
112. Bilal M, Oyedele LO, Munir K et al (2017) The application of web of data technologies in building materials information modelling for construction waste analytics. *Sustain Mater Technol* 11:28–37
113. Zhou P, El-Gohary N (2017) Ontology-based automated information extraction from building energy conservation codes. *Autom Constr* 74:103–117
114. Xiao Y-Q, Hu Z-Z, Lin J-R (2019) Ontology-based semantic retrieval method of energy consumption management. In: *Advances in informatics and computing in civil and construction engineering*. Springer, pp 231–238
115. Pauwels P, Corry E, O'Donnell J (2014) Representing SimModel in the web ontology language. In: *Computing in civil and building engineering*
116. Pauwels P, Corry E, O'Donnell J (2014) Making SimModel information available as RDF graphs. In: *eWork and eBusiness in architecture, engineering and construction: ECPPM*, pp 439–445
117. Kang TW, Hong CH (2018) IFC-CityGML LOD mapping automation using multiprocessing-based screen-buffer scanning including mapping rule. *KSCE J Civ Eng* 22(2):373–383
118. Hu Z, Zhang J, Yu F et al (2016) Construction and facility management of large MEP projects using a multi-Scale building information model. *Adv Eng Softw* 100:215–230
119. Preidel C, Daum S, Borrmann A (2017) Data retrieval from building information models based on visual programming. *Vis Eng* 5(1):18
120. Liebich T (2001) XML schema language binding of EXPRESS for ifcXML. International alliance for interoperability
121. W3C. World wide web consortium[EB/OL](2019)[2019–11–06]. <https://www.w3.org/>
122. Zhang C, Beetz J, De Vries B (2018) BimSPARQL: domain-specific functional SPARQL extensions for querying RDF building data. *Semantic Web (Preprint)*, pp 1–27
123. W3C. Sparql for rdf[EB/OL](2008)[2019–11–06]. <https://www.w3.org/TR/rdf-sparql-query/>
124. Horrocks I, Patel-Schneider PF, Boley H et al SWRL: a semantic web rule language combining OWL and RuleML[EB/OL](2004) [2010–07–12]. <http://www.w3.org/Submission/SWRL/>
125. Vilgertshofer S, Amann J, Willenborg B et al (2017) Linking BIM and GIS models in infrastructure by example of IFC and CityGML. In: *Computing in civil engineering*
126. Simeone D, Cursi S, Acierno M (2019) BIM semantic-enrichment for built heritage representation. *Autom Constr* 97:122–137
127. Liu X, Akinci B, Bergés M et al (2013) Domain-specific querying formalisms for retrieving information about HVAC systems. *J Comput Civ Eng* 28(1):40–49
128. Tao J, Sirin E, Bao J et al (2010) Extending OWL with integrity constraints. In: *Proceedings of the 23rd international workshop on description logics (DL-10)*
129. Terkaj W, Šojić A (2015) Ontology-based representation of IFC EXPRESS rules: an enhancement of the ifcOWL ontology. *Autom Constr* 57:188–201
130. Jiang L, Leicht RM (2014) Automated rule-based constructability checking: case study of formwork. *J Manag Eng* 31(1):A4014004

131. Zhang S, Sulankivi K, Kiviniemi M et al (2015) BIM-based fall hazard identification and prevention in construction safety planning. *Saf Sci* 72:31–45
132. Mikulakova E, König M, Tauscher E et al (2010) Knowledge-based schedule generation and evaluation. *Adv Eng Inform* 24(4):389–403
133. GhaffarianHoseini A, Zhang T, Naismith N et al (2019) ND BIM-integrated knowledge-based building management: inspecting post-construction energy efficiency. *Autom Constr* 97:13–28
134. Zhang L, Wu X, Ding L et al (2016) Bim-based risk identification system in tunnel construction. *J Civ Eng Manag* 22(4):529–539
135. Nickel M, Murphy K, Tresp V et al (2015) A review of relational machine learning for knowledge graphs. *Proc IEEE* 104(1):11–33
136. Shen W, Wang J, Han J (2014) Entity linking with a knowledge base: issues, techniques, and solutions. *IEEE Trans Knowl Data Eng* 27(2):443–460
137. Socher R, Chen D, Manning CD et al (2013) Reasoning with neural tensor networks for knowledge base completion. In: *Advances in neural information processing systems*
138. Wang Q, Wang B, Guo L Knowledge base completion using embeddings and rules. In: *Twenty-fourth international joint conference on artificial intelligence*
139. Apache Software Foundation. Apache Jena[EB/OL](2011) [2019–11–06]. <http://jena.apache.org/index.html>
140. Haarslev V, Hidde K, Möller R et al (2012) The RacerPro knowledge representation and reasoning system. *Semantic Web* 3(3):267–277
141. Sirin E, Parsia B, Grau BC et al (2007) Pellet: a practical owl-dl reasoner. *Web Semant Sci Serv Agents World Wide Web* 5(2):51–53
142. Tsarkov D OWL:FaCT++[EB/OL](2007)[2019–11–06]. <http://owl.man.ac.uk/factplusplus/>
143. Friedman-Hill E Jess, the rule engine for the java platform[EB] (2008)
144. Muñoz-La Rivera F, Mora-Serrano J, Valero I et al (2021) Methodological-technological framework for construction 4.0. *Arch Comput Methods Eng* 28(2):689–711
145. MacCallum KJ (1990) Does intelligent CAD exist? *Artif Intell Eng* 5(2):55–64
146. Maher ML, Brown DC, Duffy A (1994) Machine learning in design. *AI EDAM* 8(2):81–82
147. Chi H-L, Wang X, Jiao Y (2015) BIM-enabled structural design: impacts and future developments in structural modelling, analysis and optimisation processes. *Arch Comput Methods Eng* 22(1):135–151
148. Inyim P, Rivera J, Zhu Y (2014) Integration of building information modeling and economic and environmental impact analysis to support sustainable building design. *J Manag Eng* 31(1):A4014002
149. Nour M, Hosny O, Elhakeem A (2015) A BIM based approach for configuring buildings' outer envelope energy saving elements. *J Inf Technol Constr (ITcon)* 20(13):173–192
150. Liu S, Meng X, Tam C (2015) Building information modeling based building design optimization for sustainability. *Energy Build* 105:139–153
151. Wang L, Leite F (2013) Knowledge discovery of spatial conflict resolution philosophies in BIM-enabled MEP design coordination using data mining techniques: a proof-of-concept. In: *Computing in civil engineering*
152. Pärn EA, Edwards DJ, Sing MCP (2018) Origins and probabilities of MEP and structural design clashes within a federated BIM model. *Autom Constr* 85:209–219
153. Yarmohammadi S, Pourabolghasem R, Castro-Lacouture D (2017) Mining implicit 3D modeling patterns from unstructured temporal BIM log text data. *Autom Constr* 81:17–24
154. Chen L, Pan W (2015) A BIM-integrated fuzzy multi-criteria decision making model for selecting low-carbon building measures. *Procedia Eng* 118:606–613
155. Costa A, Keane MM, Torrens JJ et al (2013) Building operation and energy performance: monitoring, analysis and optimisation toolkit. *Appl Energy* 101:310–316
156. Sidani A, Dinis FM, Sanhudo L et al (2021) Recent tools and techniques of BIM-Based virtual reality: a systematic review. *Arch Comput Methods Eng* 28(2):449–462
157. Merrell P, Schkufza E, Koltun V (2010) Computer-generated residential building layouts. In: *ACM transactions on graphics (TOG)*. ACM, p 181
158. Fisher M, Ritchie D, Savva M et al (2012) Example-based synthesis of 3D object arrangements. *ACM Trans Graph (TOG)* 31(6):135
159. Eastman C, Lee J, Jeong Y et al (2009) Automatic rule-based checking of building designs. *Autom Constr* 18(8):1011–1033
160. Borrmann A, Hyvärinen J, Rank E (2009) Spatial constraints in collaborative design processes. In: *Proceedings of the international conference on intelligent computing in engineering (ICE'09)*. Berlin, Germany
161. Yang QZ, Xu X (2004) Design knowledge modeling and software implementation for building code compliance checking. *Build Environ* 39(6):689–698
162. Bouzidi KR, Fies B, Faron-Zucker C et al (2012) Semantic web approach to ease regulation compliance checking in construction industry. *Future Internet* 4(3):830–851
163. Dimyadi J, Pauwels P, Spearpoint M et al Querying a regulatory model for compliant building design audit. In: *32rd international CIB W78 conference*
164. Solihin W (2004) Achieving automated code checking with ease, a white paper, novaCITYNETS pte ltd. Singapore, 2004
165. Fiatch (2013) An overview of existing BIM standards and guidelines: a report to Fiatch AutoCodes Project. The University of Texas
166. Han CS, Kunz JC, Law KH (2002) Compliance analysis for disabled access. In: *Advances in digital government*. Springer, pp 149–162
167. Lee P-C, Lo T-P, Tian M-Y et al (2019) An efficient design support system based on automatic rule checking and case-based reasoning. *KSCE J Civ Eng* 23(5):1952–1962
168. Kadolsky M, Baumgärtel K, Scherer RJ (2014) An ontology framework for rule-based inspection of eeBIM-systems. *Procedia Eng* 85:293–301
169. Hu P (2016) BIM and ontology based automatic fire safety checking in building design. Tianjin University
170. Baumgärtel K, Kadolsky M, Scherer RJ (2014) An ontology framework for improving building energy performance by utilizing energy saving regulations. In: *Proceedings of European conference on product and process modelling*
171. De Soto BG, Adey BT (2016) Preliminary resource-based estimates combining artificial intelligence approaches and traditional techniques. *Procedia Eng* 164:261–268
172. Tixier AJ-P, Hallowell MR, Rajagopalan B et al (2017) Construction safety clash detection: identifying safety incompatibilities among fundamental attributes using data mining. *Autom Constr* 74:39–54
173. Marzouk M, Abubakr A (2016) Decision support for tower crane selection with building information models and genetic algorithms. *Autom Constr* 61:1–15
174. Wang J, Zhang X, Shou W et al (2015) A BIM-based approach for automated tower crane layout planning. *Autom Constr* 59:168–178
175. Sigalov K, König M (2017) Recognition of process patterns for BIM-based construction schedules. *Adv Eng Inform* 33:456–472

176. Faghihi V, Reinschmidt KF, Kang JH (2014) Construction scheduling using genetic algorithm based on building information model. *Expert Syst Appl* 41(16):7565–7578
177. Chen Y-J, Feng C-W, Wang Y-R et al (2011) Using BIM model and genetic algorithms to optimize the crew assignment for construction project planning. *Int J Technol* 3:179–187
178. Liu H, Al-Hussein M, Lu M (2015) BIM-based integrated approach for detailed construction scheduling under resource constraints. *Autom Constr* 53:29–43
179. De Soto BG, Rosarius A, Rieger J et al (2017) Using a tabu-search algorithm and 4D models to improve construction project schedules. *Procedia Eng* 196:698–705
180. Moon H, Kim H, Kim C et al (2014) Development of a schedule-workspace interference management system simultaneously considering the overlap level of parallel schedules and workspaces. *Autom Constr* 39:93–105
181. Hu X, Lu M, AbouRizk S (2014) BIM-based data mining approach to estimating job man-hour requirements in structural steel fabrication. In: *Proceedings of the 2014 Winter simulation conference*. IEEE Press, pp 3399–3410
182. Ganbat T, Chong H-Y, Liao P-C et al (2019) A cross-systematic review of addressing risks in building information modelling-enabled international construction projects. *Arch Comput Methods Eng* 26(4):899–931
183. Vigneault M-A, Botton C, Chong H-Y et al (2019) An innovative framework of 5D BIM solutions for construction cost management: a systematic review. *Arch Comput Methods Eng* 27:1–18
184. Song S, Marks E (2019) Construction site path planning optimization through BIM. In: *Computing in civil engineering 2019: visualization, information modeling, and simulation*. American Society of Civil Engineers Reston, VA, pp 369–376
185. Zhang S, Boukamp F, Teizer J (2015) Ontology-based semantic modeling of construction safety knowledge: towards automated safety planning for job hazard analysis (JHA). *Autom Constr* 52:29–41
186. Li N, Becerik-Gerber B, Krishnamachari B et al (2014) A BIM centered indoor localization algorithm to support building fire emergency response operations. *Autom Constr* 42:78–89
187. Tang L C M, Cho S Y, Xia L (2013) Intelligent HVAC information capturing system for smart building information modelling. In: *2013 5th International conference on power electronics systems and applications (PESA)*. IEEE, pp 1–4
188. Alshibani A, Alshamrani OS (2017) ANN/BIM-based model for predicting the energy cost of residential buildings in Saudi Arabia. *J Taibah Univ Sci* 11(6):1317–1329
189. McGlinn K, Yuce B, Wicaksono H et al (2017) Usability evaluation of a web-based tool for supporting holistic building energy management. *Autom Constr* 84:154–165
190. Corry E, O'Donnell J, Curry E et al (2014) Using semantic web technologies to access soft AEC data. *Adv Eng Inform* 28(4):370–380
191. Motawa I (2017) Spoken dialogue BIM systems—an application of big data in construction. *Facilities* 34:787–800
192. Martinez-Gil J, Chasparis G, Freudenthaler B et al (2014) Realistic user behavior modeling for energy saving in residential buildings. In: *2014 25th International workshop on database and expert systems applications*. IEEE, pp 121–125
193. Tomic S, Fensel A, Schwanzer M et al (2012) Semantics for energy efficiency in smart home environments. In: Sugumaran, V., Gulla, J.A. (eds.) *Applied Semantic Technologies: Using Semantics in Intelligent Information Processing*, Taylor and Francis, 429–454
194. Tang Y, Ciuciu IG (2012) Semantic decision support models for energy efficiency in smart-metered homes. In: *2012 IEEE 11th international conference on trust, security and privacy in computing and communications*. IEEE, pp 1777–1784
195. Katipamula S, Brambley MR (2005) Methods for fault detection, diagnostics, and prognostics for building systems—a review, part I. *Hvac&R Res* 11(1):3–25
196. Liang J, Du R (2007) Model-based fault detection and diagnosis of HVAC systems using support vector machine method. *Int J Refrig* 30(6):1104–1114
197. Liu X, Akinci B, Garrett James HJ et al ((2011)) Requirements for an integrated framework of self-managing HVAC systems. In: *Computing in civil engineering*
198. Liu X, Akinci B, Bergés M et al (2013) Extending the information delivery manual approach to identify information requirements for performance analysis of HVAC systems. *Adv Eng Inform* 27(4):496–505
199. Yang X, Ergon S (2015) Leveraging BIM to provide automated support for efficient troubleshooting of HVAC-related problems. *J Comput Civ Eng* 30(2):4015023
200. Motamedi A, Hammad A, Asen Y (2014) Knowledge-assisted BIM-based visual analytics for failure root cause detection in facilities management. *Autom Constr* 43:73–83
201. Dong B, O'Neill Z, Li Z (2014) A BIM-enabled information infrastructure for building energy Fault Detection and Diagnostics. *Autom Constr* 44:197–211
202. Golabchi A, Akula M, Kamat V (2016) Automated building information modeling for fault detection and diagnostics in commercial HVAC systems. *Facilities* 34(3/4):233–246
203. Kazmi AH, Ogrady MJ, Delaney DT et al (2014) A review of wireless-sensor-network-enabled building energy management systems. *ACM Trans Sens Netw (TOSN)* 10(4):66
204. De Paola A, Ortolani M, Lo Re G et al (2014) Intelligent management systems for energy efficiency in buildings: a survey. *ACM Comput Surv (CSUR)* 47(1):13
205. McGlinn K, Jones K, Cochero A et al Usability evaluation of an activity modelling tool for improving accuracy of predictive energy simulations. In: *Proceedings of the 14th conference of international building performance simulation association (IBSA)*
206. Ramaji JJ, Messner JJ, Mostavi E (2020) IFC-based BIM-to-BEM model transformation. *J Comput Civ Eng* 34(3):4020005

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.