Developments in the Built Environment 4 (2020) 100011

Contents lists available at ScienceDirect

Developments in the Built Environment

journal homepage: www.editorialmanager.com/dibe/default.aspx

Building Information Modelling, Artificial Intelligence and Construction Tech

Rafael Sacks, PhD [a,b,*], Mark Girolami, PhD [b], Ioannis Brilakis, PhD [b]

[a] Seskin Virtual Construction Lab, Faculty of Civil and Environmental Engineering, Technion - Israel Institute of Technology, Haifa 32000, Israel
[b] Division of Civil Engineering, Department of Engineering, University of Cambridge, 7a JJ Thomson Avenue, Cambridge CB3 0FA, UK

A R T I C L E I N F O

Keywords:
Artificial intelligence
Building information modelling
Construction technology
Digital construction
Digital twins
Innovation

A B S T R A C T

Adoption of digital information tools in the construction sector provides fertile ground for the birth and growth of companies that specialize in applications of technologies to design and construction. While some of the technologies are new, many implement ideas proposed in construction research decades ago that were impractical without a sound digital building information foundation. Building Information Modelling (BIM) itself can be traced to a landmark paper from 1975; ideas for artificially intelligent design and code checking tools date from the mid-1980s; and construction robots have laboured in research labs for decades. Yet only within the past five years has venture capital actively sought startup companies in the 'Construction Tech' sector. Following a set of digital construction innovations through their known past and their uncertain present, we review their increasingly optimistic future, all through the lens of their dependence on digital information. The review identifies new challenges, yielding a set of research topics with the potential to unlock a range of future applications that apply artificial intelligence.

1. Introduction

Researchers in architecture, engineering and construction have long dreamed of applying information technology, robotics and other new technologies to design and construction. Yet invariably, their conceptual understandings of what could be done, and hence their visionary views of the future of construction, far outstripped the practical, technical, commercial, cultural and/or organizational constraints that had to be overcome for their fulfilment.

Eastman, for example, conceived of a computerised Building Design System (BDS) with all the functionality of what we now know as Building Information Modelling (BIM) (Eastman, 1975). The basic BIM functions took 25 years to reach the market, and some – such as Eastman's prediction that "*Later, one can conceive of a BDS supporting automated building code checking in city hall or the architect's office*" – have yet to be realized in full. Indeed, Gholizadeh et al. (2018) found that, as late as 2017, of the 14 BIM functions whose adoption they investigated, only three were in widespread use. Similarly, Warszawski and Sangrey (1985) wrote that "*Implementing robotics in construction may follow several paths. One approach will involve an evolution of robotic and computer technology into existing procedures. The second approach will be more dramatic with the combination of robotics and CAD-CAM providing the basis for entirely new building systems—the construction of the future.*" Construction robotic machines of the first type are only now beginning to become practical and economical, and none have achieved the revolutionary change they contemplated in their second mode. For many researchers with foresight and a good conceptual grasp of potential implementations, automation in construction has at times proved to be a frustratingly difficult goal from the point of view of implementation in industry.

Within the last five years, however, there has been a steady influx of new, innovative companies specializing in application of a variety of information and automation technologies, developed in other industries, to construction. These startup companies are supported by venture capitalists, academic research and public and private incubator programs, together with which they form an ecosystem commonly called 'Construction Tech' (echoing the name 'High Tech' used for the information and automation technology industry). In the US, the amount of venture capital invested in Construction Tech annually is reported to have grown from circa $250 m in 2013 to well over $1,000 m in 2018 (Andersen and Forr, 2018). Most of the new companies owe their newfound practicality not only to the maturation of their core technologies, but equally to the comprehensive building information available in BIM environments. In

* Corresponding author. Seskin Virtual Construction Lab, Faculty of Civil and Environmental Engineering, Technion - Israel Institute of Technology, Haifa 32000, Israel.

E-mail addresses: cvsacks@technion.ac.il (R. Sacks), mag92@hermes.cam.ac.uk (M. Girolami), ib340@cam.ac.uk (I. Brilakis).

https://doi.org/10.1016/j.dibe.2020.100011
Received 15 March 2020; Received in revised form 4 May 2020; Accepted 11 May 2020
Available online 16 May 2020

this context, a *BIM environment* encompasses the various BIM tools, platforms, servers, libraries and workflows within a project or across an organization (for definitions of the terms, see Sacks et al., 2018, Chapter 2.3). In this paper, we use the term BIM in its broadest sense, to include all software applications that generate, manipulate, store and deliver building information, rather than restricting it to the narrow set of model authoring platforms.

Building information incorporates both the building product designs and the construction process plans. BIM environments provide this information, and it is the foundation upon which the new Construction Tech applications build and deliver value. Broadly speaking, there are four types of Construction Tech applications:

1. Software tools for design and construction management, most of which function within or in close relationship with model authoring platforms.
2. Software and hardware systems for delivering information from the design to the field – we call this group ***BIM-to-field*** tools. This group includes tools that deliver information via users' mobile devices as well as tools that deliver product and process information directly onto workspaces in the field (such as the site laser projection solution shown in Fig. 1).
3. Robotic applications for executing construction operations on site.
4. Software and hardware systems for gathering information from the site and delivering it to controlling functions – we call this group ***field-to-BIM*** tools. These Construction Tech solutions provide copious amounts of data, but that data is of no value in and of itself – one must compare the planned to the actual conditions to derive valuable information, as shown in Fig. 2.

Naturally, some applications will have more than one use – an app that delivers process information to a user's mobile device may also solicit information to update current status, for example.

BIM is an integrating technology that provides an information backbone that transcends organizational boundaries within projects (Sacks et al., 2018). As such, BIM environments support Construction Tech innovations by mediating the gap between the information intensive technology innovations and the traditionally information poor and fragmented construction project organisations. However, BIM platforms,

Fig. 1. Projection of a partition layout directly onto a concrete slab. Image courtesy of Lightyx Inc.

tools and processes are not yet able to support most technology innovations “out of the box”. Our goal in this paper, therefore, is to explore and elaborate the key aspects of the information and data processing barriers to Construction Tech, with special focus on the aspects of BIM environments and Artificial Intelligence (AI) that require basic research to underpin commercial development.

In the following section, we review research and development trends in the field of construction over the past half-century, with examples of R&D for automated code-compliance checking, construction layout, construction robotics and automated project performance monitoring and control. We then consider the last five years, leading up to the present – the growth of Construction Tech based on BIM, field monitoring, robotics, AI and other technologies. In this context, we propose a conceptual ‘House of Construction Tech’ framework that explains how the different theoretical, technical, commercial and conceptual foundations underpin the growth of innovative Construction Tech startup companies. Finally, we present and discuss four foundational challenges for the research community.

2. Known past: research and development trends

Following the research and development (R&D) history of three broad areas of technological innovation in construction, we trace their paths to the present day. Our goal is not to extrapolate into the future, but to identify key research challenges for continued development – to identify the essential R in R&D. The areas selected for review represent three of the four applications types listed in the introduction: (1) software tools for design and planning within BIM environments; (2) BIM-to-field tools; and (4) field-to-BIM tools, which are beginning to enable digital twins for construction.

2.1. Automated design and code-compliance checking

Automating design and code-compliance checking for building construction has been a goal of research and commercial development since the ideas were floated in Eastman's landmark 1975 paper envisaging BIM (Eastman, 1975). In the absence of BIM, researchers proposed stand-alone expert systems (Hayes-Roth et al., 1983), and later, systems that used CAD drawings to represent the buildings. The former, such as HI-RISE for preliminary structural design of tall buildings (Maher and Fenves, 1985), SPEX for sizing structural cross-sections (Garrett and Fenves, 1987) and EIDOCC for design of reinforced concrete columns (Sacks and Buyukozturk, 1987), used symbolic AI methods. These were typically rule-based systems that sought to elicit expert knowledge, capture it in design software, and apply it to automate or to review design. The need to input building designs explicitly and completely for each analysis, the limitations on knowledge elicitation, and the capacity of the computing technology made these systems impractical for commercial application. The advent of 2D CAD did not improve matters much, because CAD's graphic representations of design are fundamentally different to the semantic, object-oriented representations required for processing rules.

Initial optimism that design standards themselves could be expressed as rules and applied to evaluate building designs (e.g. Hakim and Garrett, 1993) proved unfounded, as experiments revealed the challenges posed by the lexical and logical complexity of building code provisions (Kiliccote and Garrett, 1995). Later, natural language processing (NLP) was applied to building design codes and regulations, resulting in some progress, but not in commercial application (Song et al., 2018; Zhang and El-Gohary, 2017).

With the introduction and adoption of BIM, automated design and code-checking became more practical. Commercial model checking systems with limited but valuable and viable functionality were developed (examples include Solibri Model Checker, BIM Assure and SMARTreview). While they are able to use BIM models, they require users to *normalize* model data before use, and the repertoire of code clauses they

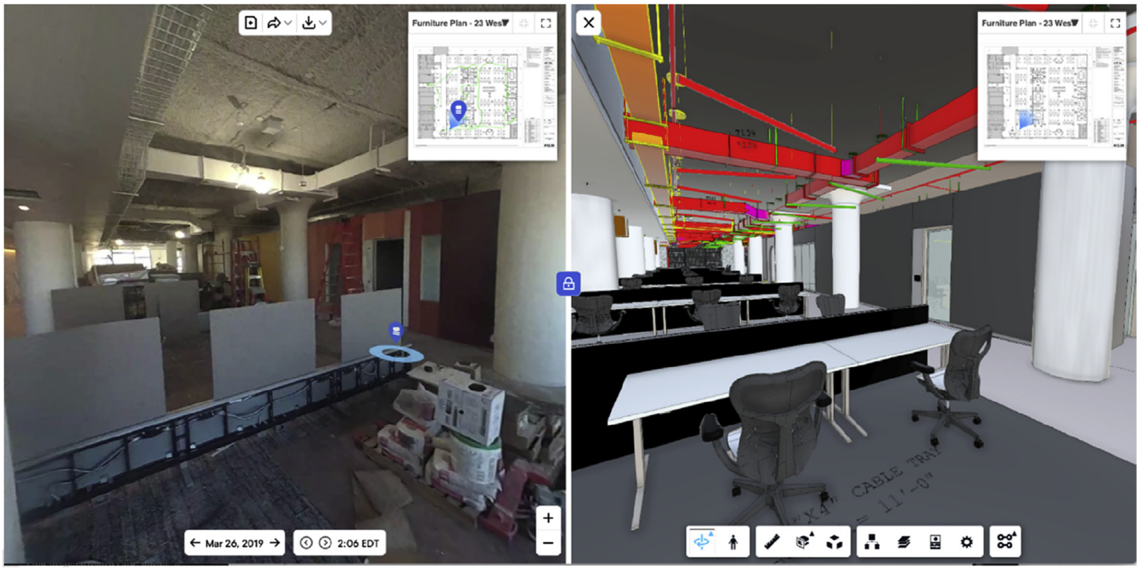

Fig. 2. Image captured with 360° cameras at left and a view of the same scene in a BIM model at right. Image courtesy of Openspace and Lee Kennedy Construction.

can check is limited to clauses that can be expressed as symbolic IF-THEN rule sets (normalization is the task of pre-processing a BIM model for symbolic code-checking. Users manually add or edit objects, parameter values and relationships to conform to the naming conventions and object typing required by the rule sets).

2.2. Construction set out

Setting out construction work on site is laborious and error prone. The challenge is to interpret design information within the context of partially completed scenes and apply physical markings to surfaces with the required precision. The state-of-the-art method is robotic total station survey layout, in which an operator localizes the total station using known points in the scene and then ‘shoots’ a laser beam to locate layout points. This is followed by manual mapping of more complete design information from the points onto other surfaces using chalk-lines, laser plane projectors, and other tools.

The scale of direct effort and subsequent rework in case of error have prompted R&D efforts to build automated layout systems that deliver BIM information directly to the field. Three types have been proposed:

a) Augmented reality (AR) systems, in which an image of the intended design is superimposed onto an image of the site recorded with a camera (Chi et al., 2013; Wang et al., 2013; Woodward et al., 2014). AR systems require special glasses or masks, or a tablet computer or other device on which the images are projected. Users must then translate that information onto the work surface. These systems are particularly useful for locating hidden system objects behind finished surfaces for building or facility operation and maintenance tasks (Lee and Akin, 2011; Sacks et al., 2018).
b) Robotic marking systems, in which a robot localizes itself and then travels the work area applying paint or other marking material directly onto the surface (Casale, 2013; Prouty, 2013). These are generally restricted to environments where the floors are clean and clear, for marking and for travel, and where the quantity of layout work is large enough to justify their setup costs.
c) Robotic systems that project BIM information directly onto the work surface. For example, Degani et al. (2019) developed a prototype in which images from a BIM model are projected directly onto a work surface. The apparatus consisted of a laser range scanner, an angled adjustable projector, and a camera. The system localizes itself using the LIDAR and the BIM model using Simultaneous Localization and Mapping (SLAM) and projects images containing the information onto the work surface. It calibrates the projection keystone correction parameters using image analysis.

With the growing capabilities of laser scanning and imaging technologies, improved accuracy of localization, and sophisticated projection tools, this area appears to offer opportunities for rapid development of new commercially viable tools. The Lightyx system depicted in Fig. 1 is a good example of a startup development path for Construction Tech in which innovators with expertise from other industries apply their knowledge of advanced technologies to solve construction problems. It is also an example of an innovation that fits entirely within current construction practice, automating an isolated operation.

2.3. Automated project performance monitoring and control

The concept of automated project performance control (APPC) was proposed as a way to provide managers with the real-time feedback necessary for application of the ‘thermostat’ model of control (Navon, 2005). The idea was to close the control loop by reporting leading performance indicators, such as labour and equipment productivity by monitoring the movements of workers and materials in real-time (Navon and Sacks, 2007; Sacks et al., 2006). This line of research might have presaged the new concept of ‘Construction 4.0’ or of digital twins for construction, but it encountered technical and conceptual barriers:

- From a technical standpoint, there was no platform available to integrate the necessary production process information for comparison with monitored data. Researchers developed sophisticated methods to extract information from construction site documents and images (Al Qady and Kandil, 2010; Brilakis et al., 2005), but these were not linked to any integrated data management system.
- From a conceptual standpoint, the thermostat model proved to be inappropriate and ineffective for planning and controlling production in construction, and it has been replaced over time with methods based on pull production planning and control (Ballard, 2008; Kenley and Seppänen, 2010).

The notion of automating monitoring work on site originated from the observation that engineers in the field spent a lot of time collecting performance data (McCullouch, 1992; McCullouch and Lueprasert, 1994). A variety of technologies have been proposed for data collection, including computer vision (Brilakis and Haas, 2020), GPS, laser scanning, radio-frequency ID tags and Bluetooth low energy (Bekkelien et al., 2012; Costin et al., 2012). Yet except for systems for monitoring heavy earthworks machinery, field-to-BIM automation has not been adopted in the construction market.

3. Uncertain present: BIM, AI and Construction Tech

A common thread runs through all three examples of Construction Tech innovations detailed in the previous section: dependence on the availability of digital building information. With the benefit of hindsight, we observe that it is not simply a matter of digitising information, but one of making information freely available across BIM platforms, with both syntactic and semantic interoperability. The current lack of direct access inhibits the application of AI. In this section we review the current status of BIM and AI in the construction industry (circa 2020), with reference to the same three Construction Tech application areas.

3.1. Automated design and code-compliance checking

Although the advent of BIM has made commercial code-checking applications viable, their core technology has not changed fundamentally from that envisaged in the 1980's. They all use symbolic AI methods, primarily rule-inferencing, which restricts their scope to relatively simple prescriptive clauses (Bloch et al., 2019). The challenge posed by the large numbers of applicable design and building codes, and the frequency with which they are updated, has not been solved (Nawari, 2017). The commercial applications still require explicit representation of the building information (Dimyadi et al., 2016), and the effort required for normalization limits their use to isolated milestone points in design processes.

Breakthrough progress in code-checking will require overcoming these barriers, and new approaches and technologies will be needed. Among the most promising:

- Semantic enrichment of BIM models, using AI methods to automatically supplement models with explicit information derived using algorithms trained to recognize and infer predefined sets of target concepts within patterns of building data, may offer a way to remove the need for normalization (Belsky et al., 2016).
- Application of machine learning algorithms to evaluate designs on the basis of the training data of known results from human experts (Sacks et al., 2019).
- Graph representations of BIM models may offer the explicit representations needed, in particular for making the relationships between building objects and abstract concepts explicit (Nahar, 2017). They are also more amenable to the types of pattern recognition algorithms that may enable semantic enrichment (Jin et al., 2018) and training of ML algorithms.

We note also that all of the companies offering commercial BIM code-checking applications are startup companies (Compliance Audit Systems Limited, Daima, Invicara, SMARTreview) or began as startup companies (Solibri, recently acquired by Nemetschek).

3.2. Construction set out

BIM-to-field information delivery has largely been solved with regard to delivering BIM information to personnel via mobile computing devices. All the major BIM platform companies offer solutions, most of which originated with disruptive Construction Tech startup companies whose solutions were acquired by the established BIM companies (e.g. PlanGrid, Trimble Connect, Solibri). Tools that present model information using augmented reality are also available (such as Trimble's XR10 with Hololens 2; Trimble, 2020), although these still suffer from practical problems such as narrow fields of vision, indistinct display in bright environments, encumbering workers, etc.

Despite ten years or more of industrial R&D, there are not yet commercially viable solutions for setting out directly onto work surfaces. One of the key challenges is to project or mark information on irregular, intermediate as-built work surfaces in the real world, because they do not correspond directly to the ideal as-designed surfaces of finished products that exist in the virtual world of BIM models. Ironically, this problem might be overcome if the 'field-to-BIM' technologies were able to build accurate, virtual digital twin representations of site conditions. Yet this too remains a challenge, as we describe next.

3.3. Automated project performance monitoring and control

This area is rife with solutions offered by both established software and hardware vendors and startup companies. Applications range from (i) inspection systems, allowing inspectors access to data before, during and after the inspection process and access to recording functionality to collect site data, to (ii) control systems, that enable the control of safety, site traffic, resource and storage utilisation, and others, to (iii) planning and measuring systems, for site logistics and layout planning, production monitoring, and others. Some investment has gone into this space; yet all applications are single track, functioning as information islands. They use one or few data acquisition technologies and interpret that as best they can into useful construction management information (e.g. Siteaware, Disperse.io, Holobuilder, Smartvid.io, Versatile Natures, Openspace.ai, Genda, and others). The information they provide is not always reliable and needs manual review and intervention, which often invalidates their automation-borne benefits. Their limited approach also limits what conclusions can be drawn.

Essentially, there is a need for complex event processing (Buchmann and Koldehofe, 2009). This would entail the merging of information from multiple monitored data sources with already existing information about the as-is status of a construction site and the production plan, to deduce accurate information about what has been built, how, and what resources were used and where merging/fusing data from multiple sources to compile comprehensive information about project status (in terms of both product and process status). However, complex event processing is only possible on the basis of well-integrated and reliable data, something we still lack. Although data is available in apparent abundance, the current lack of comprehensive, accurate and reliable linked data and information naturally restricts the opportunities to properly exploit the technologies.

4. Optimistic future: The House of Construction Tech

In theory, BIM models of buildings and infrastructure are ideally suited to manipulation by smart software tools that incorporate computer vision, rule-inferencing, machine learning, case-based reasoning and other AI strategies. The range of potential applications is wide, including – but by no means limited to – smart tools for:

- Design support and/or automation, topology optimization, generative design.
- Design review, checking compliance to standards and codes.
- Building performance simulations and engineering analyses.
- Construction planning, site layout design, supply chain management.
- Digital delivery of design and construction method directly to workers on site.
- Real-time measurement, assessment and interpretation of project status.
- Quality assurance and control.
- Production control, resource assignment, material and information flow control.
- Safety planning and control.
- Sustainability and life-cycle costing assessments.
- Acquisition of BIM models from point cloud data, photo- and videogrammetry.

- Facility operations and management using digital twins.

Researchers of computing in the Architecture, Engineering and Construction (AEC) industry across the world have sought to realize such tools since the ideas behind AI developed. In early efforts in the 1980's and 1990's, people attempted to apply expert systems and case-based reasoning to some of the tasks listed above. It soon became apparent that CAD technology was not suited to such applications because its representation of building information was graphic and symbolic, rather than object-oriented. This led to an intense effort to solve the representation challenge, which resulted in the BIM model authoring platforms that are now ubiquitous across the industry, and in an open object-oriented schema for representing buildings and infrastructure (the IFC data model) (ISO, 2013).

With the information representation challenge apparently solved, the stage appeared set for commercial implementation of innovations, and a wave of technological innovation began. As Andersen and Forr (2018) and other reviewers (Azevedo, 2019; Blanco et al., 2017) of these developments have noted, the majority of innovation financing has been provided to startup companies. There are two main reasons for this: corporate/organizational fragmentation within the industry, and the need for expert knowledge and experience with technologies adapted from other industrial domains.

Hall et al. (2019) provide compelling evidence of innovations that develop outside construction project organisations due to fragmentation of the industry: vertical fragmentation (professional and trade specialization), horizontal fragmentation (multiple small firms competing with one another), and longitudinal fragmentation (high turnover of suppliers and clients from project to project). In this environment, systemic innovations tend to disrupt existing commercial or organizational boundaries and therefore require wholly new vertically and longitudinally integrated organisations, with high startup costs and significant risk (Katila et al., 2018). Within this context, it is not surprising that many Construction Tech innovators fail to overcome the regulatory, commercial, cultural, organizational and technological barriers (Chowdhury et al., 2019), despite inventing and developing cutting edge technology applications in the construction domain. Given these risks, almost all the innovators adopt an incremental approach to change in the construction industry, as their top priority is to achieve a minimal viable product and being to generate income.

In addition to the fundamentals of entrepreneurship (ideas, investment and implementation), all Construction Tech innovators require at least three essential things: 1) a real process need in the industry, 2) an application of a new technology that fulfils the need, and 3) a workable business model. These are the pillars of the 'House of Construction Tech', which we propose as a model to explain the components essential for success in the sector (shown in Fig. 3). Entrepreneurship provides the beams that support the roof, which is the pinnacle of success – adoption in the construction industry market. The BIM environment, in its broadest sense as technology, process and people, sits at the base of the house. BIM technology is the hardware and the software that generate and store the information about a construction project, including its physical aspects (a building's design) and its process aspects (construction plans). BIM processes are the information management aspects – standards, such as ISO 19650 (ISO, 2018) and IFC (ISO, 2013); organization and project level BIM execution plans; level of detail (LOD) definitions; etc. The people are those capable of implementing the processes using the technology, including not only employees of the innovator (designers, programmers, etc.), but no less important, employees of the customers (architects, engineers, and construction managers) skilled in working within BIM environments.

Building information in a form that can be readily manipulated by software is essential for almost all Construction Tech innovations, and hence placement of the BIM environment at the base of the house. All

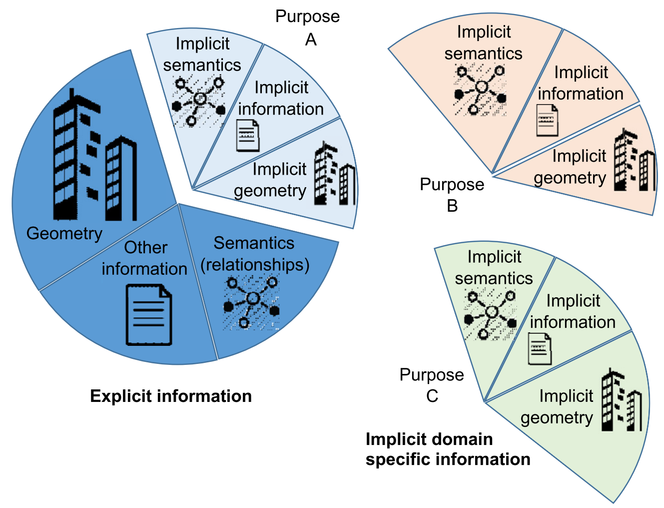

Fig. 3. Conceptual 'House of Construction Tech' model.

four application types (design and management, BIM-to-field, field automation, and field-to-BIM) are dependent on information in one form or another. The maturation of BIM environments and their broad adoption is the one key common denominator supporting the growth of Construction Tech within the last decade. Equally, however, BIM technology and processes still have severe limitations that constrain the long-promised growth and success of some of the Construction Tech applications. Among the key limitations: inadequate interoperability of information, difficulties in framing model data for machine learning applications for design and management, and the need for an intelligent digital twin platform to support integration of field-to-BIM tools.

The House of Construction Tech must be underpinned by a comprehensive understanding of the theoretical aspects of design, of information and data science, and of production in construction. Design theory encompasses design cognition (Winograd and Flores, 1986), philosophy of design (Galle, 2002), and design strategies such as top-down knowledge-based design (Mitchell et al., 1990; Sacks et al., 2003) and design optimization (Gero, 2012). Information and data science include methods for representing building information (Braid, 1973; Sacks et al., 2004; Turk et al., 1994) and methods for artificially intelligent processing of data, including machine learning and pattern recognition (Bishop, 2016; Efron and Hastie, 2016; Rogers and Girolami, 2017). Production theory concerns our understanding of the products, the processes and the operations in the context of construction (Ballard, 2000; Koskela, 2000; Sacks, 2016).

The 'House of Construction Tech' can serve as a 'checklist' for construction startup companies, and as a predictor of success or failure, by considering whether a company has successfully incorporated the columns, base and foundations. At the foundation level, for example, an AI tool for automated construction scheduling using machine learning cannot provide real value for construction managers if its authors restrict their tool to master planning using the Critical Path Method, ignoring the conceptualisation of production in construction as flows of work, products and resources that underpins essential more detailed layers of planning (Koskela, 2000). At the base level, applications that use 2D printed drawings rather than BIM models as their main input will find their scope severely limited. A company for whose innovation these aspects are relevant and yet chooses to ignore them, is unlikely to succeed in the long run. At the level of the columns, innovators must identify business process need to avoid the common trap of solutions looking for problems. For example, proponents of a virtual reality telepresence technology must identify the business process use case that will underlie market demand for their solution before developing the application.

5. Realistic future: foundational research challenges

While the outlook for innovation in Construction Tech is promising, there are still challenges to be met. Some of the technological constraints identified in the sections above, that describe the known past and the uncertain present, reveal that the information representation is still inadequate. Significant foundational challenges remain to be solved before Construction Tech innovators can begin broad incorporation of AI techniques within smart software tools that manipulate BIM models. One of the key problems is that models are commonly discipline-specific, representing buildings with the semantics of a single professional view (such as architecture, structural engineering, or MEP systems). As such, multi-disciplinary collaboration using models is difficult, with most teams using federations of separate models. Another key problem is that many object relationships and properties are still implicit in BIM models, left to the intelligent interpretation of their human users. A third problem is that even where information is complete, the object-oriented representations are not suited to the representations of objects and their features that are the standard input of existing AI techniques. Similarly, design specifications and building codes commonly define parameters that are complex compilations of geometric and other constraints that are very difficult to express using if-then rule sets.

These problems are common to all aspects of tool development for the applications discussed above. Some of the problems can be solved through semantic enrichment, in which professional knowledge is applied to infer the missing information from the purpose-specific viewpoint of the smart tools receiving data from models. In general terms, this would make the information *complete*. Others require development of data engineering strategies to make the information *compatible* for AI processing.

At an implementation level, we identify two distinct research challenges for development of foundational information processing methods for BIM models which, if solved, would greatly facilitate development of a wide range of smart BIM and AI tools for design and construction:

1. Combined, optimal use of topological rule inferencing and machine learning modules for **semantic enrichment**
2. **Representations of BIM models suitable for AI applications**, with emphasis on machine learning

5.1. Semantic enrichment

Semantic enrichment of BIM models is a process in which algorithms apply expert domain knowledge to infer any and all information needed for a given specific application that is absent from the explicit data in the model (Belsky et al., 2016). Fig. 4 illustrates the idea that models are incomplete representations of buildings. The explicit set of geometry, properties and relationships in a model is insufficient input for most receiving applications from sub-domains other than the one in which it was generated. The information inferred can be added to the model, enriching it and facilitating its use for any given receiving application. Semantic enrichment encompasses classification of building objects, aggregation and grouping of objects in functional groups or systems, unique identification of building elements, generation of missing objects, properties and relationships, and reconstruction of occluded objects in the case of application to models compiled from point cloud data (PCD) (Sacks et al., 2017).

Semantic enrichment is by no means a mature technology. It is a relatively new area of research and the literature on the subject is limited. Much work has been done toward intelligent semantic query of BIM models (Borrmann et al., 2006; Mazairac and Beetz, 2013; Wülfing et al., 2014), and these efforts have begun to exploit the meaningful topological constructs of information that is implied in models, but not stored explicitly. They use the implicit meaning, but no information is added to enrich the model during their operations.

Semantic enrichment software modules are goal driven. The missing or misrepresented information that is to be supplemented is defined in a domain specific Model View Definition (MVD). The first experimental applications applied rule inferencing, a subset of AI methods in which the system logic is defined as a set of rules, usually in 'IF-THEN' form. In forward-chaining solutions, rules are processed iteratively, adding new information to a model whenever a rule is evaluated as TRUE. Iteration ends when no new facts can be deduced. Backward-chaining is also possible, where rules are run to verify hypotheses, seeking evidence within the existing information. More recent research has shown that some aspects of semantic enrichment may be better performed using machine learning. For example, a supervised machine learning algorithm was successfully applied to classify room types in residential apartments (Bloch and Sacks, 2018). Unsupervised machine learning has been previously applied to BIM models to detect anomalies and misclassified

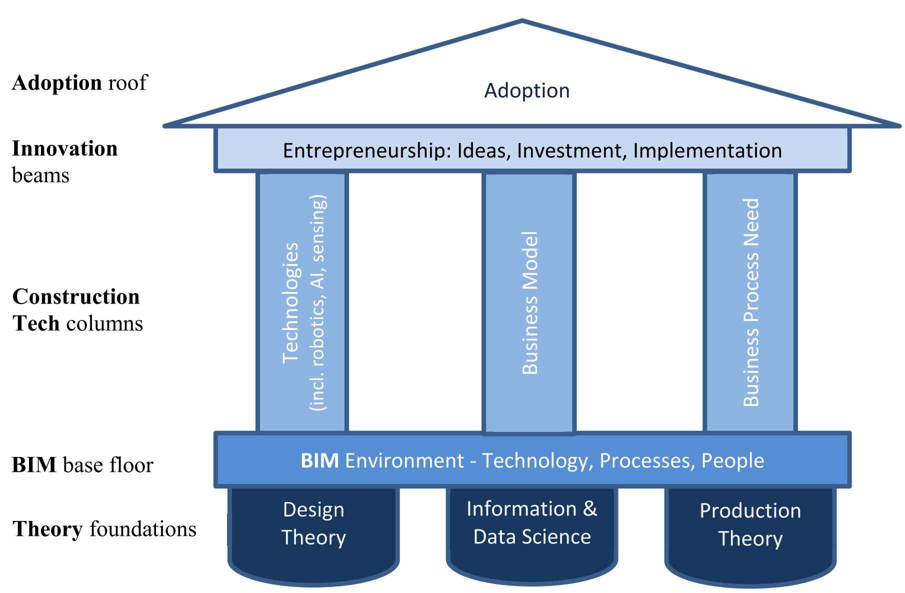

Fig. 4. Explicit and implicit information in BIM models. Each sub-domain (e.g. structural analysis, material take-off, construction detailing) has its own set of implicit concepts.

objects (Krijnen and Tamke, 2015). The discrepancies in an IFC based data exchange have also been tackled by application of machine learning algorithms such as Support Vector Machine (SVM) (Koo et al., 2019; Koo and Shin, 2018) to detect misclassifications of specific objects in IFC files. Similarly, Wu and Zhang (2019) applied learning algorithms to identify common building objects in IFC files.

Basic research is needed to classify BIM information objects according to their suitability for semantic enrichment using different AI methods (primarily topological rule inferencing, supervised machine learning or deep learning) and to design suites of semantic enrichment software methods that could be compiled to build applications in modular fashion. If these goals could be achieved, the basic problem of BIM interoperability might be solved, as standard models could be enriched for almost any purpose. This would provide the technological foundation for Construction Tech startups to implement a wide range of applications (some of which are listed in the Conclusions section below).

5.2. BIM model representations for learning applications

Building models are stored in one of three ways – in proprietary file formats specific to particular BIM authoring platforms, in open IFC format files based on EXPRESS (ISO, 2004), or more recently, in proprietary cloud databases. None of these formats are directly compatible with pattern recognition and/or machine learning algorithms. In all existing applications of AI that use BIM information, whether as complete models, partial assemblies from models, or model components, users must extract and compile the relevant information anew for every use. Extraction generally requires export of tabular schedules of objects and their properties or parsing of IFC files.

A core problem with these methods is that meaningful information is lost in translation. Perhaps most significant is the loss of relationships between objects that are available in the native BIM software representation, but that cannot be exported easily in tabular schedules and are generally absent entirely from IFC export files. Examples are the embedding relationships between windows and walls and the structural support connectivity relationships between beams and columns. The result is that solutions either use the narrow aspects of the features of building objects in isolation, or that extensive programming is needed to extract and express the building object relationships as features of individual objects.

This is a severe limitation when one considers that the networks of relationships between building objects form quite distinct patterns. For example, the rule-based methods developed to classify bridge components in the SeeBridge project relied heavily on topological relationships between the objects. In that case the information was parsed from IFC files and the relationships were inferred by the rule processing system from the physical locations of the objects relative to one another before they could be classified (Sacks et al., 2017).

The difficulty in expressing the model objects and their relationships also points to a promising solution – expression of building models as property graphs. In pioneering work, Khalili and Chua (2015) developed a set of tools to compile a graph data model from an IFC file of a building model. Their method evaluated the geometry of the building objects to determine topological relationships and supplement the graph model with connectivity, containment, separation, and intersection relationships, which were modelled as weights on the graph's edges. Ismail et al. (2017) used a more suitable property graph representation, compiling both an IFC schema graph from EXPRESS and IFC model graphs for individual instances of building models stored in IFC files. They did not convert geometry, nor did they infer topological relationships from the geometry, sufficing with expressing information contained explicitly in the IFC files. Although neither of these attempted semantic enrichment using domain specific knowledge beyond the five basic topological relationships (Nguyen and Oloufa, 2001), the graph representations and the use cases they demonstrated with those representations strongly support the idea that property graphs are appropriate for modelling buildings for applications that require machine learning and/or pattern matching.

6. Conclusion

A review of three specific areas of Construction Tech, representing design and planning, BIM-to-field and field-to-BIM applications, reveals that the broad adoption of BIM environments in the construction industry is an insufficient condition to enable effective exploitation of the information they contain, or to leverage the potential of AI in this context. The problem is that the information in models is incomplete and inaccessible. Among the many technological challenges facing Construction Tech entrepreneurs, we have identified two specific research challenges that concern development of foundational information processing methods for digital building information models which, if solved, would greatly facilitate development of smart BIM and AI tools for design and construction. They are:

1. Combined, optimal use of topological rule inferencing and machine learning modules for **semantic enrichment**
2. Encoding **representations of building information** in forms that are amenable to machine learning

With regard to the nature of innovation in Construction Tech, our review of the areas of application supports researchers' predictions that technology innovation in construction is more likely to stem from disruptive startup companies than from the traditional project oriented construction companies (e.g. Katila et al., 2018). The growth of investment in Construction Tech startup companies demonstrates that the market shares this view. The 'House of Construction Tech' model may help investors and innovators alike in evaluating the soundness of their startups' technology and business strategies. Note that the model is applicable to incremental innovation; a complete rethinking of the construction business model may require rethinking of the foundational technologies too.

Naturally, we cannot claim to have identified all possible technological challenges to implementation of AI and BIM applications. There may be others, and presumably new problems will arise even as solutions to semantic enrichment and graph representations of BIM models are developed and implemented. We are confident, however, that these two are key to progress, and thus deserving of the attention of researchers.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgement

This work was supported by the Centre for Digital Built Britain, a partnership between the University of Cambridge and the Department for Business, Energy and Industrial Strategy, and the Construction Innovation Hub. The funding was provided through the Government's modern industrial strategy by Innovate UK, part of UK Research and Innovation.

References

Al Qady, M., Kandil, A., 2010. Concept relation extraction from construction documents using natural language processing. J. Construct. Eng. Manag. 136 (3), 294–302.
Andersen, K., Forr, T., 2018. The State of Construction Technology. Jones Lang LaSalle IP, Inc., p. 12
Azevedo, M.A., 2019. Investor Momentum Builds for Construction Tech. *Crunchbase News*, San Francisco, CA.
Ballard, G., 2000. The Last Planner System of Production Control. PhD Dissertation. University of Birmingham, Birmingham, U.K.
Ballard, G., 2008. The lean project delivery system: an update. Lean Constr. J. 2008, 1–19.

Bekkelien, A., Deriaz, M., Marchand-Maillet, S., 2012. "Bluetooth Indoor positioning." *Master's Thesis*. University of Geneva.
Belsky, M., Sacks, R., Brilakis, I., 2016. Semantic enrichment for building information modeling. Comput. Aided Civ. Infrastruct. Eng. 31 (4), 261–274. https://doi.org/10.1111/mice.12128.
Bishop, C.M., 2016. *Pattern Recognition and Machine Learning*. Information Science and Statistics. Springer, New York, New York, NY.
Blanco, J.L., Mullin, A., Pandya, K., Sridhar, M., 2017. The New Age of Engineering and Construction Technology. McKinsey & Company, Philadelphia, PA.
Bloch, T., Katz, M., Yosef, R., Sacks, R., 2019. Automated Model Checking for Topologically Complex Code Requirements – Security Room Case Study. In: O'Donnell, J. (Ed.), 2019 European Conference on Computing in Construction. EC3, Chania, Crete, Greece.
Bloch, T., Sacks, R., 2018. Comparing machine learning and rule-based inferencing for semantic enrichment of BIM models. Autom. ConStruct. 91, 256–272. https://doi.org/10.1016/j.autcon.2018.03.018.
Borrmann, A., Van Treeck, C., Rank, E., 2006. Towards a 3D spatial query language for building information models. In: Proc. Joint Int. Conf. of Computing and Decision Making in Civil and Building Engineering (ICCCBE-XI).
Braid, I.C., 1973. Designing with Volumes. Cantab Press, Cambridge University, Cambridge UK.
Brilakis, I., Haas, C., 2020. Infrastructure Computer Vision. Butterworth-Heinemann, Oxford.
Brilakis, I., Soibelman, L., Shinagawa, Y., 2005. Material-based construction site image retrieval. J. Comput. Civ. Eng. 19 (4), 341–355.
Buchmann, A., Koldehofe, B., 2009. Complex event processing. IT Inf. Technol. 51 (5).
Casale, D., 2013. Project Lion - A DPR/Trimble Automated Layout Robot. DPR Construction, Redwood City, CA. https://www.youtube.com/watch?v=k-lfL3dA1S.
Chi, H.-L., Kang, S.-C., Wang, X., 2013. Research trends and opportunities of augmented reality applications in architecture, engineering, and construction. Autom. ConStruct. 33, 116–122. https://doi.org/10.1016/j.autcon.2012.12.017.
Chowdhury, T., Adafin, J., Wilkinson, S., 2019. Review of digital technologies to improve productivity of New Zealand construction industry. J. Inf. Technol. Construct. 24, 569–587. https://doi.org/10.36680/j.itcon.2019.032.
Costin, A., Pradhananga, N., Teizer, J., 2012. Leveraging passive RFID technology for construction resource field mobility and status monitoring in a high-rise renovation project. Autom. ConStruct. 24, 1–15.
Degani, A., Li, W.B., Sacks, R., Ma, L., 2019. An automated system for projection of interior construction layouts. IEEE Trans. Autom. Sci. Eng. 16 (4), 1825–1835. https://doi.org/10.1109/TASE.2019.2897135.
Dimyadi, J., Pauwels, P., Amor, R., 2016. Modelling and accessing regulatory knowledge for computer-assisted compliance audit. Spec. Track Compl. Check. 21, 317–336. ITcon, Special issue CIB W78 2015.
Eastman, C.M., 1975. The use of computers instead of drawings in building design. J. Am. Inst. Archit. 63 (3), 46–50.
Efron, B., Hastie, T., 2016. Computer Age Statistical Inference: Algorithms, Evidence, and Data Science. Institute of Mathematical Statistics monographs, Cambridge University Press, New York, NY.
Galle, P., 2002. Philosophy of design: an editorial introduction. Des. Stud. 23 (3), 211–218.
Garrett, J.H., Fenves, S.J., 1987. A knowledge-based standards processor for structural component design. Eng. Comput. 2 (4), 219–238.
Gero, J., 2012. Design Optimization. Elsevier Science.
Gholizadeh, P., Esmaeili, B., Goodrum, P., 2018. Diffusion of building information modeling functions in the construction industry. J. Manag. Eng. 34 (2), 04017060.
Hakim, M., Garrett, J.H., 1993. Using description logic for representing engineering design standards. J. Eng. Comp. 9, 108–124.
Hall, D.M., Whyte, J.K., Lessing, J., 2019. Mirror-breaking strategies to enable digital manufacturing in Silicon Valley construction firms: a comparative case study. Construct. Manag. Econ. 1–18.
Hayes-Roth, F., Waterman, D.A., Lenat, D.B., 1983. Building Expert Systems. Addison-Wesley, Reading, Massachusetts.
Ismail, A., Nahar, A., Scherer, R., 2017. Application of graph databases and graph theory concepts for advanced analysing of BIM models based on IFC standard. In: European Group for Intelligent Computing in Engineering, vol. 12. Nottingham, UK.
ISO 10303-11, 2004. Industrial automation systems and integration — Product data representation and exchange — Part 11: description methods: The EXPRESS Language Reference Manual. International Standards Organisation. https://www.iso.org/standard/57620.html.
ISO, 2013. ISO 16739:2013 Industry Foundation Classes (IFC) for Data Sharing in the Construction and Facility Management Industries. International Standards Organization.
ISO, 2018. ISO 19650-1, Organization and Digitization of Information about Buildings and Civil Engineering Works, Including Building Information Modelling (BIM) — Information Management Using Building Information Modelling — Part 1: Concepts and Principles. International Standards Organisation.
Jin, C., Xu, M., Lin, L., Zhou, X., 2018. "Exploring BIM Data by Graph-Based Unsupervised Learning:" *Proceedings of the 7th International Conference on Pattern Recognition Applications and Methods*. SCITEPRESS - Science and Technology Publications, Funchal, Madeira, Portugal, pp. 582–589.
Katila, R., Levitt, R.E., Sheffer, D., 2018. Systemic innovation of complex one-off products: the case of green buildings. In: Joseph, J., Baumann, O., Burton, R., Srikanth, K. (Eds.), Advances in Strategic Management. Emerald Publishing Limited, pp. 299–328.
Kenley, R., Seppänen, O., 2010. Location-Based Management for Construction: Planning, Scheduling and Control. Spon Press.
Khalili, A., Chua, D.K.H., 2015. IFC-based graph data model for topological queries on building elements. J. Comput. Civ. Eng. 29 (3), 04014046.
Kiliccote, H., Garrett, J.H., 1995. "Obstacles to the Development of Computable Models of Design standards." *Workshop on Computing and Information in Construction: Modeling of Buildings Through Their Lifecycle*. Stanford University, Stanford, CA.
Koo, B., Shin, B., 2018. Applying novelty detection to identify model element to IFC class misclassifications on architectural and infrastructure Building Information Models. J. Comput. design Eng. 5 (4), 391–400.
Koo, B., La, S., Cho, N.-W., Yu, Y., 2019. Using support vector machines to classify building elements for checking the semantic integrity of building information models. Autom. ConStruct. 98, 183–194.
Koskela, Lauri, 2000. An Exploration towards a Production Theory and its Application to Construction. D. Tech, Helsinki University of Technology, Espoo.
Krijnen, T., Tamke, M., 2015. In: Thomsen, M.R., Tamke, M., Gengnagel, C., Faircloth, B., Scheurer, F. (Eds.), "Assessing Implicit Knowledge in BIM Models with Machine Learning." *Modelling Behaviour*. Springer International Publishing, Cham, pp. 397–406.
Lee, S., Akin, Ö., 2011. Augmented reality-based computational fieldwork support for equipment operations and maintenance. Autom. ConStruct. 20 (4), 338–352.
Maher, M.L., Fenves, S.J., 1985. HI-RISE - an expert system for the preliminary structural design of high rise buildings. In: Gero, J.S. (Ed.), Knowledge Engineering in Computer-Aided Design. North-Holland, Amsterdam, pp. 125–135.
Mazairac, W., Beetz, J., 2013. "BIMQL ? An open query language for building information models. Adv. Eng. Inf. 27 (4), 444–456.
McCullouch, B.G., 1992. "Automated construction field-data management system. J. Transport. Eng. 118 (4), 517–526.
McCullouch, B.G., Lueprasert, K., 1994. "2D bar-code applications in construction. J. Construct. Eng. Manag. 120 (4), 739–752.
Mitchell, W., Liggett, R., Tan, M., 1990. Top-down knowledge-based design. In: McCullough, M., Mitchell, W., Purcell, P. (Eds.), The Electronic Design Studio: Architectural Knowledge and Media in the Computer Era. MIT Press, Cambridge, MA, pp. 137–148.
Nahar, A., 2017. "Applying Graph Theory Concepts for Analyzing BIM Models Based on IFC standards." Master's Thesis. Technische Universität Dresden, Dresden, Germany.
Navon, R., 2005. Automated project performance control of construction projects. Autom. ConStruct. 14 (4), 467–476.
Navon, R., Sacks, R., 2007. Assessing research issues in automated project performance control (APPC). Autom. ConStruct. 16, 474–484.
Nawari, N.O., 2017. Building Information Modeling: Automated Code Checking and Compliance Processes. CRC Press, Boca Raton.
Nguyen, T.-H., Oloufa, A.A., 2001. Computer-generated building data: topological information. J. Comput. Civ. Eng. 15 (4), 268–274.
Prouty, J. M. (2013). Robotic Construction Site Marking Apparatus. Chula Vista, CA. US Patent Office Application no. 13/471,883. Filed May 15th 2012. 11 pages.
Rogers, S., Girolami, M., 2017. A First Course in Machine Learning. Chapman & Hall/CRC Machine Learning & Pattern Recognition Series. CRC Press, Taylor & Francis Group, a Chapman & Hall book, Boca Raton London New York.
Sacks, R., 2016. What constitutes good production flow in construction? Construct. Manag. Econ. 34 (9), 641–656. https://doi.org/10.1080/01446193.2016.1200733.
Sacks, R., Buyukozturk, O., 1987. Expert interactive design of concrete columns under Bi-axial bending. ASCE J. Comp. Civil Eng. 1, 69–81.
Sacks, R., Eastman, C.M., Lee, G., 2003. Process improvements in precast concrete construction using top-down parametric 3-D computer-modeling. J. Prec. Prest. Conc. Inst. 48, 46–55.
Sacks, R., Eastman, C.M., Lee, G., 2004. Parametric 3D modeling in building construction with examples from precast concrete. Autom. ConStruct. 13, 291–312.
Sacks, R., Eastman, C.M., Lee, G., Teicholz, P., 2018. BIM Handbook: A Guide to Building Information Modeling for Owners, Designers, Engineers, Contractors and Facility Managers. John Wiley and Sons, Hoboken, NJ.
Sacks, R., Bloch, T., Katz, M., Yosef, R., 2019. Automating design review with artificial intelligence and BIM: state of the art and research framework. In: Behzadan, A.H., Cho, Y., Leite, F., Wang, C. (Eds.), Computing in Civil Engineering 2019: Visualization, Information Modeling, and Simulation. American Society of Civil Engineers, Atlanta, GA, USA, pp. 353–360.
Sacks, R., Ma, L., Yosef, R., Borrmann, A., Daum, S., Kattel, U., 2017. Semantic enrichment for building information modeling: procedure for compiling inference rules and operators for complex geometry. J. Comput. Civ. Eng. 31 (6), 04017062 https://doi.org/10.1061/(ASCE)CP.1943-5487.0000705.
Sacks, R., Navon, R., Brodetskaia, I., 2006. Interpretation of automatically monitored equipment data in construction control. J. Comput. Civ. Eng. 20, 111–120. https://doi.org/10.1061/(ASCE)0887-3801(2006)20:2(111).
Song, J., Kim, J., Lee, J.-K., 2018. NLP and Deep Learning-Based Analysis of Building Regulations to Support Automated Rule Checking System, 7.
Trimble, 2020. "Trimble XR10 with Hololens 2." Trimble mixed reality. https://mixedreality.trimble.com/.
Turk, Z., Isakovic, T., Fischinger, M., 1994. Object-oriented modeling of design system for RC buildings. J. Comput. Civ. Eng. 8, 436–453.
Wang, X., Kim, M.J., Love, P.E.D., Kang, S.-C., 2013. Augmented Reality in built environment: classification and implications for future research. Autom. ConStruct. 32, 1–13.
Warszawski, A., Sangrey, D.A., 1985. Robotics in building construction. J. Construct. Eng. Manag. 111 (3), 260–280.
Winograd, T., Flores, F., 1986. Understanding Computers and Cognition: a New Foundation for Design. Intellect Books, Bristol, UK.
Woodward, C., Kuula, T., Honkamaa, P., Hakkarainen, M., Kemppi, P., 2014. Implementation and evaluation of a mobile augmented reality system for building

maintenance. In: Dawood, N., Alkass, S. (Eds.), Proceedings of the 14th International Conference on Construction Applications of Virtual Reality, pp. 306–315. Sharjah, UAE.

Wu, J., Zhang, J., 2019. New automated BIM object classification method to support BIM interoperability. J. Comput. Civ. Eng. 33 (5), 04019033.

Wülfing, A., Windisch, R., Scherer, R., 2014. "A visual BIM query language." eWork and eBusiness in Architecture, Engineering and Construction. ECPPM 2014, 157.

Zhang, J., El-Gohary, N.M., 2017. Integrating semantic NLP and logic reasoning into a unified system for fully-automated code checking. Autom. ConStruct. 73, 45–57.