Model-driven Software Engineering for Construction Engineering: Quo Vadis?

Sebastian Götz\textsuperscript{a} Andreas Fehn\textsuperscript{a} Frank Rohde\textsuperscript{a} Thomas Kühn\textsuperscript{b}

\textsuperscript{a}. Technische Universität Dresden, Software Technology Group
\textsuperscript{b}. Karlsruhe Institute of Technology, Software Design and Quality Group

\textbf{Abstract} Models are an inherent part of the construction industry, which leverages from the steady advancements in information and communication technology. One of these advancements is Building Information Modeling (BIM), which denotes the move from 2D drawings to having semantically rich models of the objects subject to construction. Additionally, the way stakeholders collaborate in construction projects and their organization is revisited. This is commonly denoted as Integrated Project Delivery (IPD). Both BIM and IPD originate from the basic principles of Lean Construction, the vision to minimize waste, increase value, and continuous improvement. The application of Model-driven Software Engineering (MDSE) to BIM is a natural choice. Although several approaches utilizing MDSE for BIM have been proposed, so far no structured overview of the current state of the art has been conducted. Such an overview is vitally needed, because the existing literature is fragmented among multiple research areas. Consequently, in this paper, we present a systematic literature review on the application of MDSE to BIM, IPD and Lean Construction resulting in a systematically derived taxonomy, which we used to classify 97 papers published between 2008 and 2018. Based on the taxonomy, we provide an analysis of the classified research showing (a) where the discourse on model-driven construction engineering currently is, (b) the state of the art of model-driven techniques in construction engineering and (c) open research challenges.

\textbf{Keywords} Systematic Literature Review, Building Information Modeling, Integrated Project Delivery, Lean Construction
1 Introduction

Model-driven Software Engineering (MDSE) is a highly interdisciplinary research area with known applications in many domains, such as the automotive industry, avionic systems, traditional enterprise software systems, and many more. Surprisingly, although a considerable body of literature on the application of MDSE to the field of construction engineering exists, a structured overview from an MDSE perspective is missing.

Throughout the ever increasing capabilities of information technology in general, the use of computer-aided automation techniques in construction engineering increased. This development lead to three main research initiatives, which are highly relevant to the MDSE community: Building Information Modeling (BIM), Lean Construction and Integrated Project Delivery (IPD). Figure 1 shows, based on data provided by SCOPUS, how these initiatives evolved between 2005 and 2018 in terms of publication activity. Two conclusions can be drawn: (a) Lean Construction was the first initiative among the three, and (b) BIM emerged to be the most active initiative by now.

Lean Construction [Kos92] aims at improving project management in construction by applying the principles of the Toyota Production System [Mon11] to construction, i.e., reduction of waste, increase of customer’s value and continuous improvement. The goal to reduce waste implies the goal to increase reuse of development and design artifacts, as each artifact that is not reused can be considered waste. Notably, increased reusability is among the goals of MDSE [SVC06].

The vision of BIM [ETSL11] is to move from 2D drawings to semantically rich models of the objects subject to construction. This idea has already been discussed in the 70’s by Charles Eastman [Eas75] using the term Building Description Systems. The term BIM in its current meaning has been first used by Robert Aish in 1986 [Ais86] and was coined as a term in 1992 by van Nederveen and Tolman [vNT92]. By now, the vision of BIM includes more than the move from drawing to modeling, but captures the effects of this change to the whole lifecycle of construction projects [EETS11]. As stated by Succar [Suc09] the maturity of BIM adoption can be measured in four levels:
Level 0 (Pre-BIM): engineers use 2D CAD tools and typically exchange their drawings by mail.

Level 1 (Object-based): 2D and 3D CAD tools and a common data environment, hosted by the main contractor, are used.

Level 2 (Model-based): 3D CAD tools are used and the models are exchanged using a common, standardized data format (e.g., Industry Foundation Classes (IFC)).

Level 3 (Network-based): 3D CAD tools are used and all project partners work on a common, shared model.

According to a recent study by the National Bank of Scotland\(^1\), most companies are currently between level 1 and 2.

IPD\(^2\) [MH05] is the third initiative and represents the eventual goal of BIM [Suc09]. IPD focuses on the process of how the individual stakeholders in a construction project communicate. IPD aims to overcome contractual problems between the stakeholders of a construction project with a team-based approach, where all partners work on the same model. For this, advanced model-driven techniques like multi-level modeling [dLG10], multi-paradigm modeling [MV04] and mega-modeling [BJV04] are required.

Notably, over the last decade, several approaches utilizing model-driven techniques to realize the vision of BIM and, ultimately, IPD, have been proposed. But in most cases, as we will show in this study, these approaches have not been developed by software engineers but by construction engineers. A plethora of literature surveys on Building Information Modeling has been published over the last decade. Many of them are referenced and aggregated in a recent survey by Hosseini et al. [HMA+18]. Surprisingly, no consolidated literature review from a model-driven perspective exists yet.

Hence, in this paper, we present a systematic literature review on the application of MDSE to BIM, IPD and Lean Construction, with the aim to provide an introduction to this application domain for the MDSE community.

To this end, we present a systematically derived taxonomy, which we used to classify 97 papers on BIM, IPD and Lean Construction published between 2008 and 2018, representing an abstract overview of, and thus, an introduction to the research area. Based on the taxonomy, we provide an analysis of the classified research showing (a) where the discourse on model-driven construction engineering currently is, (b) the state of the art of model-driven techniques in construction engineering and (c) open research challenges for the MDSE community.

To conduct our study, we used the Systematic Literature Review Toolkit [Göt18], which is publicly available as open-source software. We have put all data of this study on the web\(^3\) including the full list of included papers, the taxonomy and the classification.

The remainder of this paper is structured as follows. Section 2 elaborates on the research method used to collect relevant literature. The taxonomy is presented and analyzed in Section 3. A discussion is provided in Section 4 reflecting on the adoption of MDSE results in the AEC community. Section 5 discusses potential threats to validity for our study and Section 6 presents potential future research topics for construction.

---

1. https://www.thenbs.com/knowledge/nbs-national-bim-report-2017
2. http://info.aia.org/SiteObjects/files/IPD_Guide_2007.pdf
3. https://github.com/sebastiangoetz/slr-toolkit/tree/master/examples/bim-ecmfa20
engineering from a model-driven software engineering perspective. Finally, Section 7 concludes the paper summarizing our answers to the identified research questions.

2 Method

We conducted a systematic literature review according to the procedure described by Kitchenham [Kit04]. In this section, we elaborate on how we instantiated the generic process to cover the topic of this paper. For this, we first present the research questions that we aim to answer with our study. Next, in Section 2.1, we describe our process of collecting relevant literature and present general insights gained from the included literature. The general process is depicted in Figure 2. To define the scope of our survey, we formulated the following research questions:

(RQ1) Where (conferences, journals, etc.) is the discourse on MDSE for BIM, IPD and Lean Construction currently conducted?

(RQ2) Which model-driven techniques have already been applied to BIM, IPD and Lean Construction?

(RQ3) Which model-driven techniques should be applied in the future?

2.1 Collection of Relevant Literature

The body of literature on BIM, Lean Construction and IPD is large. A simple search on Google Scholar in October 2019 returned approximately 31,000 hits for the keyword “Building Information Modeling”, more than 21,000 for “Lean Construction” and almost 7,000 for “Integrated Project Delivery”.

To get a first overview on the topic, we started our survey by searching for the term BIM on Google Scholar and investigated the first 60 papers listed in the search results. Based on these findings, we conducted our study using the following steps:

• **Query**: We searched for the keywords “Building Information Modeling”, “Lean Construction” and “Integrated Project Delivery” combined with the keywords “model-based” and “model-driven”. Additionally, we used the keywords “bim collaboration format” and “model view definition”. Thus, in total, we executed eight queries against Google Scholar. We narrowed our search by only investigating work of the last ten years, i.e., since 2008 until 2018. For each query, we
Table 1 – Overlap between Individual Query Result Sets

|        | Model-based (MB) | Model-driven (MD) | BCF | MVD |
|--------|------------------|-------------------|-----|-----|
|        | BIM   | IPD   | LC  | BIM   | IPD   | LC  | BCF | MVD |
| MB     | 3722  | 497   | 404 | 163  | 17    | 16  | 42  | 189 |
|        | 497   | 748   | 237 | 21   | 29    | 8   | 20  | 42  |
|        | 404   | 237   | 1164| 18   | 10    | 30  | 16  | 18  |
|        | 163   | 21    | 18  | 359  | 32    | 32  | 4   | 29  |
| MD     | 17    | 29    | 10  | 32   | 49    | 12  | 1   | 9   |
|        | 16    | 8     | 30  | 32   | 12    | 64  | 3   | 5   |
|        | 42    | 20    | 16  | 4    | 1     | 3   | 180 | 51  |
| BCF    | 189   | 42    | 18  | 29   | 9     | 5   | 51  | 678 |

extracted all entries found by Google Scholar on a yearly basis using the tool gsresearch\(^4\). In total, 6966 papers were found for all eight queries (see Table 2).

- **Filter (I):** We manually processed each so-derived list of papers by applying certain in- and exclusion criteria (defined below) on the papers’ title and type of publication (first filter step). As we divided this work among the co-authors of this paper, we did not eliminate duplicate entries among the yearly paper lists of the eight queries. Consequently, papers with titles, which do not adequately reflect the potential of the paper to contain work on model-driven software engineering, had a higher chance to not be excluded at this early point in the study. Notably, as shown in Table 1, the overlap between the individual queries was not negligible. After this filter step only 643 papers were remaining of which 99 were duplicated across the query result sets (see Table 2).

- **Filter (II):** The second filter step was split in two parts (in the following denoted as selection and inclusion).
  - **Selection:** First we determined for each paper whether the abstract gave hints about model-driven software engineering aspects. Alongside, we also created the initial taxonomy and removed the 99 duplicated papers.
  - **Inclusion:** In the second part, we studied each of the remaining papers in detail and created the final taxonomy. We did not perform the selection on a per-query basis, but merged the result sets on a yearly basis. The count of filtered and included papers per year is detailed in Table 3.

2.1.1 Inclusion Criteria

We included papers, which describe an approach using an MDSE-related technique (e.g., model transformation, model comparison, etc.). Additionally, we included surveys, which potentially reference papers meeting the above criteria. For the final list of included papers, these were excluded again.

\(^4\)https://github.com/Eden-06/gsresearch
Table 2 – Metadata-based Filtering of Papers per Query

| Query | Raw | Filtered |
|-------|-----|----------|
| BIM   | 3722| 264      |
| MB    | 748 | 43       |
| LC    | 1164| 81       |
| BIM   | 359 | 89       |
| MD    | 49  | 14       |
| LC    | 64  | 22       |
| BCF   | 182 | 14       |
| MVD   | 678 | 116      |
| Total | 6966| 643      |

Table 3 – Inclusion of Papers on a Yearly Basis, whereas 99 Papers were Duplicates

| Year | Filtered | Included |
|------|----------|----------|
| 2008 | 22       | 3        |
| 2009 | 44       | 4        |
| 2010 | 36       | 6        |
| 2011 | 43       | 7        |
| 2012 | 73       | 10       |
| 2013 | 66       | 13       |
| 2014 | 60       | 8        |
| 2015 | 60       | 14       |
| 2016 | 59       | 11       |
| 2017 | 55       | 14       |
| 2018 | 26       | 7        |
| Total| 544      | 97       |

2.1.2 Exclusion Criteria

We excluded papers, which are: patents, not written in english, books (only individual chapters are included) and doctoral or master’s theses (too coarse-grain, if they were valuable, there would have been publications by the author on subtopics).

2.2 Centers of Scientific Discourse

Responding to Research Question 1 we determined the centers of scientific discourse regarding model-driven construction engineering. From 97 included papers 50 have been published at scientific forums that are concerned with digital innovations in construction engineering, only 9 have been published at more general forums concerned with digital innovations in engineering, and still 9 have been published at general or more specific computer science forums, e.g., concerning applied informatics (general) and semantic computing (specific). The other papers have been published at forums of numerous related research directions, e.g., AEC in general (5 papers), industrial applications of digital technologies (2 papers) and geoinformation (3 papers). Most papers have been published at the Elsevier journal Automation in Construction (13 papers), the ASCE Journal of Computing in Civil Engineering (8 papers) and the Elsevier journal Advanced Engineering Informatics (6 papers). One paper, namely [SDD11], was published at the International Conference on Model Transformation, which has been organized by the MDSE community. But this paper was contributed by the MDSE expert Jim Steel and not by an original member of the AEC research community.

In conclusion, in reply to Research Question 1 we state that research on MDSE for BIM, IPD and Lean Construction focuses on forums that originate in the AEC research community and are primarily concerned with digital innovations within this field. No single author from the AEC research community has presented progress in this matter to the MDSE community at one of the leading conferences, such as, MODELS and ECMFA.
3 Taxonomy

In this section Research Question 2 is answered by describing the final taxonomy we derived after completing the Selection step. The purpose of this taxonomy is to provide an overview on the current state of the art. For this, we introduced 5 top level categories: technologies from construction engineering, modeling techniques, the metalevel at which modeling techniques were used, interfacing research areas, and application domains. While the purpose of the first two categories is to provide an overview of existing approaches, the three last categories are meant to provide insight into (a) how advanced employed modeling techniques are, (b) which other research areas are involved and (c) if and where the approaches have been evaluated, i.e., how mature the conducted research is. An overview over the taxonomy can be seen in Figure 3.
3.1 Common Technologies in the Construction Industry

This section provides an overview of technologies used in construction engineering, which we identified through our survey, i.e., technologies linked to model-driven engineering. Although most papers clearly focus on a particular set of technologies, we found four publications not targeting any concrete technology, i.e., [CH11, CAB11, LCL+12, CLC13]. Chiang et al. [CLC13], for instance, proposed to extend building information models with additional information for disaster mitigation, but did not restrict their approach to any particular technology.

In the following, each identified technology is outlined separately, whereas the number of papers, which referred to this technology, is shown in parentheses.

3.1.1 Industry Foundation Classes (77)

Industry Foundation Classes (IFC), developed by buildingSMART, is an open standard (ISO 16739:2013), which can be used to model extensive information from various domains concerning a building and its management during all of its lifecycle phases. IFC is an object-oriented data model and defines diverse entities of multiple domains to ease the process of modeling a building. Arbitrary properties can be assigned to entities through generic references and objectified relations to capture the physical and logical traits of building objects [FFG+99]. IFC has become the standard exchange format for BIM applications, therefore it is used by many publications, e.g., Tauscher et al. [TBS16] developed a generic approach to query BIM data from IFC models.

3.1.2 Model View Definition (23) and Information Delivery Manual (14)

While IFC can indeed be used to exchange information about buildings, it lacks a description of which data shall be included in exchanges needed for specific data use cases. This can be remedied by using the ISO 29481–1 standard Information Delivery Manual (IDM). An IDM is used to specify which information should be exchanged at specific points during a construction project [EJSK09, LPH13].

A Model View Definition (MVD) on the other hand is a specification, developed by buildingSMART, that can be used to document or describe a subset of an IFC schema. MVDs have two main uses. First buildingSMART uses them to generate the user documentation of IFC4; and secondly schema subsets are described to satisfy exchange requirements in the AEC sector [LPH13]. Thus, they can be used to implement IDMs. As both are closely related to each other, many approaches use IDMs to generate specific MVDs, e.g., [BHRA10, SAL+12, GGK12, LES16].

3.1.3 Open BIM Collaboration Format (2)

The Open BIM Collaboration Format (BCF) was adopted by buildingSMART and provides a way to exchange issues, proposals, and change requests in BIM data models. The location of a problem can be described very fine-grained via references to individual model elements and attached screenshots or camera perspectives. In our study both Gurtler et al. [GBS15] and Zhang et al. [ZBW15] used BCF as a neutral exchange format for communication between different applications.

3.1.4 Standard for the Exchange of Product Model Data (8)

The Standard for the Exchange of Product Model Data (STEP) is a very large ISO standard (ISO 10303) for the exchange of product data between electronic systems.

https://www.buildingsmart.org/
It consists of many parts and spans multiple product categories and lifecycle stages. The product entities exchanged with STEP are described by EXPRESS schemas (e.g., the metamodel of IFC is specified using EXPRESS). STEP contains Parts 21 and 28, which cover concrete file formats for STEP based on plain-text and XML [Pra01]. IFC models are commonly exchanged with files in one of those two formats. As such, Nour [Nou08] used STEP to generate code for reading and writing data and Steel et al. [SDD11] converted IFC into another technical space by utilising STEP.

3.1.5 City Geography Markup Language (10) and Geographic Information System (8)
The City Geography Markup Language (CityGML) is a profile of the XML-based Geography Markup Language to capture and exchange spatial information of whole cities. CityGML models are divided in different levels of detail (LOD) that range from plain areas over simple blocks to detailed models of individual buildings including their interior [KGP05].

Geographic Information System (GIS) is a broad term to describe a system designed to handle geographic data, such as coordinates and areas. The system can provide functionality, such as capturing, analyzing, and visualizing geographic data. Thus, GIS have various application domains, e.g., urban planning or evaluating specific patterns [KLR+95]. Still few approaches, e.g., [XDLM14, MN14, KH15], improve the semantics of geographic information systems by creating mappings between IFC and CityGML.

3.1.6 Web Ontology Language (16)
The Web Ontology Language (OWL) is a formal language for the semantic web. It is used (a) to describe categories of objects and how objects of these categories relate to each other and (b) to describe information about the objects themselves. As such, OWL can be used to represent both an ontology and data [BVHH+04].

There are efforts to develop an ontology for the building and construction domain to ease exchange of data by having common standards. One of those ontologies is ifcOWL [BVLDV09, PR17] which is generated by transforming the EXPRESS schema of IFC to OWL.

3.1.7 Business Process Modeling Notation (2)
The Business Process Modeling Notation (BPMN) is a graphical language to describe processes and workflows. The language is designed to be intuitively understandable by both technical and non-technical users, and thus, can be used to guide the concrete implementation in software [CT12]. Cheng et al. [CLB+10] used BPMN to describe business processes involved in supply chains for construction projects. BPMN can additionally be used to represent the output of an IDM/MVD when used in a buildingSMART specific way [GBS15].

3.1.8 Modelica (4)
Modelica is both the name of an object-oriented language and an environment for simulations. Simulations are defined by declarative equations, which the environment solves over time. Modelica can be used to simulate discrete or continuous systems over a variety of different domains [FE98]. In the literature, Modelica is mostly applied to simulate the energy efficiency of buildings, e.g., by Andriamamonjy et al. [ASK18].
3.1.9 General Model Subset Definition (4)
In [WKS03] Weise et al. proposed the General Model Subset Definition schema which is defined in EXPRESS and whose instances are used to filter IFC and other models. The definitions permit the automatic implementation of MVDs, as in [KWW+10, GG13, GGS12].

3.1.10 Technologies from Model-driven Software Engineering (18)
Interestingly, 18 out of 97 papers directly refer to technologies from MDSE. Namely, the Unified Modeling Language (11), the Eclipse Modeling Framework (5), the Systems Modeling Language (2), model transformation languages like ATL (2), QVT (2) and Tefkat (1), as well as the Meta-Programming System (1). In other words, first links between construction engineering and MDSE exist.

3.2 Modeling Techniques
This section covers the adoption of MDSE modeling techniques in the AEC community. We distinguish the modeling techniques from each other and mention several papers from the AEC research community to exemplify our decisions for their classification.

3.2.1 Model Merge (2)
In line with Lucio et al. [LAD+16], the taxonomy class Model Merge contains approaches that combine two or more models from arbitrary metamodels in such a way that all information from the models are preserved, duplicate or conflicting elements are joined together, and no additional knowledge is derived. Xu et al. [XDLM14], for instance, enriched the outdoor model of a city with indoor elements of buildings for City Information Modeling by unifying building data (IFC) with geographic data (CityGML). Rather the other way around, Mignard and Nicolle combined IFC and CityGML to build a united Urban Information Model by using an ontology [MN14].

3.2.2 Model Views (20)
Although the modeling community has a clear understanding about model views in this class we not only include papers that define views over one or more models to reduce complexity, but also those that generally strive to improve handling of model views. Despite the fact that the modeling domain allows views to be described by different metamodels than the base models [BBCW17] many of the papers under study employ MVDs, and thus, stay in one modeling space, e.g., [ASK18, PWOD+18].

Additionally, Lee et al. [LES18] verified the conformance of building information models with MVDs by converting the rules into description logic expressions and reason on them. Furthermore, Isikdag et al. [IUKAR10] proposed the use of domain-specific views around a model web service for different application aspects, whereas the views consist solely of geometric information.

3.2.3 Model Traceability (23)
The taxonomy class Model Traceability encompasses approaches to track information describing the relationships among different models oder model elements. For example, Vilgertshofer et al. [VAW+17] presented an approach to describe geographically extended structures like bridges and tunnels by integrating an IFC model with a CityGML model to put the building-specific information (IFC) into a geographical
context (CityGML). The trace links could be established automatically during model transformation or manually by the modeler. The established links are intended to be used by modeling tools to achieve model synchronization or to perform model analysis. Especially the latter requires semantically rich links that go beyond a simple mapping from one group of model elements to another, e.g., rich links that are annotated and allow for describing inter-dependent relationships like “link \( x \) is valid if link \( y \) exists and was established earlier than link \( z \”).

3.2.4 Model Querying (19)

The taxonomy class Model Querying contains approaches that present functionality that is concerned with extracting some well-specified information from an input model. For example, Mazairac and Beetz [MB13] developed a query language for building information models. The language allows for specifying queries to extract information from IFC models, e.g., a query for the set of rooms of a building whose individual area exceeds a certain value. Model queries may be performed in preparation of a model transformation to reduce the model input to a subset that is relevant for the transformation. Nevertheless, such approaches are not limited to returning information that is explicitly stated in the input model and may also return some additional information that was derived from the input model, e.g., a query may derive the emergency exit that is closest to a given room. To enhance the performance of such analytical queries Khalili and Chua [KC13] proposed a graph data model that represents building elements as nodes and topological relationships (e.g., containment, intersection) among building elements as edges.

3.2.5 Model Transformation (30)

With taxonomy class Model Transformation we classify approaches capable of translating models that conform to a metamodel \( x \) to models that conform to metamodel \( y \) or to textual expressions (e.g., source-code, documentation, configuration files). In [SDD11], for instance, Steel et al. described an infrastructure to bridge the STEP/EXPRESS and EMF technical space allowing to specify transformations for models that belong to the STEP/EXPRESS technical space using established MDSE model transformation tools. We include a wide range of approaches, effectively abstracting from the language in which the transformation is specified. Moreover, \( x \) and \( y \) may be equal metamodels. However, approaches that are typically employed for view definition or to enforce consistency among models are covered in the respective sections.

3.2.6 Model Semantics (14)

The taxonomy class Model Semantics covers approaches that strive for clarifying the very meaning of a model and for concluding further insights from the model. By concluding further insights we understand the formal or simulation-based validation of certain design rules of the (building) model, like the existence of a building height limitation to meet legal constraints, but also performing more advanced model analysis, such as determining the compliance of a detailed building model with a high-level architectural description. In preparation of a simulation-based validation Jeong et al. [JKC+14] described an approach for automatically translating an IFC model into a Modelica energy model.

By clarifying the very meaning we understand approaches to make a model more interpretable by mapping the model elements to elements of a (semantic) domain that is generally understood by the used modeling tool. For example, this is done by
Häfner et al. [HHWO13] and Fernando et al. [FSD11]. While the former describe an approach for extracting semantic information from CAD drawings into an ontology in a semi-automatic way, the latter suggest using IFC models as semantic domain.

### 3.2.7 Model Versioning (2)

The taxonomy class *Model Versioning* includes all approaches that strive to track changes inside the model. Those approaches create multiple snapshots of a whole model or individual model elements (possibly with an associated date and author). Consecutive versions are linked to each other to allow depicting a history of changes.

In our study only two papers were found that deal with versioning on a model level both merely considering IFC models. While they neglect generic model versioning concepts, they can create special approaches that are ideally adapted to the structure of an IFC model. In [JZKT15], Jaly-Zada et al. proposed to extend the EXPRESS description of IFC to include new entities recording changes in the structure of the model and changes in entity features. Oraskari et al. [OT15] developed algorithms for the generation of unique identifiers of IFC entities to enable the use of generic change management tools in construction projects.

### 3.2.8 Model Consistency (8)

The taxonomy class *Model Consistency* includes approaches that perform model transformations with the intent to keep multiple models in fulfillment of a certain, explicitly defined consistency relation. To this end, Lee et al. [LES18] presented a framework for rule validation and describe the semantics of rules that can be used for defining consistency relations for an IFC model and a view (defined using MVD) of this IFC model. Borrmann et al. [BFJ+14] described a modeling methodology that allows for the collaborative creation of building models in a way that systematically keeps their consistency using automatic model updates and locks based on specified model dependencies that have been explicitly defined by the modeler.

### 3.3 Metalevels

To gain a deeper knowledge about the way model-based approaches are used in the body of literature, we studied the meta-levels that the authors are working at. Our classification is based on the layered metadata architecture by the Object Management Group\(^6\), i.e., we consider the levels M0 (0 papers), M1 (39), M2 (46), and M3 (8) in increasing order of abstraction. For the classification we always chose the highest meta-level used. As there are also papers which cannot be classified into these categories or misclassified themselves, we added an additional class *Not applicable* (4).

As it can be argued that one needs only three meta-levels to work with IFC models and, many papers use IFC, we illustrate how we fit IFC into the above mentioned meta-model hierarchy. At the most abstract level M3 resides the EXPRESS data modeling language. Its grammar describes how to model product data on M2. Here the structure of the IFC entities is located, such as IfcWall or IfcDoor. A concrete model on M1 consists of instances of these entities, i.e., a wall with two doors. The runtime objects of a tool that operates on the model can be thought of being at level M0.

\(^6\)https://www.omg.org/spec/MOF/1.4/PDF
3.4 Interfacing Research Areas

Although most papers are working isolated in the AEC community (58), many touch other areas of active research. For the study, we identified these areas as Knowledge Engineering (18), Cloud and Service-oriented Architecture (13), Business Processes (7) and Data Warehousing (6).

3.5 Application Domains

As a last major class in the taxonomy we investigated the target domains of the studies. Many (26) do not have a specific target but represent generic approaches. Other main areas we found are (energy) Simulation (18), Facility management (18), Construction (18), Architecture (14), Process Optimization (8) and Infrastructure (3).

4 Reflections on the Adoption of MDSE Results within the AEC Community

Our study has shown that the methodology of (building-)modeling is actually an aspect of research within the AEC community, although researchers that are concerned with this aspect are not necessarily aware of the corresponding results of the MDSE community. Indeed, the majority of papers is unaware of these results and only individual papers utilize existing MDSE modeling tools and approaches for use in the AEC research domains.

For example, Stratbücker et al. [SZM17] proposed to achieve interoperability of BIM tools by setting up a model-driven toolchain that includes the use of Ecore models and the definition of ATL transformations. Nevertheless, the application of feature-rich MDSE tools from the Ecore technical space would most certainly have created the chance for improvements, e.g., Xu et al. [XDLM14] described a vision that suggests to automatically populate CityGML models with data taken from IFC models. Given a bridge to the Ecore technical space, the prototype could be based on the Eclipse plugin EMF Compare\(^7\), which allows for integrating two models with each other, but still lacks automation functionality.

Additionally, Jaly-Zada et al. [JZKT15] expanded IFC to record changes to entities and individual features of entities. Although they validated their approach by implementing a prototypical difference viewer, they could have benefited largely from the previous work done by the modeling community had they conducted their work in the Ecore technical space using EMF-based tools like EMF DiffMerge\(^8\), which is an approach that was actually even suggested by Krämer and Huhnt from the AEC community [KH14].

One surprising result of our study was the poor tool support for clash detection within building models, which could have been achieved as an add-on to a tool primarily concerned with model merging [CFJ\(^+\)16]. This could be regarded as being encouraged by the slow uptake of feature-rich MDSE modeling tools into the AEC domain.

---

\(^7\)https://www.eclipse.org/emf/compare/

\(^8\)https://wiki.eclipse.org/EMF_DiffMerge
5 Threats to Validity

There are a few threats to validity to our study. Of course, any literature survey inherently involves the risk of missing some relevant publications, if, for instance, the search terms were not correctly chosen. We mitigated this threat by using the results of a prestudy that was carried out to gain a general insight into the publications in the AEC community. From this prestudy we extracted the initial keywords and important terms. By combining them with the very generic terms “model-based” and “model-driven”, we cover as much potentially relevant literature as possible. We restricted the search to papers published between 2008 and 2018 to be able to derive results that are relatively current especially with respect to Research Question 1.

Two additional issues concern our approach to filter the found literature. Our filter process is prone to losing valuable work, because as the first step we filter the yearly results only by title and publication type, as each year was filtered by only one researcher. We decided for this approach to handle the massive body of literature, although there exists a chance papers were excluded because the title did not sufficiently reflect the usage of MDSE techniques. In the second filter step, each paper was inspected by multiple researchers, such that relevant papers remained included with a high chance. Unfortunately, some venues of the AEC community are not publicly accessible and several papers we found were excluded not because of their content, but due to access restrictions.

As all authors are software engineers, we looked at the literature from a similar perspective. To remedy this author bias in the filtering process we defined clear criteria for inclusion and exclusion as suggested by Kitchenham [Kit04] and published all material that is necessary to reproduce our results.

6 Future Research on Model-driven Construction Engineering

Although we identified the usage of eight modeling techniques, a surprising result of our study is the lack of work in seemingly obvious research areas from a model-driven software engineering perspective. Namely, multi-paradigm modeling, model evolution, runtime models and multi-level modeling. Further interesting topics include domain specific modeling languages and techniques from model-based analysis to name but a few.

In the remainder, we provide an answer to research question 3 by elaborating on each of the mentioned four research topics in relation to the vision of Building Information Modeling in more detail.

6.1 Multi-paradigm Modeling

The vision behind multi-paradigm modeling is the integration of complex simulations with modeling techniques to enable the construction of complex systems comprised of physical and digital parts [VDLM02]. For this, a large variety of formalisms both from modeling and simulation need to be supported and linked to one another.

For Building Information Modeling, solutions to this problem are vital. Despite structural drawings of buildings created using CAD tools, the integration of complex simulations is a current challenge whose solution may considerably improve the design-and planning process of building projects, especially of infrastructure construction projects, like in [BBR+17, Baz08].
6.2 Model Evolution

Building models should be up to date and reflect all relevant information of a building. Buildings, and the building models associated to them, can be in use over a long time span and new operational aspects may gain importance over time, for instance, environmental sustainability (e.g., the Barwa Commercial Avenue, a giant shopping mall in Qatar, whose construction was successfully finished in 2012 using BIM-based technology [ST18]). The building model may have to be fundamentally changed after reutilization of the building, e.g., after a change in ownership. In other words, a building model may be subject to repeated changes that are relatively profound. Over time, the aggregation of those ongoing evolutionary changes may increase the complexity and decrease the maintainability of the building model in an unacceptable way.

Research on model evolution aims to address the underlying problems, e.g., how to evolve models on a lower metalevel, when changes on a higher metalevel are performed. Currently, there are no tools and approaches available that cope with the evolution of building models. Therefore, it would be reasonable to devote future work into the development of technology that enhances the manageability of the evolution of building models.

6.3 Models@run.time

Besides the need to deal with evolving metamodels, there is the need to keep a representation of the current building state. BIM does not only target the construction phase of buildings, but is meant to improve the productivity in all phases of a buildings life cycle including its maintenance, operation, and destruction. In particular, facility management demands for up-to-date building models.

Models that have been originally created for use during the construction phase of a building should be opened up for use during the other phases, i.e., throughout the whole life cycle of the building. The same problem area is discussed within the MDSE community in the Models@run.time subcommunity [BGS19] from a software engineering point of view. In our study, not a single paper was found, which approaches this problem.

6.4 Multi-level Modeling

With the entities IfcTypeObject and IfcRelDefinesByType IFC allows for typing objects decoupled from the inheritance hierarchy and supports through the usage of the powertype pattern a restricted version of multi-level modeling [AK01]. Surprisingly, only few papers mention these entities and only one elaborates on semantics of them and their subentities [BSST15].

Considering the complexity and high volatility in construction projects, there’s a clear need for more sophisticated modeling concepts, like multi-level modeling. For example, the construction and management of a shopping mall involves dozens of individual companies, each with their individual expertise (electricity, heating, ventilation and air conditioning, waste-water management, etc.). The building information model of the whole shopping mall will be subject to many changes at least throughout the construction phase. Considering the level of detail given by IFC, modelers are likely to prepare the building information model for future adjustments using the powertype pattern.
We have not found any approach aiming for better means of typing model elements within a building information model. This and the fact that IFC has been designed to support additional typing hierarchies shows the need for future work in this area.

7 Conclusion

In this paper, we provide an overview on the current state of the art of model-driven software engineering for construction engineering. For this, we systematically collected and analyzed in total 97 papers. Our aim was to answer three research questions.

The first question asked where the discourse on MDSE for construction engineering is held. The unexpected answer, presented in Section 2.2, is that the discourse is not held at software engineering conferences or journals, but mostly at venues from civil engineering. This implies that most of the existing work has not been conducted by researchers from software engineering, which is visible in some publications, e.g., by clear misunderstandings of modeling concepts [LCL+12, Ber16].

The second research question asked for the current state of the art. The detailed answer to this question is presented in Section 3. Three main conclusions can be drawn. First, there are clearly favored technologies in this field, which are mostly managed by the buildingSMART consortium. Namely, the Industry Foundation Classes (IFC), Model View Definitions (MVD) and Information Delivery Manuals (IDM). Interestingly, a separate language family for the complete metapyramid exists with EXPRESS at M3 and IFC for all lower levels. Second, technologies developed by the MDSE community have been scarcely used. These include UML and EMF, but also model transformation languages like ATL, QVT and Tefkat. Third, we identified eight MDSE research areas, which have been addressed from the BIM perspective, i.e., model merging, traceability, querying, views, transformations, versioning, semantics and model consistency.

This leads to our third research question, which asked for open research topics and is answered in detail in Section 6. By contrasting the BIM vision with the identified eight MDSE research areas, we determined four further research areas, which haven’t been addressed, yet, despite a clear need for solutions. Namely, multi-paradigm modeling, model evolution, models@run.time and multi-level modeling.

In summary, the answer to the question posed in the title of this paper is: MDSE approaches aiming to achieve the BIM vision and, by extension the visions of Lean Construction and Integrated Project Delivery, exist, but there is plenty of room for future work and a clear need to consolidate the two technological spaces so both research fields can benefit from each others advancements.

References

[Alk86] Robert Aish. Building modelling: the key to integrated construction cad. In 5th International Symposium on the Use of Computers for Environmental Engineering Related to Buildings, volume 5, pages 7–9, 1986.

[AK01] Colin Atkinson and Thomas Kühne. The essence of multilevel metamodeling. In International Conference on the Unified Modeling Language, pages 19–33. Springer, 2001. doi:10.1007/3-540-45441-1_3.
[ASK18] Ando Andriamamonjy, Dirk Saelens, and Ralf Klein. An automated ifc-based workflow for building energy performance simulation with modelica. *Automation in Construction*, 91:166–181, 2018. doi:10.1016/j.autcon.2018.03.019.

[Baz08] Vladimir Bazjanac. Ifc bim-based methodology for semi-automated building energy performance simulation. Technical report, Ernest Orlando Lawrence Berkeley National Laboratory, Berkeley, CA (US), 2008.

[BBCW17] Hugo Bruneliere, Erik Burger, Jordi Cabot, and Manuel Wimmer. A feature-based survey of model view approaches. *Software and Systems Modeling*, pages 1–22, 2017. doi:10.1007/s10270-017-0622-9.

[BBR+17] Martin Breunig, André Borrmann, Ernst Rank, Stefan Hinz, Thomas Kolbe, Matthäus Schilcher, Ralf-Peter Mundani, Javier Ramos Jubierre, Matthias Flurl, Andreas Thomsen, et al. Collaborative multi-scale 3d city and infrastructure modeling and simulation. *ISPRS-International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, 42:341–352, 2017. doi:10.5194/isprs-archives-XLII-4-W4-341-2017.

[Ber16] Marcelo Bernal. From parametric to meta modeling in design. *Blucher Design Proceedings*, 3(1):579–583, 2016. doi:10.5151/despro-sigradi2016-815.

[BFJ+14] André Borrmann, Matthias Flurl, Javier Ramos Jubierre, Ralf-Peter Mundani, and Ernst Rank. Synchronous collaborative tunnel design based on consistency-preserving multi-scale models. *Advanced Engineering Informatics*, 28(4):499–517, 2014. doi:10.1016/j.aei.2014.07.005.

[BGS19] Nelly Bencomo, Sebastian Götz, and Hui Song. Models@run.time: a guided tour of the state of the art and research challenges. *Software & Systems Modeling*, 18(5):3049–3082, Oct 2019. doi:10.1007/s10270-018-00712-x.

[BHRA10] Manfred Breit, Fritz Häubi, Michael Raps, and Nicolas Arnold. Process oriented model based information exchange between architecture and fabrication in early project phases. In *International Conference in Computing in Civil and Building Engineering*, page 156, 2010.

[BJV04] Jean Bézivin, Frédéric Jouault, and Patrick Valduriez. On the need for megamodels. 2004.

[BSŠT15] Stefano Borgo, Emilio M Sanfilippo, Aleksandra Šojić, and Walter Terkaj. Ontological analysis and engineering standards: an initial study of ifc. In *Ontology Modeling in Physical Asset Integrity Management*, pages 17–43. Springer, 2015.

[BVHH+04] Sean Bechhofer, Frank Van Harmelen, Jim Hendler, Ian Horrocks, Deborah L. McGuinness, Peter F Patel-Schneider, Lynn Andrea Stein, et al. Owl web ontology language reference. 2004.

[BVLDV09] Jakob Beetz, Jos Van Leeuwen, and Bauke De Vries. Ifcowl: A case of transforming express schemas into ontologies. *Ai Edam*, 23(1):89–101, 2009.
Akbas, Bousang Koo, et al. Industry foundation classes for project management—a trial implementation. *Electronic Journal of Information Technology in Construction*, 4:17–36, 1999.

[FSD11] Ruwan Fernando, James Steel, and Robin Drogemuller. Using domain specific languages in the building information modelling workflow. In *Proceedings of the 16th International Conference on Computer-Aided Architectural Design Research in Asia*, 2011.

[GBS15] Mario Gürtler, Ken Baumgärtel, and Raimar J Scherer. Towards a workflow-driven multi-model bim collaboration platform. In *Working Conference on Virtual Enterprises*, pages 235–242. Springer, 2015.

[GG13] Kamil Umut Gökçe and Hasan Ufuk Gökçe. Ifc based model view definition for hybrid energy systems. In *Proceedings of the 30th CIB W78 International Conference-October*, pages 9–12, 2013.

[GGK12] Kamil Umut Gökçe, Hasan Ufuk Gökçe, and Peter Katranuschkov. IFC-based product catalog formalization for software interoperability in the construction management domain. *Journal of computing in civil engineering*, 27(1):36–50, 2012.

[GGS12] KU Gökçe, HU Gökçe, and RJ Scherer. IFC view for product catalogues in the construction management domain. *eWork and eBusiness in Architecture, Engineering and Construction: ECPPM 2012*, page 281, 2012.

[Göt18] Sebastian Götz. Supporting systematic literature reviews in computer science: The systematic literature review toolkit. In *Proceedings of the 21st ACM/IEEE International Conference on Model Driven Engineering Languages and Systems: Companion Proceedings*, MODELS ’18, pages 22–26, New York, NY, USA, 2018. ACM.

[HHWO13] Polina Häfner, Victor Häfner, Hendro Wicaksono, and Jivka Ovtcharova. Semi-automated ontology population from building construction drawings. In *KEOD*, pages 379–386, 2013.

[HMA+18] M Reza Hosseini, Mojtaba Maghrebi, Ali Akbarnezhad, Igor Martek, and Mehrdad Arashpour. Analysis of citation networks in building information modeling research. *Journal of Construction Engineering and Management*, 144(8):04018064, 2018.

[IU-KAR10] Umit Isikdag, Jason Underwood, Murat Kuruoglu, and Alias Abdul-Rahman. Geospatial views for restful bim. In *Handbook of Research on Building Information Modeling and Construction Informatics: Concepts and Technologies*, pages 473–482. IGI Global, 2010.

[JKC+14] WoonSeong Jeong, Jong Bum Kim, Mark J Clayton, Jeff S Haberl, and Wei Yan. Translating building information modeling to building energy modeling using model view definition. *The Scientific World Journal*, 2014, 2014.

[JZKT15] Aras Jaly-Zada, Christian Koch, and Walid Tizani. Ifc extensions for design change management. In *Proc., 32nd CIB W78 Conference 2015*, 2015.

Journal of Object Technology, vol. 19, no. 2, 2020
[KC13] A Khalili and DK H Chua. Ifc-based graph data model for topological queries on building elements. *Journal of Computing in Civil Engineering*, 29(3):04014046, 2013.

[KGP05] Thomas H Kolbe, Gerhard Gröger, and Lutz Plümer. Citygml: Interoperable access to 3d city models. In *Geo-information for disaster management*, pages 883–899. Springer, 2005.

[KH14] Torsten Krämer and Wolfgang Huhnt. Meta object facility to access and to integrate building information models. *EG-ICE 2011, European Group for Intelligent Computing in Engineering*, 01 2014.

[KH15] Tae Wook Kang and Chang Hee Hong. A study on software architecture for effective bim/gis-based facility management data integration. *Automation in Construction*, 54:25–38, 2015.

[Kit04] Barbara Kitchenham. Procedures for performing systematic reviews. *Keele, UK, Keele University*, 33(2004):1–26, 2004.

[KLR+95] David Koller, Peter Lindstrom, William Ribarsky, Larry F Hodges, Nick Faust, and Gregory Turner. Virtual gis: A real-time 3d geographic information system. In *Proceedings of the 6th conference on Visualization ’95*, page 94. IEEE Computer Society, 1995.

[Kos92] Lauri Koskela. Application of the new production philosophy to construction. Technical Report TR072, Stanford University, September 1992.

[KWW+10] Peter Katranuschkov, Matthias Weise, Ronny Windisch, Sebastian Fuchs, and Rainmar J Scherer. Bim-based generation of multi-model views. *CIB W78*, 2010.

[LAD+16] Levi Lúcio, Moussa Amrani, Juergen Dingel, Leen Lambers, Rick Salay, Gehan M.K. Selim, Eugene Syriani, and Manuel Wimmer. Model transformation intents and their properties. *Software and Systems Modeling*, 15(3):647–684, 2016.

[LCL+12] Chia-Ying Lin, Chien-Cheng Chou, Ya-Mei Lin, I-Chen Wu, and Hao-Wen Chang. Apply model-driven architecture to re-conceptualization of bim for extended usage. In *EG-ICE Workshop*, 2012.

[LES16] Yong-Cheol Lee, Charles M Eastman, and Wawan Solihin. An ontology-based approach for developing data exchange requirements and model views of building information modeling. *Advanced Engineering Informatics*, 30(3):354–367, 2016.

[LES18] Yong-Cheol Lee, Charles M Eastman, and Wawan Solihin. Logic for ensuring the data exchange integrity of building information models. *Automation in Construction*, 85:249–262, 2018.

[LPH13] Ghang Lee, Young Hyun Park, and Sungil Ham. Extended process to product modeling (xppm) for integrated and seamless idm and mvd development. *Advanced engineering informatics*, 27(4):636–651, 2013.

[MB13] Wiet Mazairac and Jakob Beetz. Bimql—an open query language for building information models. *Advanced Engineering Informatics*, 27(4):444–456, 2013.

[MH05] O. Matthews and G.A. Howell. Integrated project delivery an example of relational contracting. *Lean Construction Journal*, 2(1):46–61, 2005.
Clement Mignard and Christophe Nicolle. Merging bim and gis using ontologies application to urban facility management in active3d. *Computers in Industry*, 65(9):1276–1290, 2014.

Yasuhiro Monden. *Toyota production system: an integrated approach to just-in-time*. Productivity Press, 2011.

Pieter J. Mosterman and Hans Vangheluwe. Computer automated multi-paradigm modeling: An introduction. *SIMULATION*, 80(9):433–450, 2004.

Mohamed Nour. A graphical user interface for handling ifc partial model exchange. In *Proceedings of The 12. International Conference on Computing in Civil and Building Engineering (ICCCBE)*. Citeseer, 2008.

Jyrki Oraskari and Seppo Törmä. Rdf-based signature algorithms for computing differences of ifc models. *Automation in Construction*, 57:213–221, 2015.

Pieter Pauwels and Anna Roxin. Simplebim: From full ifcowl graphs to simplified building graphs. In *Proceedings of the 11th European Conference on Product and Process Modelling (ECPPPM)*, pages 11–18, 2017.

Michael J Pratt. Introduction to iso 10303—the step standard for product data exchange. *Journal of Computing and Information Science in Engineering*, 1(1):102–103, 2001.

Sergio Pinheiro, Reinhard Wimmer, James O Donnell, Sergej Muhic, Vladimir Bazjanac, Tobias Maile, Jérôme Frisch, and Christoph van Treeck. Mvd based information exchange between bim and building energy performance simulation. *Automation in Construction*, 90:91–103, 2018.

Paola Sanguinetti, Sherif Abdelmohsen, JaeMin Lee, JinKook Lee, Hugo Sheward, and Chuck Eastman. General system architecture for bim: An integrated approach for design and analysis. *Advanced Engineering Informatics*, 26(2):317–333, 2012.

Jim Steel, Keith Duddy, and Robin Drogemuller. A transformation workbench for building information models. In *International Conference on Theory and Practice of Model Transformations*, pages 93–107. Springer, 2011.

René Schumann and Jan Tulke. *BIM at HOCHTIEF Solutions*. Springer International Publishing, 09 2018.

Bilal Succar. Building information modelling framework: A research and delivery foundation for industry stakeholders. *Automation in Construction*, 18(3):357 – 375, 2009.

Thomas Stahl, Markus Völter, and Krzysztof Czarnecki. *Model-Driven Software Development: Technology, Engineering, Management*. John Wiley and Sons, 2006.

Sebastian Stratbücker, Sicheng Zhu, and Matthias Mitterhofer. Model driven engineering methods for integrated building performance optimization. In *The 15th International Conference of IBPSA Conference*.
of International Building Performance Simulation Association, pages 2683–2692. International Building Performance Simulation Association, 2017.

[TBS16] Eike Tauscher, Hans-Joachim Bargstädt, and Kay Smarsly. Generic bim queries based on the ifc object model using graph theory. In The 16th International Conference on Computing in Civil and Building Engineering. Osaka, Japan, 2016.

[VAW+17] Simon Vilgertshofer, Julian Amann, Bruno Willenborg, Andè Borrmann, and Thomas H Kolbe. Linking bim and gis models in infrastructure by example of ifc and citygml. In Computing in Civil Engineering 2017, pages 133–140. ASCE, 2017.

[VDLM02] Hans Vangheluwe, Juan De Lara, and Pieter J Mosterman. An introduction to multi-paradigm modelling and simulation. In Proceedings of the AIS’2002 conference (AI, Simulation and Planning in High Autonomy Systems), Lisboa, Portugal, pages 9–20, 2002.

[vNT92] G.A. van Nederveen and F.P. Tolman. Modelling multiple views on buildings. Automation in Construction, 1(3):215 – 224, 1992.

[WKS03] Matthias Weise, Peter Katranuschkov, and Raimar J Scherer. Generalised model subset definition schema. CIB report, 284:440, 2003.

[XDLM14] Xun Xu, Liyuan Ding, Hanbin Luo, and Ling Ma. From building information modeling to city information modeling. Journal of Information Technology in Construction (ITCon), 19:292–307, 2014.

[ZBW15] Chi Zhang, Jakob Beetz, and Matthias Weise. Interoperable validation for ifc building models using open standards. Journal of Information Technology in Construction, 20(ECPPM-2014):24, 2015.

About the authors

Sebastian Götz is a postdoctoral researcher at the Chair of Software Technology. Contact him at sebastian.goetz1@tu-dresden.de

Andreas Fehn is a research assistant at the Chair of Software Technology. Contact him at andreas.fehn@tu-dresden.de

Frank Rohde is a research assistant at the Chair of Software Technology. Contact him at frank.rohde@tu-dresden.de

Thomas Kühn is a postdoctoral researcher at the Chair for Software Design and Quality. Contact him at thomas.kuehn@kit.edu

Acknowledgments This work has been funded by the German Research Foundation within the Collaborative Research Center 912 Highly Adaptive Energy-Efficient Computing, within the Research Training Group Role-based Software Infrastructures for continuous-context-sensitive Systems (GRK 1907) and by the German Federal Ministry of Education and Research within the Project “Integrated Lifecycle Management for Wastewater Treatment” (FKZ: 01IS17007D).