



CCC 2018

Proceedings of the Creative Construction Conference (2018)
Edited by: Miroslaw J. Skibniewski & Miklos Hajdu
DOI 10.3311/CCC2018-047

Creative Construction Conference 2018, CCC 2018, 30 June - 3 July 2018, Ljubljana, Slovenia

Determination of a classification tool for IFC data models based on a predefined classification system

Robert Susset^a, Borja García de Soto^{b,c,*}, Claus Maier^d

^aSteiner AG, Zurich, 8052, Switzerland

^bNew York University Abu Dhabi (NYUAD), Saadiyat Island, Experimental Research Building, PO Box 129188, Abu Dhabi, UAE

^cNew York University (NYU), Tandon School of Engineering, 6 MetroTech Center, Brooklyn, NY 11201, USA

^dEBP, Zurich, 8032, Switzerland

Abstract

Open BIM models are generally based on the ISO-generic IFC (Industry Foundation Classes) data model. These data models only allow limited interpretations and evaluations concerning specific guidelines or standards. This paper presents a classification process for IFC data models, which allows adapting them according to country-specific classification systems. Ensuring that the different elements in a data model are adequately classified is particularly important for cost estimation during the early stages of a project. The proposed process allows assigning the proper classification even when the quality of the model is limited. The main steps include: (1) Analysis and understanding of the IFC structure, as well as a country-specific classification system, (2) creation of a classification tool for a country-specific cost structure using an IFC model, (3) creation of rules for quality assurance as well as the development of quantities according to a country-specific guideline for the development of preliminary cost estimates. The classification process was tested using IFC data models from construction projects to fit the classification standard for Construction Cost Plan for Buildings (Baukostenplan Hochbau (eBKP-H) in Switzerland). In particular, the classification was made for the “Structural Work” group of the eBKP-H using Solibri Model Checker and MS Excel. Results show that the proposed classification process was able to improve the allocation of the different elements to their correct category, reduce the number of unclassified elements, and ultimately improve the quality of the information used to do cost estimates.

© 2018 The Authors. Published by Diamond Congress Ltd.

Peer-review under responsibility of the scientific committee of the Creative Construction Conference 2018.

Keywords: Classification process; IFC; Classification tool, Building Information Modeling, Solibri Model Checker

1. Introduction

In the late 1980s, scientists from the construction industry began to ask themselves about the possibility of a higher data exchange model in the construction industry. This idea came based on what already happened in another field a few decades before. A data model that could not just be limited to its geometric description but includes semantics. Such models first appeared in the car industry in the 1970s and were initiated by the US Ministry of defense and the German car industry [1]. To develop this concept, the Industry Alliance for Interoperability (IAI) was founded in 1994 on the idea of improving interoperability of the construction industry with informatics. It was an initiative from Autodesk and was composed of twelve US companies. It was shortly after renamed International Alliance for Interoperability as it opened its door to any companies around the world. This industry consortium aimed to create an open and free standard, to improve interoperability in construction. This standard would not be based on any specific software and be free of accessibility. From this idea, the Industry Foundation Classes (IFC) was born. In 2005, the IAI was renamed BuildingSMART, and they continue supporting the development of IFC.

IFC is an open international standard for Building Information Model (BIM) data that are exchanged and shared among software applications used by the various participants in the construction or facility management sector. The

standard includes definitions that cover data required for buildings over their life cycle. The first release of IFC came out in 1997 as IFC 1.0 and was then updated through the years, with the second addendum of IFC4, released in July 2016 [2]. Open BIM models are generally based on the ISO-generic IFC (Industry Foundation Classes) data model. Since IFC4 it is covered under ISO 16739:2013. It will soon be replaced by ISO/PRF 16739-1 (currently under development), which in addition to using the EXPRESS data specification language (a compact and well-suited data definition language that includes data validation rules within the data specification), it will include the XML Schema definition language (XSD). This format enables interoperability with XML tools, and it is suitable for exchanging partial building models; however, it is not very much used in practice because it creates large files for building models. In addition, the release of the new ISO will extend the scope to include data definitions for infrastructure assets over their life cycle as well [3].

1.1. Classification system

The eBKP-H (*Baukostenplan Hochbau*), German for cost plan for building construction (SN 506 511) [4] is a Swiss standard for classification of building elements. It came into effect in May 2009. Based on the experience made in the first years of utilization this cost plan was updated in 2012. Since then, it is commonly used in the Swiss construction industry. To put it in an international context, the eBKP-H is equivalent to the Construction Specifications Institute (CSI) and Construction Specifications Canada's (CSC) UniFormat, mostly used in North America to present cost estimates during the schematic design phase [5], but it addresses specific needs of the Swiss industry.

The Swiss Research Centre for Rationalization in Building and Civil Engineering (CRB) developed this standard based on the market needs and with the idea to elaborate a code based on the entire lifecycle of a project as described in the Swiss regulation SIA 112-Service Model [6] issued by the Swiss Society of Engineers and Architects (SIA, by its initials in German). The eBKP-H aims to cope with the ever-growing pressure of the construction industry concerning costs and requirements in increasingly complex projects. The classification of this standard allows construction companies and design firms to improve the quality of their cost estimates, getting them closer to actual construction costs. In addition, the precise definition of normalized terms, costs, and quantity units increases the transparency and the efficiency in the costs estimation.

According to [4], the eBKP-H has the following key functions:

- it can be used for classification of different building types
- it is a frequent basis for designers, contractors, and planners to present costs in a uniform manner
- it is independent of the function, size, complexity, structure, and construction of a building but also from the project phase
- it can be used for new constructions but also renovations or modification of existing buildings.

Although it can be used in later phases of a project (detailed design, tender, and construction), the eBKP-H is typically used during the conceptual phases. For later phases, the NPK (Normenpositionskatalog) [7], German for catalog of standard position, is preferred because it provides information that is more detailed (equivalent to the CSI and CSC's MasterFormat).

The eBKP-H is broken down into 14 groups, from Land to Value added tax (Table 1) and it consists of the following levels:

- Main Group: The main group contains different groups of elements. They are mainly used in strategic planning.
- Group of Elements: This level gathers different isolated elements together and is generally used in the preliminary study phase of a project. It can also be used in the previous phases if a differentiated elaboration is required.
- Element: The elements are defined according to their physical and functional characteristics. This level is mainly used in the project study phase but can also be used in the earlier phases.
- Sub-element: The fourth level is only used in the group of elements "H4 hospital installation", where a more detailed level is needed.

Table 1: Main groups included in the eBKP-H [4]

Main Groups			
A	–	Land	H – User-specific installations
B	–	Preparation work	I – Building surrounding areas
C	–	Structural work	J – Furnishing / Equipment / Decorations
D	–	Technical work	V – Planning cost
E	–	Outside walls / Cladding	W – Additional implementation cost
F	–	Roofing	Y – Provisions
G	–	Interior finishing	Z – Value Added Tax

The eBKP-H can also be used with a different degree of exactitude depending on the requirements of the SIA during the different phases of a project. It is organized as a hierarchy with up to four different levels. There are costs and normed references, which are assigned to every level. Each level allows assessing standard value for costs establishment, control and management at different stages of a project according to the SIA 112 phases [6]. The grade of detail increases with the level of the classification (from the main group to the sub-element). According to the required accuracy of a cost estimate, a different level of the eBKP-H classification should be used. Table 2 shows an example of the decomposition of group C (Structural Work) into its group elements and elements.

Table 2: Example of classification for structural work (main group) and corresponding group elements and elements [4]

Main Group	Group of Elements	Elements
C - Structural work	C1 - Foundation	C1.1 - Pipes and underground utilities
		C1.2 - Sealing and isolation
		C1.3 - Continuous footing
		C1.4 - Floor slab (nonbearing)
		C1.5 - Floor slab (bearing)
	C2 - Bearing wall	C2.1 - Wall (exterior)
		C2.2 - Wall (interior)
	C3 - Column	C3.1 - Columns (exterior)
		C3.2 - Columns (interior)
	C4 - Floor & roof	C4.1 - Slab
		C4.2 - Stairs, Ramp
		C4.3 - Balcony
		C4.4 - Roofing
	C5 - Other	C5.1 -
		C5.2 -

Every level has a specific code standard. It begins with the “Main Group” defined by a letter, then the Groups of Elements, Elements, and Sub-elements are named with numbers and separated by dots (there is no separation between the Main group’s letter and the first number). For example, *H4.3.2 Equipment for surgery* belongs to the sub-element level or *C2.2 Walls Inside* to the element level.

The use of rules applied to IFC data has been investigated (e.g., applications to construction safety, [8]). Current open BIM models only allow limited interpretations and evaluations concerning specific guidelines or standards. The goal of this work is to create a classification tool for IFC data model that will identify and classify building elements according to specific cost structure. For this study, the cost structure used was the one from Switzerland, the eBKP-H classification (*Baukostenplan Hochbau* – German for cost plan for building construction) [4]. Although there is a lot of work done regarding the estimate during the early phases of the quantities required for construction projects ([9], [10]), the extensive use of BIM allows quantifying the materials required quickly and reliably to develop cost estimates from an IFC file. To be able to create an accurate cost estimate, the different elements have to be properly classified. To create a reliable classification tool, the work will begin with a good understanding of the IFC data model and the eBKP-H classification. This will allow to see what are the possibilities with an IFC data model and then classify the elements correctly according to the chosen classification. The creation of the tool itself will be made on the software

Solibri Model Checker [11] which is mainly a model checking software, but it also has a good classification capacity. This tool will also be based on some *Excel* tables.

To verify the usability of the model before classifying it and, at the end of the process, to ensure the accuracy of the classification, quality assurance tests will be done.

2. Classification process

The proposed classification process is divided into four different parts (Figure 1). Each of these parts (i.e., processes) consist of the sub-processes described in the following paragraphs.

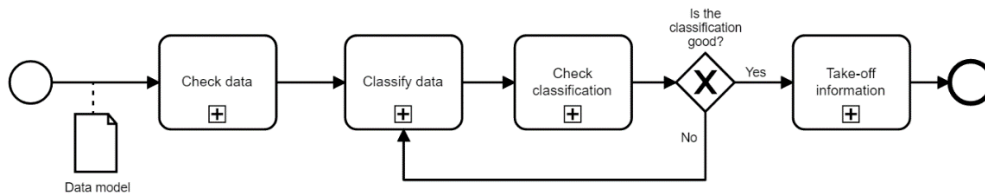


Figure 1: General classification process

2.1. Check data

The first step is to check the quality and accuracy of the IFC data model. Figure 2 shows the process for data verification. Once the format of the data is known, it is necessary to choose an appropriate software that will allow the reading and checking of data effectively. The checking rules must then be defined. It is essential that those rules are as accurate as possible to avoid unforeseen problems. Some points that can be verified are, for example, the consistency in the naming of the objects and the completeness of the model. It is also important to do a clash detection to be sure that there are no duplicated objects and objects overlapping with each other, which would lead to an overestimation of the quantities. It is useful at the end of the quality check to manually verify it through some random and independent checks. If the data model is considered good, this process is over. Otherwise, the model should be sent back to the person who created it with a list of issues to be addressed before the model can be used.

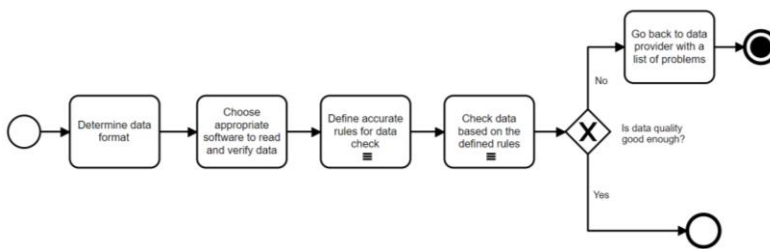


Figure 2: Sub-process for Check data

2.2. Classify data

Once the data is checked and it is verified that the model is usable, the data model can be classified according to a pre-defined classification. The first thing to do to classify data is to determine some rules to sort the data's elements accordingly. Those will mainly depend on the chosen classification (e.g., according to the eBKP classification or the CSI UniFormat). Figure 3 (a) shows the steps of the classification process.

The detailed process of the elements' classification is essentially done by the software itself. However, it is essential to fully understand how the classification works, as there will be some unclassified elements. For this reason, the classification of elements is also considered in the proposed classification process (Figure 3 (b)).

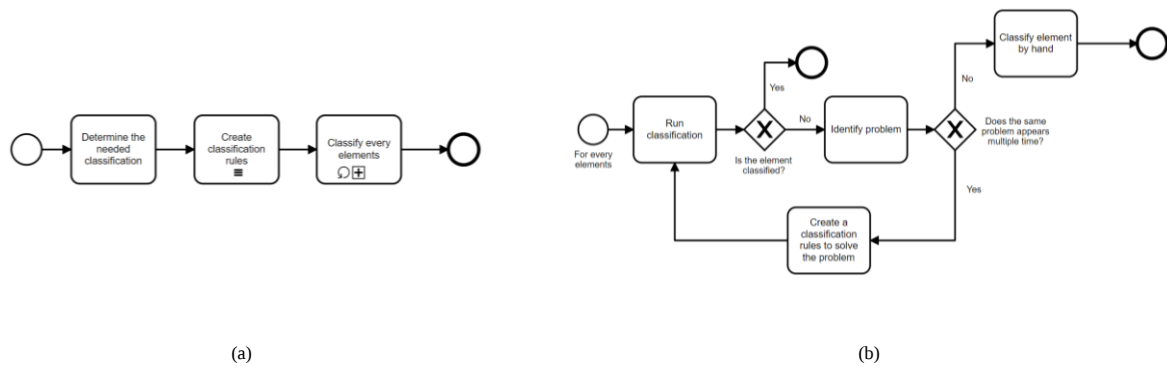


Figure 3: Second sub-process of the general classification process: Classify data (a) and sub-process: classify every element (b)

Every element will first go through the previously defined classification rules. If the element is classified, the process is done. If it is not the case, there are two possibilities. If the same problem occurs several times a new classification rule must be defined, otherwise the element has to be classified individually by hand. The probability of having the same kind or type of object not allocated in a class can easily be checked after the first classification is done by checking all the unclassified elements. It is usually easy to verify if there is a recurrence. This process has to be repeated as long as all elements belonging to a category are not classified or have been classified in the wrong category.

2.3. Check classification

When all the elements are classified, a check on the quality of the classification should be made. This step could seem redundant as the elements are sorted according to rules that were defined for the classification, but it is an important part of the process as some problems can still exist. Figure 4 (a) shows the checking process.

The first thing to do is to determine some quality assurance rules. These rules will depend on the data which are classified and the expected quality of the classification. For example, for a quite large classification with few categories during the early phases of a construction project where everything is not yet precisely defined, high precision is not needed. On the other hand, when a model is defined in detail, the classification needs to be accurate to be useful.



Figure 4: The two last sub-processes: (a) Check classification and (b) Take-off information

2.4. Take-off information

Finally, when the classification is done and the quality controlled. The output of the process needs to be determined and well presented. Figure 4 (b) shows the process to follow in order to obtain useful information. The first thing to do is to determine which quantities are needed. When they are determined, based on the software used, there can be different ways to get them. A convenient way to show data is to use Excel. If the same take-off is done multiple time, a template can be prepared to present the data.

3. Development of the Solibri/Excel tool

Based on the classification process described in the previous section, a classification tool has been developed. In this section, the specific characteristics of the eBKP-H classification and the IFC data model are used.

3.1. Check data

The data model format is IFC, and the software used Solibri. The third step, Define accurate rules for data check, is certainly the most important part of this sub-process as the rules to check the model are chosen here. The first sets of rules are based on existing rules available in Solibri classified as Structural rules. The two existing ruleset are named: “BIM Validation – Structural” and “Intersections between Structural Components.”

Those rules were chosen based on the existing rulesets in Solibri for structural model verification, and the ones that seemed useful were selected. It is always possible to add some rules based on the need of a model or a classification. The new rules can be taken from the existing ones, or new ones can be created.

After the rules are chosen, the model is checked (Checking layout of Solibri). The outcome of this process is a list of issues. This list shows all the elements that didn't follow one or more of the defined rules. There are different possible ways to deal with problems (elements not following the chosen rulesets). If the problem is of minor importance or if it does not influence the result of the classification, it can be ignored. If it does have an impact, the responsible person for this problem should be contacted and asked to make the necessary changes before the process can go further.

3.2. Classify data

The classification implemented in this study is the eBKP-H Swiss standard. For illustration purposes, only the main group *C – Structural work* was used. The rules created to classify the elements are based on this classification but also on the possibilities that the software Solibri offers. The rules are determined as follow. The first thing to do is to classify the elements according to their component, if they are defined as a wall, a slab, etc. An Excel tool is created to facilitate the creation of rules. A table is exported from the Solibri Rules, and thus new rules can easily be implemented in the excel file. Once all the rules are created, the excel table can be imported in the rules definition. To facilitate the definition of new rules, the column where the topic needs to be chosen from a list (*Classification name, Component* and every true/false property) must be filled by choosing from a defined list of possibilities. This needs to be done again to define the rules for the second level of the eBKP-H.

3.3. Check classification

When all the elements are classified, the model needs to be rechecked to make sure that the classification is good enough. This second verification follows the same process as the first checking process, except that the rules for this one are the ones for the verification of the classification. The rules for this verification are a lot more difficult to determine as it is harder to define quality sets for a classification. In addition, the aim of Solibri is not to check the quality of a classification. The first rule used ensures that there are elements in the classes (i.e., a rule will go through all the classes and make sure that at least one of them has an element on it). Another rule checks that the elements classified in the second level classification are also stored in the first level and that the group of elements is the same. Other rules could be defined and written in the rule manager window of Solibri and implemented as needed.

3.4. Take-off information

Finally, when the classification is done and verified, the information needs to be usable and easy to understand. Solibri already has a few built-in Excel templates for information take-off. One of them (Structural quantities) is used and further developed to export and present the information needed. Additional columns with the classes of every element are added to allow the calculation in Excel based on those rules. Any information, for example, volume or bottom area, can also be added to the exported information.

4. Implementation

The classification tool described in the previous section was implemented on an ongoing project consisting of a four-story office building located in an urban area with a plot size of 435 m² and a total building volume of approximately 11,000 m³.

4.1. Solibri Model Checker

The software used in this thesis is Solibri Model Checker or Solibri. It is mainly developed as a model checker that allows the visualization and combination of models from a different domain and to check the model based on predefined rules. Further, it provides a wealth of information that can be taken off (Solibri) [11].

Both aspects will be used even if the predefined rules are only utilized to verify the quality of a model or a classification, and not to combine different models. The classification tool is critical and is the main tool that will be used.

4.2. Application to a model

To evaluate the process, an example of implementation is described and will be used to discuss the usability and the quality of the proposed classification process. This application is made using the software Solibri, and the IFC data model will be classified according to the main group C (Structural work) of the eBKP-H.

The verification of the data is done using two existing rulesets in Solibri: The two existing rulesets are named: “BIM Validation – Structural” and “Intersections between Structural Components.” The first one checks the validity of the model as a structural file. The second one verifies the intersections between overlapping components. Once the verification is done, Solibri output a list of problems with different levels of severity that need to be solved before going further into the process. It is important to know that Solibri does not allow modification or editing the model. Therefore, when errors are found, the model needs to be sent back to the person responsible for corrections before any classification can be done.

Now that the classification is determined, the classification rules can be defined (second step of the data classification process, Figure 3 (a)). These rules are based on the eBKP-H but also on the classifications’ possibilities implemented in Solibri.

4.3. Result and validation

A precise verification of the quality of the classification tool is difficult to perform as it is difficult to compare it with another classification. However, to evaluate the effectiveness of the developed process, a comparison between the first classification (original IFC data model) and the improved classification at the end of the process will allow to see the improvements, as well as what the verification and problem-solving process can provide. Table 3 shows the results of the first and final classification for the structural elements according to the eBKP-H level 1.

Table 3: Amount (volume) of concrete for eBKP-H elements from original IFC data model and after implementing the classification process

		Original	After classification
Main Group	Group of Elements	Volume (m ³)	Volume (m ³)
C - Struct. work	C1 - Foundation	-	1,120.66
	C2 - Bearing wall	1,563.53	1,563.53
	C3 - Column	565.00	22.54
	C4 - Floor & roof	3,448.16	2,358.31
	C5 - Other	-	-
	Unclassified	424.83	395.19
	eBKP-T		541.29

From this comparison, the main difference is for class C1. In the first classification, the class C1 has no quantities, while in the final classification volume and area are listed, which is a great improvement. Almost all the elements that went into the *C1 Foundation* class were before classified in the *C4 Bearing floors and roofing* class. Indeed, all the slabs of the lower level were first classified as normal slabs, when in fact they should have been classified as part of the foundation (i.e., slab on grade). The second class *C2 Bearing walls* remains the same; the walls are classified correctly from the beginning. The volume of the class *C3 Column* decreases in the final classification. The reason is that all the foundation pillars had been classified as columns; however, they should not be stored in this category or any other class of this catalog (including foundation), as they belong to the eBKP-T catalog [12] (cost plan for underground construction). Therefore, another category is created (eBKP-T) so that they are not assigned to the wrong category or accounted as unclassified. Going through the whole process of classification allowed to have a better classification with less unclassified elements but also an improved division in the classes.

5. Discussion

One of the most important lessons that result from this example is the great importance of having a good model of the project. Indeed, if different types or the wrong components are used to define an element, a simple classification of this element is not possible. The quality of the model influences highly the outcome of the classification and the

time needed to conduct the process. If the elements are not defined correctly in the model, additional rules need to be defined, and the effectiveness of the classification process is compromised.

A good verification of the classification is very important. First with some pre-defined rules that allow to verify the consistency of the classified elements, but also, visual verification of the different classes can easily expose some errors that can be resolved.

The software Solibri Model Checker that was used for the classification was very helpful and allows to classify efficiently but also to check the model and the classification very accurately and quickly. The proposed classification process improves the allocation of different elements to the correct category, reduce the number of unclassified elements by 7%, and since this is used to determine the cost of a project, to ultimately improve the quality of the cost estimates.

6. Conclusion and Outlook

The link between an IFC model and the classification of its elements is very important and could guarantee the development of quick and accurate cost estimations. Indeed, based on an IFC model, an estimated quantity necessary to estimate costs can be found in little time. However, to be able to have it as fast as possible, the quality of the model needs to be sufficient. To be able to draw an optimal IFC model, some knowledge of how the format is built and its properties are essential. These abilities will then allow automated classification of a building model. From a practical perspective, the time savings using the tool developed in this study to provide a robust classification of the different elements to be used as the bases of a cost estimation according to the eBKP are attractive to companies. Within a few minutes (depending on the size of the model) the elements and the important quantities (e.g., volume and area) of a model can be classified and outputted according to the eBKP classification (or another relevant classification system). Based on this and knowing the unit costs, cost estimation can be done. Even if improvements and new rules are needed, the process delivers good results very quickly. The time needed for a reliable estimation mainly depends on the quality of the model and its size. Future work on the classification tool could include other main groups from the eBKP-H and be expanded to other classification systems (e.g., eBKP-T [12] and NPK [7]). Also, the classification process could be improved by having a catalog with all eBKP classes could be included with the IFC specification to show how elements belonging to a specific class have to be drawn. This would help to have the model ready without any adjustment to be classified according to the desired classification.

Acknowledgments

The authors would like to thank Mr. Daniel Riondel and Mr. Zafer Aytekin from CRB for their information regarding the objective and functionality of the eBKP standard. Special thanks are also given to the Solibri Support Team for providing required software for this research.

References

- [1] Borrmann, A., M. König, C. Koch, and J. Beetz (2015). *Building Information Modeling, Technologische Grundlagen und Industrielle Praxis* (Technological basis and industrial practice). Springer Vieweg. (in German)
- [2] BuildingSMART. (n.d.). Summary of IFC Releases. <http://www.buildingsmart-tech.org/specifications/ifc-releases/summary>. Accessed 2017 01-31
- [3] ISO/PRF 16739-1. Industry Foundation Classes (IFC) for data sharing in the construction and facility management industries -- Part 1: Data schema using EXPRESS schema definitions (<https://www.iso.org/standard/70303.html>) Accessed 2017 01-31
- [4] CRB (2012a). eBKP-H, Baukostenplan Hochbau, SN 506511 (cost plan for building construction). Zurich, Switzerland: CRB. (in German)
- [5] CSI (2016). Uniformat classification. <https://www.csiresources.org/practice/standards/uniformat>. Accessed: 2017-06-19.
- [6] SIA (2004). SIA 112 Modell Bauplanung (Service Model). Zurich, Switzerland: Schweizerische Ingenieur und Architektenverein. (in German)
- [7] CRB (2012b). NPK, Normpositionen-Katalog (catalog of standard position). Zurich, Switzerland: CRB. (in German)
- [8] Zhang, S., Teizer, J., Lee, J. K., Eastman, C. M., & Venugopal, M. (2013). Building information modeling (BIM) and safety: Automatic safety checking of construction models and schedules. *Automation in Construction*, 29, 183-195.
- [9] García de Soto, B., Adey, B. T., & Fernando, D. (2014). A process for the development and evaluation of preliminary construction material quantity estimation models using backward elimination regression and neural networks. *Journal of Cost Analysis and Parametrics*, 7(3), 180-218, DOI: <https://doi.org/10.1080/1941658X.2014.984880>
- [10] García de Soto, B., Adey, B. T., & Fernando, D. (2017). A hybrid methodology to estimate construction material quantities at an early project phase. *International Journal of Construction Management*, 17(3), 165-196. DOI: <https://doi.org/10.1080/15623599.2016.1176727>
- [11] Solibri (2017). Solibri model checker. <https://www.solibri.com/products/solibri-model-checker/>. Accessed: 2017 06-19.
- [12] CRB (2010). eBKP-T, Baukostenplan Tiefbau, SN 506512 (cost plan for underground construction). Zurich, Switzerland: CRB. (in German)