Similarity Evaluation of the Predeveloped MVD Concepts for Improving Reusability of Existing MVDs

Moeid Shariatfar, mshari6@lsu.edu
Ph.D. Student, Bert S. Turner Dept. of Construction Management, Louisiana State University, Baton Rouge, Louisiana

Yong-Cheol Lee A.M.ASCE, yclee@lsu.edu
Assistant Professor, Bert S. Turner Dept. of Construction Management, Louisiana State University, Baton Rouge, Louisiana (Corresponding Author)

Abstract
Industry Foundation Classes (IFC) facilitate the information sharing process by representing BIM data in a neutral format that is applicable to heterogeneous BIM software. Model View Definitions (MVD) have been broadly used to define data exchange requirements of diverse disciplines. However, the current MVD development entails the challenge of limited capability of reusing the concepts of the previously developed MVDs. To tackle this knowledge and practical gap, this study developed a concept ranking system that analyzes and identifies the previously developed concepts of the MVD library that are the most compatible with the targeted exchange model. This approach includes a similarity checking feature that estimates a measure of conformity between key entities and attributes in an exchange model, and the previously developed MVDs. The proposed method helps industry professionals and developers obtain the existing concepts including the reusable set of entities, attributes, and relationships to define new model views.

Keywords: Building Information Modeling (BIM), Interoperability, Industry Foundation Classes (IFC), Model View Definitions (MVD), Information Delivery Manuals (IDM)

1 Introduction
Building Information Modeling (BIM) has become an integral part of the Architecture, Engineering, and Construction (AEC) industries, facilitating the entire project phases and enabling to share design and project data among a variety of stakeholders and domain experts. However, the interoperability of BIM data still remains elusive because each BIM authoring tool uses their own formats that hinder a seamless data transfer among the participants. Currently, the dominant approach for sharing design and project data is to utilize sibling software supporting direct BIM data exchange. For establishing BIM data interoperability environment, Industry Foundation Classes (IFC) have been developed as a neutral data format compatible with BIM applications that enables seamless information flow and exchange across a wide range of domains and software platforms. Model View Definitions (MVD) have been proposed as a subset of the IFC schema to specify the data exchange requirements of a specific domain. Business processes and exchange models are defined in the Information Delivery Manual (IDM) in a human readable format to be transformed into the IFC schema. The current process for IDM/MVD mapping and development is considerably burdensome and full of complexities, requiring a great deal of time and cost (Jeon and Lee, 2018). In addition, the need for representing diverse objects and relationships that are iteratively defined throughout the entire project phase has led MVDs to become redundant and inconsistent (Lee, Eastman and Solihin, 2018). To relieve this redundancy, a concept that includes combinations of entities, attributes, and relationships has been proposed. It is designed to be reused in several model views to facilitate the MVD development process and also prevent possible redundant definitions of data exchange requirements.

Diverse MVDs have been developed for establishing BIM data interoperability environment for several domains. However, reusing the MVDs and their concepts for developing a new model view is challenging because of: (1) difficulties in accessing and utilizing the previously developed MVDs; and

(2) lack of a logical method for identifying and retrieving compatible concepts according to defined data exchange requirements of new IDMs. To address the first issue, the authors have proposed a new approach to establishing the MVD library, which accumulates various MVDs based on entities (Lee *et al.*, 2020). This study mainly focuses on addressing the second issue by providing a new framework that facilitates concept mapping between the MVD library and new IDM by adopting a similarity ranking system. This approach helps figure out concepts that have the most similar structure and contents to given concepts of IDM. Organized and structured mapping of predeveloped concepts and new IDM is expected to not only facilitates MVD development, but also alleviate the issue of redundancy, helping reduce human errors and required time and effort.

2 MVD development process and challenges

The process of MVD development includes a transition of data exchange models in IDM to the IFC schema by generating appropriate IFC entities, attributes, relationships, and properties in concepts. BIM data exchange requirements are organized in the IDM table, and in the next step, the contents of the IDM table are transformed into the IFC's entities and attributes. Concept templates are composed of entities and attributes without specific values and are utilized to represent each exchange requirement of IDM. The translated concepts are the concept blocks, which are the components of an MVD. Figure 1 below shows the overall MVD development process. In this process, the primary obstacle is the limited capability to reuse the concept blocks of the previously developed MVDs. The correlation of domains throughout the project lifecycle, made it necessary for the MVD components to be sharable and reusable among different domains. For example, concrete, precast concrete, and steel domains have their unique MVDs and have to refer to each other to share commonly-agreed specifications of BIM data exchange. However, the previous MVDs were not sufficiently considered because of technical issues and complicated consensus processes. These limitations lead to inconsistencies in the structure of the developed MVDs and result in heterogeneity and redundancy of MVD concepts. Thus, it is crucial to establish an environment boosting the reuse of the existing MVDs and their concepts for new MVD development.

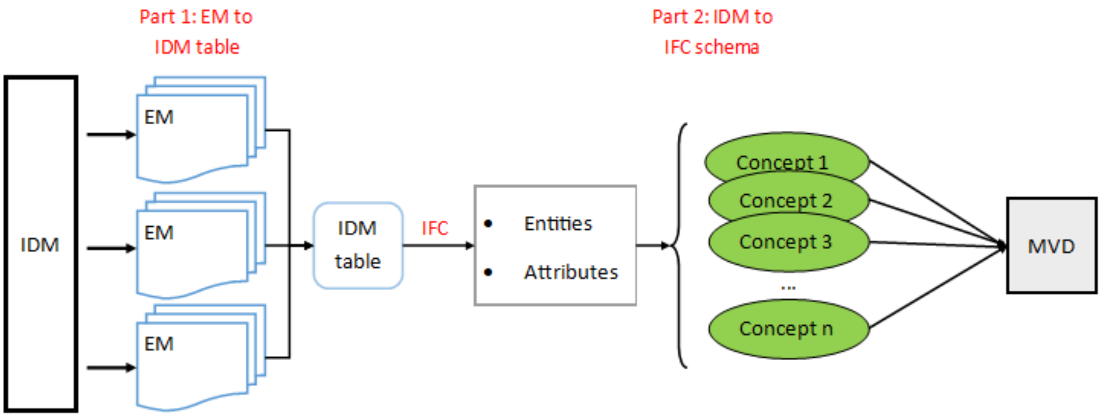

Figure 1: overall MVD development process

3 Literature review

To improve MVD development processes, previous studies proposed diverse new approaches. Lee et al. (2016) provided an ontology-based approach for logically creating model views. They utilized ontological principals for generating IDM to enable an automated transformation to the OWL (Web Ontology Language) format, which allowed a formal linking to MVD concepts through semantic reasoning (Lee, Eastman and Solihin, 2016). This approach can reduce possible semantic mismatch between objects and relationships in exchange models and IFC entities and attributes. However, it still requires a certain level of knowledge in IFC and a manual process in building ontological structures. Extended Process to Product Modeling (xPPM) was introduced by Lee et al. (2013) as a method for enhancing the efficiency of IDM and MVD development and reducing its complexity. First, they reduced the overwhelming number of shapes defined in the ISO 29481 standard for BPMN (Business Process Modeling Notation) to 22 essential shapes. The second step eliminates the separate process of IDM and MVD development by removing the roll of functional parts using the xPPM tool (Lee, Park and Ham, 2013). In addition, Le and Jeong (2017) created a keyword-driven MVD generation method that helps map the key words provided by the users to semantic-equivalent classes and attributes in the LandXML data schema (Le and Jeong, 2017). This method required a database of civil engineering terms that allowed finding the equivalent key words extracted from the exchange models. Even though these previous efforts have improved some aspects of MVD development processes, facilitating the reusability of the existing concepts still remain elusive.

In terms of improving the reusability of the IDM, one proposed framework remodified the existing IDM by dividing it into smaller IDM packages according to the project's WBS (Work Breakdown Structure) (Mondrup *et al.*, 2014). The framework was based upon the OmniClass construction classification system that allowed a hierarchical decomposition of different AEC disciplines. Although this study facilitated the reusability of IDM, the process of transforming IDM to MVD and reusability issue of MVD concepts remains as a challenge. Lee et al. (2020) also created an entity-based integration framework that provided the database of previously developed model views, which allowed developers to search for existing MVDs from a collection of different concepts and to reuse them for developing a new MVD (Lee *et al.*, 2019). The framework adopts XML representation of MVD (mvdXML) for querying entities, attributes, and properties in the concepts stored in the MVD library. This framework has been a baseline for the authors in the proposed study to build a similarity ranking system that can improve the efficiency of concept query and mapping process.

4 Objectives and Methodology

The translation of IDM to the IFC schema requires adequate level of knowledge and expertise in the complex structure and details of IFC. The purpose of this study is to facilitate this process by providing the predeveloped concepts of the MVD library identified by a similarity ranking system. The MVD library provides a framework that can include a collection of concepts of the previously developed MVDs and can provide critical references and supporting materials for generating a new MVD. Since concepts are integral parts of an MVD, facilitating concept creation with existing specifications can lead to facilitating MVD development. The MVD library used in this study currently contains model views of PCI (Precast/Prestressed Concrete Institute), ACI (American Concrete Institute), AISC (American Institute of Steel Construction), General Usage, Reference View, and Design Transfer View.

Exchange concepts are reusable packages of information that are provided as a technical solution for defining BIM data exchange requirements. A group of selected concepts with defined relationships can represent or create an MVD. The objective of the proposed study is to develop a robust data mapping system that can identify the most compatible concepts in the MVD library according to the information of the IDM and help reuse them to create a new MVD. The following two main steps are required to achieve this goal: (1) key entities, attributes, and properties should be extracted from the IDM to be defined in the IFC schema, and (2) the extracted entities, attributes, and properties should be evaluated with ones of the existing concepts by a conformity measurement rate. Figure 2 shows the proposed process including these two steps.

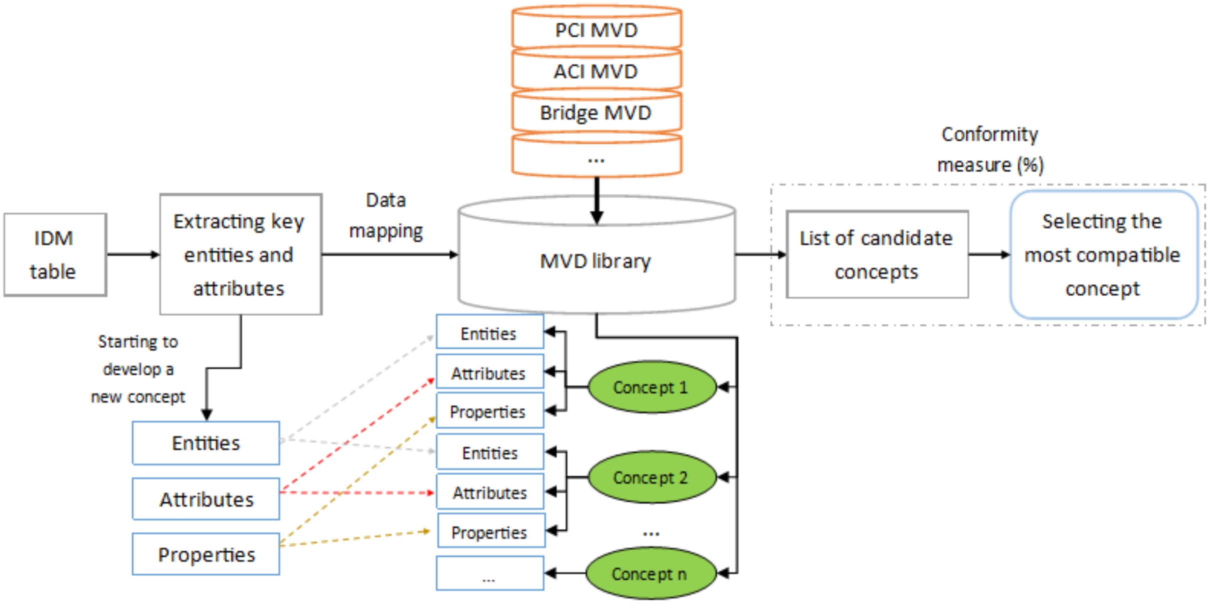

Figure 2: The process of finding the most compatible concept based on IDM data

The IFC schema is comprised of the three root entities and other entities are subsets of these three (buildingSmart, 2021b): (1) IfcObjectDefinition in which the objects are described (2) IfcRelationship that defines relationships between objects, and (3) IfcPropertyDefinition which describes properties associated with entities and others. The IFC schema has an intuitive naming structure allowing objects to be identified simply by their original terms: For example, a wall can be identified in the IFC schema as IfcWall. Other examples are IfcSite, IfcBuilding, IfcBeam, IfcColumn, and so on. The same can be applied for some of entities used to define values and properties such as IfcAreaMeasure that is used to define the area of an object. Other examples are IfcDate, IfcPowerMeasure, or more general entities such as IfcText, IfcLabel, etc. Therefore, key entities and attributes can be extracted from the IDM only through matching the terms and names used in IFC. The first step is extracting IFC entities and attributes from the human readable exchange models presented in the IDM.

One of the main problems in transforming a human readable exchange model into the IFC format is identifying and defining the necessary relationships between the target entity and other objects or entities. The complexities in defining the relationships often lead to inaccuracies in defining correct relations between objects, and even if the relationships are defined correctly, there is still a high chance of redundancy since there are several ways of defining relationships. Being able to investigate the previously developed concepts can help developers to reuse them if they are compatible with the desired use case, and then, create new concepts by slightly modifying the previously developed ones, instead of defining new objects and relationships in the IFC schema from the scratch. To accomplish this task, the proposed study has a new feature that helps identify the concepts that comply with the use case defined in the IDM. As illustrated in Figure 2, the extracted key entities and attributes are mapped into the MVD library and the concepts that contain those entities and attributes are identified. Each concept is comprised of a set of entities and attributes, therefore, the number of matching elements for each queried concept is the main criterion for the concept conformity measure, which is calculated by using Equation 1. Figure 3 shows the usage of this formula in the process of matching entities and attributes in the MVD library.

$$C(\%) = \frac{A_m + E_m}{A_t + E_t} \qquad (1)$$

Where for each queried concept; C = conformity measure in percentage, A_m = Number of matching attributes, A_t = Total number of attributes, E_m = Number of matching entities, E_t = Total number of entities.

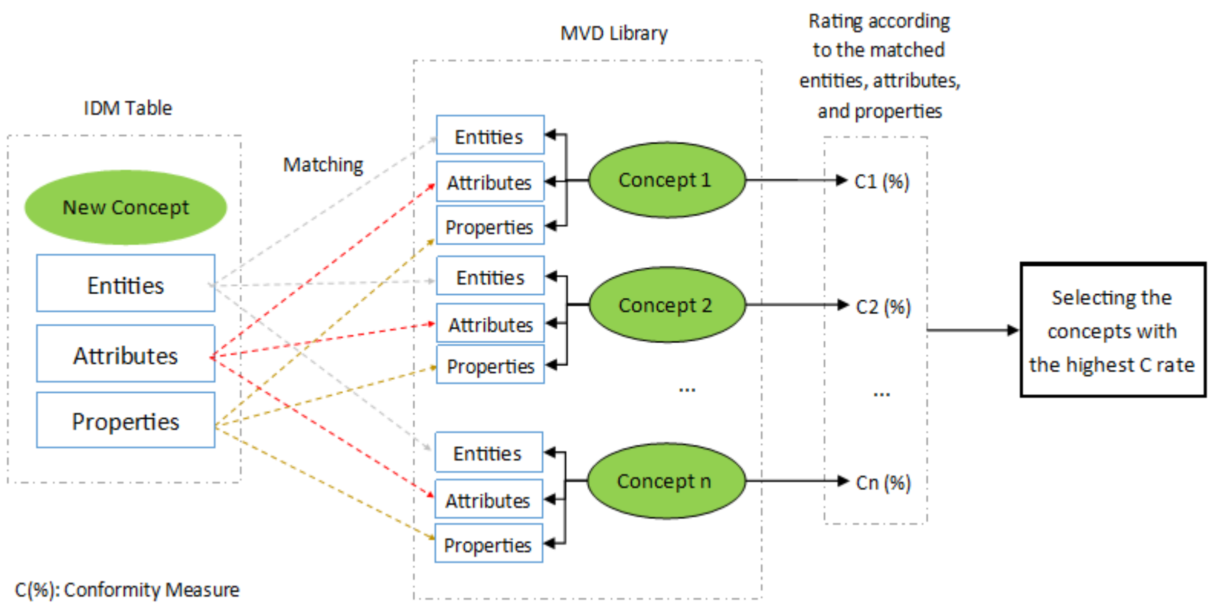

Figure 3: concept matching and rating process

Conformity measure indicates the amount of similarity between the IFC entities extracted from the IDM and each concept in the MVD library. For each concept in the MVD library, a C value is calculated and the concepts that adopt entities and attributes similar to the ones extracted from the IDM, will be selected as candidate concepts. The rate of the conformity measure (C) determines the most similar concept among the candidates and the developer can choose the most compatible concept among these candidates. Model evaluation section illustrates the process through examples of reusing concepts and creating new ones. It must be noted that this paper is focused on identifying the relevant concepts based on the extracted IFC entities and attributes. The process of extracting IFC entities from IDM is not in the scope of this paper and currently is performed manually.

5 Model evaluation

One case study is selected at the outset of evaluating the framework to provide an understandable representation. Further detailed examples are also used to examine the framework with more complicated information of concepts. The first example includes one of the data exchange requirements that defines the building area. The quantities related to IfcBuilding that indicate a floor area should be identified for this requirement as follow: "The quantities relating to the IfcBuilding are defined by the IfcElementQuantity" (buildingSmart, 2021a). According to this statement, the concepts in the MVD library that have IfcElementQuantity in their structure can be a necessary reference.

For defining IfcElementQuantity, we need to determine necessary type of quantity. This can be shown by the attribute "Quantities" presented by the entity "IfcPhysicalQuantity". This entity, IfcPhysicalQuantity, has the subtype of "IfcPhysicalSimpleQuantity", which is the supertype of "IfcQuantityArea". In addition, we need to provide a name for this quantity that indicates the area. This can be defined by IfcLabel with the Name attribute. We also need to define the relation between the object (in our case building) and the property set (the area quantity in our case). As a result, the entity, IfcRelDefinesbyProperties, with the attribute RelatingPropertyDefinition must be selected. In conclusion, the following entities and attributes shown in Table 1 are the mandatory components that should exist in the selected concepts.

Table 1: Key entities and attributes extracted from the example exchange model

No.	Entity	Attribute
1	IfcRelDefinesbyProperties	IsDefinedBy
2	IfcElementQuantity	RelatingPropertyDefinition
3	IfcLabel	Name
4	IfcQuantityArea	Quantities

By searching for these entities and attributes in the MVD library, the related concepts can be identified. The following concepts are the retrieved ones that contain the abovementioned entities in the MVD library. PCI's "Precast Element Quantity Assignment", AISC's "Quantities on Occurrences", General Usage's "Quantity Sets".

The compatibility of each concept will later be investigated by the Conformity Measure formula (C) to allow MVD developers to identify the best and most appropriate concepts for such use case. Figure 4 shows the structure of the concept named "Quantity Sets" that originally belongs to the General Usage model view, in the XML format.

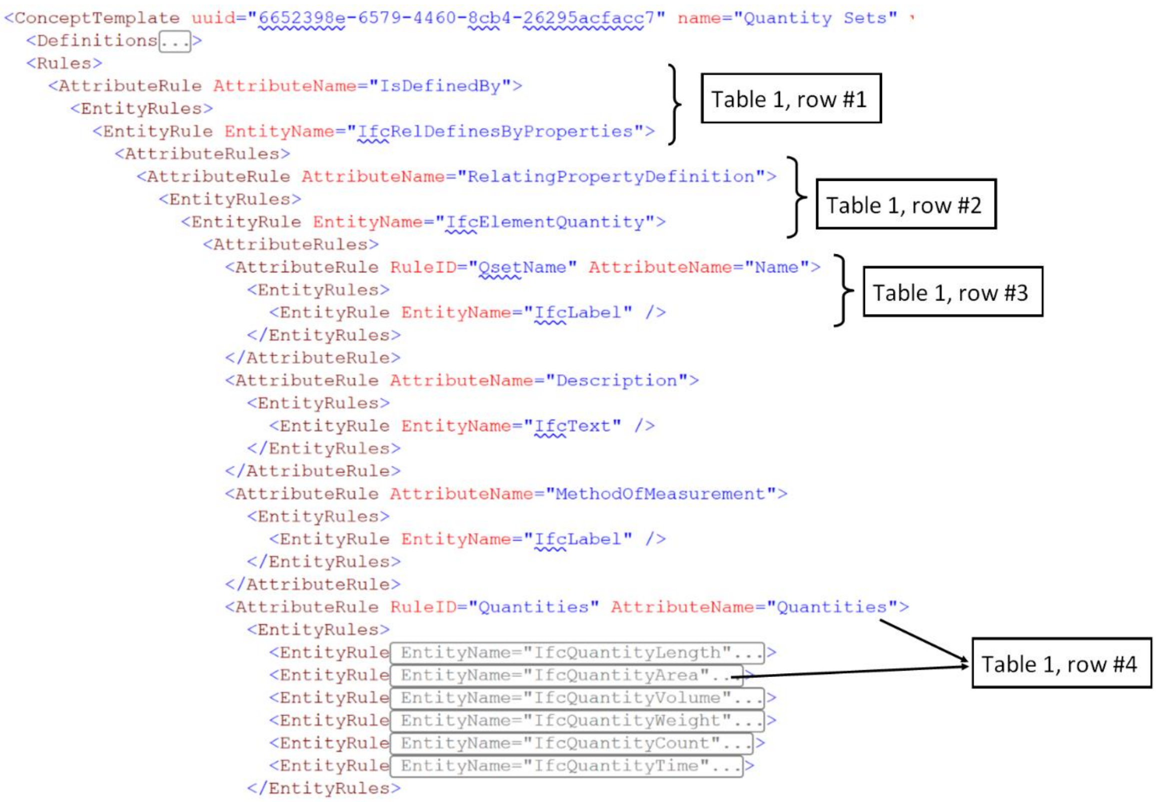

Figure 4: The XML format of the concept "Quantity Sets" and content mapping with Table 1

As evident, all of the target entities and attributes shown in Table 1 exist in this concept. However, the concept has additional entities and attributes that a developer may not need for that specific use case such as IfcText, IfcQuantityLength, etc. Table 2 demonstrates the components of the Quantity Sets concept and identifies the matching entities and attributes.

Table 2: The components of the concept Quantity Sets and the matching entities and attributes

Attribute	Matching Att	Entity	Matching Ent
IsDefinedBy	✓	IfcRelDefinedByProperties	✓
RelatingPropertyDefinition	✓	IfcElementQuantity	✓
Name	✓	IfcLabel	✓
Description		IfcText	
MethodOfMeasurement		IfcLabel	
Quantities	✓	IfcQuantityLength	
		IfcQuantityArea	✓
		IfcQuantityVolume	
		IfcQuantityWeight	
		IfcQuantityCount	
		IfcQuantityTime	

According to the matching entities and attributes identified in Table 2, the Conformity measure (equation 1) of the concepts can be calculated as shown in Table 3. The numbers in Table 3 are attained from the matching and total entities and attributes in Table 2, and also the total entities and attributes of Table 1.

Table 3: The conformity measure for the concept Quantity Sets according to the target entities and attributes

	Total attributes	Matching attributes	Total entities	Matching entities
Matching Concept: "Quantity Sets"	6	4	11	4
New Concept	4	4	4	4
Total	10	8	15	8
Conformity measure		64.00%		

The same process can be conducted for the other two concepts (PCI's "Precast Element Quantity Assignment", and AISC's "Quantities on Occurrences"), and the one with the higher rate of conformity measure can be considered as the best concept. Table 4 illustrates the final results of this process for each candidate concept.

Table 4: The conformity measure for each candidate concept

Candidate Concepts	Conformity measure
Quantity Sets	64.00%
Precast Element Quantity Assignment	30.77%
Quantities on Occurrences	70.00%

Table 4 shows that the concept Quantities on Occurrences is the concept that has the most similar attributes to the example use case. Since this is a comparative score, it does not necessarily mean that the other two concepts are not suitable. For example, the reason that the concept Quantity Sets attained lower rank compared to Quantities on Occurrences, was that the Quantity Sets concepts adopted entities such as IfcQuantityLength, IfcQuantityVolume, etc. along with the required entity IfcQuantityArea as shown in Table 2. These are the subtypes of IfcPhysicalSimpleQuantity that instead are used in the

Quantities on the Occurrences concept. This indicates that the two concepts are not different in nature although their scores are slightly different. This example shows the importance of candidate concepts, which can be an opportunity for developers to examine other close concepts even if they have not acquired the highest score. It also shows that the redundancy and inconsistency issue exists even in the MVD library, which currently contains only six different model views.

Another example includes one concept used to identify a location of an object in relation to other objects such as a location of a building inside a site, or a location of a beam inside the building. This case study also has Cartesian coordinate data and direction of the object required according to the IDM. The process of a concept ranking system was applied similar to the previous example in order to identify the concepts that are most compatible with the contents of the IDM. Table 5 shows the key IFC entities and attributes required for creating concepts in the exchange model of the case study.

Table 5: Key entities and attributes for identifying object's location, coordination, and direction

No.	Entity	Attribute
1	IfcLocalPlacement	ObjectPlacement
2	IfcAxis2Placement3D	RelativePlacement
3	IfcCartesianPoint	Location
4	IfcDirection	RefDirection

By conducting a data mapping process between the entities and attributes in Table 5 and the entities and attributes in the MVD library, the similar concepts can be extracted. The following list include the extracted concepts: PCI's "Building Contained in Site", PCI's "Mechanical Attributes", PCI's "Absolute Placement", PCI's "Relative placement", AISC's "Placement", ACI's "Product Local Placement".

Table 6 was created for AISC's "Placement" concept to identify the resembling entities and attributes that are considered as the criteria for assessing the conformity measure. In other words, Table 6 matches the entities and attributes of the AISC concept "Placement" illustrated in Table 5. Table 7 shows the calculation of the conformity measure that determines the level of similarity of each selected concept with the target concept.

Table 6: Matching entities and attributes for the AISC concept "Placement"

Attribute	Matching Att	Entity	Matching Ent
ObjectPlacement	✓	IfcLocalPlacement	✓
PlacementRelTo		IfcLocalPlacement	
RelativePlacement	✓	IfcAxis2Placement3D	✓
Location	✓	IfcCartesianPoint	✓
Axis		IfcDirection	
RefDirection	✓	IfcDirection	✓

Table 7: The conformity measure for the concept Placement according to the target entities and attributes

	Total attributes	Matching attributes	Total entities	Matching entities
Matching Concept: "Placement"	6	4	6	4
New Concept	4	4	4	4
Total	10	8	10	8
Conformity measure		80.00%		

For each of the remaining concepts (other 5 concepts) tables similar to the tables 6 and 7 should be created. Table 8 is the result of conducting the same process for all of the abovementioned concepts to identify their compatibility with the components of the IDM or the exchange model.

Table 8: Results of calculating conformity measures for each candidate concepts

No.	Candidate Concepts	Conformity measure
1	PCI Building Contained in Site	15.38%
2	PCI Mechanical Attributes	14.29%
3	PCI Absolute Placement	75.00%
4	PCI Relative placement	25.00%
5	AISC Placement	80.00%
6	ACI Product Local Placement	57.14%

The results in Table 8 shows that the concept "Placement" has the highest conformity to the example exchange model compared to other concepts in the MVD library. The case studies show the importance of candidate concepts in the compliance process and show the inconsistency and redundancy issue even in the MVD library. This example also illustrates another benefit of this rating system, which is identifying the possible missing entities and attributes in the transformation process of IDM to IFC. The example shows that those entities and attributes that were not matched in Table 6, are not necessarily useless, despite their incompatibility. They can be used as suggestions of the entities and attributes that were missed during the manual translation process of IDM to IFC. For example, in Table 6, the attribute PlacementRelTo is not matched although it is necessary for defining the relative placements of the object with other references. This means that the developer has missed this entity for defining the relative placement in table 5. However, by assessing the candidate concepts, the developer can recognize such entities and attributes that were possibly missed at the beginning of the process (IDM to IFC translation) and add them to the final concept. This advantage can be useful especially for developers with limited knowledge of the complex structure of IFC schema.

6 Conclusion

The study provides a new framework of a ranking system for measuring the conformity of the concepts with IFC entities and attributes extracted from the IDM. The proposed method allows to match and compare entities, attributes, and properties extracted from the IDM with the entities and attributes in the structure of the concepts collected in the MVD library and to provide a conformity measure for each concept to distinguish the most similar concepts from the less similar ones. The case studies conducted in the model evaluation phase also shows that the conformity measure plays a pivotal role as a reliable score that can detect the concepts relevant to the target IDM. However, concepts with slight differences in conformity measure still need to be examined manually because of the possibility of creating concepts with entities that are sub or super types of the target entities extracted from the IDM. This issue was rectified by maintaining the candidate concepts that allow MVD developers to investigate the detailed structure of the concepts with lower conformity measure scores. The underlying problem also indicates that the inconsistency gap exists even among the six different MVDs stored in the MVD library. Combining additional model views into the MVD library obviously can increase the inconsistencies, and if the MVD development process continues with the current methods, the inconsistencies and redundancies will continue to grow. Facilitating the reuse of the previously developed MVDs significantly affects resolving this problem. The proposed framework is expected to not only facilitate the MVD development but also prevent the redundancy and inconsistency issue that exists in creating IFC model views. In this study, the matching process and conformity measure calculations were conducted manually. The IFC version used in this study was IFC 4. However, MVD library can contain MVDs developed with previous IFC versions which can cause additional inconsistency problems. Although the authors did not encounter such problem in the current MVD library, it is possible that the issue occurs by adding more MVDs. The problem can be rectified by manually modifying or removing the concepts with older versions of entities and attributes. Future studies will be designed and conducted towards an automated concept matching process along with providing suggestions about possible missed entities and attributes in the IDM to IFC transformation process.

Acknowledgements

This publication was made possible by the NPRP grant (NPRP 12S-0304-190230) from the Qatar National Research Fund (a member of Qatar Foundation). The statements made herein are solely the responsibility of the authors.

References

buildingSmart (2021a) *IfcBuilding*. Available at: https://iaiweb.lbl.gov/Resources/IFC_Releases/R2x3_final/ifcproductextension/lexical/ifcbuilding.htm (Accessed: 23 February 2021).

buildingSmart (2021b) *Industry Foundation Classe*. Available at: https://standards.buildingsmart.org/IFC/RELEASE/IFC4/ADD2_TC1/HTML/ (Accessed: 4 April 2021).

Jeon, K. and Lee, G. (2018) 'Information Delivery Manual (IDM) Configurator : Previous Efforts and Future Work Information Delivery Manual (IDM) Configurator : Previous Efforts and Future Work', in *18th International Conference on Construction Applications of Virtual Reality*. Auckland University, New Zealand. Available at: https://www.researchgate.net/publication/332826128_Information_Delivery_Manual_IDM_Configurator_Previous_Efforts_and_Future_Work.

Le, T. and Jeong, D. (2017) 'Keyword-Driven Model View Generation for Civil Infrastructure Projects', *Computing in Civil Engineering*, pp. 254–261. doi: https://doi.org/10.1061/9780784480823.031.

Lee, G., Park, Y. H. and Ham, S. (2013) 'Extended Process to Product Modeling (xPPM) for integrated and seamless IDM and MVD development', *Advanced Engineering Informatics*. Elsevier Ltd, 27(4), pp. 636–651. doi: 10.1016/j.aei.2013.08.004.

Lee, Y.-C. *et al.* (2019) 'Entity-Based MVD Concept Module Generation for Development of New BIM Data Exchange Standards', in *Computing in Civil Engineering 2019*. Atlanta, Georgia: ASCE, pp. 464–472. doi: https://doi.org/10.1061/9780784482421.059.

Lee, Y.-C. *et al.* (2020) 'Generation of Entity-Based Integrated Model View Definition Modules for the Development of New BIM Data Exchange Standards', *Journal of Computing in Civil Engineering*, 34(3), p. 04020011. doi: 10.1061/(asce)cp.1943-5487.0000888.

Lee, Y. C., Eastman, C. M. and Solihin, W. (2016) 'An Ontology-based Approach for Developing Data Exchange Requirements and Model Views of Building Information Modeling', *Advanced Engineering Informatics*, 30(3), pp. 354–367. doi: 10.1016/j.aei.2016.04.008.

Lee, Y. C., Eastman, C. M. and Solihin, W. (2018) 'Erratum: Logic for ensuring the data exchange integrity of building information models (Automation in Construction (2018) 85 (249–262) (S0926580517307276) (10.1016/j.autcon.2017.08.010))', *Automation in Construction*. Elsevier, 93(July), pp. 388–401. doi: 10.1016/j.autcon.2018.06.002.

Mondrup, T. F. *et al.* (2014) 'Introducing a new framework for using generic information delivery manuals', *eWork and eBusiness in Architecture, Engineering and Construction - Proceedings of the 10th European Conference on Product and Process Modelling, ECPPM 2014*, pp. 295–301. doi: 10.1201/b17396-52.