From a meagre 'testing for interoperability' to the 'measurement of interoperability' in BIM

Joseph Jabin[1,*], Johannes Dimyadi[2] and Robert Amor[1]
[1]University of Auckland
[2]Compliance Audit Systems Limited
[*]email: jkan585@aucklanduni.ac.nz

Abstract

Current conformance tests for BIM (Building Information Modelling) data exchange standards are not reliable, as reported in the literature. It is understood that no conformance test can check for 100% error-free operation, as it is a test of the ability to exchange information rather than a measure of the quality of the information exchange. This research has investigated if the current conformance test methodology is in accordance with established measurement theories and therefore can be used to measure the interoperability of a BIM software tool against a data exchange standard. It was found that current conformance tests do not conform to measurement theory and are not able to accurately convey the results and any limitations of the measurement to the end-user. To investigate how other domains have measured interoperability and whether any of their metrics are useful for measuring interoperability in BIM, a systematic literature review (SLR) was conducted. The SLR has gathered 28 interoperability measurement models. They were analysed in terms of established principles in the measurement theory and from that current deficiencies in the interoperability measurement for BIM were identified. Suggestions are given on upgrading the current conformance tests to be compliant with the measurement theory so that they will convey more accurate information to the end-users.

Keywords: BIM, Conformance test, Certification, Interoperability, Measurement theory

1. Introduction

Conformance tests (or compliance tests in some literature) for BIM (Building Information Modelling) are conducted by mainstream organisations like buildingSMART, Green Building Council, US National Institute of Building Sciences, etc. to determine how a BIM software tool interoperates with data exchange standards like IFC, gbXML, COBie, etc. These tests are used to certify the ability of various BIM software tools to interoperate with data exchange standards. This gives the perception that a conformance test is a de-facto process for measuring interoperability. The tests check if a native representation of BIM data in a BIM software tool can map to various objects and attributes in a data exchange standard. The result of the conformance test is a list of mapped attributes along with information on whether each attribute was successfully mapped from the native representation to that of the data exchange standard and vice versa (buildingSMART, 2019). Upon successful completion of the test, the software is awarded a certification, and the software vendor can display the appropriate logo (e.g., IFC2x3 Certified) on their promotional materials (buildingSMART, 2010). Researchers have pointed out that the certification process is not reliable as the certified software tools still have interoperability issues with the data exchange standard (Amor & Dimyadi, 2010; Kiviniemi, 2008; Lipman, Palmer, & Palacios, 2010). There are arguments that no certification scheme can guarantee completely error-free operations (Steinmann, 2018), and the certification process is a test of the ability of a software product to exchange information with the data exchange standard rather than the quality of the exchange (Lipman et al., 2010). These limitations are not conveyed through the certification logos. At the same time, the inherent data exchange issues are not aligned with end-user expectations that a certified software tool will exchange data with 100% accuracy against a data exchange standard (Lipman et al., 2010). The perceived conflict phenomenon between the certification body and the end-users is well known in the realm of measurement theory. According to Fenton and Pfleeger (1997, p. 14), *“No matter how measurements are used, it is important to manage the expectations of those who will make measurement-based decisions. Users of the data should always be aware of the limited accuracy of prediction and of the margin of error in the measurements”*. Therefore, a fundamental problem with the current conformance tests and certification process is the failure to accurately convey the capability of a software tool to interoperate with the data exchange standard through to the end-users.

More insights on the problem can be gained by analysing the conformance test methodologies with respect to the measurement theory. According to the representational theory of measurement, measurement is defined as *“a mapping from the empirical world to the formal, relational world”* (Fenton & Pfleeger, 1997, p. 28). There are two types of measurements, direct measurements and indirect measurements. Direct measurements are made when an attribute of an entity can be directly mapped into a measure (e.g. measuring the length of an object). Indirect measurements are made when a measure can only be derived from more than one attribute of the entity (e.g. the density can only be measured indirectly by dividing mass by volume of the object). An indirect measure needs a ‘model’ to represent the relation between the attributes measured. In the above example [density=mass/volume] is the model. A model is an abstraction of reality that can strip away the details and allows one to view an entity from a particular perspective (Fenton & Pfleeger, 1997). As per the representational theory of measurement, an indirect measurement process should have three components: 1) a process for the actual measurement of the attributes; 2) a model to define the mapping of the measured attributes; and 3) a measure derived from the model represented in a particular scale. Representing the mapping in a scale enables one to compare and manipulate the various measurements made.

In the case of conformance tests in the BIM domain, the needed measure is the ‘interoperability’ between a software tool and the data exchange standard. A BIM interoperability measure can be defined as the degree of accuracy in preserving the syntactic and semantic correctness of the BIM model when data is exchanged between a BIM software tool and a data exchange standard. Interoperability cannot be measured directly from one single attribute. Hence, it is an indirect measurement, needing all three components recommended by the measurement theory. Current conformance tests only have the first component, i.e. the actual measurement of the attributes. They lack a model that defines interoperability and lack a scale which allows one to convey the level of the measured interoperability. Therefore, the current conformance testing methodologies in the BIM domain do not fully comply with the measurement theory, and they do not qualify as interoperability measurement models. This is one reason for the failure of current approaches to accurately convey the interoperability of BIM software tools.

It is worthwhile to understand how interoperability is measured and expressed in other domains to investigate how closely they conform to the measurement theory and whether any of their approaches are useful for BIM. This study sets out to give insights into transforming the current conformance tests from a meagre ‘testing for interoperability’ to the ‘measurement of interoperability’ in BIM. Initially, a systematic literature review (SLR) on interoperability measurement was conducted. This SLR identifies and analyses existing interoperability measurement methodologies from all domains. The review has identified 28 interoperability measurement models that are from the military, government agencies, and enterprise sectors, though none was found in the AEC (Architecture, Engineering, and Construction) domain. Although these models are from different domains, the definition of interoperability given by each of the models is very similar to that used in the BIM domain. Hence, these models can provide significant insights on adapting or developing a new, interoperability measurement methodology for the BIM domain. Section 2 of the paper explains the SLR methodology and lists out the results. Section 3 analyses and categorises the 28 interoperability measurement models. Section 4 discusses the possibility of adapting one, or more, of these approaches to enhance or complement the current conformance testing methods in the AEC domain with the ultimate objective of developing a true interoperability measurement model for BIM.

2. Systematic literature review methodology

The SLR methodology was formulated as per the guidelines of Kitchenham (2004). The research questions for the SLR were derived to address three viewpoints: population (systems of entities where interoperability issues exist); intervention (interoperability measurement models); and outcome (how these models can quantify interoperability) (Kitchenham, 2004). Three research questions were derived from the above-mentioned viewpoints, where RQ1 combines the population and intervention viewpoints and RQ2 and RQ3 are related to the outcome. The questions are as follows:

RQ1 What measurement models are currently available for interoperability measurement in general?

RQ2 How do existing interoperability measurement models quantify interoperability?

RQ3 Are there any frameworks or methodologies aligned towards interoperability measurement in the BIM or AEC/FM sector. If yes, how are they aligned?

A scoping study and search were conducted on the topic of 'interoperability measurement'. A scoping study allows one to rapidly map the key concepts and the main sources of the proposed research areas (Arksey & O'Malley, 2005). Four review papers were identified, and a snowball search was conducted on those review papers. Fifteen papers which proposed an interoperability model were gathered as the result of the snowball search. Then the databases where those fifteen papers were published were selected as the SLR search databases. The selected databases are: 1) Defence Technical Information Center; 2) IEEE Xplore; 2) Wiley online library; 3) Gridwiseac.org; 5) Springer link; and 6) Science Direct. The keywords for the main search were derived from the tiles and abstracts of those fifteen papers. The keywords were: Interoperability; Framework; Measurement; Spectrum; Quantify; Model; Assess; Metric; Federate; Evaluate; Maturity; and Standard. The search terms were constructed and modified with appropriate wild card characters as per the requirements for each database. For example, the search string for IEEE Xplore was: (("Document Title":interop*) AND ("Abstract":interop*) AND ("Abstract":Framework OR "Abstract":Measure* OR "Abstract":Spectrum OR "Abstract":Quanti* OR "Abstract":Model OR "Abstract":Assess OR "Abstract":Metri* OR "Abstract":Federate OR "Abstract":Evaluate OR "Abstract":Maturity OR "Abstract":Standard))

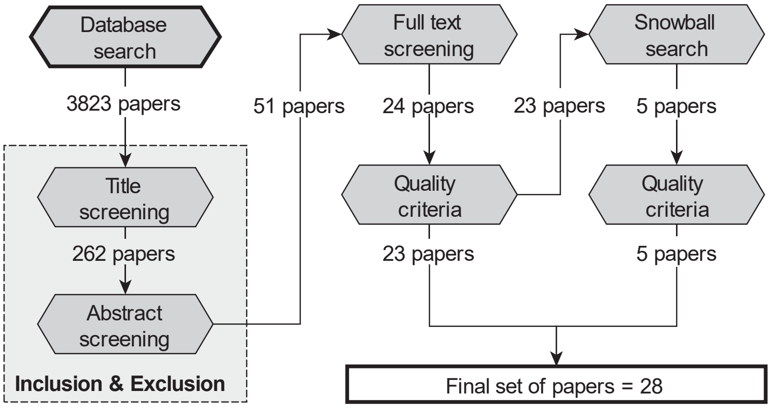

Figure 1: SLR search process

The search process is shown in **Error! Reference source not found.**. The search of the six databases returned 3,823 papers, published between 1980 and 2017. A title screening was performed on the search results and 262 papers were shortlisted. After reading the abstracts of the shortlisted papers, 51 papers were selected for the full-text screening. Finally, 24 papers were selected after the full-text screening. Then, quality criteria were applied on the final papers and one paper was rejected, leaving 23 papers as the final result of the database search. Then, a snowball search was performed on the final 23 papers to make sure no papers were missed. The snowball search was able to identify another 5 papers and they all passed the quality criteria test. Additionally, 2 other papers were identified but their sources are not currently accessible and so they have been excluded from the result. Therefore, 28 papers were selected as the final set after the entire SLR process. The final 28 interoperability measurement models are listed in

Table 1 in chronological order of their publication dates. Details about each of the models are described in a technical report (Jabin, Dimyadi, & Amor, 2019).

Table 1: Interoperability measurement models

	Model name	**Model**
	Spectrum of Interoperability Model (Lavean, 1980)	SoIM
	Quantification of Interoperability Methodology (Mensh, 1989)	QoIM
	Military Communications and Information Systems Interoperability (Amanowicz & Gajewski, 1996)	MCSI
	Interoperability Assessment Methodology (Leite, 1998)	IAM
	Levels of Information System Interoperability (C4ISR, 1998)	LISI
	Organizational Interoperability Maturity Model for C2 (Clark & Jones, 1999)	OIM
	Stoplight (Hamilton Jr, Rosen, & Summers, 2002)	SL
	Layers of coalition interoperability (Tolk, 2003)	LCI
	Levels of Conceptual Interoperability Model (Tolk & Muguira, 2003)	LCIM
0	System-of-Systems Interoperability (Morris, Levine, Meyers, Place, & Plakosh, 2004)	SoSI
1	Non-Technical Interoperability Framework (Stewart, Cremin, Mills, & Phipps, 2004)	NTI
2	Organizational Interoperability Agility Model (Kingston, Fewell, & Richer, 2005)	OIAM
3	The Layered Interoperability Score (Ford & Colombi, 2007)	i-Score
4	GridWise Interoperability Context-Setting Framework (GridWise Architecture Council, 2007)	GwICSF
5	Government interoperability maturity matrix (Sarantis, Charalabidis, & Psarras, 2008)	GIMM
6	Enterprise Interoperability Framework (Chen, Vallespir, & Daclin, 2008)	EIF
7	Maturity model for enterprise interoperability (Guédria, Chen, & Naudet, 2009)	MMEI
8	Business Interoperability Quotient Measurement (Zutshi, Grilo, & Jardim-Goncalves, 2012)	BIQMM
9	Customizable interoperability assessment methodology (Cornu, Chapurlat, Quiot, & Irigoin, 2012a)	CIAM
0	Interoperability assessment in the deployment of technical processes (Cornu, Chapurlat, Quiot, & Irigoin, 2012b)	IADTP
1	Semantic interoperability assessment (Yahia, Aubry, & Panetto, 2012)	SIA
2	A reliability-based measurement of interoperability for systems of systems (Jones, Domerçant, & Mavris, 2013)	RBMoL
3	Ultra-large-scale systems interoperability framework (Rezaei, Chiew, & Lee, 2013)	ULSSIF
4	Testing VM interoperability at an OS and application level (Lenk et al., 2014)	TIOSA
5	A novel approach IMA of interoperability measurement (Koulou, El Hami, Hmina, Elmir, & Bounabat, 2016)	IMA

	Model name	Model
6	Disaster Interoperability Assessment Model (Da Silva Avanzi, Foggiatto, Dos Santos, Deschamps, & Loures, 2016)	DIAM
7	INTERO - an Interoperability Model for Large Systems (Spalazzese, Pelliccione, & Eklund, 2017)	INTERO
8	Semantic interoperability evaluation model for devices in automation systems (Dibowski, 2017)	SIEMoD

Table 2: Type of model versus Scale

Scale / Type	Nominal	Ordinal	Ratio	Absolute
Maturity levels		SIEMoD, GwICSF, LCI, LCIM		
Maturity matrix		LISI, OIM, GIMM, OIAM		
Interoperability Matrix	TIOSA	SoIM, EIF		
Graded Interoperability Matrix			RBMoL, MMEI, ULSSIF	
Stoplight estimates	SL			
Scores	NTI		SIA, CIAM	i-Score, QoIM, IMA, IADTP
Coordinate estimates			MCISI	
Multivariate estimates		INTERO	DIAM	

3. Analysis

An initial analysis on the 28 models investigated which ones had incorporated all three components of the measurement process. Most of the papers focused on the second and third components of the measurement process. LISI was the only model that gave a clear description of the first component, i.e. how the actual interoperability of entities was measured. Since this study aims to gain an insight into how to represent the measured interoperability between a software tool and a data exchange standard in conformance tests in the BIM domain, the analysis of the 28 models focused on their measurement outcome, i.e. how they represented the interoperability between entities. The measurement theory recognises five major scales, namely nominal, ordinal, interval, ratio, and absolute (Fenton & Pfleeger, 1997). Three measurement models (IAM, SoSI, and BIQMM) analysed were framework models that did not provide a plot of the measurement outcome in terms of any scale. These models suggested that the analysts themselves could select an appropriate scale to represent the outcome. A major difference in representing a measurement outcome on different scales is the ability to undertake different levels of comparison between measured entities. The analysis of the 28 models found that the models represented interoperability on all the scales except for the interval scale. However, it was observed there were major differences in the type of representation of measurement outcomes for the models which fell into the same scale type. Hence, this study developed a 'type of model versus scale' categorisation to find out the number of unique ways the available interoperability measurement models have represented interoperability.

Table 2 lists out the remaining 25 models which have represented their outcome in one of the four major scales against different types of representation. Thus, it is found that interoperability has been represented in twelve unique ways as per the interoperability models gathered in this SLR. The following subsections discuss each of the observed scales.

3.1. Nominal scale measures

Nominal scales are the most primitive form of measurement that consists only of different classes with no notion of ordering among the classes. Any distinct symbolic or number system is an acceptable measure and there is no notion of the magnitude of difference associated with the numbers or symbols (Fenton & Pfleeger, 1997). Three models represent interoperability on a nominal scale. The simplest of the models is SL (Hamilton Jr et al., 2002) and comes under the stoplight category. This type of representation is called stoplight because it cautions the user about the current state of interoperability similar to traffic lights. In this model, each interoperability level is given four colours: red, green, yellow, orange, which denotes different states of interoperability between the entities. TIOSA (Lenk et al., 2014) represents interoperability in a matrix format. In this type of representation, the entities being tested are listed in row and column headers, and the interoperability between them is denoted by the intersecting cell. TIOSA denotes interoperability using three states: 1) Successful; 2) Warning; and 3) Failure. The NTI (Stewart et al., 2004) classifies the measurement outcome into six different scores, which are 1, 2, 4, 8, 12, and 16. These are arbitrary numbers known as the multinational forces co-operability index (identifiers for the classification) and do not signify any order or precedence of the represented interoperability classes.

3.2. Ordinal scale measures

An ordinal scale is basically a nominal scale that has information about the ordering of the classes or categories and enables analysis which is not possible with the nominal scale. The numbers represent ranking only. Hence addition, subtraction and other arithmetic operations have no meaning (Fenton & Pfleeger, 1997). There are eleven models under the ordinal scale, making it the most selected scale for measurement. The ordinal scale may be considered as the ‘sweet spot’ among the scales, because it combines the ease of representation of the nominal scale, as well as providing useful information about the relative ranking between different classes. SIEMoD, GwICSF, LCI, and LCIM represent interoperability as maturity levels. This type of representation classifies the measured interoperability between entities into various ordered levels. A standard format observed in the models is: the higher the level, the more interoperable. LISI, OIM, GIMM, and OIAM represents measured interoperability in a maturity matrix. A maturity matrix is an extension of a maturity level and can express more information than an individual maturity level. The matrix will have maturity levels represented as row headers and other measured attributes as the column headers, and the elements measured will be listed in the appropriate cells. SoIM and EIF represent interoperability as an interoperability matrix. It is the same kind of representation as in the interoperability matrix type on a nominal scale, except that the values shown in the cells will also convey the level of interoperability. The INTERO model represents interoperability using a radial chart where each arm of the chart represents the interoperability level of each measured attribute.

3.3. Ratio scale measures

A ratio scale is the most accurate scale, which enables more analysis than either the nominal or ordinal scales on the measurement outcome. For example, one can infer from a ratio scale that one software tool is twice as interoperable as another. There is a zero element, representing the total lack of the attribute. The measure starts at zero and increases in equal intervals known as units. It preserves order, the size of intervals between entities, and the ratio between entities (Fenton & Pfleeger, 1997). There are seven models under this scale. The model for representing interoperability on a ratio scale is relatively more complex than the other two scales because a ratio scale model must derive precise numerical values that capture and preserve more details about the measured attributes than other two scales where the model only classifies the attributes into broad classes. RBMoL, MMEI, and ULSSIF express interoperability as a graded interoperability matrix. This type of model assigns a number between 0 and 1 (ratio) in each cell of the matrix. SIA and CIAM represent the interoperability in percentage scores. DIAM represents interoperability for multiple parameters as a number between 0 and 1 on a radial chart. MCISI uses theories from geometry to plot the interoperability as points in multidimensional space, and the distance between the points represents the magnitude of interoperability between the entities. A zero distance means full interoperability and interoperability reduces as the distance increases.

3.4. Absolute scale measures

An absolute scale is a simple scale where the measurements are counts of elements in the entity set. The models in this category count some aspects of the measured entities. An absolute scale always takes the form "number of occurrences of x in the entity" (Fenton & Pfleeger, 1997). There are four measurement models under this scale which are i-Score, QoIM, IMA, and IADTP. All the models represent interoperability as scores. However, it was observed that even though the final output of these models are in an absolute scale, the measurement model does additional calculations on the counts to convey more information about the measurement. For example, QoIM further calculates a value for each component measured called Rx, which is the ratio of the total number of positive events out of the total number of events. This calculation brings in aspects of a ratio scale, but the model does not combine multiple Rx values into one final ratio, which leaves the output as individual absolute scores for each interoperability component measured.

4. Discussion

BIM is no longer an experimental technology and is being used in residential through to mega projects across the world. Being a certification body for BIM is a serious responsibility, especially when the technology being certified is used in critical applications where a failure of the certified system can result in the loss of life. It is a risky proposition when *"no certification scheme can guarantee completely error-free operation"* (Steinmann, 2018), and the certification process fails to convey this message to the end-user who may expect 100% error-free operation from a certified system (Lipman et al., 2010). Data accuracy while exchanging building designs is extremely crucial. For example, the information about a load-bearing structure getting altered while sending a 'good for construction' design to the contractor may cause a catastrophic failure of the structure. The test reports from individual researchers who have tested the accuracy of the data exchange of certain certified software tools are disturbing. Researchers have found a change in position and orientation for structures like columns and found missing load-bearing components after an import-export cycle (round-tripping) of the test BIM models (Lee, Won, Ham, & Shin, 2011; Ma, Ha, Chung, & Amor, 2006). When end-users are unaware that the certified software tools still have interoperability problems, they will not exercise necessary precautionary measures to counter the problems, and this has the potential for disastrous consequences. Therefore, it is necessary to advance research on upgrading the current conformance testing methods in the BIM domain into a true interoperability measurement methodology that can convey the results and the limitations of the measurement outcome without any ambiguity for the end-users.

The current conformance tests for BIM publish the test outcomes as a list of conforming and non-conforming attributes (buildingSMART, 2019), or as the number of attributes that cause some error (NIBS, 2014). As per the measurement theory, these conformance test results can only be classified as a 'quantified indication' (the first component of the measurement process). According to the model-based account of measurement a quantified indication is *"not yet a claim about any aspect of the object or event intended to be measured, but only a mathematical description of the final state of the measuring apparatus [the conformance tests, in this context]"* (Tal, 2017). Hence, the current conformance tests need to be further enhanced to become interoperability measurement methodologies. To achieve this, conformance tests need a measurement outcome and a model. A measurement outcome is defined as *"a knowledge claim associating one or more parameter values with the object or event being measured"* (Tal, 2017). Tal (2017) also specifies that to obtain the status of a measurement outcome, a knowledge claim must be abstracted away from the concrete method of the actual measurement and should typically be expressed in some unit on a particular scale. The abstraction can be done by developing a model of the measure to be represented, i.e. the second component of the measurement process (Fenton & Pfleeger, 1997). Once the measurement outcome is derived using the model from the quantified indications, it should be represented on a scale. One set of quantified indications may be represented in many scales depending upon the model created. In the measurement theory, this is called the ‘uniqueness problem’. The problem arises from the question: what should be done when there are several different possible scales (representations) for the same set of quantified indications? (Fenton & Pfleeger, 1997). To answer this question in the context of interoperability measurement in BIM, one should study the major scales used in the measurement theory and analyse the various possibilities of representing interoperability in those scales. Hence, the 28 interoperability models gathered through SLR were analysed from the angle of how they represented interoperability measurement outcomes. With the insights gained, two probable methods for representing IFC conformance tests within the existing processes are suggested by this study.

Since IFC already have a conformance test conducted by buildingSMART that provides the quantified indications (buildingSMART, 2019), this study uses it as an example to show how the quantified indications can be transformed into a measurement outcome. In the case of the current conformance test for IFC the quantified indications are a list of attributes tested for interoperability and each attribute after testing is assigned one of the following three values: 1) supported; 2) restricted; or 3) not supported (buildingSMART, 2019). This test may be upgraded into a proper measurement by converting the quantified indication into a measurement outcome. The simplest possible method to upgrade the current conformance test into a measurement method is to adopt the nominal scale combined with aspects from the absolute scale. This can be easily done because the quantified indications are already classified into three categories (supported, restricted, and not supported), and the total number of occurrences of these three attributes are already counted and displayed in the summary page. For example, the grand total of these assigned values may be calculated and displayed as a single score under three distinct categories. The rules defining how to calculate the total sum will become the model of this measurement method. Each sum behaves like values of an absolute scale, which are then classified into three distinct classes represented on the nominal scale. This final classification could be shown along with the certification logo. Displaying the measurement outcome along with the certification logo will convey to the end-user that the software tool is certified with some limitations, such as a number of attributes are not supported or are restricted (buildingSMART, 2019).

The drawback of the measure described above is that it does not convey the criticality of the unsupported and restricted attributes. For example, an unsupported attribute like ‘Placement Relative’ (name of an attribute tested) might be more critical than an unsupported ‘Geometry Axis’ in *IfcFastener* class (buildingSMART, 2019). Also, the final measurement outcome does not convey any information regarding the overall level of interoperability of the software tool. This problem can be solved by developing a different model that assigns a weighting for each attribute according to its criticality. For example, the equation [1 - (weights of unsupported attributes / total weights)] will give the level of interoperability of the tested software tool with a given IFC schema. A similar calculation can also be carried out with the restricted attributes (using different weights), and then both can be combined to give the final interoperability value. The resulting value will be on a ratio scale that gives the end-user more clarity on the overall interoperability of the software tool. This shows that it is possible to represent the measurement outcome in different ways by developing different models for the same quantified indicators.

5. Conclusion

The current conformance tests in the BIM domain do not conform to the measurement theory and consequently are not capable of accurately conveying the measurement outcome and limitations of the measure to the end-users. Other domains have used different representations for the metrics of interoperability. These can be adapted by the current conformance test processes to make them a true interoperability measurement methodology. To ensure that end-users are provided with the precise meaning of the conformance test, there needs to be further work on how current conformance tests can be revamped into true interoperability measurement methodologies. The suggested method of adapt measurement methods from other domains can serve as a starting point for further research.

A limitation of this work is that the suggested measures in this study are only an example of what could be done to upgrade the current BIM conformance test methodologies. As measurement theory is an established field in science, there is great potential to further explore the development of various models to define interoperability in BIM and represent it in an appropriate scale using techniques and methods from the measurement theory. Further research is being conducted by the authors to develop the best possible model to measure and represent interoperability of a BIM software tool against a selected data exchange standard.

References

Amanowicz, M., & Gajewski, P. (1996). *Military communications and information systems interoperability*. Proceedings of MILCOM '96 IEEE Military Communications Conference, McLean, VA, USA. *1*, 280-283.

Amor, R., & Dimyadi, J. (2010). *An Open Repository of IFC Data Models and Analyses to Support Interoperability Deployment*. Proceedings of 27th CIB W78 International Conference, Cairo, Egypt, 16-18.

Arksey, H., & O'Malley, L. (2005). Scoping studies: towards a methodological framework. *International journal of social research methodology, 8* (1), 19-32.

buildingSMART. (2010). IFC Certification 2.0 Announcement. Retrieved April 25, 2019, from http://www.buildingsmart-tech.org/certification/ifc-certification-2.0/ifc-certification-2.0-announcement

buildingSMART. (2019). Certified Software. Retrieved April 7, 2019, from https://www.buildingsmart.org/compliance/certified-software/

C4ISR. (1998). *Levels of information systems interoperability (LISI)*. Retrieved April 3, 2019 from http://web.cse.msstate.edu/~hamilton/DODAF/LISI.pdf

Chen, D., Vallespir, B., & Daclin, N. (2008). *An Approach for Enterprise Interoperability Measurement*. Proceedings of MoDISE-EUS, Montpellier, France, 1-12.

Clark, T., & Jones, R. (1999). *Organisational interoperability maturity model for C2*. Proceedings of the 1999 Command and Control Research and Technology Symposium, Newport, RI, USA, 1-13.

Cornu, C., Chapurlat, V., Quiot, J.-M., & Irigoin, F. (2012a). Customizable interoperability assessment methodology to support technical processes deployment in large companies. *Annual Reviews in Control, 36* (2), 300-308.

Cornu, C., Chapurlat, V., Quiot, J.-M., & Irigoin, F. (2012b). *Interoperability Assessment in the Deployment of Technical Processes in Industry*. Proceedings of INCOM 2012, The 14th edition of IFAC's triennal symposium INCOM, Bucharest, Romania, F20-F25.

Da Silva Avanzi, D., Foggiatto, A., Dos Santos, V. A., Deschamps, F., & Loures, E. d. F. R. (2016). A framework for interoperability assessment in crisis management. *Journal of Industrial Information Integration, 5*, 26-38.

Dibowski, H. (2017). *Semantic interoperability evaluation model for devices in automation systems*. Proceedings of Emerging Technologies and Factory Automation (ETFA), 22nd IEEE International Conference, Limassol, Cyprus, 1-6.

Fenton, N., & Pfleeger, S. (1997). *Software Metrics: A Rigorous Practical Approach* (2nd ed.). Boston, MA: PWS Publishing Company.

Ford, T., & Colombi, J. (2007). *The interoperability score*. Proceedings of the 5th conference on systems engineering research, Hoboken, NJ, USA, 1-10.

GridWise Architecture Council. (2007). GridWise Interoperability Context- Setting Framework. *The GridWise Architecture Council*, 1-52.

Guédria, W., Chen, D., & Naudet, Y. (2009). *A maturity model for enterprise interoperability*. Proceedings of OTM Confederated International Conferences On the Move to Meaningful Internet Systems, Vilamoura, Portugal, 216-225.

Hamilton Jr, J. A., Rosen, J. D., & Summers, P. A. (2002). An interoperability road map for C4ISR legacy systems. *Acquisition Review Quarterly-Winter 2002*, 17-31.

Jabin, J., Dimyadi, J., & Amor, R. (2019). *Systematic literature review on interoperability measurement models* [Technical report] 1-40. doi: 10.13140/RG.2.2.33957.35047

Jones, W., Elizabeth, Domerçant, J. C., & Mavris, D. N. (2013). *A reliability-based measurement of interoperability for systems of systems*. Proceedings of Systems Conference (SysCon), 2013 IEEE International, Orlando, Florida, USA, 408-413.

Kingston, G., Fewell, S., & Richer, W. (2005). *An organisational interoperability agility model*. Proceedings of 10th ICCRTS, Virginia Beach, VA, USA, 1-18.

Kitchenham, B. (2004). *Procedures for performing systematic reviews* [Technical report] 1-28. Retrieved from Keele University website http://www.inf.ufsc.br/~aldo.vw/kitchenham.pdf,

Kiviniemi, A. (2008). *IFC Certification process and data exchange problems*. Proceedings of the 7th European Conference on Product and Process Modelling (ECPPM 2008), Sophia antipolis, France, 517-522.

Koulou, A., El Hami, N., Hmina, N., Elmir, A., & Bounabat, B. (2016). *A novel approach IMA of interoperability measurement*. Proceedings of Information Science and Technology (CiSt), 2016 4th IEEE International Colloquium on Information Science and Technology, Tangier, Morocco, 665-669.

Lavean, G. (1980). Interoperability in defense communications. *IEEE transactions on communications, 28* (9), 1445-1455.

Lee, G., Won, J., Ham, S., & Shin, Y. (2011). Metrics for Quantifying the Similarities and Differences between IFC Files. *Journal of Computing in Civil Engineering, 25* (2), 172-181.

Leite, M. J. (1998). *Interoperability assessment*. Proceedings of the 66th Military Operations Research Society Symposium, Monterey, CA, USA, 1-13.

Lenk, A., Katsaros, G., Menzel, M., Revelant, J. R., Skipp, R., Leon, E. C., & Gopan, V. (2014). *TIOSA: Testing VM Interoperability at an OS and Application Level-A Hypervisor Testing Method and Interoperability Survey*. Proceedings of Cloud Engineering (IC2E), 2014 IEEE International Conference, Boston, Massachusetts, USA, 245-252.

Lipman, R., Palmer, M., & Palacios, S. (2010). Assessment of conformance and interoperability testing methods used for construction industry product models. *Automation in Construction, 20*, 418-428.

Ma, H., Ha, K. M. E., Chung, C. K. J., & Amor, R. (2006). *Testing semantic interoperability*. Proceedings of Joint International Conference on Computing and Decision Making in Civil and Building Engineering (ICCCBE), Montreal, Canada, 1216-1225.

Mensh, D. R. (1989). The quantification of interoperability. *Naval Engineers Journal, 101* (3), 251-259.

Morris, E., Levine, L., Meyers, C., Place, P., & Plakosh, D. (2004). System of Systems Interoperability (SOSI) Final Report. *Integration of Software-Intensive Systems Initiative*, 1-53.

NIBS. (2014). January 2014 bSa Challenge: Graphisoft. Retrieved April 16, 2019, from https://www.nibs.org/page/bsa_chall14_gs

Rezaei, R., Chiew, T. K., & Lee, S. P. (2013). An interoperability model for ultra large scale systems. *Advances in Engineering Software, 67*, 22-46.

Sarantis, D., Charalabidis, Y., & Psarras, J. (2008). Towards standardising interoperability levels for information systems of public administrations. *The Electronic Journal for E-commerce Tools & Applications (eJETA) Special Issue on Interoperability for Enterprises and Administrations Worldwide, 2*, 1-15.

Spalazzese, R., Pelliccione, P., & Eklund, U. (2017). INTERO: an Interoperability Model for Large Systems. *IEEE Software*, 1-11.

Steinmann, R. (2018). IFC Certification of BIM Software. In *Building Information Modeling Technology Foundations and Industry Practice* (pp. 139-153). Cham, Switzerland: Springer Nature.

Stewart, K., Cremin, D., Mills, M., & Phipps, D. (2004). *Non-technical interoperability: the challenge of command leadership in multinational operations.* Proceedings of the 10th International Command and Control Research and Technology Symposium: The Future of C2, McLean, VA, USA, 1-9.

Tal, E. (2017). A Model-Based Epistemology of Measurement. In *Reasoning in Measurement* (pp. 233-251). New York, NY: Routledge.

Tolk, A. (2003). *Beyond technical interoperability-introducing a reference model for measures of merit for coalition interoperability.* Proceedings of 8th International Command and Control Research and Technology Symposium, Washington DC, USA, 1-46.

Tolk, A., & Muguira, J. A. (2003). *The levels of conceptual interoperability model.* Proceedings of the 2003 fall simulation interoperability workshop, Orlando, Florida, USA. *7*, 1-11.

Yahia, E., Aubry, A., & Panetto, H. (2012). Formal measures for semantic interoperability assessment in cooperative enterprise information systems. *Computers in Industry, 63* (5), 443-457.

Zutshi, A., Grilo, A., & Jardim-Goncalves, R. (2012). The business interoperability quotient measurement model. *Computers in Industry, 63* (5), 389-404.