Evaluating the Challenges of Data Management in COBie Datasheet and Mitigation Measures

Vishal Kumar[1], Evelyn Ai Lin Teo[1]

[1]Department of the Built Environment, College of Design and Engineering, National University of Singapore, Singapore

*Email: bdgkuma@nus.edu.sg

Abstract. BIM to FM has gained considerable attention. Nevertheless, BIM adoption for FM is still low. Issues such as lack of standardized process to develop FM data during design and construction stages, Standardized data format for data transfer, and fragmented databases are highlighted in multiple BIM-FM research. COBie defines when, how, and what data needs to be captured for FM purposes. However, previous research combined with explorative studies highlighted several challenges with handling the COBie datasheet, especially its widely used spreadsheet format. This study aims to identify the issues associated with COBie handling (especially its spreadsheet format) and propose a COBie Dataset Management System framework (CDMS) to help solve these issues. In developing the proposed CDMS framework, a critical review of the published articles related to the COBie datasheet has been conducted. An exploratory study using a BIM Model was conducted along with the literature review to understand the key challenges highlighted in the reviewed articles. Based on the identified key issues, underlying reasons were recognized, and key ideas for the framework have been developed, potentially solving these key challenges. The research finding will help develop COBie-centric applications and enhance the entire COBie data capturing workflow.

1. Introduction

The way buildings are designed is revolutionized by the advanced computerization of the entire construction process, starting from concept development to Operation and Maintenance (O&M) [1]. Building Information Modelling (BIM), a disruptive technology, has enabled the digital environment for sharing and storing useful data related to a building's design, operation, and management. It can store critical and useful product and asset data for effective and efficient building information management.

However, tactical skills and tenacious strategies are expected from the facility management team to capture the building's intricate and large volume of complex asset data during the entire lifecycle of the building [2]. The skills required may include (but are not limited to) strategic management of building assets, interior operations, computer system analysis, plant operations, etc. Regardless of many researchers pointing out the benefits of BIM-FM integration, there is limited evidence in the literature of standardized FM procedures and processes applied to BIM-FM integration and accrued tangible benefits. To deal with barriers associated with BIM to FM data transfer, such as data interoperability, data integration, and reliance on recapturing data into FM system from the paper-based manuals, COBie (Construction Operations Building information exchange) has been developed. COBie is a vendor-neutral data exchange format, which can be delivered in various open file formats such as STEP, XML, and Spreadsheet. Currently, the most commonly used data transfer file format is spreadsheets due to its familiarity with most construction industry professionals.

However, various published articles highlighted multiple challenges with COBie-related data handling, such as issues in interpreting, capturing, querying, and finding the data inside the COBie sheet [3-8]. This research study wants to identify the challenges associated with handling COBie datasheet and proposes mitigation strategies. This research article aims to highlight the challenges related to COBie datasheet information capturing workflow. Furthermore, the article presents a COBie Dataset Management System (CDMS) framework that can help mitigate these challenges. This study's findings can help future researchers develop applications based on the proposed framework and develop use cases to enhance COBie adoption in the industry for BIM to FM data transfer.

2. Literature Review

2.1 COBie Workbook, Structure and Data Drops

A COBie workbook consists of 20 worksheets (including an instruction worksheet), and each worksheet is defined to have a specific type of information. The structure of the worksheets is predefined, and the location of columns in each worksheet is fixed. The type and data format units that can be used are specified in the "Picklist" worksheet. The information or data between the worksheets is connected using color coding [9]. The process of capturing and checking the data at defined stages is termed "data drops.” Data drops are defined as generating COBie files during different developmental stages of a project to track, control, and build an efficient FM data deliverable during the handover stage [10]. The process consists of extracting COBie data at predefined stages and checking them against the set benchmarks and requirements documents. Different authors and organizations have described the stages of data drops in different ways. For example, East [11] identified the data drops stages as as-planned, as-designed, as-constructed, as-occupied, as-built, and as-maintained (six stages).

Moreover, East and Carrasquillo-Mangual [9] mentioned only four stages for data drops. They defined them as design development deliverables (35% design), construction document design deliverables (100%), beneficial occupancy construction deliverables, & as-built construction deliverables. Similarly, the BIM task group from London proposed five data drop stages with elaborate descriptions of expected outcomes [12]. For better clarity, they linked the data drops with the RIBA outline plan of work (2007) stages.

2.2 Challenges in handling COBie datasheet

Various studies are done around COBie. These studies have highlighted different challenges associated with COBie handling. Understanding the reasons behind these challenges is important to mitigate these issues. However, there is a paucity in the literature that goes beyond highlighting these challenges to solving them. Figure 1 highlights these challenges vis-à-vis a COBie datasheet information capturing workflow.

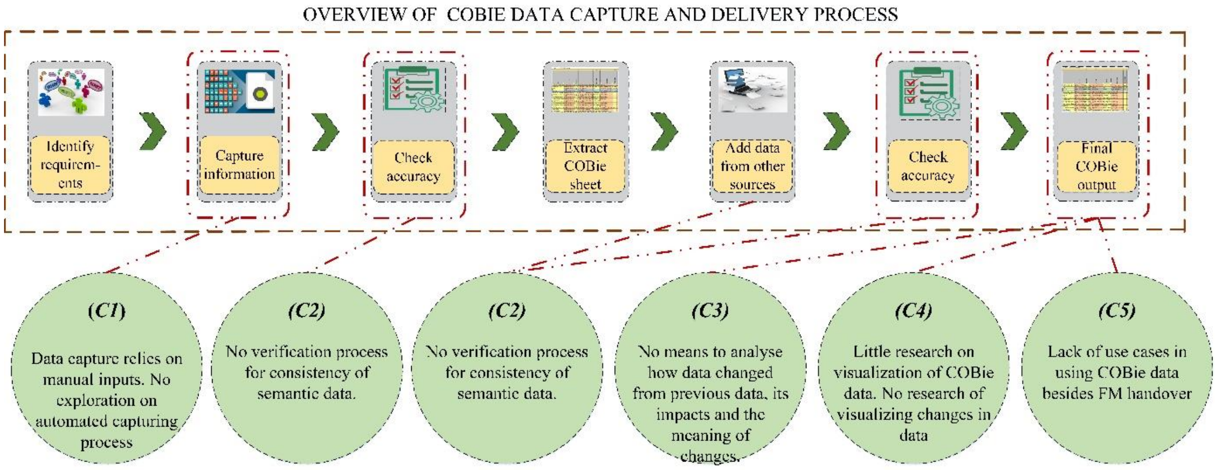

Figure 1: Challenges in COBie data capturing and delivery process.

2.2.1 Data Capture (C1). Data capture is the critical part of a COBie deliverable. Data is captured at various stages of a project and regularly verified. However, some authors highlighted critical issues associated with it. Data inside the COBie datasheet come from the BIM model (partially), and the rest are filled manually for multiple worksheets. This manual entry is prone to error, and data duplication can result in data mismatch between the BIM model and the COBie sheet [7, 13, 14]. Additionally, such manual entry can miss the entire relationship of data to be captured in the COBie workbook (one data inside the COBie datasheet is linked to various other data inside different worksheets) [15, 16].

COBie guide defines all attributes captured for each type of element, such as equipment, spaces, and so on. However, the current COBie tools don't create these attributes automatically based on the model element [17]. Most COBie tools are designed to create common parameters. Nevertheless, the COBie guide defines additional parameters which are different and specific to different types of building elements. This lack of automation has profound implications. It adds extra efforts for the project team first to understand what all attribute parameters need to be created, and then their respective values for each element need to be captured by referring to the COBie guide. Besides these, the user needs to constantly check for data consistency and identify data errors. No previous efforts have been made to look into this gap holistically.

2.2.2 Data consistency (C2). Various authors have highlighted issues with data consistency inside COBie [4, 10, 18, 19]. This data inconsistency can arise because COBie allows data modification after being exported from BIM software. Such errors can be in data deletion, missing data dependency, data changes, and the wrong format of data [4, 20]. How to tackle these errors has not been discussed, which adds an extra burden on the BIM modeler for verification. Moreover, besides data error, another problem is verifying the semantic connection of data. Various data inside the COBie datasheet are related semantically, which needs to be verified logically (for example, the "Jobs" worksheet has a column with predecessor activities to be mentioned. However, any error in this data cannot be identified by any existing COBie verifiers (because they mainly check the format of data and simple links) such as the ONUMA COBie validator or the COBie QC tool. Most of these quality checkers don't go beyond checking values and connections. However, it is essential not to look at the COBie datasheet independently but as a progressively developed datasheet. Having insight into how the data is changing inside the COBie datasheet can help the project team understand the modification of asset changes in a project.

2.2.3 Data Changes (C3). COBie-related studies have not emphasized how COBie datasheet changes as the project progress. East and Carrasquillo-Mangual [9] highlighted that more data needs to be added to the existing data as the project progresses. However, data inside COBie can change significantly between the two phases. Multiple data inside the COBie datasheet can be deleted, added, changed in values, or changed in a relationship. This issue with COBie has not been reported in previous studies. The purpose of using COBie data drops (i.e., generating COBie datasheet at regular intervals to check the data) is to monitor the consistency of COBie data and analyze how data is changing and identify discrepancies. However, this doesn't consider comparing and tracking the changes and making meaning out of them. To make a meaning out of data drop process, it is imperative to look at it beyond the simple process of checking the data but to understand it as a process to see how data is changing in a project and whether data captured in one phase is missed in another data drop phase because COBie is a compilation of both BIM extracted and outside BIM data.

2.2.4 Data visualization (C4). Lack of visualization and query capability has been identified as a significant issue while handling COBie datasheets [5, 7, 18]. Yalcinkaya and Singh [21] have addressed the issue of visualizing the COBie datasheet by proposing the application of a node-link diagram over the COBie datasheet. Farias *et al.* [5] applied OWL language to develop a query tool for COBie. The primary issue with the previous works is that they lack to address the changing data inside COBie and how to visualize them and treat COBie datasheet as an independent. Visualizing the patterns and changes inside the COBie data over various data drops is necessary. This visualization can help understand how project changes affect the data inside COBie. While visualizing the patterns and changes, such visualizers can highlight the outliners in the COBie sheet (for example, data without any connection means supporting data is missing, which is an error in COBie output).

2.2.5 Use cases beyond FM handover (C5) . Some previous studies have demonstrated its use in doing LEED (Leadership in Energy and Environmental Design) related calculations [22] or using COBie to perform HVAC analysis [23]. Still, there is a lack of substantiative case studies showing the COBie datasheet's utility beyond FM handover use. There is a lack of understanding about the real usefulness of the COBie datasheet, how they benefit FM handover, and its usefulness beyond FM [20, 24, 25] . Since COBie expandability has been questioned, owners use their traditional methods to capture FM data and use their own methods, leading to duplication of efforts [25, 26]. Therefore, use cases around COBie as a central datasheet are needed to showcase its usefulness and convince owners that using COBie can help them control and manage facilities-related datasets efficiently. By maintaining such data, users can perform multiple applications.

3. Research Phases

This article presents a part of a larger study conducted in 6 phases, as shown in Figure 2. However, in this article, we are presenting a discussion about the first three phases (phase 1 to 3, dark black lines), and in the forthcoming articles, we will present the discussion and development of the last three phases (phase 3-6, red dotted lines) in more detail as they need the attention of their own. In phase 1, preliminary literature review was conducted to identify the issues related to the COBie datasheet handling and outline the research questions, aim, and objectives.

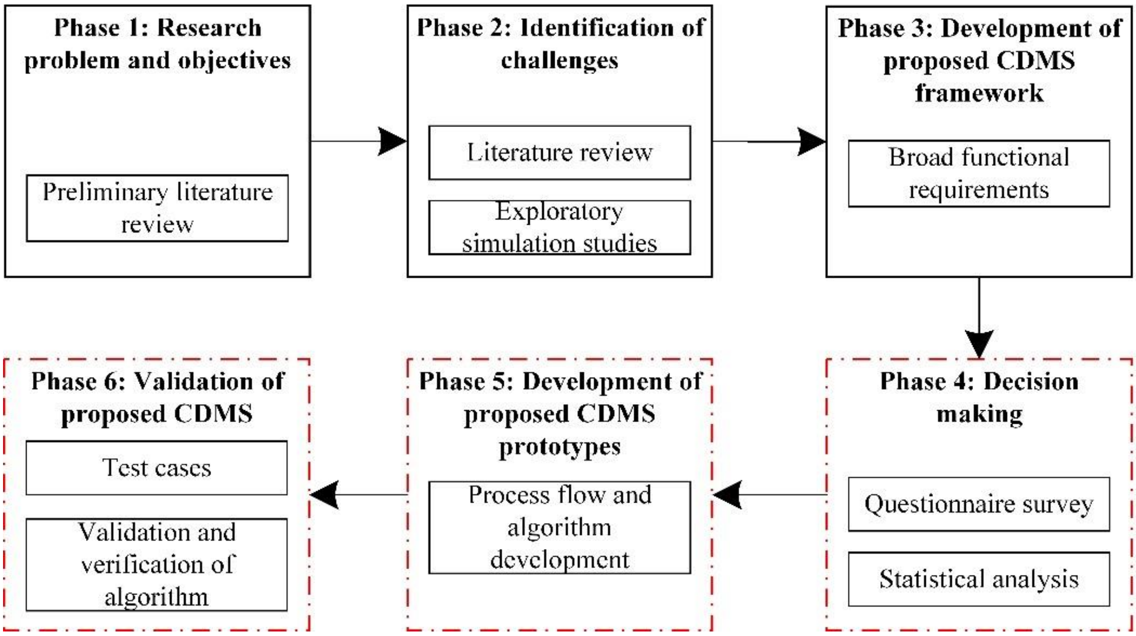

Figure 2: Research Phases

In phase 2, COBie datasheet data capturing workflow challenges were identified through a detailed literature study and an exploratory simulation study using a BIM model. This exploratory study aimed to find answers to some of the following questions: a) How is the data captured for COBie purpose; b) How is the COBie data extracted; c) How is the data represented in the COBie output; d) How are the elements of the same system represented in the output; e) How to track changes inside the COBie sheet; f) How a change in one data affect the various worksheets inside the COBie datasheet. In phase 3, mapping of the challenges with ideated mitigation strategies was conducted. This mapping is discussed in detail in our previous publication [19]. This led to the development of the proposed CDMS framework, which will be discussed in section 3.

4. Findings and Discussion

4.1 CDMS framework

This research proposes a framework for COBie Dataset Management System (CDMS) to tackle the challenges associated with the COBie datasheet information capturing workflow (Figure 3). This research study proposes organizational interventions and tools/modules developed as part of this CDMS. In the article, we only discuss the tool development part. These tools or module development criteria are generated based on the challenges identified during literature reviews and challenges faced during the exploration study. We identified three modules in CDMS (COBie Automator, COBie Evaluator, and COBie Visualizer) that can help the entire COBie datasheet capturing process. Another module (fourth) is conceptualized (COBie for Recycling) to showcase how the COBie datasheet can be used for other purposes during the building lifecycle. We are currently developing other use cases around COBie to strengthen the usefulness of the COBie datasheet.

4.1.1 COBie Automator . Section 2.2.1 highlights challenges with capturing data for COBie. There are several challenges associated with it. First, the parametric values required in the BIM model need to be created manually for various elements, which is quite a challenging task. It relies on the BIM modelers to conduct due diligence. On a practical level, many modelers are not even aware of whether the element needs to be part of COBie capture or not. Also, they might need to create custom families for various elements and select suitable families to input data. The second challenge is the manual capturing of data for each element. This is a tedious task on the part of BIM modelers where they need to look at all the product specification sheets and extract the data for every element. In a traditional workflow, most product specification sheets are in PDF format without structure for processing these data automatically through natural language processing or other such techniques. Furthermore, inputting data manually can lead to unintentional errors, which can be hard to detect.

In this research, we identified ways to automate the process through ideations. For this automation purpose, we explored how the Product Data template (PDTs) can be used to automate the information capturing part. Several international organizations are developing a standardized format in which products should be specified. Some of the important development in this is: Specifiers' Properties information exchange (SPie) developed under the aegis of National Institute of Building Sciences, Product Data Templates (PDT) developed by Chartered Institution of Building Services Engineers, Product Data Templates (PDT) by National Building Specification (NBS) and International standardization bodies CEN and ISO (under development). SPie development stopped at the pilot stage; however, PDTs by NBS and CIBSE are being developed continuously and used by the industry. This research identified the need to create a tool to extract data from PDTs and automatically store and assign it to the elements without any manual data input.

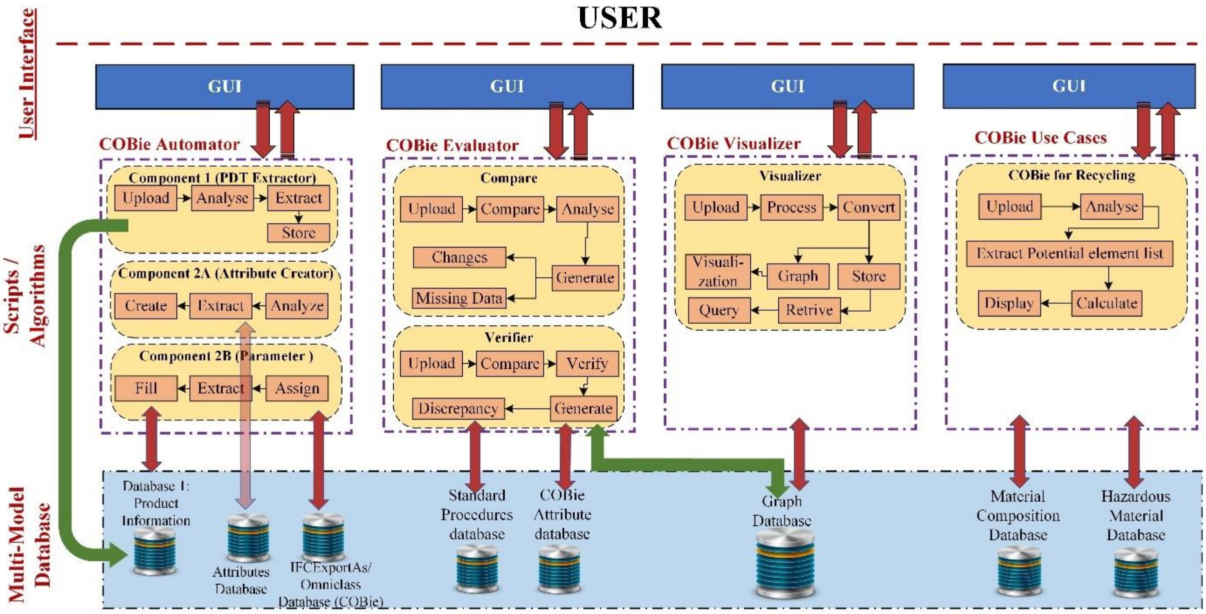

Figure 3: Proposed CDMS framework

Following are the requirements set for this tool development: a) The tool should be able to extract COBie-specific information from PDSs stored in file formats such as spreadsheets and Portable Data Format (PDF); b) The tool should be able to store extracted data in a structured format that can be retrieved and reused; c) The tool should be able to identify the equipment type and automatically create minimum required attributes (parameters) required for data capturing as defined by the COBie guide [9] using identifiers such as IFC value or Omniclass name and number; d) The tool should be able to fill COBie related attributes (parameter) data (value) automatically inside the BIM Model by assigning the product and its model. The details of the tool development will be presented in the upcoming articles. However, this tool is validated using existing PDTs and provides a use case for developing PDTs. The current PDT development is not extensive and needs more government bodies' efforts. Standardized templates of manufacturer data can help build open-source tools for automated capturing of data.

4.1.2 COBie Evaluator. COBie Evaluator is the second tool of the CDMS designed to cater to concerns over the verification and comparison of data between the COBie datasheets, as highlighted in sections 2.2.2 and 2.2.3. It was identified that the evaluation process of COBie needs to be done at several levels. Therefore, the ideation process highlighted the need for two sub-modules in the module: a) first, a part of the module should be able to compare two COBie sheets to identify changes in data and, b) a more comprehensive verification tool that not only verify data formats but also verify data from previous COBie datasheets during data drops, verify logical reasonings and prompt missing data based on historical references. Based on these issues, the following functional requirement was set for this tool development: a) The module should be able to track changes inside the COBie sheet while compared with any previous stage COBie sheet; b) The module should be able to interpret the changes inside the COBie sheet to let professionals understand what the meaning of these changes is; c) The module should be able to help in COBie datasheet consistency at a comprehensive level. The details of the COBie Evaluator module have been discussed in detail in our previous publication [13, 19, 27].

4.1.3 COBie Visualizer. Most of the usability issues are associated with the static nature of the COBie spreadsheet [4, 5]. The exploratory study substantiated these finding as COBie spreadsheet format lack visible linkage between data and dependency links. Users need to go through multiple levels of filtering and move from tabs to tabs to get a series of data required. To address this issue as highlighted in section 2.2.4, ideations highlighted that a node-link graph is an appropriate technique for efficiently visualizing such interlinked data. Furthermore, the property graph model can help store more data over the links and nodes, making it easier to store more data compactly. Based on the highlighted challenges in the literature, the following requirements were set for the tool development: a) The visualizer should be able to show the relationship between different data; b) The visualizer should be able to show the dependency of data; c) The visualizer should help in querying the COBie datasheet with logical reasoning; d) The visualizer should help in showing the changes inside the COBie datasheet; e) The visualizer should be able to show and store historical changes in data. This module needs a database where the data can be stored and later queried. Moreover, it is conceptualized to take the output from the COBie Evaluator module to visualize changes inside the graph representation. The visualizer development has been discussed in detail in our previously published article [28].

4.1.4 COBie for Recycling. The literature highlighted that owners do not see the usefulness of COBie (section 2.2.5). While COBie's primary purpose is to support the handover of BIM to FM data, the usefulness of COBie data beyond its intended use can help build a strong justification for COBie adoption. Therefore, one use case has been developed whereby the COBie datasheet has been used for recycling (Urban Mining). This module takes the COBie datasheet as input, demonstrates COBie data usage for Urban-Mining calculation, and demonstrates how COBie data can be used for hazardous equipment identification. This development has been presented in a previously published article [19].

5 Conclusions and Further Research

Most of the applications developed for COBie are for the compilation of data for COBie. Existing plugins of BIM applications can export COBie datasheets from software like Autodesk Revit or ArchiCAD. Additionally, there are other open-source applications developed for COBie exports. Furthermore, some verifier tools verify the consistency of the COBie datasheet. All these developed tools are related to exporting COBie data or verifying COBie data. All these tools lack looking at the challenges faced by COBie-related data holistically at different stages of the project. The novelty of the CDMS framework lies in addressing these challenges. The development of CDMS is based on identifying issues associated with COBie datasheet handling through an extensive literature review.

Moreover, the CDMS system is designed to address issues that were not addressed before in COBie-related research and existing tools. The CDMS framework is designed to address the issues such as automation of COBie data, verification of COBie data at a more semantic level, extracting more value from the COBie data, comparing COBie data to understand how COBie datasheet and project has changed, enhancing visualization of data through the application of graph representation, and utilizing COBie data for building demolition recovery of precious metals (Urban Mining). Various other potentials, such as carbon calculation, design feedback from the maintenance cycle, etc., can be explored, which can help demonstrate the use of developing, capturing, and maintaining COBie data for the whole lifecycle. More use cases using COBie as the central data source is in progress and will be reported in future articles. This study also set the stage for future research in developing a more comprehensive system to enhance the usability of the COBie datasheet. Furthermore, this framework can be further enhanced by forging more use cases centric around COBie.

6 References

[1] Eastman C M, et al. 2011 *BIM handbook: A guide to building information modeling for owners, managers, designers, engineers and contractors*. John Wiley & Sons

[2] Codinhoto R, et al. 2013 *BIM-FM implementation: an exploratory invesigation.* International Journal of 3-D Information Modeling (IJ3DIM). **2**(2): p. 1-15

[3] Liu R and Issa R 2013 *Issues in BIM for facility management from industry practitioners' perspectives*, in ASCE International Workshop on Computing in Civil Engineering Computing in Civil Engineering. p. 411-418

[4] Yalcinkaya M, Singh V, Nenonen S, Junnonen JM 2016 *Evaluating the usability aspects of construction operation building information exchange (COBie) standard*. In CIB World Building Congress 2016. TUT–Tampere University of Technology

[5] Farias M, Roxin A, and Nicolle C 2015 *Cobieowl, an owl ontology based on cobie standard*. in *OTM Confederated International Conferences" On the Move to Meaningful Internet Systems"*. Springer

[6] AEC Magazine 2014 *The problem with COBie*. [cited 12.08.2017]; Available from: http://aecmag.com/technology-mainmenu-35/598-the-problem-with-cobie

[7] Wang Z, Bulbul T, and Mccoy S 2015 *Comparative case study for using COBie in project turnover*. in *Proc. of the 32nd CIB W78 Conference*

[8] Kumar V and Teo E A L 2020 *Perceived benefits and issues associated with COBie datasheet handling in the construction industry.* Facilities, Vol. **39** No. 5/6, pp. 321-349

[9] East B and Carrasquillo-Mangual M 2013 *The COBie Guide: a commentary to the NBIMS-US COBie standard.* Engineer Research and Development Center, Champaign, IL

[10] Love P E, et al. 2014 *A benefits realization management building information modeling framework for asset owners.* Automation in construction. **37**: p. 1-10

[11] East B 2013 *Using COBie*, in *BIM for Facilities Managers*, P. Teichloz, Editor. 2013, John Willey & Sons: New Jersey

[12] BIM task group 2012 *COBie Data Drops-Structure, uses, and examples*. COBie – Data Drops for Standard MoJ Cell [cited 06.07.2017]; Available from: http://www.bimtaskgroup.org/cobie-data-drops/

[13] Kumar V and Teo E A L 2020, *Conceptualizing "COBieEvaluator": A rule based system for tracking asset changes using COBie datasheets.* Engineering, Construction and Architectural Management, Vol. **27** No. 5, pp. 1093-1118

[14] McPhee A 2013. *To COBie or not to COBie* [cited 20.12.2018]; Available from: http://practicalbim.blogspot.com/2013/08/to-cobie-or-not-to-cobie.html

[15] Day M. 2014 *The problem with COBie*.[cited 10.01.2019]; Available from: https://aecmag.com/technology-mainmenu-35/598-the-problem-with-cobie

[16] Yalcinkaya M and Singh V 2016. *Evaluating the Usability Aspects of Construction Operation Building Information Exchange (COBie) Standard*. in *CIB World Building Congress*. Tampere, Finland

[17] Revit 2019. *Autodesk COBie extension for Revit*; Available from: https://www.biminteroperabilitytools.com/cobieextensionrevit.php

[18] Anderson A, et al. 2012 *Construction to operations exchange: Challenges of implementing COBie and BIM in a large owner organization*. in *Construction Research Congress 2012: Construction Challenges in a Flat World*. West Lafayette, Indiana
[19] Kumar V and Teo E A L 2019 *Towards a more Circular Construction Model: Conceptualizing an open-BIM based Estimation Framework for Urban Mining*, in *CIB World Building Congress 2019*. CIB: Hong Kong
[20] Stewart G H. 2014 *COBie: The practicalities of Application*. [cited 10.06.2017]; Available from:http://iug.buildingsmart.org/resources/itm-and-iug-meetings-2014-stockholm/conference/building/cobie-the-practicalities-of-application/view
[21] Yalcinkaya M and Singh V 2019 *Exploring the use of Gestalt's principles in improving the visualization, user experience and comprehension of COBie data extension.* Engineering, Construction and Architectural Management. **26** (6): p. 1024-1046
[22] Biswas T and Krishnamurti R 2012 *Data Sharing for Sustainable Building Assessment.* International Journal of Architectural Computing. **10**(4): p. 555-574
[23] Xia L, Adkins D A, and Cheshmehzangi A. 2013 *Smart Building Information Modelling: The Use of ANN-COBie for HVAC Information Capture and Exchanged*. in *30th CIB W78 International Conference*. 2013. China: CIB
[24] NBS, *National BIM Report, 2018*. 2018: London
[25] Abdirad H and Dossick C S 2019 *Normative and descriptive models for COBie implementation: discrepancies and limitations.* Engineering, Construction and Architectural Management. **26** (8): p. 1820-1836
[26] Teicholz P 2013 *BIM for facility managers*. John Wiley & Sons
[27] Kumar V and Teo E A L 2019 *Conceptualizing "COBieEvaluator": An Application for Data Mining COBie Datasets to Track Asset Changes Throughout Project Lifecycle*, in *International Congress and Conferences on Computational Design and Engineering 2019 (I3CDE 2019)*, Society for Computational Design & Engineering: Penang, Malaysia
[28] Kumar V and Teo E A L 2021 *Development of a rule-based system to enhance the data consistency and usability of COBie datasheets.* Journal of Computational Design and Engineering. **8**(1): p. 343-361