



# KEYPHRASES ANALYSIS OF BIM STANDARDS THROUGH OCCURRENCE OF MOST COMMON BIM USES

Igor Osmokrović<sup>1\*</sup>, Sonja Panić<sup>1</sup>, Đorđe Nedeljković<sup>1</sup>

<sup>1</sup>Faculty of Civil Engineering, University of Belgrade, Serbia

\*Authors to correspondence should be addressed via email: igoro.grf@gmail.com

**Abstract:** *Building Information Modeling (BIM) does not represent only the virtual model of the facility but a comprehensive approach consisting of technology, processes, stakeholders' behavior and accompanying standards. Given the fast evolution of BIM, this paper is analysing trends of development of BIM standards throughout the years by applying the keyphrases analysis method, for some of the most common BIM uses and recognizable phrases in BIM industry.*

**Key Words:** *Building Information Modeling (BIM) / BIM Standards / BIM Uses / keyphrases analysis*

## 1. INTRODUCTION

In 2007 Building Information Modeling (BIM), was defined as a digital representation of physical and functional characteristics of a facility. [1] In 2012 the same definition is applied to term Building Information Model [2]. Between 2007 and 2012, Dodge Data & Analytics registered the increase of percentage of BIM users in North America from 28% to 71%. [3] The change in definitions implies that BIM concept has significantly evolved along with its expansion and nowadays Building Information Modeling approach should be defined as a workflow that applies the technology in developing a comprehensive virtual model of facility, while integrating all processes and aspects of the project having an important contribution to entire project life cycle through encouraging all project stakeholders (owners, engineers, contractors, suppliers etc.) to collaborate in delivering the project result. For successful BIM implementation it is required to successfully integrate all three - processes, technologies, and behavior. [4]

In parallel with the evolution of BIM concept, development of adequate standards is necessary as a fourth component needed for successful BIM implementation.

This paper is analysing trends of development of BIM standards throughout the years by applying the keyphrases analysis method, for some of the most

common BIM uses and recognizable phrases in BIM industry.

The goal of the analysis in this paper is to provide an insight to the level of presence and elaboration of major BIM uses in BIM Standards, which will be intersected with the analysis of market requirement and experts' judgment in the following phases of a research and define the recommendations for the future steps of the standards development.

## 2. ANALYSED DOCUMENTATION

One of the first countries to implement BIM in their construction industry and consequently to develop accompanying standards for BIM implementation were US, Nordic countries, the UK and Singapore. Bearing that in mind, subset of BIM standards covering the period from 2007 to 2015 is chosen, and includes standards listed in the table 1.

Table 1. *BIM Standards included in research subset*

BIM Standard	Version
US National Building Information Modelling Standard (NBIMS)	Version 1 (2007), Version 2 (2012), Version 3 (2015)
Norwegian Home Builders' Association Bim User Manual	Version 1 (2011), Version 2 (2012)
Statsbygg BIM Manual	Version 1.2.1 (2013)
BSI PAS	1192-2 (2013)
BSI PAS	1192-3 (2014)
BSI PAS	1192-5 (2015)
Singapore BIM Guide	Version 1 (2012), Version 2 (2013)
COBIM v1.0	Version 1.0 (2012)
AEC (UK/Canada) Standards	2010, 2012, 2015

A BIM Use is defined as a method of applying Building Information Modeling during a facility's lifecycle to achieve one or more specific objectives.[2]

Most common BIM uses [5], [6] and common terms and phrases in the BIM industry that were analysed include:

- BIM execution plan
- Design authoring / BIM Modeling
- Design Review
- Drawing Production
- Conflict Analysis / Multi-trade Coordination
- Quantities Extraction
- 4D Construction Sequencing
- 5D Cost Estimating
- Status / Progress Monitoring
- Record Modeling
- Facility Operation Planning
- Energy Analysis

### 3. METHODS

In this paper the different BIM uses are represented as a set of keyphrases from a text corpus, consisting of two or more words (i.e. BIM execution plan, procedures, process maps, risk assesment, etc.). Keyword analysis is a type of content analysis that uses quantitative description to analyse the content of written text. The analysis of words occurring together enables tracking of the evolution of a research field along consecutive time periods [7].

Bearing in mind that keyphrases consisting of multiple words convey more information than plain keywords, multiple statistical measures of association between words to extract keyphrases are utilized. The keyphrase of order two is defined as a pair of two consecutive words. In order to determine whether two words occurring together were a valid candidate for a keyphrase, the probability of word co-occurrence is compared to the probability of the same event in a randomly generated document. In the presented approach, following measures for word association were used:

- PMI [8]
- PMI<sub>sig</sub> [9]
- sPMI<sub>d</sub> [10]
- Dice [11]
- G2 [12]

After an extraction of highest ranked word pairs according to multiple association measures, a keyphrase network (KN) was created, with the keyphrases as nodes and relations based on shared sentence level keyphrase occurrence. Keyphrases of higher order were constructed from cliques of keyphrases in KN which always appeared in the same sequence on a sentence level.

For each BIM use, an initial set of terms is used describing them to detect BIM use relevant keyphrases (BURK). The BURK extraction process included an examination of its relatedness to other keyphrases in KN and an overview of different textual contexts in which BURK appeared.

Finally, the KN was enriched with nodes representing BIM Standards and relations to keyphrases contained in them. Each BIM Standard – keyphrase relation contained information about the location of the keyphrase occurrence in the particular Standard. The expanded KN was implemented using Neo4J (<https://neo4j.com>) as an open source graph database. Neo4j's Cypher query

language was used for analyzing BIM uses (via BURK) in Standards text corpus.

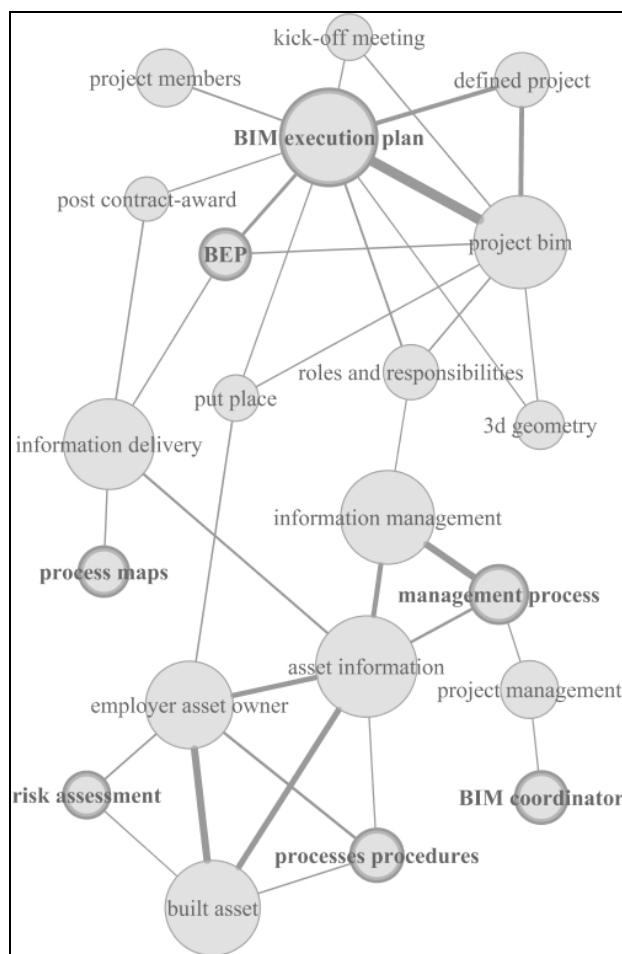


Fig. 1. BIM execution Keyphrases Sub-network

### 4. RESULTS

Analysis results with the sum number of paragraph-level occurrences for all BURKs representing one considered BIM use. Occurrences sums are broken down per year but considering that volumes of input Standards differs between years, results had to be normalized before showing the wanted trends.

Normalization is done by the volume of the Standards for each year, and it introduces the occurrence index values. The number of "meaningful paragraphs" is adopted as a measure of the volume of considered data input. "Meaningful paragraphs" number exclude those paragraphs that are not covering the topic of the text (ex. content page, headers, etc.).

After the normalization, the values of occurrence index represent the legitimate measure of elaboration of the considered BIM use (or common term) topic, within the analysed industry Standards published each year.

The values of the occurrence index for the term of "BIM execution plan" are shown in Figure 2. Similar histograms are generated for all analysed BIM uses and common terms listed in the Chapter 2.

### 5. DISCUSSION

Generated histogram for "BIM Execution plan" term gives a solid insight into the rising trend of discussion

regarding the considered term. Given the fact that "BIM execution plan" is a document that regulates entire BIM implementation on the project, it gained on the value together with the development of BIM implementation. Consequently, the observed trend is not a surprise and was fully anticipated, especially in the period 2007-2012 when the BIM implementation exploded as mentioned in the introduction.

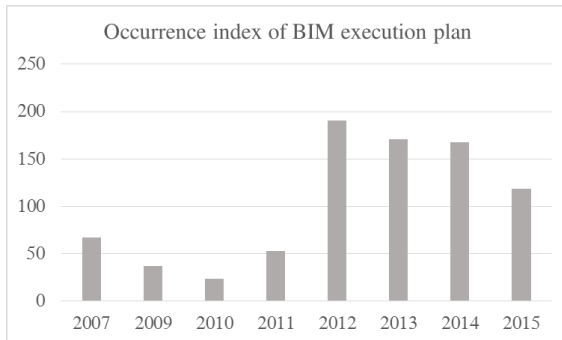


Fig. 2. Occurrence index per year histogram, for the "BIM Execution plan" term

The recognized trends of elaborating the common BIM uses within industry Standards, in theory should be in compliance with the common practice of BIM industry, which proved to be a case for the example of BIM execution plan.

This paper is the starting point of the wider research that should check, confirm or deny the compatibility of the trends obtained from BIM Standards with the current industry practice that will be examined in the following stages of the research.

## 5. REFERENCES

- [1] National Institute of Building Sciences buildingSMART alliance, "National BIM Standard – United States Version 1", Part 1, US, 2007, pp. 21.
- [2] National Institute of Building Sciences buildingSMART alliance, "National BIM Standard – United States Version 2", US, 2012, Ch. 3, pp. 3.
- [3] Dodge Data & Analytics, "Smart Market Report", *Measuring the Impact of BIM on Complex Buildings*, 2015, pp. 6.
- [4] B. Hardin, & D.McCool, "BIM and construction management: proven tools, methods, and workflows", 2<sup>nd</sup> Edition, John Wiley & Sons, 2015.
- [5] R.Kreider, J.Messner, and C.Dubler. "Determining the frequency and impact of applying BIM for different purposes on projects." *Proceedings 6th International Conference on Innovation in Architecture, Engineering and Construction (AEC)*. 2010., pp.5-9.
- [6] Construction, McGraw Hill. "Smart Market Report." *The Business Value of BIM for Construction in Major Global Markets: How Contractors Around the World Are Driving Innovation With Building Information Modeling*, 2014, pp. 32-34
- [7] E. Garfield, "Scientography: Mapping the tracks of science", *Current Contents: Social & Behavioural Sciences*, 1994, Vol. 7, No. 45, pp. 5–10.
- [8] K.W. Church and P. Hanks. "Word Association Noms, Mutual Information, and Lexicography", *Proceedings of the 27th Annual Conference of the Association for Computational Linguistics*, 1989, pp. 22–29.
- [9] J. Washtell and K. Markert "A Comparison of Windowless and Window-Based Computational Association Measures as Predictors of Syntagmatic Human Associations", *Proceedings of the 2009 Conference on Empirical Methods in Natural Language Processing*, Singapore (August), 2009, pp. 628–37.
- [10] O.P.Damani and S.Ghonge. "Appropriately Incorporating Statistical Significance in {PMI}", *Proceedings of the 2013 Conference on Empirical Methods in Natural Language Processing*, 2013, pp. 163–69.
- [11] L.R. Dice, "Measures of the Amount of Ecologic Association Between Species" *Ecology*, 1945, Vol. 26, No. 3, pp. 297–302.
- [12] T. Dunning, "Accurate Methods for the Statistics of Surprise and Coincidence", *Computational Linguistics* 19, 1993, pp. 61–74.