BIM Search Engine: Exploiting Interrelations between Objects when Assessing Relevance

Peter Demian[1]

1) Reader in Building Information Management, School of Architecture, Building and Civil Engineering, Loughborough University, Leicestershire, UK. Email: P.Demian@lboro.ac.uk

Abstract:
An increasing amount of information is packed into Building Information Models (BIMs), with the 3D geometry intended to serve as a central index leading to other information. The Three-Dimensional Information Retrieval (3DIR) project investigated information retrieval from such environments, with the aim of developing a search engine for searching and retrieving information from a building model. Here, the 3D model of the building can be exploited to formulate queries, compute the relevance of information items to a given query, and visualize search results. The focus of this paper is the computing of relevance. Literature in BIM/CAD and information retrieval was reviewed as a precursor to developing the search engine. Based on earlier research which identified the needs and aspirations of the users of BIMs, a graph theoretic formulation is proposed here to inform the emerging retrieval mechanisms of a BIM search engine. This formulation distinguishes between 3D and textual information in the model (the vertices in the graph), and between different types of relationships linking model objects (the edges in the graph). The value is tested of exploiting a 3D object's relations to other 3D objects when assessing that object's relevance to a query. For example, if a user is searching for "glazing door internal wall", such a holistic/contextual search would rate the relevance of a "glazing panel" object more highly if it was touching "internal wall" or "door" objects. This notion was tested using an Autodesk Revit model from an architectural industry partner, augmented with the 3DIR search toolset. The model contained just under 7k 3D elements. Relationships between the objects were either *hosting*, *touching* or *intersecting* relationships. A comparison of the retrieval performance for a handful of test queries with and without this holistic/contextual search function does not decisively highlight the benefit but demonstrates the promise of this approach particularly for more complex multiple search term queries, as well as the value of the underlying graph theoretic formulation for studying and developing such systems.

Keywords: Building Information Modelling, Search Engine, Information Retrieval, 3DIR.

1. INTRODUCTION

In building modelling environments, information is increasingly being crammed into 3D building and product models. This is particularly true given the rise of Building Information Modelling (BIM). For example, in a group discussion convened by *Construction Manager* magazine in 2017, a panel of industry experts all agreed that the exponential rise of construction information was becoming impossible to manage (Kenny 2017). The Three-Dimensional Information Retrieval (3DIR) project investigated information retrieval from these environments, where information or documents are linked to a 3D building model. In these situations, the 3D visualisation or 3D geometry of a building can be exploited when formulating information retrieval queries, computing the relevance of information items to the query, or visualizing search results. This paper focuses on the computation of relevance of items in the BIM to a search query when using keyword searches to retrieve information.

Information retrieval is associated with documents, and a critic might argue that documents are relics from the pre-BIM age that are no longer relevant in the era of BIM. However, the challenge of information retrieval is pertinent whether we are dealing with documents which are coarse grains of information or building object parameters/attributes as finer grains of information. Demian and Fruchter (2005) demonstrated that traditional retrieval computations can be applied with good results to 3D building models where textual or symbolic data are treated as very short documents. In this sense, it is almost a question of semantics whether the information being retrieved comes from object properties embedded in the BIM, or from external documents linked to the BIM. The challenge remains of retrieving non-geometric or textual information, and this is the focus of 3DIR.

This paper describes recent developments of the 3DIR project whose aim was to improve information retrieval when retrieving information or documents linked to a 3D artefact, or retrieving non-geometric information embedded in the model of the artefact. It proposes a formulation based on graph theory as a useful theoretical lens for research and software development for information retrieval from BIMs. The central objective was to improve retrieval performance when searching building models by exploiting links between 3D building elements.

2. RELATED WORK

Building design, construction and operation are information intensive activities. For example, even over a decade ago in the UK construction industry, on average, one computer-aided design (CAD) document was produced for every 9 m^2 of building floor space (Gray and Hughes 2001). Several researchers (Leslie, 1996; Veeramani and Russell, 2000; Ugwu, 2005) have reported the problem of "information overload" in the construction sector.

BIMs are following this general trend and becoming more information-rich. Regarding volumes of information specifically in BIMs, Demian and Walters (2014) identified BIM platforms as a particularly favourable communication medium in construction, compared to extranets, email and Enterprise Resource Planning systems. Charalambous et al. (2013) reported the advantages of BIM over documents and extranets. Although no absolute measures of the quantities of information were found, the implication from studies such as those is that BIMs are increasingly information-rich.

Information retrieval techniques have been used in construction to retrieve reusable designs (Demian and Fruchter 2005). Beyond text, Brilakis and Soibelman (2008) automatically identify particular features in construction site photographs with a view subsequently to using information retrieval techniques to manage collections of photographs. Bridging textual and geometric content, Caldas et al. (2002) propose techniques for automatically classifying construction documents based on project CAD components. Lin and Soibelman (2009) augment standard information retrieval techniques with formal representations of domain knowledge to improve the performance of a search engine for online product information. Rezgui (2006) similarly uses domain knowledge to formulate an ontology that informs the indexing and retrieval of construction content. These studies demonstrate how standard retrieval computations can be complemented when applied to building design and construction.

None of the studies encountered in the literature specifically exploit 3D data for information retrieval. This approach lies at the intersection of two academic fields: (1) BIM and CAD and (2) information retrieval.

2.1 BIM/CAD

The state of the art in digital content management in building design and construction projects is being transformed by the emergence of Building Information Modelling (Eastman et al. 2011). Whereas CAD models classically attempted to model the geometry of buildings or building components in two or three dimensions (e.g. Eastman 1999), Building Information Models include non-geometric content as well. This content includes the non-geometric attributes of physical building components (such as the cost of a component) as well as non-geometric entities. For example, Building Information Models can include entities to model the processes of design (Austin et al. 2000) and construction (Koo and Fischer 2000) and the organizations (i.e., teams and individuals) that execute those processes (Kunz et al. 1998). In addition, BIM is not limited to the design and construction phases but can be extended to cover the entire life cycle of constructed facilities, from briefing/programming, through design, to facilities management and even disposal. In the context of the 3DIR project, it is noteworthy that, although as noted above, CAD and BIMs nowadays include both geometric and non-geometric information, the geometric 3D model of the building is central, and is often expected to serve as a visual index that leads to the additional non-geometric content. This approach often fails, because such systems do not exploit human abilities in spatial cognition and visual memory. Non-geometric content does not leave enough *information scent* (Pirolli and Card 1999) in the geometric CAD model that enables the information forager to find it. This concept serves as an important point of departure for the 3DIR project.

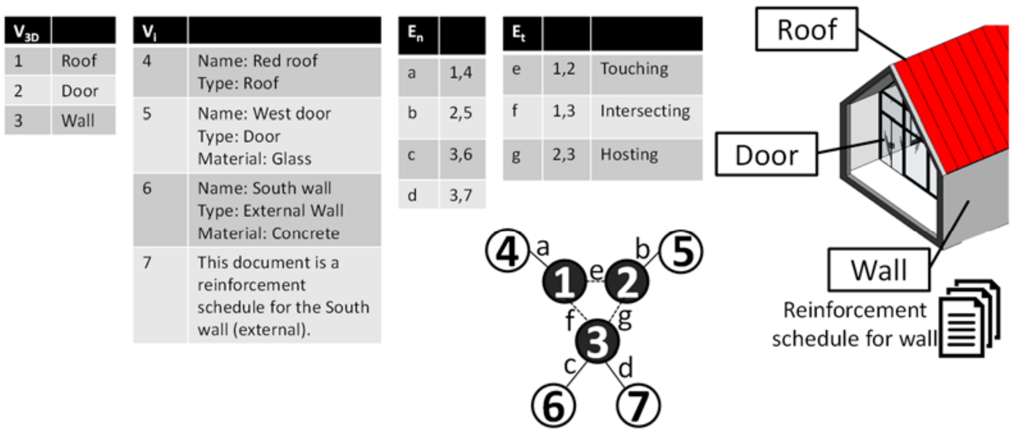
2.2 Information Retrieval

Information retrieval (IR) is concerned with systems that help users to fulfil their information needs. IR computations can quantify the relevance of information items based on user queries (Dominich 2008). Demian and Fruchter (2005) demonstrated that traditional IR techniques could be applied to retrieve information from BIMs and product models; the semantic information attached to 3D objects could be treated as very short documents and standard text document computations employed, giving reasonable retrieval results. In that case, information retrieval was used to retrieve reusable designs from a large "corporate memory" of previous construction projects. In this current research, information retrieval is used more generically to locate relevant information in an imperforation-rich BIM. As noted in the introduction to section 2 and in section 2.1, information retrieval has recently been applied in managing the vast volume of information accumulated in building design, construction and operation.

3. A GRAPH THEORETIC FORMULATION OF INFORMATION LINKED TO 3D MODELS

Graph theory provides a useful theoretical lens for studying information-rich 3D models and retrieving information in these environments. This theoretical lens can inform software research and development in this area. A *graph* in this context is a series of vertices connected by edges. Each edge joins exactly two vertices. Any graph X can be modelled mathematically by listing its set of vertices V(X) and set of edges E(X) (Aldous and Wilson (2003) give an introduction to graph theory). In the case of an information-rich BIM, it is possible to distinguish between the set of 3D vertices V_{3D}, which are the 3D objects in the model, and information vertices V_i, which are linked information items, whether properties of the 3D objects treated as short documents or linked full text documents.

Similarly, for the edges in the graph theoretic formulation of a 3D model, it is possible to distinguish between two types of edges. The first more obvious type of edge is the edge joining a 3D object to one of its properties, i.e. an edge between a vertex in set V_{3D} and a vertex in set V_i. The set of this natural type of edge can be called E_n. It arises simply from the fact that 3D objects and their properties (or linked documents) are modelled as separate (but linked) objects. The second, more subtle, type of edge is that edge which encodes some topological relationship between two 3D objects (Demian et al 2016), i.e. an edge joining two related V_{3D} vertices. The set of this type of *topological* edge can be called E_t. Such edges and the topological relations they model are one of the focal points of recent developments in the 3DIR project. (If it were not for the edges in set E_t, the emerging graph would consist of two disjoint sets of 3D and information vertices, with each natural edge in E_n connecting an item in one set to an item in the other, i.e. a *bipartite graph*.) Figure 1 gives an example of this formulation for a simple 3D model.



V_{3D}	
1	Roof
2	Door
3	Wall

V_i	
4	Name: Red roof Type: Roof
5	Name: West door Type: Door Material: Glass
6	Name: South wall Type: External Wall Material: Concrete
7	This document is a reinforcement schedule for the South wall (external).

E_n	
a	1,4
b	2,5
c	3,6
d	3,7

E_t		
e	1,2	Touching
f	1,3	Intersecting
g	2,3	Hosting

Figure 1. An example graph theoretic formulation of a simple 3D model consisting of a Roof, a Door and a Wall (all three objects having properties, and the Wall being linked to an external document)

In this research, a *topological* edge can represent any relationships between 3D building elements in a model. These relationships might be strictly topological and concerned with interior/boundary/exterior of 3D components, more general spatial/directional relationships (Borrmann and Rank, 2009), or even relationships as they occur in a very general semantic sense, albeit linked to spatial topology (Lin, 2013). In this sense, any two objects in a model sharing the same attribute (for example: two components supplied by the same manufacturer) can be said to be related. As an initial proof of concept, only the *hosting*, *touching* and *intersecting* relationships are represented as topological edges.

4. METHOD

A logical consequence of the formulation in Figure 1 is that topological relationships (or edges in E_t) can be exploited for information retrieval. To test this, the 3DIR Revit app was used with a Revit model provided by a 3DIR project industry partner. The model was of a large floor-plus-three storey office development, as shown in Figure 2. When first used, the 3DIR tool creates an index of all the text terms (form parameters and linked documents) in the model. Each entry in the index is essentially a V_i vertex. Each V_i object in the index has a reference to a V_{3D} object (i.e. following the E_n edge), as well as weights of textual terms, calculated based on term frequency using the Apache Lucene library (Lucene 2018).

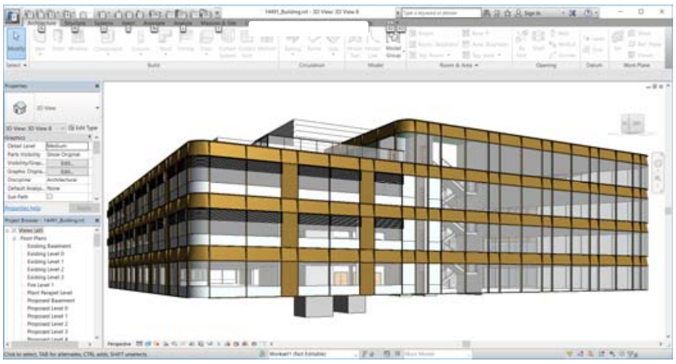


Figure 2. The building model of an office development used for the retrieval experiments reported in this paper

At the indexing stage, 3DIR identifies a list of related *Neighbours* for each V_{3D} object referenced in the index. For the 3D object being indexed, those Neighbours are either other 3D objects *hosting* or *hosted by* the object in question, or objects *touching* or *intersecting* the object in question (Demian et al. 2016). The index of V_i objects is exported to a text file, and a spreadsheet application used to create an index of V_{3D} objects. Each 3D object in that index has links to its constituent parameters (V_i objects), as well as its Neighbouring V_{3D} objects. Spreadsheet functionality is used to create a list of *Neighbours-of-Neighbours* for each 3D object.

Along with displaying search results on screen, the 3DIR tool was modified to export search results to a text file when a search is conducted. As standard 3DIR functionality, the tool retrieves V_i objects which contain the query text term(s), and assigns a score based on term frequency. This standard scoring of the V_i object is referred to as:

$S(V_i)$

For each 3D object in the index, a score $S(V_{3D})$ is calculated, as a simple mean of the $S(V_i)$ scores in relation to the same query of the constituent parameters of the object in question.

$$S(V_{3D}) = \overline{S(V_i)} \quad (1)$$

where $S(V_i)$ are the scores of the constituent V_i parameters in object V_{3D} and the overline indicates the mean of these scores.

In the same way (by averaging parameter V_i scores), a score is calculated for each *Neighbour* and *Neighbour-of-Neighbour* object linked to the 3D object in the index. This score represents the relevance to the query of that *Neighbour* or *Neighbour-of-Neighbour* object. The mean pf all *Neighbour* and *Neighbour-of-Neighbour* objects is calculated for each object in the index:

$$S(V_{3D\text{-}N}) = \overline{S(V_{3D\text{-}N})} \quad (2)$$

$$S(V_{3D\text{-}NN}) = \overline{S(V_{3D\text{-}NN})} \quad (3)$$

where $S(V_{3D\text{-}N})$ and $S(V_{3D\text{-}NN})$ are the scores of the 3D object's *Neighbours* and *Neighbours-of-Neighbours* respectively calculated as in Equation (1), and the overline indicates the mean of these scores.

Armed with these computations, an attempt was made to improve the standard 3DIR scoring (i.e. the ranking of search results) as in Table 1:

Table 1. Relevance measures tested

Name	Equation	Rationale
"V_i" Relevance	$S(V_{3D})$	Standard V_i Lucene score
"V_i+V3D" Relevance	$C_1S(V_{3D}) + C_2S(V_{3D})$	Also accounting for relevance of 3D object as a whole
"V_i+V_{3D}+N" Relevance	$C_3S(V_{3D}) + C_4S(V_{3D}) + C_5S(V_{3D\text{-}N})$	Also accounting for relevance of *Neighbours*
"V_i+V_{3D}+N+NN" Relevance	$C_6S(V_{3D}) + C_7S(V_{3D}) + C_8S(V_{3D\text{-}N}) + C_9S(V_{3D\text{-}NN})$	Also accounting for relevance of *Neighbours-of-Neighbours*

In Table 1, C_1-C_9 are constants. Those were heuristically set by the researcher as follows:

C_1=0.7, C_2=0.3
C_3=0.5, C_4=0.3, C_5=0.2
C_6=0.4, C_7=0.3, C_8=0.2, C_9=0.1

The four relevance measures were tested on three variants on a simple single keyword query: *glazing*, *glazed*, and *glaz**. This was essentially a test of 3DIR's simple keyword search. 3DIR does not currently use stemming but does accept text wildcards. In addition to the three simple queries, a more complicated multiple search term query was also tested: "*internal wall door glaz**" (query entered without quotes, disregarding term order). For each query, a list of relevant V_i objects was identified (with the help of the spreadsheet word search to filter the large number of V_i objects, but exercising extensive human judgment of relevance to further draw out relevant items from the shortlist). This was undertaken by the researcher on a dichotomous relevant/irrelevant basis.

The model used (Figure 2) contained just under 7k V_{3D} objects, which translated into just over 20k V_i objects (i.e. an average of just under three indexed parameters per 3D model object). Of the 6833 3D objects, 5167 had *Neighbours*. Of those with *Neighbours*, the average number of *Neighbours* was 2.52. A total of 606 3D objects had *Neighbours-of-Neighbours*. Of those with *Neighbours-of-Neighbours*, the average number was 8.72.

5. RESULTS

5.1 Single Keyword Queries

Table 2 summarises the results for the single keyword queries, Queries 1a, 1b and 1c. The intention behind those test queries was to test the basic search functionality of 3DIR and establish a robust research protocol. Search performance is usually assessed using measures of *Recall* (the proportion of relevant items which are retrieved) and *Precision* (the proportion of retrieved items which are relevant). As search hits are ranked by their relevance score, more relevant search results can be considered retrieved, giving Precision at increasing Recall levels. Because of the unusual retrieval results, it is not informative in this case to plot the standard Precision-Recall curves.

Table 2. Single keyword query test results

Query →	Query 1a	Query 1b	Query 1c
Query Terms	*glazing*	*glazed*	*glaz**
Relevant V_i items (according to human expert)	9	3092	3101
V_i items retrieved by 3DIR	8	250 (3DIR maximum)	250 (3DIR maximum)
"V_i" Relevance performance	3DIR successfully retrieved 8 of the 9 relevant items. The precision was 1 at all recall levels.	3DIR has a maximum of 250 search hits, which means the maximum possible recall is 0.08, and this was achieved using this basic relevance measure. Precision was 1 at all levels.	As expected, the set of relevant items for this query is the union of the relevant sets for Queries 1a and 1b. The results were roughly the same as for Query 1b.
"V_i+V3D" Relevance performance	The ranking of search hits did not change from above.	Although there were minor differences to the items retrieved and their rankings, the maximum of 250 search hits and the large number of relevant items meant that maximum precision was still 0.08, again with no irrelevant items retrieved.	Roughly the same as for Q1b.
"V_i+V_{3D}+N" Relevance performance	The ranking of search hits did not change from above.	Same as above: slightly different search hits and ranking, but no change in recall and perfect precision.	Roughly the same as for Q1b.
"V_i+V_{3D}+N+NN" Relevance performance	The ranking of search hits did not change from above.	Same as above: slightly different search hits and ranking, but no change in recall and perfect precision.	Roughly the same as for Q1b.

The results from the above demonstrate that the holistic/contextual search measures add little value in terms of retrieval performance for simple single keyword queries. It is noteworthy that for Queries 1b and 1c, the search rankings did change slightly when the holistic/contextual relevance measures were applied, but improvements in performance could not be measured because of the upper limit of 250 in the 3DIR search hits and the large number of relevant items for these queries.

5.2 Multiple Search Term Queries

The holistic/contextual search function is expected to be useful particularly for multiple search term queries. For example, if a user is searching for "glazing door internal wall", such a holistic/contextual search is intended to rate the relevance of a "glazing panel" object more highly if it is touching "internal wall" or "door" objects. Table 3 summarises the retrieval results for the query "*internal wall door glaz**" (query entered without quotes). As the holistic/contextual search measures did not affect the retrieval of the results, only their ranking, the Precision and Recall measures are the same for all four relevance measures. Precision-Recall curves did not clearly differentiate between the four measures. Table 3 attempts to distinguish the performance of the four measures by noting the top ranking *irrelevant* search result and the bottom ranking *relevant* search result.

Table 3. Multiple search term query test results

	Query 2			
Query Terms	*internal wall door glaz**			
Relevant V_i items (according to human expert)	238			
Items retrieved	250 (3DIR maximum)			
Maximum Recall	0.567			
Average Precision (averaged over 250 retrieved search hits)	0.871			
Relevance Measure → Performance Criterion ↓	**"V_i" Relevance**	**"V_i+V3D" Relevance**	**"V_i+V_{3D}+N" Relevance**	**"V_i+V_{3D}+N+NN" Relevance**
Top Rank of Irrelevant Retrieved Search Hit	134	134	134	134
Bottom Rank of Relevant Retrieved Search Hit	150	250	250	250

From the results, the contextual/holistic measures do not deliver any observable improvement in search performance. Indeed, the bottom row of the table shows that for all three contextual/holistic measured, a relevant search result is pushed down to the lowest rank in the 250 search hits.

6. DISCUSSION

The results highlight the need to remove the 250-limit on search hits in 3DIR. For the single search term queries (particularly Queries 1b and 1c), measures of Precision and Recall might not be able to distinguish subtleties in retrieval performance when the number of relevant items is vastly greater than the number of retrieved results. The 250-limit on search hits imposed by 3DIR might have obscured the benefits of the contextual/holistic relevance measures.

For the multiple search term query, the assessment of the retrieval performance is more nuanced. For Query 2, the number of relevant items was just within 3DIR 250-limit. It was hoped that the contextual/holistic relevance measures would push the irrelevant retrieved results to the bottom of the list; this did not clearly occur. More sensitive indicators of retrieval and ranking performance will be explored in future research. By observing the distribution of relevance scores of retrieved results, it is clear that the holistic/contextual measures played an important role in distributing relevance scores. The raw $S(V_i)$ scores are based on term frequency. With the limited text in 3D object parameters, most of the retrieved objects had identical scores. The effect of the contextual/holistic measures in scattering those scores based on the relevance of other related parameters and 3D objects is undoubtedly beneficial. However, more sensitive indicators are needed to assess this benefit.

A significant improvement in retrieval performance remains elusive, but the mechanisms proposed here are a significant first step in developing a dedicated search engine for BIMs. This work is an important extension of earlier work by Demian and Fruchter (2005) and demonstrates that the sparse texts in commercial BIMs can be used for retrieval. The graph theoretic formulation presented is a possible theoretical lens for developing and assessing such BIM search engines.

7. CONCLUSIONS

The 3DIR project investigates information retrieval from environments such as BIM, where information is linked to a 3D artefact. This paper presents a graph theoretic formulation for studying and developing BIM search engines. This formulation distinguishes between 3D and textual information in the model (the vertices in the graph), and between different types of relationships linking model objects (the edges in the graph). A logical consequence of this formulation is that relationships between 3D objects can be used when measuring the relevance of a 3D object to a search query. This paper presents three such contextual/holistic search measures which, in addition to the retrieved parameter (information), consider *all* of an object's parameters, as well as the relevance of other 3D objects related to the object in question.

Standard retrieval performance assessments using Precision and Recall could not highlight the benefit of those contextual/holistic measures above the standard practice of considering each object in isolation. However, those contextual/holistic measures played an important role in distributing otherwise clustered relevance scores, thereby allowing a more informative ranking. More sensitive retrieval performance indicators are needed to measure this benefit.

A number of limitations can be noted. The relevant items for each query were identified by a single researcher. In a more rigorous approach, multiple human experts would independently identify relevant items, and inter-rater agreement would be reported. For the feasibility of the spreadsheet analysis, the number of Neighbours of each 3D object was limited to 20. Fifty-two of the 6833 3D objects in the model had 20 Neighbours in the index and it is possible that they actually had more but that the number of Neighbours was truncated.

Future research can fine-time the relevance measures by adjusting the values of the constants in the equations in Table 1, perhaps by using a genetic algorithm to optimise retrieval performance. Informal sensitivity analysis shows that even minor reconfigurations of those constants significantly affects the ranking of the search results. Further research can also explore other possible relationships between 3D objects beyond the hosting, touching and intersecting relationships used here.

ACKNOWLEDGMENTS

The 3DIR add-in for Revit is developed by Graitec UK Ltd and is available as a free download from the Autodesk App Store. Mr Neil Sutton of Graitec provided guidance on the 3DIR source code. Mrs Victoria Boyle of Walter's Bakery provided support in analysing the search results in Microsoft Excel. Colleagues Dr Russell Lock and Dr Panayiotis (Panos) Balatsoukas looked at drafts of this paper; shortcomings are the author's responsibility.

REFERENCES

Aldous, J. M., & Wilson, R. J. (2003). *Graphs and applications: an introductory approach*. Springer Science & Business Media.

Austin, S., Baldwin, A., Li, B., & Waskett, P. (2000). Analytical design planning technique (ADePT): a dependency structure matrix tool to schedule the building design process. *Construction Management & Economics*, 18(2), 173-182.

Brilakis, I. K., & Soibelman, L. (2008). Shape-based retrieval of construction site photographs. *Journal of Computing in Civil Engineering*, 22(1), 14-20.

Borrmann, A., & Rank, E. (2009). Specification and implementation of directional operators in a 3D spatial query language for building information models. *Advanced Engineering Informatics*, 23(1), 32-44.

Caldas, C. H., Soibelman, L., & Han, J. (2002). Automated classification of construction project documents. *Journal of Computing in Civil Engineering*, 16(4), 234-243.

Charalambous, G., Thorpe, T., Demian, P., Yeomans, S. G., Doughty, N., & Peters, C. (2013). Collaborative BIM in the Cloud and the Communication tools to support it. *Proc. 30th CIB W78 Int. Conference on Applications of IT in the AEC industry*, Beijing, China, 9-12 October 2013, pp. 58 - 67.

Demian, P., & Fruchter, R. (2005). Measuring relevance in support of design reuse from archives of building product models. *Journal of Computing in Civil Engineering*, 19(2), 119-136.

Demian, P., & Walters, D. (2014). The advantages of information management through building information modelling. *Construction Management and Economics*, 32(12), 1153-1165.

Demian, P., Ruikar, K., Sahu, T., & Morris, A. (2016). 3DIR: exploiting topological relationships in three-dimensional information retrieval from BIM environments. *Proceedings of 2016 16th International Conference on Computing in Civil and Building Engineering (ICCCBE2016), Osaka, Japan, 6-8 July 2016.*

Dominich, S. (2008). *The modern algebra of information retrieval*. Heidelberg: Springer.

Eastman, C. M. (1999). *Building product models: computer environments, supporting design and construction*. CRC press.

Eastman, C., Teicholz, P., Sacks, R., and Liston, K. (2011). *BIM Handbook: A Guide to Building Information Modeling for Owners, Managers, Designers, Engineers, & Contractors* (2nd ed.). John Wiley & Sons.

Gray, C., & Hughes, W. (2001). *Building design management.* Routledge.

Kenny, J. (2017). Have we reached information overload? *Construction Manager* magazine, November 2017. Retrieved from http://www.constructionmanagermagazine.com/insight/round-table-have-we-reached-information-overload/, accessed on 15 January 2018.

Koo, B., & Fischer, M. (2000). Feasibility study of 4D CAD in commercial construction. *Journal of construction engineering and management*, 126(4), 251-260.

Kunz, J. C., Christiansen, T. R., Cohen, G. P., Jin, Y., & Levitt, R. E. (1998). The virtual design team. *Communications of the ACM*, 41(11), 84-91.

Lin, C. J. (2013). Smart Spatial Ontology: Bridging Semantic Ontology to Spatial Topology. *Computer-Aided Design and Applications*, 10(3), 489-497.

Lin, K. Y., & Soibelman, L. (2009). Incorporating domain knowledge and information retrieval techniques to develop an architectural/engineering/construction online product search engine. *Journal of Computing in Civil Engineering*, 23(4), 201-210.

Leslie, H. G. (1996). Strategy for information in the AEC industry. *INCIT 96 Proc.: Bridging the Gap*.

Lucene 2018. Apache Lucene. Retrieved from https://lucene.apache.org/, accessed 30 January 2018.

Pirolli, P., & Card, S. (1999). Information foraging. *Psychological review*, 106(4), 643.

Rezgui, Y. (2006). Ontology-centered knowledge management using information retrieval techniques. *Journal of Computing in Civil Engineering*, 20(4), 261-270.

Ugwu, O. O. (2005). A service-oriented framework for sustainability appraisal and knowledge management. *International Journal of IT-Construction ITCon* (Electronic Journal), 10, 245-263.

Veeramani, D., & Russell, J. (2000). Preparing the AEC industry for the knowledge economy. *In Berkeley-Stanford Construction Engineering and Management Workshop: Defining a Research Agenda for AEC Process/Product Development in 2000 and Beyond*.