INTELLIGENT RETRIEVAL
Open Journal

Review

A Natural-language-based Approach to Intelligent Data Retrieval and Representation for Cloud BIM

Jia-Rui Lin; Zhen-Zhong Hu*; Jian-Ping Zhang; Fang-Qiang Yu
Department of Civil Engineering (Tsinghua University), Haidian District, Beijing 100084, PR China

****Corresponding author***
Zhen-Zhong Hu
Department of Civil Engineering (Tsinghua University), Haidian District, Beijing 100084, PR China

Article information
Received: September 30th, 2023; **Revised:** December 30th, 2023; **Accepted:** January 5th, 2024; **Published:** January 27th, 2024

Cite this article
Lin J-R, Hu Z-Z, Zhang J-P, Yu F-Q. A natural-language-based approach to intelligent data retrieval and representation for cloud BIM. 2024; 2(1).
doi: https://doi.org/10.70705/ppp.ir.2024.v02.i01.pp4-10

ABSTRACT

A building's BIM (Building Information Model) grows in size when data from many fields is integrated over an AEC project's lifespan. Because of this issue, consumers will have a hard time getting the information they want on a mobile device, which has limited area for interaction. Employees who aren't BIM professionals or who don't have a deep understanding of Industry Foundation Classes (IFC) will have it much worse. An intelligent data retrieval and representation strategy for cloud BIM applications based on natural language processing was developed to increase the value of the huge data of BIM. To begin managing the massive amounts of BIM data, methods for storing data and speeding up queries using the widely used cloud-based database were investigated. Then, to describe the user's expectations in a natural-language phrase, the notions of "keyword" and "constraint" were suggested to capture the important objects and their specifications. The International Framework for Dictionaries (IFD) allows for the mapping of keywords and restrictions to entities or characteristics inside the International Financial Center (IFC). Data retrieval and analysis were made possible by route finding in a graph built from the IFC schema, which established the link between the user's demand and the IFC-based data model. Lastly, the structure of the obtained data was used to depict the analyzed and summarized outcomes of the BIM data. Using data acquired during the construction of Kunming Airport's terminal—the biggest single structure in China—a prototype application was created to test the suggested strategy. The following were shown by the case study: (1) user needs and data are linked; (2) cloud-based BIM allows for automated retrieval and aggregation of user-related data; and (3) data is delivered in an appropriate format for both visual representation and detailed reporting. Users may greatly benefit from information requests using this technique, and BIM's value will be boosted..

Keywords
Building information model (BIM); Industry foundation classes (IFC); International financial center (IFC).

INTRODUCTION

According to Azhar (2011), the building information model (BIM) promotes the integration of all project stakeholders' responsibilities and is a new paradigm in the AEC business. According to studies conducted by McGraw-Hill Construction and Pike Research, more and more businesses in the construction sector, as well as individual projects, are embracing building information modeling (BIM) as a tool for technical assistance. The rate of BIM usage among AEC organizations will keep going up, according to research (Jennifer, 2012).

All through an AEC project's lifespan, people have worked to further the development, dissemination, and incorporation of building information modeling (BIM), in addition to teamwork and knowledge management (KM). For BIM-based interoperability and cooperation, two distinct approaches were taken:

Ontology-and semantic web-based methods: Anumba et al. (2008) found that explicit domain ontologies help with decision-making and collaboration in the design and construction process by creating a shared understanding of the problem-solving domain across disciplines. The current body of research demonstrates promising advances in the study of ontologies and information sharing, which should be the first focus of our proposal. According to Sørensen et al. (2010), who analyzed existing ontologies, there should be ideas for integrating virtual models and construction components using ontologies so that information may be shared in building operation management and construction projects. Additionally, Kong et al. (2005) laid forth a method for making online building product catalogs work with one another.

Industry Foundation Classes (IFC)-based Methods: The International Alliance of Interoperability (IAI, also known as BuildingSMART) proposed IFC as a solution to the inadequate interoperability problem (Laakso et al., 2012). IFC is a complete model of the building along with the rules and protocols that define building data. An analysis of the theoretical and practical needs for using

Additionally, the use of an IFC-based BIM-server as a platform for interdisciplinary cooperation was covered (Singh et al., 2011). Eastman et al. (2009) laid out the steps to take in order to create an Information Delivery Manual (IDM) that would allow for accurate data transmission via IFC. Also offered for bridge management was a mobile model based on IFC and a 4D model (Hammad et al., 2006).

The number of stakeholders, the amount of data integrated into the BIM, and the size of the data will all rise in tandem with the number of suggested ways and tools for cooperation and interoperability. Users would have challenges in accessing the information they really want due to this predicament. The information flood necessitates that even developers and professionals in building information modeling (BIM) software use several manipulations (Jasper, 2011). Users will have a harder time finding relevant, focused information, and more time will be required to learn the program as BIM data sizes and capabilities continue to grow. The time-consuming and error-prone human processing might be replaced by automated data extraction and processing (Boukamp and Akinci, 2007). The use of mobile terminals to access the web is also on the rise; however, these devices have far smaller displays than desktops and laptops, necessitating an easy-to-navigate interface that makes the most efficient use of the space available. When it comes to making information requests, voice assistants like Cortana, Google Now, and Siri have established a new standard. Experts and novices alike need efficient and adaptable data capture from massive BIM systems, and practical ways to visualize or represent data will increase BIM's value.

An intelligent system for retrieving and representing BIM data on the cloud is presented in this research. The first step in working with BIM data was to set up a database on the cloud. The next step was to start processing user input using natural language processing (NLP). This included extracting keywords and then mapping them to entities or properties in IFC. This procedure set the stage for data retrieval in a BIM data model organized according to IFC standards. Lastly, according to user expectations, the analyzed data was displayed in a variety of ways, including charts, tables, animations, or a mix of these.

In Section 2, we take a look at the relevant literature and talk about the potential for intelligent BIM data representation and retrieval in BIM applications. on Section 3, we lay down some guidelines for building the database and storing data on the cloud. In Section 4, the International Framework for Dictionaries (IFD)-supported techniques for keyword extraction using natural language processing (NLP) and its mapping mechanism are discussed. Section shows the process of data retrieval, analysis, and representation. Section 6 presents a real-life example of Kunming Airport, China's biggest single structure, to confirm if the strategy is sound. Section 7 concludes the discussion and moves on to future research.

2 LITERATURE REVIEW

1.1 Related works

1. Interoperability based on IFC and IFD

The National Building Information Modeling Standards Committee of the United States explains that BIM encompasses information throughout the life cycle of a facility, multidisciplinary collaboration and decision making (NBIMS, 2012). IFC is a data protocol proposed for sharing information between different software. Backed by years of development and improvement, the IFC standard is widely accepted for information management. An IFC-based system (Lee et al., 2003) was proposed for design information management, and some IFC-based BIM servers (Beetz et al., 2010; Jørgensen et al., 2008; Kang and Lee, 2009) were also developed for information sharing, extraction and integration. Based on BIMserver.org (Beetz et al., 2010), an open query language for BIM called supporting BIMQL was developed (Mazairac and Beetz, 2013), providing flexible data retrieval interface that is domain specific and platform independent. An open repository for IFC model analysis to facilitate the interoperability of building information was also presented (Amor and Dimyadi, 2010). Meanwhile, experiences with issues of model-based interoperability in exchanging BIM between various tools were reported (Steel et al., 2012). Research on extending IFC by eXtensible Markup Language (XML), and information exchange with the help of model view (Eastman et al., 2009; Fu et al., 2006) or IDM (Zhang et al., 2012; Kim et al., 2010) were also discussed. Therefore, IFC, with feasible extending mechanism and numerous related tools, lays a solid foundation for BIM interoperability.
However, as stated in the IFD white paper (BuildingSMART, 2008), a controlled vocabulary of construction terminology, which IFC lacks, is essential to support data exchange. With the support of multi-language terminology, IFD provides a mapping method from concepts to IFC entities and attributes (Zhang et al., 2012), supporting distinguishing concepts from specific linguistic instances. Shayeganfar (et al., 2008) conducted a case study on how to implement an IFD library using semantic web technologies, which bridge the gap between BIMs and web services. Thus, with the support of IFD library and semantic web, terminology or its synonym in a specific language can be mapped to the entity in a data schema like IFC.

2. Cloud and BIM

The use and sharing of BIM data generated by standalone systems becomes more challenging as BIM grows in size; however, cloud computing will facilitate data manipulation (Chuang et al., 2011). Examination performed Some businesses are already working on cloud BIM systems, according to research by Redmond (Redmond et al., 2012), and this is because cloud computing improves data visualization and sharing. First, cloud-based building information modeling (BIM) increases the likelihood of interdisciplinary collaboration (Redmond et al., 2012). Second, web services provide convenient and adaptable access to data and resources. Third, an extensible platform makes it feasible to manage massive amounts of data.
Many cloud-based BIM applications have already been launched, in-

cluding BIM+, BIMcloud, and BIM 360 Glue. Data visualization, collaboration, and sharing are the key focuses of these products. Extending goods in data storage, manipulation, and display becomes challenging due to the absence of information on data storage and access. It is worthwhile and doable to investigate how IFC models may be saved and shared using these NoSQL databases, since constructing cloud computing platforms that use non-relational, distributed databases like MongoDB is now trendy (Membrey et al., 2010).

3. Data retrieval and visualization

The information retrieval process for an online product was made easier with the use of domain expertise to aid in the acquisition of the information (Lin and Soibelman, 2009). For better data retrieval, researchers suggested two systems: one for project management that answers questions (Cheng et al., 2002) and another for building performance management that uses context sensitive information (Keller et al., 2008). Various methods, like BIMQL (Mazairac and Beetz, 2013) and graph-based data retrieval (Langenhan et al., 2013), were suggested to make BIM data retrieval easier. Yet, the former offers a query language that is comparable to SQL, which the majority of consumers are not acquainted with. A non-specialist will have a hard time using a query language similar to SQL. The second method is more suited to handling geographical data. There was no practical data representation interface provided by either method.

Data representation and visualization are crucial tools for conveying information in a clear and effective manner. Various studies on data visualization were covered, including those dealing with categorical data (Chang and Ding, 2005) and time-oriented data (Klimov et al., 2010). To help with design collaboration in 3D virtual environments, a visual representation was given, which supports several subsystems of design representation and how they interact with one another (Gu and Tsai, 2010). In order to facilitate building maintenance, a method was suggested for simulating the visual degeneration of flooring systems (Khosrowshahi et al., 2014). A visual method for planning construction led to better coordination and higher-quality work from all trade contractors (Tan et al., 2005).

1.2 Discussion

IFC, the object-oriented and semantic schema, provides a solid foundation for BIM interoperability. Most BIM applications now adopt IFC as a data exchange standard, a large amount of data is accumulated and will continue to increase. It is of potential to carry out a data retrieval and analysis service based on the IFC-structure data model. Meanwhile, related studies show that implementing a cloud BIM is possible and will improve the collaboration and interoperability based on BIM. And an intelligent data retrieval approach based on natural language will provide a flexible way to retrieve BIM data for a wide range of stakeholders. A data representation method will enhance the value of the data retrieval approach.

Natural Language Processing (NLP) is widely used and deeply explored as a computerized approach to analyze text based on both a set of theories and a set of technologies (Liddy, 2001). Tools including Stanford parser (De Marneffe et al., 2006), and NLTK (Bird et al., 2009) were already developed to process text in different languages.

In conclusion, an urgent demand and a good foundation exist for intelligent BIM data retrieval and representation based on the requirements of non-experts. With a large amount of BIM data collected from different BIM applications, further promoting the value of cloud BIM based on flexible retrieval and feasible representation is important.

3 CLOUD-BASED DATA STORAGE

The cloud makes it possible for individuals from all over the world to work together efficiently by giving them access to shared data. To make it easier for diverse users to get information and to alter the massive BIM, cloud-based BIM data storage is used.

Owners, contractors, and other customers have access to data manipulation and information services provided by the cloud's several clusters, as shown in Figure 1. Metadata describes the kinds of data saved in a particular cloud cluster, and a collection of NoSQL databases that persist data in the IFC format BIM make up each cloud cluster. Most cloud implementations come with a MapReduce framework, which is a set of related programming models and tools for processing data in parallel across massive computer clusters (Dean and Ghemawat, 2008). Two sections make up the framework: map and reduce. The first one is a function that divides up tasks across the dispersed cluster's nodes, and the second one is another function that gathers all the results and returns them as a single number. The cloud platform's MapReduce foundation will improve data analysis and manipulation for quick access.

The BIM server now has a columnar database interface (Beetz et al., 2010), so users may easily migrate to other NoSQL databases or even relational databases.

setting up the interface for the BIM database. To meet these needs, however, a separate database is required since the BIM server does not provide an appropriate interface for integrating the MapReduce framework or comparable functions offered by distributed databases like HBase or MongoDB.

(5) Other: A flexible serialization technique and drivers for many programming languages make MongoDB suitable for a diverse group of developers. Object serialization may be easily controlled in this way.

3.1 Serialization strategies for IFC-based BIM data

In MongoDB, objects of the same type are stored in a collection and each object is serialized as a bson (similar to json) document in the collection. The attribute of an object can be serialized as a single value if it is a simple type, such as string or integer. While for a complex type, it should be serialized as a nested sub-document or as a document in another collection (Figure 2).

For an object-oriented data model, such as IFC, objects in the model are referenced to one another, which in some situation causes self-reference, resulting in problems when serializing the entities into a database. With a deep inherit chain, serializing an object of the IFC data model to a MongoDB document may also result in a document with deep nested sub-documents. It will lead to inefficient data retrieval and analysis. Given that IFC is designed for data exchange, the geometric information of an IFC data model is too complicated and inefficient for display in a 3D view. Thus, the geometric information should be converted into a display friendly format and stored separately.

Therefore, different serialization strategies should be adopted when designing the IFC-based database on MongoDB. Entities in an IFC data model can be classified into five parts:

al., 2012; Wang et al., 2014), and numerous NLP tools (such as NLTK and Stanford parser) were already developed. In the current research, among the many NLP tools proposed, the Stanford parser (De Marneffe et al., 2006) was selected to parse the words into a syntactic tree structure. The detail of this process is further discussed in the following part.

4.1 Keyword extraction based on NLP

Consider the natural sentence "quantity of beams of second and third storey" as an example. Following the above-mentioned steps, the user intention in a sentence processed by NLP tools will be parsed to a probabilistic context-free grammar structure similar to Figure 5. The sentence is split into words, and all the words are tagged by

Penn Treebank POS tag set (Bies et al., 1995). The final syntactic structure is shown as a tree, where ADJP, NP, and PP represent adjective, noun and preposition phrase respectively, and IN, NNS, NN, CC, and JJ stand for preposition, noun (plural), noun (single or mass), coordinating conjunction, and adjective, respectively.

From the root of the tree, a path passes through all nodes only tagged as NP/NN/NNS to the leaf (like left blue part of Figure 5), which is considered the keyword that represents what concerns the user. For each sub-tree whose root is tagged as noun (or other noun phrase), the path to the leaf with all nodes tagged as noun or its equivalent tag determines the keyword of this sub-tree, while the child node tagged as adjective or preposition phrase performs as a constraint for the keyword node that share the same parent with it (middle green part of Figure 5). The order of constraints (top of Figure 5) for different keywords can be determined through recursive analysis of the sub-tree of the structure, which is the inversion of the analysis order. Based on the extracted keywords, it is clear that the key concepts are "quantity", "beam", and "storey". Since adjective and preposition phrases provide more information about the keywords, with the order of constraints, it can be determined that what the user wants to know is the quantity constraint to beams, which were contained in a storey whose name is equal to second or third. These grammatical relationships are important in the subsequent data retrieval and representation.

However, two adjacent leaf nodes with the same parent node may both be tagged as NN in some sentences, for instance "construction progress of the check-in zone" (Figure 6). In this situation, a type dependency analysis based on an NLP toolkit (like Stanford parser) is required, and the dependency between "construction" and "progress" in the sentence is expressed as NN (progress-2, construction-1), which means that the keyword should be "progress" and should be constrained by "construction".

Finally, considering the above-mentioned situations, main steps for extracting keywords from a tree structure of a sentence should be as follows: (1) execute a type dependency analysis for the tree structure; (2) traverse the tree for a path connecting a leaf node tagged as NN to the root then return the leaf node; and (3) if two leaf nodes share a same parent, then obtain the keyword based on results of the type dependency analysis.

4.2 IFD-based keyword mapping

Concepts and phrases are defined, characterized semantically, and assigned a unique identification number in the open library that is the IFD Library. Two primary categories of concepts exist: There are two types of concepts:(1) subjects (or words) and (2) characteristics (or attributes). A subject is anything that is distinct from other things and can be easily differentiated. According to NBIMS (2012), there are six distinct categories into which these notions fall: behavior, environmental effect, function, measure, property, and unit. "Acts upon," "composes," "specializes," "assigns properties," "assigns measures," and "assigns values" are some of the categories into which relations between subjects and characteristics sometimes fall. A global unique identifier (GUID) is issued to each notion so that they may be easily identified and reused. But the IFD library is more of an idea database than an instance database. IFD just specifies the names of object types and the attributes that characterize them.

the construction management system was used to model the other aspects of the project, such as the various render styles for activities, and the Microsoft Project file was used to offer text notes on the various tasks. This method laid the groundwork for the project-wide construction BIM. Figure 9.1 shows a portion of the model used for the installation of the MEP integrated unit.According to Ma et al. (2011), the quantity bill approach was used to simulate the amount takeoff data. The construction BIM included several sorts of quantities, including truss mass, pipe and cable length, water cannon count, and number of cameras.

2. A cloud database built on MongoDB was used to store all of the construction BIM data. The Intelli-BIM system was equipped with apps that enable intelligent data retrieval and representation.

A total of three clusters make up the cloud, with one configuration server and two data servers constituting each cluster. Approximately 150,000 objects make up the design BIM of each average floor, resulting in a total file size of roughly 1.5 GB. The overall file size for the whole structure is 10 GB, and there are 1,100,000 items in it. The final BIM is about 50 GB when you include in all the related images, documentation, schedule data, and quantity takeoffs.

The majority of the BIM data was saved in the IFC format during export. All seven stories of the building were serialized as IfcBuildingStorey objects with unique names and elevations; the "Name" property was used to keep track of the storey numbers. There were many zones that made up each floor. In this study, IfcSpace was used to store zones. The "Name" element of IfcSpace was used to differentiate between several zones, like zone A, zone B, the check-in zone, and so on. The IfcRelContainedInSpatial-Structure function was used to tie the building components to the zone in which they were located.

After the user entered their intention in a natural language sentence, Intelli-BIM extracted, analyzed, and represented the user-relevant data in various combination forms based on the aforementioned approach, since all BIM data from the design and construction phases were integrated into it. To illustrate, consider the following sentence: "construction progress of the check-in zone." After running it through natural language processing (NLP) for keyword extraction and mapping through IFD, it provides instructions on retrieving data from the cloud database. Lastly, it displays a multi-aspect information representation based on Microsoft's Windows form controls. See Figure 10 for a visual representation of the query's response; read on for an explanation:

1. A data visualization for resources, a comprehensive list view for the construction schedule, an animation for the simulation of the building process, and a timeline dashboard for the summary of significant activities throughout the construction were all used to display the feedback.

1. For resource usage calculated by aggregating the quantities of building elements related to the tasks, a chart was utilized for result representation. Different types of resources, as previously modeled, were shown in the chart. For resources used in a short time span, a colored bar was chosen, and for long duration resource usage a line strip was adopted. To obtain a feasible representation of different data of different types distributed in a large range, double Y-axis was adopted and configured (i.e., logarithmic and common axes were used, and minimum and maximum values were selected). Thus, users could easily identify how many types of resources were used and their corresponding quantities.
2. A collapsible list view was utilized for details of construction tasks. Finished works were collapsed with a total finished amount as default, and working tasks were displayed with start/finish time as well as resource quantities of the elements related to the task. Since quantities of different resources were displayed in a single column, resource types were not specified in the column header. Works planned in the future were also represented similar to working tasks. Obtaining an overall view of tasks involved in the construction of the check-in zone was flexible for users.
3. For representation of the construction progress, a timeline-based dashboard was proposed. The form of the representation was composed of three parts: a timeline view showed important tasks (such as milestones) as a flag with the task name; a snapshot of the selected task and a title for the picture were provided; and finally, a summary was given including the name, start time, finish time, and duration of the task. Furthermore, related contractors and resources used were presented in the dashboard by the above-mentioned data analysis supports. With the timeline- based dashboard, users could obtain a full view of the construction progress, resulting in an easy and efficient construction management. Such a method could be easily extended to other phases of the project.
4. In summary, with the combination of animations, charts, list views and a timeline-based dashboard, users can easily obtain the whole information for the construction of the check-in zone, proving that: (1) the cloud-based database is suitable for data storage and the pre-join of the data based on the MapReduce framework works fine; (2) the user intention in a natural langue sentence is properly processed, and data retrieval and representation are correctly implemented; and (3) a multi-aspect view that integrates different representations of the results provides users abundant feedback on what they want.

More queries, such as "quantity of beams of second and third storey"(Lin et al., 2013), and "quantity of steel columns of the check-in zone", were tested in the system, the results were shown and displayed as expected. Query time of these queries ranges from 1.5s to 3.5s, keywords extraction and mapping takes about 0.5s, while data retrieval and representation takes the left part of the time. Investigation of each step of the approach shows that: 1) keywords extraction is based on traversal of the syntactic tree provided by the NLP tools, results are mainly limited by these tools; 2) text searching in the IFD library was adopted for keywords mapping, since total number of the concepts is less than 800, it is inevitable that related concepts cannot be found for some keywords, thus resulting in a query failure; 3) data retrieval mainly depends on the relationship finding based on the graph generated from IFC schema. Different time is needed for different number of keywords; 4) only simple sentences without verbs and pronouns were considered, it is inadequate for complex sentences currently.

7 CONCLUSION AND FUTURE WORKS

This research delves at the ways in which cloud computing may alleviate the burden of managing massive amounts of BIM data, as well as how non-experts might benefit from BIM by making use of data retrieval and representation based on natural language. First, a distributed data storage cloud based on MongoDB was set up, and the MapReduce framework was used to automatically pre-join two collections, so that processing a massive BIM and getting a quick query could be supported. Secondly, in order to identify the most important things related to the user, the terms "keyword" and "constraint" were suggested. The extraction of keywords from the user's natural language intent was enhanced using an NLP-based approach. Third, an enhanced approach was developed to use IFD to map keywords to IFC entities or characteristics. The IFC schema may be used to construct a graph that can be used to establish relations between various IFC items or characteristics. Data relevant to the user may be accessed from the BIM database once the property restrictions have been evaluated. Lastly, the data was organized, summarized, and lastly shown according to its format, which may be tables, charts, animations, or a mix of these. Automatic retrieval, analysis, and representation of user-concerned data is possible with semantic comprehension of the user's purpose in natural language, as shown by the practical application results in construction management. This method improves the value of building information modeling (BIM) and makes it easier for businesses to utilize it, even without highly technical users.

Other areas that might benefit from the proposed method to streamline human-computer interaction include facilities management and cost management. The following restrictions, nonetheless, are associated with the suggested method. 1) At this time, it is not possible to use the data representation results for custom configuration. An adaptable data representation may be attained with further effort. 2) We only support basic phrases at the moment. Sentences that need complicated calculations, include verbs, or include operators (*, /) are not not yet back up. Mapping verbs and actions to calculation functions should be further improved to ensure correct support. 3) The IFD library must to include adequate AEC domain ideas and be regularly enhanced because it is used to link keywords to IFC entities or attributes. Cloud BIM makes it easy to access database data; however, the time it takes to extract keywords and map user meaning in plain language is where the system slows down. As a result, we need to look at other ways to enhance it, such as caching frequently used queries and parallelizing NLP over many clusters. Fifthly, the data representation strategies discussed in this research still do not fit on mobile devices. We need better ways to portray things. Integration of features like full-text search of unstructured documents (e.g., dxf and pdf files) is essential for maximizing the benefits of building information modeling (BIM).

REFERENCES

In the proceedings of the 27th annual International CIB W78 Conference, held in Cairo, Egypt, Amor and Dimyadi (2010) presented an open repository of IFC data models and analyses to support interoperability deployment.

Publication: Construction Innovation: Information, Process, Management, Volume 8, Issue 3, Pages 218–239, Authors: Anumba, C.–J., Issa, R.–R., Pan, J., & Mutis, I. (2008).

Leadership and Management in Engineering, vol. 11, no. 3, pp. 241–242, Azhar, S. (2011), Building information modeling (BIM): Trends, advantages, dangers, and difficulties for the AEC sector. 252.

In the proceedings of the 27th annual International CIP W78 conference in Cairo, Egypt, Beetz, J., van Berlo, L.-E.-O., de Laat, R., and van den Helm, P. (2010) presented BIMserver.org, an open source IFC model server.

In their 1995 publication, "Bracketing guidelines for Treebank II style Penn Treebank project," Bies, Ferguson, Katz, MacIntyre, Tredinnick, Kim, Marcinkiewicz, and Schasberger outlined the procedures to be followed while creating brackets for the Penn Treebank project.

O'Reilly Media, Inc. published Bird, Klein, and Loper's (2009) Natural Language Processing with Python.

Automation in Construction, 17(1), 90-106, Boukamp, F., & Akinci, B. (2007). Automated processing of construction requirements to assist inspection and quality control.

In 2013, Bromberg compared the fundamental performance of MongoDB with SQL Server. The article can be found at: http://www.nullskull.com/a/1498/mongodb-vs-sql-server-basic-speed-tests.aspx. The page was visited 201 times. Friday, February 3, 2015.

This document is available online at: http://www.ifd-library.org/images/IFD Library and was authored by BuildingSMART in 2008.

"White Paper 2008-04-10 I" (.pdf), retrieved May 20, 2013. An Introduction to Natural Language, CCSI (2012)

Data processing, May 20, 2013, obtained from http://www.mind.ilstu.edu/curriculum/protothinker/natural_language_processing.php.

Data & Knowledge Engineering, 53 (3), 243-262, Chang and Ding (2005), Categorical data visualization and grouping utilizing subjective criteria.

Using a question-answering system for use in project management applications, Cheng, Kumar, and Law (2002) published in Advanced Engineering Informatics, 16(4), 277-289.

In 2012, Chiarcos, Ritz, and Stede published a paper. With all these beautiful mementos... Language resources and evaluation, 46(1), 53-74, merging contradictory tokenizations.

In the proceedings of the 28th International Symposium on Automation and Robotics in Construction, held in Seoul, Korea, from 144 to 149, Chuang, T., Lee, B., and Wu, I. (2011) discuss the use of cloud computing technologies for building information modeling (BIM) visualization and manipulation.

In the Proceedings of the International Conference on Language Resources and Evaluation, held in Genoa, Italy, from 449 to 454 pages, De Marneffe, MacCartney, and Manning (2006) presented a method for producing typed dependency parses using phrase structure parses.

Simplified Data Processing on Large Clusters: MapReduce, by Dean and Ghemawat (2008), in Communications of the Association for Computing Machinery, 51(1), 107–113.

"Exchange model and exchange object concepts for implementation of national BIM standards" was published in the Journal of Computing in Civil Engineering in 2009 and was co-authored by Eastman, Jeong, Sacks, and Kaner. The article can be found in volume 24, issue 1, pages 25-34.

An IFC model viewer to facilitate the use of nD models was published in the 2006 issue of Automation in Construction by Fu, Aouad, Lee, Mashall-Ponting, and Wu. The article can be found on pages 178 to 185.

The article "Interactive Graphical Representation for Collaborative 3D Virtual Worlds" was published in 2010 by Gu and Tsai in the journal Computer-Aided Civil and Infrastructure Engineering and covers pages 55 to 68.

Computer-Aided Civil and Infrastructure Engineering, vol. 21, no. 7, 2006, pp. 530-547, is authored by Hammad, A., Zhang, C., Hu, Y., and Mozaffari, E.

Automation in Construction, 20 (2), 167-180 (2011). Hu, Z.-Z., & Zhang, J.-P. (2011). Building Information Modeling (BIM) and 4D-based integrated solution of analysis and management for disputes and structural safety concerns during construction: 2. Development and site testing.

At http://www.aconex.com/blogs/2011/04/bim-, you may find Jasper's (2011) article titled "BIM: Managing the Data Deluge."

on May 20, 2013, via the website managing-the-data-deluge-part-1.html.

This information was retrieved from Jennifer A.'s (2012) article "Expected BIM Trends in 2013" at http://www10.aeccafe.com/blogs/aecsanjay/201 2/12/20/expected-bim-trends-2013/.

In their 2008 publication, "Modeling collaboration possibilities in practice," Jørgensen, Skauge, Christiansson, Svidt, Sørensen, and Mitchell discuss the practical use of IFC model servers.

In the Proceedings of the 2009 ICCEM/ICCPM conference in Seoul, Korea, Kang and Lee detailed the creation of an object-relational IFCS server.

Integrating the specification, collection, and processing of building performance information is the subject of an article published in 2008 by Keller, O'Donnell, Menzel, Keane, and Gokce in the journal Tsinghua Science and Technology. The article is numbered 1–6.

In the field of building maintenance, Khosrowshahi, Ghodous, and Sarshar (2014) developed a method to visualize the predicted deterioration of flooring systems. Their work was published in Computer-Aided Civil and Infrastructure Engineering, volume 29, issue 1, and spans pages 18-30.

Authors Kim, Kwon, You, and Lim (2010) used Proceedings of the 27th International Symposium on Automation and Robotics in Construction, Bratislava, Slovakia, includes the Information Delivery Manual (IDM) for effective transmission of building design information.

Artificial Intelligence in Medicine, 49(1), 11–31, Klimov, D., Shahar, Y., & Taieb-Maimon, M. (2010). Intelligent display and exploration of time-oriented data of many patients.

The authors of the article "Web services enhanced interoperable construction products catalogue" (Kong, S.-C.-W., Li, H., Liang, Y., Hung, T., Anumba, C., & Chen, Z., 2005) wrote about it in Automation in building, volume 14, issue 3, pages 343–352.

Journal of Information Technology in Construction, 17, 134-161, Laakso, M., Kiviniemi, A., & Others (2012), The IFC standard—a overview of history, development, and standardization.

In order to aid in the early phases of design, Langenhan, C., Weber, M., Liwicki, M., Petzold, F., and Dengel, A. (2013) published an article in Advanced Engineering Informatics titled "Graph-based retrieval of building information models" (27 (4), 413-426).

"A Core System for Design Information Management Using Industry Foundation Classes" was published in the journal Computer-Aided Civil and Infrastructure Engineering in 2003 by Lee, K., Chin, S., and Kim, J., and can be found on pages 286-298.

New York: Encyclopedia of Library and Information Science, Liddy, E.-D. (2001), Natural language processing, 2nd edition. The company known as Marcel Decker, Inc.

In the proceedings of the 30th CIB W78 International Conference on Applications of IT in the AEC Industry, held in Beijing, China, Lin, Hu, and Zhang (2013) presented intelligent data mining and representation with a focus on building information modeling.

In their 2009 article published in the Journal of Computing in Civil Engineering, Lin and Soibelman provide an online product search engine tailored to the architectural, engineering, and construction industries by combining domain knowledge with information retrieval methods.