# BIM - search for information

Igor Svetel
University of Belgrade, Innovation Center, Faculty of Mechanical Engineering, Serbia
isvetel@mas.bg.ac.rs ORCID: https://orcid.org/0000-0001-7821-1412

**Abstract:** The digitalization in the AECO (Architecture, Engineering, Construction, and Operations) sector is often heralded as a revolution in the way of working. However, it turned out to be a gradual evolution of the process. An industry accustomed to monolithic representations containing all necessary data for a certain project phase has a hard time changing its habits. The first digitalization step was creating those documents using a computer. The second step is BIM technology, which has the ambition to remain the final form of digitalization in the AECO sector.

The IFC scheme, which represents basis of the openBIM concept enabling connection among all BIM applications, also represents one monolithic digital model. Conceived to represent one AECO project, the IFC scheme contains all data about the single project. Additional technologies such as Data Dictionary and Data Templates, which add more data structures to the basic model, have been developed to enrich IFC.

Currently, the focus in the BIM community is on models containing all the necessary data that users require. Different kinds of information requirements define what data should be included in models. Most of the effort is focused on including the data in the model, and little on how the data will be extracted from the model. The fact that well-structured data does not represent the whole information is recognized in the ISO 7817 standard, which states that when creating the Level of information need, attention should be paid to who, when, why and to what extent will use information, although it does not specify how to use these prerequisites to search for information.

The paper analyzes the ways to search BIM data structures from the simplest browsing using the BIM viewer to more complex techniques. Existing proposals are considered, and original ones based on the latest BIM developments are given.

**Keywords:** BIM; information; browsing; query

## 1 Introduction

At the beginning of the 1980s, the digitalization of the AECO sector began with the appearance of personal computers and CAD programs that worked on them, which made this technology accessible to many. The advantages of creating traditional design documents using computers have been relatively quickly accepted, but the transition to a process based on a fully digital environment has been much slower. Already in the second half of the 80s, programs appeared that used the 3D model of the building as a basis for the automatic creation of 2D plans, sections and views of the object, which would become the basis for the development of BIM technology together with the development of parametric modeling of geometry.

In the mid-90s, the beginning of the development of the IFC specification marked the beginning of more intensive industrial application of BIM technology. Developed both as a conceptual data schema and an exchange file format IFC enabled the creation of a common, vendor-neutral digital representation of the built asset, including buildings and civil infrastructure (buildingSMART 2024a). Today, IFC represents the core of the so-called openBIM approach, a collaborative process of creating, sharing and using digital information that spans the entire lifecycle of an asset and that enables seamless collaboration among all project participants by creating a common language to export and import data. Managed by buildingSMART International openBIM today represents a set of standardized technologies and processes that determine information management in the AECO sector.

Despite these great achievements in the digitalization process of the AECO sector, there still lingers a document-based thinking among the participants in the process. The transition from a document-based process to a model-based process is a key aspect of the digitalization of the AECO sector (Svetel 2020). A significant part of the participants in the AECO process views BIM only as a tool for creating traditional project documentation, the only thing is that instead of paper, they use digital formats to display drawings, images, texts or tables. However, even among the advocates of a purely digital approach to the AECO process, there are still traces of document-based thinking.

The IFC scheme is a monolithic digital data model designed to represent one AECO project and to contain all the necessary data. Over time, it turned out that due to the specifics of the professions that are included in the life of a built asset, it is much more practical to use separate models for each individual purpose. However, each model respects the entire object-oriented structure of the IFC scheme, forming a monolithic entity that contains all the necessary elements, their properties and values (Figure 1).

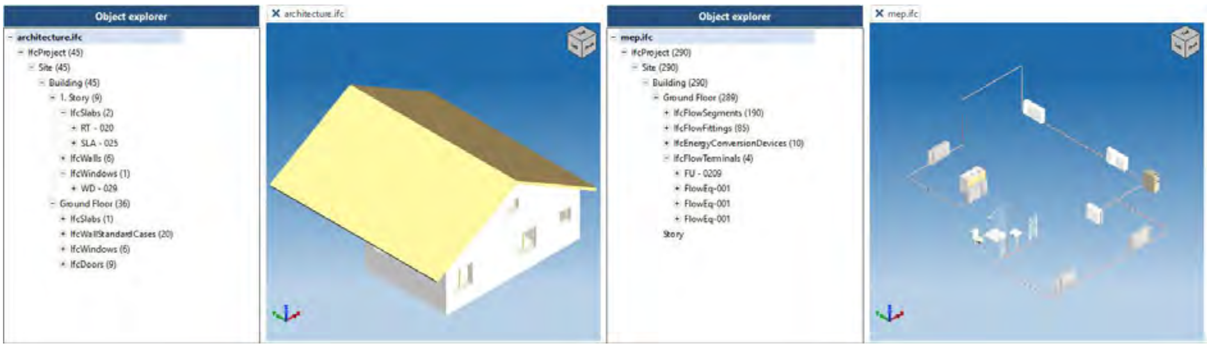


**Figure 1** The structure of the IFC scheme for models of different disciplines (screenshot Open IFC Viewer)

The document-centric mindset is also reflected in the BIM community's current preoccupation with models that contain all the necessary data required by users, which brings us back to the traditional paper based method in which it was important to create plans, schedules and details that contain all the necessary data. Different technologies have been developed that enable a structured representation of user requirements and enable computer processing of them (Svetel 2023). But mostly these technologies are used to define the structures of digital models for certain specialized purposes or to check whether all the required data is present in the model and whether the values are properly formatted. Most of the effort is focused on incorporating the data into the model, and little on how the data will be extracted from the model.

The same influence of documentation is felt in the definition of what information is in a BIM environment. The ISO 19650-1 standard defines information as “reinterpretable representation of data in a formalized manner suitable for communication, interpretation or processing” (ISO 2018). The newer standard ISO 7817-1 defines information simply as “meaningful data” (ISO 2024). Both of these definitions are consistent with the established concept of the DIKW (data, information, knowledge, wisdom) pyramid (Ackoff 1989) where data is defined as set of raw, unorganized symbols that denote something in the real world, while information assumes organization of data in structure that gives additional meaning that goes beyond basic references to objects or concepts. In accordance with that concept that dominates the ICT world, a highly structured IFC scheme represents information about a building or infrastructure. The shortcomings of the IFC scheme have been overcome by the development of new highly structured data models such as Data Dictionary or Data Templates that enrich IFC with additional data.

However, if we return to the mathematical theory of communication in the first paper that explicitly dealt with the concept of information (Hartley 1928), information is related to communication. In communication, participants use symbols, abstract patterns that convey meaning to the parties in communication. When the sending party selects one symbol and sends it over the communication channel, it excludes all other available selectable symbols. As more symbols are added, more and more possible symbol combinations are ruled out and the information becomes more precise. However, in order to obtain a mathematical definition, the process had to be separated from people and reduced to an abstract theory, which was finally elaborated in a work of Claude Shannon (1948) that became the basis of the development of ICT technologies. The problems brought by the abstract theory were noticed even before the explosive development of ICT technologies.

Peter F. Drucker (1970) emphasized that formal and impersonal information processes will separate people and will require much greater efforts to restore the human relationship reflected in communication. He also noted that the person who receives the information is the one who communicates. Without receiver there is no communication, no information, but only physical manifestation of sending signals. The conclusion is that communication is on the receiving end and that it is a reaction and not an action. In order for communication to take place, both parties must respect a prior agreement on the code used. The perception, experience, expectations, involvement and motivation (values, beliefs, aspirations) of the receiver to whom the message is intended must also be taken into account. Information should be relevant to the receiver; otherwise it is treated as noise. Transferred to the realm of BIM technologies, a set of highly structured wall data expressed in standardized values that represent valuable information to someone doing an energy efficiency analysis means nothing to someone who just wants to drive a nail into that wall.

The fact that the quality of information depends on the recipient has been known since classical antiquity. In the field of rhetoric, the principle of the "Seven Elements of Circumstance" originated from the works of Aristotle (Sloan 2010). It states the need to answer to seven questions: Why, Where, When, What, With Whom, By Whom, and How in order to provide worthwhile argument. The same concept is also present in the Five Ws

principle in journalism, which defines the questions of who, what, when, where and why as essential in gathering information.
The shift away from understanding information as solely a task of creating a highly organized data structure is also noticeable in the BIM environment. The ISO 19650-1 standard (ISO 2018) states that before starting a project, the party initiating it should clearly define its information requirements, specifications on what, when, how and for whom the information will be produced. The process of satisfying information requirements is called information exchange and is defined in the standard ISO 19650-4 (ISO 2022). To facilitate and automate the information exchange process, the Information Delivery Specification (IDS), a computer-interpreted format for specifying and verifying information requirements, was developed (buildingSMART 2024b). The IDS is designed to support only the IFC scheme and contains metadata and list of specifications. Each specification describes one entity in the model (ifcWall, ifcWindow, etc.) and states its applicability and requirements. Applicability defines the information required for the specification to be applicable to the entity and requirements define information that an entity needs to have. The metadata has two parts. The first part refers to the entire specification and which specifies who will use the IDS, why it is created, for what projects, what the IDS will achieve and the phase of the project to which the IDS applies. The second part refers to each specification and explains why a particular element is important to the project and gives instructions on who is responsible for providing the information and how. The IDS also represents a shift away from monolithic schemes. A valid specification can only refer to a single element in the model, and it can also describe complex requirements for a specific discipline, phase in a project, or even an entire facility. A new member in the collection of BIM standards ISO 7817 - Level of information need (ISO 2024) states that attention should be paid to who, when, why and to what extent will use information. The development of a data model and derived XML schema for defining the Level of information need in software applications is planned.
In accordance with the assumption that the information is on the receiver's side, this paper provides an overview of possible ways to search BIM models. As the openBIM approach requires the model to be saved in an open format, IFC schema search procedures are presented starting from simple model browsing, through structured model browsing, to BIM queries.

**2 Simple model browsing**
Considering that both formats for the representation of the IFC scheme (STEP and XML) use plain text notation the simplest way to search would be to use a text editor. However, the way the data is formatted is not simple. Each item in the IFC STEP file is defined in one line that starts with an ID number followed by a name and a list of properties. If the property is simple, it is specified in the list, and if it is complex, the ID of the item that defines it is given (Figure 2). The similar approach is used in XML format. Therefore, in order to get a complete overview of the model, it is necessary to parse the file in some software.

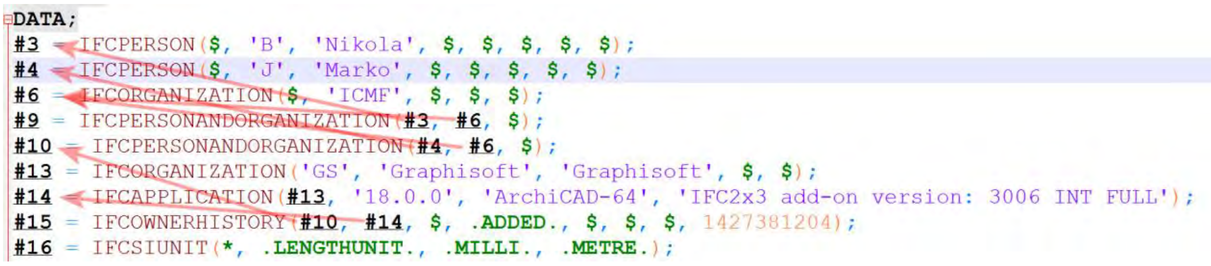

**Figure 2** A snippet of an IFC STEP file that demonstrates item cross-referencing (screenshot Notepad++)

IFC viewers began to be produced quite early, at a time when neither the content of IFC files nor the method of application had yet been determined. Since it was not possible to resolve whether the data transfer problem was occurring in the application that was exporting the data or the one that was importing it, IFC viewers were a way to see the content directly.
Each IFC viewer contains a minimum of three panels, one showing the structure of the IFC schema, the second showing the 3D view of the model, and the third showing the parameters (attributes and properties) and their values for the selected model element. In order to get some information from the viewer, it is necessary to know for which element the data is requested, and then select that element either in the structure panel or in the 3D view panel, after which the data about that element appears in the properties panel.
The quality of information that can be obtained in this way depends on how deep the parsing of the file goes and on what parameters the software developers consider important. For example, IfcOwnerHistory, which was previously used to monitor who and with which software created and changed objects in the IFC model,

and which, with the advent of the CDE system that monitors and records those changes, became an optional parameter, is most often not loaded into new IFC viewers.

**3 Structured model browsing**

Using the earliest IFC viewers required knowledge of the IFC schema structure. To bring them closer to a wider number of users, software developers have introduced additional ways to represent the structure of the IFC schema. Most often, the structuring of the model according to the IFC types of elements was applied. Other frequently applied structuring of the model is according to layers, groups (systems, zones) and classification (Table 1).

| | IFC schema | IFC type | layer | system/group | classification | search |
|---|---|---|---|---|---|---|
| ACCA usBIM | X | X | X | X | | elem. name |
| BIMcollab Zoom | X | X | | | | full |
| BIMvision | X | X | X | X | X | (paid) |
| Dalux BIM Viewer | select object or search | | | | | full |
| DDScad Viewer | X | X | | X | | |
| Open IFC Viewer | X | | | | | |
| Solibri Anywhere | X | X | X | | X | partial |
| Trimble Connect | custom | | X | | | data table |

**Table 1** Overview of alternative IFC schema views in free IFC viewers

Structured viewing of the model provides advantages compared to classic IFC views, enabling easier finding of the desired elements, especially groups of elements, which is most often the goal of the search for information. Certain IFC viewers allow searching within the model structure display panel, where the terms that can be searched for are only those that appear in that panel. This approach allows for a much better search for groups of elements in the model, but it implies knowledge of the parameters displayed by some software, which may differ from the standard IFC scheme. As with classic IFC viewers, the volume and quality of the information that can be obtained depends on the depth of parsing and support for the elements of the IFC scheme that individual software developers provide.

**4 BIM query**

The oldest way to realize a BIM query, which is still in use, is to choose categories and parameters in the model and to export them in the spreadsheet format, most often as MS Excel file, and then to use traditional spreadsheet query methods. Although practical and comprehensible to all participants in the AECO industry, this approach loses the advantages of structured data organization that the IFC scheme provides. In order to overcome this deficiency, methods were developed for direct querying of the IFC scheme.

Some IFC viewers that are shown in the previous chapter (Table 1) allow a full and highly structured data search. It is possible to add multiple criteria, include or exclude elements based on criteria, define criteria based on types of elements, properties and values. Values can be selected based on their equality, inequality, partial match, existence or non-existence in a set of values. Because the search results can be seen immediately during the creation of the query itself, this approach enables the creation of complex queries even for less experienced users. The disadvantage of this approach is that the search results cannot be exported from the application. In order to overcome this, it is necessary to use BIM query, a system that enables finding a well-specified subset of the entire building model and its export.

The BimQL (Mazairac and Beetz 2013) works on top of the BIMserver, an open source building information server that stores IFC data as objects, not as files. It is a query language for IFC-based models that allows selection and partial modification of model instances. The language enables flexible and ad hoc generation of partial datasets from IFC models supporting everyday practical information requirements. It was developed with the goal of making query writing easy, with grammar and syntax that users can quickly grasp. The basic complexities of the model are hidden through domain-specific constructions.

The QL4BIM adds the possibility of forming queries based on geometry data (Daum and Borrmann 2014). The language provides metric (Distance, CloserThan, etc.), directional (Above, Below, NorthOf etc.), and topological (Touch, Within etc.) operators for defining search filters. Since spatial searching of 3D geometry is computationally expensive, QL4BIM facilitates optimized geometric algorithms and spatial indexing. The language was also developed with the aim of enabling query development for less experienced users.

The IfcOpenShell (2024) is an open source toolkit for working with BIM models represented in IFC format. It provides C++ and Python API. Using its selector query syntax, it is possible to perform complex IFC schema searches, and since IfcOpenShell supports full IFC schema parsing, it is possible to search all data without limitations. Unlike BimQL and QL4BIM, which mostly remained at the level of research projects, IfcOpenShell is an active project that is gaining more and more popularity.

**5 Conclusion**

In the past years, BIM has achieved enviable progress in the field of defining digital data structures and standardization of accompanying processes, thus becoming the official method of digitalization of the AECO sector. At the same time, the traditional document-based way of thinking still dominates the sector, so that the model is still viewed as an aid to the creation of project documentation, and information is looked for only in the documentation. The only difference is that the documentation is not paper based but digital. In accordance with such an approach, the focus of the entire process is placed on the creation of the model and its verification and validation, while the ways of using the model are assumed to be predetermined.

Contrary to this attitude, more and more comments are appearing on the topic of the usability of BIM models and the possibility of finding information in them.

In his recent book Alexander Koutamanis (2022) notes that the problem with new technologies does not lie in the amount of information, but in the way in which it is accessed in a new digital way. People usually rely on easily available information, even if it is wrong, than on highly structured data.

In his article, John Ford (2023) recognizes that the technical part of ISO 19650 ('the what') is well supported within AECO sector, but the 'why' has been lost or completely ignored. Therefore, information requirements do not reflect the actual needs of the clients, and in order to overcome this, it is suggested that the question of why the information is needed be asked before determining how it will be structured.

The DATA Home study (Chimni 2022) aimed to determine the requirements of homeowners for digital information. The study recognized that homeowners expect greater access and control of data about aspects of their lives related to their home. In a world where smart environment control applications are regularly used, it is not sustainable for a homeowner to not have access to digital home information tailored to their needs. Starting with questions that owners may have about their homes, the study determined the types of data and their scope for all the proposed scenarios.

A recently published 'the golden thread' directive (UK Government, 2024) determines that for higher-risk buildings it is necessary that information related to security must be stored in digital form, available when someone needs them, presented in a usable way and accessible in a simple format that is easy to understand.

The buildingSMART also started work on Regulatory Information Requirements (buildingSMART 2024c) because it recognized that regulatory bodies have different information needs than participants in the BIM creation process. The goal is to improve communications with regulatory bodies and implementation of improved compliance checks.

By themselves, the BIM model search techniques described in this paper are not sufficient to provide a technical basis for the trends described above. Although developed with a different purpose, IDS represents a technology that could be used in combination with query techniques to find relevant information, at the right time, for the right person and in the required quantity. Instead of being used to automatically check that all data is present in the IFC model and that values are expressed in the required way, IDS could be used to guide information retrieval. In doing so, it would be necessary to create IDS specifications for each individual information requirement and then write a request for a model search guided by that specification. Based on what is shown in this paper, the most promising candidate for the realization of such an intention is IfcOpenShell.

From all of the above, it becomes clear that the digitalization of the AECO sector not only requires changes in work habits based on documents and a shift to work based on digital models, but also learning to create computer code in order to fully exploit the potential of these new technologies.

**Acknowledgments:** The results presented here are the result of research supported by the Ministry of Science, Technological Development and Innovation of the RS under Contract 451-03-66/2024-03/200213 dated 05.02.2024.

**References**

Ackoff, R. (1989) ‘From Data to Wisdom’, *Journal of Applied Systems Analysis*, 16, pp. 3–9.
buildingSMART (2024a) IFC 4.3.2.0 (IFC4X3_ADD2), https://standards.buildingsmart.org/IFC/RELEASE/IFC4_3/, accessed 7 September 2024.
buildingSMART (2024b) Information Delivery Specification (IDS), https://www.buildingsmart.org/standards/bsi-standards/information-delivery-specification-ids/, accessed 7 September 2024.
buildingSMART (2024c) Regulatory Information Requirements, https://www.buildingsmart.org/standards/calls-for-participation/regulatory-information-requirements/, accessed 7 September 2024.
Chimni (2022) The DATA Home - The Digital Information Required by a Homeowner When Buying a New Home, https://chimni.net/data-home-download, accessed 7 September 2024.
Daum, S. and Borrmann, A. (2014) ‘Processing of Topological BIM Queries using Boundary Representation Based Methods’, *Advanced Engineering Informatics*, 28(4), pp. 272-286.
Drucker, P.F. (1970) *Technology, Management & Society.* New York: Harper & Row.
Ford, J., (2023) ‘Why is the ‘why’ of ISO 19650 being lost?’ *BimPlus*, 22.08.23 https://www.bimplus.co.uk/why-is-the-why-of-iso-19650-being-lost/ , accessed 7 September 2024.
Hartley, R.V.L. (1928). ‘Transmission of information’. *Bell System Technical Journal*, 7(3), pp. 535-563.
IfcOpenShell (2024) IfcOpenShell The open source IFC toolkit and geometry engine, https://ifcopenshell.org/ , accessed 7 September 2024.
ISO (2018) ISO 19650-1:2018 Organization and digitization of information about buildings and civil engineering works, including building information modelling (BIM) — Information management using building information modelling — Part 1: Concepts and principles, https://www.iso.org/standard/68078.html, accessed 7 September 2024.
ISO (2022). ISO 19650-4:2022 Organization and digitization of information about buildings and civil engineering works, including building information modelling (BIM) — Information management using building information modelling — Part 4: Information exchange, https://www.iso.org/standard/78246.html, accessed 7 September 2024.
ISO (2024) ISO 7817-1:2024 Building information modelling — Level of information need Part 1: Concepts and principles, https://www.iso.org/standard/82914.html, accessed 7 September 2024.
Koutamanis, A. (2022) *Building Information - Representation and Management: Principles and foundations for the digital era.* Delft: TU Delft OPEN, doi:10.5074/T.2022.001.
Mazairac, W. and Beetz, J. (2013) ‘BIMQL – An open query language for building information models’, *Advanced Engineering Informatics*, 27(4), pp. 444-456.
Shannon, C. (1948) ‘A mathematical theory of communication’. *Bell System Technical Journal*, 27(3), pp. 379-423.
Sloan, M. (2010) ‘Aristotle's Nicomachean Ethics as the Original Locus for the Septem Circumstantiae’, *Classical Philology*, 105(3), pp. 236–51.
Svetel, I. (2020) ‘Document vs. Model Based Digital Design Process’, in *On Architecture: Learning Architecture,* Belgrade: STRAND, pp. 35-41.
Svetel, I. (2022) ‘Information in Architecture in the Digital Age’, *in On Architecture Philosophy of Architecture Proceedings*, Belgrade: STRAND, pp. 71-78.
Svetel, I. (2023) ‘Toward the Utility of BIM Models’, *in On Architecture – Challenges in Design Proceedings*, Belgrade: STRAND, pp. 272-277.
UK Government (2024) Keeping information about a higher-risk building: the golden thread, https://www.gov.uk/guidance/keeping-information-about-a-higher-risk-building-the-golden-thread, accessed 7 September 2024.