DOCUMENT VS. MODEL BASED DIGITAL DESIGN PROCESS

Dr Igor Svetel
Innovation Center, Faculty of Mechanical Engineering, University of Belgrade
Kraljice Marije 16, 11000 Belgrade, Serbia, isvetel@mas.bg.ac.rs

ABSTRACT

Beginning in the mid-20th century, numerous studies have pointed to the limitations of the human cognitive system and provided insight into human reliance on external structures as a way to overcome these limitations. In architecture, such external representations are drawings, tables, mock-ups and other types of project documents. With the introduction of digital technology in architectural design, a new type of architectural project representation was introduced into the design process - the digital model. The development of commercial architectural design software was based on the development of both information models that support computer processing and digital representations of traditional documents.

A document is a representation intended for interpretation by humans, which takes into account all the shortcomings of the human cognitive system. Its value is based on permanence and immutability. On the other hand, the digital model enables computer-based processing, and it is free from the limitations of the human cognitive system. In order for a person to understand the digital model, it is necessary to translate it into the form of a document. The fact that a human cannot directly understand a digital model reduces the impact of computer technology and saves traditional working methods.

The article discusses the differences between data and information and documents and digital models and how these concepts are implemented in the current BIM technology. The overview of the international standards BIM standards is given and their impact on the use of documents and models in the BIM process has been studied.

Draft – published at: On Architecture: Learning Architecture Proceedings, R. Bogdanovic (ed.) 2020, Belgrade, pp. 35-41. ISBN 978-86-89111-23-1

INTRODUCTION

Starting around the Renaissance period, a traditional role of master builder, who was in charge of both the design and construction of the building, is divided into two disciplines: architecture and construction. The architects produced project documents as the representation on the basis of which the builders constructed actual building.

Representations have played a significant role in human history. The term "representation" takes many meanings. Common to all is a notion of something standing in for something else. More accurately, it is a physical entity (material or energy pattern) that stands for some other physical entity or some abstract concept. Representation enables us to step out of the world, and to reason about alternatives, thus escaping potentially dangerous consequences of our actions. It can be said that representation enables thinking. We can identify two classes of representation: analogical and symbolic (Boden 1988). The representation in which there is some significant similarity (interpretative mapping or some isomorphism) between structure of the representation and the thing that is represented belongs to a class of the analogical representations (Sloman 1978). Symbolic representation can be defined as the abstract pattern that by agreement stands for some other thing. This other thing can be some real-world object or event, some abstract concept or idea, or some other symbolic representation. An important aspect of the symbolic representation is its ability to be combined and manipulated to make more complex, molecular, structures (Kirsh 1991). These complex structures often represent different things than their constituent elements. The creation of representation is the characteristic of the whole thinking (Clark and Chalmers 1995) not only design process. Humans have a rich representational capacity. During their history people developed many artificial aids to thinking such as notation, language, models and formal systems. The most recent development in that tradition is the universal information processing machine - the computer.

It was only in the mid-20th century when preliminary psychological studies have pointed to the limitations of the human cognitive system (Miller 1956). The observed limitation of human short-term memory to seven plus minus two chunks, which was later reduced to four plus minus one (Cowan 2001), demonstrated that people can not keep too much information in mind at the same time while solving problems. One school of thought thinks that people developed representations to compensate this limitation. On the other hand, we can guess that during human evolution the internal processing capacities have not evolved because thinking process is based on the ability to use external structures. If we abandon the dualistic distinction between mind and real world, we can treat thinking as a single process of the mind/world interaction. We can trace the evolution of this process from the perception through social interaction and participation in culture (Burke 1995). Thinking, according to this model, does not happen just in the person's mind, but is a two-way interactive process resembling perception. Instead of dealing with the whole world, as perception does, thinking is oriented toward representations embodied in linguistic, discursive, or expressive media.

From the very beginning of the development of architecture as an independent practice, architects have used documents as the representations of their work. Over time, because of their reliability and permanence analogical representations like plans, sections, elevations and symbolic one like schedules became legally valid documents in the entire architectural, engineering and construction (AEC) practice. That situation did not change for centuries, until the recent inclusion of digital representations in AEC professions.

ABOUT DIGITAL REPRESENTATION MODELS

Computers have evolved in 1950-es from traditional calculating machines. Most calculating machines had limited scope of computation, but in 1830s Charles Babbage conceived first universal computing machine in the form of his analytical engine. In the seminal paper Turing (1936) depicted universal computing machine which became the blueprint for the later development of all digital computers. The Turing machine consists of the infinite tape on which 1s or 0s are written by the read/write device that moves in both directions along the tape and according to the states written in action table either changes value on the tape or leaves it unmodified (Petzold 2008). The program that defines what machine is doing, data on which machine operates, and results are all written as strings of 1s and 0s on the tape and can be all modified by the read/write device. And precisely this process of reading and modifying 1s and 0s on the tape is what computers do.

A common view is that the model is a simplified representation of a phenomenon expressed in some media. Since in all modern digital computers representations are based on strings of 1s and 0s the ICT field is preoccupied with the problem how to use them to model phenomena in the real world. For example, the ASCII standard is the model that defines how to use sequence of seven ones and zeros to represent letters of Latin alphabet, numbers, mathematical and other symbols. The Unicode standard defines how to use sequences of ones and zeros to represent all conceivable letters and symbols. Bitmap image file format is the model that specifies how to represent images using binary information. The 3D model defines how to represent coordinates in space, vertices, edges and surfaces of some real world body and establish oblique, axonometric or perspective representation on the screen or other media using strings of ones and zeros (Rogers and Adams 1976). Based on these basic models, it is possible to make more complex models. For example, HTML (WHATWG 2020) combines Unicode, bitmap, 3D models an many other digital formats into a new model that allows the diverse data present on the Internet to be interconnected and displayed as a combined document on a computer.

Unlike traditional representations that required man as an agent who interprets them and creates new representations based on his/her reasoning, strings of 1s and 0s in computers serves both as data that is manipulated and as commands that defines how computer processes data. This means that digital representations have a dual nature which is necessary to respect when creating digital representations. On the one hand, it is necessary to create models that determine how they will be displayed to people and function like traditional documents, and on the other hand, it is necessary to create models that determine how the information will be processed by computers. For example, HTML is information model intended for human use and it creates documents that are easy for people to understand, but in order for the information contained in them to be processed by computers, it was necessary to create a new information model - the Semantic Web (W3C 2015).

DATA VS. INFORMATION

It is clear that, from the ICT point of view, all digital representations are some kind of models. Without those data models we could not decode strings of 1s and 0s and obtain anything meaningful from ICT technology. One of the key aspects of digital representations is very often overlooked - the difference between data and information and the role of data models.

There are many opinions on the topic of exact definition of data and information but certain general characteristics can be determined (Zins 2007). Data is unstructured fact that needs to be processed to become meaningful and useful. The data needs to be interpreted, by humans or machines to acquire meaning. Humans learn from their birth how to interpret data. Simplest digital data models, like ASCII, Unicode, bitmap, wav, etc. provide instructions how to decode strings of 1s and 0s and get some real life data, like letters, words, picture, sound, etc. To obtain information it is necessary to provide structure that organizes, defines relationships and gives a context to data.

In computer systems, information can be expressed in two ways. We can use strings of 1s and 0s to model a traditional form of document, but in digital form, like PDF or HTML. The use and interpretation of such structures depends on people, computers can only extract unrelated data from them and cannot further process them without human intervention.

Another way to express information involves the ability of computers to process data. We can either create program procedures where the computer code defines the types of input and output data, or open data models that define the structures and relationships between data. In the first case, the data structure is locked into the application itself. The user must provide the required input data and use the application to obtain the requested information. Only the people who programmed the application know the mechanism of combining data and creating information. In the second case, digital data models are created that enable their processing. The advantage of this approach is that the structure of data model and the process that combines data to produce information are public and enable people to use the model independently of the individual computer application.

Strict distinction between data and information is not present in the ICT field either leading to a proliferation of ambiguity in other professions as well. Standard modelling languages like IDEF1X and EXPRESS are denoted as “data modelling languages” although they define structure and context for data and enable the creation of information.

DOCUMENTS VS. MODELS

People have developed many data structures like language, documents, mathematical formulas etc. to represent and convey information. It takes years during formal education to learn to use them. The documents are information structures represented in different material media, mostly in paper, developed to express and transmit information among people. They have reached a high level of expression and are widely accepted due to their durability and immutability. Most professions have developed elaborated forms of documents on which the legal foundations of profession rely. The structure of documents presupposes human interpretation and takes into account most human cognitive abilities and their imperfections.

Recently, with the advent of computer technology, it has become possible to create digital models that enable computer processing of data and automatic information creation. In the presentation so far, we have seen that from the aspect of ICT professionals, all computer representations represent a kind of model. But from the aspect of engineering professions, the model represents the simplification of a real phenomenon with the aim of explaining it, enabling the prediction of its behaviour, and understanding it (Frigg and Hartmann 2020). Over time, people have developed numerous types of models that have helped them understand the world around them and accomplish numerous engineering achievements. The ICT revolution brought new kind of models to the engineering field – digital models. The main difference from traditional models is that information structures do not require human processing but are organized to enable computer processing. The process of computer processing itself is not understandable to people and in order to present the results of computer processing to people, it is necessary to present them in the form of traditional documents. This aspect of computer technology still confuses many, and leads many engineers to feel isolated in relation to computers and to feel that they are in some way replacing them in their professions.

There is nothing true in the statement that computers will replace people. All applications and data models are created by people for people. It is necessary to accept computer technology as the advancement of all human activities and to accept that documents are information structures intended only for human processing and that it is necessary to improve our traditional work based on documents with digital models and the ability to process information by computers.

BUILDING INFORMATION MODELLING

Building information modelling (BIM), as defined by BIM Dictionary (2019), is a “set of computer technologies, processes and standards that enables collaborative design, construction and operation of built assets” which currently grows into the dominant mode of operation in construction industry. BIM has gradually evolved from older CAD and 3D modelling technologies that aimed to enable the creation of technical drawings and basic 3D geometric models using computers. The first BIM authoring application on the market, ARCHICAD, was not conceived as a BIM application, but as a program that creates consistent 2D architectural and construction documentation based on a 3D model. A special feature of this application was the orientation to building modelling, unlike previous applications that were intended for all engineering fields. To facilitate the creation of complex 3D building models, the program used parametric 3D objects that model individual building elements such as walls, ceilings, beams, etc. By adjusting the parameters, the user was able to adapt the generic objects to the specific needs of their own model. An additional simplification of modelling was the fact that the program was intended for building models, so that the objects included information about the real functioning of these components, such as materials, so that the program could automatically connect and trim elements of the model. The next application to hit the market, Revit, introduced a definition of functional relationships between components that allows for faster model revisions and automatic propagation of changes made to a single component across all related components. When these applications began to be accepted in the AEC industry, other companies decided to adapt their 3D model engines to the above described functioning, and soon this mode of operation was called BIM. In addition to numerous advantages, the new technology has also brought new problems.

The fact that each commercial application is based on a different 3D engine and that each discipline requires a different 3D model prevents the existence of one general BIM model. Also, the legal foundation of AEC professions on traditional documents requires that they be part of the whole process. The ISO 19650-1 standard (ISO 2018b) describes the IT structure on which BIM is based as “shared digital representation” which implies that the BIM model includes both documents and data. The US National BIM Standard (NIBS 2019) also describes BIM model as “digital representation of physical and functional characteristics”. This heterogeneous environment that combines traditional processes and modern technology often confuses users.

As we have seen before, digital representation of documents in most cases involves human interpretation. This means that the information in them is understandable only to humans, while for computers they represent only a set of data. Although this did not seem like a big problem to people, thanks to computer technologies, it is possible to create an enormous amount of documents today. Without the ability to process the information stored in them by computers, people are no longer able to control this process.

Draft – published at: On Architecture: Learning Architecture Proceedings, R. Bogdanovic (ed.) 2020, Belgrade, pp. 35-41. ISBN 978-86-89111-23-1

Related to this problem is the common approach among AEC practitioners to treat BIM as a tool. Traditionally, the tool is used to operate on the product, and since only product in AEC is document this view fosters digital documents. If we look at BIM this way, it means that people use their abilities to convert one document into another using a computer.
Another limiting view is that BIM is a 3D model with additional data describing the elements of the object, a kind of 3D database. Although this view accepts the term model, it remains limited to the basic level of the IT model that allows sequences of 1's and 0's to be transformed into a digital representation of bodies in space and thus still supports the traditional document-based process.
An attempt to solve these problems was the development of BIM standards.

BIM STANDARDS

The BIM standard with the longest tradition of application is Industry Foundation Classes (IFC) - ISO 16739 (ISO 2018a). The IFC is both information model (data scheme) and exchange file format that provides highest level of interoperability among various BIM applications. The standard defines all information structures which make up the model of the building through all the cycles in the life of the building. Recently its scope is extended to include definitions for infrastructure assets. Because the scope of the standard is so broad, it is impractical to use whole information structure in each information exchange. For the purpose of efficient data exchange, the model exchange specifications have been developed, known as Model View Definitions (MVD) (bSI 2020c). MVD precisely defines information needs for each particular case of model exchange and defines criteria for IFC exchange software certification.
Since interoperability based on the IFC format requires the exchange of complete models, day-to-day issue management based on these principles would be impractical. For these needs, the BIM Collaboration Format (BCF) standard has been developed (bSI 2020a). The BCF enables file based or server based issue exchange between different BIM applications. Most important aspect of BCF is that it enables a human created issue report to be integrated with BIM model by referencing the issue to IFC model using IFC GUIDs (globally unique identifiers) of the elements in the model where issue is detected. That way a human constructed document is linked to the model enabling designer to understanding a full context of the detected problem.
The IFC has managed to standardize a large number of elements and features that make up the building. But the number of elements and features was limited by the basic task of exchanging information between BIM applications and supporting all possible applications. In order to realize the possibility of developing a computer model of all elements of the building, their properties and relations, new information models have been developed. The ISO 12006 Part 3 (ISO 2007) is an information model standard that provides means to describe all concepts in the AEC field and their properties, groupings and relations. It provides computer interpretable structure that links building classification systems and model oriented structures for information, objects, and processes. It enables creation of data dictionaries that provide computer interpretable information about construction works. Using ISO 12006-3 it is possible to create many different specialized data dictionaries that suit national standards or specific company needs. To provide method to connect all these dictionaries an ISO 23386 (ISO 2020a) and ISO 23387 (ISO 2020b) standards were developed. In order to have consistent and interoperable data dictionaries it is necessary to develop a unique way to define the machine interpretable properties of all construction products. The ISO 23386 provides rules for defining properties and groups of properties through sets of mandatory and optional attributes. It also gives methodology for authoring and maintaining data dictionaries to provide high quality of content and establish seamless interoperability. The ISO 23387 defines a specific data model to create machine interpretable data templates for construction products.
The whole process described in ISO 12006, 23386 and 23387 standards requires centralized development and maintenance to ensure high quality of content. Toward this goal a buildingSMART Data Dictionary (bSDD) (BSi 2020b) is under construction. It is an online service to host different classifications modelled in accordance to above standards and provide linking between concepts in those classification systems. The foundation of this service on the mentioned ISO standards enables all concepts to be uniquely defined through GUID codes, which enables the connection of all dictionaries into a single network of meanings.
The basis of the BIM processes on a large number of diverse applications and digital formats, as well as the need to provide support for the traditional way of working have created a diverse and complex working environment. In order to enable the effective use of these new technologies, it was necessary to define BIM processes. Toward this goal a set of ISO 19650 standards is under development.
The ISO 19650 Part 1 (ISO 2018b) provides concepts and principles for information management of built assets during their entire life cycle using BIM technologies. The standard recognizes different perspectives

on asset information among different involved parties. To meet these needs, different types of information requirements are defined. Stakeholders' needs are defined as organizational information requirements (OIR) and project information requirements (PIR), while contractual obligations are covered in asset information requirements (AIR) and exchange information requirements (EIR). The deliverables that have to meet those requirements are project information model (PIM) that is used during delivery phase and asset information model (AIM) that provides support for information management during asset's operational phase. The standard describes all requirements that information about the asset must meet during its life cycle. The standard also describes information delivery cycle and describes general principles how to specify information requirements and plan for information delivery.

The ISO 19650 Part 2 (ISO 2018c) defines a specific information management process during design and construction phases that are collectively referred as delivery phase. It gives precise sequence and content of all plan-do-check-act cycles of information management. The process starts with the determination of information management functions, information requirements, information standards, delivery milestones, production methods, references and shared resources, common data environment, and information protocol. Results of these activities serve to make invitation to tender and then to create pre-appointment BIM execution plan and to evaluate responses to tender. When all project roles are appointed, a BIM execution plan (BEP) is created that defines all information processes and responsibility matrix. Next phases are: mobilization, collective production of information, information model delivery and project closing.

STATUS OF DOCUMENTS AND MODELS IN CURRENT BIM

There is an obvious great effort in the BIM community to improve the whole process. The IFC, BCF and Data Dictionary technologies and related standards create conditions for the creation and exchange of project information that are machine interpretable, thus reducing the need for human intervention and reinterpretation of information in any interaction among project participants. The set of ISO 19650 standards tries to create unique guidelines for the process of managing all project information. However, at the moment there are a lot of legacies from the previous period that are not clearly demarcated in relation to the new approach, which leads to confusion and stagnation in development.

According to a 19650-1 standard, information is defined as formalized representation of data, suitable both for human and machine processing. The information container is any named and permanent collection of structured and unstructured information retrievable from file, cloud storage or application. The structured information implies that information is machine interpretable and unstructured covers human interpretable one. The information model is a collection of structured and unstructured information containers, which means that it includes both machine interpretable information and digital representation of documents. Also, the term "common data environment" is used to describe the IT structure that supports the BIM process and "data" denotes basic unstructured form of digital representation, as it is shown in this paper.

It is clear that the existing legislation requires reliance on documents, but the lack of a clear distinction between human and computer interpretable information creates ambiguities that lead to inefficiencies in the whole process. Deceived by the lack of a clear distinction, many may assume that it is enough to just digitize existing documents and that this directly leads to an efficient BIM process. But without the ability for computers to interpret information, we only come to the point that computer technologies create huge amounts of information that people have to interpret and then re-structure the data according to the needs of certain BIM applications. That way, people becomes slaves of machine generated data instead of masters of information.

The process of information requirements determination represents a significant part of the 19650-2 standard. It requires all participants in the process to think through all the information that the asset will require during its lifetime. As will the final digital asset models increasingly replace traditional documentation, so will raise the understanding of the necessity of computer-interpretable information. The requirements for the existence of digital twin models that enable the management of smart buildings are also contributing to this transition.

CONCLUSION

The lack of understanding of the role of the digital model and computer-interpretable information multiplies problems that were present in the design and construction of the building. Instead of the model helping the process, it becomes a problem in itself. BIM analysis often does not refer to the quality of the building being designed, but to the quality of the BIM model itself, whether the elements are repeated,

whether they have the necessary information, etc. Any use of documents complicates this process because it involves the interpretation of information by people and then their translation into structures that computers can process.
It is obviously necessary to keep the documents due to legal regulations. Also, in order for a person to understand the digital model, it is necessary to translate it into the document. However, in order to realize all the potentials that computer technology brings, it is necessary to clearly understand the distinctions between documents and digital models. It is time to stop thinking about documents and file formats and start thinking about effective workflows among people enabled by computer interpretable information and model based processes. The future of AEC competence lies in the ability to create models rather than documentation.

REFERENCES

BIM Dictionary. (2019). *Building Information Modelling (BIM)* https://bimdictionary.com/en/building-information-modelling/1 published July 31, 2019, accessed 29 September 2020.

Boden, M.A. (1988). *Computer Models of Mind*. Cambridge: Cambridge University Press.

bSI. (2020a). *BIM Collaboration Format (BCF),* https://technical.buildingsmart.org/standards/bcf/ accessed 29 September 2020.

bSI. (2020b). *buildingSMART Data Dictionary,* https://www.buildingsmart.org/users/services/buildingsmart-data-dictionary/ accessed 29 September 2020.

bSI. (2020c). *Model View Definition (MVD),* https://technical.buildingsmart.org/standards/ifc/mvd/ accessed 29 September 2020.

Burke, T. (1995). Dance floor blues: The case for a social AI. *Stanford Humanities Review,* 4: (2). pp. 221–248.

Cowan, N. (2001). "The Magical Number 4 in Short-Term Memory: A Reconsideration of Mental Storage Capacity." *The Behavioral and brain sciences,* 24. pp. 87-114.

Frigg, R. and Hartmann S. "Models in Science." in Zalta, E.N. (ed.), *The Stanford Encyclopedia of Philosophy (Spring 2020 Edition),* https://plato.stanford.edu/entries/models-science/ First published Mon Feb 27, 2006; substantive revision Tue Feb 4, 2020, accessed 29 September 2020.

ISO. (2007). *ISO 12006-3:2007 Building construction — Organization of information about construction works — Part 3: Framework for object-oriented information*, https://www.iso.org/standard/38706 html accessed 29 September 2020.

ISO. (2018a). *ISO 16739-1:2018 Industry Foundation Classes (IFC) for data sharing in the construction and facility management industries — Part 1: Data schema*, https://www.iso.org/standard/70303 html accessed 29 September 2020.

ISO, (2018b). *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 29 September 2020.

ISO, (2018c). *ISO 19650-2:2018 Organization and digitization of information about buildings and civil engineering works, including building information modelling (BIM) — Information management using building information modelling — Part 2: Delivery phase of the assets,* https://www.iso.org/standard/68080.html accessed 29 September 2020.

ISO, (2020a). *ISO 23386:2020 Building information modelling and other digital processes used in construction — Methodology to describe, author and maintain properties in interconnected data dictionaries,* https://www.iso.org/standard/75401.html accessed 29 September 2020.

ISO, (2020b). *ISO 23387:2020 Building information modelling (BIM) — Data templates for construction objects used in the life cycle of built assets — Concepts and principles* https://www.iso.org/standard/75403.html accessed 29 September 2020.

Kirsh, D. (1991). Today the earwig, tomorrow man. *Artificial Intelligence,* 47: (1-3). pp. 161-184.

Miller, G. A. (1956). "The magical number seven, plus or minus two: Some limits on our capacity for processing information." *Psychological Review,* 63, pp. 81–97.

NIBS, (2019). *The National BIM Standard-United States® (NBIMS-US™ V3*), https://www.nationalbimstandard.org/ accessed 29 September 2020.

Petzold, C. (2008). *The Annotated Turing: A Guided Tour Through Alan Turing's Historic Paper on Computability and the Turing Machine*. Indianapolis: John Wiley & Sons, Inc.

Rogers, D.F. and Adams, J.A. (1976). *Mathematical elements for computer graphics,* New York: McGraw-Hill.

Sloman, A. (1978). *The Computer Revolution in Philosophy: Philosophy, Science, and Models of Mind.* Brighton: Harvester Press.

Turing, A.M. (1936), "On Computable Numbers, with an Application to the Entscheidungsproblem", *Proceedings of the London Mathematical Society,* 2, (42), pp. 230–65.

WHATWG. (2020). *HTML Living Standard*, https://html.spec.whatwg.org/ Last Updated 28 September 2020, accessed 29 September 2020.

W3C, (2015). *Semantic Web,* https://www.w3.org/standards/semanticweb/ accessed 29 September 2020.

Zins, C. (2007). Conceptual approaches for defining data, information, and knowledge. *Journal of the American Society for Information Science and Technology*, 58: (4). pp. 479-493

Draft – published at: On Architecture: Learning Architecture Proceedings, R. Bogdanovic (ed.) 2020, Belgrade, pp. 35-41. ISBN 978-86-89111-23-1