Automation in Construction 80 (2017) 22–36

Contents lists available at ScienceDirect

Automation in Construction

journal homepage: www.elsevier.com/locate/autcon

A shared ontology approach to semantic representation of BIM data

Mehrdad Niknam [a,*], Saeed Karshenas [b]

[a] EnTech Engineering, P.C. (EnTech), 11 Broadway, 21st Floor, New York, NY 10004, United States
[b] Department of Civil, Construction, and Environmental Engineering, Marquette University, Haggerty Hall, Milwaukee, WI 53233, United States

ARTICLE INFO

Article history:
Received 14 November 2015
Received in revised form 24 February 2017
Accepted 15 March 2017
Available online xxxx

Keywords:
Building information modeling
BIM
AEC-FM
Interoperability
Semantic web
Ontology
RDF/OWL

ABSTRACT

Architecture, engineering, construction and facility management (AEC-FM) projects involve a large number of participants that must exchange information and combine their knowledge for successful completion of a project. Currently, most of the AEC-FM domains store their information about a project in text documents or use XML, relational, or object-oriented formats that make information integration difficult. The AEC-FM industry is not taking advantage of the full potential of the Semantic Web for streamlining sharing, connecting, and combining information from different domains. The Semantic Web is designed to solve the information integration problem by creating a web of structured and connected data that can be processed by machines. It allows combining information from different sources with different underlying schemas distributed over the Internet. In the Semantic Web, all data instances and data schema are stored in a graph data store, which makes it easy to merge data from different sources.

This paper presents a shared ontology approach to semantic representation of building information. The semantic representation of building information facilitates finding and integrating building information distributed in several knowledge bases. A case study demonstrates the development of a semantic based building design knowledge base.

1. Introduction

With the advent of digital modeling, physical and functional characteristics of a building can be digitally represented, which is also referred to as a Building Information Model (BIM). In a BIM platform (e.g. Autodesk Revit), a user can digitally represent physical elements (e.g. wall, door, column, and slab) that exist in a building project and define element properties such as its type, location, level, material, geometry, and relations to other elements.

A user of a BIM platform is limited to the platform schema and is not able to define information that falls out of the platform schema [1]. IFC representation of BIM data was developed to facilitate interoperability among AEC-FM software applications. However, in order to exchange information between two applications, both applications must understand the IFC schema. To overcome the limitations of the IFC data model for information exchange, a use case approach has been developed [2] that requires an Information Delivery Manual (IDM) and a Model View Definition (MVD) prepared by domain experts. In order to use this approach, an IDM and a MVD must be prepared that specify the data that needs to be exchanged between two applications. The IDM and MVD approach provides a static definition of the data that can be exchanged among two applications and do not support rule-based automated extraction of information from several sources [3].

Architecture, engineering, construction and facility management (AEC-FM) projects involve several experts in different domains who collaborate with each other to deliver a project. Fig. 1 shows information flow in construction projects and how experts in different construction domains (e.g. estimating, scheduling, supply chain, and fabrication domains) query a BIM as a centralized 3D model of the project and use the extracted data to create the information that represents their views of the project. When the information about various project views are stored in heterogeneous file formats, integrating the information about a building element requires accessing several project documents and manually extracting and combining the necessary information.

The Semantic Web technology allows anyone to express a piece of data about some entity in a way that can be combined with information that other sources provide [4]. Therefore, a semantic representation of a building information model would allow anyone involved in a building project to express his/her information about a building element in a way that can be easily combined with information provided by others. This paper presents a semantics-based approach to modeling building information that would facilitate sharing information among various AEC-FM domains.

2. Integrating distributed sources of data

Lack of information sharing causes delay, duplication, and increases construction cost. Several IT applications have been developed to help

* Corresponding author.
E-mail address: mehrdad.niknam@gmail.com (M. Niknam).

http://dx.doi.org/10.1016/j.autcon.2017.03.013
0926-5805/

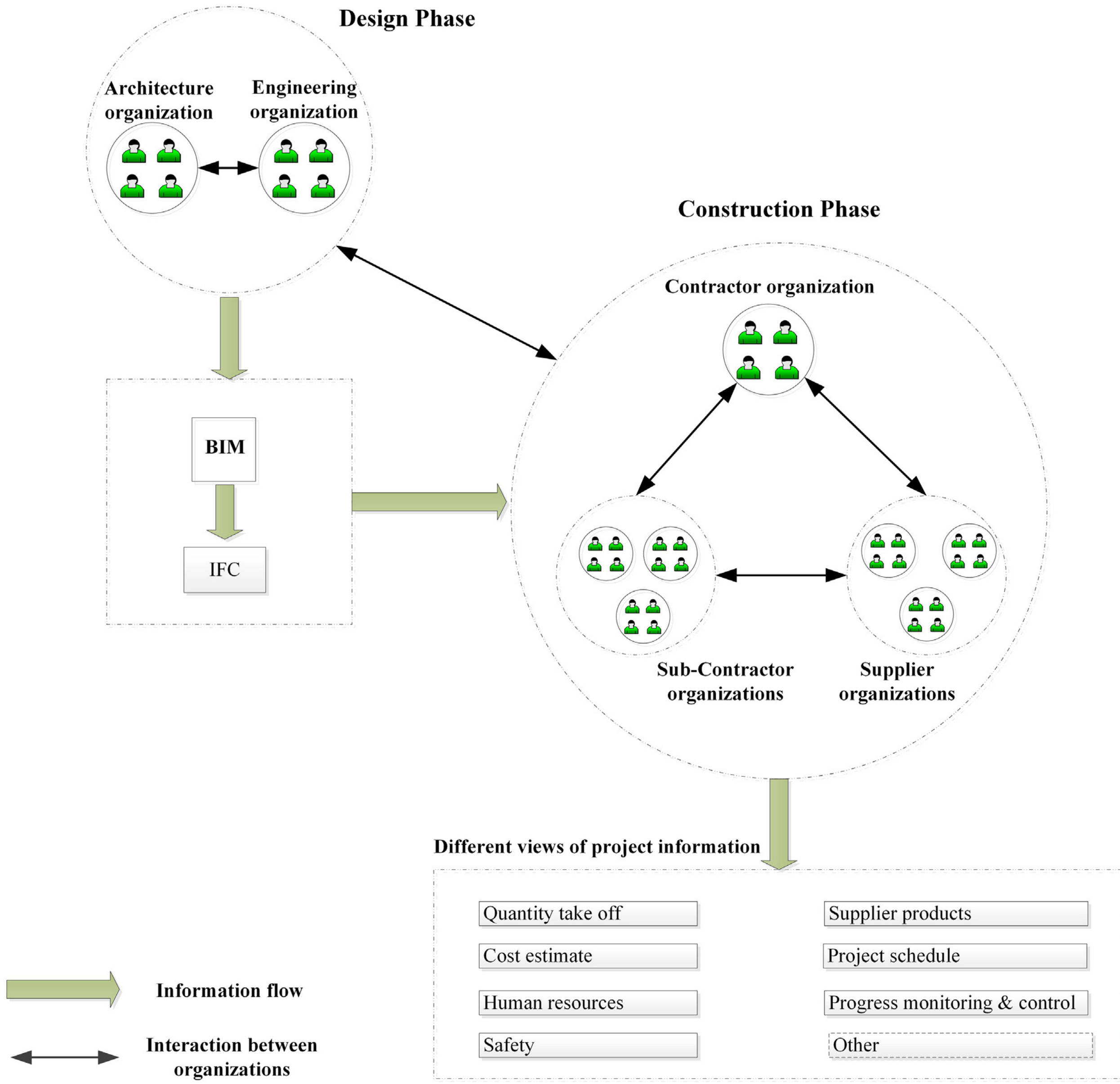

Fig. 1. Information flow in construction projects.

the construction industry in design (e.g., Autodesk and Bentley), project planning and scheduling (e.g., Microsoft project, Primavera), cost estimating (e.g., WinEstimator), collaboration (e.g., Project extranets), and field technologies (e.g., RFID, Mobile phones). Most of these tools serve a single business function and do not provide means for information flow across functional and organizational boundaries [5].

Currently, different domains in AEC-FM industry store their data about a project in heterogeneous data formats [3]. For example, BIM data are stored in object (e.g., IFC), XML (e.g., ifcXML, gbXML), or relational (e.g., ODBC) databases; cost estimating and project schedule data are kept in relational databases; and material suppliers' product data are usually provided in text, HTML, XML, or relational formats.

When different sources of data are stored in heterogeneous data formats, computers cannot easily integrate the data. This is due to a number of factors including: (1) the schemas of the data sources are local and cannot be shared on the Internet among computer applications, (2) different data sources may use different vocabularies to refer to the same entity or a word may have different meanings in two databases, and (3) it is not easy to dynamically modify database schemas since data schemas map to the object models of the computer programs that use them. In other words, each domain develops its own schema to represent domain properties for the same objects. For example, the same building element is modeled using two different class hierarchies in the estimating and the scheduling domains. This makes integration of the schedule and cost properties of the same building elements very difficult. The problems associated with integrating data stored in relational or object-oriented databases are discussed in detail by W3C [6].

Extensible Markup Language (XML) is a serialization format that addresses some of the problems associated with enabling different programs and computers to communicate with each other [7]. An XML schema defines the structure of an XML document. However, XML documents defined by a schema are not extensible [8]. Adding a simple attribute to an XML document requires rewriting all applications using the document in order to be able to read the modified document.

The Semantic Web technology provides a common framework that allows data to be shared and reused across applications, enterprise and community boundaries [9]. In the World Wide Web, anybody can say anything about any topic and publish it as a web page. The Semantic Web extends principles of the Web from documents to data. Data would be related to one another just as documents are already [4,10].