LDAC 2012

First International Workshop on Linked Data in Architecture and Construction

Workshop Report

http://multimedialab.elis.ugent.be/LDAC2012/
UGent Multimedia Lab - UGent SmartLab
Ghent University, Ghent, Belgium
March 28-29, 2012

Abstract

The first International Workshop on Linked Data in Architecture and Construction is a two-day workshop that addresses the usage and role of linked data in the context of architecture, engineering and construction (AEC). This workshop gathers researchers working on this specific topic, thereby aiming to bring together diverse ideas about ways in which linked data and semantic web technologies can enhance information exchange in the AEC domain.

The first edition of this workshop was sponsored and hosted by the Interdisciplinary Institute for Broadband Technology (IBBT), and took place on 28 and 29 March 2012 in Ghent, Belgium. The workshop focused specifically on the issue of interoperability in the AEC domain, which is typically encountered when information from multiple domains is merged. The encountered issues are listed and discussed, and suggestions are made on how to address these issues using semantic web technologies.

Workshop Chair

Pieter Pauwels
UGent SmartLab, Ghent University, Ghent, Belgium

Organising Committee

Pieter Pauwels
UGent SmartLab, Ghent University, Ghent, Belgium

Davy Van Deursen
UGent Multimedia Lab, Ghent University - IBBT, Ghent, Belgium

Erik Mannens
UGent Multimedia Lab, Ghent University - IBBT, Ghent, Belgium

Programme Committee

Edward Corry
Informatics Research Unit for Sustainable Engineering (IRUSE), National University of Ireland, Galway, Ireland

Gonçal Costa
Architecture Representation Computation (ARC) research group, Universitat Ramon Llull, Barcelona, Spain

Edward Curry
Digital Enterprise Research Institute (DERI), National University of Ireland, Galway, Ireland

Leif Granholm
Tekla Finland, Finland

Nam Vu Hoang
Laboratory of Software Technology, Aalto University, Helsinki, Finland

Oraskari Jyrki
Laboratory of Software Technology, Aalto University, Helsinki, Finland

Marcus Keane
Informatics Research Unit for Sustainable Engineering (IRUSE), National University of Ireland, Galway, Ireland

Leandro Madrazo
Architecture Representation Computation (ARC) research group, Universitat Ramon Llull, Barcelona, Spain

James O'Donnell
Simulation Research Group, Lawrence Berkeley National Laboratory (LBNL), Berkeley, CA, USA

Seppo Törmä
Laboratory of Software Technology, Aalto University, Helsinki, Finland

Programme

Wednesday 28 March 2012

9.00	Welcome and Coffee
9.30	IFC-to-RDF: Adaptation, Aggregation and Enrichment
	Pieter Pauwels, Davy Van Deursen
10.30	Linked Data Projects in the ARC Research Group
	Leandro Madrazo
11.30	Open Product Modelling and Interoperability in the AEC sector
	Gonçal Costa
12.30	Lunch Break
14.00	Building Optimisation using Scenario Modeling and Linked Data
	Edward Corry, Edward Curry
15.15	Coffee Break
15.45	Distributed Transactional Building Information Management
	Seppo Törmä, Oraskari Jyrki, Nam Vu Hoang
17.00	End
19.00	Workshop Dinner

Thursday 29 March 2012

9.30	Welcome and Coffee
10.00	Position Paper(s) Session
12.00	Lunch Break
14.00	Discussion and Workshop Conclusion
	Teleconference Session
16.00	End

Contents

IFC/RDF: Adaptation, Aggregation and Enrichment

Pieter Pauwels
Department of Architecture and Urban Planning
Ghent University, Ghent, Belgium
pipauwel.pauwels@ugent.be

Davy Van Deursen
Multimedia Lab
Ghent University - IBBT, Ghent, Belgium
davy.vandeursen@ugent.be

Abstract

The usage of semantic web technologies might enable addressing existing interoperability issues in the domain of architecture, engineering and construction, because they allow linking diverse information models. We indicate here how IFC models can be made available as RDF graphs in the semantic web and how they can be linked to other information. As such, their capabilities in addressing interoperability issues can be assessed.

The Industry Foundation Classes, Interoperability and the Semantic Web

The domain of architecture, engineering and construction (AEC) involves all kinds of information, including material characteristics, legal regulations, three-dimensional parameters, and so forth. Traditional information systems typically represent only part of this information in their underlying information structures. Additionally, each of these information systems has its own way of describing and managing the considered information. As a result, a significant number of different information models is typically available for one and the same AEC project.

It is hard to correctly combine these available information models using traditional techniques, resulting in a low level of interoperability among information systems. Consequently, information often needs to be remodelled in an AEC context. This in turn leads to (1) a significant loss of time and resources, and (2) to an increased risk of construction errors and misconceptions in the design (Gallagher et al. 2004). Diverse strategies exist to address interoperability issues, including the usage of semantic web technologies for making direct links between diverse information models. We presented such a 'linked data' strategy earlier (Pauwels, De Meyer, and Van Campenhout 2010). The success of this strategy depends on the extent to which alternative information models can be represented and interrelated so that they form one combined 'semantic web'. In this paper, we briefly outline the software built to test this strategy and give an indication of what this software can be used for.

The IFC-to-RDF Conversion Service

In the above research context, we decided to build an ontology of the Industry Foundation Classes (IFC) (Liebich et al. 2012) using the web ontology language (OWL) (McGuinness and van Harmelen 2009). This ontology can then serve as an alternative representation of the EXPRESS schema of IFC. Similar work was done earlier by (Schevers and Drogemuller 2005, Beetz, Van Leeuwen, and De Vries 2009). Using this OWL ontology of IFC, one is theoretically able to represent building information according to the IFC schema in a semantic web format and link this to alternative representations of building information. As such, it can be tested to what extent alternative information models can be represented and interrelated when using the linked data strategy outlined earlier. As such, it may lead to an appropriate assessment of the extent to which semantic web technologies can address the interoperability issue for the AEC domain.

The IFC Ontology

To construct an IFC ontology in OWL, we followed an approach that is largely similar to the approach discussed for EXPRESS schemas in (Beetz, Van Leeuwen, and De Vries 2009). In this approach, the EXPRESS file is mapped as good as possible onto an OWL ontology, meaning that every EXPRESS element is mapped onto its nearest equivalent in OWL. This mapping process is not straightforward, as is the case for any mapping process. However, we are not targeting a perfect mapping, we are targeting a sufficiently good mapping, of which results are usable within the above research context.

In our conversion process of the EXPRESS schema to an OWL ontology, the three following steps can be distinguished.

Generation of Classes and Properties For each `ENTITY` definition in the EXPRESS schema, a corresponding `owl:Class` is generated with the appropriate properties. EXPRESS attributes are converted into the corresponding OWL properties (`owl:DatatypeProperties` and `owl:ObjectProperties`). For the transformation of simple data types (`REAL`, `INTEGER`, `STRING`, and so forth), we use an `owl:DatatypeProperty`.

This apparently is a straightforward process. However, as (Beetz, Van Leeuwen, and De Vries 2009) also outlines, special care is needed regarding naming conflicts. Attributes in EXPRESS are local to their entities, whereas properties in OWL are global in the ontology. Hence, attributes with the same name often relate to different entities, resulting in conflicts in an OWL context. In our case, we decided to append integers to the OWL property names for which a naming conflict is detected.

Generation of Basic Restrictions for Classes and Properties After generation of the `owl:Classes` and `Properties`, a second step is started in the conversion process, in which the appropriate basic restrictions are generated and added to the ontology. Using the `rdfs:subClassOf` construct, a hierarchical ontology structure is created for supertype and subtype relations in EXPRESS. Currently, `ABSTRACT SUPERTYPE` constructs in EXPRESS are represented by the appropriate combination of `rdfs:subClassOf` and `owl:disjoint` constructs in OWL.

In this step, also the `rdfs:Range` and `rdfs:Domain` constructs are completed for the generated properties. Simple `owl:DatatypeProperties` are completed with the corresponding XML schema data type in their ranges. Properties that represent `SELECT` types in the EXPRESS schema are completed with an `owl:unionOf` construct in their range. This is not the most elegant solution, because ... However, this was considered the best mapping to an OWL construct currently available. For `ENUM` types in the EXPRESS schema, `rdfs:subClasses` are used in combination with the `owl:oneOf` construct. `LIST` types are translated into `rdf:Lists`.

A certain level of semantics is lost in the direct conversion of simple data type entities into corresponding `owl:DatatypeProperties`. Namely, the name of the EXPRESS relation (e.g. `MainPlaneAngle`) is preserved, but the attribute name (e.g. `IfcPlaneAngleMeasure`) is replaced by a simple data type indication only (e.g. `xsd:double`). This is considered a minor change in the overall conversion procedure, however, because the name of the EXPRESS relation is still available.

Generation of Advanced Restrictions for Classes and Properties A last step in the conversion process is the generation of advanced restrictions for classes and properties. This step is currently not implemented. These advanced restrictions represent some of the more advanced features of the EXPRESS schema. This includes, for instance, cardinality restrictions and value restrictions for properties. Also the `OPTIONAL`, `UNIQUE` and `DERIVE` keywords are currently not included in our OWL ontology and could be considered as advanced restrictions for classes and properties.

A difficulty in implementing these features is the difficulty of choosing between OWL restrictions and rules expressed with a rule language. One of the rule languages available in the semantic web could be used (e.g. RIF, N3Logic, SWRL). It is also questioned to what extent all EXPRESS restrictions *can* be converted into OWL restrictions and/or rules.

Instantiating the IFC Ontology

Although additional work is obviously needed, the available IFC ontology in OWL can be used to generate RDF instances that instantiate this ontology. One could of course implement a modelling environment that relies completely on the IFC ontology and that allows modelling a building directly using the IFC ontology. However, considering our initial research question, we decided to rely on existing building information modelling (BIM) environments and their available IFC export mechanisms for the modelling part. Our work can then be limited to building a file-based conversion service that uses exported IFC files and converts them to RDF graphs.

Also in converting from IFC instances to RDF instances, we use a relatively straightforward method. The generated RDF class instances are named using the applicable EXPRESS `ENTITY` name and the line number in the IFC file. For instance, the line in IFC `#4796 = IFCAXIS2PLACEMENT3D(#3,$,$);` is converted into the RDF concept `ifcAxis2Placement3D_4796`. This is not the optimal conversion method when changes in IFC files need to be tracked, because every time an IFC file is exported from a modelling environment, line numbers in the IFC file change and an obvious link cannot be made.

Using the Resulting IFC/RDF Graph

The RDF-to-IFC conversion service is available online (UGent Multimedialab 2012a). After uploading an IFC file, a conversion process can be initiated. During this conversion process, resulting RDF triples are stored in an online RDF triple store. This triple store and its contents are available through an online SPARQL endpoint (UGent Multimedialab 2012b).

There are diverse ways in which the result can be reused. The most obvious and easiest way is to browse through the information and/or query for information in the RDF graph, in both cases using the SPARQL endpoint. Alternatively, the RDF graph of the building model can be used as a basis for further aggregation and enrichment with related information models. For instance, building elements represented in the IFC/RDF graph can be linked directly to material information or product information described in separate, specialized RDF graphs. This was tested previously for the aggregation of IFC information and specific 'building element information', 'design information', and 'project information' (Pauwels, De Meyer, and Van Campenhout 2010). A similar test was made with specific material information, thereby targeting at the automation of acoustic building performance checking with rule sets in N3 Logic and a reasoning engine (Pauwels et al. 2011).

Discussion and Conclusion

This research indicates that it is possible (1) to represent alternative information models with RDF graphs and (2) to interrelate these RDF graphs into an aggregate or enriched RDF graph. Consequently, the linked data strategy appears a valid approach for addressing existing interoperability issues in the AEC domain.

Care has to be taken, however, of the mapping process between existing information models and RDF representations of these models. This is made clear by the above documentation of (1) our mapping process between the EXPRESS schema of IFC and an OWL ontology of IFC, and (2) our mapping process between native IFC files and corresponding IFC/RDF graphs. A comparison of both mapping processes with previous approaches, e.g. (Beetz, Van Leeuwen, and De Vries 2009), further strengthens this argument.

Nonetheless, even with the resulting shortcomings in the resulting IFC/RDF graphs, significant new possibilities emerge because of the notable aggregation and enrichment possibilities offered by semantic web technologies.

Acknowledgments

The authors gratefully acknowledge the funding support from the Research Foundation - Flanders (FWO).

References

Beetz, J.; Van Leeuwen, J.; and De Vries, B. 2009. ifcowl: A case of transforming express schemas into ontologies. *Artificial Intelligence for Engineering Design, Analysis and Manufacturing* 23:89–101.

Gallagher, M.; O'Connor, A.; Dettbar, J.; and Gilday, L. 2004. Cost analysis of inadequate interoperability in the U.S. capital facilities industry. Technical Report NIST Report GCR 04-867, National Institute of Standards and Technology (NIST).

Liebich, T.; Adachi, Y.; Forester, J.; Hyvarinen, J.; Karstila, K.; Reed, K.; Richter, S.; and Wix, J. 2012. Industry Foundation Classes IFC2x edition 3 technical corrigendum 1. Available online: http://www.buildingsmart-tech.org/ifc/IFC2x3/-TC1/html/index.htm. Last accessed on 22 February 2012.

McGuinness, D., and van Harmelen, F. 2009. OWL 2 Web Ontology Language document overview. W3C recommendation. Available online: http://www.w3.org/TR/owl2-overview/. Last accessed on 22 February 2012.

Pauwels, P.; Van Deursen, D.; Verstraeten, R.; De Roo, J.; De Meyer, R.; Van de Walle, R.; and Van Campenhout, J. 2011. A semantic rule checking environment for building performance checking. *Automation in Construction* 20(5):506–518.

Pauwels, P.; De Meyer, R.; and Van Campenhout, J. 2010. Interoperability for the design and construction industry through semantic web technology. In *SAMT*, 556–560. Washington, DC: IEEE Computer Society.

Schevers, H., and Drogemuller, R. 2005. Converting the industry foundation classes to the web ontology language. In *Proceedings of the First International Conference on Semantics, Knowledge and Grid*, 556–560. Washington, DC: IEEE Computer Society.

UGent Multimedialab. 2012a. IFC-to-RDF service. Available online: http://ninsuna.elis.ugent.be/IfcRDFService/. Last accessed on 22 February 2012.

UGent Multimedialab. 2012b. IFC/RDF SPARQL endpoint. Available online: http://ninsuna.elis.ugent.be/-SPARQLEndpoint/. Last accessed on 22 February 2012.

Open Product Modelling and Interoperability in the AEC sector

Leandro Madrazo and **Gonçal Costa**
Department of Architecture
Enginyeria i Arquitectura La Salle, Barcelona, Spain
madrazo@salle.url.edu, gcosta@salle.url.edu

Abstract

The integration of Information Technologies in the AEC sector can contribute towards improving the industrialized working process. It can make the industrialized working process more open and transparent to enable the participation of multiple actors, and increase the efficiency of the design and construction. Although BIM technologies aim at fulfilling both objectives, there are some challenges associated with interoperability that still should be overcome. In particular, there are some important limitations of the IFC standard with regard to its semantic expressiveness which make the exchange of data across applications, difficult.

This document offers an overview of the alternatives to interlink data generated around different BIM domains using ontologies and semantic web technologies following the Linked Open Data initiative. Specifically, the potential of applying ontologies for product modelling is discussed.

Introduction

A BIM model allows different actors participating in the design and building process to exchange information throughout the whole building lifecycle. This model is a unique representation of the building from which different kinds of information are extracted in the specific format required by each specialist. In this process of exchange, there may intervene different BIM models, created with software from different vendors (Revit, Archicad, Allplan), as well as specific applications (energy simulation, structural analysis, and others) that can either be part of the suite of a particular vendor (e.g. Robot, a structural analysis program for Revit) or not. This information exchange generates problems when coming to both, the interoperability among the BIM models, and among them and the applications that process the extracted data from the model.

Interoperability Issues

In order to achieve this interoperability across BIM platforms, the International Alliance for Interoperability (IAI) created the open standard IFC (Industry Foundation Classes). However, and since the IFC is a unique and neutral scheme, it can only represent the information created under the rules and hierarchy established by a particular BIM software. Nevertheless, the rules embedded in a BIM software are not always captured by the IFC standard, giving rise to problems in the translation between BIM models from different vendors, for example, between Revit and Archicad. For instance, a wall is described in a different way depending on the program: in Revit, the parameters of a wall have five levels of detail in contrast with the fifteen levels in Archicad.

To overcome the limitations which are intrinsic to the IFC standard, some alternative strategies can be considered. The use of ontologies in combination with semantic web technologies is an alternative approach for improving this interoperability bringing a higher level of expressiveness. These technologies provide a powerful mechanism to enhance the information of a building model adding semantics in the form of concepts, properties and rules. Instead of relying on a centralized standard as the IFC, ontologies would enable to connect IFC-based models among them and with other models. In this way, it may be possible to combine the standardization that IFC provides with the flexibility ontologies facilitate.

Ontologies and Semantic Web Technologies

With semantic technologies it is possible to achieve certain level of commitment between a common language and the free interpretation of it (Berners-Lee, Hendler, and Lassila 2001). These technologies integrate OWL (Web Ontology Language), a language designed for publishing and sharing data that is specified by ontologies, and RDF (Resource Description Framework) which provides the basis on which to build the OWL. In this way, whereas RDF provides the necessary structure and syntax for exchanging data, OWL enables to deploy ontology graphs to represent building information with a higher level of expressivity than in the one facilitated by a BIM/IFC model. Interoperability among applications could be improved, depending on the predisposition and the available mechanisms facilitated by a BIM software which enable the integration of semantic web languages and technologies.

Another possibility to define an interoperability framework is to design an upper (or foundational) ontology that sets a rule base to gain a semantic control of the models which will be subject to these rules. There are some issues to consider in this scenario. One of them is how to create an appropriate vocabulary to define generic concepts for an upper ontology layer. This vocabulary should be accepted by all; or, at least, should count with the widest possible acceptance. The inconsistent use of terminology has long been a significant barrier to effective communication between and among both, persons and computers (Uschold 2000). Assuming that a set of terms can be standardized, an upper ontology should be ensuring that any specification of knowledge can be embedded in a model that is under the control of this layer.

Linked Open Data (LOD)

Semantic web languages such as OWL or RDF provide some kind of data structure which is not always open or linked. Linked data is a concept that arises within the paradigm of semantic web. It is based on the idea of creating a common database through the publication of distributed databases which are linked together. The purpose of the Open Data is to make data available on the Internet to any user with no kind of license or restriction. The union of these two concepts constitutes the Linked Open Data (LOD).

Applying the principles of LOD to the AEC sector would enable to have interlinked BIM models accessible on the web using semantic technologies. This could give rise to new datasets and new functionalities which are not possible with the existing BIM-centered technologies.

Product Modelling: State of the Art

The use of semantic technologies in the AEC sector could be especially relevant for product modelling. In particular, they could be used to create component catalogues whose contents are linked to BIM models. A representative example of the application of ontologies to product modelling is the SWOP project (Böhms, Bonsma, and Bourdeau 2009), which aims to develop a semantic web-based open engineering platform. The Product Modelling Ontology (PMO) developed by TNO has sufficient capabilities to make an end-user product ontology for any parametric/configurable product type.

On the other hand, Autodesk Seek (seek.autodesk.com) provides a large database of components described by means of CAD/BIM files, 2D drawings, visual images, and product specification data. Even though it is possible to upload and download these files, this catalogue does not provide mechanisms to set up connections between the components and the BIM model where they are inserted.

BIMobject (bimobject.com) is another example of a commercial online catalogue of components with information about building products designed and manufactured by different companies. The product information stored in the cloud is viewable on the web and can be downloaded in various BIM formats. Furthermore, it is possible to place component models directly within the actual schema of a BIM program, such as Revit.

Conclusions

The capabilities of BIM programs to assure interoperability across different vendor's platforms and applications are limited. In order to overcome the inherent limitations of a centralized model, such as the IFC standard, an alternative is to use ontologies and semantic web technologies as a mechanism to obtain a more flexible data modelling.

An alternative for interoperability based on the IFC standard passes through semantic web technologies. In this context it is possible to define a framework for interoperability through an upper ontology design comprised of a set of terms and rules. However, there is no clear solution for this scenario in the BIM context since their semantic integration remains a challenge. Difficulties arise on how to integrate different heterogeneous information in a global domain. A way to lighten this problem is to define a methodology for the model's specification process. Another approach is to define a more specific framework domain, for example, focused on building components.

Acknowledgements

This report is a result of the presentation made by Gonçal Costa, from the research group ARC Enginyeria i Arquitectura, La Salle, during the Linked Data in Architecture and Construction Workshop, held on 28th-29th March 2012, at the University of Ghent (Belgium).

References

Berners-Lee, T.; Hendler, J.; and Lassila, O. 2001. The semantic web. *Scientific American* 284:35–43.

Böhms, M.; Bonsma, P.; and Bourdeau, M. 2009. Semantic product modeling and configuration: Challenges and opportunities. *Journal of Information Technology in Construction* 14:507–525.

Uschold, K. 2000. Creating, integrating and maintaining local and global ontologies. In W., H., ed., *14th European Conference on Artificial Intelligence, ECAI'00.* IOS Press.

Building Optimisation using Scenario Modeling and Linked Data

Edward Curry
Digital Enterprise Research Institute
NUI Galway, Galway, Ireland
ed.curry@deri.org

James O'Donnell
Building Technologies Group
LBNL, Berkeley, CA, USA
jtodonnell@lbl.gov

Edward Corry
IRUSE
NUI Galway, Galway, Ireland
edward.corry@nuigalway.ie

Abstract

As buildings become more complex, it becomes more difficult to manage and operate them effectively. The holistic management and maintenance of facilities is a multi-domain problem encompassing financial accounting, building maintenance, human resources, asset management and code compliance, affecting different stakeholders in different ways. One technique, called scenario modelling, customises data-driven decision support for building managers during building operation. However, current implementations of scenario modeling have been limited to data from Building Management Systems with little interaction with other relevant data sources due to interoperability issues. Linked data helps to overcome interoperability challenges to enable data from multiple domains to be merged into holistic scenario models for different stakeholders of the building. The approach is demonstrated using an owner-occupied office building.

Introduction

The holistic management and maintenance of facilities is a complex problem encompassing financial accounting, building maintenance, human resources, asset management and code compliance, affecting different stakeholders in different ways. The type of information needed by each stakeholder is different and varies in the level of complexity required. The skill-set of each stakeholder is also markedly different and each has very different motivations for accessing information on the building. Navigating through this mass of data in a coherent manner to derive information and tailoring this output to specified end-users is a challenge. Different stakeholders need different views of the information. For example, the Financial Controller is concerned with cost metrics, the Human Resource function is concerned with issues like occupancy patterns, building occupants are primarily concerned with comfort, whilst the owner is motivated by the overall efficiency of the building. There are a variety of measurement methodologies that can be utilized to quantify each of these, but when considered independently, it is difficult to get a complete picture of the building. There is a clear need to define the building's operational strategy in a comprehensive and structured manner with decision support that provides relevant information to stakeholders from each domain.

Decision Support Systems & Building Operation

Unavailable, unreliable, and inaccurate performance information is a major cause of inefficient building operation (O'Donnell 2009). Information used by building managers must be trustworthy, but there are no standards currently available for analyzing and transforming building performance data and information. In addition, current methods and tools fail to account for the profile of building managers, both in terms of the operational context of their role, and their typical technical and educational background (O'Donnell 2009). Building operators tend to lack the level of information necessary to make fully informed decisions and routinely make decisions based on intuition and experience, rather than on quantifiable metrics. Consequently, decisions taken by building managers are often ad-hoc, arbitrary, and incomplete (Neumann and Jacob 2010).

Scenario Modelling

Scenario Modelling allows for the holistic analysis of building performance using quantifiable performance metrics incorporated into a functional model. Scenario modeling enables the explicit and unambiguous coupling of building functions with other pivotal aspects of building operation, see Fig. 1, in a method that specifically considers the education and technical expertise of building managers. This new method captures, transforms, and communicates the complex interdependencies of environmental and energy management in buildings through an easily navigable, holistic, and reproducible checking mechanism that compares actual performance with predicted performance and completes the "plan-do-check- act" cycle for building managers (ISO 2011). Scenario modelling forms the basis for effective decision support for building managers.

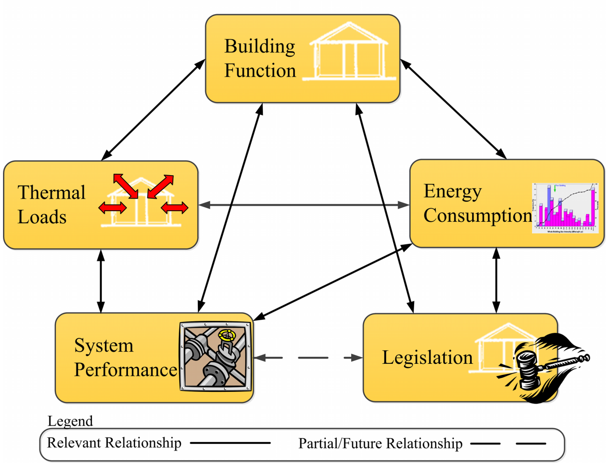


Figure 1: Holistic building performance analysis relies on an understanding of five key performance aspects.

The power of this technique lies in the manner a formalised measurement and assessment framework can be created around disparate data sources. However, the lack of informational interoperability has restricted the level of automated evaluation possible for particular performance metrics.

The task based nature of the Building Lifecycle (BLC) leads to information being gathered independently several times for specific tasks across various domains. The type of cross domain analysis necessary to drive decision support systems becomes increasingly difficult with a lack of interoperability.

Challenges with Interoperability

Building Information Models (BIM) are a relatively new concept and offer a mechanism through which a complete set of project data can exist in one easily accessible format for building managers (Eastman et al. 2011). However, within the wider-context of the organization a BIM is only one silo of static information, usually captured during the design and construction phase of the BLC. Other relevant information must also be utilized to optimize both the building and organization itself. This information, which may also exist in separate data silos, includes payroll, human resources, production systems, ordering systems, resource-planning systems, etc. Each of these systems may be implemented with different incompatible technologies and data formats, making it difficult to interoperate the data. Due to the critical lack of information interoperability amongst these data silos, it is quite difficult to get a complete cross-domain view of a building (Gallagher et al. 2004).

The open Industry Foundation Classes (IFC) standard promotes interoperability within the building and construction domain, and for BIMs in particular. However, IFC by itself is not sufficient to enable interoperability with systems outside of the Architecture Engineering Construction (AEC) domain, or with systems that are dynamically producing data during the operational phase (i.e. sensor/meters). If powerful decision support solutions like Scenario Modelling are to be implemented a more flexible approach to data interoperability is needed.

Linked Building Data

Semantic Web technologies and standards are playing an important role to simplify the sharing of large quantities of data on the Web based upon W3C standards. The Resource Description Framework (RDF) standard provides a common interoperable format and model for data linking and sharing on the Web. Linked Data is a best practice approach used to expose, share, and connect data on the Web. Linked Data has the following characteristics:

- *Open:* Linked data is accessible through an unlimited variety of applications because it is expressed in open, non-proprietary formats.
- *Modular:* Linked data can be combined (mashed-up) with any other pieces of linked data. No advance planning is required to integrate these data sources as long as they both use linked data standards.
- *Scalable:* It's easy to add and connect more linked data to existing linked data, even when the terms and definitions that are used change over time.

Linked data provides a mechanism through which all silos can exist in a homogeneous format. Most importantly, linked data principles identify common elements between silos, and where possible interlink silos. Representing building data, such as a BIM, within the linked data format, will allow it to be combined with linked data from other relevant silos. In doing so, organizations can generate and extract additional value from current stand-alone repositories, across multiple domains. The resulting merged cross-domain data provides a holistic view of the building's operations, which can have added value for domain stakeholders throughout the organization. Linking building data together can build a holistic view of the building, allowing broader context to be used within scenario modelling decision support. The remainder of this paper briefly describes the approach and demonstrates the concept with an owner-occupied building.

DERI Building Use Case

In order to support the argument for Linking Building Data a proof-of-concept has been developed for the Digital Enterprise Research Institute office building at the National University of Ireland, Galway. The approach was implemented using the Linked dataspace for Energy Intelligence (LEI). This section discusses the LEI system architecture, and a building energy management application built using the resulting linked building data.

Architecture

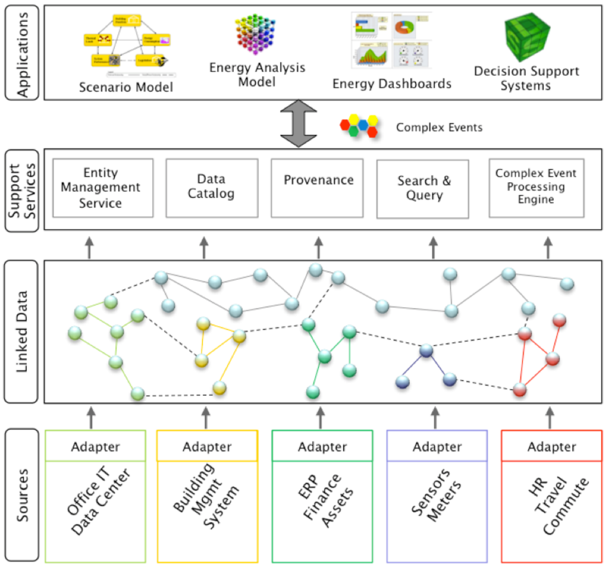


Figure 2: Linked dataspace for Energy Intelligence.

Fig. 2 illustrates the architecture of the OEIP, the main components are:

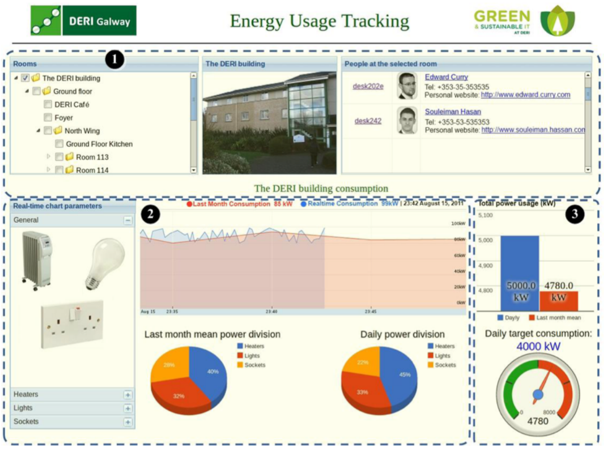
- *Data Sources:* At the bottom of the architecture are the data sources coming from legacy information systems, as well as sensor networks in the building. There are a variety of data sources from multiple domains. Linked data adapters are used to expose these sources using the linked data principles through the use of unique URIs, RDF, etc.
- *Linked Data Cloud:* The resulting outcome of linked data adapters over data sources is a cloud of interlinked web resources. This linked data cloud is rich with knowledge and semantics about the building, building-related performance indicators, and other contextual data. The linked data cloud forms the basis for real-time building analytics and other decision support applications with the help of some support services.
- *Support Services:* Developing applications that use the linked data cloud is simplified through the use of support services. Services are available for provenance information (Freitas et al. 2011), entity reconciliation, data curation, retrieval, and discovery of data and data sources. A

CEP engine (Hasan et al. 2011) supports real-time aggregation and abstraction over dynamic streams (i.e. energy sensors) in the linked data cloud.

- *Applications & Analytics:* At the top of the architecture are a range of applications and analytics built using the linked data cloud. Within DERI a number of applications utilize the linked data cloud including, data center energy management (Curry et al. 2012), IT-energy management, and personal energy management. A Building Energy Explorer application was also developed.

Building Energy Explorer

The Building Energy Explorer allows users to understand the cause and effect of energy consumption within the DERI building. The objective of the explorer is to help users identify energy leaks and non-ecological actions within the DERI building. The explorer makes extensive use of the merged data within the linked data cloud (i.e. people, projects, teams, building layout, etc) and combines it with energy consumption sensor data. It then presents it in an actionable manner that requires minimal effort for users to leverage within energy-related decision-making.



desk202e Edward Curry
Tel: +353-35-353535
Personal website: http://www.edward.curry.com

desk242 Souleiman Hasan
Tel: +353-53-535353
Personal website: http://www.souleiman.hassan.con

Figure 3: Facilities energy tracking dashboard.

The main screen of the dashboard is presented in Fig. 3, within box (1) data from rooms, people, and groups can be seen; it is used to add context to the energy consumptions readings. In (2) historical usage along with real-time instant measures from the energy sensors are shown, along with a breakout for consumption type (lights, heat, sockets). The interface also displays the output of the Energy Situation Awareness Service via a widget in (3). The service is based on a scenario model to perform energy situation assessment by comparing the accumulative consumption with historical usage data, and usage targets, to detect high usage situations. In the widget, two bars are shown aside to show the daily-accumulated energy usage in comparison with the monthly average which gives an idea about the amount of deviation in the consumption pattern. Other scenarios can be easily defined based on different performance criteria including cost, occupancy comfort, and average energy usage per occupant.

Conclusions

The combination of Scenario Modelling and Linked Data offers a promising approach for the design of building decision support systems. Scenario modelling helps buildings operators to understand the complex interdependencies of environmental and energy data in buildings. Linked data overcomes interoperability challenges to enable data from multiple domains to be merged into holistic views for stakeholders of the building. The approach was demonstrated within a proof-of-concept in an owner-occupied building. Future research will focus on the definition of more complex scenario models and a user evaluation of the resulting decision support applications.

Acknowledgments

The work presented in this paper is funded by Science Foundation Ireland under Grant No. SFI/08/CE/I1380 (Lion- 2) and the IRCSET Embark initiative.

References

Curry, E.; Hasan, S.; White, M.; and Melvin, H. 2012. An Environmental Chargeback for Data Center and Cloud Computing Consumers. In Huusko, J.; de Meer, H.; Klingert, S.; and Somov, A., eds., *First International Workshop on Energy-Efficient Data Centers*. Madrid, Spain: Springer Berlin / Heidelberg.

Eastman, C.; Teicholz, P.; Sacks, R.; and Liston, K. 2011. *BIM Handbook: A Guide to Building Information Modeling for Owners, Managers, Designers, Engineers and Contractors*. Wiley.

Freitas, A.; Knap, T.; O'Riain, S.; and Curry, E. 2011. W3P: Building an OPM based provenance model for the Web. *Future Generation Computer Systems* 27(6):766–774.

Gallagher, M. P.; O'Connor, A. C.; Dettbarn, J. L.; and Gilday, L. T. 2004. Cost Analysis of Inadequate Interoperability in the U.S. Capital Facilities Industry. Technical report, National Institute of Standards and Technology (NIST).

Hasan, S.; Curry, E.; Banduk, M.; and O'Riain, S. 2011. Toward Situation Awareness for the Semantic Sensor Web: Complex Event Processing with Dynamic Linked Data Enrichment. In *4th International Workshop on Semantic Sensor Networks 2011 (SSN11)*, 69–81.

ISO. 2011. ISO 50001:2011 - Energy Management Systems – Requirements with Guidance for Use.

Neumann, C., and Jacob, D. 2010. Results of the Project Building EQ. Tools and Methods for Linking EPBD and Continous Commissioning. Technical report, Fraunhofer ISE,.

O'Donnell, J. 2009. *Specification of Optimum Holistic Building Environmental and Energy Performance Information to Support Informed Decision Making*. Ph.d, University College Cork.

Distributed Transactional Building Information Management

Seppo Törmä and **Jyrki Oraskari** and **Nam Vu Hoang**
Department of Computer Science and Engineering
Aalto University, School of Science
seppo.torma@aalto.fi, jyrki.oraskari@aalto.fi, nam.vuhoang@aalto.fi

Abstract

Impressive results have been gained in the field of building information modeling – most notably the standardized schema (IFC) and advanced discipline-specific design tools – but the *management* of building information is still in its infancy. Building information is produced in a distributed and parallel fashion by different discipline experts throughout a building project. These partial models are *interrelated* as they describe the same building, although from different perspectives. How can the interdependencies be represented and the contents of the partial models be exchanged, combined, and modified in a systematic manner? We propose an approach based on *Linked Data*: the partial models are converted into Linked Data datasets (utilizing URIs, RDF, and OWL) and published on the Web (allowing browsing and querying with SPARQL). We describe an ongoing research project to determine the feasibility of this approach. The work is based on a IFC-to-RDF converter, and use of RDF store to maintain the resulting dataset. The research problems – link-type modeling, link generation, change discovery, change management, and information scope management – are described.

Introduction

In every significant building project a huge volume of information both about the building and the construction process is produced in a distributed and loosely coupled manner. Different parties produce *partial models* that describe the building from different perspectives, including requirements, architecture, structural design, mechanical systems, fabrication, and construction process.

As the partial models are representations of the same physical end result, they are obviously *interrelated*. Inconsistencies between the partial models will result in potentially costly problems in the construction phase: spatial collisions, omissions, and other incompatible design decisions. The problems can inflict the schedule of the project and have long-lasting effects on the eventual quality of the building.

Building information modeling (BIM) aims to make the partial models more coordinated. An important part of work in the field has centered on the definition of a common data model for building data. The result is a widely used standard IFC (Industry Foundation Classes) currently supported by major BIM authoring tools. Although BIM tools still maintain the models in their own native formats, they have a capability to export their models as IFC files. The tools can thus map their native entities into a common data model. However, this does not mean that the entities in different models — e.g., the wall in an architectural model and a precast element in a structural model — would have any direct relations in that model. When an entity needs to be modified, the information about its links could aid the management of the change across the models. It could be used to provide the modifier an awareness of a dependency, or to generate notifications or change requests to other parties.

The current way to share building information models is typically based on simple exchange of IFC files. While this approach is compatible with the distributed and loosely-coupled nature of building projects, it does not provide a framework for managing the interlinking of models. The relations of models can be interpreted only by human designers, which limits the support that BIM systems could give to management of cross-model dependencies. Model data could be shared also through other mechanisms — centralized repositories, distributed event-based systems, cloud, federated databases, and so on — but none of these directly addresses the problem of cross-model relations either.

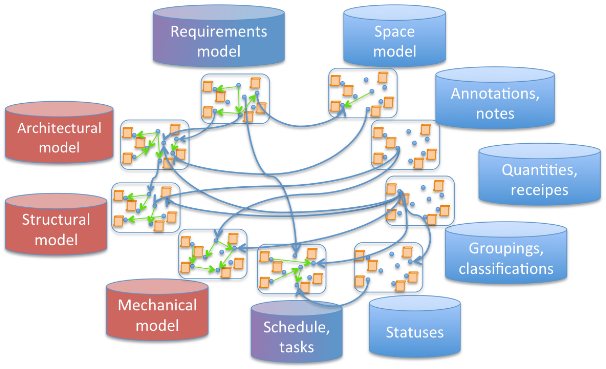


Figure 1: Linked Building Data.

The DRUM project (Distributed Transactional Building Information Management) studies the mechanisms for distributed information management in building information. The focus is on the interrelations of partial models, and the support for cross-model change management.

The starting point of DRUM is to respect the distributed and discipline-centered way of information production and maintenance. That is, the exported IFC models are considered to be shared in a read-only mode and *all changes are always made to native models*, and then exported again to IFC. This is most natural way of working since the maintenance of a model requires discipline-specific expertize and tools. Moreover, changes made to an exported IFC model are difficult to import back to native models, because there is *no proper conversion roundtrip* between native and IFC models. A conversion would loose information and restructure the model, making it very tedious to parse it back into a valid native model.

DRUM studies the problem of model interrelations based on *Linked Data*, a technology that directly addresses the relations of co-existing datasets (Berners-Lee 2006, Bizer, Heath, and Berners-Lee 2009, Heath and Bizer 2011). In addition, by providing a Web-oriented and open linking framework, Linked Data can support the connection of external information sources (annotations, documents, project management tools, etc) to IFC-based models (Fig. 1).

Linked Building Data

Linked Data is based on the Semantic Web technologies designed for the *representation, publication, and browsing of structural data on the Web*. In this approach identifiers are represented with URIs (Uniform Resources Identifiers), data with RDF (Resource Description Framework), schema with OWL (Web Ontology Language), and queries with SPARQL (SPARQL Protocol and RDF Query Language).

The application of the Linked Data in BIM means that partial models are published by their producers in the Web as RDF datasets. The GUID-identified entities in IFC models are identified through URIs. The properties and relations of entities are represented as RDF statements. The conversion to RDF can be only partial: for instance, it can be useless to convert the geometrical information in a model — usually its largest part — since for efficient geometrical manipulation it needs to be converted to a more suitable geometrical representation, which can be done directly from IFC.

In Fig. 1 there are a set of partial models at the outer skirt, covering both models based on IFC (on the left) and external information sources not represented in IFC (on the right). The important objects from all these models are exposed in Web by giving them a URI. In addition, the essential internal structures of models are exported as RDF statements and essential data is provided as literals in RDF. The linking between these models is provided by linksets in RDF.

The conversion of IFC files into RDF datasets is based on IFC ontology which is produced by converting the IFC schema into OWL (Beetz 2009). Both of these conversions are relatively straightforward as the OWL and RDF are mostly more flexible and expressive than EXPRESS (the IFC schema language) and STEP Part21 format (the IFC data representation format). There are some minor issues with the conversion — e.g., the procedural constraints in EXPRESS or global property constraints in RDFS — but they have little practical impact.

A publicly accessible converter for IFC data is available at the Ghent University (Pauwels, De Meyer, and Van Campenhout 2011), and an efficient IFC-to-RDF converter has also been implemented in the DRUM project.

While the partial models need to be converted into RDF datasets, the links can be provided directly as linksets in RDF format, since there are really no other existing representations for them. There can potentially be many independently created linksets between two models. The question of how to make the users of the datasets aware of these linksets. This belong to the area of information scope management discussed more below.

Once an IFC model is converted into an RDF dataset, it can be stored in an RDF store. The contents of the model can be queried with SPARQL queries or browsed with a Web browser.

Research Problems

While Linked Data can provide a framework in which to represent the interrelations of models, there are many open questions about how to apply the available representations and technologies. At least the following research problems can be identified:

Link Type Modeling *What kinds of relations can there be between models?* The first thing to observe is that there can be links at two different levels: between the *models* on the whole and between *entities* within the models. The model-level relations are often defined in advance, already when the models to be created in a project are specified. Once the model-level links are known, they provide the basis for looking for specific types of entity-level links across the models. The model-level links can be classified into:

- *Sequential relations*: A model *is based on* another, such as structural design is based on architectural model.
- *Parallel relations*: Two models *compete* of a shared resource (like common space) or *complement* each other (like designs of two wings of a building).

The types of entity-level links are different between different kinds of entities: between a requirements and a building object, two building objects, or a building objects and an activity. Definition of relevant link types also depends on the cross-model functionalities to be supported. In DRUM a working categorization of links has been developed. There is a need for thourough experimentation and refinement to come up with an *ontology of BIM links*.

Link Generation *How to generate the actual links at the entity level?* This is a broad problem that will quite probably require a combination of multiple automatic, manual, and semi-automatic methods. Obviously the role of manual methods has to be limited: it is an error prone task and an additional burden to designers.

There are some clear categories of links that can be identified automatically, most notably those based on geometric information of building objects. These kinds of links can potentially be generated by any tool that can import two or more models at the same time, check spatial clashes, and provide detailed information about clashes. We are currently studying the application of Solibri Model Checker and Tekla Structures to this problem.

There are link discovery frameworks for detecting which nodes in two RDF datasets represent same real-world entities. Examples are Silk (Volz et al. 2009) and LIMES (Ngomo and Auer 2011).

Change Discovery *How to discover changes between different versions of a same model?* RDF is based on graph data model which means that two graphs need to be compared to identify added, removed, and modified nodes. The task is complicated by two problems: nodes with *no identity*, and with *changing identity*.

The anonymous nodes - referred in RDF as blank nodes - cannot be directly mapped to corresponding nodes across model versions as they only have an identity within one version. In RDF graphs converted from IFC files, approximately 80-95% of nodes are blanks. This complicates the application of existing methods for RDF change detection such as RDFSync (Tummarello et al. 2007). RDFSync is based on the computation of *Minimal Self-contained Graphs (MSGs)* of the model, and on the efficient detection of changes using checksums for MSGs. Unfortunately, the large ratio of blanks results in very large MSGs that connect most of the nodes in the model. An area of active research in DRUM is to come up with efficient strategies to give at least *partially stable identities* for blank nodes.

Changing identities have different reasons. Firstly, not all BIM tools can maintain the identities of all entities across model versions, since some exported entities have no corresponding objects in their native models. Secondly, it is typical practice for a designer who wants to change an existing entity in a model to first delete the old entity and then design a new entity in its place. This is common when more detail is added into existing entities, such as reinforcements to concrete. For economy of work, a detailed design can then be copied and pasted into the place of many other similar entities. However, maintaining the identities of nodes in essential to prevent the cross-model links to break unnecessarily. Methods to overcome this problem have been studied in DSNotify system (Popitsch and Haslhofer 2010), which attempts to match recently removed entities with newly created ones.

Change Management *How to coordinate changes that affect multiple models?* Assuming that IFC models - such as an architectural model and a structural model - have been converted into RDF datasets and that cross-model links have been defined between entities that represent same real-world entities, there are many levels at which change management can be supported:

1. *Preventive*: Awareness of the connection can be provided to prevent superfluous changes.
2. *Reactive*: Change notifications to the affected parties can be automatically generated.
3. *Proactive*: Change requests to potentially affected parties can be created when a change is needed.
4. *Transactional*: Using distributed versioning the parties can commit their tentative model versions when designs have reach a consistent state.

When the architect is planning a change in her model - for instance, the position of a wall - the tool can provide her with an awareness of the structures that have been designed inside the wall. Once she makes a change, a notification to dependent models can be sent. Or if she desires, she can send a change request while doing her own changes. Finally, she and the other designers can work on the change collaboratively. They can create new tentative versions of their models to create new designs and once everyone has finished the work, commit the changes together.

Information Scope Management *How to represent and manage information about what parties, datasets, and linksets belong to a project? How to implement that in a distributed manner?* At the model-level this is a question of model metadata that can be represented using — and perhaps extending — dataset description ontologies such as VoID (Alexander and Hausenblas 2009). At the entity-level the problem can be addressed with protocols that support registration of links between models. Each link is a new linkset can registered with the datasets whose entities it refers to. When the dataset is afterwards accessed, the registered links will be available for traversal.

Some of the above-mentioned problems — especially the problem of changing identities — should eventually be solved in BIM tools but before that happens, their effects are encountered in the information management solutions. If the change in the identity of an entity cannot be detected, the links associated with the entity will be broken. A large number of broken links can severely decrease the usefulness of linksets. In addition, the change cannot be discovered in a proper manner if identities cannot be equated. For now, to study Linked Building Models, these problems need to be solved outside BIM tools.

Summary

DRUM studies the application of Linked Data technologies to management of building information. Linked Data provides a framework that focuses on the central question in exchange, combination, and modification of co-evolving partial models of a building: the interlinking of the models. However, the previous applications of Linked Data are quite different from building information management, and many questions of how to use the technologies need to be solved.

Linked Building Data is an emerging area of research. Previously the IFC data model has been translated to an OWL ontology. There are translators that can convert IFC models into RDF datasets, one implemented in DRUM project. There are initial working categorization of BIM link types but a more principled ontology needs to be worked out. The development of efficient methods for change discovery, models versioning, and link generation are currently under work. Information scope management and collaborative change management methods are future research topics in DRUM.

References

Alexander, K., and Hausenblas, M. 2009. Describing linked datasets-on the design and usage of void, the'vocabulary of interlinked datasets. In *In Linked Data on the Web Workshop (LDOW 09), in conjunction with 18th International World Wide Web Conference (WWW 09*. Citeseer.

Beetz, J. 2009. *Facilitating distributed collaboration in the AEC/FM sector using Semantic Web Technologies*. Ph.D. Dissertation, PhD Thesis, Eindhoven University of Technology. ISBN 978-90-6814-618-9.

Berners-Lee, T. 2006. Linked data-design issues. *W3C* 2009(09/20).

Bizer, C.; Heath, T.; and Berners-Lee, T. 2009. Linked data-the story so far. *International Journal on Semantic Web and Information Systems (IJSWIS)* 5(3):1–22.

Heath, T., and Bizer, C. 2011. Linked data: Evolving the web into a global data space. *Synthesis Lectures on the Semantic Web: Theory and Technology* 1(1):1–136.

Ngomo, A., and Auer, S. 2011. Limes-a time-efficient approach for large-scale link discovery on the web of data. In *Proceedings of IJCAI*.

Pauwels, P.; De Meyer, R.; and Van Campenhout, J. 2011. Interoperability for the design and construction industry through semantic web technology. *Semantic Multimedia* 143–158.

Popitsch, N., and Haslhofer, B. 2010. Dsnotify: handling broken links in the web of data. In *Proceedings of the 19th international conference on World wide web*, 761–770. ACM.

Tummarello, G.; Morbidoni, C.; Bachmann-Gmür, R.; and Erling, O. 2007. Rdfsync: efficient remote synchronization of rdf models. In *Proceedings of the 6th international The semantic web and 2nd Asian conference on Asian semantic web conference*, 537–551. Springer-Verlag.

Volz, J.; Bizer, C.; Gaedke, M.; and Kobilarov, G. 2009. Silk–a link discovery framework for the web of data. In *Proceedings of the 2nd Linked Data on the Web Workshop*. Citeseer.

Discussion and Workshop Conclusions: Supporting Decision-Making in the Building Lifecycle using Linked Data

Pieter Pauwels
Department of Architecture and Urban Planning
Ghent University, Ghent, Belgium
pipauwel.pauwels@ugent.be

Davy Van Deursen
Multimedia Lab
Ghent University - IBBT, Ghent, Belgium
davy.vandeursen@ugent.be

Leandro Madrazo and **Gonçal Costa**
Department of Architecture
Enginyeria i Arquitectura La Salle, Barcelona, Spain
madrazo@salle.url.edu, gcosta@salle.url.edu

Leif Granholm
Tekla Finland, Finland

Edward Corry
IRUSE
NUI Galway, Galway, Ireland
edward.corry@nuigalway.ie

Edward Curry
Digital Enterprise Research Institute
NUI Galway, Galway, Ireland
ed.curry@deri.org

James O'Donnell
Building Technologies Group
LBNL, Berkeley, CA, USA
jtodonnell@lbl.gov

Seppo Törmä and **Jyrki Oraskari** and **Nam Vu Hoang**
Department of Computer Science and Engineering
Aalto University, School of Science
seppo.torma@aalto.fi, jyrki.oraskari@aalto.fi, nam.vuhoang@aalto.fi

Abstract

Diverse approaches were presented during the workshop for addressing the interoperability challenge in the domain of architecture, engineering and construction (AEC). Additionally, multiple contexts were presented in which linked data and semantic web technologies could generate an important added value to experts in the AEC domain. However, also in these contexts, the eventually targeted functionality significantly relies on the extent to which the interoperability challenge is really addressed and information becomes shareable among information systems in the design and construction process. In this concluding article, we outline our conclusions regarding the extent to which semantic web technologies can address the interoperability challenge, and how information management in such a context might be realised in practice. Finally, we give an initial outline of some anticipated use cases in the building lifecycle in which the usage of semantic web technologies may generate considerable advantages over existing technologies and methods.

Introduction

Projects in the AEC domain typically involve diverse parties, each bringing specific information into these projects. Client information needs to be combined with the information of the architectural design firm; electrical engineering information needs to be combined with facility management information; plumbing information needs to be combined with sensor information; and so forth. Also after the construction phase, building information needs to be accessible for a range of diverse users, including the facility director, in-house machinery and systems, renovation specialists, technicians, and so forth. Because users involved in AEC projects typically rely on specialized information systems, a large number of information models is typically available for one and the same building. This naturally leads to the following questions:

- How can all this information be combined so that it is comfortably accessible to the diverse parties involved?
- Can the experience of the diverse users be improved by a combination of various information resources?

The first question hereby refers to the well-known interoperability issue in the AEC domain. The second question relates to human-computer interaction (HCI). In the AEC context, this is referred to as a 'functionality mismatch issue' in (Pauwels 2012). In (Curry, O'Donnell, and Corry 2012), this is described as the inability of current methods and tools *"to account for the profile of building managers, both in terms of the operational context of their role, and their typical technical and educational background (O'Donnell 2009)"*.

Building information modelling (BIM) environments (Eastman et al. 2011) appeared to bring about improved facilities for information management in AEC projects. Although a lot of improvements have been generated by the usage of such BIM environments, they are not entirely successful in addressing the above questions. Difficulties persist regarding information interoperability, also when relying on the Industry Foundation Classes (IFC) (Liebich et al. 2012) as a standard for information exchange. Also the relation between end user and information system did not improve, because BIM environments typically prove not to be flexible enough to house the specific kinds of information of the diverse parties in an AEC project outlined above. In the end, also a BIM environment thus provides only one silo of information to the end user, with the contained information often not being customized or tailored to the needs and requirements of the end user.

Semantic web technologies, as they were suggested in (Berners-Lee, Hendler, and Lassila 2001), might provide better answers to the above questions. Namely, these technologies currently lie at the basis of a global Linked Open Data (LOD) cloud (Bizer, Heath, and Berners-Lee 2009, Bizer, Jentzsch, and Cyganiak 2011). As such, they might well allow to effectively connect the diverse information models in AEC projects. With this global source of cross-domain information, also the end user experience might eventually be improved, because applications can theoretically rely on a larger and more diverse information source (see also (Cyganiak and Jentzsch 2011)).

In this concluding article, an overview is given of the workshop conclusions regarding the focus points of this workshop.

1. investigating the possibilities of using linked data in AEC projects
2. distinguishing difficulties or barriers in answering the above questions with linked data technologies
3. outlining future research directions to facilitate an appropriate use of linked data in AEC projects

We first look into the ways in which semantic web technologies can help in integrating information models in the AEC domain. Then we look into the ways in which the connecting links between information models can be created and managed. Finally, a short overview is given of anticipated use cases for deploying and benefiting from linked data in the building lifecycle.

Information Integration in Architecture and Construction

Semantic web technologies lie at the basis of the LOD cloud. These technologies are designed for the representation, publication and usage of structured data on the World Wide Web (WWW). Information resources are in this context represented with the Resource Description Framework (RDF) and identified with Unique Resource Identifiers (URIs) (Manola and Miller 2004). The Web Ontology Language (OWL) enables the representation of ontologies or vocabularies that can be used for structuring RDF graphs (McGuinness and van Harmelen 2009). The Simple Protocol and RDF Query Language (SPARQL) allows querying the RDF graphs (Prud'hommeaux and Seaborne 2008).

With these possibilities, semantic web technologies and the resulting LOD cloud can be considered a useful set of technologies for addressing the initial research questions of this workshop. They apparently promise to connect various information resources on a global scale and make the result easily available to various services and application types (Pauwels and Van Deursen 2012, Madrazo and Costa 2012).

Integration within the Construction Project

The approaches presented and discussed in this workshop typically start from the IFC schema. This is a neutral and standard schema for information exchange among BIM environments. Using the IFC schema, one should be able to represent a BIM model so that other applications are able to use this information as well, for instance, for simulation and visualisation purposes. Although researchers in the AEC domain are pointing more and more towards the limitations of IFC regarding information exchange, this information schema is currently one of the best approaches currently available and used to address interoperability issues in the AEC sector.

It was shown in this workshop how an IFC-to-RDF conversion service can be implemented that converts IFC information into an RDF representation (Pauwels and Van Deursen 2012). This is not a straightforward process, because diverse mapping schemas are typically available for mapping between an IFC file and an RDF graph, or between an IFC schema in EXPRESS and an IFC ontology in OWL. Especially the more advanced features of the EXPRESS schema of IFC, such as rule functionality and cardinality restrictions, can be translated into diverse RDF constructs. A comparison of the three conversion procedures used in (Pauwels and Van Deursen 2012), (Beetz, Van Leeuwen, and De Vries 2009) and (Törmä, Oraskari, and Huang 2012) illustrates this situation. One might thus conclude that a 'perfect conversion procedure' does not exist for converting IFC information into RDF representations. Rather, there exist various 'flavours' of conversion procedures, each resulting in a specific kind of RDF representation. One conversion procedure might result in a simple, compact and straight-forward RDF graph, a second procedure might result in a complete but impractical RDF graph, and yet another procedure might provide an RDF graph fit for specific reasoning purposes, for instance.

Assuming that similar RDF conversion services can be implemented for other (neutral or proprietary) schemas typically deployed in an AEC context, one can easily imagine diverse information models available as RDF graphs within the same building project (Fig. 1). By relying on linked data principles and techniques, these information models, which can be considered separate 'silos of information' (Curry, O'Donnell, and Corry 2012), might well be linked together, resulting in a linked data cloud for each AEC project. Direct links are thus available among simulation information, CAD information, architectural information, visualisation information, and so forth. A distinction can be made between links among models (sequential or parallel) and links among model entities (Törmä, Oraskari, and Huang 2012).

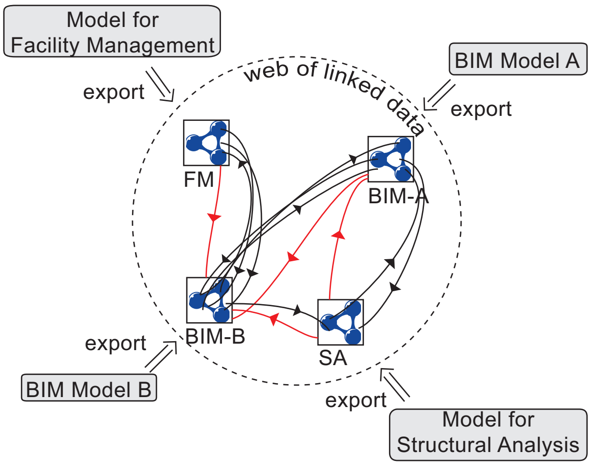


Figure 1: Diverse conversion services might enable to make diverse partial information models available in RDF graphs. These graphs might then be linked using semantic web technologies.

When considering the future prospect shown in Fig. 1, however, one has to seriously keep in mind our earlier remark regarding the diverse flavours of conversion procedures. There are diverse conversion routines possible in each step from an AEC application towards the linked data cloud for an AEC project. One can easily understand the resulting information management difficulties by considering:

- the number of schemas available in the AEC domain,
- the number of conversion routines between schemas and OWL ontologies, and between information models and RDF graphs,
- the number of linking possibilities between two RDF models.

Integration outside the Construction Project

Semantic web technologies additionally allow to link the linked data clouds of AEC projects (Fig. 1) to information outside the construction project (Fig. 2). As such, external information may be deployed for specific purposes in an AEC project. This includes annotations, documents, project management tools, geographical information, demographic information, and so forth. With this information, more advanced services and applications may be targeted, in which diverse resources of information are combined (Curry, O'Donnell, and Corry 2012). For example, an outline of cost efficiency statistics related to usage statistics of a building might bring about significant new insights to the building owner.

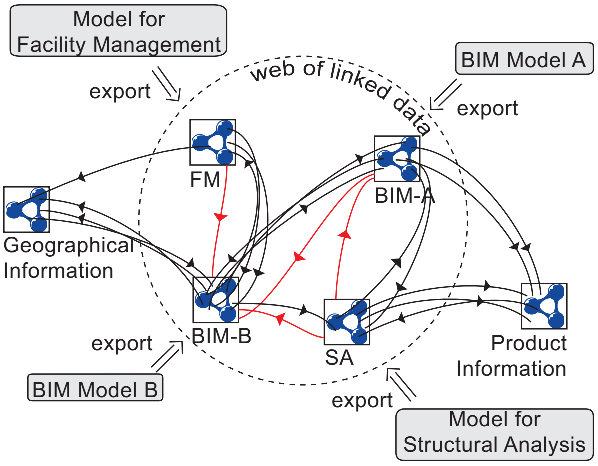


Figure 2: The linked data cloud for the AEC project can be further enriched with additional links to external resources of information.

Note that, also in this context, the same difficulties need to be taken into account regarding the management of mapping and linking procedures among information models. This can be related to the difficulties outlined in the semantic web domain regarding the usage of the `owl:sameAs` construct (Halpin et al. 2010). Although it might be valid in one context to link entities in different information models or application domains, these links might not be equally obvious or valid in other contexts.

How to create and manage the links

In this section, we handle in some more detail the challenges related to creating and managing the links between diverse information models in RDF. This question has been dealt with before in the AEC domain, although not relying on semantic web technologies. For example, the usage was suggested of 'view models' that are integrated or that communicate through a 'model kernel', which is formed by the overlapping of the view models (Van Nederveen 1993). Alternatively, the definition was proposed of 'views' as 'functional contexts' for the diverse partners or disciplines in an AEC project (Rosenman and Gero 1996). These views can then be linked by the addition of explicit relations. An implementation with relational database technologies was furthermore suggested (Rosenman and Gero 1996).

These and other approaches have thus been suggested for dealing with the creation and integration of partial models. Initiatives that tend to fail, are initiatives in which the original information is converted or translated into an alternative information schema, possibly combined with discarding the original information. In these cases, a valuable amount of information is lost. Approaches that appear to have higher chances to succeed, are approaches that enable users not only to create partial models, but also to *maintain* these partial models as such. In these cases, the partial model is provided to interested third parties, but the original information is kept intact. Crucial in this approach, is to maintain the link or the mapping schema between the original partial model and the related partial model, whether this model be a follow-up model (sequential) or an alternative model (parallel).

In the following subsections, we look into the diverse considerations that have to be made in realizing the latter approach for AEC information in a linked data context with semantic web technologies. We make a distinction between technical considerations, practical considerations, usability considerations and maintenance considerations.

Usability considerations

Information in an AEC project is always represented by a specific partner in this AEC project. Not only is this partner supposed to be qualified for representing this information, this partner is typically also considered responsible and representative for this information. This is important metadata that should be taken into account when giving access to the information that is represented by this partner, not only for reasons of rights and ownership, but also for reasons of representability, trustworthiness, and usability. To what extent is the represented information *correct* and *trustworthy* and to what extent can it consequently be *used*? This consideration relates to the issue of information scope management that is outlined in (Törmä, Oraskari, and Huang 2012).

The system that we propose here combines partial models that are maintained and used. Each of these partial models has its creator, who is considered representative for the information in the partial model. Reference to this creator is not only available in the URIs used for representing the entities in the partial model, but should also be available in metadata that is associated to the partial model. Both the URIs and the metadata allow other end users to check the trustworthiness and usability of the information. In this approach, anyone can provide information, and certain information becomes more or less trustworthy depending on the number of similar assertions and the status or expertise of the partner that has provided the information.

Practical considerations

As became clear of the diverse contributions to this workshop, one has to deal with the presence of diverse partial models. Each of these partial models is typically *produced by one specific party*, for instance one of the many architectural design firms, and *with one specific design tool*, for instance one of the many CAD modelling applications. These partial models implicitly include the original schema and understanding of the representation, both in terms of the design situation (specific people and context) and the used design tool (specific information structure).

When developing linked data environments for accommodating such partial models, this context needs to be taken into account, because it has a considerable impact on the system. For instance, information within one and the same partial model reflects the information from one specific subdomain in the AEC project and is thus typically more coherent and more closely related. Links among entities within this partial model can be made with less effort. Links among diverse partial models, for instance a structural model and a client requirements model for the building, are harder to realise. The system thus, for instance, needs to take into account that considerably more links are available within one and the same partial model, and less are available among diverse partial models.

Additionally, within this context of partial models with all their own information structures and their own contexts, it makes sense to allow building all partial models within their specific modelling environments (modelling applications, simulation applications, visualization applications, and so forth), and link them together only at a read-only level (see also (Törmä, Oraskari, and Huang 2012) and Fig. 1). As a result, only one conversion service is needed instead of a roundtrip through two conversion services. If one would want to integrate all information once and for all into one complete all-containing model, which we do not suggest here, proper conversion roundtrips would be necessary.

When linking diverse partial models into a linked data cloud for an AEC project, the following link types can be considered:

- links between partial models
  1. sequential links
  2. parallel links
- links between entities of a partial model
  1. links among objects
  2. links between objects and requirements
  3. links between objects and activities

Technical considerations

Considerable technical considerations were outlined in this workshop regarding the generation of links among and within the diverse information models (Törmä, Oraskari, and Huang 2012). It seems infeasible to rely on either automatic or manual methods. A semi-automatic method thus seems most promising. In such a method, an initial set of links is generated among (entities of) the considered information models, after which the generated links are returned to the end user for further modifications. The usage of clash detection and link discovery software could be considered as aids in the link generation process (Törmä, Oraskari, and Huang 2012).

One needs to take into account that, after generating links among and within models, these links should be easily maintained and managed. How this maintenance and management of links can be realised, is handled in the following subsection, which briefly deals with change discovery and change propagation. It should be enough to note here that a practical and realistic change discovery and change propagation relies on the availability of persistent and unique identifiers for the many available entities represented in the RDF graphs (see also (Törmä, Oraskari, and Huang 2012)). In this regard, we initially suggest using the following procedure for generating links among and within the information models.

1. Identify the IDs:
   Upon conversion from the initial software environment into an RDF graph, the diverse IDs that are used in the original software environment are retrieved, so that the diverse entities and concepts in the RDF graph can be given unique URIs that relate to the IDs in the original software environment. By doing so, future changes to the partial model can be propagated into the linked building data graph.
2. Link the IDs among and within models:
   When the URIs of the entities in the diverse partial models are available, (entities within) the partial models can be linked in a semi-automatic manner using the outlined link types (sequential-parallel or object-object / object-requirement / object-activity). When one of the linked partial models is modified, a reasonable decision should be made by the partners who are in charge of the partial models about whether or not to maintain or modify the specific links between the entities or the modified models.
3. Add more information:
   Further information can be added to the diverse partial models, with the information coming from various domains of practice, also outside the AEC domain. The existing LOD cloud (Cyganiak and Jentzsch 2011) provides an important available resource to make such links from and to.
4. Provide an interface to access links:
   The information models, the entities in the information models and the links among both should finally be made available to the end users who have the appropriate access rights. Using the metadata that is added for information scope management, an interface with the appropriate levels of security and rights administration can be realised.

In terms of the proposed LOD system for this approach, the above procedure could be realised as follows:

1. Identify the IDs:
   The GUIDs that are being used by the diverse software applications producing partial models are converted into appropriate URIs. These URIs could additionally take into account information about corresponding owner, project and partial model. This could thus result in the following URI design for entities in the partial information models: `http://owner.country/project/-partialmodel/guid`.
2. Link the IDs among and within models:
   The (entities in the) partial information models are available as (entities in) RDF graphs, and can thus be linked by additional RDF links using semi-automatic methods.
3. Add more information:
   Extra links to information in external RDF graphs can similarly be added with additional RDF links.
4. Provide an interface to access links:
   The complete graph is published in an online RDF store with an appropriately accessible SPARQL endpoint. Using the metadata that is added to the partial models, an appropriate user interface can be implemented on top of this SPARQL endpoint, giving partners in the AEC project access to the information for which they have access rights.

Maintenance and management considerations

A realistic maintenance and management of the generated LOD cloud for the AEC project requires important considerations in terms of change discovery and change propagation. We suggested earlier to initially rely on an approach in which existing software (e.g. BIM software, simulation software, and so forth) is used by partners in an AEC project to build partial models, after which these partial models are exported into RDF graphs (see also Fig. 1). The conversion of GUIDs into URIs, which is central in the presented procedure for generating links among models and entities in those models, should allow to appropriately update/replace the available RDF graphs with the newly exported partial models.

The most important issue then becomes the maintenance and management of the links that were previously made among (entities in) the partial models that are being replaced. Whether these links be stored internal or external to the models they belong to, one does not want to end up with hanging, missing or wrong links. Note that a change in one of the partial models can propagate all the way up to the final construction plan, so this is a crucial part of realising a linked data system that gives support for AEC projects.

The three following strategies were outlined during the workshop regarding link change management:

- Reactive change propagation across models:
  The other parties are notified about a change so that they can restore the consistency.
- Proactive change protocols:
  Collaborative protocols are used that enable taking into account the views of different parties affected by a change. There are different possible protocols based on change proposals, counterproposals, and so on.
- Transactional change management protocols:
  Protocols are used that take the advantage of the distributed versioning capabilities of the participating models.

Anticipated Use Cases in the Building Lifecycle Context

In the workshop, finally, diverse use cases were anticipated in the building lifecycle context in which the usage of a linked data approach, as briefly sketched above, might provide additional benefits to the diverse partners in the AEC project. These use cases focus on building optimisation, information management and support for the design and construction process. Central in these use cases is the idea that a linked data approach has the possibility to enable a more holistic view on information about the building, as well in the design and construction phase as in the maintenance phase.

Design and Construction Phase

The resulting improvements to information management are a key reason for adopting a linked data approach in the AEC domain. Consequently, main use cases in which improvements can be expected are situated in the design and construction phases of AEC projects. In this context, the most important improvement is expected to be generated by the change management features of the suggested approach. Namely, assuming that appropriate links can be made among and within diverse partial models stemming from diverse partners in the AEC project, and assuming that the appropriate metadata is added in this process, the system can presumably give better indications of how changes in a certain partial model affect the linked partial models in the AEC project. These indications can be used by that specific partner in the AEC project to make better informed decisions. By using a linked data approach, more partial models can be reached in this way than is traditionally the case, because of the ease of linking diverse partial models in an RDF graph. A more holistic view of the AEC project is thus obtained than is traditionally the case.

A concrete use case that was anticipated in the workshop, focuses on *energy performance and usage evaluation* in the design and construction phase. In this case, three partial models are combined into one RDF graph (Fig. 3). The first partial model represents the building using the terms of a specifically devised space topology vocabulary. We could in this case rely on the room vocabulary ontology that was devised in DERI and that is available at `http://vocab.deri.ie/rooms`. Alternatively, we could deploy the topology ontology that was built as part of the AIM framework at Ghent University.

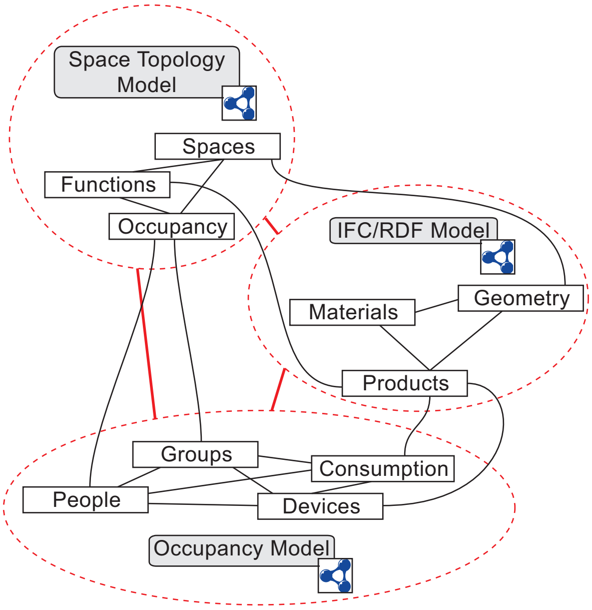


Figure 3: A combination of three partial models for a use case in performance evaluation in the design and construction phase.

The second partial model, that should be closely linked to the first partial model, represents the building using the terms available in the IFC ontology, including geometric properties of the building and, to some extent, product information. It would presumably be a good test for the change propagation and change discovery features of the system to see to what extent changes in the IFC/RDF model of the building can be propagated into the space topology model (Fig. 3).

A third partial model finally represents people, groups, devices and energy consumption using the terms available in the FOAF ontology and the DERI Energy ontology. By linking this third partial model to the two other partial models, one can test to what extent the links between these models can inform the designer or construction firm about the energy performance effects that are inferred by certain changes in the space topology, the product choices or the building geometry, for instance. Alternatively, it might be possible to make strategic choices in the occupancy model and see if and how one should change the room topology or building geometry to accommodate the desired performance level.

Maintenance Phase

A second use case focuses on the maintenance phase of the building. The use case that is anticipated in the workshop literally extends the first use case, in the sense that additional partial models are added upon completion of the building that take into account sensor information, operations and maintenance manuals, financial information, weather data, and so forth (Fig. 4). By making links among these partial models, one can perform very specific queries over the merged graph, thereby enabling a better informed or more holistic view on the overall performance level and the usage of the building. Such a use case was already started in the context of an exploratory test of the possibilities of a linked data approach. This was in part presented in (Curry, O'Donnell, and Corry 2012). The use case suggested in this workshop extends this initial test.

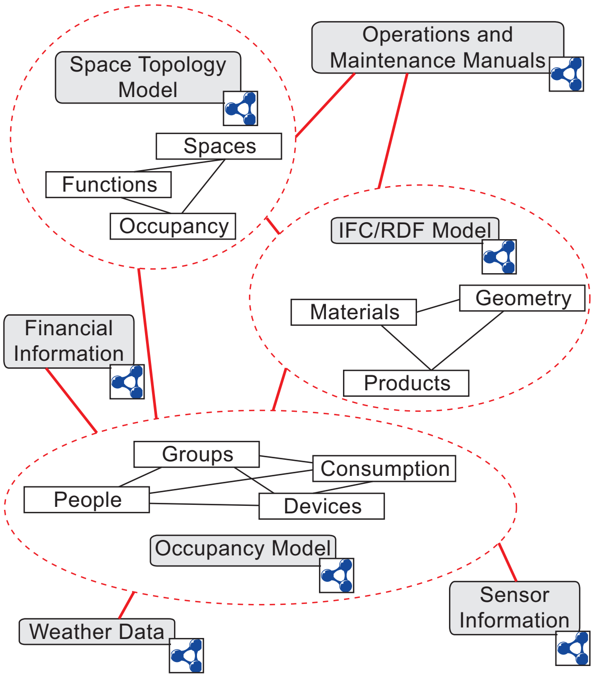


Figure 4: Extending the linked data cloud for the AEC project with links to sensor information, operations and maintenance manuals, financial information, weather data should allow improving building performance in the maintenance phase of the building lifecycle.

Conclusion

It was concluded in this workshop that significant improvements can be made regarding the management and usage of information in AEC projects by relying on a linked data approach. The main anticipated improvement is situated in the context of interoperability of information in the AEC domain. The second anticipated improvement is situated in the context of human-computer interaction in the AEC domain. Because of the sheer amount of information available in the Linked Open Data cloud, a more holistic view on the information of the building could be made available to the end users, thereby enabling them to make better informed decisions.

In this article, an indication is given of how information in AEC projects can be integrated using a linked data approach. Additionally, significant considerations are outlined regarding the creation and management of the links within and among the diverse (partial) information models. Finally, this article has given an initial outline of possible use cases in the design and construction phase of an AEC project and in the maintenance phase of an AEC project.

Acknowledgments

The work presented in this article could not be presented without funding support from the Interdisciplinary Institute for BroadBand Technologies (IBBT), who made this workshop possible.

References

Beetz, J.; Van Leeuwen, J.; and De Vries, B. 2009. ifcowl: A case of transforming express schemas into ontologies. *Artificial Intelligence for Engineering Design, Analysis and Manufacturing* 23:89–101.

Berners-Lee, T.; Hendler, J.; and Lassila, O. 2001. The semantic web. *Scientific American* 284(5):35–43.

Bizer, C.; Heath, T.; and Berners-Lee, T. 2009. Linked data - the story so far. *International Journal on Semantic Web and Information Systems* 5(3):1–22.

Bizer, C.; Jentzsch, A.; and Cyganiak, R. 2011. State of the LOD cloud. Available online: http://www4.wiwiss.fu-berlin.de/lodcloud/state/. Last accessed on 22 February 2012.

Curry, E.; O'Donnell, J.; and Corry, E. 2012. Cross-domain building optimisation using scenario modeling and linked data. In *International Workshop on Linked Data in Architecture and Construction*, 7–9.

Cyganiak, R., and Jentzsch, A. 2011. The linking open data cloud diagram. Available online: http://lod-cloud.net/. Last accessed on 22 February 2012.

Eastman, C.; Teicholz, P.; Sacks, R.; and Liston, K. 2011. *BIM Handbook: A Guide to Building Information Modeling for Owners, Managers, Designers, Engineers and Contractors*. Wiley.

Halpin, H.; Hayes, P.; McCusker, J.; McGuinness, D.; and Thompson, H. 2010. When owl:sameAs isn't the same: an analysis of identity in linked data. *Lecture Notes in Computer Science: ESWC Part I* 6496:305–320.

Liebich, T.; Adachi, Y.; Forester, J.; Hyvarinen, J.; Karstila, K.; Reed, K.; Richter, S.; and Wix, J. 2012. Industry Foundation Classes IFC2x edition 3 technical corrigendum 1. Available online: http://www.buildingsmart-tech.org/ifc/IFC2x3/TC1/html/index.htm. Last accessed on 22 February 2012.

Madrazo, L., and Costa, G. 2012. Open product modelling and interoperability in the aec sector. In *International Workshop on Linked Data in Architecture and Construction*, 5–6.

Manola, F., and Miller, E. 2004. RDF Primer. W3C Recommendation. Available online: http://www.w3.org/TR/rdf-primer/. Last accessed on 22 February 2012.

McGuinness, D., and van Harmelen, F. 2009. OWL 2 Web Ontology Language document overview. W3C recommendation. Available online: http://www.w3.org/TR/owl2-overview/. Last accessed on 22 February 2012.

O'Donnell, J. 2009. *Specification of Optimum Holistic Building Environmental and Energy Performance Information to Support Informed Decision Making*. Ph.d, University College Cork.

Pauwels, P., and Van Deursen, D. 2012. Ifc/rdf: Adaptation, aggregation and enrichment. In *International Workshop on Linked Data in Architecture and Construction*, 2–4.

Pauwels, P. 2012. *Reconsidering Information System Support for Architectural Design Thinking*. Ph.D. Dissertation, Department of Architecture and Urban Planning, Ghent University, Ghent, Belgium.

Prud'hommeaux, E., and Seaborne, A. 2008. SPARQL query language for RDF. W3C recommendation. Available online: http://www.w3.org/TR/rdf-sparql-query/. Last accessed on 22 February 2012.

Rosenman, M., and Gero, J. 1996. Modelling multiple views of design objects in a collaborative cad environment. *Computer-Aided Design* 28:193–205.

Törmä, S.; Oraskari, J.; and Huang, N. 2012. Distributed transactional building information management. In *International Workshop on Linked Data in Architecture and Construction*, 10–13.

Van Nederveen, S. 1993. View integration in building design. In *Management of Information Technology for Construction*, 209–221. Singapore: World Scientific Publishing Company.