<table>
<thead>
<tr>
<th><strong>Title</strong></th>
<th>Making SimModel information available as RDF graphs</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Authors(s)</strong></td>
<td>Pauwels, Pieter, Corry, Edward, O'Donnell, James</td>
</tr>
<tr>
<td><strong>Publication date</strong></td>
<td>2014-09-19</td>
</tr>
<tr>
<td><strong>Conference details</strong></td>
<td>ECPPM 2014: 10TH European Conference on Product and Process Modelling, Vienna, Austria, 17-19 September 2014</td>
</tr>
<tr>
<td><strong>Publisher</strong></td>
<td>Taylor &amp; Francis</td>
</tr>
<tr>
<td><strong>Item record/more information</strong></td>
<td><a href="http://hdl.handle.net/10197/11024">http://hdl.handle.net/10197/11024</a></td>
</tr>
</tbody>
</table>

The UCD community has made this article openly available. Please share how this access benefits you. Your story matters! (@ucd_oa)
Making SimModel information available as RDF graphs

P. Pauwels
Ghent University, Ghent, Belgium

E. Corry
National University Ireland Galway, Galway, Ireland

J. O’Donnell
School of Mechanical and Materials Engineering, Electricity Research Centre (ERC) and UCD Energy Institute, University College Dublin, Dublin 4, Ireland

ABSTRACT: Many building energy performance Simulation (BEPS) on tools use custom schema definitions as opposed to standardised schema definitions (defined in XSD, EXPRESS, and so forth). A Simulation Domain Model (SimModel) was therefore previously developed and is representative of a new interoperable XML-based data model for the building simulation domain. In this paper we document our ongoing efforts to make building simulation data more interoperable with other building data. In order to better integrate SimModel information with other building information, the authors propose to represent this information in the Resource Description Framework (RDF). A specialised conversion service parses the SimModel ontology, which is in the form of linked XSD schemata, and output a SimModel ontology in OWL. In this article, we further outline how the conversion service now also parses input SimModel XML files and outputs SimModel instances as RDF graphs. We briefly indicated how those SimModel RDF graphs can subsequently be used.

1 INTRODUCTION

1.1 Interoperable building energy performance data

Projects in the domain of architecture, engineering and construction (AEC) typically involve diverse parties, each bringing specific information into these projects. Therefore, a well-functioning information flow throughout the complete building life-cycle is crucial. This is also the case for the information flow within one of the subdomains of the AEC domain, namely the domain of building energy performance Simulation (BEPS) (Bazjanac et al. 2011a). In this article, we will therefore look into one of the diverse options to make information for BEPS increasingly interoperable.

Many BEPS tools, such as EnergyPlus and DOE-2, use custom schema definitions (IDD and BDL respectively) as opposed to standardised schema definitions (defined in XSD, EXPRESS, and so forth). A Simulation Domain Model (SimModel) was therefore proposed earlier (O’Donnell et al. 2011), representative for a new interoperable XML-based data model for the building energy performance simulation domain. Its ontology aims at moving away from tool-specific, non-standard nomenclature by implementing an industry-validated terminology aligned with the Industry Foundation Classes (IFC – ISO 2013, Liebich et al. 2009). As such, it can also be aligned with many of the existing applications for BEPS.

In this article, we document our ongoing efforts to make BEPS data more interoperable with other building data through SimModel. We document the origin and background of SimModel. Furthermore, we outline how SimModel aims to combine diverse schema definitions for BEPS tools, such as EnergyPlus, DOE-2, IFC, gbXML, and so forth. We then elaborate on the way in which we can extend this initial level of interoperability to other building domain data sources using semantic web technologies (Berners-Lee et al. 2001). In order to be able to better integrate SimModel information with other building information, we have aimed at representing SimModel information in the Resource Description Framework (RDF) data model (Manola & Miller 2004), thereby relying on earlier work on the usage of semantic web technologies in the AEC domain (Pauwels et al. 2010, 2011a). By doing so, the BEPS information in a SimModel can be made available as directed labelled graphs. A conversion service has been built that is able to parse the SimModel ontology in the form of XSD schemata and output a SimModel ontology in the Web Ontology Language (OWL - W3C OWL Working Group 2012). This work was previously documented and presented in Pauwels et al. (2014). In this article, we first briefly document this effort and give an indication of what the resulting SimModel ontology in OWL can be used for.
1.2 Generating and using RDF graphs of SimModels

Furthermore, we show how the SimModel ontology can be used to implement a conversion service that is able to convert SimModel instances, which are available in XML files, into RDF graphs. Using this conversion service, one is able to make the SimModel information of specific buildings available for semantic web technologies. These semantic web technologies, for instance, allow advanced SPARQL queries (W3C SPARQL Working Group 2012) on the BEPS information that is stored in SimModel. They also allow us to combine the SimModel information with custom or specialised inference rules and a standard reasoning engine that is, like all semantic web technologies, based on description logics (DL - Baader & Nutt 2003). An example combination of RDF data with inference rules and a standard reasoning engine was previously discussed in Pauwels et al. (2011b).

Finally, making SimModel information available to semantic web technologies allows building energy performance information to link with information that is available from various other sources. Among these sources is the Linked Open Data (LOD) cloud (Linkeddata.org 2014, Bizer et al. 2009), which combines all kinds of information available worldwide (geographical, general knowledge, material properties, services, and so forth). Furthermore, the BEPS information in the RDF graphs of SimModels can be linked to IFC building models that are converted through the IFC-to-RDF converter service (LBD 2013, Törmä 2013, Pauwels & Van Deursen 2012). Using the IFC and SimModel conversion services, a web of 'linked building data' may thus be targeted, which combines diverse sources of information for particular buildings with the aim of improving information exchange throughout the building life-cycle.

2 THE SIMMODEL: BACKGROUND & CONTEXT

2.1 Building Energy Performance Simulation-BEPS

SimModel is primarily used as an internal data model by the Simergy software developed at LBNL (LBNL 2013, See et al. 2011), the successor of the Mojito platform presented by (Bazjanac et al. 2011a). The Mojito tool was first conceived as a platform that facilitates data flow to and from BEPS tools to and from potentially any building modelling tool (Bazjanac et al. 2011a). Data flow is possible to and from BIM models in IFC, software or tools that use the DOE-2 engine, EnergyPlus, and tools with gbXML export. These tools are typically used for BEP simulations. Data from any of these environments can be mapped to and from the SimModel data model using the Simergy software (LBNL 2013).

SimModel can thus be placed among other AEC domain tools as displayed in Figure 1.

Simergy is a comprehensive graphical user interface (GUI) for the US Department of Energy's (DOE) building energy simulation program EnergyPlus (LBNL 2013, O’Donnell et al. 2013, See et al. 2011). The application provides an intuitive schematic editor for HVAC systems. The editor hands on a space-based building model, which can be imported from the diverse sources shown in Fig. 1 (O’Donnell et al. 2011).

The underlying SimModel is an object-oriented data model which defines all object / attribute / relationship sets used for BEPS. “The primary objective of SimModel is to accommodate the existing input data requirement of EnergyPlus, while allowing mapping from/to other domain data models and easy incorporation of new definitions.” (Bazjanac et al. 2011b). At the core, SimModel is represented using the XML markup language (O’Donnell et al. 2011). This representation is closely aligned to the IFC data model, so that it can be linked to incoming or outgoing IFC information (Fig. 1).

2.2 The SimModel ontology scheme

An indication of the SimModel hierarchy and structure was documented before (Bazjanac et al. 2011a, O’Donnell et al. 2011). The model has a ‘project’ node at the top of its hierarchy, which is decomposed by various design alternatives. Each design alternative is then eventually decomposed by a ‘building’ node, which includes ‘building elements’, ‘zones’, ‘HVAC systems’ and ‘other systems’. The nodes ‘building elements’, ‘zones’, and ‘HVAC systems’ hereby map directly to the diverse domain models displayed in Figure 1. The node ‘other systems’ is included to enable future extensions of the SimModel (e.g. addition of electrical systems).

This SimModel hierarchy is contained within five main XSD files, each representing a part of the
SimModel data model, and one to-level XSD file, which combines all classes and properties of the other five XSD files into one SimModel class. The five XSD files are simcore.xsd, simbldg.xsd, simres.xsd, simgeom.xsd, simmep.xsd, and the top-level XSD file is the simmodel.xsd file. All XSD files are tightly linked together with cross-references to each other’s classes. SimModel contains a few hundred classes and it is hierarchically structured using classes, types and subtypes (O’Donnell et al. 2011).

2.3 SimModel instances

Whenever a SimModel is modelled and stored according to the SimModel ontology scheme recorded in the XSD files, one XML-file results with numerous instances of the classes and properties that are made available through the SimModel ontology scheme. These instances are all interlinked, resulting in one complete and consistent model.

3 BUILDING AN OWL ONTOLOGY FOR SIMMODEL

3.1 Towards a web of linked building data

We previously suggested to work towards a linked building data approach for the BEPS domain (Pauwels et al. 2014, in press). In this approach, diverse information sources used within a building project are linked together as needed (Fig. 2), and a management system is set up on top of those linked information sources for managing the links between the diverse models (Pauwels et al. 2010, 2011a, Bazjanac et al. 2011a, 2011b), a SimModel was proposed with the aim of bringing such models (IFC, gbXML, EnergyPlus, DOE-2) together into one centralised schema. It remains to be seen whether or not the SimModel ontology will eventually indeed be used in a centralised manner or not. If it is not, it still makes sense to combine SimModel information with other building models available in the AEC project and with data outside the AEC project (e.g. more static references, such as material information and geographic information). For this purpose, we will test the ‘linking’ approach documented above (Fig. 2).

This approach has its issues and challenges, like any other approach. Namely, this approach is not easily combined with an information exchange approach in which information is centralised (e.g. BIM approach). Instead, the linked building data approach consists mainly of pairwise links, resulting in loosely connected data that is often shared on-the-fly, when necessary and in the format that is required at a particular moment in time. As a result, it is not a straightforward matter to provide a management system and appropriately manage the diverse links in an intuitive manner.

Nevertheless, the linked building data approach appears to be promising. If it can be properly combined with centralised information exchange approaches, as they are promoted by BIM applications and most IFC-compatible applications, the decentralised information exchange approach might provide answers and opportunities to some of the issues of information exchange that also occur in the BEPS domain.

When considering the domain of BEP simulations, many different models are used in parallel, and not often is there one central BEPS model. For example, EnergyPlus and DOE-2 represent two such parallel schemata, which are not often used in combination within a project, let alone in a centralised model strategy, in which one of them precedes the other in importance. In Bazjanac et al. (2011a, 2011b), a SimModel was proposed with the aim of bringing such models (IFC, gbXML, EnergyPlus, DOE-2) together into one centralised schema. It remains to be seen whether or not the SimModel ontology will eventually indeed be used in a centralised manner or not. If it is not, it still makes sense to combine SimModel information with other building models available in the AEC project and with data outside the AEC project (e.g. more static references, such as material information and geographic information). For this purpose, we will test the ‘linking’ approach documented above (Fig. 2).

At the core of the linked building data approach are semantic web technologies (Berners-Lee et al. 2001). Central is the Resource Description Framework (RDF), which is a data model for the representation of any information in RDF graphs (Manola and Miller 2004). An RDF graph is a directed labelled graph, following a triple structure for each statement (object – predicate – subject). Using this structure, one can represent near to any information. The Web Ontology Language is based on the RDF data model and allows to represent ontologies (W3C OWL Working Group 2012). We already suggested in Pauwels et al. (2014, in press) how the SimModel ontology, as represented in the set of six XSD files, that information in one format / application can be related to information in another format / application.

Törmä 2013). The links made between building models are often pairwise.

Figure 2. In a linked building data strategy, information keeps following the format that is required by the application that is using it (nodes outside dashed circle). Additionally, it is tightly linked together on a data-level (nodes inside dashed circle), so
can be converted into an OWL ontology. We briefly document the resulting ontology here.

3.2 SimModel in OWL

In our approach to convert the SimModel into an OWL ontology, we started from the XML Schema Definition (XSD) files that represent the SimModel. Each class and property in these XSD files is converted into a corresponding class and property using the applicable OWL constructs. In the end, each class or subtype is converted into an OWL class (owl:Class), referring to an upper class when required, as indicated in Pauwels et al. (2014). For each class, the required properties are generated as owl:DatatypeProperty or owl:ObjectProperty declarations, resulting in a complete representation of the SimModel in six ontology files. As an example, for the simmep:SimFlowMover_Fan class and its property simmep:SimFlowMover_SimFlowMover_Name, the corresponding OWL representation is given in Figure 3.

![Figure 3. OWL representation for the simmep:SimFlowMover_Fan class and the datatype property simmep:SimFlowMover_SimFlowMover_Name, in the Notation-3 (N3) syntax.](image)

4 SIMMODELS AS RDF GRAPHS

4.1 The purpose

Once the SimModel ontology is available in OWL, one is able to represent SimModel information in well-structured RDF graphs, so that these graphs can be used by more advanced semantic web technologies, such as reasoning engines with a basis in Description Logics (DL) or SPARQL query interfaces. Furthermore, the possibility of representing SimModel information in RDF graphs enables addressing interoperability issues in the AEC domain using a linked building data approach. In this approach, the SimModel information of the BEP simulation domain can be more easily integrated with information outside the BEP simulation, to enable, for instance, a better overall building life-cycle management.

4.2 The conversion process

Diverse options can be followed in converting structured information into RDF graphs. Most of these conversion options have already been investigated in attempts to convert data in existing relational databases into RDF graphs. The main attempts in this regard is R2RML, a "language for expressing customized mappings from relational databases to RDF datasets" (Das et al. 2012), and the Direct Mapping approach (Arenas et al. 2012). Both approaches became W3C recommendations in 2012. Diverse implementations exist of both approaches (cfr. RDF2RDF 2012), including for instance D2RQ (D2RQ 2013), a platform that generates a virtual (read-only) RDF graph from a relational database and mapping document.

A lot of effort has also been made in the conversion of CSV files (CSV2RDF) and XML files (XML2RDF). For CSV2RDF mappings, the XLWrap mapping language (Langegger & Wöss 2009), the Mapping Master’s M2 (O’Connor et al. 2010) and TARQL (Cyganiak 2014) can be used. For XML2RDF mapings, the GRDDL mechanism (GRDDL Working Group 2007), Krextor (Lange 2009), and ReDeFer (Garcia & Gil 2007) can be used. Many other converter tools are available for a myriad of structured file formats. Most of these file formats and corresponding file formats are listed in W3C (2014). Several projects also aim at converting tabular data in general, not limited to CSV data or XML data, to reach semantically structured and interlinked data. Examples are Any23 (Apache 2014), Triplify/Sparqlify (Auer et al. 2009), Tabels (CTIC Foundation 2014), RDF Refine (DERI 2014). Furthermore, Dimou et al. (2014) aim at the definition of a source-independent mapping language, RML, which can be combined with a generic processor for converting basically anything into RDF graphs.
Although these approaches are highly promising and successful, most of them make conversions on a syntax level using rather generic conversion routines. For a complex schema like the SimModel schema, it seemed beneficial to interpret the SimModel schema and its instances to a sufficient level of detail and provide a custom mapping or conversion process that remains as faithful as possible to the original schema and instances.

In Figure 4, an indication is given of the structure of a sample XML fragment in a SimModel. This XML fragment displays the representation of a SimBldgStoryType_BuildingStory_Default element. The RefId attribute is associated, in the SimModel XSD files, to all simres:SimResourceObject and simcore:SimRoot elements, which are the two top-level classes of the SimModel. So, every instance in a SimModel.xml has a RefId property. Cross-references are made between XML entities throughout the XML document using these RefId properties, as can be seen in Figure 4 for the simres:TypeDefCreationParams and simres:ContainedBldgElementTypeDefs properties.

```
<SimBldgStoryType_BuildingStory_Default RefId="ID21540">
  <simres:SimUniqueId>
    94cd7da5-61ac-43f5-b77f-7908a99467c8
  </simres:SimUniqueId>
  <simres:SimModelType>
    BuildingStory
  </simres:SimModelType>
  <simres:SimModelSubtype>
    Default
  </simres:SimModelSubtype>
  <simres:SourceModelSchema>
    Simergy v0.9.500
  </simres:SourceModelSchema>
  <simres:TypeDefCreationParams>
    ID22182
  </simres:TypeDefCreationParams>
  <simres:ContainedBldgElementTypeDefs>
    ID21541 ID21542 ID21543 ID21544
  </simres:ContainedBldgElementTypeDefs>
</SimBldgStoryType_BuildingStory_Default>
```

Figure 4. XML representation for a SimBldgStoryType_BuildingStory_Default element.

In our converter application, the complete XML document is parsed and deserialised into JAVA class instances and property instances. The JAVA classes and properties were previously generated during the conversion of the XSD schema into OWL ontology files. At the end of the serialisation process, a SimModel instance is available in-memory. These instances are then iterated again and converted into corresponding RDF instances using JENA. The resulting JENA model can then be exported into files or triple stores as required.

4.3 The resulting RDF graph

Figure 5 displays an extract of the RDF graph that was generated for the XML fragment displayed in Figure 4. An instance inst:SimBldgStoryType_BuildingStory_Default_ID21540 is generated, of the type simres:SimBldgStoryType_BuildingStory_Default and with a number of properties associated (UniqueId, SimModelType, SimModelSubtype, TypeDefCreationParams, and so forth).

```
inst:SimBldgStoryType_BuildingStory_Default_ID21540 rdf:type simres:SimBldgStoryType_BuildingStory_Default ;
simres:SimUniqueId "94cd7da5-61ac-43f5-b77f-7908a99467c8"^^xsd:string ;
simres:SimModelType "BuildingStory"^^xsd:string ;
simres:SimModelSubtype "Default"^^xsd:string ;
simres:SourceModelSchema "Simergy v0.9.500"^^xsd:string ;
simres:TypeDefCreationParams SimBldgStoryParams_BldgStory_Default_ID22182 ;
```

Figure 5. RDF graph for the SimBldgStoryType_BuildingStory_Default element that was displayed in XML in Figure 4.

The URI simres:SimBldgStoryType_BuildingStory_Default uniquely identifies the instance in Figure 5. Furthermore, the instance is directly linked to the SimModel ontology (rdf:type), its datatype properties (e.g. SimUniqueID), and its object properties. Through its object properties, links are made to other elements in the SimModel, including, for instance, the instance simbldg:SimWindowType_Window_Default_ID21541. Hence, a directed labelled (RDF) graph is created that represents the instances in the SimModel, as they were described in the initial XML model.

5 A WEB OF LINKED BUILDING DATA

Once an RDF graph is available of the SimModel, concepts from other RDF graphs can be linked to the concepts / instances in the SimModel instance graph. As such, it is possible to link the SimModel instances to instances following the Industry Foundation Classes (IFC) ontology (Pauwels & Van Deursen...
2012). Similarly, links can be made to concepts in Geographical Information Systems (GIS) or material information, for instance. As such, a web of linked building data can be generated, including partial models with different types of information (GIS, IFC, SimModel, and so forth) for diverse buildings. After building this web of linked building data (LBD, 2013), query technologies and reasoning technologies can be deployed, thereby realising the benefits that can be offered by semantic web technologies at their maximum (Pauwels et al. 2011a, 2011b).

6 CONCLUSION

In this article, we documented the origin of the SimModel. We briefly indicated how the SimModel aims to combine diverse schema definitions in Building Energy Performance Simulation (BEPS) tools, such as EnergyPlus, IFC, gbXML, and so forth. It is conceived as an interoperable XML-based data model for the building simulation domain. As such, it can be aligned with many of the existing applications for BEP simulation.

In order to be able to better integrate SimModel information with other building information, we have aimed at representing SimModel information in RDF graphs. A first step towards this aim was documented earlier and resulted in a conversion service of the XML schema definition (XSD) files of the SimModel schema into corresponding OWL ontologies. With this set of ontologies, one is able to represent SimModel information in well-structured RDF graphs. The current paper presents our implementation of a SimModel-to-RDF conversion service, which generates RDF graphs using the given SimModel XML files. The resulting graphs can be used by more advanced semantic web technologies, such as reasoning engines with a basis in Description Logics (DL) or SPARQL query interfaces. Furthermore, the possibility of representing SimModel information in RDF graphs enables building a web of linked building data. In such a web of building data, the SimModel information of the BEPS domain can be more easily integrated with information outside the BEP simulation, to enable for instance a better overall building lifecycle management.

7 REFERENCES


