proceedings

Proceedings

Energy-Related Data Integration Using Semantic Data Models for Energy Efficient Retrofitting Projects †

Álvaro Sicilia * and Gonçal Costa

ARC, La Salle Engineering and Architecture, Ramon Llull University, 08022 Barcelona, Spain; gcosta@salleurl.edu

* Correspondence: asicilia@salleurl.edu; Tel.: +34-932-902-449

† Presented at Sustainable Places 2017 (SP2017) Conference, Middlesbrough, UK, 28–30 June 2017.

Published: 5 December 2017

Abstract: Energy efficient retrofitting projects of urban areas require to analyze data from multiple sources and domains—BIM, GIS, statistics, energy data, and climate. An interoperability solution is needed to overcome the semantic and structural heterogeneity of the data sources. Within OptEEmAL project, we have design and implemented a District Data Model which integrates multiple data sources and makes them interoperable with several simulation tools (Energy plus, Nest, CitySim) using Semantic Web technologies, namely, ontologies and SPARQL construct queries.

Keywords: energy efficiency retrofitting; semantic web; linked data; data integration; ontologies; SPARQL

1. Introduction

In order to carry out energy efficient retrofitting projects of urban areas, it is necessary to analyze data from multiple sources—BIM, GIS, statistics, energy prices, climate—and domains—urban planning, energy, economic—for the analysis of the building energy performance. This involves calculating different district performance indicators (DPIs) from different domains related to retrofitting projects in order to support stakeholders in decision-making activities. To automate this process, a Web-based platform for district energy-efficient retrofitting is currently being developed within the context of the OptEEmAL project funded by the Horizon 2020 programme. In this platform, data from multiples domains and scales, and tools (e.g., Energy Plus and Nest) are integrated into a district data model. This makes it easier to calculate DPIs from different domains (e.g., energy, comfort, environmental, economic, urban, and social) [1]. One of the main challenges in this scenario is to make the data and the simulation tools interoperable so that they can overcome two existing barriers:

- Data sources needed for district retrofitting projects are generated and maintained by multiple experts (planners, designers) using different tools (BIM authoring tools and GIS software).
- A variety of simulation tools—with different input formats and data requirements—need to have access to this data to carry out assessments and analysis.

In this context, an interoperability solution is required to overcome these two barriers simultaneously: (1) to integrate data from multiple heterogeneous sources and (2) to ensure the communication between the integrated data and an open set of simulation tools.

Standard data models such as IFC and CityGML aim at solving interoperability by providing centralized models that can grow to embrace new domains by means of extension mechanisms. However, this approach has proved to have some limitations: (i) difficulties to reach a consensus among a community of users; (ii) lack of flexibility of the data models to adapt to changes; and (iii) the loss of information after exporting and importing data through applications. An alternative approach to centralized standard data models is the integration of data using Semantic Web technologies. In this context, the role of Semantic Web technologies is to provide bridges between multiple models by means of ontologies. Such ontologies do not need to be created from scratch but they can be based on standards like ISO or CEN. This way, semantic-based interoperability brings together the best of both worlds: a standardization based on the ontologies—rather than on the data models—and a decentralization of the data models and tools which are interlinked through the ontologies [2].

We have devised and implemented a technological solution based on standard data models such as IFC and CityGML intertwined with Semantic Web technologies in the so-called District Data Model which lies in the core of the OptEEmAL platform.

2. A district Data Model for Retrofitting Projects

The District Data Model (DDM) developed within OptEEmAL project provides a semantically integrated data model (including information about the geometry, materials, energy conservation measures, climate, social data, and equipment, at the building and urban scales) that the simulation tools need to calculate the district performance indicators.

The goals of the DDM is to integrate the native data provided by users to carry out the retrofitting project and to make this integrated data interoperable with the simulation tools which have different input data requirements and formats. To achieve these goals, the DDM has been designed as a set of simulation data models based on domain ontologies (e.g., urban, energy simulation, social, energy carriers) that each of one encompasses the data required to conceptualize each domain (Figure 1). For example, the ontology for the energy simulation domain includes the concepts, properties and attributes required by any kind of energy simulation tool such as Energy Plus and Nest. Since finding or creating an ontology to encapsulate the requirements of all existing simulation tools is a task that requires a great deal of resources, we have limited the tools that the ontologies have to consider to the ones to be used within the OptEEmAL project. Namely, Energy Plus, Nest, an Economic tool, and a HVAC tool.

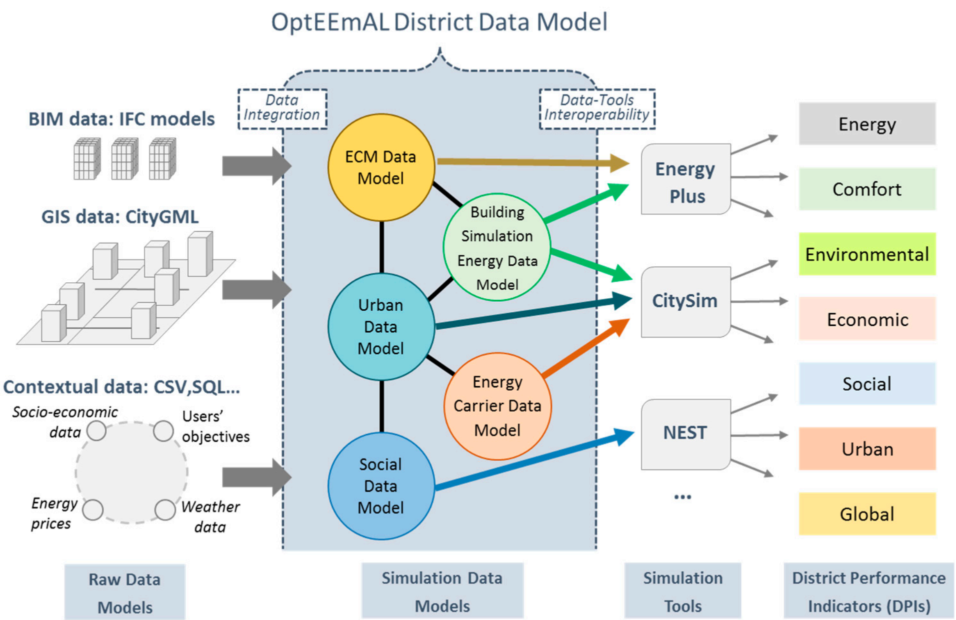

Figure 1. Conceptualization of the District Data Model.

The simulation data models that are considered in the DDM are the following:

- Urban Data Model: a model to represent the district. Buildings, district energy systems, their "physical" relations (e.g., buildings connected to district energy systems).
- Building Simulation Energy Data Model: a model to represent the energy simulation of a building. It includes the geometric representation, material properties, building energy systems, cast shadows.
- ECM Data Model: a model to represent an Energy Conservative Measure and its application in a district or building. It includes costs (e.g., sales, maintenance, operational), global warming potential impact, and where it has been applied.
- Energy Carrier Data Model: a model to represent the energy carriers of a project. It includes energy carrier costs, global warming potential impact, and emissions coefficients.
- Social Data Model: a model to represent the social data of a district. It includes number of inhabitants and their income.

The ontologies used to represent the simulation data models of the domains have been created by reusing parts and concepts from different existing ontologies. One of them is the Semanco ontology [3] which contains concepts and properties for urban energy models at different scales. To represent the data of building energy simulations, the SimModel ontology [4] was selected since it comprises all the concepts and properties to represent an energy simulation model.

3. Creating District Data Models with Semantic Web Technologies

In OptEEmAL, Semantic Web technologies have been applied to integrate data from multiple sources and to make it interoperable with a variety of simulation tools. The methodology for data integration using SPARQL Constructs presented in [5] has been applied to integrate the following data sources:

- District information: a CityGML model including the building models of the district.
- Building information: an IFC model for each building to be retrofitted is provided.
- Contextual data: climate data and energy carrier prices (e.g., electricity, gas, biomass).

The methodology consists of three steps. The first step is to select the ontologies to represent the data of each possible source. For example, in the case of BIM data (i.e., IFC file) the selected ontology is ifcOWL [6]. The second step is to transform the input data sources, described according to the corresponding ontologies selected, into RDF graph models. We have used a tool to transform IFC data in STEP format into RDF according to ifcOWL [7]. In the third step, the RDF data corresponding to each input data source are integrated into simulation data models that represent the domains described in the previous section. This step involves the implementation of SPARQL Construct queries which are executed over the RDF graph models generated in the second step.

Through this methodology, the structural and semantic heterogeneity between the data models can be overcome. An example of structural and semantic differences between ifcOWL and SimModel OWL can be seen in Figure 2. In ifcOWL, the height of a building story value is an attribute of the class *ifc:IfcLengthMeasure* which is connected to the class *ifc:IfcBuildingStorey*. However, in SimModel OWL, the height of the story is directly an attribute of the *sim:SimBuildingStory* class. Moreover, the height units in ifcOWL are meters and in SimModel OWL in millimeters. The SPARQL Construct in Figure 2 transforms the data according to ifcOWL to SimModel OWL, solving structural and semantic differences.

To carry out the transformation of the data sources between ifcOWL and SimModel ontology, 50 queries have been implemented, with an average size of approximately 80 lines. Those queries have been applied in the case of an IFC model of a real building: 11-storey building with 284 rooms. The whole transformation process takes less than 2 min to complete on a regular desktop computer (OS: Windows 7 Professional 32-bits, Processor: Intel® Core™2 Quad CPU Q9400 @ 2.66 GHz, RAM memory: 4 GB).

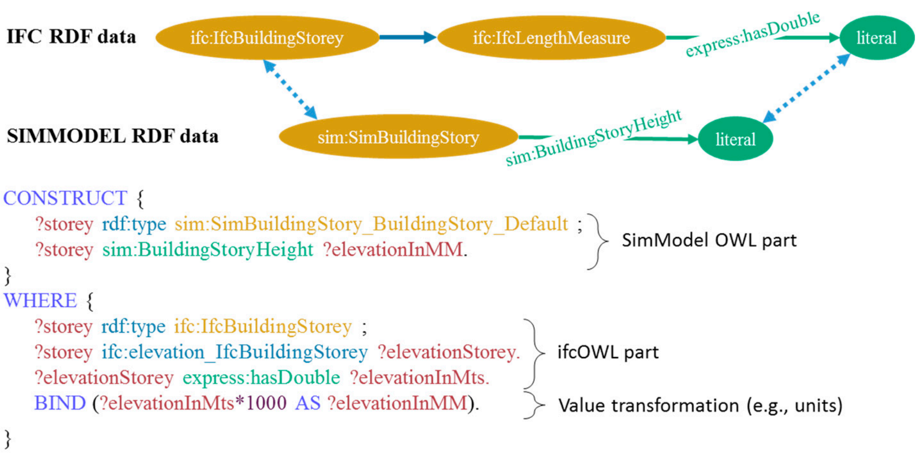

Figure 2. Example of overcoming semantic and structural heterogeneity.

4. Conclusions

In this paper we have presented the District Data Model developed within OptEEmAL project which solves interoperability using Semantic Web technologies between data from multiples domains and sources and simulation tools. The transformation rules to integrate data sources can be maintained and extended by the community since it has been coded in a declarative language such as SPARQL instead of using procedural programming. The creation of the transformation rules using SPARQL Construct queries requires knowledge of the source (e.g., IFC) and target (e.g., SimModel) models. The further work in the next phase of the project will be to complete the mappings and transformation rules as well as to optimize the existing ones. For example, to include active systems.

Acknowledgments: The work presented in this paper is based on research conducted within the project "Optimised Energy Efficient Design Platform for Refurbishment at District Level", which has received funding from the European Union Horizon 2020 Framework Programme (H2020/2014–2020) under grant agreement No. 680676.

Author Contributions: Álvaro Sicilia and Gonçal Costa designed and implemented the District Data Model described in the paper; Álvaro Sicilia and Gonçal Costa wrote the paper.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Costa, G.; Sicilia, A.; Lilis, G.N.; Rovas, D.V.; Izkara, J.L. A comprehensive ontologies-based framework to support retrofitting design of energy-efficient districts. In Proceedings of the European Conference on Product & Process Modelling, Limassol, Cyprus, 7–9 September 2016.
2. Sicilia, A.; Madrazo, L.; Pleguezuelos, J. Integrating multiple data sources, domains and tools in urban energy models using semantic technologies. In Proceedings of the 10th European Conference on Product & Process Modelling, Vienna, Austria, October 2014.
3. Nemirovski, G.; Nolle, A.; Sicilia, Á.; Ballarini, I.; Corrado, V. Data integration driven ontology design, case study smart city. In Proceedings of the 3rd International Conference on Web Intelligence, Mining and Semantics, Madrid, Spain, 12–14 June 2013; p. 43.
4. Pauwels, P.; Corry, E.; O'Donnell, J. Representing SimModel in the web ontology language. *Comput. Civ. Build. Eng.* **2014**, 2271–2278, doi:10.1061/9780784413616.282.
5. Costa, G.; Sicilia, Á. Methodology for data integration using SPARQL Constructs in the AEC industry. In Proceedings of the 5th Linked Data in Architecture and Construction Workshop (LDAC2017), Dijon, France, 13–15 November 2017.

6. Pauwels, P.; Terkaj, W. EXPRESS to OWL for construction industry: Towards a recommendable and usable ifcOWL ontology. *Autom. Constr.* **2016**, *63*, 100–133.
7. Pauwels, P. Implementation of IFC-to-RDF Conversion by Ghent University (Multimedia Lab—SmartLab) and Aalto University. 2015. Available online: https://github.com/mmlab/IFC-to-RDF-converter (accessed on 29 September 2017).