RDB2Onto: Relational Database Data to Ontology Individuals Mapping

Michal Laclavík[1]

Institute of Informatics, Slovak Academy of Sciences, Dúbravská cesta 9, 845 07 Bratislava, Slovakia,
laclavik.ui@savba.sk,
WWW home page: http://ikt.ui.sav.sk/

Abstract. In this paper we present the approach for creating semantic metadata from relational database data. When building ontology based information systems, it is often needed to convert or replicate data from existing information systems such as databases to the ontology based information systems, if the ontology based systems want to work with real data. RDB2Onto converts selected data from a relational database to a RDF/OWL ontology document based on a defined template. Such filled in templates can be then stored to the ontology based knowledge memory.

1 Introduction

Building ontology based information systems, it is frequently necessary to convert or replicate data from existing information systems such as databases to the ontology based information systems, if the ontology based systems want to work with real data. Usually data in existing information systems are stored in a Relational Database. Such problem arises also in the NAZOU project [9] where some knowledge acquisition and maintenance tools store results data only in the RDB database. Due to the common presentation frame work [8] the result data need to appear also in its ontology form and this is the place where RDB2Onto plays its role. In addition, a large quantity of data can be found on the web automatically generated from relational databases, often referred to as the Deep Web [1]. If we want to create web content based on semantic web technologies such as OWL [5], we have to solve conversion of RDB data to ontology data, while several approaches exist. Most of them are usually based on creating new quite complicated mapping languages like D2R MAP [3], D2R [4] or R2O [2]. In our approach we try to give a more simple solution based on SQL queries and RDF/OWL templates which are filled in with results of SQL query.

2 Approach

The goal of the tool is to provide Relational Database Data to Ontology Individuals Mapping. The tool works on a domain ontology model and a relational

database. The overall idea is to map SQL query to RDF/OWL XML template. Such OWL data are then sent to an ontology model. The tool is being implemented in Java using Jena [6] or Sesame [7] library for ontology manipulation and MySQL database for testing but it is possible to use any other relational database using JDBC connector.

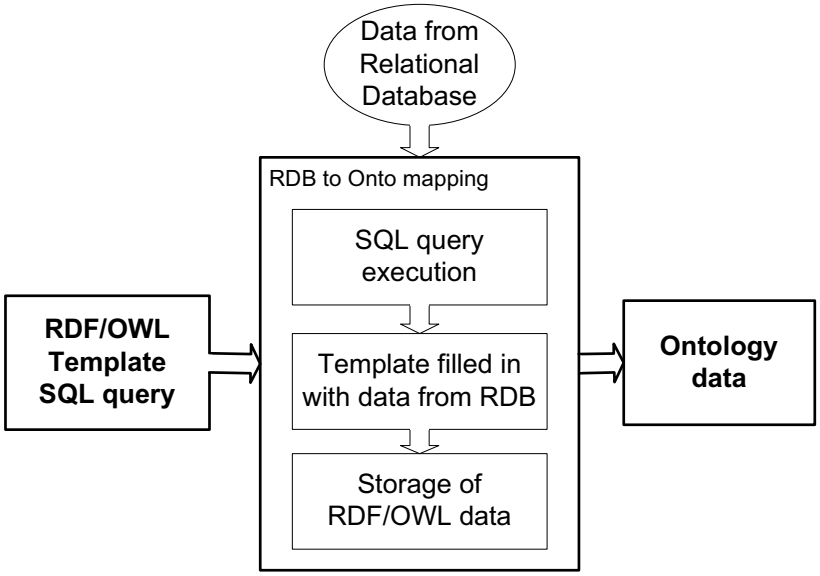

Fig. 1. RDB2Onto Architecture

Architecture of the tool is shown on figure 1. It contains 3 basic steps which are explained on the following example: There is a *document* table with following fields: *id*, *title*, *author_id* and *url* in relational database. Moreover there is a table *author* with the following fields: *id* and *name*. There is a *Document* class with properties *hasTitle* and *hasUrl* in ontology. Furthermore an *Author* class with property *hasName* is defined in the ontology In this example SQL (1) query will look as follows:

$$\begin{aligned} &SELECT \\ &\quad d.id\ AS\ d\_id,\ title,\ url,\ a.id\ AS\ a\_id,\ name \\ &FROM \\ &\quad document\ d,\ author\ a \\ &WHERE \\ &\quad d.author\_id\ =\ a.id \end{aligned} \tag{1}$$

The SQL query is executed and for each row of the query results it fills in the XML-based OWL template 2. Each element enclosed with {} brackets is replaced with adequate value from SQL query for a given row and composed OWL data are stored to the ontology model.

$$
\begin{aligned}
&<?xml version = "1.0"?>\\
&< rdf:RDF\\
&\quad xmlns:rdf = "http://www.w3.org/1999/02/22-rdf-syntax-ns\#"\\
&\quad xmlns:xsd = "http://www.w3.org/2001/XMLSchema\#"\\
&\quad xmlns:rdfs = "http://www.w3.org/2000/01/rdf-schema\#"\\
&\quad xmlns:owl = "http://www.w3.org/2002/07/owl\#"\\
&\quad xmlns = "http://onto.ui.sav.sk/\#"\\
&\quad xml:base = "http://onto.ui.sav.sk/" >\\
&\quad < Documentrdf:ID = "document\_\{d\_id\}" >\\
&\quad\quad < hasAuthorrdf:resource = "\#author\_\{a\_id\}" / >\\
&\quad\quad < hasTitlerdf:datatype = "http://www.w3.org/2001/XMLSchema\#string"\\
&\quad\quad > \{title\} < /hasTtitle >\\
&\quad\quad < hasURLrdf:datatype = "http://www.w3.org/2001/XMLSchema\#anyURI"\\
&\quad\quad > \{url\} < /hasURL >\\
&\quad < /Document >\\
&\quad < Authorrdf:ID = "author\_\{a\_id\}" >\\
&\quad\quad < hasNamerdf:datatype = "http://www.w3.org/2001/XMLSchema\#string"\\
&\quad\quad > \{name\} < /hasName >\\
&\quad < /Document >\\
&< /rdf:RDF >
\end{aligned}
\tag{2}
$$

3 Conclusions

Such solution is quite simple but powerful. It is possible to receive any complicated data from relational database upon an SQL request. The advantage of the solution is its simplicity and easy configuration for concrete data mapping. The approaches as [4] [3] [2] provide infrastructure and languages for relational data mapping by setting data dependencies between RDB data and ontology. This can be sometimes too complicated and heavy solution which require too much effort to learn, set up and use the solution. RDB2Onto requires only knowledge of SQL and RDF/OWL and thus can be applied with minimum effort.

Acknowledgment This work was partially supported by the Slovak State Programme of Research and Development "Establishing of Information Society" under the contract No. 1025/04, K-Wf Grid EU RTD IST FP6-511385 and RAPORT APVT-51-024604 projects.

References

1. Bergman MK.: The Deep Web: Surfacing hidden value. White paper. Sept 2001
2. Barrasa J., Corcho Ó., Gómez-Pérez A.: R2O, an Extensible and Semantically Based Database-to-ontology Mapping Language, Second Workshop on Semantic Web and Databases (SWDB2004). Toronto, Canada. August 2004.
3. Bizer C.: D2R MAP - A DB to RDF Mapping Language. 12th International World Wide Web Conference, Budapest. May 2003
4. Barrasa J, Corcho O, Gómez-Pérez A.: FundFinder - A case study of Database-to-ontology mapping. Semantic Integration Workshop, ISWC 2003. Sanibel Island, Florida. Sept 2003
5. W3C: Web Ontology Language OWL, 2005, http://www.w3.org/TR/owlfeatures/
6. HP Labs and Open Source Community, Jena Semantic Web Library, 2005, http://www.sf.net/
7. OpenRDF.org, Sesame RDF Database, 2006, http://www.openrdf.org/
8. Návrat, P. - Bieliková M. - Rozinajová, V.: Methods and Tools for Acquiring and Presenting Information and Knowledge in the Web. In: CompSysTech 2005, B. Rachev, A. Smrikarov (Eds.), Varna, Bulgaria, June 2005. - pp. IIIB.7.1-IIIB.7.6.
9. NAZOU Project Website, 2006, http://nazou.fiit.stuba.sk/