SEMANTIC WEB TECHNOLOGIES IN EEG/ERP DOMAIN
Software Solution

Roman Mouček, Petr Ježek and Václav Papež
Department of Computer Science and Engineering, University of West Bohemia
Univerzitní 8, 306 14 Pilsen, Czech Republic

Keywords: Semantic web, RDF, OWL, Relational database, Automated transformation, Electroencephalography, Event-related potentials, EEG/ERP portal, Jena, D2RQ, OWL API.

Abstract: The paper shortly describes EEG/ERP domain and the system for storage and management of data and metadata from EEG/ERP experiments. The system has been developed using common technologies and data structures (object oriented code, relational database). A registration of the system as a recognized data source requires representation of its data and metadata by semantic web resources. A short overview of differences between common data structures and semantic web representations is given and possible approaches to automated transformation of data and metadata from common data structures are mentioned. The existing semantic web frameworks and tools, widely tested for our purposes, are listed. Finally, the integration of selected tools and the software solution for automated transformation from the relational database to a semantic web representation is described.

1 INTRODUCTION

Our research group at Department of Computer Science and Engineering, University of West Bohemia in cooperation with other partner institutions (Czech Technical University in Prague, University Hospital in Pilsen, Škoda Auto, Inc....) specializes in the research of attention, especially attention of drivers, children with developmental coordination disorders and seriously injured people. We widely use the methods of electroencephalography (EEG) and methods of event-related potentials (ERP).

We are members of International Neuroinformatics Coordinating Facility (INCF) (INCF, 2010). Among other things this organization supports collaboration of national nodes for neuroinformatics and standardization of data and metadata structures in the domain of neuroinformatics.

We have developed a software tool for a long term storage and management of EEG/ERP experiments. This tool is based on a relational database and modern Java technologies. A registration of this data source within the neuroinformatics community requires providing data and metadata in a form of ontology. This led us to the idea of using semantic web technologies. However, our goal is to provide this ontology automatically from the existing structures, not to build a separate ontological model.

The aim of this paper is to provide readers with our approach in usage of semantic web resources, techniques, frameworks and tools in the EEG/ERP domain. Finally, a presentation of our software solution is given.

The paper is organized as follows. At first we briefly introduce EEG/ERP domain. Then the system for storage and management of EEG/ERP experiments is shortly described. Possibilities for transformation of data and metadata from common data structures are mentioned and some known approaches to a mapping of knowledge from common data structures (relational database, object oriented code) to semantic web languages are summarized. Existing mapping tools, which were widely tested, are listed and the final software solution for a transformation of relational data to a semantic web representation is described.

Mouček R., Ježek P. and Papež V..
SEMANTIC WEB TECHNOLOGIES IN EEG/ERP DOMAIN - Software Solution.
DOI: 10.5220/0003291406180621
In *Proceedings of the International Conference on Health Informatics* (HEALTHINF-2011), pages 618-621
ISBN: 978-989-8425-34-8

2 EEG/ERP DOMAIN

EEG and ERP methods are widely used in research of brain processes. EEG and ERP experiments take usually long time and produce extensive data. Experiments made in our laboratory were initially recorded and managed using a common folder structure (commercial software tools are used during recordings). Because an appropriate software tool for a long-term storage and management of data obtained during EEG/ERP experiments did not exist, we have developed a custom software solution. This software tool, EEG/ERP portal, is now used in our department and it will be provided to neuroinformatics community after more extensive testing.

The basic ideas and features of the system are described in (Ježek, 2010). From that time the software incorporated a number of changes and improvements. Not only EEG/ERP experiments but also data and metadata structures in various formats, scenario templates of EEG/ERP experiments, groups of people involved in experiments, etc. are now covered by the features of EEG/ERP portal. However, these changes and improvements are not important for the aim of this paper.

3 EEG/ERP PORTAL

EEG/ERP portal enables community researchers to store, update and download data and metadata from EEG/ERP experiments. The system is developed as a standalone product; the database access is available through a web interface.

The system essentially offers the following set of features (the set of accessible features depends on a specific user role):

- User authentication
- Storage, update, and download of EEG/ERP data and metadata
- Storage, update and download of EEG/ERP experimental design (experimental scenarios)
- Storage, update and download of data related to testing subjects
- Sharing of knowledge and working in groups

The three layer architecture (MVC pattern) was used to build the EEG/ERP portal. It consists of the persistent layer (relational database), the application layer (object oriented code, object relational mapping from persistence layer) and the presentation layer (JSP).

4 MAPPING TO SEMANTIC WEB LANGUAGES

Scientific papers often describe a domain using a semantic web language and consider this kind of domain modelling as a crucial point of software solution. However, real software applications use up the underlying data structures such as relational databases or object classes. Then a mapping from common data structures to a semantic web language is necessary.

Concerning the architectural layers of EEG/ERP portal there is a question which layer is more feasible for mapping into a semantic web language (RDF, OWL). We considered two possibilities:

- Mapping from the persistence layer (relational database)
- Mapping from the application layer (object oriented code)

The mapping from the relational database is straightforward while the mapping from the application layer to a semantic web language requires the precedent object relational mapping (provided by Hibernate framework in our case). Since our goal is to generate an ontology, we will consider OWL as the target semantic web language.

However, it is important to emphasize that OWL is neither a database framework nor a programming language. There are fundamental differences in richness of semantics between OWL (Description Logic based system) and relational database or object oriented systems. If some information is not present in a database, it is considered to be false (closed-world assumption). By contrast, if some information is not present in an OWL document, it may be missing and possibly true (open-world assumption). On the other hand, there are several approaches how to bridge at least some of the semantic gaps. A more detailed overview you can find in (Mouček, 2010).

5 SEMANTIC WEB FRAMEWORKS AND TOOLS

There are many frameworks and software tools, which are considered to generate RDF or OWL output from a relational database or an object oriented code. Some of these frameworks and tools exist only as initial proposals or prototypes described in scientific papers, while some of them have been really implemented. The following list includes existing frameworks and software tools, which were considered to be used for our purposes. More detailed information about these frameworks and tools and our experience with them was already given in (Mouček, 2010):

- Jena
- D2RQ
- Virtuoso
- SquirrelRDF
- MetaMorphoses
- JRDF
- SPASQL
- Sommer
- JenaBean
- Java2OWL-S
- OWL API

6 MAPPING FROM RELATIONAL DATABASE

As we already mentioned our goal is to register our system as the recognized data source and to provide the system ontology. It means that we need to perform only one-sided mapping from a relational database (object oriented code) to OWL; then we need to use only a subset of semantic richness of RDF(S) and OWL.

We decided to try out two parallel approaches. The first approach includes the transformation of relational database into ontology using D2RQ tool, Jena and OWL API. The second approach includes the transformation from object oriented code to OWL using JenaBean tool.

We focus on the first approach; the second approach to transformation process is out of scope of this paper.

6.1 Integration of Selected Tools

We selected three tools, which were finally integrated to provide the domain ontology from the relational database of EEG/ERP portal. Jena, the most widespread tool, was selected for inner representation of RDF graph. Sesame framework was rejected because of non-standard approach to stored data. Moreover, the next selected tool, D2RQ, cooperates with Jena.

EEG/ERP raw data are stored in the database as binary files. D2RQ as the second selected tool had difficulties with BLOB and CLOB data types. The transformation process generated program exceptions in this case. We solved it by modification of D2RQ source code. BLOB data types (raw data) are processed but they are not included into RDF graph. CLOB data types (e.g. XML scenarios) are also processed but they are included into RDF graph. OWL API was used for storing RDF in various formats.

The integration of selected tools and the transformation process can be described as follows. D2RQ connects to the relational database, determines its structure and loads this structure to the virtual RDF graph, which is derived from Jena graph. Jena transforms this graph into RDF/XML format. OWL API parses RDF/XML format and converts it to a format required by a user. Then it is, for example, possible to transform relational data to ontology provided in OWL/XML format. It is clear that this ontology is semantically as rich as the corresponding relational database.

6.2 Software Description

The final software tool is divided into two parts – the application library and the graphical user interface. The application library defines a method for transformation of the relational database to RDF graph and a method, which stores this graph into a resulting file according to a selected format. The UML diagram of both parts is shown in Figure 1.

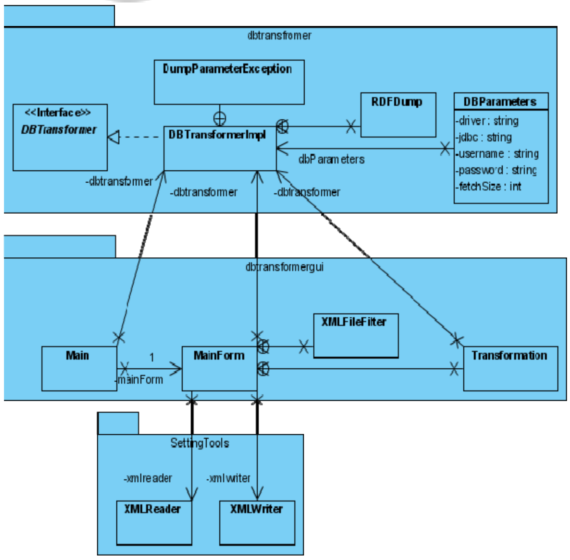

Figure 1: Software tool for transformation of relational database to semantic web representation - UML diagram.

The graphical user interface provides a simple and transparent way to control the application (Figure 2).

6.3 Formats of Output Files

It is possible to generate the system ontology in the following formats: RDF/XML, OWL/XML, Turtle, Manchester OWL syntax, KRSS2, Latex, and OWL functional syntax. The following example shows a record of one person from table PERSON in Turtle syntax.

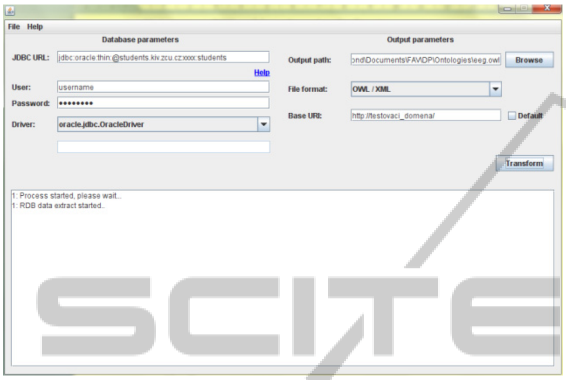

Figure 2: Software tool for transformation of relational database to semantic web representation - graphical user Interface.

```
###http://eeg_domain/database/EEGERP.PE
RSON/22

<http://eeg_domain/database/EEGERP.PERS
ON/22>
      rdf:type
      <http://localhost/vocab/EEGERP.PER
      SON> ,
        owl:NamedIndividual ;

      rdfs:label "EEGERP.PERSON #22" ;

<http://localhost/vocab/PERSON_PERSON_I
D> "22"^^xsd:decimal ;
<http://localhost/vocab/PERSON_SURNAME>
"Papez" ;
<http://localhost/vocab/PERSON_GIVENNAM
E>"Vaclav" ;

<http://localhost/vocab/PERSON_AUTHORIT
Y>"ROLE_ADMIN" ;
<http://localhost/vocab/PERSON_PASSWORD
>"bf2bc2545a4a5f5683d9ef3ed0d977e0";
<http://localhost/vocab/PERSON_USERNAME
>"vpapez" ;

<http://localhost/vocab/PERSON_EMAIL>"v
papez@kiv.zcu.cz";
<http://localhost/vocab/PERSON_GENDER>"
M" .
```

Currently the system ontology is automatically generated from 23 tables, which are created and maintained in the relational database of the EEG/ERP portal.

7 CONCLUSIONS

The presented paper shortly describes the system for storage and management of data and metadata from EEG/ERP experiments. This EEG/ERP portal has been developed using common technologies and data structures (object oriented code, relational database). A registration of the portal as the recognized data source requires representation of its data and metadata using semantic web resources.

A short description of differences between common data structures (relational databases and object oriented code) and semantic web representations (OWL) was given. The existing semantic web frameworks and tools, widely tested for our purposes, were listed. The integration of selected tools and the final software application that allows automated transformation from the relational database to a final semantic web representation were described.

In the near future we consider finishing a transformation from an object oriented code to a semantic web representation. We will also investigate possibilities of both side mapping between common structures and widely used semantic web representations.

ACKNOWLEDGEMENTS

This work was supported by Grant no. 2C06009 Cot-Sewing.

REFERENCES

International Neuroinformatics Coordinating Facility (INCF). (2010). Retrieved October 22, 2010, from http://www.incf.org/.

Ježek, P., Mouček, R., 2010. Database of EEG/ERP experiments. In *HEALTHINF 2010 - Proceedings of the Third International Conference on Health Informatics*. Valencia, Spain.

Mouček, R., Ježek, P., 2010. System for storage and management of EEG/ERP experiments – generation of ontology. *Proceedings of the 12th International Conference on Enterprise Information System*. Madeira, Portugal.