PROCEEDINGS

of the Union of Scientists - Ruse

Book 5

Mathematics, Informatics and Physics

Volume 12, 2015

RUSE

PROCEEDINGS
OF THE UNION OF SCIENTISTS - RUSE

Union of Scientists - Ruse
16, Konstantin Irechek Street
7000 Ruse
BULGARIA
Phone: (++359 82) 828 135,
(++359 82) 841 634
E-mail: suruse@uni-ruse.bg
web: suruse.uni-ruse.bg

Contacts with Editor
Phone: (++359 82) 888 738
E-mail: zzdravkova@uni-ruse.bg

PROCEEDINGS
of the Union of Scientists – Ruse

ISSN 1314-3077

ISSN 1314-3077

Proceedings

of the Union of Scientists – Ruse

Contains five books:

1. Technical Sciences
2. Medicine and Ecology
3. Agrarian and Veterinary Medical Sciences
4. Social Sciences
5. Mathematics, Informatics and Physics

The Ruse Branch of the Union of Scientists in Bulgaria was founded in 1956. Its first Chairman was Prof. Stoyan Petrov. He was followed by Prof. Trifon Georgiev, Prof. Kolyo Vasilev, Prof. Georgi Popov, Prof. Mityo Kanev, Assoc. Prof. Boris Borisov, Prof. Emil Marinov, Prof. Hristo Beloev. The individual members number nearly 300 recognized scientists from Ruse, organized in 13 scientific sections. There are several collective members too – organizations and companies from Ruse, known for their success in the field of science and higher education, or their applied research activities. The activities of the Union of Scientists – Ruse are numerous: scientific, educational and other humanitarian events directly related to hot issues in the development of Ruse region, including its infrastructure, environment, history and future development; commitment to the development of the scientific organizations in Ruse, the professional development and growth of the scientists and the protection of their individual rights.

The Union of Scientists – Ruse (US – Ruse) organizes publishing of scientific and popular informative literature, and since 1998 – the “Proceedings of the Union of Scientists- Ruse".

BOOK 5

"MATHEMATICS, INFORMATICS AND PHYSICS"

VOLUME 12

CONTENTS

Mathematics

Informatics

BOOK 5

"MATHEMATICS, INFORMATICS AND PHYSICS"

VOLUME 12

Physics

Applications

web: suruse.uni-ruse.bg

TRANSLATING A SQL APPLICATION DATA TO SEMANTIC WEB

Desislava Baeva

Angel Kanchev University of Ruse

Abstract: *Semantic web technology is more and more often applied to a large spectrum of applications where domain knowledge is conceptualized and formalized due to Ontology, as a support for diversified processing operated by machines. A crucial factor settling the Semantic Web success is the development of methods for transforming relational databases into Semantic Web accessible data. This article describes an ontology view, which is based on an RDB information.*

Keywords*: semantic web, ontology*

INTRODUCTION

Creative ideas in the computer industry are so rapidly developed nowadays that innovations are every day introduced and what was futuristic a year or two ago appears a good practice at present. Yet, all innovations can be unified by a common denominator – their aim to make computer technologies more powerful and smarter, getting closer to the specificity of human potentials to logically analyze the incoming information.

An approach to that idea is the Semantic Web. According to the W3C, "The Semantic Web provides a common framework that allows data to be shared and reused across application, enterprise, and community boundaries". The purpose of the Semantic Web architecture is to build a standardized model of all data types (RDF, which encodes arbitrarily labeled directed graph) by adding metadata encoded as ontology (in RFB or OWL), as well as data connectivity ontologies and other data on the network. The standardized protocol of services (SPARQL) helps the network to seem as a giant database and allows distributed query execution [3].

CONTEXT

A crucial factor settling the Semantic Web success is the development of methods for transforming relational databases into Semantic Web accessible data. The modern trends in the website design impose the thesis that working out automated methodology for integrating relational databases with the Semantic Web is of major importance.

The methods developed for direct automatic mapping transform mechanically the SQL database schema and its contents into ontology, usually by using translation rules, and the relational data is presented as RDF instances of the generated ontology.

W3C provides recommendations for implementing relational database translations into RDF graph. The core idea is the relational database information to be transformed into ontology by translating the table names into classes, the columns names - into data contents, the key relationships - into object properties. [3]

When transforming relational database into ontology the following problems have to be considered:

- Data losses in result of transformation - the original data must be adequately specified.
- Semantic loss – in some cases the transformation cannot be implemented without data losses in result of which a relational database construct cannot be translated into an ontology construct. Therefore, it is necessary to analyze the quality of transformation in order to avoid semantic losses.

- Focus on structures – besides structures mapping (i.e. tables, columns, etc.) mechanisms for data mapping (i.e. fields) should be provided.
- Focus on data – data mapping must include data types.
- Application - in some cases the transformation is not really an integrity, meaning that the application is rather constrained. For instance, if translation is allowed only for exotic relational databases and it is not applied in practical situations then the transformation suffers from the problem of application.
- Correctness – the transformation must have provable correctness.

The advantages of creating a relational data RDF view comprise:

- Integration – the data from different RDBs can be integrated with the help of RDF semantics and mechanisms. In this context, the constructed Semantic Web can be represented as an enormous database. Moreover, the information in the database can be easily integrated with the information flow from other data sources.
- Data retrieval – unlike relational databases the queries can involve different data sources and more powerful retrieving methods can be constructed.

Tirmizi et al. presented a type of layer cake mapping methodology [2] that was achieved between the Open Biomedical Ontology language and the Semantic Web stack. The organization of the Semantic Web regarding the assignment of expressive power to each of its respective layers embodies a larger organizing principal.

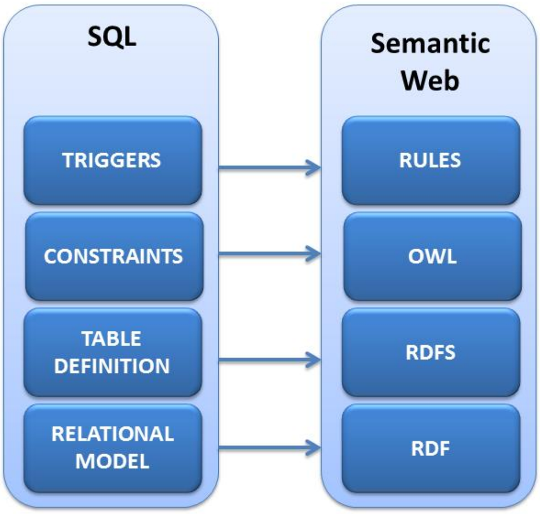

Fig.1. Layer Cake Correspondence betwen SQL and the SemanticWeb

The systems that translate SQL databases to the Semantic Web have different goals with respect to the extent of the expressive power of SQL-DDL is targeted, and similarly the amount of expressive power of the Semantic Web is exploited. Furthermore, the choices made by individual projects are as dependent on the time frame of the efforts as they are on the research itself. OWL became a W3C recommendation in 2004, so the efforts that pre-date that necessarily do not consider OWL although some of them consider other ontology languages.

On the other hand, classes may be regarded as a way for an abstract representation of grouped resources in accordance with specified common signs. An OWL class is associated with a set of individuals, named instances of classes. Therefore, a class has an intentional meaning, which is related but not equal to its class extension. In this aspect, two classes may have the same class extension, but still be different classes.

The core focus of the efforts in the data interpretation is laid on the semantic representation of the combinations of primary and secondary keys. Each mapping system must foresee any possible relationship that can be formed by an external key and the combinations of all possible components.

A good demonstration is the implicit semantics that are in SQL-DDL and how they can be made explicit. This paper presents an example of an university SQL schema and its corresponding ontology, thus demonstrating how a SQL schema and ontology of the same domain are related.

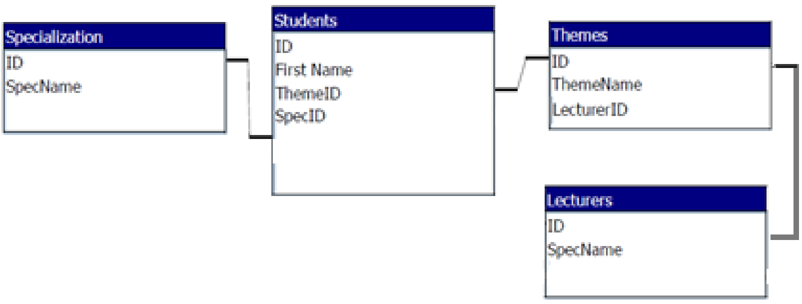

Fig.2. A relational database schema

According to the relationship within the relational database, it is clear that one lecturer can lead more than one student, while a student is assigned only one thesis to work on. The description of these relationships may be traced by an ontology construct.

This is a simple example for how the semantics presented in the relational schema can be translated to ontology. In result, some relationships within the relational schema may have explicit transformations of the ontology.

Every ontology is an implementation of an ontological model. This model includes constructs for specifying classes, properties, data types, inheritance, restrictions, and other semantics. But the ontology does not need to include all constructs of the ontological model (i.e. it can use only a portion of the ontological model). Domains are specified in classes and subclasses providing a hierarchical model presenting all the knowledge fields that are included in the relational model. There is also a number of properties denoting the relationship between classes [1].

In the presented ontology the relation Thesis – Student can be represented with the triplate Student 1 –>IsCreatorOf->Thesis 1. In this case the classes StudentName и StudentSpec remain as sub-classes of the class Student and can be regarded as its inheritors.

OWL distinguishes two main categories of properties that can be specified under o*bject properties* link individuals to individuals and *data type properties* link individuals to data values. On Figure 3 the properties are demonstrated with dashed arrows.

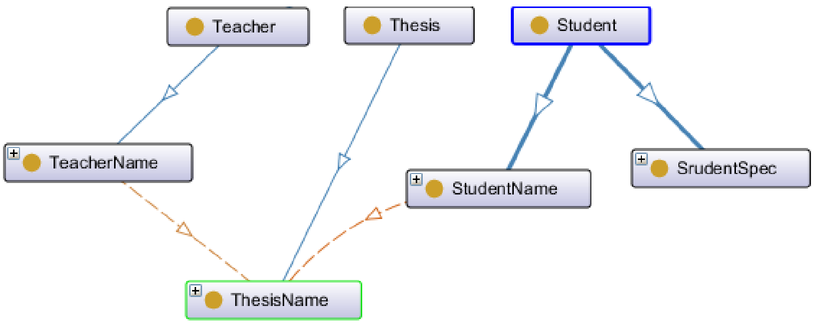

Fig.3. The ontology properties

The transformation of relational databases to ontologies is based on a set of rules called mapping rules that specify how to map constructs of the relational model to the ontological model. These rules are then applied to a relational database (source) to produce an ontology (target). Since the rules are specified on the model level, they are applicable to any relational database that conforms to the relational model.

The proposed approach is able to provide a transition system to produce a conceptual model from relational database based on graph theory which leads to product of database graph, and also with transforming of the graph obtained, final ontology is generated. The result of this transition system is conversion of SQL into the OWL ontology structure. With more components involved in the transition process, this method has succeeded to show richer semantics in the target ontology. With Using conceptual graph model the final output is obtained independent of the physical implementation of database and database management system.

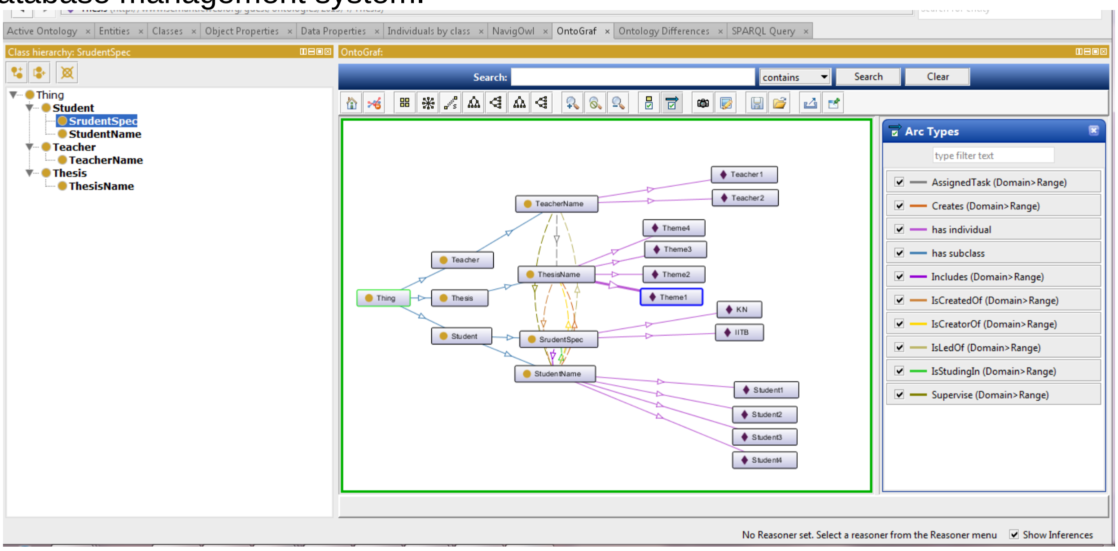

Fig.4. OntoGraph created with Protégé 5.0.0.

A general problem when creating an ontology is the lacking completeness. On the other hand, the creation of even relatively small ontologies, takes a quite high amount of time and resources.

Creation ontology process needs two main ingredients: On the one hand, it needs to generate some semantic data and on the other hand, this knowledge needs then to be connected to already existing ontologies.

Protégé 5 is a software used to create ontologies. The first step is to establish a connection to the data mining results and the ontology.The second step regards the need for mapping between different domain description ontologies. This mapping is necessary, since each domain ontology represents semantics of different knowledge domains. Given that layer functionalities differ significantly, similar differences may be present to the corresponding ontologies. Since integration needs require that information is seamlessly passed among the different layers, domain ontology mapping is absolutely necessary.

CONCLUSION

The most essential advantages of the Semantic Webs are the simplicity, unpretentiousness, the visual and clear representation. The specific drawbacks are linked with the inconvenient presentation of random n-relations, insufficient power of expression, dimmed semantics, impeded conduct of various operations, difficult inheritance control.

The trouble with ontologies is that they are complicated and hard to create, implement and maintain. Depending on their scope, they become enormous, defining a wide range of concepts and relationships. Some developers prefer to focus on logic and rules than on ontologies because of these difficulties. Furthermore, the disagreements regarding the roles the rules should play may be a potential pitfall for the Semantic Web.

REFERENCES

[1] Berners-Lee T., J. Hendler, O. Lassila, The Semantic Web, Scientific American, May 2001, p. 28—37.

[2] Tirmizi S. H., J. Sequeda, and D. P. Miranker. Translating SQL Applications to the Semantic Web. In DEXA, pages 450–464, 2008.

[3] W3C OWL Working Group. OWL 2 Web ontology language document overview. W3C Recommendation May 2015, http://www.w3.org/TR/owl2-overview/.

CONTACT ADDRESS

Pr. Assist. Desislava Baeva, PhD
Department of Informatics and Information Technologies
Faculty of Natural Sciences and Education
Angel Kanchev University of Ruse,
8 Studentska Str., 7017 Ruse, Bulgaria
Phone: (++359 82) 888 214
E-mail: dbaeva@ami.uni-ruse.bg

ПРЕОБРАЗУВАНЕ НА ДАННИ НА SQL ПРИЛОЖЕНИЕ В СЕМАНТИЧНА МРЕЖА

Десислава Баева

Русенски университет “Ангел Кънчев”

Резюме: *Семантичните уеб технологии все по-често се прилагат в широк спектър от приложения, където знанията са концептуализирани и формализирани, съгласно онтология, като средство за разширение и разнообразяване на процесите, управлявани от машини. Решаващ фактор за успеха на семантичните мрежи е разработването на методи за трансформиране на релационни бази данни в семантично достъпни данни. Тази статия описва онтология, която се основава на информация, получена от релационна база от данни.*

Ключови думи: *семантичен уеб, онтология*

Requirements and guidelines for the authors - "Proceedings of the Union of Scientists - Ruse" Book 5 Mathematics, Informatics and Physics

The Editorial Board accepts for publication annually both scientific, applied research and methodology papers, as well as announcements, reviews, information materials, adds. No honoraria are paid.
The paper scripts submitted to the Board should answer the following requirements:
1. Papers submitted in English are accepted. Their volume should not exceed 8 pages, formatted following the requirements, including reference, tables, figures and abstract.
2. The text should be computer generated (MS Word 2003 for Windows or higher versions) and printed in one copy, possibly on laser printer and on one side of the page. Together with the printed copy the author should submit a disk (or send an e-mail copy to: vkr@ami.uni-ruse.bg).
3. Compulsory requirements on formatting:

- font - Ariel 12;
- paper Size - A4;
- page Setup - Top: 20 mm, Bottom: 15 mm, Left: 20 mm, Right: 20mm;
- Format/Paragraph/Line spacing - Single;
- Format/Paragraph/Special: First Line, By: 1 cm;
- *Leave a blank line under Header - Font Size 14;*
- Title should be short, no abbreviations, no formulas or special symbols - Font Size 14, centered, Bold, All Caps;
- *One blank line - Font Size 14;*
- Name and surname of author(s) - Font Size: 12, centered, Bold;
- *One blank line - Font Size 12*;
- Name of place of work - Font Size: 12, centered;
- *One blank line;*
- abstract – no formulas - Font Size 10, Italic, 5-6 lines ;
- keywords - Font Size 10, Italic, 1-2 lines;
- *one blank line;*
- text - Font Size 12, Justify;
- references;
- contact address - three names of the author(s) scientific title and degree, place of work, telephone number, Email - in the language of the paper.

4. At the end of the paper the authors should write:

- The title of the paper;
- Name and surname of the author(s);
- abstract; keywords.

Note: The parts in item 4 should be in Bulgarian and have to be formatted as in the beginning of the paper.
5. All mathematical signs and other special symbols should be written clearly and legibly so as to avoid ambiguity when read. All formulas, cited in the text, should be numbered on the right.
6. Figures (black and white), made with some of the widespread software, should be integrated in the text.
7. Tables should have numbers and titles above them, centered right.
8. Reference sources cited in the text should be marked by a number in square brackets.
9. Only titles cited in the text should be included in the references, their numbers put in square brackets. The reference items should be arranged in alphabetical order, using the surname of the first author, and written following the standard. If the main text is in Bulgarian or Russian, the titles in Cyrillic come before those in Latin. If the main text is in English, the titles in Latin come before those in Cyrillic. The paper cited should have: for the first author – surname and first name initial; for the second and other authors – first name initial and surname; title of the paper; name of the publishing source; number of volume (in Arabic figures); year; first and last page number of the paper. For a book cited the following must be marked: author(s) – surname and initials, title, city, publishing house, year of publication.
10. **The author(s) and the reviewer, chosen by the Editorial Board, are responsible for the contents of the materials submitted**.

Important for readers, companies and organizations

1. Authors, who are not members of the Union of Scientists - Ruse, should pay for publishing of materials.
2. Advertising and information materials of group members of the Union of Scientists – Ruse are published free of charge.
3. Advertising and information materials of companies and organizations are charged on negotiable (current) prices.

Editorial Board

ISSN 1314-3077