UNIVERSITY OF SOFIA "St. KLIMENT OHRIDSKI"
FACULTY OF MATHEMATICS AND INFORMATICS

Abstract

University of Sofia "St.Kliment Ohridski"
Department of Information Technologies
Faculty of Mathematics and Informatics

Thesis title:

Web application for creation and edition of web pages with metadata for the Semantic Web.

Graduate student: : Rumyana Angelova, subject Distributed Systems and Mobile Technologies, faculty number M-21362

Supervisor: assoc.prof Sylvia Ilieva, PhD, Sofia University, Department of Information Technologies, Faculty of Mathematics and Informatics

Defense of master's thesis date: 14.02.2007

Key words: Semantic Web, web application, web editor, RDF, RDFa, OWL, Java EE 5.0, JSF, Session EJB 3.0, Java Persistence API, JAX-WS.

Annotation:

The main purpose of the Semantic Web is to apply semantics on the WWW content. Without semantics this content presents variety of web documents which are hyper-linked. These documents contain texts in natural languages (Bulgarian, English and so on), graphics, multimedia and additional metadata that is used for presentation of the content. While looking into the pages, the man is able to perceive easily different types of information (text, graphics) using his senses. He can make conclusion even from a partly information, because he associates it with facts which are already known. But the information in WWW becomes larger and larger and people are not able to look into all the data. Thus, it is highly necessary that the processes of searching, making conclusions and retrieving the information become automated.

The machines must understand the meaning of the whole information in WWW in order to make conclusions in real time and present the results quickly. Now the searching in WWW is implemented using text matching. The Semantic Web enhances searching using characteristics of the searched object. For instance when one searches for "the oldest university in Bulgaria", the result in the Semantic Web will be "Sofia University Sv. Kliment Ohridski" although this phrase has not been used in a web page.

The organisation W3C (World Wide Web Consortium) has been specifying the standards for the Semantic Web for several years so far. The main idea is that this Semantic Web is an extension of the current WWW, not something new and abstract. The protocols which describe the semantics of the data in WWW are RDF and OWL. SPARQL is a protocol for making queries in the semantic data described according to RDF and OWL. Now all these protocols have been finalized (RDF and OWL since 2004, SPARQL since 2006) and the current information in WWW has been being adapted – not only universities which are doing experiments, but also many commercial companies are consolidating the data according to the protocols of the Semantic Web.

The purpose of the master thesis is to support the process of making web pages with semantic content. This master thesis presents a web application,

which provides user-friendly interface. It hides the complexity of the RDF and OWL syntax from the user. He is not meant to be an expert in the area of Semantic Web but he will be able to create and edit web pages with text and graphics and their semantics.

The master thesis starts with presentation of the main ideas and notions in the Semantic Web. There are some examples of commercial and non-commercial applications for the Semantic Web. After that are described the technologies used for the implementation of the web editor (the technologies for the Semantic Web and the technologies in Java EE 5.0 for development of web applications). Next are described the requirements, architecture, design, coding and testing of the web application. At the end of the master thesis are proposed several directions for future development of the application.