ANALELE ŞTIINŢIFICE ALE UNIVERSITĂŢII „ALEXANDRU IOAN CUZA” DIN IAŞI
Număr special **Ştiinţe Economice** **2010**

ADVANTAGES OF SEMANTIC WEB TECHNOLOGIES IN THE KNOWLEDGE BASED SOCIETY

Simona Elena VARLAN
Alexandru Ioan Cuza University of Iasi, Faculty of Economics and Business Administration
Iasi, Romania
varlan_simona@yahoo.com

Abstract

As communication devices and the Internet become ubiquitous the world continues its progression toward a knowledge-based society in which what you know equals the degree of your success. Knowledge sharing breaks down now time and distance barriers and makes people work in virtual offices and communicate and interact with each other around the world allowing them to learn from one another and share the same information no matter the geographical location they are. The inclusion of the web means that what once was done by direct implication between humans now occur using technology and computers. Semantic Web technologies has developed very fast in the last past years and continues to grow as the importance of knowledge and technologies working together for human benefits becomes a necessary part in all the known domains: IT&C, economic, social, health and even political. The paper tries to show the importance of Semantic Web Technologies in a Knowledge Based Society providing some discussions on the technical background followed by examples of knowledge representation based languages.

Keywords: Semantic Web, RDF, OWL, XBRL, SKOS
JEL classification: O14, O31, O32, O33

1. INTRODUCTION

A major importance in research field is centered on the way of making computers to understand data and relations between data as well as peoples do. Making computers to realize the semantic of every data processed is the biggest challenge that researchers, in particular from artificial intelligence field, are trying to achieve. In the same way members from W3C consortium have look to accomplish an idea developed more then 10 year ago “which seeks to give computers the ability--the seeming intelligence--to understand content on the World Wide Web” [4]. In other words everything that means knowledge for humans one day can have the same meaning for a computer.

The big project was called Semantic Web and the W3C community’s vision was a Web of “meaning” where computers after processing online data can perform reasons and assist humans in resolving problems and take decisions. Much of it still needs to be fulfilled but the basic technologies are already developing and the opportunities are growing for all the domains involved: IT&C, economic, social and even political. In other words the Semantic Web is affecting all the sectors where technology and knowledge work together for human benefits [8].

The paper tries to show the importance of knowledge representation for an automated elaboration by any software application. In this meaner web documents can become machine-readable not only human-readable. We will try to identify the specific semantic web languages and some applications based on semantic languages as well as the future web trends.

2. KNOWLEDGE REPRESENTATION AND SEMANTIC WEB

Knowledge representation is about designing and model the real world into a machine-made form [5]. The idea is to analyze knowledge about the real world and then create a model/standard upon stabled rules and relation types to translate the human (natural) language in a machine and human readable language.

Because of tremendous grow of Internet the web was more important than ever with its new web applications and sites appearing with surprising regularity letting humans directly interacting and communicating with each other. The computer became only a communication environment without “understanding” the knowledge people shared about real life through web pages that could be easily accessed from anywhere. The result was an amount of knowledge stored in a chaotic and unstructured way [2]. The machines are “unschooled”, they don’t know what to do with all the data so most of the information remain unusable; they do not know how to distinguish an image from a video file or to make connections between data.

In conclusion we needed to classify and organize online data such as text, pictures, videos, or database entries in a system with logical connections between data representing the knowledge shared by people.

Knowledge is defined according the Oxford English Dictionary as expertise and skills acquired by a person through experience or education; what is known in a particular field or in total; facts and information or awareness or familiarity gained by experience of a fact or situation. So we can conclude that knowledge is a perception of reality, a description of the way things are and the relationships between them. The question that arises is how can we model reality to describe how things really are from the perspective of computer and web expansion? How can we create a medium for human expression, or in other words how can we represent knowledge so that computer can “understand” it? For that we need specialized web languages capable to describe concepts and concepts scheme.

The vision about Semantic Web or Web 3.0, as Sir Tim Berners-Lee states, is to become an extension of the current web, in which information has a well defined meaning and it is understood and processed by computers. John Markoff defined Web 3.0 as a set of technologies that offer efficient new ways to help computers organize and draw conclusions from online data.

We consider Semantic Web, in an abstract way, the alliance between knowledge representation and tools powerful enough to enable reasons about online date, and in strict way, alliance between semantic web languages and semantic web applications to share, analyze and process data.

We have specialized web languages as XML markup language and all the derived languages: RDF, OWL, XBRL, SKOS etc. Semantic Web was built on XML language capacity to define ordinary schemes much closer to data representation. But the language that did marked the beginning of a real web of data was the RDF language with its “triplets”: subject, property and object (represented by URI-s) to form a direct, labeled graph which “connects” data. The RDF specifications have become the general representation knowledge method used together with other syntax formats.

Every web language has its own specification and applicability as well as advantages and disadvantages.

3. FROM XML TO OWL

The Extensible Markup Language (XML) is a standard for creating markup languages and beeing extensible allows users to define their own elements in a structured way. XML and all the XML based languages have attributes like human-readable, machine-readable, object-oriented and hierarchical structured model, extensible, and easy to query and processed by applications. Because it is user defined language, XML documents become perfectly valid for human standpoint in any format but if we need to describe vocabularies or different sets of terms and relationships between them to be “understood” by computers, words with a fixed meaning, than we need specialized knowledge representation languages able to communicate the meaning of the data.

The first created was Resource Description Framework (RDF) which is a “simple metadata representation framework that uses URIs to identify Web-based resources and a graph model for describing relationships between resources”[1].

Because RDF does not offer more complex restraints regarding resources and their properties another semantic web standard – the Ontology was developed. Ontologies, in a general way, are a form of knowledge representation about the world, and in a particular way a formal representation of a set of concepts and the relationships between those concepts. The OWL Web Ontology Language is designed for use by applications that need to process the content of information instead of just presenting information to humans. OWL facilitates greater machine interpretability of Web content than that supported by XML, RDF by providing additional vocabulary along with a formal semantics [9]. At this moment OWL has three sublanguages: OWL Lite, OWL DL, and OWL Full.

An important work was conducted for developing so called Digital Libraries. The ideea was to create a shared catalogue which can be used by all libraries and search engines across the Web. This requires the use of common metadata to describe the field of the catalogue and common controlled vocabularies to allow subject identifiers to be assigned to publications [1]. To create controlled vocabularies like taxonomies, glossaries, classification schemes and thesauri, it was developed the Simple Knowledge Organisation System (SKOS) language, an extensible RDF language, to describe concept and content of concept schemes. In other words SKOS represent a set of formats and tools for describing controlled vocabularies and classifications schemes. Because SKOS Core is also an RDF vocabulary it can be used in combination with other semantic web vocabularies.

Next we will present an example of combination between SKOS and OWL language. In the example we defined two concepts: OperaMusic and PopMusic which are disjoint from each other and concept PopMusic has also defined three elements: R&B, Disco and Blues.

```
<rdf:RDF
  xmlns:rdf="http://www.w3.org/1999/02/22-rdf-syntax-ns#"
  xmlns:rdfs="http://www.w3.org/2000/01/rdf-schema#"
  xmlns:owl="http://www.w3.org/2002/07/owl#">
   <skos:Concept
            rdf:about=”http://www.myexample.com/OperaMusic#”>
      <skos:prefLabel> About opera music. </skos:prefLabel>
      <owl:disjointWith rdf:resource=” http://www.myexample.com/PopMusic# “/>
</skos:Concept>
<skos:Concept rdf:about=”http://www.myexample.com/PopMusic#”>
       <skos:prefLabel> About opera music.</skos:prefLabel>
       <owl:disjointWith rdf:resource=” http://www.myexample.com/OperaMusic# “/>
                <owl:oneOf rdf:parseType="Collection">
                              <owl:Thing rdf:about="#R&B"/>
                                             <rdfs:label>R&B</rdfs:label>
                              <owl:Thing rdf:about="#Disco"/>
                                             <rdfs:label>Disco</rdfs:label>
                              <owl:Thing rdf:about="#Blues"/>
                                             <rdfs:label>Blues</rdfs:label>
                </owl:oneOf>
            </skos:Concept>
</rdf:RDF>
```

4. CREATING A USER MANUAL STANDARD WITH SKOS

To better understand the utility of semantic web technologies we propose here a standard for user manuals made in SKOS. Everybody knows the importance of a user manual as a mechanism to present necessary information about software application. A user manual it is based on concepts with definitions and labels. To cover all the requirements and characteristics on how to represent a user manual we can use the SKOS vocabulary in combination with other semantic languages like OWL, RDF, FOAF. SKOS offers semantic relations between concepts like skos:narrower, skos:broader and skos:related. To represent the supplementary information about the document like general notes, source dates, historic dates and examples we have properties like skos:note, skos:editorialNote, skos:definition, skos:scopeNote, akos:changNote and skos:historyNote. Because SKOS language and user manuals are both concept orientated models choosing SKOS as a way of representation is an optimal solution. All these properties can of course be adapted to correspond with the requirements of a standard representation for user manuals. Feature versions of the SKOS language can bring new elements which can improve information representation for a user manual. For further studies we will implement a standard conceptual scheme to represent a user manual.

5. THE FUTURE WEB TRENDS

One of the primary considerations in an open environment like Internet where anyone and anywere can access and provide information is the credibility and reliability of available information. We can not have a total control. We also can presume that the reliability and credibility wanted could be provided by corporate status like government institutions. But even then information could be approximate or uncertain. It is well known that humans have a natural capacity to make a good discernment when they use web documents but computers don't. That means we need a computer-discernment to realize an automated validation of online data.

This is the future vision about a *web of trust*.

Of course there are some problems to resolve, like how to establish the credibility of data or how to determine the reliability of data. This is the most important part because we can not know the exact source of the information and also we can not control which information should be part or not of the Web. In order to have a little control of what is on the Web maybe a solution would be to create a system of validation, in the way the detecting e-mail spasms system work, where every web resource to have a signature given by an authorized party. But the inconvenience of that solution is that it costs a lot of time and effort.

Other vision about Semantic Web is a strong and bigger collaboration between researchers and business corporations for industry solutions and products which use semantic web technology to increase profits and reduce costs; a web with better defined semantic languages with an increase expressivity and a wide area of covered domains used everywhere in the simple possible way in different corporations by non-expert users without even realizing what they use.

We can not miss out the online video/TV trend that has already becoming a daily possible activity for anyone especially because of the success of YouTube. For this trend we believe that in few years TV will be available on mobiles through the Internet everywhere with higher quality pictures than now, and in a personalized way for every user after the country, age, culture or other properties that matter for each individual in part.

6. CONCLUSIONS

Right now the semantic web techniques cannot replace a human. He still must validate all the results that a computer generates. Still the human is the one to formal define concepts, things, and events, real live and presented in a machine-understandable form.

Event if the vision about the Web of trust can be still far way, we have to point out the important steps already achieved: RDF and OWL standards have been completed; many semantic web applications have bee developed in the last years making collaboration with corporations much stronger, and given the right benefits to semantic web technologies.
There is much to be fulfilled but the opportunities are big because of the incredible capacity humans have, called knowledge.

References

[1] Brian Matthews, *Semantic Web Technologies*, CCLRC Rutherford Appleton Laboratory

[2] Conrad Barski, *How to tell stuff to a computer*, see http://www.lisperati.com/tellstuff/index.html. Last accessed 22 April 2009

[3] Ivan Herman, *Tutorial on the Semantic Web*, W3C, 23 March 2009

[4] Borland John., *A Smarter Web*, Technology Review, Vol. 2 March/April 2007

[5] http://www.technologyreview.com/infotech/18395/. last accessed on 22 April 2009

[6] John F. Sowa, *Knowledge Representation: Logical, Philosophical, and Computational Foundations,* Pacific Grove, CA: Brooks/Cole, ISBN 0-534-94965-7, 2000 Reviewed by Stuart C. Shapiro Computational Linguistics June 2001, Vol. 27, No. 2: 286–294 at http://www.mitpressjournals.org/doi/pdf/10.1162/089120101750300544?cookieSet=1. Last accessed 22 April 200

[7] Lassila O., Hendler J., *Embracing "Web 3.0"*, IEEE Computer Society, 2007. See http://www.mindswap.org/papers/2007/90-93.pdf. Last accessed 22 April 2009

[8] Davis Mills., *The Business Value of Semantic Technologies*, TopQuadrant Special report 2004.

[9] Roberta Cuel, Alexandre Delteil, Vincent Louis, Carlo Rizzi, *Knowledge Web Technology Roadmap "The Technology Roadmap of the Semantic Web"* , see http://knowledgeweb.semanticweb.org/o2i/menu/KWTR-whitepaper-43-final.pdf. Last accessed 22 April 2009

[10] http://www.w3.org/TR/owl-features/. Last accessed 22 April 2009

[11] www.w3c.org. Last accessed 22 April 2009