Using Semantic Web technologies for knowledge management in e-commerce applications

Simona Elena Varlan
Department of Computer Science
University "Vasile Alecsandri" from Bacau
Calea Marasesti, no. 157, Bacau
Romania
varlan_simona@yahoo.com

Abstract: The scope of the paper is to present a model for knowledge management using Semantic Web technologies in e-commerce applications. Knowledge management refers to the representation, processing and sharing of knowledge. We will represent knowledge for an IT&C e-commerce web site using the OWL semantic web language for creating the domain ontology and inserting RDFa tags into HTML files conform to the ontology created, then knowledge is processed using SPARQL query language and shared within the web site.

Key-Words: Semantic Web, Knowledge Management, OWL, Ontology, RDFa, SPARQL

1 Introduction

Today, with the development and spread of information technologies, the challenge in economic and financial capital market is no longer owned by a technological society (access to information technologies of various kinds is easy depending on the financial possibility of acquisition) but the knowledge has become the tool basis for the survival of any company in the competitions of any kind. In practice the companies that managed to create, use, storage and disseminate knowledge have experienced success in business and improve performance[1]. Therefore new technologies able to realize all that have been developed recently and have seen increasing use. We talk about Semantic Web technologies, technologies developed for use in the Web because the Internet is the one who influenced the development of information technologies and thus contributed to the emergence of a new type of organization, the virtual organization. The article proposes a way of using Semantic Web technologies for knowledge management within an e-commerce organization. We will present how these technologies can be used by applying them in a specific e-commerce site dedicated to computer components. In the next chapter we will formulate the starting hypothesis of our research, we will refer to existing applications in this area and the motives why we choused the domain of IT&C e-commerce.

2 Problem Formulation

Semantic Web started to develop very fast in the last years, and its technologies to be used widely in web applications. W3C continue to develop and promote the technologies among web developers. Beneficiaries are virtual organizations whose main purpose is to develop ways of integration and automatic processing of knowledge, as it is recognized that knowledge is the main factor for progress and performance especially in virtual organizations. From these facts we began our research with two questions about what advantages can Semantic Web technologies bring to knowledge management within virtual organizations, and how we can effectively use the existing SW technologies for an efficient KM? Starting from these questions we formulate the hypothesis that knowledge management in the context of virtual organizations may be based on the Semantic Web technologies for semantic processing by software.

2.1 Why e-commerce

The choice for IT&C e-commerce as field study was done because the direction of online market in Romania is directed primarily to e-commerce. According to the study conducted by the association Link2eCommerce [2] on Romanian Internet users conducted between December 6, 2010 - January 15, 2011, shows that more than 5% of sites are devoted

to electronic commerce in Romania and that the number of online shoppers is growing, and the areas that generate the largest volumes in Romanian electronic commerce are, in order: travel services and reservations airline tickets, web sites dedicated to mobile phones and IT&C products. Number of sites for online shopping is also growing over the last three years around 1500 sites, but many of them can not resist the changing market and fail. For better visibility on the Internet, to be in the top search results, there are necessary SEO methods. The new trend in this direction is the use of RDFa tags in web pages and ontology that define these tags. Achieving this is the scope of our research.

2.2 Existing e-commerce applications with SW technologies

The big "names", the big companies that already started to use Semantic Web technologies are Google, Yahoo and Facebook. But the company that kept out attention is the leading U.S retailer, Best Buy, which is using the Semantic Web markup language RDFa to add semantics to its web sites. The Semantic Web is a Web of added meaning, which ultimately enables smarter and more personalized web apps to be built [3]. MacManus interviewed the lead web development engineer Jay Myers which explained the idea behind. They use semantic technologies to increase the visibility of their products and services like store name, address, opening hours, reviews, etc. Data being marked up with RDFa, search engines are able to identify each of those data components more easily and put them into context. Conform to Myers Best Buy employees entered information into the blogs every day using online forms that output RDFa based on the GoodRelations[4] ontology, a Semantic Web vocabulary for e-commerce that describes product, price and company data. The result of this process was that all of the RDFa data was very visible to humans via search engines, so the search traffic increased with 30%. It was a boon because the company is "very reliant on search engines" for product discovery and store locations[3].

Even the experiment at Best Buy represents a concrete example of how we can use Semantic Web technologies it doesn't show how to process the knowledge contain inside the application.

3 Problem Solution

The web application proposed for an e-commerce site dedicated to ICT products integrates the Semantic Web technology. The idea followed was to make a semantic web application in which

knowledge representation to be made using semantic web language OWL, and knowledge processing to be performed using ontology created and SPARQL query language.
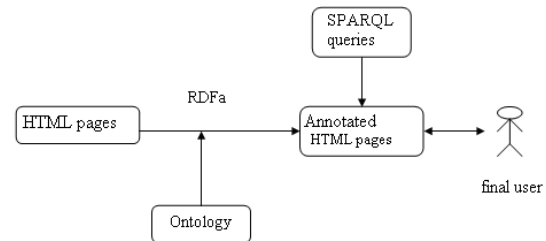General architectural scheme of the application is as follows:

Fig.1Architecture scheme

Form the figure above we can see that in the HTML page RDFa tags are inserted according to the ontology created for the specified domain. Annotated pages can then be queried using SPARQL query language. The interaction between end user and e-commerce site that contains annotated pages is in both ways, user can perform queries and view the search results.
We used the fallowing Semantic Web components:

- Knowledge representation languages for creating web ontology: OWL, RDF, etc.
- For annotating web pages according to the ontology created we used RDFa.
- To achieve the queries were used SPARQL endpoints and SPARQL query language.
- For the RDFa identification tags from the web pages was used RDFa extractor developed by the W3C.

3.1 Creating the Ontology

We choused to create our ontology and not to use an existing one (e.g. Good Relations) because the proposed web application is at this time a demo application, created with experimental purposes and as applicative part of the research. Because it is intended, for the moment, for experiments on Romanian market of e-commerce web sites, we create a specific ontology, in Romanian language according to the majority Romanian web sites on IT&C components. Were evaluated and studied several Romanian sites to determine the definition of concepts, classification and relationship between concepts. After the study we established the creation of six basic classes: Category, Product, Manufacturer, Country of production, Stock liquidations, Offers. Category is the only class that contains other subclasses according to the

classification of overall IT&C sites. Products class contains all the application's products and each product (class instance) have DatatypeProperty and is linked with other instances of different concepts through ObjectProperty. We defined a member of a class like this:

```
<owl:NamedIndividual
rdf:about="http://OntologyProduse.owl#
Acer_Aspire_AOD_255">
        <rdf:type
rdf:resource="http://OntologyProduse.owl#
Produse"/>
        <codProdus   rdf:datatype="&xsd;string">
12345AB
        </codProdus>
        <arePretValoare
rdf:datatype="&xsd;float"> 1600
        </arePretValoare>
        <laPromotie
rdf:resource="http:// OntologyProduse.owl#
Acer_Aspire_AOD_255"/>
        <areMarca
rdf:resource="http:// OntologyProduse.owl#
Acer_Company"/>
</owl:NamedIndividual>
```

The properties written in bold characters represent, in the respective order, the domain of the individual, the cod of the product, the price, offers and the mark of the product.

3.2 Using RDFa
RDFa tags can be used to describe products, to provide contact details of your online store, and other important knowledge for the site's visibility on the Internet. An example of how we can use RDFa using our ontology is presented below:

```
xmlns:pr=http://OntologyProduse.owl#

        <div                 property="foaf:name"
content="Acer_Aspire_AOD_255"
xml:lang="rom">
        </div>
        <div            property="pr:descriere"
content="descriere" xml:lang="rom">
        </div>
        <div                 rel="foaf:depiction"
resource="<img.jpg >">
        </div>
        <div             property="pr:arePretIN"
content="RON" datatype="xsd:string">
        </div>
```

```
        <div         property="pr:arePretValoare"
content="170.00" datatype="xsd:float">
        </div>

        <div                   rel="pr:areSite"
resource="laptop_uri/28054-hp_cq1000sl.html">
        </div>
        <div            property="pr :areMarca"
content="Acer" xml:lang="en">
        </div>
        <div property="pr:esteLaptop" ></div>
```

3.3 Semantic queries
We can semantically query the instances from ontology and also the RDFa tags declared in the HTML files.

To achieve the search over RDFa tags we need an RDFa extractor. This technology was developed by the W3C. It is actually an application where can be sent the URI of the web page where we will extract the tags. This application can also be used locally and the package can be downloaded without a license. URI returned by the RDFa extractor can then be used as a query request to a SPARQL endpoint together with SPARQL query format. The result is then sent backwards from the endpoint to the website in JSON format that stores data in a PHP array. This is then parsed and the results displayed in HTML format for viewing in the web page.

For this query we use an endpoint Virtuoso developed by OpenLink Software[5]. The software can be downloaded and create a local endpoint. SPARQL query format is sent to the SPARQL endpoint that returns results in JSON format, then transformed into PHP array. The array is parsed and the results displayed.

An example of a query made with SPARQL and the displayed results is presented below.

The SPARQL query where we search for the mark of a specified product is the following:

```
PREFIX pr: < http://OntologyProduse#>
SELECT ?marca
WHERE {
  ?entity foaf:name "Acer_Aspire_AOD_255" .
   ?entity pr:areMarca ?marca }
```

The figure 2 presents the displayed result of the query.

Fig.2: Displayed results

4 Conclusion

In the paper we proposed a web application where we integrated the Semantic Web technologies like ontology, RDFa and SPARQL query language for an efficient knowledge management. We demonstrate how to use these technologies for an efficient KM, for representing knowledge and processing it in the context of virtual organizations. It is efficient because of the advantages it implies: the ontology is more flexible and also offers, for example, transitive properties to model the knowledge base, than traditional database relational used in majority cases in web applications today; the RDFa tags, offers visibility in front of the search engines so users can find on the Internet the specific and exact information about the store, products and others thinks; SPARQL query it is used for semantic queries over the domain knowledge base.

The existing Romanian web applications doesn't use any of this technologies yet, so our proposed model can be seen as a demonstration on how all this can be implemented on a real market.

References:
[1] Griffith T.L., Sawyer J. E. And Neal M.A., *Virtualness and Knowledge in teams: Managing the love triangle of organizations, individuals, and information technology*, MIS Quarterly 27, no.2, 2003, pp. 265-287
[2] http://www.link2ec.ro – the official web site of Link2eCommerce Romanian association
[3] Richard MacManus, *How Best Buy is Using The Semantic Web*, july 1, 2010, http://www.wilshireconferences.com/semtech2010/RWW_070110.pdf
[4] http://www.heppnetz.de/projects/goodrelations/ - the official web site of the project developed by Prof. Martin Hepp
[5] http://www.openlinksw.com/index.htm - the official web site of Virtuoso Software