
Munich Personal RePEc Archive

Publishing Semantic Web Business Decision-Making Data

Necula, Sabina-Cristiana

Alexandru Ioan Cuza University of Iasi

June 2012

Online at https://mpra.ub.uni-muenchen.de/51599/
MPRA Paper No. 51599, posted 21 Nov 2013 06:01 UTC

PUBLISHING SEMANTIC WEB BUSINESS DECISION-MAKING DATA

Sabina-Cristiana Necula
Alexandru Ioan Cuza University of Iasi, Faculty of Economics and Business Adinistration, Department of Research
Iasi, Romania
sabina.mihalache@gmail.com

Abstract

This paper treats with examples the practical use of Semantic Web technologies for describing business data. The main problems that the present paper addresses are related to identifying the main activities/ functions of an organization that can be represented with Semantic Web technologies. We treat in the beginning of the paper the main functions of an organization that can be addressed with Semantic Web technologies. We continue by presenting the main technological aspects involved in developing Semantic Web applications in order to publish semantic web data. We present some examples and then we discuss our main ideas. The paper ends with some conclusions.

Keywords
RDF, SPARQL, semantic web, decision-making

1. INTRODUCTION

Semantic Web represents a set of technologies useful to mark-up web sources in order to retrieve information from multiple web sites in a single Semantic Web Search Browser. There are multiple case studies and multiple applications in different domains: biology & medicine, E-Government, education & research, geography, business etc. A set of case studies can be found at W3C Semantic Web Use Cases and Case Studies web site.
The most common standards which are called technologies for Semantic Web are Resource Description Format (RDF) for describing web sources as subject-predicate-object and SPARQL Protocol and RDF Query Language (SPARQL) [5], [6].
Although the business domain doesn't have many implementations yet we try in this paper to present some useful applications. In fact everything is a business but we must think at the functions of an organization in order to discuss potential Semantic Web applications [3], [4].
Starting from these functions we will give examples in order for the reader to understand better that in the knowledge society one organization must take into considerations the tremendous use of Semantic Web technologies.

The scope of this paper is to realize a demarcation between information and knowledge from the organization's point of view in what concerns its functions possible to manipulate with Semantic Web technologies.
This paper is organized as follows: in the next section we will discuss organization's functions possible to manipulate with Semantic Web technologies, in the following section we will discuss the main technological aspects involved. Section 4 presents examples and Section 5 discusses the main implications. The paper ends with conclusions and remarks.

2. ORGANIZATION'S FUNCTIONS

When someone thinks about the uses of one technology or another the main problem that arrives is: where and how I can make use of this technology? When we think to an organization we must think at its main functions: production, commercial, decision-making, human resources, accounting and financing.

We must make an inner join with semantic Web technologies. As Semantic Web technologies are of practical use for web sources we must recognize that for the production function Semantic Web has little applicability. Maybe in data integration or in describing information that is the output from the production process.

But the commercial function has many applications. An organization has the opportunity to describe its web sources concerning products and services by making use of a public ontology in order to be easily retrieved in the web space.

The decision-making function has many applications because the main important phase from the decision-making process is the information phase and the decision-makers need information from multiple sources: the internal process of production, information from the commercial function, from the external medium of organization and from the accounting and financing functions. We will focus in this paper on this function.

The human resources function has little applicability of Semantic Web technologies. As in the case of production function we must think only to the phase of describing useful information for decision-making processes.

The accounting and financing function has little applications of Semantic Web technologies if one thinks to an organization that is producing goods or is offering services. As in the above cases Semantic Web technologies are useful only to describe information that is needed in the decision-making processes. But, if one thinks to a financial organization or to the service of Revenue Taxes or to Exchange Commission or to another financial entity the benefits of Semantic Web technologies are tremendous because its main application for data integration.

Another important distinction to be made is the one between information and knowledge. It is important to make this distinction in order to one know what are the proper sources necessary to be described with Semantic Web technologies.

Every web source is in fact an html file that contains a lot of data that is not described. Semantic Web came with the solution of describing data from an html file according to a common vocabulary/ ontology available at a web address. In order for those data to be described the owner of the web site must do this mark-up.

Information represents any data that has significance for a receptor either human being or computer application. For a Semantic Web application one must have data described with Resource Description Format (RDF). Usually the significance is in decision-making process. The problem-solving situations or the decision-making processes make use of information and knowledge.

Knowledge represents the know-how, knowing the alternatives of making a decision or solving a problem. It is usually expressed with If-THEN-ELSE rules if we think to a computer-based application. Very often knowledge belongs to decision-makers, is a tacit one and the costs of formalizing knowledge might be too high.

The Semantic Web technologies are useful for data describing. They are not dedicated to formalize knowledge. The only form of represented knowledge is the one of ontology which means concepts' classification from one domain.

Very often we see slogans like: intelligent semantic web or semantic web is for representing knowledge. We wanted to make these demarcations in order to give a better representation to the uses of semantic web technologies. We discuss in the next section the main technological aspects involved.

3. THE MAIN TECHNOLOGICAL ASPECTS INVOLVED

In order to publish RDF data an web site owner must decide first what is publishable data.

The RDF Book Mashup provided an early example of publishing Linked Data related to retail and commerce. The Book Mashup uses the Simple Commerce Vocabulary to represent and republish data about book offers retrieved from the Amazon.com and Google BaseWeb APIs [2]

More recently, the GoodRelations ontology has provided a richer ontology for describing many aspects of e-commerce, such as businesses, products and services, offerings, opening hours, and prices. GoodRelations has seen significant uptake from retailers such as Best Buy and Overstock.com seeking to increase their visibility in search engines such as Yahoo! and Google, that recognize data published in RDFa using certain vocabularies and use this data to enhance search results.

If we think at what technological aspects are involved in developing a Linked Data Mashup we must remind RDF, SPARQL, RDF triple store, RDFa, and SPARQL end point.

The Resource Description Framework (RDF) is a family of World Wide Web Consortium (W3C) specifications originally designed as a metadata data model. It has come to be used as a general method for conceptual description or modeling of information that is implemented in web resources, using a variety of syntax formats [5].

SPARQL is an RDF query language, that is, a query language for databases, able to retrieve and manipulate data stored in Resource Description Framework format. It was made a standard by the RDF Data Access Working Group (DAWG) of the World Wide Web Consortium, and considered as one of the key technologies of semantic web [6].

"Triple Store" is the common name given to a database management system for RDF Data. These systems provide data management and data access via APIs and query languages to RDF Data. A triplestore is a purpose-built database for the storage and retrieval of triples, a triple being a data entity composed of subject-predicate-object [7], [8]

A SPARQL endpoint is a conformant SPARQL protocol service as defined in the SPROT specification. A SPARQL endpoint enables users (human or other) to query a knowledge base via the SPARQL language [9].

4 EXAMPLES

This section treats examples from business decision-making activity where Semantic Web technologies can be used with success.

Data sets stored in relational databases can be published through the use of relational database to RDF wrappers. These tools allow the data publisher to define mappings from relational database structures to RDF graphs that are served up on the Web.

Static input data may consist of CSV files, Excel spreadsheets, XML files or database dumps. In order to serve them as Linked Data on the Web, they must undergo a conversion process that outputs static RDF files or loads converted data directly into an RDF store.

Example of an RDF serialization with n3 is presented in Figure 1.

```
@prefix rdf: <http://www.w3.org/1999/02/22-rdf-syntax-ns#> .
@prefix rdfs: <http://www.w3.org/2000/01/rdf-schema#> .
@prefix dcterms: <http://purl.org/dc/terms/> .
@prefix foaf: <http://xmlns.com/foaf/0.1/> .
@prefix onto: <http://www.knowledgedecisionmaking.ro/vocab/onto#> .

<http://www.knowledgedecisionmaking.ro/onto.rdf#company>
rdf:type onto:Company ;
foaf:name "Alfa Company" ;
rdfs:label "Alfa Company SRL" ;
dcterms:description "Alfa Company SRL is a manufacturing company based near Iasi,
Romania recognized for its innovativeness" ;
foaf:based_near <http://sws.geonames.org/3333125/> ;
onto:hasTeam <http://www.knowledgedecisionmaking.ro/echipa> .

<http://www.knowledgedecisionmaking.ro/echipa>
rdf:type onto:Team ;
rdfs:label "The Alfa Company Team" ;
onto:isTeamOf <http://www.knowledgedecisionmaking.ro/onto.rdf#company> .
```

Fig. 1. RDF n3 serialization

This static profile document will be published at http://www.knowledgedecisionmaking.ro/onto.rdf. The fragment identifier #company is added to the URI of the document to give a URI for the company of http:// www.knowledgedecisionmaking.ro/onto.rdf#company. While this static file references URIs in the Alfa Company namespace that use the 303 URI pattern, the URI http:// www.knowledgedecisionmaking.ro/onto.rdf#company is minted within the context of the document http:// www.knowledgedecisionmaking.ro/onto.rdf and therefore must use the hash URI pattern. Once it has been created, this static RDF/XML file can be uploaded to the Alfa Company server using FTP or any other method.

Once we have data described in RDF format we can query it.

```
prefix onto: <http://www.knowledgedecisionmaking.ro/vocab/onto#>
prefix dcterms: <http://purl.org/dc/terms/>
prefix foaf: <http://xmlns.com/foaf/0.1/>
select ?s ?isTeamOf ?Team
where
{ ?s onto:isTeamOf ?Team .
}
```

Fig. 2. SPARQL query

5 DISCUSSIONS

We started our paper by discussing the main benefits for organizations. We want to draw attention to the following aspects:

- In order to benefit from this technology one must use it. For the moment Semantic Web technologies are not so well known at least not at the technical level.
- In order to benefit to maximum from this technology it must that Google or Yahoo! index web sites according to published ontologies.
- In order to be large scale application Semantic Web developers must wait until open and easy to use tools become available.

The main implications will be in the field of data visualization and data integration.

6 CONCLUSIONS

Our contributions in this paper are from the experience site. We want to test the possible uses for organizations, to identify use cases and to understand better this area of research. From the research point of view we must say that the main problems to be discussed are: aligning different ontologies, ontology matching, distributed data management, data confidence and security.

From the methodological point of view we must say that developing a Linked Data Application must follow Linked Data principles and, in the essence, the methodology means: describing data, publish data, and consume data.

From the experiment point of view we must say that we still need examples of use cases.

ACKNOWLEDGMENTS

This work was supported by CNCSIS-UEFISCSU, project number PN II-RU code 188/2010.

References

[1] Allemang, D., Hendler, J.: Semantic Web for the working ontologist: modeling in RDF, RDFS and OWL. Burlington, MA: Morgan Kaufmann, (2011).

[2] Heath, T., Bizer, C.: *Linked Data: Evolving the Web into a Global Data Space* (1st edition). Synthesis Lectures on the Semantic Web: Theory and Technology, 1:1, 1-136. Morgan & Claypool, (2011).

[3] Necula S.C.: Financial Data Sets Querying by Using Semantic Web Technologies, ISBN-13: 978-1-61275-994-4, 2011, Management Science and Engineering, Editors: Mark Zhou, în: Advances in Artificial Intelligence, Vol. 1, pp.598-603, http://www.ier-institute.org/2160-147X/978-1-61275-994-4.htm, (2011)

[4] Necula, S.C.: Using Resource Description Format in Integrating and Querying Web Data Sources for Semantic Portals, Book Chapter, Editor: Elena García-Barriocanal, Editors: Zeynel Cebeci, Mehmet C.Okur, Aydın Öztürk, Metadata and Semantic Research, in Communications in Computer and Information Science, 2011, Springer Berlin Heidelberg, ISBN: 978-3-642-24731-6, pp. 194-200, Vol. 240, (2011)

[5] Rapoza, J.: SPARQL Will Make the Web Shine. eWeek (2006)
http://www.eweek.com/article2/0,1759,1965980,00.asp

[6] Segaran, T., Evans, C., Taylor, J. Programming the Semantic Web. O'Reilly Media, Inc., 1005 Gravenstein Highway North, Sebastopol, CA 95472, (2009).

[7] Virtuoso FAQ http://virtuoso.openlinksw.com/dataspace/dav/wiki/Main/VOSRDFFAQ

[8] Semantic Web community http://semanticweb.org/wiki/SPARQL_endpoint

[9] W3C Semantic Web Standards Timeline http://www.dblab.ntua.gr/~bikakis/XML%20and%20Semantic%20Web%20W3C%20Standards%20Timeline-History.pdf