INT J COMPUT COMMUN, ISSN 1841-9836
7(5):907-915, December, 2012.

Implementing the Main Functionalities Required by Semantic Search in Decision-Support Systems

S.C. Necula

Sabina-Cristiana Necula
"Alexandru Ioan Cuza" University of Iasi
Romania, 700505 Iasi, 22 Carol I Blvd
E-mail: sabina.mihalache@gmail.com

Abstract:
This paper exploits different semantic web technologies and builds a prototype of semantic web mashup functionality based on an architecture proposed by us. The main scope is to improve decision-making processes. In this paper, we are focusing on querying different ontologies in order to improve decision-making semantic search. As a conclusion, we demonstrate that in order to improve decision-making semantic search there is a need of special constructions that a query language must support.
Keywords: semantic web, decision-making semantic search, ontology, SPARQL, SQWRL, RDF, OWL.

1 Introduction

The main problem treated by the present article is: improving Decision Support Systems (DSS) means improving search which leads to the semantic interoperability problem.

There is a main problem and, in the same time, a controversy here. When discussing enterprise decision support systems data exists, data comes from different sources (internal and external to enterprise) but they are differently described. There is an open world (the Web) and a closed world (the enterprise). If in the open world we discuss search engines and queries made by Internet users on existent data from www space, in the closed world we discuss databases and answers to queries that are priory represented. In the first case the answers to queries seems to be poorly represented in the aspect of semantics, in the second case the answers seems to be perfectly represented but not relevant. In this closed world of enterprise, usually decision-makers need answers to queries formulated by them and these queries are ad-hoc, very often cannot be anticipated. Their answers need a mixture of information that comes from the both worlds.

Even if all computer-based solutions are well intended the practical use shows that user satisfaction is low due to many reasons analyzed by studies that have been realized. There are many identified reasons: poor maintainability, poor flexibility and less reusability (see [22]). Much of the implicit information remains undiscovered thereby resulting in sub-optimal business decisions.

Web-based technologies are having a major impact on design, development, and implementation processes for all types of decision support systems (see [1]).

Therefore it is our goal to solve this problem by this article.

We consider in this study two main important reasons: (1) the possibility to apply knowledge in the decision-making moment at the decision-making place by the decision-maker; (2) semantic integration of information.

As a solution to the problem identified by us we propose practical implementations of semantic web search functionalities with existent Decision Support Systems. We present scenarios, technical requirements, and the architecture of a semantic web based application in decision support systems.

The scope of this paper is to present a method to improve search at the decision-making level. The main idea consists in using ontologies and semantic search technologies. The motivation is given by lack of interoperability and semantic consistency of different formats for the same content.

2 Related Work

In order to discuss differences of our approach we present the related achieved work in the field of adopting semantic web standards and the related achieved work in the field of developing applications in the context of decision-making.

Decision-making has been considered as one of the main concern in practice as in theory (see [19], [3], [13]and [10]).

A reference paper in the field belongs to H.K. Bhargava, D.J. Power, and D. Sun (see [2]) which presents the technological challenges that web technologies create for Decision Support Systems. They show that DSS generally require repeated interactions with a model, while the basic Web architectural model was designed for random jumps in hyperspace (see [2]). Integration is a major challenge in the context of Web-based decision computation.

Most of all previous studies focus on success factors of information systems (see [4], [5], [12]and [17]).

Describing big data sets according to schema(s) and accessing data by overlapping schemas is the big problem that Semantic Web is trying to address. Its main applications are in the field of managing the currently linked data, learning how to extract information from currently linked data (Linked Data Browsers), sense-making of events (there is a lot of life data streams like tweets) in order to provide a solution able to gather, collect and analyze in real time a large number of live data streams (e.g. twits), to extract the information contained and to map any reference to both a) geographical location/point of interest, etc. and b) domain specific facts (e.g. music events or violence/ demonstration). The goal is to identify events happening in a specific area (e.g. a specific city) in a short time (e.g. some hours). Sense making of the events is provided via mapping events over location and time. There is a lot of use for social purposes like emergency operators, governmental bodies. We must say that Semantic Web is intended to be used by people in the way which web intended to be at its origins.

Semantic web technologies are by far most often used for data integration and for improving the search (see [9]). The use of ontologies for knowledge sharing, heterogeneous database integration, and semantic interoperability has been long realized (see [8], [11]and [18]).

Metadata (data about data)/ ontology promises to overcome the problem of interoperability. An accepted standard for representing metadata should provide information about the syntax, the structure and the semantic context of data. Many standards have been proposed. For the moment the World Wide Web Consortium recommends RDF (S) (Resource Description Framework - Schema), and OWL (Ontology Web Language) for representing linked web data.

A number of query languages have been developed to query RDF and OWL. SPARQL (SPARQL Protocol and RDF Query Language) is currently the de facto standard RDF query language. Since OWL can be serialized as RDF, SPARQL can be used to query it. There is thus a need for an expressive OWL query language that supports comprehensive querying of OWL (see [14]). These authors proposed SQWRL (Semantic Query-enhanced Web Rule Language) built on SWRL (Semantic Web Rule Language).

A mashup is a Web page or application that uses and combines data, presentation or functionality from two or more sources to create new services. Mashup techniques retrieve content from several sources to create a new service or application.

3 Problem Statement

The limits of information integration means for decisions modeling the following issues: (1) input data in the decision-making model is changing. The set of models uses to change over time due to changing conditions and priorities. Thus, it becomes necessary that the conceptual structure can change as the data impose. A new proposed solution is to use ontologies; (2) at the same time, data come from multiple data sources (internal and external to organization) (see [6]and [7]). Thus, it becomes necessary to ensure semantic interoperability.

Two expensive stages appear in the management of decision-making models: 1) each data source has to have a schema and 2) there is a need to overlap different schemas.

Our research tries first to identify the main factors that are important in Decision Support usefulness through an empirical study and then to discuss how to overcome the limitations by applying Semantic Web to Decision Support Systems.

In order to clarify what we are talking about, we need to discuss what we mean when treating the problem of decision-making semantic search: (1) we don't discuss methods for representing knowledge or data structures or search algorithms in order to prove that one method or another technique is better and improves the semantics of content. Instead we the problem of decision-making semantic search as a search need to provide answers to queries formulated by enterprises decision-makers from two main worlds: the web and the internal world of enterprise; (2) we present the problem of decision-making semantic search in the light of the main technologies involved, the architecture and the main resources involved. We demonstrate it by presenting practical examples. These examples prove that by using semantic web technologies search results improves.

4 Our Approach

We propose the research model shown in Figure 1 by considering the characteristics of Decision Support Systems and also by referring to DeLone and McLeans Information Systems success model (see [4]).

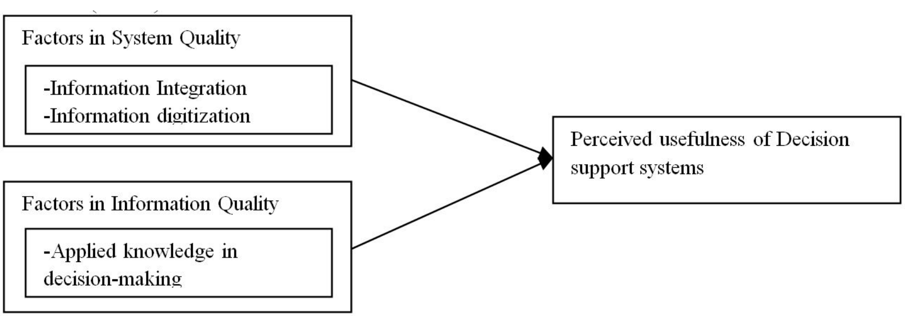


Figure 1: The research model

We developed a number of items to measure each construct from the research model by making references to previous work in the research field. We present a survey and the analysis realized on the results obtained by conducting the survey.

Table 1 provides the definition of the factors listed in Fig. 1, and indicates the number of items used to measure each construct. We developed the measures by referring to the following previous work: 1) system quality and information quality of the IS success model (see 4, 5); 2) application of the Semantic Web to KM and (3) the possibilities to apply knowledge by the decision-makers (see [15], [16]and [21]).

Factors		Definition	Number of items
Factors in system quality	Information integration Information digitization	Systems interoperability and semantic integration Degree of electronic-based information used in decision-making processes	1 2
Factors in information quality	Applied knowledge in decision-making processes	Functionalities required for a DSS in order to assist the process of applying knowledge in decision-making processes	12
User satisfaction		Degree of overall satisfaction with system use	1

Table 1: Definition of Decision Support Systems success factors

We derived two main hypotheses from the research model.

H1: applying knowledge in the decision-making moment has a positive impact on user satisfaction with the DSS H2: information integration has a positive impact on user satisfaction with DSS.

We interviewed 18 decision-makers from local public authorities that adopted and used a Decision Support System for three years to pretest the survey questionnaire. The primary goal of the pretest was to check content validity of the questionnaire. We revised a few question items. We then addressed the questionnaire to 34 decision-makers that own a business and that have invested in IS solutions. A total of 30 responses were used for statistical analysis (the software package SPSS version 17.0).

We used CrossTab correlation (Table 2) to see if is any direct relation between the degree in which it is considered that information needed in decision-making process comes from multiple sources and the degree in which the user is satisfied with DSS use (direct relation the main majority of the subjects responded affirmatively).

			q8 (DSS usefulness)			
			70%-100%	40%-69%	Under 40%	Total
q2	70%-a00%	Count	6	15	0	21
		% within q2 (multiple sources)	28.6%	71.4%	.0%	100.0%
		% within q8	100.0%	68.2%	.0%	70.0%
	25%-69%	Count	0	7	2	9
			.0%	77.8%	22.2%	100.0%
			.0%	31.8%	100.0%	30.0%
	Total	Count	6	22	2	30
		% within q2	20.0%	73.3%	6.7%	100.0%
		% within q8	100.0%	100.0%	100.0%	100.0%

	Value	df	Asymp. Sig. (2- sided)
Pearson Chi-Square	7.273[a]	2	.026
Likelihood Ratio	9.130	2	.010
Linear-by-Linear Association	6.313	1	.012
N of Valid Cases	30		

Table 2: CrossTab correlation and Chi-Square Tests

Semantic web applications can be classified into two categories: generic applications (Linked Data browsers and Linked Data Search Engines) and domain-specific applications (ESW LOD wiki; Semantic Web Use Case Studies and Use Cases).

A mashup should be developed following three steps:

1) Discover data sources that provide data by following RDF links from an initial seed URI into other data sources.

2) Download data from the discovered data sources and store the data together with provenance meta-information in a local RDF store.

3) Retrieve information from the local store using SPARQL query language.

While there are not many implementations in the field of decision-making semantic search we benefit from the existing grounding technologies like ontology, RDF/OWL descriptions, SPARQL language. Our solution consists on all of the above technologies and the proposed architecture is presented in Figure 2.

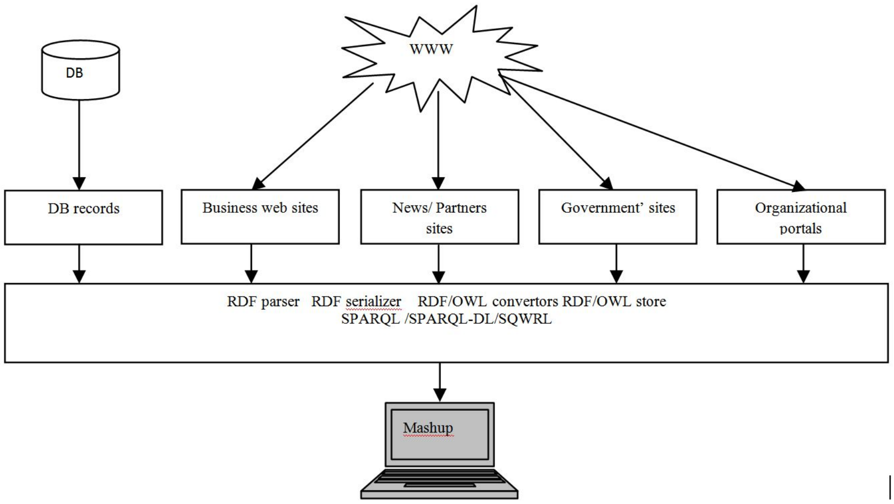


Figure 2: Semantic Web applications architecture

Our work is built upon five specifications: RDF, RDFS, OWL, SPARQL, and SWRL (SQWRL).

In Figure 3 we present a schema for describing Companies and an example of data described by respecting the defined schema/ ontology.

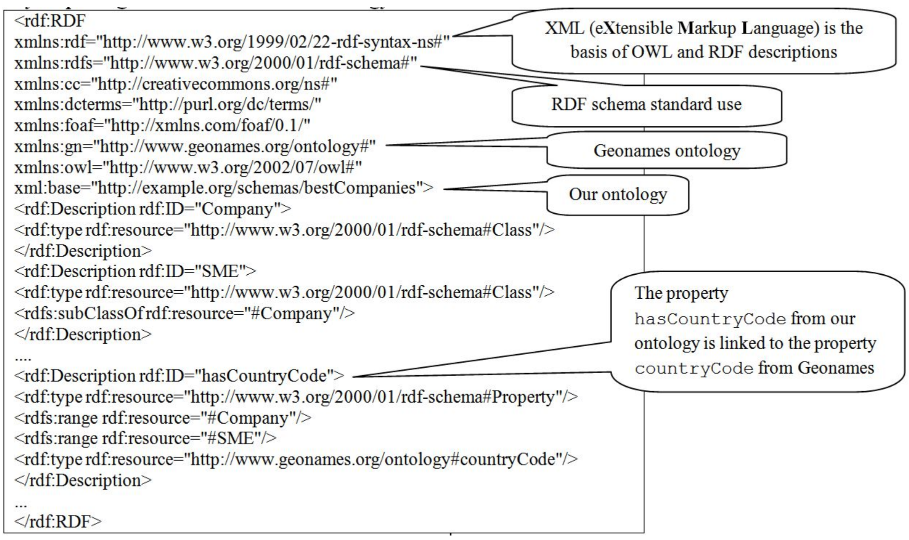
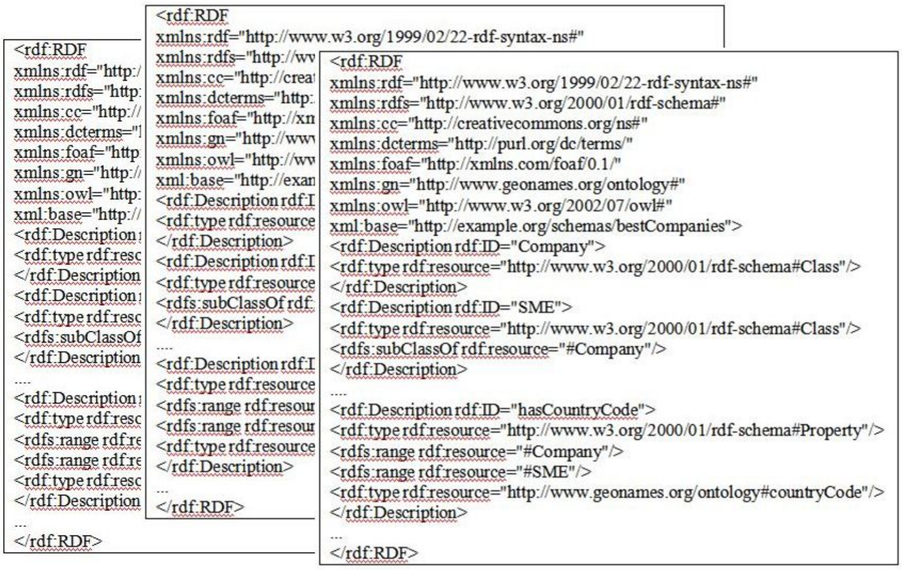


Figure 3: Companies RDF schema/ ontology

Web Data that has been cached locally is usually either accessed via SPARQL queries or via an RDF API. Using these specifications we developed a prototype Web based application to semantically improve a mashup that advises decision-makers of an organization. The Web-based application allows decision-makers to receive advice concerning their compared assets value to the ones of competitors. We present an example of different schemas/ ontologies needed to represent data stored in the RDF store (Figure 4).



```
<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:cc="http://creativecommons.org/ns#"
xmlns:dcterms="http://purl.org/dc/terms/"
xmlns:foaf="http://xmlns.com/foaf/0.1/"
xmlns:gn="http://www.geonames.org/ontology#"
xmlns:owl="http://www.w3.org/2002/07/owl#"
xml:base="http://example.org/schemas/bestCompanies">
<rdf:Description rdf:ID="Company">
<rdf:type rdf:resource="http://www.w3.org/2000/01/rdf-schema#Class"/>
</rdf:Description>
<rdf:Description rdf:ID="SME">
<rdf:type rdf:resource="http://www.w3.org/2000/01/rdf-schema#Class"/>
<rdfs:subClassOf rdf:resource="#Company"/>
</rdf:Description>
....
<rdf:Description rdf:ID="hasCountryCode">
<rdf:type rdf:resource="http://www.w3.org/2000/01/rdf-schema#Property"/>
<rdfs:range rdf:resource="#Company"/>
<rdfs:range rdf:resource="#SME"/>
<rdf:type rdf:resource="http://www.geonames.org/ontology#countryCode"/>
</rdf:Description>
...
</rdf:RDF>
```

Figure 4: Different schemas and data stored in the RDF store

SPARQL makes it possible to send queries and receive results, e.g., through Hypertext Transfer Protocol (HTTP) or Service Oriented Architecture Protocol (SOAP).

Using SPARQL consumers of the Web of Data can extract possibly complex information (i.e., existing resource references and their relationships). If we want to query what are the uri and the StockPrice of highest SockPriced companies we will write a query that look like query depicted in Figure 5.

```
PREFIX foaf: <http://xmlns.com/foaf/0.1/>
PREFIX rdfs: http://www.w3.org/2000/01/rdf-schema#
prefix dc:      <http://purl.org/dc/elements/1.1/>
prefix vcard:     <http://www.w3.org/2001/vcard-rdf/3.0#>
prefix :       <http://example.org/company/>
prefix ns:      <http://sandbox.metadataregistry.org/uri/schema/fin>
SELECT ?company ?StockPrice
WHERE {
        ?company ns:StockPrice ?StockPrice.
}
ORDER BY DESC(?StockPrice)
LIMIT 10
```

Figure 5: A SPARQL query that returns uri and StockPrice of the ten companies that have the highest Stock Price

SPARQL is particularly adequate for extracting data from ontology and, through its CONSTRUCT statement, for generating new data.

But when it comes to aggregated functions there is no aggregated function implemented by SPARQL yet. So, because of an evident need in the decision-making processes for aggregated functions we were making use of SQWRL a powerful tool to express queries with aggregated functions. Figure 6 presents a query with aggregated function.

When executing SQWRL query, first SWRL inference rules are taken into account/executed. Running SQWRL requires a rule engine, currently Jess. We were able to derive answers for

```
Clients(?c) ∧ hasBills(?c, ?b) ∧ hasProducts(?b, ?p) ∧ hasTotalValue(?b, ?v) ˚
sqwrl:makeSet(?s, ?v) ∧ sqwrl:groupBy(?s, ?p) ˚
sqwrl:avg(?avg, ?s) →
sqwrl:select(?p, ?avg)
```

Figure 6: SQWRL query that returns the average value for each sold product

questions like: what is the average value for each sold product. In order to obtain an answer we used SQWRL, because SPARQL doesnt have the necessary specifications.

5 Conclusions and Future Works

We have two contributions in this paper: (1) We propose that a combination of RDF/OWL with SQWRL to add semantic markup is useful for DSS; (2) We show how to combine them technically.

The need today is for a distributed evolution of ontologies. The overall problem for ontology engineering is that the number of ontologies which is available is currently very limited, and it is hard to validate the approaches using real ontologies.

Open issues that remains to discuss are: (1) consistency in order to meet the requirements of future real-life applications; (2) evolution of ontologies and metadata; (3) method and tools that scale up to handle a large number of networked ontologies and related metadata.

The web is effective at bringing any resource to the web user, but if the information the user needs is not represented in a single place, the job of integration belongs to the user.

How much from the intended message could be provided by using semantic web technologies? Pretty much how much it is intended to be represented. The power of represented linked data should be in discovering relations in existing represented data. The whole idea of linked data consists in consuming and publishing data. There is no intent in representing a standard/ de facto ontology that is the best in the modeled domain, the whole intent is to link our data with others data.

The main problem remains that of scrapping data from the Web. In order to not scrap for data, every web source provider should have his/her data represented in a standard semantic web format. In this way, semantic web application could gather data semantically described and sharing it to the user.

For the future work, (1) currently we are building software to improve mashup; (2) We plan to improve the existing ontology mediation.

Acknowledgment

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

Bibliography

[1] Bharati P., Chaudhury A., An empirical investigation of decision-making satisfaction in web-based decision support systems, *DECISION SUPPORT SYSTEMS*, 37: 187-197, 2004.

[2] Bhargava H.K., Power D.J., Sun D., Progress in Web-based decision support technologies, *DECISION SUPPORT SYSTEMS*, 43: 1083-1095, 2007.

[3] Chu P.C., Spires E.E., The joint effects of effort and quality on decision strategy choice with computerized decision aids, *DECISION SCIENCES*, 31 (2): 259-292, 2000.

[4] Delone W.H., McLean E.R., Information system success: The quest for the dependent variable, *INFORMATION SYSTEM RESEARCH*, 3(1): 60-95, 1992.

[5] Delone W.H., McLean E.R., The DeLone and McLean model of information systems success: A ten-year update, *JOURNAL OF MANAGEMENT INFORMATION SYSTEMS*, 19(4): 9-30, 2003.

[6] Eppler M., Mengis J., The concept of information overload: A review of literature from organization science, accounting, marketing, MIS, and related disciplines, *THE INFORMATION SOCIETY*, 20(5): 325-344, 2004.

[7] Farhoomand A.F., Drury D.H., Managerial information overload, *COMMUNICATION OF THE ACM*, 45(10): 127-131, 2002.

[8] Gruber T., The Role of Common Ontology in Achieving Sharable, Reusable Knowledge Bases, in *Principles of Knowledge Representation and Reasoning*, J. Allen, R. Fikes, and E. Sandewall, eds., Morgan Kaufman, San Mateo, CA, pp. 601-602, 1991.

[9] Janev V., Vrane S., *Applicability assessment of Semantic Web technologies*, *INFORMATION PROCESSING AND MANAGEMENT*, Elsevier, 2010.

[10] Janjua N. K., Hussain F. K., Web@IDSS Argumentation-enabled Web-based IDSS for reasoning over incomplete and conflicting information, *KNOWLEDGE-BASED SYTEMS*, 2011, Article in press.

[11] Kashyap V., Sheth A., Semantics-based Information Brokering, in *Proceedings of the Third International Conference on Information and Knowledge Management (CIKM)*, pp. 363-370, 1994.

[12] Kulkarni U., Ravindran S., Freeze R., A knowledge management success model: Theoretical development and empirical validation, *JOURNAL OF MANAGEMENT INFORMATION SYSTEMS*, 23(3): 309-347, 2007.

[13] Lau H.C.W., Tsui W.T., An iterative heuristics expert system for enhancing consolidation shipment process in logistics operations, in: Z. Shi, K. Shimohara, D. Feng (Eds.), *INTELLIGENT INFORMATION PROCESSING*, Springer, Boston, pp. 279-289, 2006.

[14] Connor M., Das A., SQWRL: a Query Language for OWL, OWL: Experiences and Directions (OWLED 2009), *Fifth International Workshop*, Chantilly, VA, 2009.

[15] Pomerol J.-C., Scenario development and practical decision making under uncertainty, *DECISION SUPPORT SYSTEMS*, 31 (2):197-204, 2001.

[16] Ramirez R., Melville N., Lawler E., Information technology infrastructure, organizational process redesign, and business value: An empirical analysis, *DECISION SUPPORT SYSTEMS*, 49(4): 417-429, 2010.

[17] Seddon P.B., A respecification and extension of the DeLone and McLean Model of IS Success, *INFORMATION SYSTEMS RESEARCH*, 8(3): 240-253, 1997.

[18] Sheth A., Changing Focus on Interoperability in Information Systems: From System, Syntax, Structure to Semantics, in *INTEROPERATING GEOGRAPHIC INFORMATION SYSTEMS*, M. Goodchild, M. Egenhofer, R. Fegeas, and C. Kottman, Eds., Kluwer Publishers, 1998.

[19] Simon H.A., Administrative Behavior. A Study of Decision-Making Processes in Administrative Organization, *The Free Press*, New York, 1976.

[20] Turban E., Aronson J.E., Decision Support Systems and Intelligent Systems, 6th ed, Prentice Hall, Upper Saddle River, NJ, 2001.

[21] Wagner E.L., Scott S.V., Galliers R.D., The creation of 'best practice' software: Myth, reality and ethics, *INFORMATION AND ORGANIZATION*, 16(3): 251-275, 2006.

[22] Xie Y., Wang H., Efstathiou J., A research framework for Web-based open decision support systems, *KNOWLEDGE-BASED SYSTEMS*, 18 (7): 303-319, 2005.

[23] The W3C Semantic Web Development Tools website http://www.w3.org/wiki/SemanticWebTools