Categorization of Semantic Web Applications – The Basis for Defining Semantic Web Application Development Process

LIDIA ROVAN, TOMISLAV JAGUŠT, MIRTA BARANOVIĆ
Faculty of Electrical Engineering and Computing
University of Zagreb
Unska 3, 10000 Zagreb
CROATIA
{lidia.rovan, tomislav.jagust, mirta.baranovic}@fer.hr http://www.fer.hr/{lidia.rovan, tomislav.jagust, mirta baranovic}

Abstract: - At the moment, the field of Semantic Web is lacking methodologies for development of Semantic Web solutions. We argue that existing methodologies are no longer relevant because of the constant technological changes in the field of Semantic Web as well as in the Web development. We also argue that providing unique development methodology for all Semantic Web applications is not satisfactorily and that specialization of methodologies is needed. To establish foundation for such methodologies we conducted an analysis of Semantic Web solutions resulting in a categorization of Semantic Web applications. In this paper we present our categorization scheme and provide precise definitions for each of the identified categories.

Key-Words: - Semantic Web, categorization, development process, development methodology, analysis, linked data, architecture

1 Introduction

From its beginning, Semantic Web is a very live and changing scientific field. For this reason the results of work in the area of Semantic Web application development tends to become obsolete in a short period of time. Also, since its original appearance, the Semantic Web has seen many changes in its core architectural layers, and even today, many of core building elements are still not standardized. Furthermore, there was a change of focus, from the original idea to build semantic agents [3], which induced the research in the Semantic Web as platform (eg. semantic web services standards) to the latest development directions that are focused on the "Web of data" or "Linked data" principles [2]. We made an overview of the current trends and standards in the field of the Semantic Web. We also made an overview of the current trends in the Web application development process. It showed that most of the basic building elements needed for Semantic Web application development are standardized, and therefore can be used as a cornerstone for future Semantic Web applications. Trough analysis of existing Semantic Web applications that we conducted, a number of typical use-cases were identified, allowing us to propose categorization scheme of the Semantic Web applications. Categorization is an essential prerequisite for proposal of Semantic Web application development process as we believe that defining general architecture and development process for all Semantic Web applications would be at the too high a level of abstraction Therefore, it is necessary to identify categories of Semantic Web applications and propose software architecture and development process for each of them.

The paper is organized as follows. We start out, in Section 2 by defining Semantic Web application and elaborating the motive for the creation of the methodologies supporting Semantic Web application development process. In Section 3 we provide an overview of the related work in this field. In Sections 4 and 5 we present our conducted analysis and discuss obtained results. Finally, we conclude in Section 6 with a brief summary and a short outlook to future work.

2 Semantic Web Application Development Process – Need for the Methodology

2.1 Semantic Web Application

First, it is important to give a precise definition of the Semantic Web application since there are different understandings of that term. A number of existing papers [11], [19], [20] and available literature introduced various definitions of a Semantic Web application. The problem is that many of them are built upon a different, or not enough precise definition of a Web application. For that reason, we introduce definitions that we follow throughout our research:

Based on a Connalen definition [6], we define: Definition 2.1: Web application as a Web system in which a user action can change the state of a business according to the defined business rules.

Definition 2.2: Semantic Web application is a Web application that depends on the Semantic Web standards for its successful execution.
To classify a Web application as a "semantic" it is sufficient for it to use semantic technologies in at least one of its functional components.

2.2 Development methodology

In general, engineering is defined as a process of physical system implementation, starting from the creation of high-level abstractions and logical model independent of the implementation [9].
Most mature Web design methodologies emphasize the importance of separating implementation of data model, behavior model and navigation (user interface). Clear separation of tasks which the solution has to meet improves the quality of design, which is crucial for achieving the re-usability of software modules as well as to ease the evolution and maintenance of application [14]. After the Model-View-Controller (MVC) form was recognized as a suitable solution for the Web development, many Web frameworks were based on it (e.g. Struts, Tapestry, Spring MVC). Although there were many changes in the modeling and application development principles trough years, a number of recent trends in the application development is still based on the MVC. At the moment, due to their immaturity, Rich Internet applications (Ajax programming frameworks) are mostly not based on MVC pattern. However, research shows the need for such solutions, so their number is expected to increase in the near future [17].
Semantic Web Engineering (similar to the Web engineering) is engineering applied to the Semantic Web applications. Although it is considered to be a cloned Web engineering, Semantic Web engineering contains new approaches, principles, methods, tools, techniques and guidelines that can meet the specific needs of a Semantic Web application. It is important to set up a strong methodology basis for the future development in the field of Semantic Web field in order to avoid a crisis similar to the "Web crisis" [10] that struck the Web development in its beginnings.
It is important to emphasize that the development of Semantic Web applications significantly differs from the development of Web applications, and brings a number of additional challenges. Ordinary Web applications are primarily oriented toward human users, operate using unstructured data with informal logic and links between documents. On the other hand, Semantic Web applications are oriented toward human users and machines (software agents), operate with structured formal statements, and use a formal descriptive logic with the links between data.
We believe that proposing general architecture and development process for all Semantic Web applications would be at the too high level of abstraction, and as such not useful. Therefore, as a result of this research we identify categories of Semantic Web applications, for which we will later in the research separately propose architecture and development methodology that would be in accordance with the current Web application development methodologies.

3 Related work

Previous research in the area of the Semantic Web applications development resulted in a number of development methodologies proposals [8] [13], [20]. We argue that those researches are no longer relevant, mostly because of the constant changes of the Semantic Web technologies, but also because of the changes in the Web applications development process itself.
Hera methodology [13] divides Semantic Web application in three layers: semantic, application and presentation layer. They do not propose general solution guidance; but only focus on the presentation layer generation. Since their work was created before appearance of the relatively new Semantic Web technologies, for example RDFa (Resource Description Framework – in – attributes) which significantly changed the way of presentation layer generation, Hera does not meet the current standards. Also, it supports only restricted navigational capability, which makes the methodology applicable only to the small set of applications, with restricted number of different use case scenarios.
Corcho et al. [8] proposed extension of the ODESeW framework [7] for Semantic Web portal development. The solution uses a specific MVC framework and non-standard technologies (RQL - RDF Query Language) and therefore cannot be used as a general guideline for the Semantic Web portal development.
Heitmann [12] conducted a research similar to ours - analysis of Semantic Web applications with the aim of discovering common design problems. But, his analysis included applications which do not satisfy the definition of the Web application. The relevance of that research is also questionable regarding the fact that analyzed Semantic Web solutions were developed before 2007 and only for the scientific purposes. In addition, the research focused on identifying only the most frequent architectural components, regardless of the category in which they usually appear and manner in which the components are interconnected in such cases. However,

our research suggested that each category has a specific set of functionalities and method of applying technologies for achieving such functionality.

4 Semantic Web Solution Analysis

As a part of annual European and world conference on the Semantic Web, there are contests for the best Semantic Web application. Also, W3C (World Wide Web Consortium) maintains a list of existing Semantic Web solutions. We analyzed Semantic Web applications based on the following criteria:

- Active projects from the W3C list
- Projects from the "Scripting for the Semantic Web" contest from 2005 to 2009
- Projects from the "Semantic Web challenge" contest, from 2003 to 2008

The main goal of the analysis was to categorize the existing Semantic Web applications, but also to recognize the most frequent use cases, and typical challenges that are being solved by semantic technologies.

Also, it was necessary to identify the functionalities typical for the particular category. Their implementation in various applications differs because of the specific real life situations the applications are built for. We tried to recognize the most frequent design patterns, so by applying them we could suggest the generalization of functionalities.

Besides fulfilling the mentioned basic goals set on the analysis, the results also showed that the Semantic Web is moving towards the modern research directions, and tends to materialize the "Linked data" ("Web of data") and "Web of tags" principles.

The sample set included 128 solutions (Fig. 1).

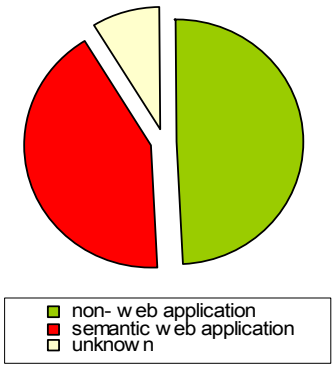

Fig. 1 The sample set

For 11 solutions, we could not find the documentation nor was the solution itself available, so the analysis is based on the remaining 117 solutions. About half of them (54) satisfy our definition of the Semantic Web application. The rest mainly consists of standalone applications, application frameworks etc.

During the analysis, we tried to match each solution with the corresponding Semantic Web research direction. The solution timeline confirms our assumption that the Semantic Web development moves toward new research areas such as "Linked data". Fig. 2 shows the distribution of solutions belonging to different research directions.

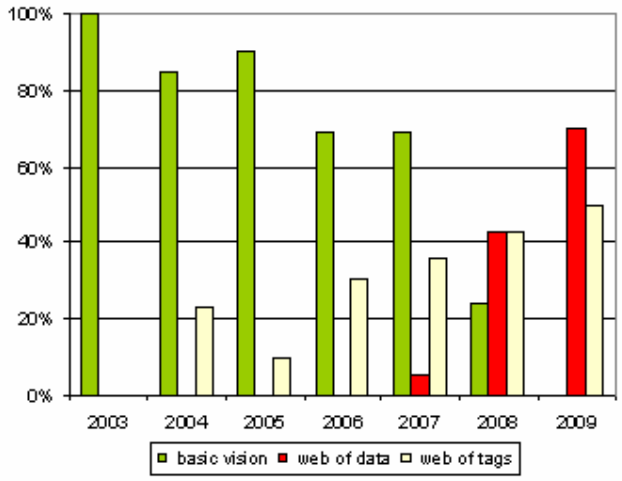

Fig. 2 Number of solutions in each research direction over the years

For solutions which are a part of multiannual projects, and were in different contests over the years, we considered only the most recent year, assuming that over the years the solution adopted itself to the trends. Also, it is possible that one solution belongs to more than one research direction (eg. "Web of data" and "Web of tags").

The solution is a part of a research trend if it has the following characteristics:

- The basic vision – solutions that use ontologies with high-level logic as their data model, and are focused on gaining extra value from the model using reasoning. Typical applications examples of this research direction are the ones that support development of the main vision of the Semantic Web (eg. Swoogle – RDF and OWL search engine).
- Web of data – program solutions built upon the rules from the "Linked data" concept.
- Web of tags – the solution does not use ontology or "Linked data" model; it is aimed on tagging the data, usually using low-level logic. The tagging is typically performed by users, trough some custom interface.

5 Semantic Web Application Categories

Before we started with the analysis of available material and online applications, the preliminary categories were defined. For the list of preliminary categories we did an overview of the similar researches in this area. There are a number of papers that deal with the categorization of classical Web applications [23]. The problem was that most of them were not specialized in Web applications (they also examined the Web sites). Furthermore, the applications considered in previous studies were made before 2002 and as such did not include the recent trends in Web development that enabled the creation of the new categories, for example social applications.

In his research, Heitmann [12] proposed the categorization of Semantic Web solutions according to three criteria: application domain, architecture type and application type. He divided the architecture into a centralized server with Web frontend, decentralized network of servers, standalone application and a peer-to-peer network. Evidently, Heitmann also takes into account the solutions that do not match the definition of a Web application. He determines the application type by functionalities that the application has, and claims that one application can belong to several types at the same time. The application type can be one of the following: a semantic portal, semantic annotation, semantic repository, semantic authoring, semantic desktop application and semantic scripting language. Unlike Heitmann's assumptions, the hypothesis of our research is that the Semantic Web applications can be categorized according to their main purpose, regardless of the domain to which they belong. Also, it is assumed that each of these categories contains particular set of functionalities, and that the purpose and functionalities define the application architecture.

Finally, on the basis of previous research, we established following preliminary categories: Semantic Web portal, semantic knowledge management system and semantic recommender application. In line with the trends in the Web development, the category "social Semantic Web applications" was also added to the list.

In accordance with the literature and theoretical foundations, we made definitions for categories. The requirement that the category must meet the theoretical frame means that it the assessors evaluated to what extent the applications met the requirements arising from the category definition given below. Every definition clearly outlines the purpose and functionality that is expected from the application.

The first step was to categorize the test sample, which resulted in revising the categories list. Then the categorization of all applications selected for analysis was conducted. In the analysis, we ignored the categorization claims by the authors themselves, since it seemed that they often differ from the theoretical frames. Categorization was performed by two different assessors, which led to yet another adjustment of the categories. In the end, the following categories were formed (Fig. 3 Frequency of the analyzed applications by defined categories):

- Semantic Web portal
- Semantic knowledge management system
- Semantic Web recommender application
- Social Semantic Web applications
- Semantic Web expert system
- Applications that use semantic technology for content tagging
- Applications that use semantic technology to improve search results

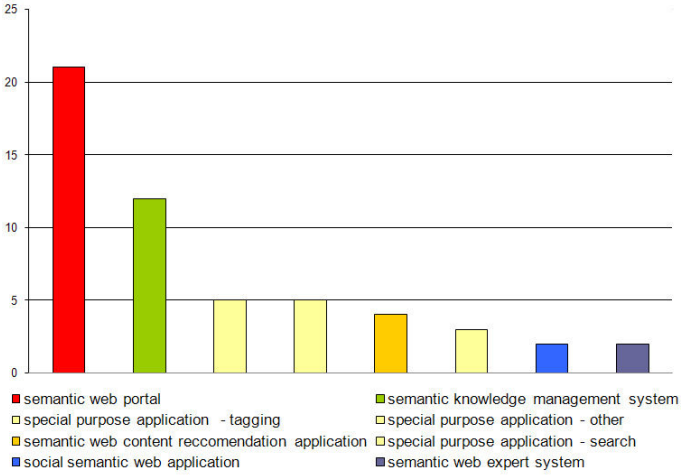

Fig. 3 Frequency of the analyzed applications by defined categories

5.1 Semantic Web portal

Based on the research that Smith [22] conducted for the purpose of giving general definition of a Web portal:

Definition 5.1 A: The Web portal is an infrastructure that provides secure, customizable, personalized, integrated access to dynamic content from a variety of sources, in a variety of source formats, wherever it is needed.

In analogy to the definition of the Semantic Web application:

Definition 5.1 B: The Semantic Web portal is a Web portal that depends on the Semantic Web standards for its successful execution.

Since the portal serves as a unique interface for gathering different heterogeneous data sources, the functionality it must have is the integration. Personalization and customization are less important, since their level of implementation in the portal fully depends on the domain the portal covers. Semantic Web portal usually uses semantics to improve integration, browsing, search and personalization.

5.2 Semantic knowledge management system

In the literature, there are many papers on the topic of systems for knowledge [1], [16] and documents and content [4] management. Mostly according to [16], a knowledge management system is considered a superset of all afore-mentioned forms and is defined as:

Definition 5.2 A: A knowledge management system is an application and a set of related processes that allows the identification, storage and retrieval of the intellectual capital to the particular organization (any type of organization) or a group of users that share a same interest.

The study conducted by Joo and Lee [15] identified the basic technical limitations of knowledge management systems that the Semantic Web technology could affect, namely the multiple data source integration, search implementation, user dissatisfaction with the system usability and the lack of consistency and completeness of knowledge. The study noted that Semantic Web technologies could provide a solution to the problems detected. In accordance with the research from Joo and Lee, the implementation of semantic technologies in knowledge management systems became one of the most frequent Semantic Web application categories. According to Joo and Lee research we give a definition for the semantic knowledge management system:

Definition 5.2 B: Semantic knowledge management system is a knowledge management system that uses Semantic Web technologies to improve integration, search results, user satisfaction and accuracy of knowledge that the system handles.

Knowledge management systems are the systems used by the group of users with the aim of mutual knowledge exchange. Their main functionality, which makes them different from the other applications, is the ability to store and handle a variety of multimedia content. Semantic technologies in these systems are commonly used for improved document tagging, which enhances their retrieval.

5.3 Semantic Web recommender application

According to Resncick and Varian, a typical recommender system is defined [21]:

Definition 5.3 A: The recommender system is a system in which user provides recommendations as inputs, which the system then aggregates and directs to appropriate recipients.

From basic definitions we exclude the part which states that people are necessarily the ones who provide recommendations; the present technology allows a variety of ways to identify recommendations without explicit human input. The value of these applications is in the design of recommendation aggregation algorithm and in finding the proper relationship between the recommendations and those who search for them. Applying semantic technologies on these two key issues may lead to operational improvements. In the simplest cases of semantic usage in recommender systems, content and user descriptions are stored in semantic format (ontology) allowing identification of additional data by reasoning over the ontology. More complex semantic procedures use semantic descriptions also in the process of recommending content, by widening opportunities to analyze data and users. The potential of semantic processes lies in the possibility to discover new knowledge about users, content, and their mutual relationship, which can be used in the process of content recommendation. In accordance with explanation we introduce the following definition:

Definition 5.3 B: Semantic Web recommender application is a Web recommender system which uses semantic technologies for the purpose of conducting the analysis on which the user finds the recommended content.

Web portals that include personalization often implement a recommender system as well. In such circumstances, this category can be considered as a subset of the Web portals. In other words, that kind of Web portal is also a recommender system.

5.4 Social Semantic Web applications

A social Web application is a term that has different meanings, usually broader than those defined in this research.

Often the term Web 2.0 is considered equivalent to social Web application [18]. Here we introduce stricter and narrower definition of social applications - oriented to the social networks. Therefore, the social Web application is focused on communication and collaboration among users. All other applications, where interaction between users is not the main goal of an application, are excluded from this category. An example is Wikipedia, which is often cited as a social application, but in fact it belongs to the knowledge management category, since it does not contain any form of user interaction.

Definition 5.4 A: Social Semantic Web application is a Web application oriented to the social network that uses semantic technology to enhance cooperation/interaction among its users.

Social Web applications are a newer form of applications. These applications are aimed at any form of interaction between individual users, as well as within the user groups. Since the FOAF (Friend of a friend) gained huge popularity in the Semantic Web area, it is often used as a tool that allows distributed user profiles and implementation of social networks.

5.5 Semantic Web expert system

According to Brooks [5], the expert system is defined as:

Definition 5.5 A: An Expert System is a program containing a generalized inference engine and a rule base, designed to take input data and assumptions and explore the logical consequences through the inferences derivable from the rule base, yielding conclusions and advice, and offering to explain its results by retracting its reasoning for the user.

Thus, the expert system is a computer system that simulates the process of expert thinking while resolving a complex problem. It is used in problem solving or decision making tasks. Its purpose is to replace an expert in the particular field, her knowledge and the work she does. The defined rules represent the expert's knowledge whereas reasoning on the same rules represents an expert's work.

While building an expert system, Semantic technologies can be used to define data, a base of rules and reasoning over them, for example by using OWL. Also, since the data is semantically tagged, their interpretation is improved.

Definition 5.5 B: The Semantic Web expert system is an expert system implemented as a Web application that uses semantic technology to define the rules, the inference over these rules and to improve interpretation of data.

This category is mainly used in very narrow domains, since it replaces the specific field experts and is not expected to be useful in a wide area of problems. Typically, the applications from this category are developed in the medical field.

5.6 Special purpose applications

The remaining categories consist of Web applications that are designed to resolve specific business problems, thus it is not possible to generalize them into one of the categories above.

Such applications are further divided into categories depending on the way in which they implement the semantic technologies.

- Applications that use semantic technology for content tagging (the data is tagged using the top-down approach)
- Applications that use semantic technology to improve search results

Definition 5.6 A: Semantic special purpose application is a Web application that is oriented towards solving specific business problems of a domain using Semantic Web technologies.

6 Conclusion and future work

The main goal of our research is to propose of development process for Semantic Web applications. In order to achieve that goal, it was necessary to explore what characterizes the Semantic Web application and how it is defined. Also, it was necessary to do a review of available Semantic Web technologies and Web application development trends in order to find out are the Semantic Web technologies mature enough to be building blocks of a modern Web application. Finally, based on our hypothesis that providing general architecture and development process for all Semantic Web applications would be inadequate, we conducted an analysis of Semantic Web solutions in order to identify typical use cases. In this paper we shortly presented the methodology used throughout our research. We elaborated the results of our analysis providing definitions for the identified frequent Semantic Web application categories.

Once we have established categorization scheme, and defined basic functionalities for each of the categories, we set foundation for future work, which will consist of defining architecture for the categories and later, the development process for the Semantic Web applications.

References:

[1] Alavi, M., Leidner, D., Review: Knowledge management and knowledge management systems: Conceptual foundations and research issues, *MIS Quarterly*, pp. 107-136, 2001

[2] Berners-Lee, T.: 'Linked data', http://www.w3.org/DesignIssues/LinkedData.html, 02.09.2010.

[3] Berners-Lee, T., Hendler, J., and Lassila, O.: 'The semantic web', *Scientific American*, 284, (5), pp. 34-43, 2001

[4] Boiko, B., *Content management bible 2nd edition*, Wiley Publishing, Indianapolis, USA, 2005

[5] Brooks, F., No silver bullet: Essence and accidents of Software engineering, *IEEE computer*, 20, (4), pp. 10-19, 1987

[6] Conallen, J.: 'Modeling Web application architectures with UML', Commun. ACM, 42, (10), pp. 63-70, 1999

[7] Corcho, O., Gómez-Pérez, A., López-Cima, A., López-García, V., Suárez-Figueroa, M.: 'ODESeW. Automatic generation of knowledge portals for Intranets and Extranets', *Lecture Notes in Computer Science 2870*, pp. 802-817, 2003

[8] Corcho, O., López-Cima, A., Gómez-Pérez, A.: 'A platform for the development of semantic web portals'. *In: Proceedings of International Conference on Web Engineering*, , pp. 145-152, 2006

[9] Demeyer, S., Ducasse, S., Nierstrasz, O., *Object-oriented reengineering patterns*, Morgan Kaufmann, 2003
[10] Ginige, A., and Murugesan, S., Web engineering: an introduction, *Multimedia, IEEE*, 8, (1), pp. 14-18, 2001
[11] Hausenblas, M., Slany, W., Ayers, D.: A performance and Scalability metric for virtual RDF graphs. *3rd Workshop on Scripting for the Semantic Web*, Innsbruck, Austria, 2007
[12] Heitmann, B., Transitioning Web application frameworks towards the Semantic Web, *Master's Thesis*, Universität Karlsruhe, 2007
[13] Houben, G., Barna, P., Frasincar, F., Vdovjak, R., HERA: Development of Semantic Web information systems, *Lecture notes in computer science*, pp. 529-538, 2003
[14] Jacyntho, M., Schwabe, D., Rossi, G., A Software Architecture for Structuring Complex Web Applications, *Journal of Web Engineering*, 1, (1), pp. 37-60, 2002
[15] Joo, J., Lee, S., Adoption of the Semantic Web for overcoming technical limitations of knowledge management systems, *Expert Systems With Applications*, 36, (3P2), pp. 7318-7327, 2009
[16] Maier, R., Knowledge management systems: Information and communication technologies for knowledge management, *Springer-Verlag*, New York Inc, 2004
[17] Mycroft, A., Controlling Control Flow in Web Applications, *Electronic Notes in Theoretical Computer Science (ENTCS)*, 200, (3), pp. 119-131, 2008
[18] O'Reilly, T., What is Web 2.0: Design patterns and business models for the next generation of software, http://www.oreillynet.com/pub/a/oreilly/tim/news/2005/09/30/what-is-Web-20.html, 02.09.2010
[19] Oren, E.: *Algorithms and Components for Application Development on the Semantic Web*, National University of Ireland, 2008
[20] Reif G., *WEESA – Web Engineering for Semantic Web Applications*. Ph.D. thesis, Technische Universität Wien, 2005
[21] Resnick, P., Varian, H., Recommender Systems, *Communications of the ACM*, 40, (3), pp. 56-58, 1997
[22] Smith, M.A., Portals: toward an application framework for interoperability, Communications of the ACM, 47, (10), pp. 93-97, 2004
[23] Ziemer, S., Web Engineering, http://csgsc.idi.ntnu.no/2004/data/svenz/WE_csgsc.pdf, 02.09.2010.