

The Semantic Web: A mythical story or a solid reality

Jorge Cardoso

SAP AG, Germany, jorge.cardoso@sap.com

Abstract The Semantic Web vision has drove hundreds of practitioners to research and develop a new bread of applications that could take the full potential of the Web to the next level. While there is a fairly clear understanding of where Web 1.0 and Web 2.0 stand now a day, the current status and position of the Semantic Web, also known as Web 3.0, is not as clear and well defined. Therefore, in this paper we present a landscape that illustrates and captures the trends in the Semantic Web with the purpose of guiding future developments.

1. Introduction

The vision for a Semantic Web has been termed as *The Next Big Thing* for information system. The publication of the Scientific American Magazine article “The Semantic Web” in May, 2001, by Tim Berners-Lee, James Hendler and Ora Lassila has triggered a strong wave of research worldwide. Thousands of research papers have been written and hundreds of applications have been implemented. But what are the research trends. That is what we will address in this paper by presenting a summary of a survey (Cardoso, 2007) to gives an account of current Semantic Web practices. The findings reported in this article are based on 627 surveys that were filled in and conducted from 12 December 2006 to 26 January 2007.

2. Web.X – A story in 3 chapters

The Semantic Web has been coined Web 3.0. While previous Web technologies, such as Web 1.0 and Web 2.0, had a remarkable success, there is no consensus that Web 3.0 had, or will have in the future, the same levels of success. While some researchers believe that the Web 3.0 is still to come to delight users and take

them to a new paradigm, others believe that it has been such a long time since the Semantic Web vision was introduced that it is becoming almost mythical.

The Web 1.0 classifies the very first applications for the Web which had the main objective to only provide information content to end users. Web 2.0 brought interactivity to users. Features such as tagging, social networks, online media, content aggregation and syndication (RSS), mashups and folksonomies became common on the Web.

Web 1.0 and Web 2.0 successes are undeniable. Success stories are available at the distance of a mouse click. Web 1.0 has delivered innovative solutions such as DoubleClick, mp3.com, Britannica Online, personal websites, screen scraping, and content management systems. Web 2.0 has engineered landmarks such as Google AdSense, Flickr, BitTorrent, Napster, Wikipedia, and del.icio.us.

But what about the Web 3.0? When will the Semantic Web arrive? The building blocks are already available: RDF, OWL, and microformats are a few of them. During the 2007 Semantic Technology Conference, several new and robust start-ups companies based on – Semantic Web technologies – have demonstrated their interest and commitment to this new concept for their business. But while it seems that we are getting close to deliver the promised automated applications, we are still a few years away from unleashing the killer-app. Interesting companies that are actively trying to implement the Semantic Web include Twine, Garlik, and Hakia. Twine is a real application of the Semantic Web that gives users a superior way to share, organize, and find information with people they trust. It uses the Semantic Web, natural language processing, and machine learning to make your information and relationships more expressive. Garlik is using Semantic Web technologies to help users monitor their personal information online and protect themselves against identity theft. Hakia is a semantic search engine with a focus on natural language processing methods to try and deliver search results based on analyzing semantics and not keywords.

According to Google search engine, the number of queries for “Semantic Web” keywords has dropped since 2004. Figure 1 shows a trend in Semantic Web popularity. What does this empirical trend represent the semantic research field?



Fig 1. Volume of search queries for the keywords “Semantic Web”

On the one hand, this can be explained by the fact that the hype is fading away and that people are losing interest in Semantic Web solutions. On the other hand, it can be explained by the fact that the concept is already accepted and understood by many people and, therefore, it does not need to be searched for. To better understand the current research and industrial efforts to make the Semantic Web a reality, and not a mythical story, we have carried out a survey in 2006 and 2007. The results are summarized in the following section.

3. Current trends of the Semantic Web

The survey carried out was divided into five categories: Demographics, Tools and Languages, Ontology, Ontology Size, and Production. The complete results of the survey can be found at (Cardoso, 2007) and are here presented as a summary of the most interesting points.

Ontology Editors. We asked respondents to tell us which ontology editors they were currently using in their organizations. The editor most frequently cited was Protégé with a market share of 68.2%. Approximately equal numbers of respondents use SWOOP (13.6%), OntoEdit (12.2%) (OntoEdit is now called OntoStudio), and Altova SemanticWorks (10.3%). A good survey of the most popular ontology editors can be found in (Escórcio and Cardoso, 07).

Ontology Languages. Our study revealed more than 75% of ontologists have selected OWL to develop their ontologies and that more than 64% rely on RDF(S). Curiously, Description Logic and FLogic are also being used with a penetration rate of 17% and 11.8%, respectively. A recent language, WSML (Web Service Modeling Language), has also gained some popularity (3.7%).

Reasoning Engines. We asked all respondents to indicate the reasoning engines they were using. The largest segment (53.6%) indicated that they were using Jena (McBride, 2002). Smaller groups indicated they were using Racer (28%) and Pellet (23.7%). FaCT++ (13.3%) and OWLJessKB (8.1%) have also gained preference by a small group of participants.

Ontology Domains. To determine the actual trend in the development of ontologies for particular domains, we asked all respondents to indicate for which industries they were representing knowledge with ontologies. Education and Computer Software are the best represented industries (31% and 28.5%), followed by Government (17%) and Business Services (17%).

Methodologies. We asked all respondents to indicate which methodology or method they were employing to develop their ontologies. Sixty percent develop

ontologies without using any methodology. The methodologies with greatest adoption are METHONTOLOGY (13.7%) followed by On-To-Knowledge methodology (7.4%) and Uschold and King's method (4.2%).

Purpose of using ontologies. We asked participants to tell us the reasons that motivated them to use ontologies. The vast majority of participants (69.9%) answered "to share common understanding of the structure of information among people or software agents, so models can be understood by humans and computers".

Techniques used with ontologies. We asked respondents to tell us if they used any specific technique to manipulate ontologies. The largest segment of respondents (67.1%) indicated that they use ontology mapping. Roughly equal numbers of participants indicated that they were integrating ontologies (40.7%) and merging ontologies (33.9%). Twenty six percent stated that they align ontologies.

Ontologies size. We asked each respondent to indicate the average size of the smallest, typical, and biggest ontologies they were working with. A vast majority of respondents (72.9%) indicated that their smallest ontologies had less than 100 concepts. When asked about typical ontologies, forty four percent of respondents stated that such types of ontology had between 100 and 1000 concepts.

Production. Each respondent was asked to indicate if they would put their systems into production. More that 70% of all respondents indicated that they planned to adopt ontology-based systems, while 27.9% indicate that they do not have any plans to use such types of systems in the future.

4. Conclusions

Web 1.0 and Web 2.0 had, and still have, an important impact on society and electronic commerce. The next step for the evolution of the Web is the Semantic Web, or Web 3.0. The information presented in this article constitutes a checkpoint to frame the current status and trends in the Semantic Web. It depicts a paradigm change from an idea to the development and use of real and concrete semantic solutions.

5. References

- Cardoso, J., "The Semantic Web Vision: Where are We?" IEEE Intelligent Systems, September/October 2007, pp.22-26, 2007.
- Escórcio, L. and Cardoso, J. "Editing Tools for Ontology Construction", in Semantic Web Services: Theory, Tools and Applications, Idea Group, March 2007. ISBN: 1-59904-045-X.