Information Processing and Management 47 (2011) 507–517

Contents lists available at ScienceDirect

Information Processing and Management

journal homepage: www.elsevier.com/locate/infoproman

Applicability assessment of Semantic Web technologies

Valentina Janev *, Sanja Vraneš

Mihajlo Pupin Institute, Volgina 15, 11060 Belgrade, Serbia

ARTICLE INFO

Article history:
Received 19 March 2009
Received in revised form 22 September 2010
Accepted 4 November 2010
Available online 27 November 2010

Keywords:
Semantic Web technologies
Applicability
Assessment
Industry
Adoption
Applications
W3C

ABSTRACT

The Semantic Web is one of the fastest developing fields within the Information and Communication Technology sector and, as such, under constant examination by scientists and IT professionals. This article aims to provide a better understanding of the applicability of Semantic Web tools and technologies in practice. This aim will be achieved by surveying the recommended and emerging W3C standards, presenting an overview of the state-of-the-art in the Semantic Web research in the European Union, analysing the W3C collection of Case studies and Use Cases, and discussing the extent of adoption of Semantic Web technologies. The overall technology maturity level assessment has shown that Semantic Web technologies are finding their ways into real-world applications, and that, rather than being merely a fashionable research issue, the Semantic Web, slowly but surely, becomes our reality.

1. Introduction

New technologies, such as Semantic Web (SW) technologies, are usually subjected to experimentation, refinement, and increasingly realistic and exhaustive testing. This kind of information gathering, which aims to look beyond the immediately obvious and at analysing the ramifications of a given technology in as wide-ranging and far-sighted a manner as possible, is known as *technology assessment* (Braun, 1998). Technology assessment is usually based on different forecasting methods including extrapolation, expert opinion (the Delphi method) and modelling, cost-benefit analysis, cross-impact analysis, and others.

The term "Semantic Web" refers to the World Wide Web Consortium's vision of the Web of linked data (called also the Web of Data) as "…an extension of the current Web in which information is given a well-defined meaning, better enabling computers and people to work in cooperation" (Berners-Lee, Hendler, & Lassila, 2001). There are few international standardization organizations (associations or consortia) relevant for assessment of information technologies such as IEEE-SA (see The Institute of Electrical and Electronics Engineers Standards Association, http://standards.ieee.org/), OASIS (see The Organization for the Advancement of Structured Information Standards, http://www.oasis-open.org/), OMG (see The Object Management Group, http://www.omg.org/) and W3C (see The World Wide Web Consortium, http://www.w3.org/).

In order to determine the achievements in the Semantic Web field, especially its maturity status and the adoption of these technologies by the industry, as well as to predict the future development of SW technologies, we have studied and analyzed a substantial amount of various sources. These range from deliverables from projects financed by the European Commission within its 6th and 7th Framework Programmes (EC FP6 + FP7), through scientific papers from prestigious international journals and conferences, Semantic Web technologies and tools from the industry top vendors and open source communities

* Corresponding author. Tel.: +381 11 2774 024; fax: +381 11 2775 835.
E-mail addresses: Valentina.Janev@impcs.com (V. Janev), Sanja.Vranes@institutepupin.com (S. Vraneš).

0306-4573/$ - see front matter © 2010 Elsevier Ltd. All rights reserved.
doi:10.1016/j.ipm.2010.11.002

(Janev, 2008), to Web resources of the W3C standardization body. This paper clarifies the applicability of Semantic Web technologies and identifies the major trends in the Semantic Web field, both from researchers' point of view and from the perspective of early adopters.

The article is organized as follows. After identifying the work related to Semantic Web technologies assessment in Section 2, we point to existing and emerging SW standards and discuss their adoption for developing new applications and for meeting different technical and application challenges. Using the W3C collection of Case Studies and Use Cases, we infer relations between the SW technologies and the gained benefits. Section 4 considers the state-of-the-art of SW technologies from several viewpoints such as development of SW languages, availability of data, content and ontologies on the Semantic Web, data and service integration, while in Section 5 we point to open issues such as stability of SW languages, interoperability issues, and scalability of SW applications.

2. Related work to Semantic Web technologies assessment

In literature and in practice, SW tools and technologies are named by using different keywords: ontology design/management/maintenance tools, semantic data management and integration platforms (Ahmad & Colomb, 2007), RDF triple storage systems, web services, SOA middleware platforms, semantic annotation tools (Reeve & Han, 2005), content indexing and categorization tools, semantic search and information retrieval technologies (Andrews, 2009; Crestani, Lalmas, Van Rijsbergen, & Campbell, 1998; Mangold, 2007), NLP, linguistic analysis and text mining algorithms, collaboration and other social networking technologies (Correndo & Alani, 2007; Gootzit, Phifer, Valdes, & Knipp, 2009), knowledge visualization/presentation technologies, ontology mediated portals, ontological querying/inference engines, rule-based engines, ontology learning methods (Gómez-Pérez & Manzano-Macho, 2003), ontology reasoners, etc. Exploring the business value of semantic technologies provided by 50 commercial companies (Davis, Allemang, & Coyne, 2004) identified the following four major functions: discover, acquire, & create semantic metadata; represent, organize, integrate, and inter-operate meanings & resources; reason, interpret, infer, & answer using semantics; and provision, present, & communicate, and act using semantics. Analysing the functionalities of more than 50 SW tools (see the Web4WeB repository of the Sixth Framework Program "Web Technologies for the West Balkan Countries", at www.web4web.org/portal/Semantic_Web_Tools) provided by 30 different commercial vendors and open source communities, Janev and Vraneš (Janev, 2008) established the following classification of key semantic technology segments: semantic modelling and development, semantic annotation, semantic data management and integration, semantic search and retrieval, semantic collaboration including portal technologies, learning and reasoning. According to this analysis, vendors have made recognizable progress towards the specification and acceptance of semantic standards, but still lack efficient reasoning support which is crucial for realization of the Semantic Web vision.

The literature review has showed that most of the scientific studies usually analyze a single aspect of semantic technologies, and that just a few studies provide the key trends in the Semantic Web field (see, e.g., Davis et al., 2004; Cali et al., 2005; Cardoso, 2007; d'Aquin et al., 2008) or investigate the adoption of semantic technologies in various scientific and industrial domains. While some surveys take an industrial economic approach towards semantic technologies (Davis et al., 2004; Provost, 2008), others concentrate on the evaluation of specific technologies and their potential success within specific application areas (Davis et al., 2004; Cuel, Delteil, Louis, & Rizzi, 2007; Pellegrini et al., 2009). In the "The Technology Roadmap of the Semantic Web" white paper the authors deliver a comprehensive analysis maturity, applicability and adoption of the SW technologies (Cuel et al., 2007) based on a survey conducted with selected experts from industry and academia. Using the Gartner Hype Cycle Curve (Linden & Fenn, 2003), they present expert opinions, i.e. results of estimation of years to mainstream adoption for different SW technologies and applications both from the researchers' and the business point of view. They found out that: "research community considers that developments from the past 10 years have resulted in some tools and standards, which are reliable and mature enough to be transferred to industry and successfully integrated into SW applications. The developers' community, however, is not yet fully aware of the availability of such tools, which, consequently, has to be promoted further, together with the innovative functionalities they can provide to software applications". In the "The Semantic Web Awareness Barometer" (Pellegrini & et al., 2009), the authors use statistical methods to summarize the results of the survey and the Chi-square test to compare expert opinions from industry and academia. The analysis indicated that "the expectations in Semantic Web technologies were very high in both groups. Generally both groups believe that organizational culture is not yet ready for the Semantic Web. Additionally, application-oriented participants believe that Semantic Web technologies are too complex. On the contrary, research-oriented participants believe that the lack of success stories, a lack of quality of available software, the problem of quantifying the benefits, the costs of implementation and the general heterogeneity of information are the biggest obstacles to the application of Semantic Web technologies."

3. Analysis of the W3C collection of Case Studies and Use Cases

3.1. Semantic Web standards

In the time since the SW's initial conception (Berners-Lee et al., 2001), the Semantic Web Activity Group of the World Wide Web Consortium (W3C) has accepted numerous Web technologies as standards or recommendations. In an attempt