RECIIS
Electronic Journal of Communication Information & Innovation in Health
[www.reciis.cict.fiocruz.br]
ISSN 1981-6286

Book Reviews

Semantic Web services, processes and applications

Jorge Cardoso and Amit P. Sheth (Eds.)

Semantic Web and semantic Web services

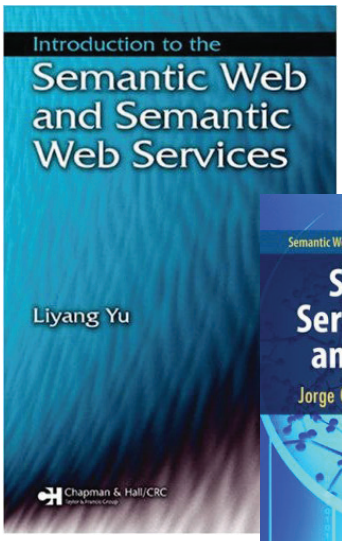
Liyang Yu

DOI: 10.3395/reciis.v3i1.246en

Laura Cristina Simões Viana
Coordination of Human Resources in the Research Field of the Vice-Presidency for Technological Research and Development, Oswaldo Cruz Foundation, Rio de Janeiro, Brazil
laura@fiocruz.br

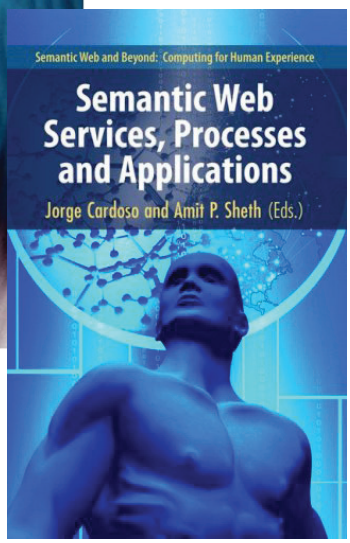
Tim Berners-Lee developed, in 1989, the basic tools required by the Web as we currently know it: the Hypertext Markup Language – HTML; the Hypertext Transfer Protocol – HTTP; and a system for locating objects on the Web, the Universal Resource Locator – URL. Twelve years later[1], with a fully functional Web, Berners-Lee published an original article, with the challenging proposal of a "Semantic Web" (Berners-Lee 2001).

According to his proposal, the Semantic Web would be an extension of the current one, in which information would be given well-defined meaning, better enabling computers and people to work in cooperation. Thus the current Web, basically comprised of documents presented by computers and read by man, would also include data and information that would automatically be handled by agents and utilities. Back then, Berners-Lee (2001) argued that advancements in the Semantic Web would require the development of a language able to express data and data reasoning rules, in addition to enabling the any knowledge representation system to be exported to the Web. Up to that time, two key technologies for the achievement of the Semantic Web had been developed: the eXtensible Markup Language – XML, and the specifications family Resource Description Framework – RDF, the latter intended for the description or modeling of information implemented in Web resources.



New York, Springer, 2006

ISBN 978-0387302393



USA, Chapman & Hall/ CRC, 2007

ISBN 978-1584889335

In 2006, although the Web standards that express shared meaning were still under development, Shadbolt (2006) published another article entitled "The semantic Web revisited", where He stated that the Semantic Web would basically remain in the sphere of ideas, until such Standards were well-established and agreed upon. Shadbolt (2006) emphasized that, similarly to what happened to the HTTP protocol, whose pioneer utilization by the physics community paved the road for the success of the Web, the growing use of ontologies by the e-science community could also lead the Semantic Web to the same degree of success achieved by the Web as we know it. According to the author, in addition to social factors and design decisions, the success of the Semantic Web can be partially credited to the sequence of specifications (Universal Resource Identifier - URI, HTTP, RDF, ontologies etc.) and of records (URI scheme, internet contents of the Multipurpose Internet Mail Extensions – MIME type or multifunctional extensions for Internet messages), which provide the means for constructions such as ontologies to derive meaning from a URI identifier.

The book by Jorge Cardoso & Amit Sheth, published in 2006, Liyang Yu's book, published in 2007, are practical texts that approach the Semantic Web subject with all the multitude of Standards, languages, applications and processes that the theme provokes. Amit Sheth has a Doctors' Degree in Computer and Information Sciences from the Ohio State University, and was a professor at the University of Georgia between 1994 and 2006. He published over 250 peer revised articles and eight books; he is currently a professor of Computer Engineering and Science, and director of the Kno.e.sis Center, both at Wright State University. Jorge Cardoso took his Doctorate in Computer Sciences along with Amit Sheth at the University of Georgia, and currently works at the Research Department of German company SAP. He is also a professor at Universidade da Madeira, Portugal. Cardoso has published over 80 articles in peer revised newspapers in the fields of workflow management systems, Semantic Web, and has edited three books on the Semantic Web and Web services. Liyang Yu took Doctorate at Ohio State University, and currently works for Delta Technologies in the United States.

The book edited by Jorge Cardoso and Amit Sheth is a collection of contributions by researchers from universities and companies about semantics, divided into three parts: the first deals with Semantic Web services; the second presents Semantic Web processes; and the third shows applications for the Semantic Web. The set of chapters approach standards that combine semantics and Web services technology according to three knowledge requirement levels: beginner, intermediate, and advanced. Liyang Yu's book, divided into four parts, presents the subjects in a similar fashion to that of Cardoso and Sheth's book. The first part of Liyang Yu's text deals with the understanding of, the justifications for and the potential value of the Semantic Web; the second part presents detailed technical aspects and technologies of the Semantic Web; the third part shows real examples and practical applications for the Semantic Web, and the final part documents the evolution of Web services towards Semantic Web services. Similarly to Cardoso & Sheth's edition, Liyang Yu's book is intended for developers, graduate and post-graduate students, and researchers, being of interest to communities involved in the application of Semantic Web services technology.

Both books aim at occupying the position of a Semantic Web textbook. Concerning the text edited by Cardoso & Sheth, the end of each chapter presents issues for discussion and suggestions for additional reading. As for Lyiang Yu's book, despite its being an introduction to the world of the Semantic Web, the comprehensiveness of its scope requires the reader to possess basic knowledge of HTML language, and solid knowledge of the XML and Java languages and Web servers. Along the chapters of her book – which does not call for linear reading – Liyang Yu gives examples of the integration of systems, applications and Web services; search tools and data mining applications, which are currently the Web's main practical applications.

The first part of Cardoso & Sheth's text, which deals with semantic Web services, starts with a discussion on the Semantic Web and its applications, and shows that heterogeneity on the Web has been known to the database distributed systems community for a long time. According to Cardoso & Sheth (2006), heterogeneity results from discrepancies on the meaning, interpretation and intended use of data; semantic heterogeneity considers the contents of an information item and the intended meaning. The solution suggested by the authors for semantic heterogeneity is to rely on the technological foundations of the semantic Web which, in short, tags documents with semantic metadata, which are in turn understood by tools for subsequent extraction, for instance, by means of ontologies. Chapters 2, 3 and 4 of the first part of Cardoso and Sheth's book present and discuss the evolutionary path of Web services towards semantic Web services: semantic annotation in Web services; broadening of the semantic scope of Web services standards; and methods for the discovery and publication of Web services. Chapter 5 suggests a methodology for the enlargement of the semantic description of a service with temporal properties, which enable one to obtain inferences on behavior across time.

The first part of Liyang Yu's book introduces the reader to the world of the semantic Web. The first chapter contends that the integration of systems, applications and Web services, as well as of search tools and data mining lack a software agent that is capable of large-scale processing on the Web. Similarly to Cardoso & Sheth, Liyang Yu believes that it the introduction of semantics in the Web is required, highlighting the importance of metadata for the tools. The second chapter of Liyang Yu's text shows the behavior of search tools and the changes required in adapting them to the world of the semantic Web, in addition to highlighting the benefits resulting from the evolution of said search tools. The second part of Liyang Yu's book is divided into four chapters and deals with the technical foundations of the semantic

Web: **RDF, Resource Description Framework** *Schema* **- RDF** *Schema*, **Web Ontology Language - OWL, taxonomy and ontology, and of assistant tools for the semantic Web, such as the validation of OWL ontologies.**

The second part of Cardoso and Sheth's book discusses the semantic Web processes. Thus, chapter 6 presents the ideas and principles that direct the choreography, that is, the modeling of the visible behavior of cross-service interactions. Chapter 7 explains how to introduce semantics in Web services standards by using WSDL – S, which is an extension of the specification of the Web Services Description Language – WSDL. Chapters 8 and 9 cover advanced topics and discuss the components of Web services based on non-functional properties by employing multi-purpose optimization techniques, in addition to presenting a general overview of the standardization and mapping of semantic Web processes.

The third part of both books presents applications and examples of the semantic Web in the real world. Cardoso & Sheth provide examples of semantic Web applications in the tourism, government, bioinformatics, and Web services fields. Chapter 10 of Cardoso and Sheth's text illustrates and describes the construction of an ontology for the e-tourism by using the OWL language, and works as a starting point for acquiring advanced OWL knowledge. Chapter 11 presents a pilot project developed with the intention of attaining semantic interoperability and the integration of semantic data in the governmental area. Chapter 12 discusses the application of Web processes and services, as well as the role of semantics in bioinformatics. This chapter is a basic literature for understanding the applications of semantic Web services in life sciences and associate bioinformatics. According to Sahoo (2006), although numerous Web services provide access to biologic resources, many such resources are separate computing tools, as interoperability among these resources is insignificant. This chapter describes computer genomics, computer proteomics and structural bioinformatics, in addition to presenting a case study in glycoproteomics. The penultimate chapter of Cardoso & Sheth's book deals with the project, development and implementation of services oriented by semantic business services for the commercialization of agricultural produce. The last chapter of Cardoso and Sheth's book presents the frameworks underlying the programmatic development of OWL ontologies, and discusses those most used by the developers' community, such as Jena, Protégé-OWL API and WonderWeb OWL API, all available for Java. This chapter provides details on how the semantic Web applications can be developed by using the Jena framework.

Liyang Yu's examples and applications for the semantic Web differ from those described in Cardoso and Sheth's book, and show search tools and ontologies, with the purpose of introducing readers to the semantic Web tools: RDF, RDF *schema* and OWL. Thus, chapter 7 describes Maryland Universty's Swoogle search project, a search tool for the semantic Web on the Web. Swoogle visits a large number of Web pages, and opens every, or almost every hyperlink on that page. Liyang Yu shows the architecture, the data flow, and examples of how to use Swoogle to find semantic documents on the Web. Chapter 8 deals with the Friend of a Friend – FOAF project, which intends to create a Web of pages of machine-readable descriptions of individuals, the links between them, and what they create and do. In this chapter, Liyang Yu discuses the concept and ideas behind the FOAF project, including project-related ontologies, and presents some examples, including how to create and publish your own FOAF document on the Web and insert it in your circle of friends. Liyang Yu, based on the practical examples of Swoogle and FOAF, underscores the need for a connection between the world of semantics and the world of the Web. At this point, like Cardoso & Sheth (2006), Liyang Yu argues that such connection by means of the tagging of Web pages. Chapter 9 provides examples of how to add semantics to Web documents. In the next chapter, the author returns to the semantic Web search tools, this time as an example of the use of metadata combined with semantic tagging.

The last part of Liyang Yu's book examines how Web services can benefit from the semantic Web, by focusing on the discovery of services. Thus, chapter 11 presents semantic Web services along with a detailed review of current Web service standards:. WSDL, Simple Object Access Protocol - SOAP and *Universal Description, Discovery, and Integration* - UDDI. It concludes that, in order to facilitate the automatic discovery, composition and monitoring of Web services, semantics must be added to current standards. Chapter 12 presents the idealizations and features of the OWL-S tagging language, which may be used to formally express the semantics of a Web service. On chapter 13, Liyang Yu presents two approaches to the introduction of semantics in Web services descriptions: one such approach consists of adding semantics to Web services, inserting semantic annotations in the current Web service standards, such as WSDL-S or UDDI; the second is a more comprehensive solution and employs a high-level ontology, OWL-S. One advantage of adopting WSDL-S is the reuse of currently available standards and tools, such as WSDL. By using OWL-S, an automatic agent will have enough information to discover, invoke, compose and monitor services, as any service can be described by using high-level ontologies. This chapter also shows how to map an OWL-S document for UDDI data frameworks, resulting in a semantically extended UDDI record, which works as a central repository, facilitating the automatic finding of requested Web services. An alternative proposal is necessary, however, due to the unavailability of public UDDI records. By using Java programming along with Jena Application Program Interfaces – API, chapter 14 of Liyang Yu book suggests the construction of a semantic Web services search tool that manages its own records and does not rely on public UDDI records. The last chapter of Liyang Yu's book summarizes previous chapters and suggests additional reading for those

willing to further their research on semantic Web and semantic Web services.

Both books offer a comprehensive view of the semantic Web and prepare the reader for the next round of standards, languages, and specifications, similarly to the development of the data query language in RDF SPARQL Protocol and RDF Query Language – SPARQL. This language is regarded as a relevant innovation for the Semantic Web, as it can be used in direct or mediated queries in several data sources.

Before concluding, it should be underlined that, although the fight for standards, languages, specifications and other techniques and technologies required by the Web's new evolutionary level is favorable to the Semantic Web proposal, there are specialists who disagree. Thus, Lévy (2006) believes that semantic interoperability is equivalent to the development of digital-base collective intelligence, contending that the Semantic Web does not tackle the problem of semantic interoperability as concept annotation in natural language is arbitrary, and also because the countless ontologies are incompatible.

Berners-Lee (2008) recently acknowledged that the Semantic Web idea as presented in 2001 somehow belonged to science fiction, as it was based on a still very distant future. People imagined that implementation of the Semantic Web would allow people to do everything that artificial intelligence-type systems could do. To Berners-Lee (2008), the Semantic Web is made possible by interoperability, such as the integration of data within and between companies, and the integration of scientific data, in addition to the possibility of querying the integrated data, like the Linked Data[2] initiative. However, since the Semantic Web is a set of diverse technologies that must be able to accomplish different tasks for different communities, the necessary developments are equally different. Likewise, the requirements for the life sciences community to be able to use its protein data in a Semantic Web environment must be different from the steps required to attain interoperability among the data repositories of libraries and museums.

Notes

1. Shadbolt (2006) states that in 1994 Tim Berners-Lee had already outlined a view of the Semantic Web.

2. http://linkeddata.org

Bibliographic references

Berners-Lee T, Hendler J, Lassila O. The semantic web: a new form of web content that is meaningful to computers will unleash a revolution of new possibilities.

Scientific American. May 2001: 34-43

Shadbolt N, Berners-Lee T, Hall W. The semantic Web revisited. IEEE Intelligent Systems. 2006; 21(3): 96-101.

Cardoso J, Sheth AP. The semantic Web and its applications. In: Semantic Web services, processes and applications. Nova Iorque: Springer; 2006: 3-33

Sahoo SS, Sheth A. Bioinformatics applications of Web services, Web process and role of semantics. In: Cardoso J, Sheth AP (editors), Semantic Web services, processes and applications. Nova Iorque: Springer; 2006: 306-22.

"Sir Tim Berners-Lee Talks About the Semantic Web", transcrição de entrevista em Podcast com Paul Miller. Feb. 2008. Available at: http://talis- odcasts.s3.amazonaws.com/twt20080207_TimBL.html. Accessed in: 21 Jan. 2009.

Lévy P. IEML: computational semantics in the service of collective intelligence. Otawa: CRC/FRSC; 2006: 9 (translated by Michele Healy). Available at: <http://www.ieml.org/IMG/pdf/visionieml-Initiativeen. pdf>. Accessed in: 8 Apr. 2007.