

Semantic Web services, processes and applications

Jorge Cardoso and Amit P. Sheth (Eds.)

Semantic Web and semantic Web services

Liyang Yu

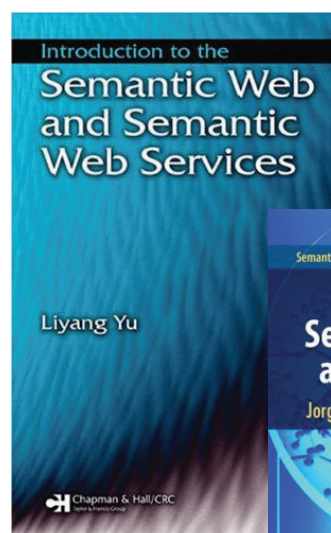
DOI: 10.3395/receis.v3i1.246pt

Laura Cristina Simões Viana

Coordenação de Recursos Humanos em Pesquisa da Vice-Presidência de Pesquisa e Desenvolvimento Tecnológico da Fundação Oswaldo Cruz, Rio de Janeiro, Brasil
laura@fiocruz.br

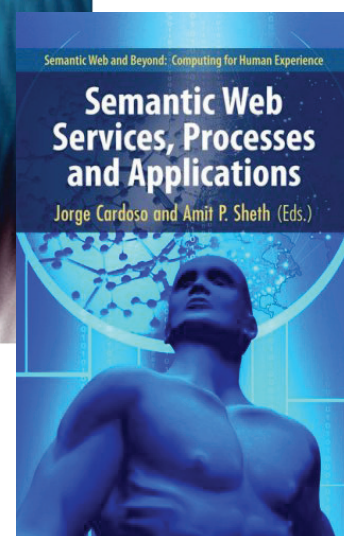
Tim Berners-Lee criou, em 1989, as ferramentas básicas necessárias ao funcionamento da Web, tal como a conhecemos atualmente: a linguagem HyperText Markup Language – HTML ou linguagem de marcação de hipertexto; o protocolo HyperText Transfer Protocol – HTTP ou protocolo de transferência de hipertexto, e o sistema de localização de objetos na Web, Universal Resource Locator – URL ou localizador universal de recursos. Dose anos após¹, com a Web funcionando plenamente, Berners-Lee publicou um artigo original com uma proposta desafiadora de “Web semântica” (Berners-Lee, 2001).

De acordo com essa proposta, a Web semântica seria em uma ampliação da web atual, na qual a informação passaria a ter um significado definido, viabilizando o trabalho cooperativo entre homens e máquinas. Assim, a Web atual, que é formada, basicamente, por documentos apresentados por computadores e lidos pelo homem, passaria a incluir dados e informações que seriam automaticamente manipulados por agentes inteligentes e utilitários. Naquela época, Berners-Lee (2001) argumentou que o progresso da Web semântica necessitaria do desenvolvimento de uma linguagem que expressasse dados e regras de raciocínio sobre os dados, e que também permitisse a exportação para Web das regras de qualquer sistema de representação do conhecimento. Até então, encontravam-se desenvolvidas duas tecnologias funda-



New York,
Springer, 2006

ISBN
978-0387302393



USA, Chapman & Hall/
CRC, 2007

ISBN
978-1584889335

mentais para a realização da Web semântica: a linguagem de marcação eXtensible Markup Language – XML e a família de especificações Resource Description Framework – RDF, esta última para descrição ou modelagem de informação implementada em recursos da web.

Em 2006, embora os padrões Web que expressam significado compartilhado estivessem em franco desenvolvimento, Shadbolt (2006) publicou outro artigo intitulado “The semantic Web revisited”, no qual afirmava que a Web semântica permaneceria, basicamente, no plano das idéias até que esses padrões fossem bem estabelecidos e acordados. Shadbolt (2006) ressaltou que, a exemplo do protocolo HTTP, cujo uso pioneiro pela comunidade da área da física abriu caminho para o sucesso da Web, o uso crescente de ontologias pela comunidade científica eletrônica (“e-science”) também poderia levar a Web semântica ao sucesso tal como hoje conhecemos a Web. Segundo o autor, além de fatores sociais e de decisões de projeto, parte do sucesso da Web semântica reside na sequência de especificações (Universal Resource Identifier - URI, HTTP, RDF, ontologias, etc.) e de registros (esquema URI, conteúdos de Internet do tipo Multipurpose Internet Mail Extensions – MIME ou extensões multifunção para mensagens de Internet), os quais fornecem meios para que construções como uma ontologia derivem significado de um identificador URI.

O livro de Jorge Cardoso & Amit Sheth, publicado em 2006, e o livro de Liyang Yu, publicado em 2007, são textos práticos, que abordam a Web semântica com toda a riqueza dos padrões, linguagens, aplicações e processos que o tema provoca. Amit Sheth fez seu Doutorado em Ciências da Computação e Informação na Ohio State University e foi professor da Univesidade da Geórgia entre 1994 e 2006, tendo publicado mais de 250 artigos revisados pelos pares e oito livros; atualmente, é professor de Engenharia e Ciência da Computação e diretor do centro Kno.e.sis, ambos na Wrigt State University. Jorge Cardoso fez seu Doutorado em Ciências da Computação, com Amit Sheth, na Univesidade da Geórgia e trabalha atualmente no Departamento de Pesquisa da empresa alemã SAP, além de manter seu vínculo de professor com a Universidade da Madeira, em Portugal. Cardoso tem mais de 80 artigos publicados em periódicos revisados pelos pares nas áreas de sistemas de gestão de workflow, Web semântica e editou três livros sobre Web semântica e serviços Web. Liyang Yu fez seu Doutorado na Ohio State University e, atualmente, trabalha para a Delta Technologies, nos Estados Unidos.

O livro editado por Jorge Cardoso e Amit Sheth é uma coletânea de contribuições de pesquisadores de universidades e empresas sobre semântica, organizada em três partes: a primeira trata dos serviços Web semânticos, a segunda apresenta processos da Web semântica e a última parte do livro mostra aplicações da Web semântica. O conjunto dos capítulos aborda padrões que combinam semântica e tecnologia de serviços Web segundo três níveis de exigências de conhecimentos: iniciante, intermediário e avançado. Organizado em quatro partes, o livro de Liyang Yu apresenta os assuntos de

modo semelhante ao livro de Cardoso e Sheth. A primeira parte do texto de Liyang Yu é sobre o entendimento, as justificativas e o valor potencial da Web semântica; a segunda detalha aspectos técnicos e tecnologias da Web semântica, a terceira parte do livro apresenta exemplos reais e aplicações práticas de Web semântica e, ao final, mostra a evolução de serviços Web para serviços Web semânticos. A exemplo da edição de Cardoso e Sheth, o livro de Liyang Yu destina-se a desenvolvedores, estudantes de graduação e pós-graduação, e pesquisadores, atendendo às comunidades interessadas nas aplicações de tecnologia de serviços Web semânticos.

Ambos os livros almejam o posto de livro texto sobre Web semântica. No caso do texto editado por Cardoso e Sheth, ao final de cada capítulo constam questões para discussões e sugestões de leitura adicional. O livro de Liyang Yu, embora tenha um caráter introdutório ao mundo da Web semântica, a abrangência da abordagem exige do leitor conhecimentos básicos sobre a linguagem HTML e sólidos conhecimentos das linguagens XML, Java e de servidores Web. Ao longo dos capítulos do seu livro, cuja leitura não precisa ser linear, Liyang Yu apresenta exemplos de integração de sistemas, aplicativos e serviços Web; de máquinas de buscas e de aplicações de mineração de dados, que são as principais utilizações da Web, atualmente.

A primeira parte do texto de Cardoso e Sheth, que aborda os serviços Web semânticos, se inicia com uma discussão sobre a Web semântica e suas aplicações, e mostra que a heterogeneidade na Web é conhecida de longa data da comunidade de sistemas distribuídos de bancos de dados. Segundo Cardoso e Sheth (2006), a heterogeneidade surge de discordâncias sobre o significado, a interpretação e o uso pretendido dos dados; a heterogeneidade semântica leva em conta o conteúdo de um item da informação e o significado pretendido. A solução proposta pelos autores para a heterogeneidade semântica é basear-se nos fundamentos tecnológicos da Web semântica que, em resumo, marca os documentos com metadados semânticos, os quais são compreendidos por máquinas, para posterior extração, por exemplo, através de ontologias. Os capítulos dois, três e quatro da primeira parte do livro de Cardoso e Sheth apresentam e discutem os caminhos da evolução de serviços Web para serviços Web semânticos: anotação semântica em serviços Web, ampliação semântica dos padrões de serviços Web, e esquemas para descoberta e publicação de serviços Web. O quinto capítulo do livro propõe uma metodologia para expandir a descrição semântica de um serviço com propriedades temporais, que permitam a inferência sobre o comportamento ao longo do tempo.

A primeira parte do livro de Liyang Yu introduz o leitor no mundo da Web semântica. O primeiro capítulo argumenta que a integração de sistemas, de aplicativos e de serviços Web, assim como as máquinas de buscas e a mineração de dados se ressentem de um agente de software que seja capaz de realizar o processamento em larga escala, na Web. Do mesmo modo que Cardoso e Sheth, Liyang Yu entende que é necessário introduzir se-

mântica na Web, destacando a importância de metadados para as máquinas. O segundo capítulo do texto de Liyang Yu mostra o funcionamento das máquinas de busca e as mudanças necessárias para adaptá-las ao mundo da Web semântica, além de destacar os benefícios desta evolução das máquinas de busca. A segunda parte do livro de Liyang Yu está estruturada em quatro capítulos e trata dos fundamentos técnicos da Web semântica: RDF, Resource Description Framework *Schema* – RDF *Schema*, Web Ontology Language - OWL, taxonomia e ontologia, e das ferramentas acessórias para a Web semântica, como a validação de ontologias OWL.

A segunda parte do livro de Cardoso e Sheth aborda os processos da Web semântica. Assim, o capítulo seis apresenta as idéias e os princípios que norteiam a coreografia, ou seja a modelagem do comportamento visível das interações entre serviços. O capítulo sete mostra como introduzir semântica em padrões de serviços Web, utilizando WSDL-S, que é uma extensão da especificação da linguagem Web Services Description Language – WSDL. Os capítulos oito e nove cobrem tópicos avançados e discutem a composição de serviços Web com base em propriedades não funcionais, utilizando técnicas de otimização multiobjetivo, além de apresentar um quadro genérico de harmonização e mapeamento de processos Web semânticos.

A terceira parte de ambos os livros apresenta aplicações e exemplos da Web semântica no mundo real. Cardoso e Sheth ilustram aplicações de Web semântica nas áreas do turismo, governo, bioinformática e serviços Web. O capítulo 10 do texto de Cardoso e Sheth ilustra e descreve a construção de uma ontologia para o turismo “eletrônico” (e-turismo) utilizando a linguagem OWL, e serve como ponto de partida para aquisição de conhecimentos avançados em OWL. O capítulo 11 apresenta um projeto piloto desenvolvido com o propósito de alcançar interoperabilidade semântica e integração de dados semânticos na área governamental. O décimo - segundo capítulo do livro discute a aplicação de serviços e processos Web, bem como o papel da semântica na bioinformática. Esse capítulo é uma leitura básica para a compreensão das aplicações de serviços Web semânticos nas ciências biológicas e a bioinformática associada. Segundo Sahoo (2006), embora existam numerosos serviços Web que oferecem acesso a recursos biológicos, muitos destes recursos são ferramentas computacionais isoladas, pois a interoperabilidade entre estes recursos é mínima. Esse capítulo descreve a genômica computacional, a proteômica computacional e a bioinformática estrutural, além de apresentar um estudo de caso em glicoproteômica. O penúltimo capítulo do livro de Cardoso e Sheth trata do projeto, desenvolvimento e implementação de sistemas orientados por serviços de negócios semânticos para o marketing de produtos agrícolas. O último capítulo do texto de Cardoso e Sheth apresenta estruturas que suportam o desenvolvimento programático de ontologias OWL, e discute aquelas mais utilizadas pela comunidade de desenvolvedores, como Jena, Protégé-OWL API e WonderWeb OWL API, todas disponíveis para a linguagem Java. Esse capítulo detalha como as aplicações

Web semântica podem ser desenvolvidas utilizando a estrutura Jena.

As aplicações e os exemplos de Web semântica apresentados no livro de Liyang Yu diferem daqueles relatados no livro de Cardoso e Sheth e mostram máquinas de busca e ontologias, com o objetivo de familiarizar o leitor com as ferramentas da Web semântica: RDF, RDF *schema* e OWL. Assim, o capítulo sete detalha o projeto de pesquisa Swoogle, da Universidade de Maryland, que é uma máquina de busca para a Web semântica, na Web. Swoogle visita um número grande de páginas na Web e abre todos ou quase todos os *hyperlinks* da página visitada. Liyang Yu mostra a arquitetura, o fluxo de dados e exemplos de como utilizar Swoogle para encontrar documentos semânticos na Web. O oitavo capítulo trata do projeto *Friend of a Friend* – FOAF, que se propõe criar uma Web de páginas que descrevem pessoas, os elos entre as pessoas e aquilo que essas pessoas criam e fazem, e que podem ser lidas por máquinas. Neste capítulo, Liyang Yu discute o conceito e as idéias do projeto FOAF, inclusive as ontologias que se relacionam com o projeto, e apresenta alguns exemplos, inclusive como criar e publicar seu próprio documento FOAF na Web e como inseri-lo no círculo de amigos. Com base nos exemplos práticos Swoogle e FOAF, Liyang Yu enfatiza a necessidade de conexão entre o mundo da semântica e o mundo da Web. Neste ponto, assim como Cardoso e Sheth (2006), Liyang Yu argumenta que esta conexão se dá através da marcação das páginas Web e apresenta, no nono capítulo, exemplos de como adicionar semântica aos documentos Web. No capítulo seguinte, o autor retorna às máquinas de busca Web semânticas, desta vez como um exemplo da utilização de metadados adicionados por marcação semântica.

Na última parte de seu livro, Liyang Yu investiga como os serviços Web podem beneficiar-se da Web semântica, concentrando-se na descoberta de serviços. Assim, o décimo - primeiro capítulo apresenta os serviços Web semânticos, junto com uma revisão detalhada dos padrões atuais para serviços Web: WSDL, Simple Object Access Protocol - SOAP e *Universal Description, Discovery, and Integration* - UDDI, e conclui que, para facilitar a descoberta automática, a composição e o monitoramento de serviços Web, a semântica deve ser adicionada aos padrões atuais. No décimo – segundo capítulo, o autor apresenta as idealizações e as características da linguagem de marcação OWL-S, que pode ser utilizada para expressar formalmente a semântica de um serviço Web. No décimo – terceiro capítulo, Liyang Yu mostra duas abordagens para introdução de semântica na descrição dos serviços Web: uma das abordagens consiste em adicionar semântica em serviços Web inserindo anotações semânticas nos padrões atuais de serviço Web, tais como WSDL-S ou UDDI; a outra solução é mais completa e utiliza uma ontologia de alto nível, OWL-S. A vantagem do caminho que adota WSDL-S é reutilizar padrões e ferramentas atualmente disponíveis, como WSDL. Com a utilização de OWL-S, um agente automático terá informação suficiente para descobrir, invocar, compor e monitorar o serviço, pois qualquer serviço pode ser des-

crito utilizando-se ontologias de alto nível. Este capítulo ainda mostra como mapear um documento OWL-S para estruturas de dados UDDI, resultando em um registro UDDI estendido semanticamente, que funciona como um repositório centralizado, facilitando a descoberta automática de serviços Web requisitados. Entretanto, com a indisponibilidade de registros UDDI públicos, surge a necessidade de uma proposta alternativa. Utilizando programação Java, junto com Jena Application Program Interfaces – API, Liyang Yu propõe, no décimo – quarto capítulo, a construção de uma máquina de busca de serviço Web semântico, que gerencia seus próprios registros e não depende de registros UDDI públicos. O último capítulo do livro de Liyang Yu resume os capítulos anteriores e sugere leitura adicional para aqueles que quiserem continuar os estudos sobre Web semântica e serviços Web semânticos.

Ambos os livros fornecem uma visão abrangente da Web semântica e preparam o leitor para a próxima rodada de padrões, linguagens e especificações, a exemplo do desenvolvimento da linguagem de consulta a dados em RDF SPARQL Protocol and RDF Query Language – SPARQL. Esta linguagem é considerada uma inovação relevante para Web Semântica, pois pode ser utilizada em consultas diretas, ou via mediação, a diversas fontes de dados.

Antes de concluir, convém lembrar que, embora a disputa por padrões, linguagens, especificações e demais técnicas e tecnologias necessárias ao próximo estágio evolutivo da Web seja favorável à proposta contida na Web semântica, outros especialistas discordam. Assim, Lévy (2006), entende que a interoperabilidade semântica equivale ao desenvolvimento da inteligência coletiva de base digital, e argumenta que a Web semântica não resolve a questão da interoperabilidade semântica porque a notação de conceitos em linguagem natural é arbitrária e também porque as inúmeras ontologias são incompatíveis.

Concluindo, recentemente, Berneres-Lee (2008) admitiu que a visão da Web semântica apresentada em 2001 era de alguma forma, ficção científica, pois estava baseada em um futuro ainda muito distante. Imaginava-se que a Web semântica seria implementada e as pessoas fariam tudo que os sistemas do tipo inteligência artificial poderiam fazer. Para Berners-Lee(2008), a prática da Web semântica acontece a partir da interoperabilidade, a exemplo da integração de dados intra e entre empresas e da integração de dados científicos, além da possibilidade

de consultas aos dados integrados, a exemplo da iniciativa Linked Data². Entretanto, como a Web semântica é um conjunto variado de tecnologias, que deverá ser capaz de realizar tarefas diferentes para comunidades diferentes, os desenvolvimentos necessários são igualmente diferentes. Assim, as necessidades da comunidade de ciências da vida para que ela seja capaz de utilizar seus dados sobre proteínas em um ambiente de Web semântica deverão ser diferentes dos passos necessários para se alcançar a interoperabilidade entre repositórios de dados de bibliotecas e de museus.

Notas

1. Shadbolt (2006) informa que em 1994, Tim Berners-Lee já havia articulado uma visão da Web Semântica.
2. <http://linkeddata.org>

Referências bibliográficas

- 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*. Mai. 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. Fev. 2008. Disponível em: http://talos-odcasts.s3.amazonaws.com/twt20080207_TimBL.html. Acessado em: 21 Jan. 2009.
- Lévy P. *IEML: computational semantics in the service of collective intelligence*. Ottawa: CRC/FRSC; 2006: 9 (translated by Michele Healy). Disponível em: <<http://www.ieml.org/IMG/pdf/visionieml-Initiativeen.pdf>>. Acessado em: 8 Apr. 2007. 