

Tecnologias da Web Semântica na arquitetura da informação*

Caio Saraiva Coneglian

Doutorando em Ciência da Informação. Mestre em Ciência da Informação, Universidade Estadual Paulista. Bacharel em Ciência da Computação, Centro Universitário Eurípides de Marília. Marília-SP – Brasil.
caio.coneglian@gmail.com
orcid.org/0000-0002-6126-9113

Sandra Milena Roa-Martinez

Doctoranda em Ciência da Informação, Universidade Estadual Paulista. Magister en Ingeniería énfasis Electrónica, Universidad del Valle. Especialista en Redes de Comunicación, Universidad del Valle. Ingeniera de Sistemas, Universidad Industrial de Santander. Profesora titular tiempo completo, Universidad del Cauca. Popayán – Colombia. smroa@unicauca.edu.co
orcid.org/0000-0002-2271-6101

Ana Maria Jensen Ferreira da Costa Ferreira

Doutoranda em Ciência da Informação. Mestre em Ciência da Informação. Arquivologia, Universidade Estadual Paulista. Marília-SP – Brasil. anajfferreira@gmail.com
orcid.org/0000-0003-2484-303X

Silvana Aparecida Borsetti Gregório Vidotti

Doutora em Educação, Universidade Estadual Paulista. Mestre em Ciência da Computação, Universidade de São Paulo. Licenciatura em Matemática. Professora Assistente Doutora, Universidade Estadual Paulista. Marília-SP – Brasil.
vidotti@marilia.unesp.br
orcid.org/0000-0002-4216-0374

José Eduardo Santarem Segundo

Doutor em Ciência da Informação. Mestre em Ciência da Informação, Universidade Estadual Paulista. Superior de Tecnologia em Processamento de Dados, Universidade de Marília. Professor Assistente Doutor, Universidade de São Paulo. São Paulo – Brasil.
santarem@usp.br
orcid.org/0000-0003-3360-7872

Resumo

Os estudos da Arquitetura da Informação e da Web Semântica se mostram cada vez mais necessários, pela exigência de compreender o contexto da informação, tanto da perspectiva dos usuários humanos, quanto dos usuários não-humanos/máquinas. Esses dois campos de estudos contribuem para aprimorar a forma como os conteúdos informacionais são apresentados e utilizados nos ambientes digitais. Assim, este trabalho identificou e analisou as relações entre os sistemas da Arquitetura da Informação e os recursos da Web Semântica, apresentando como cada sistema da Arquitetura da Informação pode ser aprimorado com a incorporação de tecnologias da Web Semântica. Para tal, utilizou-se uma metodologia qualitativa, de cunho exploratório e descritivo, relacionando os cinco sistemas da Arquitetura da Informação com os conceitos, as tecnologias e as aplicações da Web Semântica. Como resultados, foi possível verificar que todos os sistemas da Arquitetura da Informação podem ser enriquecidos com as tecnologias da Web Semântica. Desta forma, concluiu-se que a aplicação de tecnologias da Web Semântica nos sistemas da Arquitetura da Informação contribui para estruturar semanticamente os ambientes informacionais digitais.

Palavras-chave: arquitetura da informação, web semântica, sistemas da arquitetura da informação, conceitos e tecnologias da web semântica, ambientes informacionais digitais.

Cómo citar este artículo: Coneglian, C. S., Roa-Martinez, S. M., Ferreira, A. M. C. F., Vidotti, S. A. B. G., & Santarem-Segundo, J. E. (2019). Tecnologias da Web Semântica na arquitetura da informação. *Revista Interamericana de Bibliotecología*, 42(1), 23-35. doi: 10.17533/udea.rib.v42n1a03

Recibido: 2017/06/07/ **Aceptado:** 2018/08/26

* O presente texto é um progresso do projeto “Arquitetura da Informação e Experiência do Usuário em Ambientes Informacionais Digitais” de la línea Información e Tecnología del Programa de Pósgraduação em Ciência da Informação de la Universidade Estadual Paulista (Unesp). financiado pelo Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq).



Semantic Web Technologies in Information Architecture

Abstract

The researches of Information Architecture and Semantic Web are increasingly necessary, due to the need to have an understanding of the context of the information, both from the perspective of human users and non-humans/machines users. These two study fields contribute to improving the way informational content is presented and used in digital environments. Thus, this paper identified and analyzed the relationships between Information Architecture systems and Semantic Web resources, presenting how each Information Architecture system can benefit from the incorporation of Semantic Web technologies. For this, a qualitative methodology was used, with a descriptive and exploratory approach, relating the five systems of Information Architecture with the concepts, technologies and applications of the Semantic Web. As results, it was possible to verify that all systems of Information Architecture can be enriched with Semantic Web technologies. In this way, it was concluded that the application of Semantic Web technologies in Information Architecture systems contributes to a semantic structure of digital information environments.

Keywords: Information architecture, semantic web, information architecture systems, semantic web concepts and technologies, digital information environments.

Tecnologías de la Web Semántica en arquitectura de la información

Resumen

Los estudios de Arquitectura de la Información y Web Semántica son cada vez más necesarios, debido a la exigencia de comprender el contexto de la información desde la perspectiva de usuarios humanos y usuarios no humanos/máquinas. Estos dos campos de estudio contribuyen a mejorar la forma en que son presentados y utilizados los contenidos en ambientes informacionales digitales. Por lo tanto, este trabajo identificó y analizó las relaciones entre los sistemas de la Arquitectura de la Información y los recursos de la Web Semántica, mostrando cómo cada sistema de la Arquitectura de la Información puede mejorarse con el uso de las tecnologías de la Web Semántica. Para ello, se utilizó una metodología cualitativa, de carácter exploratorio y descriptivo, relacionando los cinco sistemas de Arquitectura

de la Información con los conceptos, tecnologías y aplicaciones de la Web Semántica. Como resultados, fue posible verificar que todos los sistemas de la Arquitectura de la Información pueden ser enriquecidos con las tecnologías de la Web Semántica. De esta forma, se concluyó que la aplicación de las tecnologías de la Web Semántica en los sistemas de la Arquitectura de la Información contribuye para una estructura semántica de los ambientes informacionales digitales.

Palabras clave: arquitectura de la información, web semántica, sistemas de arquitectura de la información, conceptos y tecnologías de la web semántica, ambientes informacionales digitales.

1. Introdução

A World Wide Web (Web) e as formas de acesso e uso de informações apresentadas nos ambientes informacionais digitais tem sido um campo de estudo relevante devido às constantes transformações tecnológicas, sociais e culturais. Nesse sentido, surge a necessidade de interseções entre os diferentes arcabouços teóricos, bem como de técnicas e de tecnologias a fim de compreender e contribuir na interação dos indivíduos com as informações disponíveis na Web.

Neste contexto, com as Tecnologias de Informação e Comunicação (TIC) cada vez mais presentes no cotidiano das pessoas, surgem novos desafios para diferentes áreas de conhecimento que se preocupam com soluções para abordar as problemáticas relativas a organização dos objetos digitais em ambientes informacionais buscando promover eficiência e eficácia no momento de acesso e uso de informações.

Apontamos, neste trabalho, dois campos de estudos que estão avançando com pesquisas neste quesito: a Arquitetura da Informação (AI) e a Web Semântica (WS), que ganham destaque no que tange aos processos relativos às informações digitais na Web.

A Arquitetura da Informação tem como objetivo o planejamento de ambientes digitais procurando atender os requisitos relacionados com o contexto, conteúdo e usuário; enquanto que a Web Semântica foca na criação de estruturas que possibilitem a compreensão dos conteúdos digitais pelos usuários e pelos agentes computacionais.

Desta forma, questiona-se: como e quais tecnologias da Web Semântica podem ser inseridas nos sistemas da Arquitetura da Informação Digital de modo a enriquecer o planejamento e o desenvolvimento de ambientes informacionais digitais?

Na busca de solucionar tal questão, o presente trabalho objetivou identificar e analisar as relações entre os sistemas da Arquitetura da Informação e os recursos da Web Semântica, apresentando como cada sistema da Arquitetura da Informação pode-se beneficiar da incorporação de tecnologias da Web Semântica. Para ilustrar, são apresentadas as aplicações semânticas no contexto da Arquitetura da Informação.

2. Arquitetura da informação

Rosenfeld, Morville e Arango (2015) consolidam a expressão Arquitetura da Informação, associando-a especificamente aos ambientes Web e aproximando-a de forma inquestionável ao domínio da Ciência da Informação e da Biblioteconomia. Por quanto, estruturar, representar, organizar e rotular, conseqüentemente, são tarefas características da AI.

A partir da publicação do livro *Information Architecture for the World Wide Web* (Rosenfeld & Morville, 1998), pesquisas relacionadas foram realizadas no âmbito da Ciência da Informação, por exemplo, na aplicação em websites organizacionais, bibliotecas digitais, repositórios institucionais, blogs e redes sociais. Posteriormente, tais autores conceituaram a AI como:

O desenho estrutural de ambientes informacionais compartilhados; A combinação de sistemas de organização, rotulagem, busca e navegação em websites e intranets; [...] Uma disciplina emergente e comunidade de prática focada em trazer princípios de design e arquitetura para o ambiente digital. (Rosenfeld, Morville & Arango, 2015, p. 24, tradução)

Nos conceitos apresentados, percebe-se a aplicação dos princípios de design e de arquitetura no projeto de ambientes informacionais digitais, que norteiam o conjunto de sistemas que, metaforicamente, compõem a anatomia da AI e apontam para a estrutura de um ambiente informacional digital. Esses sistemas são de:

Organização, Rotulagem, Busca e Navegação; os autores supracitados apresentam ainda outros elementos como os vocabulários controlados, tesouros, metadados e ontologias que são considerados por Oliveira e Vidotti (2012) como um novo sistema, o sistema de Representação.

Assim, baseados em Rosenfeld e Morville (1998), Vidotti e Sanches (2004) apontam que:

A Arquitetura da Informação é constituída por sistemas de organização, navegação, rotulagem e busca, visando à criação de estruturas digitais que priorizam a organização descritiva, temática, representacional, visual e navegacional de informações, em consonância com o conteúdo, o contexto e o usuário, com objetivos bem definidos, adequando assim o dimensionamento e o direcionamento dos serviços e dos produtos informacionais aos usuários potenciais. (p. 2)

Nesse cenário, o Sistema de Organização é responsável pela estruturação dos conteúdos informacionais em esquemas e estruturas que melhor satisfaçam às necessidades do usuário na localização de informação, em consonância com os sistemas de navegação e de busca. Os esquemas e as estruturas de organização devem ser representativos para as comunidades no contexto da organização do conhecimento.

O Sistema de Rotulagem é responsável pela representação do conteúdo informacional estruturado, por exemplo, em categorias significativas para o usuário, por meio de rótulos textuais e/ou iconográficos, que devem estar associados com os aspectos cognitivos, linguísticos e/ou emocionais do usuário.

O Sistema de Navegação busca promover os diferentes caminhos de acesso ao conteúdo informacional e contemplar as experiências do usuário na busca de informações facilitando a interação do usuário no ambiente informacional.

A localização das informações, além da navegação, é contemplada pelo Sistema de Busca que visa atender as necessidades informacionais do usuário de forma mais direta.

Os metadados e os instrumentos de organização do conhecimento, tais como: vocabulários controlados,

tesauros, taxonomias, folksonomias e ontologias são abordados no Sistema de Representação de modo a propiciar a interação do usuário com o sistema e a interoperabilidade entre diferentes sistemas.

Neste contexto, o Sistema de Representação por meio do uso de ontologias pode aprimorar a interoperabilidade, ao fornecer um conhecimento compartilhado e formal capaz de possibilitar uma melhor comunicação entre diferentes sistemas. Pesquisas como a de Martin et al. (2007) e Grasic e Podgorelec (2010) expressam como as ontologias se constituem em elementos importantes para a promoção da interoperabilidade e da comunicação entre diferentes serviços e aplicações.

Atualmente, os estudos da Arquitetura da Informação têm como foco os ambientes informacionais digitais utilizados tanto pelos usuários humanos quanto pelos não-humanos/ máquinas, e nesse contexto, a Web Semântica apresenta tecnologias para a estruturação e a recuperação de conteúdo com significado.

3. Web Semântica

A Web Semântica desde a sua concepção, no início dos anos de 2000, tem influenciado uma série de serviços e aplicações de ambientes informacionais digitais. Visando tornar a Web uma plataforma capaz de promover a compreensão dos conteúdos por agentes computacionais, a WS foi proposta como uma extensão da Web, não tendo a intenção de ser uma substituta desta (Berners-Lee, Hendler & Lassila, 2001).

Vale destacar que há uma separação entre os conceitos, as tecnologias e as aplicações da Web Semântica. Neste âmbito, a partir das concepções teóricas delineadas por Berners-Lee et al. (2001), a Ciência da Informação e a Ciência da Computação buscaram definir um arcabouço teórico comum e desenvolver tecnologias e aplicações em ambientes informacionais digitais visando materializar a WS.

Neste sentido, algumas tecnologias possuem papel central na implementação de propostas da WS, com destaque para: *eXtensible Markup Language* (XML), *Resource Description Framework* (RDF), *Web Ontology Language* (OWL), *SPARQL Protocol and RDF Query Language* (SPARQL) e *Semantic Web Rule Language* (SWRL).

A XML, o RDF e a OWL contemplam especialmente formas de descrever e de representar uma determinada informação ou contexto. Enquanto que, a linguagem de consulta SPARQL possibilita a recuperação da informação e a SWRL é uma linguagem baseada em regras que permite fazer inferências.

A XML pode ser compreendida como a sintaxe utilizada para a disponibilização das informações, que permite agregar semântica aos documentos e possibilita que cada aplicação crie uma interpretação da marcação atribuída ao conteúdo. Assim, “Esta abordagem amplia significativamente as possibilidades do uso das linguagens de marcação, entre elas a capacidade de definir Metadados [...]” (Campos, Santachè & Teixeira, 1999, p. 51).

Por outro lado, o RDF pode ser entendido como o modo de associar, interligar e representar os dados. Segundo o World Wide Web Consortium (2014), o RDF é modelo padrão para intercâmbio de dados na Web e atua como uma linguagem para representar informações visando fornecer a interoperabilidade dos dados e contribuindo, assim, com a recuperação das informações dos recursos na Web.

No tocante às ontologias, a OWL permite a inserção de características semânticas que possibilitam a compreensão de um contexto por um agente computacional. Destaca-se que a linguagem OWL é reconhecida, como o principal padrão em linguagens para ontologia e recomendada pelo consórcio W3C para a construção de ontologias (Santarem -Segundo, 2010).

Com relação à recuperação dos dados, o SPARQL apresenta-se como uma função central, que permite a recuperação dos dados em formatos compatíveis com a WS, especialmente o RDF e a OWL. Esse protocolo permite realizar inferências dos dados, pelas associações entre as triplas RDF, ademais, provê formas de verificar informações como a quantidade e a soma, dentre outros operadores, no momento que são realizadas as consultas (Ducharme, 2013).

Ainda, referente a possibilidade de realização de inferências, a SWRL é uma linguagem, recomendada pelo W3C, que se baseia na OWL e utiliza um conjunto de

axiomas para possibilitar a construção das regras lógicas (Horrocks et al., 2004).

As tecnologias citadas anteriormente foram desenvolvidas visando tornar implementável e materializável a Web Semântica. O uso dessas tecnologias culminou no desenvolvimento de aplicações capazes de aprimorar a interface e a representação dos conteúdos construídos para as pessoas e os sistemas computacionais. Como exemplos temos a OWL e o RDF como tecnologias que possibilitam materializar a Web Semântica, e o *Linked Data* como uma das aplicações que incorporaram as tecnologias da Web Semântica para tornar real esta materialização.

Ainda, as tecnologias da Web Semântica têm sido utilizadas no projeto da Arquitetura da Informação para o desenvolvimento de ambientes informacionais digitais. Destacamos os trabalhos de Hinkelmann, Maïse e Thönssen (2013), que por meio de uma ontologia conecta os objetos informacionais numa arquitetura empresarial, e os trabalhos de Santos, Bax e Kalra (2010), de Al-Sudairy e Vasista (2011) e de Gurupur e Tanik (2012), que apresentam arquiteturas da informação com incorporação de tecnologias da Web Semântica para a integração de dados.

Vale ressaltar que, no que tange a relação entre a Web Semântica e a Arquitetura da Informação, serão discutidos e apresentados, a seguir, como os conceitos, as tecnologias e as aplicações da Web Semântica estão contribuindo com a Arquitetura da Informação, e como essas tecnologias se relacionam com os sistemas da AI.

A partir dos pressupostos teóricos apontados acerca da AI e da WS, na próxima seção se expõem os procedimentos metodológicos utilizados neste trabalho.

4. Procedimentos metodológicos

O presente trabalho foi desenvolvido seguindo uma metodologia qualitativa, de cunho exploratório e descritivo. Esse processo foi realizado relacionando os cinco sistemas da Arquitetura da Informação descritos anteriormente (sistema de organização, sistema de navegação, sistema de rotulagem, sistema de busca e sistema de representação), com os conceitos, as tecnologias e as aplicações da Web Semântica.

Para explanar como a WS se relaciona com cada sistema da AI, foram consideradas as tecnologias apontadas pelo W3C (XML, RDF, RDF Schema, OWL, SPARQL e SWRL), bem como algumas aplicações que utilizam tecnologias da WS e que são objetos de estudo da Ciência da Informação (*Knowledge Graph*, *Rich Snippets*, *Linked Data*, *Explorator* e *Europeana Data Model – EDM*).

Cada um dos cinco sistemas da Arquitetura da Informação foi analisado sob as perspectivas dos conceitos e das tecnologias da Web Semântica apresentados anteriormente. Posteriormente, a cada sistema foi associada e discutida uma aplicação que utiliza essa(s) tecnologia(s) da Web Semântica.

5. Resultados e discussões

A seguir são apresentados os resultados e as discussões alcançadas enfocando a WS (conceitos, tecnologias e aplicações) e os sistemas da AI (organização, rotulagem, navegação, busca e representação).

5.1. Sistema de organização

No Sistema de Organização para estruturar os conteúdos informacionais (Esquema/Estrutura) podem ser utilizadas as seguintes tecnologias da Web Semântica: XML, RDF, RDF Schema, OWL, SPARQL e SWRL, e que podem ser percebidas na aplicação do *Knowledge Graph* no Google.

Neste sistema, pode-se utilizar os esquemas de dados estruturados em XML, RDF, RDF Schema e as ontologias construídas em OWL, com relações explícitas capazes de possibilitar que um sistema computacional identifique os relacionamentos entre os dados e as necessidades dos usuários. Ainda, os esquemas de dados estruturados com semântica formal, possibilitam a criação de sistemas de organização dinâmicos, tornando a identificação das relações mais significativas.

Adicionalmente, para desenvolver estruturas dinâmicas de organização, considerando o contexto dos usuários, são necessárias as consultas por meio da linguagem SPARQL, que tem a função de descobrir as relações existentes, bem como, de encontrar uma forma de esquematar as informações em um sistema de acordo com o contexto do usuário e o contexto dos dados, que

são descobertos a partir das ontologias, dos modelos de dados em RDF e das regras da SWRL.

Em suma, a incorporação das tecnologias RDF, OWL e SPARQL no Sistema de Organização permitirá a identificação do significado dos conteúdos informacionais e do perfil de cada usuário, e propiciará a personalização do ambiente para cada tipo de indivíduo, para cada tipo de informação, e/ou para cada informação especificamente.

Uma aplicação dessas tecnologias pode ser percebida nos sistemas de organização dinâmicos do *Knowledge Graph*, que vincula dados a uma gama de variáveis que perpassam desde a quantidade de buscas por um determinado elemento, por registros anteriores de um usuário, até a sua localização final.

O *Knowledge Graph* é uma aplicação de busca do Google que possibilita a pesquisa por um determinado conjunto de caracteres, ou seja, a busca por entidades. Em princípio, o Google tenta localizar uma entidade que representa a busca feita pelo usuário, e apresenta essa entidade em uma estrutura de painel, contendo informações de relações que um determinado objeto possui.

Para Monteiro (2015), o *Knowledge Graph* trata-se: “[...] de uma engenharia de Recuperação da Informação, [...] que tem como objetivo a aprendizagem da máquina e a semantização dos resultados de busca” (p. 7).

A organização da informação utilizada pelo *Knowledge Graph* varia de acordo com o tipo de informação que será apresentada, considera as características específicas do usuário que realiza a busca e os detalhes da informação. A variação do sistema de organização possibilita que o ambiente informacional digital de busca, considere características do domínio de cada informação e do usuário que está realizando a busca.

Visando demonstrar a diferenciação que ocorre no sistema de organização desta aplicação, realizou-se duas buscas por entidades classificadas de forma idêntica pelo *Knowledge Graph*: ex-presidentes do Brasil. A Figura 1-A, referente ao resultado da busca por “Fernando Henrique Cardoso” e a Figura 1-B, por “Luiz Inácio Lula da Silva”. As duas entidades por mais que tenham papéis semelhantes para o buscador, são apresentadas

de forma visivelmente diferentes, com tipos de informações distintas, como pode ser visto na Figura 1, a seguir.

A comparação apontada na Figura 1, demonstra com exatidão as diferenças das informações apresentadas das duas entidades. Subentende-se que diferenças semânticas entre os dados buscados dos dois ex-presidentes, causaram esta diferença. Por exemplo, as obras escritas por Fernando Henrique Cardoso são de relevância dentro de um universo de pesquisa, sendo uma informação que por vezes alguns usuários devem buscar. Contudo, Luiz Inácio Lula da Silva não é um escritor, não havendo motivo para apresentar tais informações sobre este segundo ex-presidente. Tais considerações demonstram que o buscador considerou elementos semânticos dessas duas personalidades, visando apresentar de uma forma personalizada as informações aos usuários.



Figura 1. Sistemas de organização dinâmicos do KG 6.

5.2. Sistemas de Rotulagem

O sistema de rotulagem possui uma relação simbiótica com a organização, devido à influência que os rótulos possuem na promoção do sistema de organização da AI.

As tecnologias da WS, tais como, XML, RDF, OWL, Microdados e RDFa apoiam o Sistema de Rotulagem e podem auxiliar na representação dos conteúdos infor-

macionais por meio de rótulos, construindo categorias significativas e semanticamente formais. Os Microdados são partes do código HTML (ou XHTML) que permitem estruturar informações aproveitando os atributos *id* (identifica um elemento) e *class* (identifica um grupo de elementos) empregados em algumas etiquetas. O RDFa é um conjunto de extensões do XHTML que permite introduzir semântica neste tipo de documento.

A Web Semântica possui um vínculo com este sistema, uma vez que é necessário que as informações apresentem significado e estejam contextualizadas semanticamente ao serem expostas. Diante disso, o arquiteto da informação deve buscar tecnologias que contribuam em apresentar as informações com expressividade para os usuários. Tais rótulos, também devem contribuir para que os buscadores sejam capazes de localizar as informações com mais clareza e recuperar os dados com maior precisão.

Diante disso, algumas tecnologias da WS trazem contribuições nessa tarefa. Uma primeira tecnologia que contribui para isso, é o uso do XML. A inserção de metadados em XML nas páginas Web, permitem uma contextualização do conteúdo no qual as informações estão inseridas, contribuindo, assim, para que agentes computacionais ao se depararem com estes conteúdos sejam capazes de compreender o significado dos rótulos da página. Ainda, no contexto dos sistemas de rotulagem, destacam-se as ontologias em OWL, que aumentam a expressividade das informações de uma página da Web.

Além disso, as tecnologias da WS, como os microdados e o RDFa, estão cooperando no que tange a rotulação dos dados. Estas tecnologias são capazes de inserir dentro de uma página *HyperText Markup Language* (HTML) ou *eXtensible Hypertext Markup Language* (XHTML) elementos acerca do seu conteúdo, descrevendo computacionalmente os rótulos que por vezes só possuem significado para as pessoas.

Em síntese, essas duas tecnologias inserem marcações/*tags*, seguindo a estrutura do RDF, para descrever o conteúdo. A partir da inserção destas *tags* em uma página, os buscadores serão capazes de compreender com mais precisão o conteúdo que a página possui, com o intuito

de que o usuário, ao final do processo, possa ter uma experiência satisfatória.

Utilizar estruturas computacionais compreensíveis pelas máquinas, atualmente é tão importante quanto utilizar uma linguagem clara para os rótulos que serão apresentados aos usuários, pois muitas vezes o acesso às páginas Web ocorrem por buscadores.

Diante disto, os rótulos para usuários humanos e não humanos, devem ser uma preocupação do arquiteto da informação, que visa tornar o ambiente acessível.

O *Rich Snippets* tem contribuindo significativamente para as questões relativas aos sistemas de rotulagem, é uma aplicação que tem em sua base os conceitos e as tecnologias da WS e foi concebido inicialmente pelos buscadores para extração de informações com maior valor semântico das páginas Web. A sua estrutura se dá por meio do RDFa e do microdados.

Destaca-se que, o *Rich Snippets* é uma descrição enriquecida que contem informação adicional para os usuários, além das informações apresentadas nos *snippets*, que podem ser informações com respeito ao conteúdo, como: imagens, valorações dos usuários, tempo e datas (Rovira, Codina & Monistrol, 2013).

Assim, o *Rich Snippets* traz uma contribuição relevante ao possibilitar que o contexto das informações de uma página Web possa ser apresentado no próprio buscador, dando mais elementos para que o usuário possa selecionar uma página de seu interesse.

O Google apresenta o *Rich Snippets* juntamente com o conceito de marcação estruturada, base da descrição dos dados na WS, com o objetivo de descrever os recursos na Web, bem como as suas propriedades (Google, 2017).

Os mecanismos de *Rich Snippets* irão fazer com que um buscador obtenha na página Web uma informação descrita em RDFa, que pode estar oculta ao usuário. A Figura 2(A) apresenta os resultados de uma busca com base nas páginas marcadas com *Rich Snippets*, enquanto que a Figura 2(B) expõe como a página Web qualifica a informação para que o buscador tenha acesso aos conteúdos estruturados.



Figura 2. Rich Snippets nos rótulos de tempo e imagem.

A informação contida no HTML (Figura 2[B]) contém Rich Snippet que é utilizado pelo buscador para aprimorar a visualização dos resultados. Assim, o buscador (Figura 2[A]) apresenta as informações tradicionais de um resultado de busca, como título, link e resumo, acrescido dos elementos contidos no HTML (Figura 2[B]), como a imagem do pudim e o tempo de preparo da receita.

A inserção destes elementos aprimora a rotulagem para os agentes computacionais da busca, bem como para outros mecanismos que exploram as páginas Web, por meio das informações estruturadas. Assim, é possível afirmar que os Rich Snippets beneficiam consideravelmente a AI dos websites e dos motores de busca.

5.3. Sistema de navegação

O Sistema de Navegação determina a forma como o usuário navega em uma página Web. Os princípios da WS se encontram aderentes as propostas relacionadas aos sistemas de navegação, principalmente no que tange a permitir que o usuário percorra diferentes caminhos de acesso ao conteúdo informacional previstos neste sistema e conheça sua localização dentro de um ambiente, o que está indiretamente relacionado a contextualização que um sistema possui de seus conteúdos. Assim, contextualizar os dados de acordo com os conceitos da WS, poderá tornar os mecanismos de navegação mais expressivos.

Nesse contexto há uma série de tecnologias da WS que poderão aprimorar os sistemas de navegação, como os modelos de dados e ontologias em RDF e OWL, que possibilitam uma contextualização dos dados mais precisa. Ainda, o uso do SPARQL poderá contribuir na

localização de informações auxiliando a navegação do usuário, e assim, permitir uma navegação contextual dentro da Web.

Uma aplicação da WS que apresenta uma contribuição valiosa aos sistemas de navegação da AI é o *Linked Data*, o qual contempla a ligação entre diversas bases de dados complexas, com um alto nível de contextualização, por meio do uso de tecnologias WS, como o RDF.

O *Linked Data* pode apoiar a AI na elaboração de um sistema de navegação contextual, interligando conteúdos de diferentes bases de dados e possibilitando que o usuário tenha clareza do ambiente em que ele se encontra e da interligação desse ambiente com outros sistemas de informação e websites.

Além disso, a base Linking Open Data (LOD) pode fornecer um suporte estável para a criação de sistemas de navegação baseados no *Linked Data*, uma vez que contempla uma ampla base de dados, dos mais variados domínios (Santarem -Segundo, 2015).

Para ilustrar o exposto, foi analisado a aplicação *LodLive* (<http://lodlive.it/>), que possibilita a interligação e a contextualização de determinados ambientes com o LOD e tem como princípio permitir a localização de um termo dentro de uma base de dados.

A partir de um termo, a referida aplicação demonstra como poderia ser utilizada uma estrutura de navegação em que o usuário manipula as ligações de um recurso no *Linked Data*, para navegar e acessar outros recursos. O uso de uma estrutura semelhante ao *LodLive* poderia ser um sistema de navegação baseado na WS, capaz de contextualizar a informação apresentada dentro de um âmbito geral.

Para demonstrar como funciona o *LodLive* e sua relação com a WS, foi realizado uma busca por “Harry Potter”, uma série britânica de livros de romance, sendo explorada algumas das relações que esse recurso possui dentro da base de dados do *DBpedia* (<http://wiki.dbpedia.org/>). Os resultados obtidos estão apresentados na Figura 3, em que na parte superior é possível visualizar a tela de busca do *LodLive* e na parte inferior são apresentados as respostas da busca.

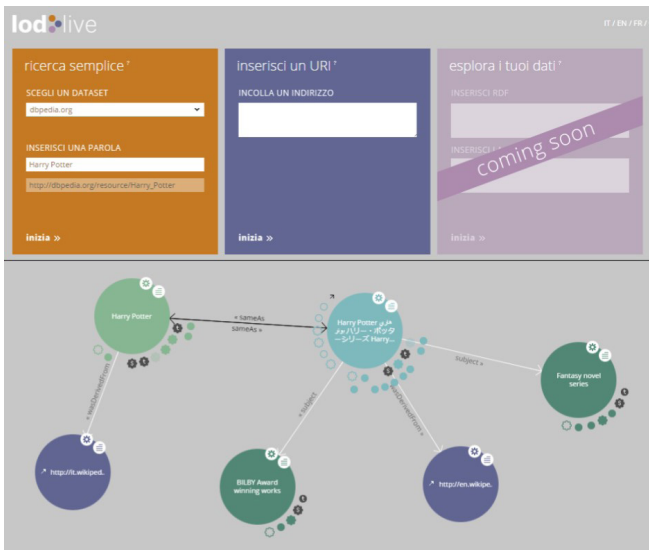


Figura 3. Relações exibidas pelo mecanismo Explorator.

Por meio do grafo apresentado na Figura 3, verifica-se que esta estrutura poderia auxiliar na navegação, sendo uma proposta de contextualizar uma informação ou um conteúdo disponível em um ambiente informacional digital, utilizando como fonte as bases de dados do *Linked Data*. Ainda, a utilização de uma estrutura como essa poderia auxiliar os sistemas de buscas.

5.4. Sistemas de busca

A Web Semântica contribui com os sistemas de buscas facilitando a compreensão das necessidades informacionais do usuário. As tecnologias da WS utilizadas neste sistema são: OWL, RDF, *RDF Schema*, SPARQL e a SWRL. Vale lembrar que algumas aplicações que fazem uso destas tecnologias e influenciam a busca foram apresentadas em outros sistemas da AI, são elas: *Knowledge Graph*, *Rich Snippets* e *Linked Data*.

As tecnologias da WS referentes aos modelos de dados e de ontologias, como RDF, *RDF Schema*, OWL, podem indicar como os dados que um usuário busca estão relacionados às informações da página e sua localização, bem como o contexto de uma determinada busca. Outro destaque são os mecanismos para as consultas e a construção de inferências, como o SPARQL e a SWRL, tecnologias que possibilitam recuperar os dados semanticamente, realizando inferências acerca das buscas realizadas pelos usuários.

A utilização destas tecnologias para a busca, se baseia na ideia de que, se um mecanismo computacional for capaz de compreender o contexto em que os dados se encontram, a recuperação da informação será mais eficiente, e levará em consideração a semântica formal dos objetos digitais. Tal questão, tem como consequência uma exploração aprofundada dos dados, tanto da parte dos usuários, quanto da parte dos mecanismos computacionais.

Uma aplicação que utiliza os princípios da WS para promover um sistema de busca, é o mecanismo de busca *Explorator* (exemplo: <http://139.82.71.26:3002/explorator>), o qual é definido como: “[...] uma ferramenta de pesquisa exploratória *open-source* para grafos RDF, implementada em uma metáfora de interface de manipulação direta.” (Araújo & Schwabe, 2009, p. 3). Essa aplicação pode ser utilizada em diversos sistemas, pois seus códigos são abertos e facilmente interoperáveis.

Na perspectiva do Sistema de Busca da AI, o *Explorator* propõe uma nova maneira de pesquisar e encontrar informações, tendo como princípio as triplas RDF (sujeito relacionado com um objeto por meio de uma propriedade). Este mecanismo possibilita ao usuário realizar uma busca e obter como resultado um recurso (uma entidade que representa algo do mundo real), bem como as suas relações com outras entidades. Desta forma, todo o processo de busca é realizado com base nas relações e nas propriedades de uma determinada informação, ao mesmo tempo que possibilita ao usuário uma forma de explorar os dados por meio dos grafos.

O *Explorator* foi capaz de construir uma forma de “buscar”, totalmente integrada aos princípios da WS, em que os processos de busca e interação são realizados mediante a navegação e exploração em ambientes estruturados em RDF. Essa proposta não segue a perspectiva de resultados de buscas apresentados em listas lineares, sem opção para o usuário expandir as informações que estão relacionadas. Assim, no *Explorator*, os resultados da busca são demonstrados graficamente, dispondo os recursos com a visualização das suas relações e propriedades, como apresentado na Figura 4.

Diante do exposto, verifica-se que a transformação nos sistemas de busca está diretamente vinculada aos sistemas de representação, outro elemento da AI, isso por-

que estes sistemas podem inserir elementos essenciais para descrever e contextualizar um domínio. A seguir são apresentadas as relações da WS com os sistemas de representação.

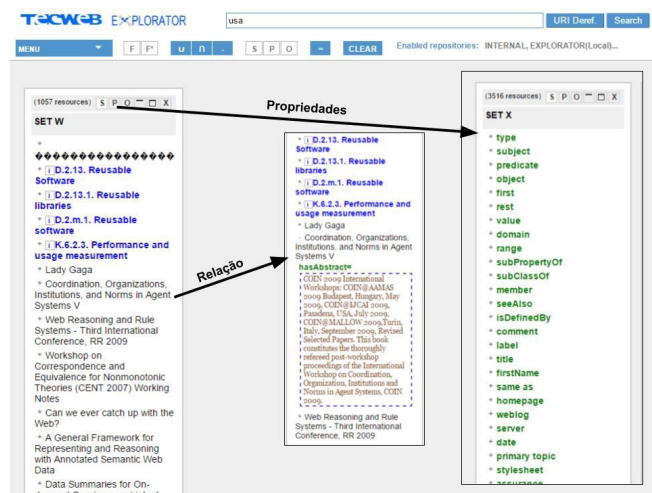


Figura 4. Relações exibidas pelo mecanismo Explorer.

5.5. Sistema de representação

Os vocabulários controlados, tesouros, taxonomias, folksonomias, ontologias e metadados contidos no Sistema de Representação, possibilitam a interação entre os usuários humanos e as máquinas.

No âmbito das tecnologias da WS, destaca-se aquelas, cuja função está na representação dos dados e das ontologias, como o RDF, a OWL e a XML. Adicionalmente, o *Simple Knowledge Organization System* (SKOS) é um modelo concebido para criar tesouros baseados nas tecnologias da WS.

Percebe-se que todas essas tecnologias podem ser utilizadas como a base de um projeto de AI no que tange ao sistema de representação. O uso destas tecnologias é capaz de representar as informações com alta expressividade, de forma que os agentes computacionais compreendam o significado dos dados expressos no website.

Um exemplo de aplicação que utiliza um sistema de representação semântico é a *Europeana* (<http://www.europeana.eu>). A *Europeana* é um ambiente digital que reúne objetos culturais de bibliotecas, arquivos, museus e galerias espalhadas pela Europa. Para Winer e Rocha (2013), os principais propósitos e motivações da

Europeana estão na preservação de todo o conhecimento registrado desde os primórdios da civilização.

A *Europeana* está estruturada com o *Europeana Data Model* (EDM), modelo que contempla as tecnologias da WS, como o RDF. Este modelo permite a existência de relações explícitas de todos os dados da *Europeana*. Ao utilizar o ambiente digital, o usuário é capaz de identificar informações relacionadas de acordo com o contexto dos dados.

Neste sentido, Coneglian e Santarem-Segundo (2017) afirmam que: “O Modelo de Dados EDM utiliza diversos vocabulários para formar um modelo que consiga expressar as ligações existentes, entre autores, obras, organizações, direitos autorais, além de outros tipos de informações, contidas em um objeto cultural.” (p. 91). Os vocabulários são construídos com o uso de SKOS, *dcterms*, dentre outros, o que permite interoperabilidade com diversas aplicações da Web.

No âmbito dos sistemas de representação da AI, o EDM possibilita com que as relações apresentadas entre os conteúdos no ambiente digital da *Europeana* tenham um significado explícito para os computadores, permitindo que a apresentação das informações para os usuários ocorra com maior precisão. Desta forma, os metadados dos conteúdos ficam descritos com base no modelo de dados EDM, o qual conduz a uma representação com base nos princípios da WS.

A Figura 5 ilustra como as relações do EDM ocorrem, ou seja, como os diversos recursos estão relacionados por meio de vocabulários expressivos para o âmbito da WS. Além disso, é possível verificar a possibilidade de realizar ligações com outras bases de dados, como por exemplo o *DBpedia*.

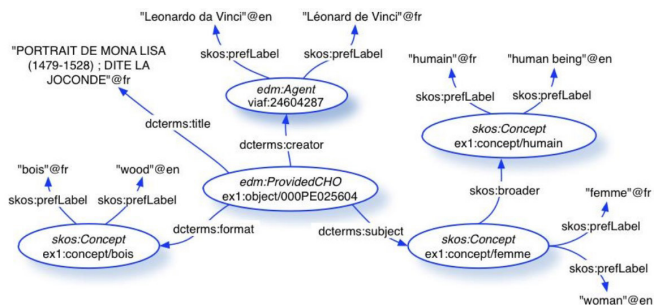


Figura 5. Exemplo das relações do EDM (Europeana, 2013, p. 15).

Os recursos apresentados, na Figura 5, possuem uma série de propriedades, que os interligam a outros recursos, permitindo um relacionamento entre os dados que favorece a compreensão de usuários não-humanos/máquinas ao explorar informações descritas com esse esquema. Desta forma, o EDM pode trazer um sistema de representação para a AI, compatível com os princípios da WS.

5.6. Sistemas da arquitetura da informação e Web Semântica

Visando sintetizar as contribuições da Web Semântica para a Arquitetura da Informação descritas anteriormente, resumiu-se no Quadro 1 as tecnologias da WS que foram analisadas e que podem ser aplicadas nos sistemas da Arquitetura da Informação. Ainda, para cada sistema da AI foi apresentada uma aplicação que utiliza tecnologias da Web Semântica.

Quadro 1. Sistemas de AI e WS.

Sistemas da AI	Tecnologias da WS	Exemplos de aplicações da WS
Organização	XML, RDF, RDF Schema, OWL, SPARQL e SWRL	Knowledge Graph
Rotulagem	XML, RDF, OWL, Microdados e RDFa	Rich Snippets
Navegação	RDF, OWL e SPARQL	Linked Data e LodLive
Busca	OWL, RDF, RDF Schema, SPARQL e a SWRL.	Explorator
Representação	RDF, OWL, XML e SKOS	Europeana Data Model (EDM)

No Quadro 1, identifica-se um conjunto de tecnologias que se encontram presentes na maioria dos sistemas, as quais permitem a incorporação de mais significado nos ambientes informacionais digitais. Dentre essas tecnologias, destaca-se RDF e OWL que estão presentes em todos os sistemas. Tal fato ocorre porque ambas estão vinculadas à representação dos dados, que perpassa pelos processos de todos os sistemas da Arquitetura da Informação.

6. Conclusões

O desenvolvimento de pesquisas que conciliam a Arquitetura da Informação e a Web Semântica são necessá-

rias, uma vez que há uma demanda por aprimorar tanto a forma como uma informação deve ser estruturada, quanto o modo de apresentação que deve ser eficiente para os usuários. Essas demandas são inerentes a AI e WS, e a interrelação entre elas pode aperfeiçoar o modo como as informações são dispostas nos ambientes es informacionais digitais.

A utilização dos conceitos e das tecnologias da WS em projetos de AI, podem fornecer marcação clara e significado bem definido dos conteúdos. Assim, o uso dessas tecnologias favorece a recuperação e aprimora a acessibilidade. Se as máquinas “compreenderem” o sentido das informações, elas serão capazes de apresentar o conteúdo de modo acessível e com mais significado para o usuário.

Para demonstrar a intrínseca relação existente entre a AI e a WS, o presente trabalho utilizou a Arquitetura da Informação Sistêmica, com abordagem aos sistemas de organização, rotulagem, navegação, busca e representação, e identificou tecnologias da Web Semântica que podem enriquecer o projeto da AI. Ainda, foram apresentadas aplicações que utilizam tecnologias da Web Semântica no intuito de esclarecer o relacionamento com a Arquitetura da Informação.

Os resultados obtidos evidenciam como cada sistema da AI pode ser aprimorado com o uso de tecnologias da WS. Foram destacadas as tecnologias com maior frequência de utilização. Além disso, foi possível identificar um conjunto de aplicações que estão transformando a forma como os usuários navegam na Web com influência direta da WS, que por sua vez promove uma melhor interação entre os usuários e os conteúdos informacionais.

Uma consequência da materialização de aplicações baseadas na WS pode levar ao desenvolvimento de projetos de AI com a utilização dos princípios da WS em todos os seus sistemas.

Finalizando, afirma-se que as relações entre a Arquitetura da Informação e a Web Semântica permitem auxiliar novos projetos de AI, que tenham estrutura com um nível de semântica formal mais elevado e que estejam integrados às recomendações da WS.

7. Agradecimentos

Os autores agradecem ao Grupo de Pesquisa Novas Tecnologias em Informação (GPNTI) da Universidade Estadual Paulista Júlio de Mesquita Filho (Brasil), ao Grupo de Investigación en Inteligência Computacional (GICO) do Departamento de Sistemas da Universidad del Cauca (Colômbia), ao Conselho Nacional de Desenvolvimento Científico e Tecnológico - CNPq, à Coordenação de Aperfeiçoamento de Pessoal de Nível Superior - CAPES e à Associação Universitária Ibero-americana de Pós-graduação - AUIP.

8. Referências

- Al-Sudairy, M. T., & Vasista, T. G. K. (2011). Semantic Data Integration Approaches. *International Journal of Web & Semantic Technology*, 2(1)-12. Recuperado de <http://citeseerx.ist.psu.edu/viewdoc/download?rep=rep1&type=pdf&doi=10.1.1.187.3045>
- Araújo, S.F.C., & Schwabe, D. (2009). Explorator: a tool for exploring RDF data through direct manipulation. *Proceedings*, 1-10. Madrid. Recuperado de http://events.linkeddata.org/ldow2009/papers/ldow2009_paper2.pdf
- Berners-Lee, T., Hendler, J., & Lassila, O. (2001). The semantic web. *The Semantic Web*, 284(5), 28-37.
- Campos, J., Santachè, A., & Teixeira, C. (1999). Visualização de modelos tridimensionais de sistemas de informações geográficas distribuídos baseados na WEB. In *Brazilian Workshop on Geoinformatics*, Campinas. São José dos Campos: INPE. 50-58.
- Congeliani, C. S., & Santarem-Segundo, J. E. (2017). Europeana no Linked Open Data: conceitos de Web Semântica na dimensão aplicada das Humanidades Digitais. *Encontros Bibli: revista eletrônica de biblioteconomia e ciência da informação*, 22(48), 88-99. doi: 10.5007/1518-2924.2017v22n48p88
- Ducharme, B. (2013). *Learning Sparql*. Sebastopol, CA: O'Reilly.
- Europeana. (2013, julho 14). *Europeana Data Model Primer. Europeana*. Recuperado de http://pro.europeana.eu/files/Europeana_Professional/Share_your_data/Technical_requirements/EDM_Documentation/EDM_Primer_130714.pdf
- Google. (2017, março 28). Introduction to Structured Data | Search. Recuperado de <https://developers.google.com/search/docs/guides/intro-structured-data>
- Grasic, B. & Podgorelec, V. (2010). Automating ontology based information integration using service orientation. *WSEAS Transactions Computers*, 9(6), 547-556.
- Gurupur, V. P., & Tanik, M. M. (2012). A system for building clinical research applications using semantic web-based approach. *Journal of medical systems*, 36(1), 53-59.
- Hinkelmann, K., Maise, M., & Thönssen, B. (2013). Connecting enterprise architecture and information objects using an enterprise ontology. *Enterprise Systems Conference (ES)*, Cape Town, South Africa: IEEE. 1-11. doi: 10.1109/ES.2013.6690088
- Horrocks, I., Patel-Schneider, P., Boley, H., Tabet, S., Grosz, B., & Dean, M. (2004). SWRL: A Semantic Web Rule Language Combining OWL and RuleML. *W3C Member Submission*. Recuperado de <https://www.w3.org/Submission/SWRL/>
- Martin, D., Burstein, M., Mcdermott, D., Mcilraith, S., Paolucci, M., Sycara, K., ...Srinivas N. (2007). Bringing semantics to web services with OWL-S. *World Wide Web*, 10(3), 243-277. Recuperado de <https://link.springer.com/article/10.1007/s11280-007-0033-x>
- Monteiro, S. D. (2015). Knowledge Graph e a significação: novos agenciamentos semióticos dos índices contemporâneos. *Tendências da Pesquisa Brasileira em Ciência da Informação*, 8(2), 203-218.
- Oliveira, H. P., & Vidotti, S. A. B. G. V. (2012). Arquitetura da Informação digital: conexões interdisciplinares dentro da abordagem sistêmica. In L. E. Cavalcante, V. B. Pinto, & S. A. B. G. V. Vidotti (Orgs.), *Ciência da informação e contemporaneidade: tessituras e olhares* (pp. 184-202). Fortaleza: Edições UFC.
- Rosenfeld, L., & Morville, P. (1998). *Information architecture for the World Wide Web* (1st ed). Cambridge; Sebastopol, CA: O'Reilly.
- Rosenfeld, L., Morville, P., & Arango, J. (2015). *Information architecture: for the web and beyond* (fourth edition). Sebastopol, CA: O'Reilly Media, Inc.

18. Rovira, C., Codina, L., & Monistrol, R. (2013). Rich snippets: información semántica para la mejora de la identidad digital y el SEO. *El profesional de la información*, 22(6), 554-561.
19. Santarem-Segundo, J. E. (2010). *Representação Iterativa: um modelo para repositórios digitais*. (Tese, Doutorado em Ciência da Informação). Faculdade de Filosofia e Ciências, Universidade Estadual Paulista, Marília.
20. Santarem-Segundo, J. E. (2015). Web semântica, dados ligados e dados abertos: uma visão dos desafios do brasil frente às iniciativas internacionais. *Tendências da Pesquisa Brasileira em Ciência da Informação*, 8(2), 219-239.
21. Santos, M. R., Bax, M. P., & Kalra, D. (2010). Building a logical EHR architecture based on ISO 13606 standard and semantic web technologies. *Stud Health Technol Inform*, 160(Pt. 1), 161-165. Recuperado de <https://www.ncbi.nlm.nih.gov/pubmed/20841670>
22. Vidotti, S. A. B. G. V., & Sanches, S. (2004). Arquitetura da informação em web sites. *Anais eletrônicos*, 2. Campinas: UNICAMP.
23. Winer, D., & Rocha, I. E. (2013). Europeana: um projeto de digitalização e democratização do patrimônio cultural europeu. *Patrimônio e Memória*, 9(1), 113-127. Recuperado de <http://pem.assis.unesp.br/index.php/pem/article/viewFile/327/598>
24. World Wide Web Consortium. (2014). Resource Description Framework.