



Universidade Estadual Paulista (UNESP),  
Faculdade de Filosofia e Ciências,  
Campus de Marília

**Jaider Andrade Ferreira**

Wikis semânticos:  
da Web para a Web Semântica

---

*Semantic Wikis: from the Web to the Semantic Web*

Marília  
2014

**Jaider Andrade Ferreira**

**Wikis semânticos:  
da Web para a Web Semântica**

---

*Semantic Wikis: from the Web to the Semantic Web*

Dissertação apresentada ao Programa de Pós-Graduação em Ciência da Informação da UNESP, Faculdade de Filosofia e Ciências, Campus de Marília, como exigência para a obtenção do título de Mestre em Ciência da Informação.

**Orientadora:** Prof.<sup>a</sup> Dr.<sup>a</sup> Plácida Leopoldina Ventura Amorim da Costa Santos.

**Financiamento:** Fundação de Amparo à Pesquisa do Estado de São Paulo (FAPESP) (processo 11/15085-6).

Marília  
2014



A obra “Wikis semânticos: da Web para a Web Semântica”  
de Jaider Andrade Ferreira está licenciada sob a Licença  
*Creative Commons Atribuição 4.0 Internacional* (CC BY 4.0)  
([http://creativecommons.org/licenses/by/4.0/deed.pt\\_BR](http://creativecommons.org/licenses/by/4.0/deed.pt_BR))  
Contato: jaideraf@gmail.com

---

F383w Ferreira, Jaider Andrade.  
Wikis semânticos : da Web para a Web semântica / Jaider Andrade  
Ferreira. – Marília, 2014.  
130 f. ; 30 cm.

Dissertação (Mestrado em Ciência da Informação) - Universidade  
Estadual Paulista, Faculdade de Filosofia e Ciências, 2014.

Bibliografia: f. 122-131.

Orientadora: Plácida Leopoldina Ventura Amorim da Costa Santos.

Apoio financeiro: FAPESP.

1. Wiki semântico. 2. Web semântica. 3. Tecnologia da informação. 4.  
Gestão da informação. 5. Sistemas de recuperação da informação. I.  
Título.

CDD 004.67

---

## **Agradecimentos**

À minha orientadora, a professora Dra. Plácida Leopoldina Ventura Amorim da Costa Santos, pela oportunidade fornecida para a realização da pesquisa, por despertar meu interesse pela investigação científica, pelas orientações, pela motivação e por confiar no desenvolvimento deste trabalho.

Aos professores do Programa de Pós-Graduação em Ciência da Informação da UNESP de Marília, por contribuírem de maneira fundamental para a minha formação acadêmica.

Aos professores da banca de qualificação e de defesa, pela leitura deste trabalho e pelas contribuições apresentadas à estrutura e ao conteúdo do mesmo.

À Fundação de Amparo à Pesquisa do Estado de São Paulo (FAPESP) pelo apoio financeiro (processo 11/15085-6).

Ao Fabrício, por tudo.

## Resumo

Com o desenvolvimento das Tecnologias de Informação e Comunicação, a Ciência da Informação precisou repensar a postura tradicional de gerenciamento informacional. O hipertexto, advindo principalmente por meio do ambiente Web, elevou ainda mais a complexidade do tratamento informacional. A popularização da Internet fez com que a Web se tornasse mais interativa e colaborativa tornando comum a utilização de sistemas wiki para o gerenciamento informacional. Idealizada por Tim Berners-Lee, surge a iniciativa da Web Semântica, na qual as máquinas são capazes de analisar os dados presentes na rede. Nesse contexto aparecem os wikis semânticos, caracterizados por fazerem uso das tecnologias da Web Semântica. Diante desse cenário, considera-se que a Ciência da Informação, que já se preocupa com o desenvolvimento da Web e da Web Semântica, deve também se preocupar com os wikis semânticos. Assim, por meio de uma pesquisa descritiva e exploratória, objetivou-se explorar, apresentar e descrever as características dos wikis semânticos com enfoque nas atividades de descrição, de recuperação e de intercâmbio de informações apoiadas em tecnologias da Web Semântica, visando a favorecer o entendimento, a discussão e o uso dessas tecnologias em ambientes informacionais digitais. Após uma apresentação sobre as raízes históricas da Web Semântica, são destacados os padrões de representação, codificação, descrição, relação e consulta de dados estruturados (URI, XML, RDF, RDFS, OWL e SPARQL) que, junto a outras tecnologias, formam a base da Web Semântica e apoiam o funcionamento dos wikis semânticos. Os wikis semânticos são apresentados e definidos como sistemas wiki que se utilizam de tecnologias da Web Semântica para incorporar conhecimento formalizado, conteúdo, estruturas e links em suas páginas. Após essa etapa, são descritas as principais atividades de descrição, de recuperação e de intercâmbio de informações no *Semantic MediaWiki*, o wiki semântico mais utilizado até o momento. Como considerações finais, entende-se que os wikis semânticos favorecem o entendimento, a discussão e o uso de tecnologias da Web Semântica em ambientes informacionais digitais.

**Palavras-chave:** Wikis semânticos. Web Semântica. Informação e tecnologia.

## **Abstract**

Due to the development of Information and Communication Technologies, Information Science has been forced to rethink the traditional posture of information management. Hypertext, arising mainly through the Web environment, further increased the complexity of the information handling. The popularization of the Internet has led the Web to a more interactive and a more collaborative environment, bringing wiki systems, for example, to manage information in a collaborative way. Conceived by Tim Berners-Lee, there is the Semantic Web initiative in which machines are able to analyze data on the network. In this context, semantic wikis arise: wikis characterized by the use of Semantic Web technologies. Therefore, we believe that Information Science, which cares about the development of the Web and the Semantic Web, should also care about semantic wikis. Thus, by a descriptive and an exploratory research, the objective is to explore, to present and to describe the characteristics of the semantic wikis on the activities of representation, retrieval and exchange of information supported by Semantic Web technologies in order to facilitate the understanding, the discussion, and the use of these technologies in digital information environments. After a presentation about the origins of the Semantic Web, we highlight the data representation, encoding, description, relation, and query standards (URI, XML, RDF, RDFS, OWL and SPARQL) which, with other technologies, form the basis of the Semantic Web and support the functioning of semantic wikis. Semantic wikis are presented and defined as wiki systems that use Semantic Web technologies in order to incorporate formalized knowledge, content, structure and links on their pages. After that, we describe the main activities for information description, retrieval and interchange on Semantic MediaWiki, the most popular and most used semantic wiki engine so far. As conclusion, we consider that semantic wikis can promote understanding, discussions, and use of Semantic Web technologies in digital information environments.

**Keywords:** Semantic wikis. Semantic Web. Information and technology.

## Lista de figuras

Figura 1 – Árvore de Porfírio: uma estrutura em árvore na modelagem conceitual.....	29
Figura 2 – Diferença entre os recursos e links tradicionais e os propostos pela Web Semântica.....	37
Figura 3 – Algumas tecnologias da Web Semântica. ....	38
Figura 4 – Um grafo RDF simples descrevendo o relacionamento entre um livro e seu autor.....	48
Figura 5 – Um grafo RDF com literais para a descrição de valores de dados.....	50
Figura 6 – Serialização em XML de um grafo RDF que descreve o relacionamento entre um livro e um autor, o título do livro e o nome do autor. ....	52
Figura 7 – Explicitação do conhecimento assertivo (RDF) e terminológico (RDFS). ....	57
Figura 8 – Exemplo de uma consulta utilizando a linguagem SPARQL. ....	66
Figura 9 – <i>Wikitexto</i> utilizado na edição de uma página wiki. ....	69
Figura 10 – Página wiki em HTML criada a partir do <i>wikitexto</i> . ....	69
Figura 11 – Página no <i>namespace</i> principal (sem prefixo).....	92
Figura 12 – Página no <i>namespace</i> Ajuda (com o prefixo Ajuda:). ....	92
Figura 13 – Página de categoria.....	94
Figura 14 – Redirecionamento de MASP para Museu de Arte de São Paulo. ....	95
Figura 15 – Redirecionamento de Museu de Arte de São Paulo Assis Chateaubriand para Museu de Arte de São Paulo.....	95
Figura 16 – Página de desambiguação.....	96
Figura 17 – <i>Wikitexto</i> de uma página wiki sobre Brasília no <i>MediaWiki</i> (primeiro parágrafo) e no <i>Semantic MediaWiki</i> (segundo parágrafo).....	99
Figura 18 – Uma página em que foram marcadas propriedades e valores.....	100
Figura 19 – Página de uma propriedade. ....	101
Figura 20 – Página de edição de dados via formulário. ....	103
Figura 21 – Página contendo dados inseridos via formulário. ....	104
Figura 22 – Exemplo de um mapeamento entre uma propriedade local do wiki e uma propriedade de um vocabulário externo por meio de seu URI.....	107
Figura 23 – O resultado de uma consulta semântica e a sua respectiva requisição. ....	111
Figura 24 – Página de busca que se utiliza de um formulário. ....	112
Figura 25 – Uma consulta SPARQL por todas as triplas (sujeito, predicado, objeto) presentes no wiki, começando pelo 15º resultado e limitando-se a 10 resultados.....	113
Figura 26 – Exemplo de exportação de dados em OWL/RDF a partir do <i>Semantic MediaWiki</i> .....	114
Figura 27 – Documento XML obtido pela API do <i>MediaWiki</i> .....	115

## Lista de quadros

Quadro 1 – Exemplos de coisas referenciadas e seus possíveis URIs.....	36
Quadro 2 – Exemplos de modelagens básicas realizadas com a OWL.....	60
Quadro 3 – Exemplos de modelagens complexas realizadas com a OWL. ....	61
Quadro 4 – Exemplo de um resultado de uma consulta SPARQL.....	66
Quadro 5 – Comparação entre sites web tradicionais e sites wiki. ....	72
Quadro 6 – Comparação entre características de wikis tradicionais e de wikis semânticos. ....	80
Quadro 7 – Sistemas de wikis semânticos classificados de acordo com seus focos.....	84
Quadro 8 – Representação dos principais elementos da linguagem OWL no <i>Semantic MediaWiki</i> ...	106
Quadro 9 – Mapeamento para conceitos externos ao wiki por meio de URIs.....	107



## Lista de siglas

ACE	Attempto Controlled English
AIFB	Institut für Angewandte Informatik und Formale Beschreibungsverfahren
API	Application Programming Interface
ARPA	Advanced Research Projects Agency
CERN	Conseil Européen pour la Recherche Nucléaire
CI	Ciência da Informação
DDC	Dewey Decimal Classification
HTML	Hypertext Markup Language
HTTP	Hypertext Transfer Protocol
JSON	JavaScript Object Notation
KIT	Karlsruhe Institut für Technologie
N3	Notation3
OWL	Web Ontology Language
PICS	Platform for Internet Content Selection
RDF	Resource Description Framework
RDFS	RDF Vocabulary Description Language: RDF Schema
SGML	Standard Generalized Markup Language
SMW	Semantic MediaWiki
SPARQL	SPARQL Protocol and RDF Query Language
TCP/IP	Transmission Control Protocol/Internet Protocol
TI	Tecnologia da Informação
TIC	Tecnologias de Informação e Comunicação
URI	Uniform Resource Identifier
URL	Uniform Resource Locator
W3C	World Wide Web Consortium
WWW	World Wide Web
XML	Extensible Markup Language

## Sumário

<b>1 Introdução</b> .....	11
1.1 Problema .....	14
1.2 Proposição.....	14
1.3 Objetivos .....	14
1.4 Justificativa e relevância .....	15
1.5 Motivação.....	16
1.6 Metodologia .....	17
1.7 Estrutura da pesquisa .....	17
<b>2 Web Semântica: histórico, definições e tecnologias</b> .....	22
2.1 A construção de modelos conceituais .....	27
2.2 A inteligência artificial.....	31
2.3 Definições de Web Semântica.....	33
2.4 Tecnologias da Web Semântica.....	38
2.4.1 Uniform Resource Identifier (URI).....	39
2.4.2 Extensible Markup Language (XML).....	41
2.4.3 Resource Description Framework (RDF).....	44
2.4.4 RDF Vocabulary Description Language: RDF Schema (RDFS).....	54
2.4.5 Web Ontology Language (OWL).....	57
2.4.6 SPARQL Protocol and RDF Query Language (SPARQL) .....	64
<b>3 Wikis semânticos: histórico, definições e tecnologias</b> .....	68
3.1 Os wikis e suas características.....	68
3.2 Definições de wiki semântico .....	74
3.3 Principais iniciativas e tecnologias.....	85
3.4 <i>Semantic MediaWiki</i> .....	87
<b>4 Descrição, recuperação e intercâmbio de informações em wikis semânticos</b> .....	90
4.1 A descrição de informações em wikis semânticos .....	90
4.1.1 Estruturação e organização do conteúdo em um wiki .....	91
4.1.2 Propriedades e sua utilização .....	98
4.1.3 Mapeamento para a linguagem OWL.....	104
4.2 A recuperação de informações em wikis semânticos.....	108
4.3 O intercâmbio de informações em wikis semânticos .....	113
<b>5 Considerações finais</b> .....	117
<b>Referências</b> .....	122

## 1 Introdução

A pesquisa aqui apresentada insere-se na linha de pesquisa Informação e Tecnologia do Programa de Pós-Graduação em Ciência da Informação da UNESP, Campus de Marília, e foi desenvolvida com o financiamento da Fundação de Amparo à Pesquisa do Estado de São Paulo (FAPESP) (processo 11/15085-6). Esta pesquisa possui como tema as tecnologias da Web Semântica e, de modo mais específico, a aplicação de tais tecnologias aos ambientes informacionais digitais do tipo wiki semântico.

A Ciência da Informação (CI), ciência social aplicada que se desenvolveu principalmente após a Segunda Guerra Mundial, é a “disciplina que investiga as propriedades e comportamentos da informação, as forças que governam o fluxo da informação e o meio de processamento da informação para uma ótima acessibilidade e usabilidade” (BORKO, 1968, p. 3, tradução nossa).

Le Coadic (2004, p. 2, grifo nosso) afirma que a Ciência da Informação

[...] nasceu da biblioteconomia, tomando, assim, como objetivo de estudo a informação fornecida pelas bibliotecas, fossem públicas, universitárias, especializadas, ou centros de documentação. A leitura pública e a história do livro foram a matéria dos primeiros estudos realizados. Depois, a informação relativa às ciências, às técnicas, às indústrias e ao Estado sobrepôs-se a esses assuntos, dinamizada pelo advento da tecnologia da informação e as necessidades crescentes de informação dos setores científicos, técnicos e industriais, e do grande público. A ciência da informação, portanto, construiu-se e se fundamenta atualmente sobre essa base informacional.

Nesse contexto, observa-se que a Ciência da Informação se desenvolveu principalmente após o avanço das tecnologias de informação. Como apontado por Le Coadic (2004, p. 25), a Ciência da Informação nasceu da Biblioteconomia e

tornou-se, portanto, numa ciência social rigorosa que se apoia em uma tecnologia também rigorosa. Tem por objeto o estudo das propriedades gerais da informação (natureza, gênese, efeito), e a análise de seus processos de construção, comunicação e uso.

Le Coadic (2004, p. 25) também destaca que, no campo da tecnologia da informação, a criação de “produtos, sistemas e serviços que permitem a construção, comunicação, armazenamento e uso da informação” são os objetos de estudo dessa área.

Borko (1968, p. 3), em seu clássico artigo “*Information Science: what is it?*”, também discute a Ciência da Informação como uma ciência interdisciplinar derivada de e relacionada

a campos como a tecnologia da computação, a comunicação, a biblioteconomia entre outras. A tecnologia aplicada à informação deu origem a uma área relativamente nova chamada de Tecnologia da Informação (TI).

Ainda sobre essa relação de interdisciplinaridade entre Ciência da Informação e Ciência da Computação, Saracevic (1996, p. 42) destaca que “a CI está inexoravelmente ligada à tecnologia da informação” e acrescenta:

a ciência da computação trata de algoritmos que transformam informações enquanto a CI trata da natureza mesma da informação e sua comunicação para uso pelos humanos. Ambos os objetos são interrelacionados e não competidores, mas complementares (SARACEVIC, 1996, p. 42).

Na literatura científica nacional, Alves et al. (2007, p. 37-38) enfatizam que, na atuação conjunta entre as duas áreas há um favorecimento na constituição de uma estrutura tecnológica e informacional, contemplando as metodologias de tratamento informacional já consolidadas na Ciência da Informação e as aplicações tecnológicas advindas da Ciência da Computação.

Pode-se, portanto, perceber que Ciência da Informação e Ciência da Computação devem se empenhar mutuamente na busca de estudos, produtos e serviços que tenham por objeto a informação e um aperfeiçoamento de qualidade na “origem, coleção, organização, armazenamento, recuperação, interpretação, transmissão, transformação e utilização” dessa informação (BORKO, 1968, p. 3, tradução nossa).

Sob essa perspectiva considera-se que

[...] a Ciência da Informação deveria ter ou criar mais espaços de investigação que permitam a compreensão das Tecnologias de Informação e Comunicação para a potencialização de competências informacionais, para a criação de arquiteturas informacionais e computacionais mais inclusivas, para a conceituação de usos da informação em ambientes informacionais digitais, para a aprendizagem de metalinguagens e para a representação da informação (SANTOS; VIDOTTI, 2009, p. 5).

Também sob essa perspectiva, Le Coadic (2004, p. 20-21) afirma que a Ciência da Informação também se relaciona com vários temas além dos relacionados ao livro e a sua história, a biblioteca, a biblioteconomia, a bibliometria, e as teorias de classificação, indexação, descrição e recuperação da informação (temas centrais para a Ciência da Informação segundo o autor). O autor alega ainda que a Ciência da Informação também apoia firmemente campos como a Informática (bases de dados relacionais, sites na rede (Web), programas para *groupware*, programas de hipertexto e cartográficos, etc.); a

Eletrônica; e as Telecomunicações (compressão, memórias ópticas, multiplexação, redes de comunicação).

Enfatizando a vertente tecnológica, Le Coadic (2004, p. 17-18) afirma que a Ciência da Informação deve trabalhar sob a influência do desenvolvimento da produção e das necessidades de informações (construção e demanda de informação); advento do novo setor das indústrias da informação (produtores e servidores de bancos de informações, sites e portais da Internet, etc.); e a influência do advento das tecnologias eletrônicas (analógicas ou digitais) e fotônicas da informação (microcomputadores, fibras ópticas, dispositivos multimídias, etc.).

Esses organismos tornaram-se depósitos de conhecimentos sobre um assunto, um objeto, de respostas a consultas, isto é, entrepostos de informações. Melhor ainda, são verdadeiros meios de comunicação de informações, que atingem um número cada vez maior de pessoas. Ou seja, sob o efeito dessas três categorias de mudanças – culturais, econômicas e tecnológicas – tornaram-se multimídias de massa, como seus colegas da imprensa escrita e audiovisual (LE COADIC, 2004, p. 18).

Com a tecnologia favorecendo a criação, a comunicação e o uso da informação, a Ciência da Informação direciona seus estudos para o ambiente em que essa transformação está ocorrendo de maneira mais significativa: a Web.

O volume de informações disponibilizadas por meio da Web vem se multiplicando exponencialmente nos últimos anos em decorrência de fatores tecnológicos, políticos e sociais (CASTELLS, 2007). Com o desenvolvimento, o barateamento e o consumo de Tecnologias de Informação e Comunicação (TIC) por todas as camadas sociais, a Ciência da Informação encara o desafio de repensar a postura tradicional de gerenciamento informacional, principalmente, em ambientes digitais.

O hipertexto, disseminado principalmente por meio da Web, elevou ainda mais a complexidade do tratamento informacional. A popularização da Internet fez com que a Web se tornasse mais interativa e colaborativa tornando comum a utilização de sistemas wiki para o gerenciamento informacional (AQUINO, 2008; FUMERO; ROCA, 2007; O'REILLY, 2005, 2006).

Tim Berners-Lee, por sua vez, idealizou a proposta da Web Semântica (BERNERS-LEE; HENDLER; LASSILA, 2001), em que as máquinas seriam capazes de analisar os dados presentes na rede. Nesse contexto aparecem os wikis semânticos (BRESLIN; PASSANT;

DECKER, 2009; BUFFA et al., 2007; SCHAFFERT et al., 2006; ZAIDAN; BAX, 2011a, 2011b), caracterizados por fazerem uso das tecnologias da Web Semântica.

Diante desse cenário, considera-se que a Ciência da Informação, que já aborda os desenvolvimentos da Web (JORENTE; SANTOS; VIDOTTI, 2009; JORENTE; SANTOS, 2010) e, inclusive, a Web Semântica (SOUZA; ALVARENGA, 2004; CARVALHO, 2009), deve também abordar os wikis e seus desenvolvimentos, como é o caso dos wikis semânticos.

A partir da caracterização da Ciência da Informação, de sua origem, de seu objeto de estudo, de sua interdisciplinaridade com outras áreas do conhecimento e de seu ambiente de pesquisa mais destacado atualmente – a Web –, apresenta-se a seguir a descrição dos principais componentes desta pesquisa: o problema (1.1), a proposição (1.2), os objetivos geral e específicos (1.3), a justificativa e a relevância científica e social (1.4), a motivação (1.5), a metodologia (1.6) e a estrutura textual desta dissertação (1.7).

### **1.1 Problema**

Partindo do pressuposto de que a Ciência da Informação deve preocupar-se com os ambientes informacionais do tipo wiki, tem-se como o problema desta pesquisa: como se configuram os processos de representação, recuperação e intercâmbio de informações em wikis semânticos?

### **1.2 Proposição**

Diante do problema exposto, propôs-se um estudo que objetivou explorar, apresentar e descrever as características dos wikis semânticos com enfoque nas atividades de representação, recuperação e intercâmbio de informações apoiadas em tecnologias da Web Semântica para, assim, favorecer o entendimento, a discussão e o uso dessas tecnologias em ambientes informacionais digitais gerenciados e propostos por profissionais da Ciência da Informação.

### **1.3 Objetivos**

A presente pesquisa tem como objetivo geral:

- Explorar, apresentar e descrever as características dos wikis semânticos com enfoque nas atividades de representação, de recuperação e de intercâmbio de informações apoiadas em tecnologias da Web Semântica para, assim,

favorecer o entendimento, a discussão e o uso dessas tecnologias em ambientes informacionais digitais.

Para alcançar o objetivo geral foram elencados os seguintes objetivos específicos:

- Apresentar brevemente conceitos de Web Semântica, wiki e wiki semântico para fornecer o entendimento básico da terminologia utilizada na pesquisa;
- Compreender os padrões de codificação, de representação, de relação e de consulta de dados estruturados relacionados ao desenvolvimento da Web Semântica (URI, XML, RDF, RDFS, OWL e SPARQL);
- Investigar as principais iniciativas que englobam as tecnologias wiki com as tecnologias da Web Semântica;
- Descrever as principais contribuições das tecnologias da Web Semântica aplicadas aos wikis para uma possível utilização dessas tecnologias em ambientes informacionais digitais.

#### **1.4 Justificativa e relevância**

Em âmbito acadêmico, a realização desta pesquisa contribui para o corpo teórico da Ciência da Informação no que diz respeito aos ambientes semânticos, em especial, aos wikis semânticos e suas tecnologias para a codificação, a representação e o intercâmbio de informações.

A pesquisa justifica-se também por favorecer o entendimento, a discussão e o uso das tecnologias da Web Semântica em ambientes informacionais digitais.

Em âmbito profissional a pesquisa mostra-se relevante por proporcionar subsídios aos profissionais da Ciência da Informação para o entendimento das tecnologias relacionadas à Web Semântica que estarão presentes nos sites Web e que nortearão a criação e o desenvolvimento de ambientes informacionais digitais interoperáveis, mais visíveis na rede e mais abertos à colaboração.

O estudo proposto torna-se exequível, pois se desenvolve tendo a supervisão do Programa de Pós Graduação em Ciência da Informação da UNESP (PPGCI/UNESP), que tem focado sua atenção aos estudos referentes ao tratamento, ao acesso e à recuperação de informações em ambientes informacionais digitais.

Dentro do PPGCI/UNESP, esta proposta foi apresentada à linha de pesquisa Informação e Tecnologia, uma vez que os wikis semânticos fazem parte das Tecnologias de

Informação e Comunicação e que tal linha desenvolve estudos sobre as preocupações da Ciência da Informação que estão associadas aos métodos e instrumentos proporcionados por essas tecnologias.

### **1.5 Motivação**

Pode-se dizer que a motivação para a proposição desta pesquisa originou-se na realização de um estudo em nível de iniciação científica intitulado “A *Wikipedia*, a Inteligência Coletiva e o conceito Web 2.0”, fomentado pela Fundação de Amparo à Pesquisa do Estado de São Paulo (FAPESP) (processo 09/16330-4).

O estudo de iniciação científica desenvolvido no período entre março de 2010 e dezembro de 2011 teve como proposição compreender os conceitos de Web 2.0, Inteligência Coletiva, wikis e *Wikipedia* no contexto da Ciência da Informação, visando analisar como esses conceitos estavam representados nas publicações científicas nacionais e internacionais de tal área.

Permeando essa temática, foi explorada a interdisciplinaridade entre a Ciência da Informação, a Ciência da Computação e a Tecnologia da Informação. Nesse ponto foi abordada a proposta do *Memex* de Vannevar Bush (1945), o hipertexto de Ted Nelson (1965) e, além desses, o conceito de ciberespaço de Pierre Lévy (2000), o surgimento da Web com Tim Berners-Lee (1989) e a mudança de paradigma trazida pelo conceito de Web 2.0.

Após a explanação desses conceitos, foi discutida a relação entre a chamada Web 2.0 e o conceito de Inteligência Coletiva proposto por Pierre Lévy. Em seguida, foi analisado o conceito de Web 2.0 dado por teóricos da Ciência da Informação, os novos paradigmas da área apontados por Le Coadic (2004) e o conceito de Biblioteca 2.0. Foram abordados também os wikis (ferramentas utilizadas para o gerenciamento de informações de modo colaborativo) e a *Wikipedia* (site wiki mais conhecido), apresentando seus conceitos e seus usos por teóricos e profissionais da Ciência da Informação.

Com o desenvolvimento da pesquisa de iniciação científica e com o amadurecimento acadêmico adquirido no período da graduação surgiram outros questionamentos que permeiam a temática dos wikis. Dentre esses questionamentos, encontra-se o que esta pesquisa se propõe a compreender: os wikis semânticos.



## **1.6 Metodologia**

Quanto à natureza, esta pesquisa pode ser identificada como teórica e exploratória, quanto aos fins como descritiva e quanto aos meios de investigação como bibliográfica (MORESI, 2003).

Trata-se de uma pesquisa exploratória porque busca compreender as características e contribuições dos wikis semânticos para a Ciência da Informação, área em que existe pouco conhecimento acumulado e sistematizado sobre tal tema. Além disso, é uma pesquisa descritiva porque apresenta as características dos wikis semânticos e dos principais padrões da Web Semântica.

Esse procedimento metodológico permite abordar os elementos presentes nos objetivos específicos definidos e concretizar o objetivo geral apresentado.

Para a coleta de dados foram utilizadas fontes primárias (obras produzidas pelo(s) próprio(s) autor(es): livros, teses, dissertações, relatórios, entre outros), fontes secundárias (documentos produzidos a partir de documentos primários: artigos de revistas, manuais, dicionários, enciclopédias, entre outros) e fontes terciárias (documentos que tem por função guiar o usuário para todas as fontes primárias e secundárias: bibliografias, catálogos, índices e outros).

Para atingir os objetivos propostos, os levantamentos bibliográficos foram realizados em acervos de bibliotecas, bases de dados nacionais e internacionais (textuais e referenciais) e na Internet.

Foram considerados os seguintes tipos de documentos nos idiomas português, inglês e espanhol e sem limitação cronológica ou geográfica: livros, artigos de periódicos, teses, dissertações, trabalhos apresentados em eventos, relatórios de pesquisa e páginas web.

Durante os levantamentos bibliográficos os documentos encontrados foram selecionados, lidos e fichados. O método de fichamento consiste em fazer apontamentos sobre o material selecionado a fim de facilitar a assimilação de seu conteúdo e tê-los registrados para o posterior uso na redação da dissertação.

## **1.7 Estrutura da pesquisa**

Esta pesquisa está estruturada em cinco capítulos que podem ser apresentados sucintamente da seguinte forma.

O capítulo 1, Introdução, apresenta o delineamento da dissertação, identificando o problema de pesquisa, a proposição, os objetivos a serem alcançados, a justificativa e a relevância científica e profissional do estudo, além da motivação e da metodologia adotada para o alcance dos objetivos.

O capítulo 2, Web Semântica: histórico, definições e tecnologias, fornece um breve histórico das origens da Web Semântica, as definições e os conceitos mais relevantes do tema para esta pesquisa e apresenta brevemente algumas das tecnologias mais importantes necessárias à concretização da Web Semântica.

O capítulo 3, Wikis semânticos: histórico, definições e tecnologias, aborda as origens dos wikis, a sua utilização, definições sobre o que seriam os wikis semânticos, as principais iniciativas relacionadas a eles e apresenta o *Semantic MediaWiki*: principal aplicação de wiki semântico utilizado atualmente.

O capítulo 4, Descrição, recuperação e intercâmbio de informações em wikis semânticos, aborda a descrição das atividades de representação, de recuperação e de intercâmbio de informações no *Semantic MediaWiki*, relaciona os wikis semânticos e a área de Ciência da Informação e apresenta as potencialidades de uso dessas tecnologias por profissionais da Ciência da Informação, favorecendo o entendimento, a discussão e o uso dessas tecnologias em ambientes informacionais digitais.

Finalmente, no capítulo 5, Considerações finais, são apresentadas as considerações possibilitadas pela pesquisa.

Para uma melhor compreensão da dissertação, a seguir, são apresentados alguns termos e como eles são entendidos e empregados no texto.

**Representação** – O termo representação, nesta pesquisa, se resume ao ato de representar uma entidade do mundo (real ou não) por meio de um ou mais identificadores (persistentes ou não) geralmente por meio de uma cadeia de caracteres (padronizada ou não). Na literatura é comum encontrar o termo representação sendo utilizado como sinônimo de descrição. Exemplos:

Entidade: Jaider Andrade Ferreira  
Identificador: 345.607.482-32

Entidade: Universidade Estadual Paulista (UNESP)  
Identificador: <http://unesp.br>

Nesses exemplos, os identificadores **representam**, e por conseguinte identificam, as entidades. O termo representação será utilizado com esse entendimento principalmente nas seções de 2.4.1 a 2.4.5, seções que tratam dos URIs, da XML, do RDF, da RDFS e da OWL.

**Descrição** – Ato de descrever uma entidade do mundo (real ou não) por meio de propriedades (também chamadas de atributos) e seus valores (existentes ou não). Exemplos:

Entidade: Jaider Andrade Ferreira  
Idade: 24 anos  
Nacionalidade: Brasileiro  
Naturalidade: São Paulo

Entidade: Universidade Estadual Paulista (UNESP)  
Idade: 38 anos  
Reitor: Julio Cezar Durigan  
Sede: São Paulo

Nesses exemplos, as propriedades idade, nacionalidade, naturalidade, reitor, sede e seus respectivos valores **descrevem** as entidades. Nesta pesquisa, o termo descrição será utilizado com esse entendimento principalmente nas seções de 2.4.3 a 2.4.5 (RDF, RDFS e OWL).

**Codificação** – Na perspectiva do processamento digital de sinais, a codificação é entendida como a modificação de características de um sinal para torná-lo mais apropriado para uma aplicação específica como, por exemplo, a transmissão ou o armazenamento de dados. Exemplos:

Entidade: Jaider Andrade Ferreira  
Idade: 24 anos  
Nacionalidade: Brasileiro  
Naturalidade: São Paulo

```
<entidade id="46602292X">Jaider Andrade Ferreira</entidade>
<idade tipo="anos">24</idade>
<nacionalidade >Brasileiro</nacionalidade>
<naturalidade >São Paulo</naturalidade>
```

Nesse exemplo, as marcações **codificam** uma descrição. O termo codificação será utilizado com esse entendimento principalmente nas seções 2.4.1 e 2.4.2 (URI e XML).

**Recuperação** – O termo recuperação é entendido nesta pesquisa como sendo a obtenção de um ou mais recursos por meio de uma estratégia (ou parâmetro) de busca em um sistema de informação. Normalmente é utilizado no contexto de banco de dados. O

termo será utilizado com esse entendimento principalmente nas seções 2.4.6 e 4.2 que tratam, respectivamente, da linguagem SPARQL e da recuperação de informações em wikis semânticos.

**Intercâmbio** – Intercâmbio é aqui entendido como troca, se referindo à troca de dados ou ao envio e ao recebimento destes por aplicações de informática. O termo intercâmbio está relacionado aos termos interoperabilidade e codificação. Nesta pesquisa, o termo intercâmbio será utilizado com esse entendimento principalmente nas seções 2.4.2, 2.4.3, 2.4.4, 2.4.5 e 4.3 (XML, RDF, RDFS, OWL e o intercâmbio de informações em wikis semânticos).

**Dado** – Nesta pesquisa, dado será entendido como um dispositivo que pode ser interpretado como um tipo de valor (quantitativo ou qualitativo) sobre alguma entidade (real ou não). Os dados podem ser manipulados tanto como valores (fixos) quanto como variáveis. Dados que são derivados do raciocínio lógico e/ou que são utilizados para uma tomada de decisão (humana ou não) serão coletivamente chamados de conhecimento. Na Ciência da Informação e na Ciência da Computação o termo dado é utilizado em campos como Metadados, Modelagem de dados, Banco de dados e *Linked Data*. O termo dado será utilizado com esse entendimento principalmente nas seções 2.4.2, 2.4.3, 2.4.5, 2.4.6 e 4.1.2 (XML, RDF, OWL, SPARQL e propriedades e sua utilização em wikis semânticos).

**Conhecimento** – O termo conhecimento será utilizado nesta pesquisa como sendo o resultado de um raciocínio lógico (realizado por humanos ou não), isto é, a conclusão lógica das premissas (verdadeiras ou falsas) em um silogismo. Na prática, a conclusão das premissas pode levar o agente a uma tomada de decisão. Na Ciência da Informação e na Inteligência Artificial o termo conhecimento é utilizado em campos como Organização do Conhecimento e Representação do conhecimento. O termo conhecimento será utilizado com esse entendimento principalmente na seção 2.4.5 (OWL).

**Informação** – O termo informação, nesta pesquisa, poderá ser utilizado de forma intercambiar, por vezes representando dado, por vezes representando conhecimento ou mesmo representando um estágio intermediário entre esses. A informação, que poderá ser verdadeira ou falsa, poderá também ser entendida como premissa em um silogismo.

Faz-se necessário salientar que os termos acima listados possuem diversos outros entendimentos. Porém, para facilitar a compreensão desta pesquisa, eles foram delineados,

recortados, diminuídos, sintetizados e, por que não, interpretados pelo autor como se seguem.

## 2 Web Semântica: histórico, definições e tecnologias

Há algumas centenas de anos que a humanidade convive com o texto escrito, ou seja, com a forma linear de comunicação e informação constituída pelo alfabeto. O texto linear, impresso e rígido foi e ainda é de grande importância para as ciências e para a sociedade como um todo. Sem o texto, o progresso científico – na história, na educação, na saúde, etc. – teria outro rumo. Um rumo desconhecido que só é possível imaginar.

Contudo, ao longo do tempo, outro tipo de texto surgiu: o hipertexto. Em princípio, não aquele em ambientes digitais, mas sim aquele em forma de remissivas, referências, indicações providas de um texto a outro.

Nas últimas décadas o hipertexto foi digitalizado, ganhando, assim, novas potencialidades. Quando em formato digital, o hipertexto ganha e fornece interatividade, rapidez, fluidez, e, até mesmo, desorientação (LEÃO, 2005).

Os nativos digitais (SÁEZ VACAS, 2007) reconhecem o hipertexto principalmente pelo ambiente *World Wide Web*, ou simplesmente Web – um dos ambientes digitais hipertextuais mais utilizados para a produção e a obtenção de informações (ALVES, 2005; RAMALHO; VIDOTTI; FUJITA, 2007; SOUZA; ALVARENGA, 2004).

O hipertexto utilizado na Web advém de um conjunto de tecnologias que foram planejadas para funcionar nos mais variados sistemas computacionais. Tais tecnologias foram criadas e melhoradas principalmente nos últimos 50 anos, sendo que a principal estrutura em que o hipertexto repousa é a estrutura de redes.

Sabe-se que o cálculo foi certamente a principal motivação para a construção dos primeiros computadores, mas o aspecto de comunicação entre eles logo se tornou um importante problema (FERNEDA, 2012). Já no final da década de 1950, os primeiros computadores já estavam distribuídos por vários locais nos Estados Unidos e vários projetos conjuntos que se utilizavam desses dispositivos precisavam transmitir seus dados uns aos outros. Redes de telecomunicações já estavam sendo utilizadas há algum tempo e foi natural o uso dessa infraestrutura tecnológica para que as primeiras transmissões de dados entre computadores acontecessem. No entanto, várias décadas de desenvolvimento foram requeridas para se chegar à rede de comunicação global e ubíqua que se conhece hoje, com a *World Wide Web* sendo o meio mais proeminente construído com base nessa estrutura (HITZLER; KRÖTZSCH; RUDOLPH, 2009).

O funcionamento das redes de computadores para a transmissão de dados progrediu significativamente na década de 1960. O cientista da computação J. C. R. Licklider (1915-1990) foi o primeiro a conceber as redes de computadores em um sentido moderno do termo, apresentando suas considerações em uma série de memorandos em 1962. Uma invenção crucial foi a ideia da comutação de pacotes (*packet switching*) – a noção de dividir a informação em pequenos pacotes que são transmitidos individualmente – que foi atribuída aos trabalhos individuais de Paul Baran, Donald Davies e Leonard Kleinrock (HITZLER; KRÖTZSCH; RUDOLPH, 2009).

Vários protocolos de comunicação foram desenvolvidos durante e após a década de 1960 permitindo que mais e mais computadores fossem interligados. Em 1969, a primeira rede de transmissão realizada pela comutação de pacotes foi estabelecida, dando origem a ARPANET: uma rede de quatro computadores mantida pela *Advanced Research Projects Agency* (ARPA) (Agência de Projetos de Pesquisa Avançados) do Departamento de Defesa dos Estados Unidos. A seguir, outras redes locais foram sendo estabelecidas e o problema da interconexão das redes logo se tornou relevante (CASTELLS, 2007).

Em 1974, a pilha de protocolos de interconexão de redes *Transmission Control Protocol/Internet Protocol* (TCP/IP) foi publicada como uma tentativa de superar a diversidade de protocolos de redes de computadores. Com mais e mais redes conectadas, a infraestrutura de comunicação global que agora é conhecida como Internet se consolidou (CASTELLS, 2007; HITZLER; KRÖTZSCH; RUDOLPH, 2009).

Foi somente em 1989 que o inglês Tim Berners-Lee, trabalhando no *Conseil Européen pour la Recherche Nucléaire* (CERN) (Conselho Europeu para a Pesquisa Nuclear) na Suíça, fez uma proposta (BERNERS-LEE, 1989) para superar os problemas de comunicação entre físicos colaboradores distribuídos pelo mundo: sua proposta visava à criação de um meio comum de troca de documentos hipertextuais interligados entre diversos sistemas de computadores. Berners-Lee chamou esse meio de *World Wide Web* e, no final do ano de 1990, providenciou as primeiras versões do *Hypertext Transfer Protocol* (HTTP) (Protocolo de Transferência de Hipertexto), da *Hypertext Markup Language* (HTML) (Linguagem de Marcação de Hipertexto), o primeiro editor e navegador de HTML e o primeiro software de servidor Web.

Berners-Lee foi influenciado pelas ideias de Vannevar Bush, Ted Nelson, Douglas Engelbart e pela criação e pelo desenvolvimento da Internet nos anos 70 por,

principalmente, os engenheiros e pioneiros no desenvolvimento de rede, Donald Davies, Paul Baran, Vint Cerf e Bob Kahn (BERNERS-LEE; FISCHETTI, 2000, p. 5-6).

Inicialmente, o objetivo da Web era facilitar a troca de informações entre pesquisadores de diversos lugares por meio de uma interface visual mais dinâmica do que a interface de linha de comando, até então utilizada na Internet.

Essa interface visual é apresentada pelo navegador (*browser*) que, por sua vez, utiliza protocolos de comunicação para receber, transmitir, codificar e decodificar dados de e para uma rede previamente estabelecida para conexão. As páginas da Web exibidas nesses navegadores são codificadas em HTML.

Segundo Berners-Lee e Fischetti (2000, p. 76, tradução nossa) a principal proposta da Web é “ser um meio hipertextual único, universal e acessível para compartilhar informação”.

Para os professores e pesquisadores Pascal Hitzler (Kno.e.sis Center at Wright State University, Dayton, Ohio, USA), Markus Krötzsch (Department of Computer Science at the University of Oxford) e Sebastian Rudolph (Karlsruhe University, Mannheim, Germany) o objetivo da Web é “tornar o conhecimento disponível amplamente acessível e aumentar a utilidade desse conhecimento por meio de aplicações avançadas de busca, navegação e avaliação” (2009, p. 1, tradução nossa).

Em 1991 foi estabelecido o primeiro servidor Web fora da Europa e o próprio Berners-Lee anunciou a Web publicamente em um grupo de notícias:

O projeto *World Wide Web* (WWW) tem por objetivo permitir que links possam ser feitos com qualquer informação e em qualquer lugar. [...] O projeto WWW foi lançado para possibilitar aos físicos de alta energia compartilhar dados, notícias e documentos. Estamos bastante interessados em disseminar a Web para outras áreas e ter servidores *gateway* para outros dados. Colaboradores são bem-vindos! (BERNERS-LEE, 1991, tradução nossa).

Durante a década de 1990, a Web se tornou o meio mais popular para se utilizar a Internet e, a partir de 1995, conseqüentemente, obteve maior relevância comercial. Como a própria WWW, aplicações Web inovadoras como os wikis e os blogs começaram a ser introduzidas, primeiramente, para o uso por comunidades privadas e públicas com fins acadêmicos e educacionais e, posteriormente, adotadas para fins comerciais (HITZLER; KRÖTZSCH; RUDOLPH, 2009).

Atualmente, a Web, tal como a Internet, é utilizada não somente para o consumo passivo de informações, mas também para a criação e a publicação de novos conteúdos,



distinguindo-se, desse modo, das mídias tradicionais. Essa característica foi impulsionada por sites populares que simplificaram os processos de contribuição pelos usuários: *Wikipedia, Flickr, YouTube* e várias redes sociais são exemplos típicos.

A Web desenvolvida com páginas estáticas, com forte apelo comercial e caracterizada pelo modelo comunicacional um-todos ou um-para-muitos, sofreu modificações com o passar dos anos.

Novas linguagens e padrões para a criação de páginas Web dinâmicas foram desenvolvidas permitindo que usuários interagissem com aplicações ou mesmo com outros usuários pela rede, gerando altíssimas quantidades de dados não estruturados e/ou semiestruturados de um modo nunca antes visto pela humanidade, intensificando um modelo comunicacional denominado todos-todos (LÉVY, 2000) ou muitos-para-muitos no qual, em larga escala, cada receptor pode transformar-se em um emissor de grande alcance.

Esses recentes desenvolvimentos ocorridos na Web – o crescimento das contribuições dos usuários e a crescente exploração de dados estruturados e/ou semiestruturados – coincide com um melhoramento em usabilidade e interatividade das interfaces gráficas utilizadas na Rede.

Atualmente, a Web é um espaço distribuído de informação que provê uma abundância de fontes de informação heterogêneas e em muitos formatos. Porém, a troca de informações na Web só é possível quando há acordo no estabelecimento de formatos de dados padrão (*standard data formats*) e na exploração dos hiperlinks que tornam os recursos distribuídos em estruturas do tipo Web. Vale destacar que essas estruturas não são somente utilizadas pelos agentes humanos para a leitura e para a navegação, mas também por mecanismos de busca para a coleta e para o *ranking* de páginas.

A falta de estruturação, isto é, a falta de descrição formal das informações presentes nas páginas Web, fez com que os mecanismos de busca clássicos se tornassem insuficientes para gerenciar a quantidade sempre crescente de conteúdo.

Mesmo com a variedade de ferramentas disponíveis e apesar de estarem sempre em constante atualização em suas técnicas de busca, muitas vezes as ferramentas atuais não conseguem atender de modo satisfatório seus usuários. Apesar de toda tecnologia algumas limitações ainda ocorrem. Os principais empecilhos estão relacionados com a crescente quantidade de informações disponibilizadas; com as próprias limitações das ferramentas de busca (suas técnicas de busca e indexação dos sites) e com a falta de representação (tratamento) adequado dos recursos informacionais na rede [...] (ALVES, 2005).

Ao invés de depender somente de buscas baseadas em texto completo, as aplicações Web têm introduzido mais modelos de organização e busca por informação. Um exemplo popular é a etiquetagem social (*social tagging*) que é tipicamente utilizada para prover buscas estatísticas e funcionalidades de navegação baseadas em simples palavras-chave (*tags*) que os usuários atribuem aos recursos. Essa abordagem tem tido sucesso particularmente na estruturação do conteúdo que não é primariamente baseado em texto – imagens, vídeos, sons ou produtos comercializados em lojas online. Tanto as abordagens de etiquetagem social em geral, quanto o resultado da etiquetagem têm sido chamadas de folksonomia (*folksonomy*), junção do termo inglês *folk* (povo, gente) e taxonomia.

Existem ainda outros tipos de atributos que os usuários utilizam e que são reaproveitados para ajudar o processo de recuperação da informação de acordo com um domínio de interesse: produtos comerciais têm um preço, certos produtos têm suas especificações técnicas, entradas de blogs têm uma data, hotéis podem ser classificados com quantidades de estrelas, etc.

Nesse contexto, há uma tendência a uma maior estruturação das informações disponibilizadas na Web. Estruturação essa que diz respeito à representação e à descrição dos recursos informacionais. No entanto, muitos dos exemplos citados não estão baseados em qualquer padrão, o que dificulta o reuso das informações por humanos e máquinas. Se cada site possui seu próprio mecanismo de representação, de descrição e de consulta das informações, somente o usuário humano, acessando especificamente o mecanismo de busca de tal site, conseguirá realizar as buscas necessárias, enquanto que um mecanismo de busca geral e externo não conseguirá “entender” sobre o quê tratam as informações disponibilizadas. Nesse sentido, o que falta na Web são descrições semanticamente mais ricas para que os mecanismos de busca possam recuperar informações com maior precisão.

Da necessidade de estruturar e de descrever o conteúdo das páginas ou mesmo de representar e de descrever uma entidade não presente no ambiente digital, surge a iniciativa da Web Semântica.

Desde seu início, o desenvolvimento da Web Semântica esteve próximo do desenvolvimento da Web, uma vez que Berners-Lee, o criador da Web, foi quem cunhou o termo Web Semântica (*Semantic Web*).

Mas, além da Web propriamente dita, a Web Semântica tem outras duas importantes raízes históricas: a construção de modelos abstratos que capturam as complexidades do

mundo em termos de ideias simples (modelagem conceitual) e a ideia de computação com o conhecimento, ou seja, o desejo de representar informações de um modo que permita que máquinas automaticamente cheguem às conclusões (inferência)<sup>1</sup> por meio do raciocínio lógico (ramo da computação conhecido como Inteligência Artificial<sup>2</sup>) (HITZLER; KRÖTZSCH; RUDOLPH, 2009).

Desse modo, pode-se dizer que os três tópicos seguintes provêm um entendimento básico das origens e da atual configuração da Web Semântica:

- A troca de informações entre computadores: a transmissão de recursos informacionais complexos entre computadores permitindo a distribuição, a ligação e a integração de informações em uma escala global;
- A construção de modelos conceituais: a questão de descrever um domínio de interesse em termos abstratos com o intuito de permitir um entendimento mais fácil de uma realidade complexa; e
- A inteligência artificial: a tentativa de construir máquinas que possuam raciocínio lógico para realizar conclusões (inferir) a partir de conhecimentos codificados.

A questão da troca de informações entre computadores foi abordada brevemente nesta seção. A fim de favorecer uma melhor compreensão acerca da construção de modelos conceituais e da inteligência artificial, apresentam-se as seções 2.1 e 2.2 que abordam esses tópicos com mais detalhes.

## 2.1 A construção de modelos conceituais

Segundo Hitzler, Krötzsch e Rudolph (2009), um modelo conceitual é uma descrição de certos aspectos da realidade, utilizado para entender, estruturar ou prever partes do

---

<sup>1</sup> “Na Web Semântica, inferência pode ser caracterizada como a descoberta de novos relacionamentos, isto é, considerando que os dados são modelados como um conjunto de relacionamentos entre recursos, inferência significa que procedimentos automatizados podem gerar novos relacionamentos baseados nos dados e em algumas informações adicionais fornecidas na forma de um vocabulário, isto é, de um conjunto de regras [lógicas]” (W3C, 2013, tradução nossa).

<sup>2</sup> As áreas de pesquisa mais proeminentes da Inteligência Artificial são: a) dedução, raciocínio, e solução de problemas realizados por máquinas; b) representação do conhecimento; c) planejamento realizado por máquinas; d) aprendizagem automática; e) processamento de linguagem natural (comunicação); e) percepção e; f) movimento e manipulação de objetos (RUSSELL; NORVIG, 2013; POOLE; MACKWORTH; GOEBEL, 1998).

mundo real. Normalmente os modelos conceituais são criados de modo científico, ou seja, por uma modelagem científica.

A modelagem científica tem suas raízes na Filosofia antiga. O filósofo grego Platão (428/427 a.C. - 348/347 a.C.) propôs respostas a algumas das questões mais fundamentais que surgem durante a modelagem conceitual: O que é a realidade? Que coisas podemos dizer que existem? Qual é a natureza verdadeira das coisas? Para Hitzler, Krötzsch e Rudolph (2009) isso marca a primeira grande contribuição para o campo agora conhecido como Ontologia.

No sentido filosófico, ontologia é o estudo da existência e dos seres como tais, as classes fundamentais e os relacionamentos entre as coisas existentes (GRUBER, 1993; HITZLER; KRÖTZSCH; RUDOLPH, 2009; SCHIESSL; BRÄSCHER, 2011; SOUZA; ALVARENGA, 2004; RAMALHO, 2006; RAMALHO, 2010).

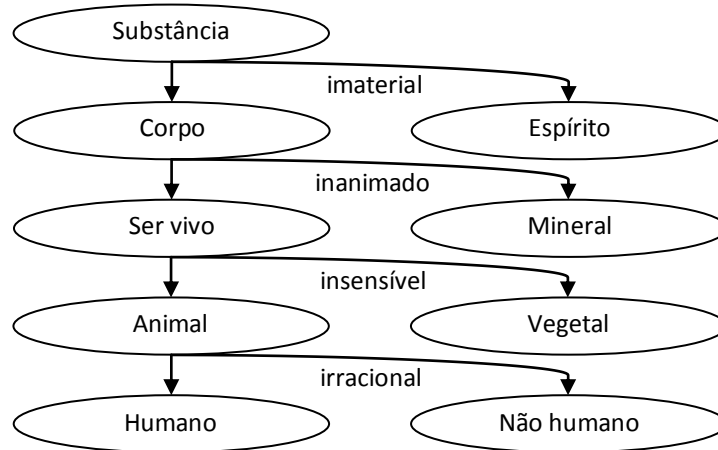
Porém, o termo ontologia, na Ciência da Computação e, por conseguinte, nas tecnologias da Web Semântica, possui outro significado: uma ontologia é uma descrição do conhecimento sobre um domínio de interesse na forma de uma especificação processável por máquinas com um significado formalmente definido (GRUBER, 1993; HITZLER; KRÖTZSCH; RUDOLPH, 2009).

O campo de ontologia, no sentido filosófico, recebeu significativas contribuições do filósofo e aluno de Platão, Aristóteles (ALMEIDA, 2002a; HITZLER; KRÖTZSCH; RUDOLPH, 2009; POMBO, 1998), que propôs dez categorias para classificar todas as coisas existentes e, para complementar, propôs também subcategorias para especificar ainda mais cada uma dessas coisas. Por exemplo, a categoria “animais”, criada por Aristóteles, é composta pelas subcategorias “animais racionais” (humanos) e “animais irracionais” (todos os outros animais). Tal como em classificações tradicionais, as subcategorias no modelo proposto pelo filósofo são exaustivas, isto é, cada coisa em uma categoria pertence a uma de suas subcategorias, e são mutuamente exclusivas, ou seja, cada coisa em uma categoria pertence somente a uma de suas subcategorias.

Hitzler, Krötzsch e Rudolph (2009) afirmam que essas antigas abordagens para a classificação científica também introduziram a utilização de estruturas na modelagem conceitual. Segundo os autores, o filósofo Porfírio (ca. 234-305) desenvolveu um pequeno modelo em forma de árvore, posteriormente chamado de Árvore de Porfírio, que captura os

relacionamentos hierárquicos de algumas das categorias propostas por Aristóteles. A Figura 1 é baseada na Árvore de Porfírio e ilustra as relações entre as categorias.

Figura 1 – Árvore de Porfírio: uma estrutura em árvore na modelagem conceitual.



Fonte: Traduzida de Hitzler, Krötzsch e Rudolph (2009, p. 3).

Estruturas em forma de árvores, hierarquias de conceitos e heranças de propriedades (características) são noções essenciais para numerosas tarefas de modelagem conceitual que ainda são encontradas em várias aplicações atualmente (HITZLER; KRÖTZSCH; RUDOLPH, 2009).

Numerosos e influentes modelos científicos foram desenvolvidos nos séculos seguintes, frequentemente construídos a partir da ideia básica da classificação encontrada nas obras de Aristóteles. Carolus Linnaeus (1707-1778), em sua obra *Systema Naturae*, forneceu a base para a classificação moderna da biologia, possibilitando, assim, classificar praticamente todas as formas de vida (BURKE, 2003; HITZLER; KRÖTZSCH; RUDOLPH, 2009). O termo taxonomia – composto pelos termos *taxis* (ordem) e *nomos* (lei, ciência) – se tornou o nome da ciência da classificação, embora seja também utilizado para se referir aos esquemas de classificação hierárquicos individuais. Outra importante classificação inclui a *International Statistical Classification of Diseases and Related Health Problems* (Classificação Estatística Internacional de Doenças e Problemas Relacionados com a Saúde - CID), a *Köppen climate classification* (Classificação climática de Köppen) e a classificação estelar, todas elas criadas no século XIX.

Como parte de um conhecimento antigo, o maior objetivo da moderna classificação ainda é o de ordenar objetos naturais para se estabelecer um melhor entendimento de sua

verdadeira natureza. Categorias cuidadosamente escolhidas geralmente são a base para se obter novos conhecimentos. Para se estudar uma nova espécie de animais, por exemplo, é necessário primeiramente classificar essa espécie para depois distingui-la de outras espécies (HITZLER; KRÖTZSCH; RUDOLPH, 2009).

Mas a classificação não é apenas importante quando se lida com objetos naturais e fenômenos. Esses mesmos métodos são similarmente úteis para se organizar objetos feitos por humanos, por exemplo, na ordenação de livros em uma biblioteca, uma prática já bastante conhecida pela Biblioteconomia, pela Documentação e pela Ciência da Informação. Um dos primeiros sistemas de classificação modernos para a organização de livros foi apresentado por Melvil Dewey (1851-1931) com sua *Dewey Decimal Classification* (DDC) – Classificação Decimal de Dewey (CDD) –, embora possam ser encontradas abordagens anteriores para a organização de livros em antigas bibliotecas na Ásia e na Europa. Tais classificações podem ser consideradas como precursoras das tecnologias semânticas para a busca e a recuperação em sistemas de informação (HITZLER; KRÖTZSCH; RUDOLPH, 2009).

Segundo Hitzler, Krötzsch e Rudolph (2009) a abundância de classificações taxonômicas sugere que estruturas hierárquicas são uma base e, possivelmente, a estrutura mais relevante na modelagem conceitual. Em muitos casos, no entanto, meras hierarquias não são suficientes para descrever um domínio de interesse e, desse modo, alguns modelos científicos não são baseados em estruturas hierárquicas. A célebre tabela periódica dos elementos químicos é um exemplo clássico de um modelo científico não hierárquico. Em outros casos, modelos hierárquicos e não hierárquicos são combinados. Um tesouro moderno, por exemplo, é um conjunto de termos em que se tem disponível a descrição dos relacionamentos entre os termos de um domínio de interesse, de modo que os mesmos estejam ordenados hierarquicamente baseados na significação mais geral até uma significação mais específica (relacionamentos do tipo hiperonímia e hiponímia) e não hierarquicamente por meio dos relacionamentos do tipo sinonímia (termo que possui o mesmo significado de outro termo) e antonímia (termo que possui o significado contrário ao de outro termo).

Diferentemente de muitas classificações taxonômicas, algumas abordagens para a modelagem científica permitem que objetos pertençam a mais de uma categoria específica. A classificação de Dewey ainda segue a abordagem tradicional: um livro, por exemplo, pode pertencer à categoria 636.8 se seu assunto principal for “gatos”, e essa classificação

determina sua posição em uma prateleira ou estante com livros sobre “animais domésticos” (636). Atribuir uma única posição para um livro é útil em uma biblioteca física, mas não muito útil em um catálogo digital. Conseqüentemente, sistemas de classificação modernos frequentemente permitem que um item seja classificado como pertencente à mais de uma categoria. Relacionada a essa abordagem está a classificação baseada em múltiplos aspectos independentes, as chamadas facetas, que permitem descrever um recurso pela combinação de critérios ao invés de uma única categoria. A navegação e a busca que ocorre nas bases de dados bibliográficos atuais, por exemplo, é frequentemente baseada em facetas, tais como título, autor e editora.

## 2.2 A inteligência artificial

Na seção anterior foram abordados alguns esforços tomados para se conseguir, por meio das classificações, armazenar e estruturar o conhecimento. No entanto, apesar de o conhecimento estar estruturado, somente os seres humanos, por meio da leitura, por exemplo, são capazes de acessar e de realizar deduções e inferências sobre esse conhecimento.

Os seres humanos possuem a capacidade de interpretar, conseguem distinguir o sentido das palavras em um determinado contexto, isto não ocorre nas ferramentas de busca e principalmente nos robôs e indexadores, responsáveis pela localização e extração de informações do conteúdo dos recursos necessárias para a construção da base de dados das ferramentas. Não conseguem distinguir semanticamente o contexto em que a informação está inserida, por este motivo que muitas vezes não recuperamos recursos que correspondem às nossas necessidades (ALVES, 2005, p. 26).

Por outro lado, Hitzler, Krötzsch e Rudolph (2009, p. 5) lembram que Aristóteles afirmou que o processo de dedução lógica, na maior parte das vezes realizado semiconscientemente pelos seres humanos, pode ser formalizado e externalizado pelo arranjo de uma série de regras que são utilizadas de um modo bastante semelhante ao processo realizado na aritmética. O exemplo de raciocínio lógico que se segue mostra que a terceira linha indica a conclusão das duas linhas precedentes a ela.

$$\begin{array}{l} \text{Todo A é B} \\ \text{Todo B é C} \\ \hline \text{Todo A é C} \end{array}$$

Essas regras, chamadas de silogismos, são independentes de domínio, ou seja, servem de base para inferir qualquer conhecimento, bastando apenas substituir os indivíduos por conceitos de um domínio qualquer, por exemplo:

$$\frac{\begin{array}{l} \text{Todo homem é humano} \\ \text{Todo humano é mortal} \end{array}}{\text{Todo homem é mortal}}$$

Assim, Aristóteles sugeriu que o pensamento lógico, assim como a matemática, tem sua própria existência, independente das opiniões e atitudes de pessoas individuais. A ideia de externalizar e padronizar o raciocínio humano apareceu esporadicamente nos séculos subsequentes, mas, somente a partir do século XIX, a captura lógica do pensamento em cálculos foi continuamente buscada, começando com George Boole (1815-1864) e sua lógica proposicional (lógica booleana) e passando por Gottlob Frege (1848-1925) e Charles Sanders Peirce (1839-1914) com a lógica de primeira e segunda ordem (HITZLER; KRÖTZSCH; RUDOLPH, 2009).

Segundo Hitzler, Krötzsch e Rudolph (2009), no verão de 1956, John McCarthy organizou uma série de brainstormings no Dartmouth College em Hanover (New Hampshire, EUA) com importantes pesquisadores da área da computação e da lógica. Inspirados pela novidade dos computadores digitais, eles exploraram a possibilidade de empregar esses dispositivos para simular ou gerar comportamento inteligente, assim, surgiu nessa época o termo inteligência artificial. A tarefa de deduzir conhecimento de fatos conhecidos (dados) foi considerada um dos problemas centrais para atingir esses objetivos.

No entanto, naquela época os dispositivos ainda eram bastante caros e limitados em termos de armazenamento, potência, rapidez, etc., por isso a tarefa de fornecer fatos (informações ou dados) aos computadores e de estes tirarem conclusões, isto é, gerar novas informações ou novos dados (conhecimento), era absolutamente inviável em larga escala (HITZLER; KRÖTZSCH; RUDOLPH, 2009).

Nesse sentido, foi a primeira vez que o problema de escalabilidade surgiu, e tal característica tem sido um dos requisitos centrais no desenvolvimento das tecnologias da Web Semântica desde seu surgimento (HITZLER; KRÖTZSCH; RUDOLPH, 2009).

Como consequência, as pesquisas focaram no alcance de objetivos mais modestos. Para domínios restritos de conhecimento, as abordagens tomaram o caminho de implementar formalismos moderadamente expressivos mas computacionalmente



processáveis. Tal abordagem funcionou muito bem dando origem aos chamados sistemas inteligentes (*expert systems*) – na grande maioria, sistemas baseados em regras, construídos para áreas altamente estruturadas como a Medicina e a Biologia (HITZLER; KRÖTZSCH; RUDOLPH, 2009).

Sistemas inteligentes proliferaram na década de 1980 e influenciaram a criação e o desenvolvimento da Web e, principalmente, da Web Semântica. Essa ligação entre a Web e a área de Inteligência Artificial é evidenciada pelo idealizador da Web, Berners-Lee, que é detentor da cátedra de fundador do *MIT Computer Science and Artificial Intelligence Laboratory* (Laboratório de Ciência da Computação e Inteligência Artificial do MIT) (SCHOROW, 2007).

O breve estudo realizado sobre as raízes históricas da Web Semântica objetivou contextualizar essa temática tendo em vista uma melhor compreensão acerca das definições e das tecnologias apresentadas a seguir.

### **2.3 Definições de Web Semântica**

Após a breve apresentação das origens da Web Semântica, convém abordar mais especificamente em que ela consiste.

O termo Web Semântica se popularizou a partir do artigo *The semantic Web: a new form of Web content that is meaningful to computers will unleash a revolution of new possibilities* publicado em 2001 na revista *Scientific American* de autoria de Berners-Lee, Hendler e Lassila, embora a ideia principal da Web Semântica (recursos e links que expressam seus relacionamentos) estivesse presente desde o início do desenvolvimento da Web (BERNERS-LEE, 1989, 1996, 1998, 1999; BERNERS-LEE; FISCHETTI, 2000; KOIVUNEN; MILLER, 2001).

Nesse artigo, a Web Semântica é descrita pelos autores como uma extensão da Web atual, na qual é dado à informação um significado bem definido, possibilitando uma melhoria no processo de cooperação entre pessoas e computadores (BERNERS-LEE; HENDLER; LASSILA 2001, p. 25).

Para que a Web Semântica funcione, Berners-Lee, Hendler e Lassila (2001, p. 26) afirmam que os computadores precisam ter acesso a coleções estruturadas de informações e a conjuntos de regras de inferência, para que possam ser utilizados para conduzir a máquina ao raciocínio lógico automatizado.

Segundo Berners-Lee, Hendler e Lassila (2001, p. 30), a estrutura da Web Semântica abrirá o conhecimento e o funcionamento da humanidade para análises significativas por agentes de software, proporcionando uma nova classe de ferramentas com as quais os seres humanos e as máquinas poderão viver, trabalhar e aprender juntos.

O principal órgão responsável pelo desenvolvimento da Web Semântica é o *World Wide Web Consortium* (W3C) – comunidade internacional liderada por Berners-Lee que desenvolve padrões abertos e que tem como principal objetivo levar a Web para seu potencial máximo, garantindo seu crescimento em longo prazo.

Segundo o W3C (2012c, tradução nossa),

Além da clássica “Web de documentos” o W3C está ajudando na construção de um conjunto de tecnologias para dar origem a uma “Web de dados”, dados esses que você encontra em qualquer banco de dados. O objetivo final da Web de dados é possibilitar que computadores realizem trabalhos mais úteis e desenvolver sistemas que possam apoiar interações confiáveis na rede. O termo “Web Semântica” se refere à visão do W3C sobre a Web de dados associados [*linked data*].

O W3C (2013c) explica que a Web Semântica é uma rede de dados. Há muitos dados que todos nós utilizamos todos os dias, mas esses dados ainda não fazem parte da Web. Por exemplo, posso ver meus extratos bancários na Web, minhas fotografias e meus compromissos em um calendário. Mas consigo ver minhas fotos em um calendário para saber o que estava fazendo quando as tirei? Posso ver as linhas dos meus extratos bancários em um calendário? Por que não? Porque não temos uma rede de dados. Porque os dados são controlados por aplicações e cada aplicação guarda esses dados para si mesma. A Web Semântica trata de duas coisas: de formatos comuns para a integração e a combinação de dados provenientes de diversas fontes e do registro das representações de como os dados se relacionam aos objetos do mundo real. Isso permite a uma pessoa ou a uma máquina começar a buscar em um banco de dados e então mover-se por entre um grande conjunto de bancos de dados que estão ligados não por cabos, mas por referências a uma mesma coisa.

De modo semelhante, Koivunen e Miller (2001, tradução nossa) afirmam que

O objetivo da [iniciativa da] Web Semântica é desenvolver padrões e tecnologias facilitadoras destinados a ajudar as máquinas a entenderem melhor as informações disponibilizadas na Web de modo que elas possam realizar descobertas mais ricas, integrar dados, navegar e automatizar tarefas. Com a Web Semântica, não só recebemos resultados mais precisos na busca por informações, mas também sabemos quando podemos

integrar informações de diferentes fontes, quais informações comparar e fornecer vários tipos de serviços automatizados em diferentes domínios, desde nossas futuras casas e bibliotecas digitais até nossos negócios eletrônicos e serviços de saúde.

Em âmbito nacional, Souza e Alvarenga (2004, p. 134) declaram que

O projeto da Web Semântica, em sua essência, é a criação e implantação de padrões (*standards*) tecnológicos para permitir este panorama, que não somente facilite as trocas de informações entre agentes pessoais, mas principalmente estabeleça uma língua franca para o compartilhamento mais significativo de dados entre dispositivos e sistemas de informação de uma maneira geral.

Para Alves (2005, p. 26), a proposta da Web Semântica é instituir um maior nível semântico na descrição dos recursos informacionais na rede, desse modo, proporcionando uma maior eficiência na recuperação de informações nos mecanismos de busca da Web.

Ramalho (2006, p. 37) afirma que a Web Semântica é

um título genérico que representa uma série de pesquisas que têm como principal objetivo possibilitar um melhor aproveitamento das potencialidades do ambiente Web, a partir do desenvolvimento de meios de organização de conjuntos específicos de informação e instrumentos de representação da informação.

Koivunen e Miller (2001) apontam que com a Web Semântica é possível associar informações descritivas semanticamente ricas, de maneira global, a qualquer recurso, tal como um metadado<sup>3</sup> que indique quem é o autor de um determinado recurso. Uma vez que essas informações descritivas estão associadas aos recursos, é possível realizar consultas como, por exemplo, quais recursos têm como autor o autor João da Silva, quais recursos são do tipo livro de ficção. Na Web Semântica atribui-se *Uniform Resource Identifiers* (URIs)<sup>4</sup> não só aos documentos, mas também às pessoas, aos conceitos e aos relacionamentos entre eles. No exemplo anterior, ao fornecer identificadores únicos para a pessoa João da Silva, para a função autor e para o conceito livro de ficção, explicita-se quem é a pessoa e qual é a relação existente entre ela e um tipo de recurso em particular. Além disso, pela atribuição de identificadores únicos globais, todas as pessoas que possuem o nome João da Silva podem ser diferenciadas umas das outras.

---

<sup>3</sup> Metadados são informações estruturadas que descrevem, explicam, proveem a localização ou, de alguma forma, tornam mais fácil a recuperação, a utilização ou o gerenciamento de recursos (NISO, 2004, p. 1, tradução nossa).

<sup>4</sup> Para mais informações sobre URIs, ver seção 2.4.1: *Uniform Resource Identifier* (URI).

Desse modo, a Web Semântica fornece meios para adicionar informações específicas à Web, para que esta possa, por exemplo, auxiliar na automação de serviços, na descoberta de e no relacionamento entre recursos. “O objetivo da iniciativa da Web Semântica é padronizar as tecnologias-chave que possibilitam o desenvolvimento não centralizado enquanto assegura que todas as peças se encaixem” (KOIVUNEN; MILLER, 2001, tradução nossa).

De acordo com Koivunen e Miller (2001), um dos princípios básicos para a construção da Web Semântica é a possibilidade de atribuição de URIs a qualquer coisa, por exemplo, pessoas, lugares e objetos. Esses URIs servem como identificadores para a menção das coisas referenciadas. A esse respeito, Ramalho (2006, p. 39) aponta que para identificar um lugar, tal como a biblioteca universitária da UNESP de Marília, por exemplo, é possível utilizar o URI de sua página Web. Para se referir a uma pessoa, é possível utilizar-se do URI da caixa de e-mail dessa pessoa. O Quadro 1 ilustra algumas coisas referenciadas e seus possíveis URIs na Web.

Quadro 1 – Exemplos de coisas referenciadas e seus possíveis URIs.

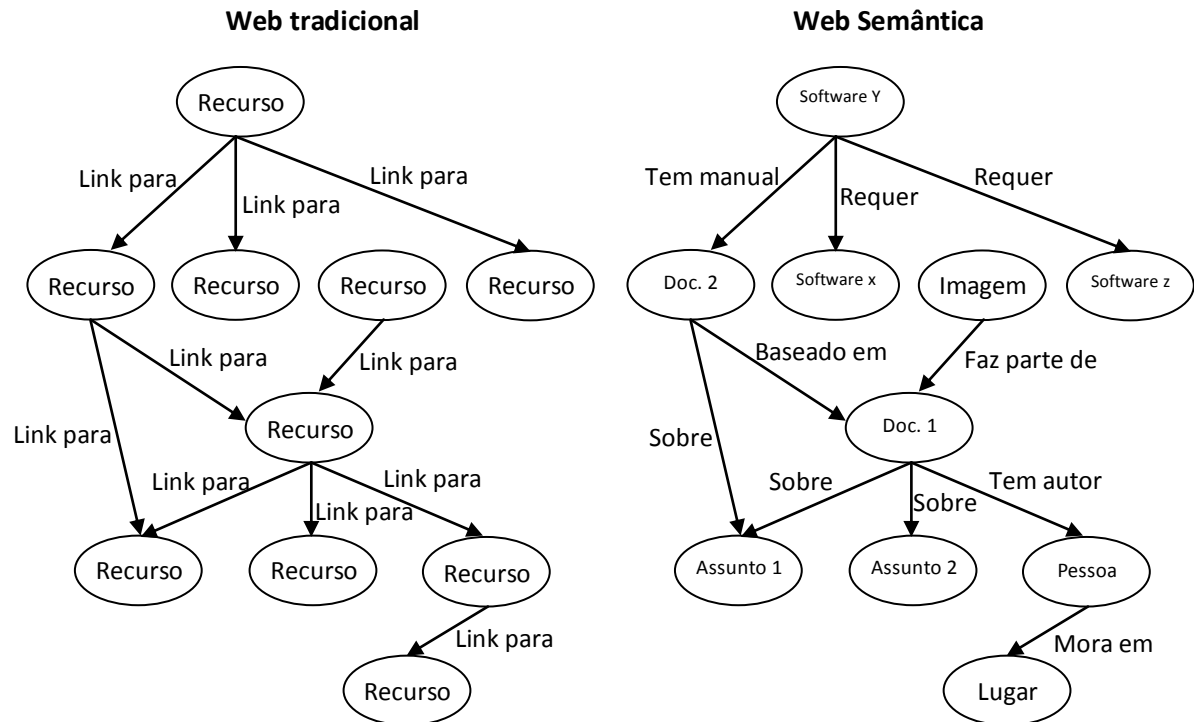
Coisa referenciada	URI
Unesp (organização)	<a href="http://www.unesp.br/#uri">http://www.unesp.br/#uri</a>
Jaider Andrade Ferreira (pessoa)	<a href="http://lattes.cnpq.br/1507095348642291">http://lattes.cnpq.br/1507095348642291</a>
Livro (conceito)	<a href="http://dbpedia.org/resource/Book">http://dbpedia.org/resource/Book</a>
Biblioteca (conceito)	<a href="http://www.wikidata.org/entity/Q7075">http://www.wikidata.org/entity/Q7075</a>
Livro Cibercultura, de Pierre Lévy (objeto)	<a href="http://wikincat.org/id/Cibercultura_(1999)">http://wikincat.org/id/Cibercultura_(1999)</a>
Publicador (conceito)	<a href="http://purl.org/dc/elements/1.1/publisher">http://purl.org/dc/elements/1.1/publisher</a>
Obra (conceito)	<a href="http://rdaregistry.info/Elements/c/Work">http://rdaregistry.info/Elements/c/Work</a>

Fonte: Elaborado pelo autor.

Ramalho (2006, p. 39), baseando-se em Koivunen e Miller (2001) e Miller (2004) aponta outra característica bastante importante na construção da Web Semântica: “links podem possuir diferentes tipos, possibilitando a definição de conceitos úteis para as máquinas”.

A Figura 2 ilustra a diferença entre os recursos e links tradicionais (chamados de *untyped links*), na Web tradicional, e os recursos e links que declaram seus relacionamentos (*typed links*), na Web Semântica.

Figura 2 – Diferença entre os recursos e links tradicionais e os propostos pela Web Semântica.



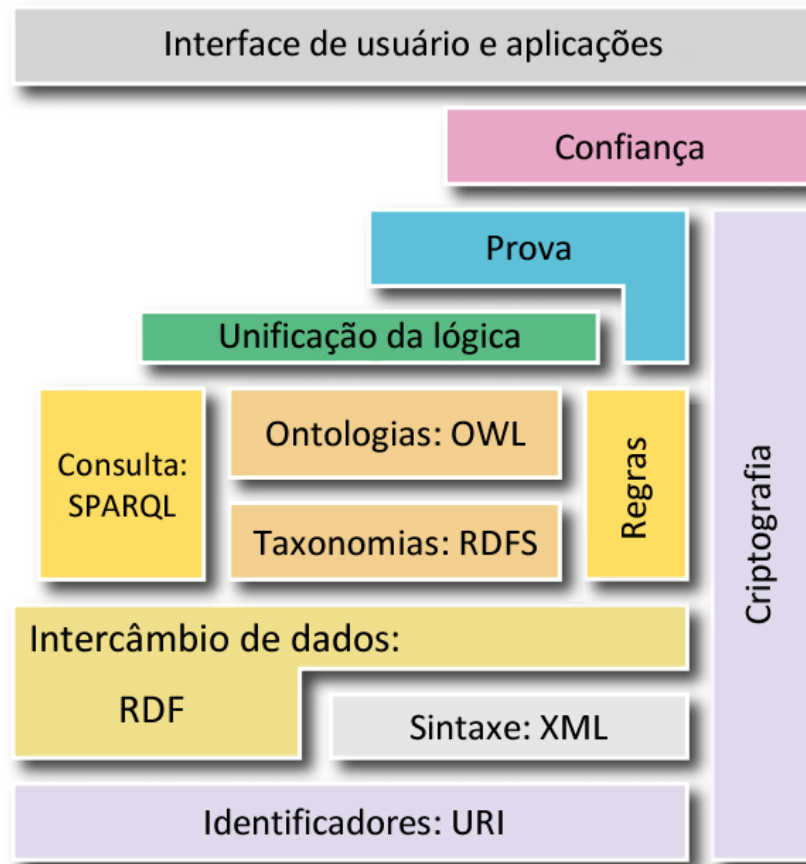
Fonte: Adaptada de Ramalho (2006, p. 40) e Miller (2004, p. 5-6).

Para que os recursos e os links entre esses recursos possam ser classificados em tipos (pessoas, documentos, lugares, etc.) o W3C tem desenvolvido uma gama de tecnologias que visam a possibilitar a codificação, a representação, a relação, a interoperabilidade, a contextualização e a segurança desses dados compartilhados na Web.

Assim, a Web Semântica é concretizada por meio da utilização de tecnologias Web e padrões abertos distribuídos em camadas com objetivos e funcionalidades específicas.

A Figura 3 ilustra uma das muitas representações acerca das camadas de tecnologias presentes na Web Semântica. Esse conjunto de tecnologias geralmente é chamado arquitetura da Web Semântica.

Figura 3 – Algumas tecnologias da Web Semântica.



Fonte: Adaptada de BRATT (2007, p. 24) e OBITKO (2007).

Embora existam várias outras tecnologias desenvolvidas e em desenvolvimento para a consolidação da Web Semântica, a Figura 3 apresenta apenas as tecnologias importantes para o prosseguimento desta pesquisa (URI, XML, RDF, RDFS, OWL e SPARQL).

O breve estudo realizado sobre as definições e características básicas da Web Semântica objetivou expandir o conhecimento geral acerca dessa importante iniciativa do W3C. Tendo em vista que as tecnologias e padrões da Web Semântica são condições indispensáveis para a consolidação de uma Web de dados, cada uma das tecnologias destacadas na Figura 3 (URI, XML, RDF, RDFS, OWL e SPARQL) será abordada a seguir.

## 2.4 Tecnologias da Web Semântica

A Web Semântica tem sido concebida como uma extensão da Web em que as máquinas podem buscar, combinar e processar os dados de uma forma inteligente, isto é, baseando-se no significado geral que esses dados possuem para os seres humanos. Para auxiliar no alcance desses objetivos, a iniciativa da Web Semântica estabelece padrões,

linguagens e estruturas que são utilizadas para representar, codificar, descrever, intercambiar e consultar dados para o seu posterior processamento em sistemas computacionais.

Segundo o W3C (2012c), as tecnologias da Web Semântica possibilitam às pessoas criarem e exporem seus dados na Web, construírem vocabulários e escreverem regras para a manipulação desses dados. Tais dados são relacionados e empoderados por tecnologias como a *Resource Description Framework (RDF)*, a *Web Ontology Language (OWL)* e a *SPARQL Protocol and RDF Query Language (SPARQL)*.

Santos e Alves (2009) argumentam que

as tecnologias da Web Semântica convergem para a área de Ciência da Informação, estabelecendo uma estreita relação na questão da representação do conhecimento, principalmente no que se refere ao uso de metadados considerados essenciais no estabelecimento dos requisitos para uma boa representação dos recursos informacionais na rede.

Tal afirmação vai ao encontro dos apontamentos de Souza e Alvarenga (2004, p. 134) sobre o propósito da Web Semântica:

Para atingir tal propósito, é necessária uma padronização de tecnologias, de linguagens e de metadados descritivos, de forma que todos [não necessariamente todos] os usuários da Web obedeçam a determinadas regras comuns e compartilhadas sobre como armazenar dados e descrever a informação armazenada e que esta possa ser “consumida” por outros usuários humanos ou não, de maneira automática e não ambígua. Com a existência da infra-estrutura tecnológica comum da Internet, o primeiro passo para este objetivo está sendo a criação de padrões para descrição de dados e de uma linguagem que permita a construção e codificação de significados compartilhados.

Algumas dessas tecnologias foram intensamente estudadas e apresentadas na literatura científica nacional e internacional da Ciência da Informação e não serão abordadas com profundidade nesta pesquisa. As seções seguintes deste capítulo realizarão uma breve apresentação dessas tecnologias, destacando apenas os aspectos relevantes para uma melhor compreensão acerca dos wikis semânticos.

#### **2.4.1 Uniform Resource Identifier (URI)**

Para representar um recurso em qualquer ambiente primeiramente é necessário identificá-lo. Na Web, a identificação de recursos é, de modo geral, realizada por meio de uma cadeia de caracteres chamada de *Uniform Resource Identifier (URI)*.

Ray (2001) aponta que existem duas categorias principais de URI: a primeira identifica um recurso com exclusividade com base em sua localização, enquanto que a segunda categoria apenas atribui ao recurso um nome exclusivo. Segundo o autor, um URI começa com um esquema – um nome curto que especifica como o recurso está sendo identificado. “Normalmente, ele é um protocolo de comunicações como HTTP ou FTP. Isso é seguido por um sinal de dois pontos (:) e uma sequência de caracteres que identifica o recurso com exclusividade” (RAY, 2001, p. 77).

Na Web Semântica é o URI que torna o recurso único e não confundível com outros recursos. Segundo Ramalho (2006, p. 71), os URIs possibilitam “uma maneira global e única de nomear itens”.

URIs englobam os URLs (*Uniform Resource Locators*)<sup>5</sup>, isto é, os endereços Web que são utilizados para acessar recursos online. Cada URL é um URI válido e, portanto, pode ser utilizado como um identificador em documentos que descrevem recursos disponíveis na Web. No entanto, Hitzler, Krötzsch e Rudolph (2009) chamam a atenção para o fato de que em várias aplicações o objetivo não é simplesmente a troca de informações ou metadados sobre recursos Web, mas sobre vários e diferentes tipos de recursos que estão, inclusive, fora da Web. Por meio da Web, é possível descrever qualquer objeto ou recurso que tenha uma clara identidade no contexto de uma dada aplicação: livros, pessoas, locais, editoras, eventos, relacionamentos entre essas coisas, conceitos abstratos e muito mais. Tais recursos podem, obviamente, não ser recuperados online e, portanto, seus URIs são utilizados exclusivamente para fins de identificação única.

Mesmo aqueles URIs que não se referem a recursos disponíveis na Web podem ser construídos seguindo a estrutura padrão de endereçamento Web (esquema://autoridade.com/caminho/a/ser/percorrido?consulta#fragmento).

Hitzler, Krötzsch e Rudolph (2009) afirmam que a principal característica de qualquer URI é a sua parte inicial, ou seja, seu esquema. Os esquemas como o *http* são tipicamente associados ao protocolo para a transmissão de informações na Web, entretanto, pode-se encontrar também tais esquemas em muitos URIs que não se referem a uma localização na

---

<sup>5</sup> Informações mais específicas sobre URIs podem ser encontradas em:

W3C/IETF. *URIs, URLs, and URNs: Clarifications and Recommendations 1.0*. Disponível em:

<<http://www.w3.org/TR/uri-clarification>>. Acesso em: 28 set. 2012.

W3C. *Cool URIs for the Semantic Web*. Disponível em: <<http://www.w3.org/TR/cooluris>>. Acesso em: 28 set. 2012.



Web. Desse modo, os detalhes do protocolo não são, obviamente, relevantes quando se utiliza um URI apenas como um nome, isto é, como um identificador. O URI `http://biblioteca-x.org/uri#livro-X`, como exemplo para essa situação, pode ser utilizado para se referir a um determinado livro em uma determinada biblioteca. O importante, no entanto, não é a recuperação ou não de algo por um navegador a partir desse URI, nem mesmo se o quê é recuperado tem ou não alguma relação com o livro em questão, mas sim a própria identificação do recurso<sup>6</sup>.

Muitas vezes os URIs são empregados na XML quando se utiliza um recurso chamado *namespace*. A XML e esse recurso são apresentados a seguir.

#### **2.4.2 Extensible Markup Language (XML)**

A *Extensible Markup Language* (XML) é uma tecnologia que descreve uma classe de objetos de dados chamados de documentos XML e que parcialmente descreve o comportamento dos programas de computador que processam esses documentos (W3C, 2008a). A XML foi criada por volta de 1998 com base na *Standard Generalized Markup Language* (SGML), uma linguagem padronizada de marcação genérica especificada na ISO 8879.

Documentos XML são construídos a partir de unidades de armazenamento chamadas de elementos. Os elementos contêm dados analisados sintaticamente que consistem de caracteres, alguns dos quais formam os dados de caracteres e alguns formam a marcação. A marcação codifica uma descrição do layout de armazenamento de um documento e sua estrutura lógica. A XML provê mecanismos para impor restrições no layout de armazenamento e na estrutura lógica do documento (W3C, 2008a).

Segundo Ray (2001, p. 1), a XML é um conjunto de ferramentas para “o armazenamento de dados, um veículo configurável para qualquer tipo de informação e um padrão aberto e em evolução”. Ray (2001, p. 1) apresenta algumas das principais características da XML:

- a XML pode armazenar e organizar praticamente qualquer tipo de informação em um formato adequado às várias necessidades;

---

<sup>6</sup> Para mais considerações acerca da utilização de URIs para a representação de recursos não recuperáveis pela Web, veja Berners-Lee (2009).

- como um padrão aberto, a XML não está ligada a qualquer empresa isolada ou acoplada a um software em particular;
- com o Unicode como seu conjunto de caracteres padrão, a XML aceita um grande número de sistemas de escrita e símbolos, desde caracteres rúnicos escandinavos até ideógrafos han chineses;
- a XML oferece várias maneiras de verificar a qualidade de um documento, com regras para sintaxe, verificação de vínculo interno, comparação com modelos de documento e tipos de dados;
- com sua sintaxe clara e simples e sua estrutura sem ambiguidades, a XML é relativamente fácil de ser lida e analisada por seres humanos e programas;
- a XML é facilmente combinada com folhas de estilo para criar documentos formatados em vários estilos desejáveis. Além disso, a pureza da estrutura da informação presente em documentos XML auxilia nas conversões de formato.

Ray (2001, p. 2) atenta para o fato de que, “apesar do nome, a XML não é uma linguagem de marcação por si só: é um conjunto de regras para a criação de linguagens de marcação.” Entende-se por linguagem de marcação “um conjunto de símbolos que pode ser colocado no texto de um documento para demarcar e rotular as partes desse documento”<sup>7</sup>.

Gama (2011, p. 46) afirma que a XML “não traz nenhuma predefinição de etiquetas para a estruturação do conteúdo a ser demarcado. Isso permite definição de elementos de metadados para qualquer conteúdo que se queira representar”.

Como exposto, a XML não é uma linguagem de marcação propriamente dita (como é o caso da HTML), mas sim um conjunto de regras sintáticas para a criação, desenvolvimento e utilização de linguagens de marcação. Segundo Ray (2001, p. 59-60), “essas regras são construções sintáticas muito simples”:

- os elementos precisam ter *tags* de início e fim, por exemplo, <nome>Texto marcado</nome>;
- a *tag* de um elemento vazio precisa ter uma barra (/) antes do sinal de maior, por exemplo, <elementoVazio/>;
- os valores dos atributos precisam estar entre aspas, por exemplo, <extensaoDoTexto tipo=“folhas”>;

---

<sup>7</sup> Para obter mais informações sobre linguagens de marcação, ver: Bax (2001) e Furgeri (2006).

- os elementos não podem se sobrepor, por exemplo, `<a>Texto de <b>exemplo</b> correto</a>`;
- caracteres de marcação isolados (`<`, `]]>`, e `&`) não podem aparecer no conteúdo analisado; quando necessários, tais caracteres são substituídos pelas devidas notações, por exemplo, `<formula>6 &lt; 8</formula>` (a notação “&lt;” equivale ao caractere “<”).

Uma regra final estipula que os nomes dos elementos só podem começar com letras e com o caractere de sublinhado e, além disso, os nomes só podem conter letras, números, hifens, pontos e sublinhados. Os sinais de dois-pontos são permitidos quando se utiliza um recurso da XML chamado *namespace*. Alguns nomes de elementos válidos são `<a></a>`, `<_a></_a>`, `<a.b></a.b>` e `<a-1></a-1>`.

Os *namespaces* em XML são largamente utilizados no contexto da Web Semântica. Segundo Ray (2001, p. 43), um *namespace* representa “um conjunto de nomes de elemento e atributos”, em outras palavras, um vocabulário. Um vocabulário, nesse contexto, é definido como um conjunto de propriedades ou termos descritivos, isto é, elementos de metadados definidos, por exemplo, por uma comunidade de descrição de recursos (MILLER, 1998). Assim, esses vocabulários de termos descritivos podem ser entendidos como padrões de metadados.

Para o W3C (2012f, p. 1, tradução nossa),

Na Web Semântica, os vocabulários definem os conceitos e os relacionamentos (também referidos como “termos”) utilizados para descrever e representar uma área de interesse. Vocabulários são utilizados para classificar os termos que podem ser usados em uma aplicação particular, caracterizam as relações possíveis e definem possíveis restrições sobre o uso desses termos. Na prática, os vocabulários podem ser muito complexos (com vários milhares de termos) ou muito simples (descrevendo um ou dois conceitos apenas).

Não há uma clara divisão entre o que é referido como “vocabulário” e o que é referido como “ontologia”. A tendência é utilizar a palavra “ontologia” para coleções de termos mais complexas e possivelmente mais formais, enquanto que “vocabulário” é utilizado quando tal formalismo estrito não é necessariamente usado ou é usado somente em um sentido mais brando. Os vocabulários são os blocos básicos para as técnicas de inferência na Web Semântica.

Um documento XML pode conter elementos e atributos, ou seja, termos, advindos de diferentes vocabulários, por exemplo, *title* (do vocabulário Dublin Core), *name* (do vocabulário FOAF), etc. Para indicar ao analisador sintático (*parser*) de qual vocabulário

advém cada elemento ou atributo é necessário anteceder seus nomes com prefixos de *namespaces* e o sinal de dois pontos (:). Desse modo, tem-se, por exemplo, *dc:title*, *foaf:name*, etc.

Cabe destacar que é necessário indicar a qual vocabulário corresponde cada prefixo de *namespace*. Essa indicação é encontrada no elemento raiz do documento XML, o qual contém o atributo *xmlns:* seguido por um prefixo de *namespace*, um sinal de igualdade e o URI do vocabulário a ser utilizado, por exemplo, *xmlns:dc= http://purl.org/dc/elements/1.1/*.

Assim, ao anexar o *namespace dc* ao elemento *creator* (`<dc:creator></dc:creator>`), por exemplo, declara-se ao analisador sintático que o termo “*creator*” está sendo utilizado como definido no vocabulário Dublin Core, não como um elemento *creator* qualquer, definido localmente.

Os *namespaces* possibilitam que elementos de padrões de metadados diferentes, por exemplo, possam ser utilizados em um mesmo “registro”, isto é, em um mesmo documento XML que descreve um ou vários recursos. Desse modo, os *namespaces* são amplamente utilizados na *Resource Description Framework* (RDF), tecnologia da Web Semântica apresentada a seguir.

### 2.4.3 Resource Description Framework (RDF)

Esta seção objetiva apresentar brevemente o modelo *Resource Description Framework* (RDF) como um importante mecanismo para a representação e a descrição de recursos.

Com o entendimento de que “a representação da informação é necessária em qualquer ambiente para proporcionar uma recuperação mais eficiente” (SANTOS; ALVES, 2009, p. 1), os desenvolvedores da Web têm procurado construir mecanismos e estruturas que possibilitem a atribuição de metadados aos recursos, visando à otimização do acesso à informação e à recuperação de recursos nas mais variadas áreas do conhecimento.

De acordo com Catarino e Souza (2012, p. 78), “o alicerce da Web Semântica está num modelo de descrição dos recursos da Web, o *Resource Description Framework* (RDF)” e, por lidar com a descrição de recursos, a Catalogação descritiva está intimamente relacionada a esse tema.

Segundo Ortega (2011), o termo Catalogação descritiva se refere aos aspectos da descrição formal dos recursos, incluindo a descrição física e a descrição dos elementos que

identificam os recursos, sendo que a atividade de descrição é conhecida na área de Biblioteconomia pelo nome da disciplina Catalogação.

Desse modo, concorda-se com Catarino e Souza (2012, p. 86) ao afirmarem que “a base da Web Semântica é a representação descritiva dos recursos a partir de um modelo em que os registros de metadados são representados por um conjunto de declarações [...]”.

Considerando o fato de que a Web oferece acesso sem precedentes às informações distribuídas globalmente e que os metadados – dados estruturados sobre dados – melhoram a descoberta e o acesso a tais informações, o uso efetivo de metadados entre aplicativos requer convenções comuns acerca de semântica, de sintaxe e de estrutura (MILLER, 1998).

A semântica dos metadados, ou seja, seu significado, é definida por cada comunidade de descrição de recursos (agências catalogadoras, por exemplo) que busca atender a sua necessidade específica de descrição. A sintaxe, disposição sistemática dos elementos de dados, facilita a troca e o uso de metadados entre vários aplicativos, enquanto que a estrutura pode ser considerada como uma restrição formal sobre a sintaxe para a representação consistente da semântica (MILLER, 1998).

Segundo Miller (1998), o RDF, desenvolvido pelo *World Wide Web Consortium (W3C)*, é uma infraestrutura que permite a codificação, o intercâmbio e o reuso de metadados estruturados. Essa infraestrutura permite a interoperabilidade de metadados por meio da concepção de mecanismos que suportam convenções comuns de semântica, de sintaxe e de estrutura. De acordo com o autor, o RDF não estipula a semântica para cada comunidade de descrição de recursos, mas sim oferece para essas comunidades a capacidade de definirem os elementos de metadados conforme as suas necessidades específicas de descrição.

O RDF pode utilizar-se da XML como sintaxe comum para o intercâmbio e o processamento de metadados. Utilizando-se da XML, o RDF impõe uma estrutura que proporciona a expressão não ambígua da semântica e, desse modo, possibilita a codificação, o intercâmbio e o processamento consistente de metadados padronizados (MILLER, 1998).

Além disso, Miller (1998) destaca que o modelo RDF suporta o uso de convenções que facilitam a interoperabilidade modular entre conjuntos separados de elementos de metadados. Essas convenções incluem mecanismos padrão para a representação que se baseiam em um simples, porém poderoso, modelo de dados. Adicionalmente, a *RDF Schema*, tecnologia abordada na seção 2.4.4, proporciona um meio para a publicação tanto de vocabulários legíveis por humanos quanto de vocabulários legíveis por máquinas.

Fazendo uso de termos de vocabulários definidos por meio de URIs, o RDF é um modelo de dados (*data model*) formal para a descrição de recursos, sendo que seu principal objetivo consiste em possibilitar que aplicações computacionais troquem (meta)dados enquanto preservam seu significado original (FURGERI, 2006; HITZLER; KRÖTZSCH; RUDOLPH, 2009).

Diferentemente da HTML [...], a principal intenção não é exibir os recursos corretamente, mas permitir o posterior processamento e a recombinação das informações que estão contidas nesses recursos, ou seja, os dados que compõem as suas representações. Desse modo, o RDF é conseqüentemente e frequentemente visto como o modelo básico de representação para o desenvolvimento da Web Semântica (HITZLER; KRÖTZSCH; RUDOLPH, 2009, p. 19, tradução nossa).

Desde o início dos anos 2000, a importância do RDF tem sido percebida pelas comunidades de descrição de recursos que, diante disso, o tem adotado para a formalização de seus termos descritivos. Como exemplo, cita-se a iniciativa *Dublin Core* (padrão ISO 15836-2009, ANSI/NISO Z39.85-2001, IETF RFC 5013-2007), comunidade internacional voltada para a descrição simples com foco na descoberta de recursos, que tem adotado o RDF desde 2002 (BECKETT; MILLER; BRICKLEY, 2002; NILSSON et al., 2008).

Segundo Miller (1998), a história dos metadados no W3C começou em 1995 com o desenvolvimento do padrão *Platform for Internet Content Selection* (PICS) que visava a possibilitar a classificação e a descrição do conteúdo de páginas Web. Com esse padrão seria possível descrever, por exemplo, que uma página Web continha conteúdo sexual, nudez, violência, palavras de baixo calão, etc. para que os pais, a partir de configurações do navegador, pudessem ter um controle sobre o que seus filhos poderiam acessar na Web. O desenvolvimento desse padrão, naquela época, foi motivado principalmente pela antecipação de restrições sobre o conteúdo da Internet nos EUA.

No entanto, no decorrer do desenvolvimento do PICS, percebeu-se a necessidade de fornecer descrições mais abrangentes acerca das páginas Web e, a partir de uma série de reuniões com a comunidade de bibliotecas digitais, limitações nas especificações do PICS foram identificadas e requisitos funcionais foram delineados para resolver o problema mais geral da associação de informações descritivas aos recursos presentes na Web. Com a percepção de que a questão da descrição de recursos era muito mais ampla do que a proposta feita inicialmente no contexto do PICS, o W3C criou um grupo de trabalho intitulado *Resource Description Framework* (RDF) para discutir especificamente uma

estrutura de descrição de recursos que fosse suficientemente abrangente para cobrir as necessidades das várias comunidades de descrição interessadas.

Nesse contexto, o processo de criação e de desenvolvimento do RDF na década de 1990 foi influenciado por várias linguagens, vocabulários de descrição já existentes e áreas do conhecimento. Uma das comunidades de descrição que influenciou a criação do RDF foi a da iniciativa *Dublin Core* (MILLER, 1998), outra influência, por exemplo, foi a XML.

Em 1999 uma especificação oficial sobre o RDF foi publicada pelo W3C (W3C, 1999), porém, a ênfase naquela época estava apenas na descrição de recursos presentes na Web e não na descrição de recursos em geral, como é o caso da especificação publicada em 2004 (W3C, 2004b).

O termo metadados geralmente se refere a “conjuntos de atributos, mais especificamente dados referenciais, que representam o conteúdo informacional de um recurso” (ALVES, 2005, p. 115). Em 1999 tais recursos eram considerados principalmente como páginas Web as quais o RDF poderia ajudar, por exemplo, na descrição de informações acerca da autoria ou do *copyright* dessas páginas. Posteriormente, com a visão da Web Semântica, o RDF foi estendido para auxiliar na descrição de qualquer recurso, indo além das simples descrições de recursos Web como sujeitos primários de tais descrições. Isso gerou, em 2004, a publicação da especificação retrabalhada e estendida do RDF (HITZLER; KRÖTZSCH; RUDOLPH, 2009; W3C, 2004b).

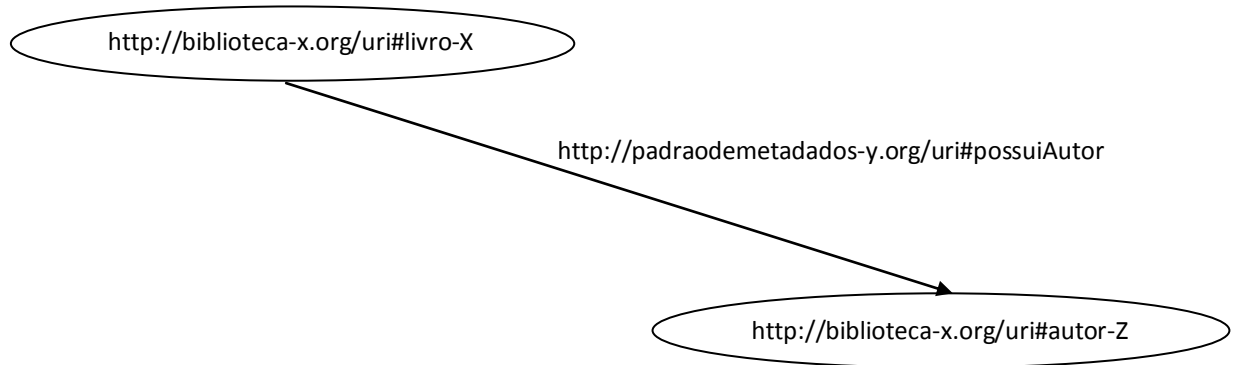
O modelo e a especificação da sintaxe RDF foram propostos em fevereiro de 1999 pelo W3C, com o intuito de possibilitar uma maior interoperabilidade no ambiente Web, oferecendo um padrão aberto para a descrição de recursos. Deste modo, o padrão RDF possibilita uma ampla gama de aplicações, permitindo que sejam feitas declarações a respeito de praticamente qualquer tipo de objeto, desde que este possa ser identificado a partir de um URI (RAMALHO, 2006, p. 70).

Devido à possibilidade de descrever recursos que não estão necessariamente disponíveis na Web – um livro impresso, por exemplo – o RDF também pôde ser utilizado para a troca de (meta)dados em diversas áreas de aplicação específicas, como é o caso da área de Biblioteconomia e, mais especificamente, da Catalogação.

Um documento RDF é estruturado em forma de grafo direcionado (também chamado de grafo orientado, grafo dirigido, dígrafo ou quiver), isto é, um conjunto de nós (ou vértices) que estão ligados por arestas direcionadas (setas). Em RDF, tanto os nós quanto

as arestas são rotulados com identificadores que os distinguem. A Figura 4 ilustra um exemplo de um grafo RDF simples com apenas dois nós e uma aresta.

Figura 4 – Um grafo RDF simples descrevendo o relacionamento entre um livro e seu autor.



Fonte: Adaptada de Hitzler, Krötzsch e Rudolph (2009, p. 20).

De acordo com Hitzler, Krötzsch e Rudolph (2009), é importante notar que, enquanto a informação codificada em XML apresenta uma estrutura em formato de árvore, o RDF apresenta a informação em forma de um grafo. O formato de árvore é perfeitamente adequado para organizar informações em documentos eletrônicos nos quais são frequentemente encontradas estruturas estritamente hierárquicas (documento > capítulo > corpo do texto > parágrafo > linha). Além disso, informações estruturadas em formato de árvore podem ser buscadas diretamente e serem processadas eficientemente. Mas, se o formato de árvore já é utilizado em XML e é tão eficiente, por que o RDF está estruturado em forma de grafos?

Uma razão importante para essa decisão é que o RDF não foi criado para a tarefa de estruturar documentos, mas sim para descrever relacionamentos genéricos entre objetos de interesse, ou seja, entre recursos de um domínio, criando declarações acerca de tais recursos (FURGERI, 2006; HITZLER; KRÖTZSCH; RUDOLPH, 2009).

A Figura 4, apresentada anteriormente, ilustra um grafo que expressa que o livro X possui como autor o autor Z. Percebe-se que o relacionamento entre o livro X e o autor Z é uma informação que não pertence a um relacionamento do tipo hierárquico (em forma de árvore) entre os recursos livro X e autor Z. Assim, o modelo de dados RDF considera tais relacionamentos como blocos básicos de informação. Quando há vários relacionamentos como esses, as imagens que se formam naturalmente geram grafos, não estruturas em



árvore, como as estruturas hierárquicas presentes em documentos XML (HITZLER; KRÖTZSCH; RUDOLPH, 2009).

Segundo os autores, outra razão pela escolha da utilização de grafos ao invés de árvores no modelo de dados RDF se deu pelo objetivo de tornar o RDF uma estrutura de descrição para os dados presentes na Web e em outras redes eletrônicas, isto é, de acordo com os autores, as informações disponíveis nesses ambientes são normalmente armazenadas e gerenciadas por meios descentralizados e, desse modo, torna-se mais fácil combinar dados que estão em RDF a partir de múltiplas fontes. Por exemplo, os grafos RDF de um site podem ser reunidos com os grafos RDF de outros sites, gerando um grafo maior que pode ou não prover novas informações úteis.

Segundo Furgeri (2006, p. 237), o RDF “elimina o problema da representação da informação em forma de árvore, criando uma estrutura mais flexível em forma de grafos, possibilitando a formação de uma cadeia de informações e estabelecendo uma rede de conhecimento”.

Como dito anteriormente, URIs permitem nomear e identificar recursos abstratos, mesmo aqueles recursos que não podem ser processados diretamente por um computador (pessoas, livros impressos, lugares, editoras, sintomas, etc.). Nesses casos, tais URIs são tratados como identificadores, referências ou nomes aos objetos pretendidos. Uma vez que os URIs podem ser tratados como nomes, a atual interpretação pretendida de um particular URI não é dada de nenhuma maneira formal e, portanto, ferramentas específicas podem ter seus próprios meios para interpretar URIs (HITZLER; KRÖTZSCH; RUDOLPH, 2009).

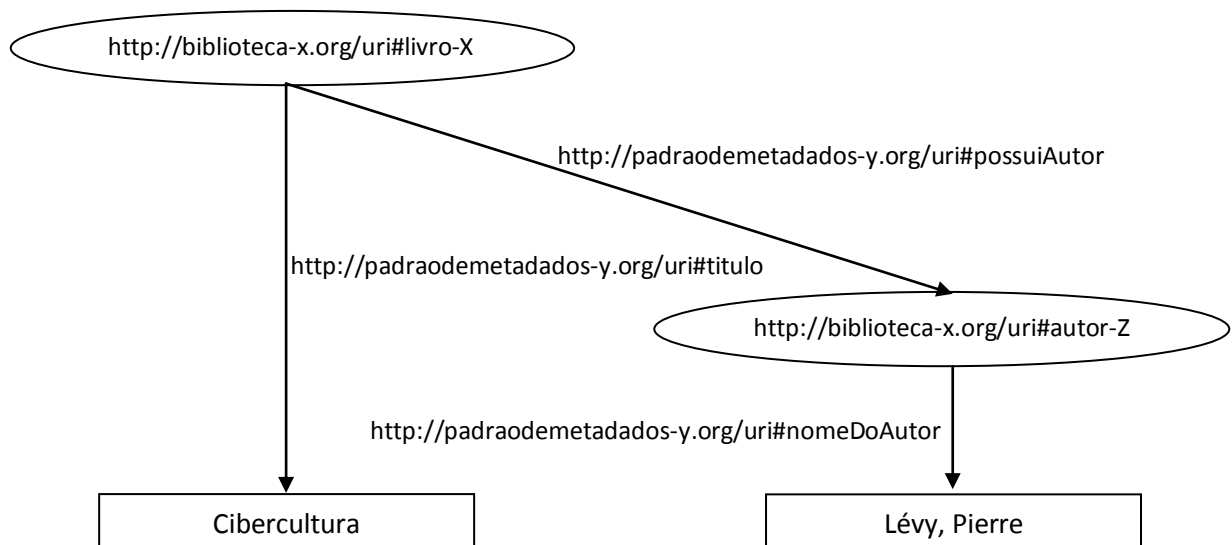
Nesse contexto, um serviço Web qualquer pode reconhecer URIs que se referem a livros e tratá-los de alguma forma especial, exibindo, por exemplo, detalhes de publicação ou mesmo os locais em que esses livros estão disponíveis para empréstimo. Segundo Hitzler, Krötzsch e Rudolph (2009), esse nível de liberdade é útil e inevitável quando se lida com recursos arbitrários. Porém, a situação se torna diferente quando se lida com valores de dados concretos como números, datas, valores absolutos ou cadeias de caracteres. Nota-se que em registros bibliográficos os dados encontrados são frequentemente cadeias de caracteres (*strings*).

Nesses casos, espera-se que cada aplicação tenha um entendimento mínimo dos significados concretos de tais valores. O número 36, por exemplo, tem a mesma interpretação numérica em qualquer contexto (HITZLER; KRÖTZSCH; RUDOLPH, 2009). Em

RDF, esses valores de dados são representados pelo que os profissionais da Ciência da Computação chamam de valores literais, ou apenas, literais. Literais são valores de dados de um certo tipo de dados (*datatype*). O valor de cada literal é geralmente descrito como uma sequência de caracteres, tais como a cadeia de caracteres composta pelos símbolos 3 e 6 do exemplo anterior. A interpretação de tais sequências realizada pela máquina é, então, baseada em um tipo de dados específico. Declarar o tipo de dados é essencial para o entendimento do significado pretendido: as cadeias de caracteres 36 e 036, por exemplo, referem-se ao mesmo número natural, no entanto, são diferentes cadeias de caracteres de texto (HITZLER; KRÖTZSCH; RUDOLPH, 2009).

Como pode ser visto na Figura 5, retângulos são utilizados para distinguir valores literais de URIs nos grafos RDF (FURGERI, 2006, p. 236).

Figura 5 – Um grafo RDF com literais para a descrição de valores de dados.



Fonte: Adaptada de Hitzler, Krötzsch e Rudolph (2009, p. 24).

Outra característica importante dos grafos RDF é que valores literais não podem ser utilizados como origem de arestas. Na prática, isso significa que não se pode fazer declarações acerca de literais, ou seja, as triplas permitidas em RDF seguem o padrão recurso-propriedade-valor, nunca valor-propriedade-valor, ou valor-propriedade-recurso. Assim, Hitzler, Krötzsch e Rudolph (2009) alertam que essa restrição precisa ser levada em consideração quando se planeja a modelagem de dados em RDF.

O modelo RDF é descrito em forma de grafos por meio de diagramas. Esse modo de representar os dados em RDF facilita o aprendizado do modelo e a leitura por humanos,

além de servir como um modo preciso para a modelagem conceitual de um domínio. Mas é evidente que os computadores não são adequados para o processamento e o intercâmbio de grafos uns com os outros. Para humanos, entender grafos visualmente funciona muito bem se os grafos são pequenos, com poucos nós, arestas e literais. Mas os gigantescos bancos de dados atuais contêm milhares de dados que gerariam grafos absurdamente grandes.

Para representar um conjunto de dados modelados em RDF, existem meios que não se utilizam de diagramas, mas sim de cadeias de caracteres que são processadas por máquinas.

Para que um grafo RDF seja transformado em cadeias de caracteres, ele precisa ser dividido em partes menores que são, então, armazenadas uma por uma. Essa transformação de estruturas de dados complexas para cadeias de caracteres lineares é chamada de serialização (HITZLER; KRÖTZSCH; RUDOLPH, 2009).

De acordo Hitzler, Krötzsch e Rudolph (2009), o primeiro passo para serializar um grafo RDF é isolar cada uma de suas relações, o que conduz a um recurso, a uma propriedade e a um valor para cada relação.

Tomando como exemplo a Figura 5, tem-se o recurso “<http://biblioteca-x.org/uri#livro-X>”, a propriedade “<http://padraodemetadados-y.org/uri#possuiAutor>” e o valor “<http://biblioteca-x.org/uri#autor-Z>”. Essas três partes distintas são chamadas de sujeito, predicado e objeto, respectivamente.

Percebe-se que cada grafo RDF pode ser completamente descrito por meio de seus nós e arestas e convertido em sujeitos, predicados e objetos. Cada conjunto sujeito-predicado-objeto é considerado uma tripla RDF, ou seja, uma declaração.

Furgeri (2006) explica que o recurso é o sujeito de uma declaração. Pode ser um website, um livro, um CD, etc. Um artigo científico é um exemplo de recurso. A propriedade é o predicado de uma declaração, é um atributo utilizado para descrever um recurso. Um artigo científico pode conter diversas propriedades: nome do autor, título do artigo, data de publicação, etc. O valor é o objeto de uma declaração e representa o conteúdo das propriedades. Seguindo o exemplo, trata-se dos conteúdos referentes ao nome do autor, ao título do artigo e a data de publicação.

Cabe observar que o sujeito e o predicado precisam ser URIs, enquanto que o objeto pode ser um URI, isto é, outro recurso, ou um valor literal (HITZLER; KRÖTZSCH; RUDOLPH, 2009).

Em relação à aparência de um grafo RDF, Furgeri (2006, p. 236) aponta que “o sujeito (recurso) é um nó [...] em forma de elipse, o predicado (propriedade) é um arco com uma seta apontando para o objeto (valor), cuja representação é um retângulo (se for um literal) ou uma nova elipse (se representar outro recurso)”.

Atualmente, a principal sintaxe utilizada para serializar um grafo RDF é a sintaxe XML (RDF/XML). Outras sintaxes possíveis são: Notation3 (N3), N-Triples, Turtle e JSON<sup>8</sup>. O grafo RDF da Figura 5 pode ser serializado com a sintaxe XML como mostrado na Figura 6.

Figura 6 – Serialização em XML de um grafo RDF que descreve o relacionamento entre um livro e um autor, o título do livro e o nome do autor.

---

```
<?xml version="1.0" encoding="utf-8"?>
<rdf:RDF xmlns:rdf="http://www.w3.org/1999/02/22-rdf-syntax-ns#"
  xmlns:cod="http://padraometadados-y.org/uri#">

  <rdf:Description rdf:about="http://biblioteca-x.org/uri#livro-X">
    <cod:titulo>Cibercultura</cod:titulo>
    <cod:possuiAutor>
      <rdf:Description rdf:about="http://biblioteca-x.org/uri#autor-Z">
        <cod:nomeDoAutor>Lévy, Pierre</cod:nomeDoAutor>
      </rdf:Description>
    </cod:possuiAutor>
  </rdf:Description>

</rdf:RDF>
```

---

Fonte: Elaborada pelo autor.

Souza e Alvarenga (2004) constatam alguns benefícios providos pela utilização do RDF, como, por exemplo, a viabilização de um ambiente consistente para a publicação e utilização de metadados por meio da infraestrutura da XML, o estabelecimento de uma sintaxe padronizada para a descrição de recursos com suas propriedades e respectivos valores e a possibilidade de aplicações agirem de modo inteligente e automatizado sobre as informações publicadas.

---

<sup>8</sup> Uma vez que este estudo visa apenas à apresentação do modelo RDF e de suas principais características, não serão apresentados detalhes técnicos de serializações ou de sintaxes para a codificação e o intercâmbio de documentos RDF.

Furgeri (2006, p. 237) destaca que o RDF

resolve o problema da diversidade na representação da informação que ocorre em XML, criando ligações únicas entre recursos e [com o auxílio do RDF *Schema*] estabelecendo vocabulários por meio de *namespaces*. A RDF elimina o problema da limitação do tamanho da estrutura enfrentada pela XML, criando ponteiros que unem documentos com estruturas menores.

Segundo o autor, o RDF é mais indicado para a criação de metadados, “pois um recurso é referenciado a um objeto por meio de um predicado com significado próprio” (FURGERI, 2006, p. 237). O autor aponta também que o RDF possui atribuições que lhe conferem diversas possibilidades de uso na área de Ciência da Informação, uma vez que esse modelo de dados permite criar relações entre recursos. O “RDF provê uma estrutura mais flexível que XML, aproximando-se da forma como os seres humanos relacionam informações, isto é, por associações” (FURGERI, 2006, p. 237).

Para Moura (2001 apud DZIEKANIAK; KIRINUS, 2004), na área de descoberta de recursos, o RDF possibilita o desenvolvimento de mecanismos de pesquisa mais eficientes. Na Catalogação, o RDF pode ser utilizado para a descrição de recursos de informação (LIBRARY OF CONGRESS, 2011, 2012; CASTRO, 2012) e, na interoperabilidade, facilita o intercâmbio de informações.

A partir dos apontamentos apresentados, pode-se considerar que a Ciência da Informação, especialmente em razão da Catalogação descritiva, está intimamente relacionada com um dos alicerces da Web Semântica: o modelo para a descrição de recursos *Resource Description Framework* (RDF).

Por meio de uma breve revisão de literatura acerca do modelo RDF, sua origem e suas principais características, percebe-se a sua importância para a atividade de descrição de quaisquer recursos, sejam eles digitais ou não, bibliográficos ou não.

Por fim, utilizando-se da XML como sintaxe comum para o intercâmbio e o processamento de metadados, o RDF colabora para a interoperabilidade entre os vários sistemas de informação e de descrição existentes contribuindo, desse modo, para a construção de mecanismos de busca mais integrados que permitirão a oferta de serviços mais especializados aos seus usuários.

#### 2.4.4 RDF Vocabulary Description Language: RDF Schema (RDFS)

Conforme visto anteriormente, com o RDF é possível expressar, de modo legível por máquinas, declarações (ou proposições) acerca de recursos únicos. Apesar de uma declaração RDF ser constituída de um sujeito, de um predicado e de um objeto, de certa forma, o objeto de uma declaração, caso não seja um valor literal, pode ser o sujeito de outra declaração. Isso leva a conclusão de que há pelo menos dois elementos essenciais em questão: um indivíduo (ora sujeito, ora objeto) e um predicado, ou, relação que conecta um indivíduo a outro.

Todavia, Hitzler, Krötzsch e Rudolph (2009) alegam que quando se pretende descrever domínios de interesse, são introduzidos novos termos não apenas para indivíduos (como “Cibercultura” ou “Pierre Lévy”) e suas relações (como “possui autor” ou “possui nome”), mas, também, para classes (como “livro”, “pessoa” ou “instituição”). Um repertório contendo termos que se referem a indivíduos, relações e classes é chamado de vocabulário, como abordado brevemente na seção 2.4.2 deste estudo<sup>9</sup>.

Quando uma pessoa utiliza-se de termos de um vocabulário (de propriedades descritivas, por exemplo, como o vocabulário *Dublin Core*), ela possui um entendimento sobre o significado desses termos. É intuitivamente claro para os seres humanos que o valor de uma propriedade intitulada criador, se refere a uma pessoa ou a uma instituição que criou algo, também é intuitivamente claro que uma instituição é composta por pessoas.

Porém, sob a perspectiva de um sistema computacional, todos os termos introduzidos são meramente cadeias de caracteres sem qualquer significado fixado previamente (SANTARÉM SEGUNDO; VIDOTTI, 2003; HITZLER; KRÖTZSCH; RUDOLPH, 2009). Assim, as inter-relações semânticas entre os termos “criador”, “pessoa”, “instituição”, etc., que são necessárias para a compreensão da declaração, precisam ser, de alguma forma, explicitamente e/ou formalmente comunicadas para o sistema, para que ele possa realizar inferências, baseadas em raciocínio lógico, que repousem nesse tipo de conhecimento subjacente, que é particularmente acessível aos humanos.

Esse tipo de conhecimento subjacente, chamado de conhecimento terminológico (ou conhecimento intensional), sobre os termos utilizados em um vocabulário pode ser

---

<sup>9</sup> Vocabulários complexos são chamados de ontologias (W3C, 2012f) (ver seção 2.4.2). Ontologias serão explicadas na seção 2.4.5.

explicitado por meio da linguagem *RDF Schema* (RDFS), uma parte adicional da especificação RDF (W3C, 2004c).

Segundo o W3C (2004c), as propriedades definidas no RDF podem ser consideradas como atributos de recursos e, nesse sentido, correspondem aos tradicionais pares atributo-valor. Além disso, essas propriedades podem representar também relacionamentos entre recursos.

No entanto, o RDF não fornece um mecanismo para descrever essas propriedades nem fornece qualquer mecanismo para descrever os relacionamentos entre essas propriedades e outros recursos. Essa é a função da linguagem de descrição de vocabulários RDF, *RDF Schema*. *RDF Schema* define classes e propriedades que podem ser utilizadas para descrever classes, propriedades e outros recursos (W3C, 2004c, tradução nossa).

Portanto, a RDFS não especifica um vocabulário de propriedades descritivas como “autor”, “título”, “publicador”, etc. Ao invés disso, ela especifica propriedades e classes que podem ser utilizadas para nomear e descrever propriedades e classes que descreverão recursos (W3C, 2004c).

Segundo Hitzler, Krötzsch e Rudolph (2009), a capacidade da RDFS de explicitar o conhecimento terminológico dos vocabulários caracteriza-a como uma linguagem de ontologia, pois oferece meios para descrever uma parte das interdependências semânticas existentes em um determinado domínio de interesse. Segundo os autores, “um documento RDFS é uma especificação processável por máquinas que descreve o conhecimento sobre algum domínio de interesse” (HITZLER; KRÖTZSCH; RUDOLPH, 2009, p. 47, tradução nossa).

Apesar da utilidade da RDFS como uma linguagem de ontologia, Hitzler, Krötzsch e Rudolph (2009) alertam que ela possui limitações e que, por isso, é às vezes categorizada como uma linguagem de representação para as chamadas ontologias leves (*lightweight ontologies*). Assim, aplicações mais sofisticadas requerem linguagens de representação mais expressivas como a *Web Ontology Language* (OWL), linguagem que será apresentada na seção 2.4.5.

Para Ramalho (2006, p. 49) a RDFS é utilizada para a descrição de vocabulários RDF, “possibilitando a definição de taxonomias de recursos em termos de uma hierarquia de classes”.

Segundo Furgeri (2006, p. 237-298), um vocabulário que faz uso da RDFS pode

[...] estabelecer relações e restrições entre os recursos. Relações tais como: o autor pertence a uma classe chamada pessoa, um site pertence a uma classe chamada recurso, um artigo que pertence a uma classe publicações deve possuir um título, cujo conteúdo deve ser um texto, e assim por diante. São criadas regras e restrições que devem ser seguidas pelos documentos RDF.

Com a utilização da RDFS é possível explicitar que um indivíduo (também chamado de instância) faz parte de um conjunto de indivíduos (classe) que compartilham alguma(s) característica(s). Um exemplo seria explicitar que o indivíduo intitulado “Cibercultura” faz parte da classe “Livro” ou explicitar que o indivíduo intitulado “Pierre Lévy” faz parte da classe “Autor”. Para esse tipo de explicitação, é utilizada a propriedade *rdf:type*.

Com a RDFS é possível criar hierarquias de classes para melhor representar uma determinada classe, melhorando o aspecto da granularidade utilizada. Assim, uma classe intitulada “Livro em braile” pode ser uma subclasse da classe intitulada “Livro” que por sua vez, pode ser uma subclasse da classe intitulada “Publicação”, e assim por diante. Para isso, utiliza-se a propriedade *rdfs:subClassOf*.

Esse tipo de explicitação permite, por exemplo, à máquina fazer inferências (isto é, derivar conhecimentos implícitos) do tipo “todos os indivíduos da classe ‘Livros em braile’ são ‘Publicações’”.

A RDFS também permite que sejam explicitados quais os indivíduos que são, na verdade, propriedades, desse modo, é possível declarar que o indivíduo intitulado “possuiAutor” é, na verdade, uma propriedade. Para essa declaração, é utilizada a propriedade *rdf:type* e a classe *rdf:Property*. Além disso, é possível explicitar que tal propriedade, quando utilizada em uma declaração, terá como sujeito um indivíduo da classe “Livro” e como objeto um indivíduo da classe “Autor”. Para isso, são utilizadas as propriedades *rdfs:domain* e *rdfs:range*.

Esse tipo de explicitação permite à máquina fazer inferências do tipo “‘Cibercultura’ é um ‘Livro’ e ‘Pierre Lévy’ é um ‘Autor’ já que ambos estão conectados pela propriedade ‘possuiAutor’”.

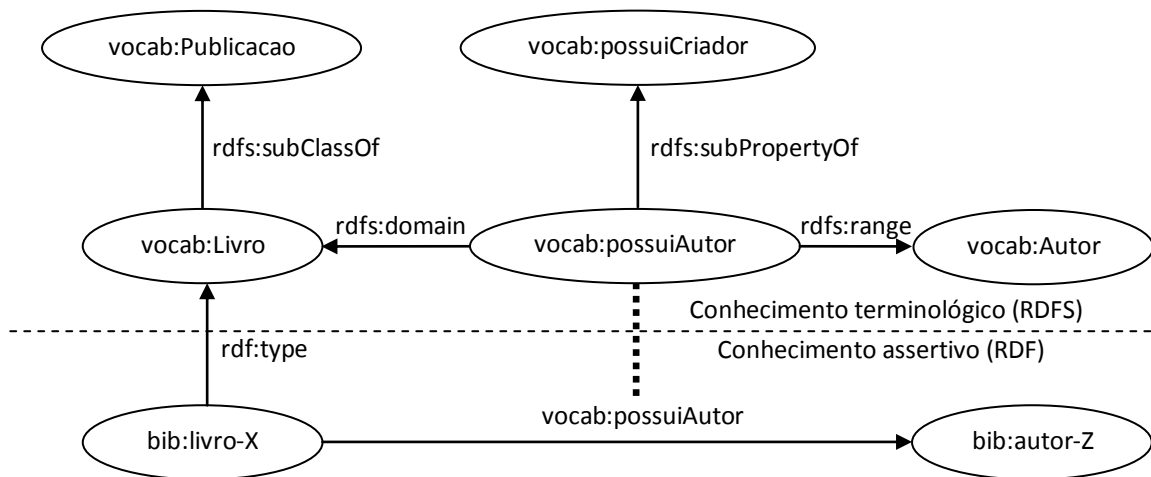
Propriedades também podem ser hierarquizadas, assim, uma propriedade intitulada “Indicação de responsabilidade” pode ser especificada como uma subpropriedade da propriedade intitulada “Área do título e indicação de responsabilidade”. Para tanto, utiliza-se a propriedade *rdfs:subPropertyOf*.



Portanto, o objetivo da especificação da RDFS é definir as primitivas requeridas para descrever classes, instâncias e relacionamentos (W3C, 2004c) e, desse modo, possibilitar uma expressividade semântica que habilita os sistemas computacionais a fazerem pequenas inferências sobre o conteúdo de um documento RDF.

A Figura 7 ilustra um pequeno vocabulário, em forma de um grafo RDF, que se utiliza da RDFS. A Figura 7 ilustra também a diferença entre o conhecimento assertivo (ou conhecimento extensional), expresso pelo RDF, e o conhecimento terminológico (ou conhecimento intensional), expresso pela RDFS. Os URIs foram abreviados com prefixos de *namespaces* de modo a facilitar a visualização.

Figura 7 – Explicitação do conhecimento assertivo (RDF) e terminológico (RDFS).



Fonte: Adaptada de Hitzler, Krötzsch e Rudolph (2009, p. 67).

Como dito anteriormente, aplicações mais sofisticadas requerem linguagens de representação mais expressivas como a OWL, linguagem apresentada a seguir.

#### 2.4.5 Web Ontology Language (OWL)

Conforme apresentado anteriormente, a RDFS permite a modelagem de ontologias simples (vocabulários, taxonomias) e a derivação de conhecimentos implícitos. Porém, a RDFS fornece um meio expressivo bastante limitado, não sendo possível, portanto, representar conhecimentos de ordem mais complexa (W3C, 2004c; BRESLIN; PASSANT; DECKER, 2009; HITZLER; KRÖTZSCH; RUDOLPH, 2009).

Segundo Hitzler, Krötzsch e Rudolph (2009), para fazer a modelagem de conhecimentos complexos, linguagens expressivas de representação baseadas na lógica

formal são comumente utilizadas. Isso permite a realização de raciocínio lógico no conhecimento utilizado e possibilita, desse modo, o acesso ao conhecimento que está somente implicitamente modelado. A linguagem *Web Ontology Language* (OWL) é uma linguagem que permite esse tipo de expressividade.

OWL é uma linguagem que teve a sua especificação publicada pelo W3C em 2004 (W3C, 2004a), com uma atualização (OWL 2) publicada em 2012 (W3C, 2012a), sendo um padrão recomendado para a modelagem de ontologias.

Segundo o W3C (2012b, tradução nossa),

*A OWL 2 Web Ontology Language* (OWL) do W3C é uma linguagem da Web Semântica projetada para representar conhecimento rico e complexo sobre coisas, grupos de coisas e relacionamentos entre coisas. OWL é uma linguagem computacional baseada na lógica, de modo que o conhecimento expresso em OWL pode ser raciocinado com e por programas de computador, tanto para verificar a consistência desse conhecimento quanto para tornar o conhecimento implícito em explícito. Documentos OWL, conhecidos como ontologias, podem ser publicados na *World Wide Web* e podem se referir a ou serem referidos por outras ontologias OWL.

Ainda segundo o W3C (2012b), OWL é uma linguagem para expressar ontologias, sendo que o termo ontologia tem uma história complexa tanto dentro quanto fora da Ciência da Computação.<sup>10</sup>

No contexto da Ciência da Computação e da Ciência da Informação, uma ontologia define um conjunto de primitivas representacionais com as quais se modela um domínio do conhecimento ou discurso. As primitivas representacionais são tipicamente classes (ou conjuntos), atributos (ou propriedades) e relacionamentos (ou relações entre os membros das classes). As definições das primitivas representacionais incluem informações sobre seus significados e restrições em suas aplicações logicamente consistentes. No contexto dos sistemas de banco de dados, uma ontologia pode ser vista como um nível de abstração do modelo de dados, análogo aos modelos hierárquicos e relacionais, mas planejado para modelar o conhecimento sobre indivíduos, seus atributos e seus relacionamentos com outros indivíduos. Ontologias são tipicamente especificadas em linguagens que permitem a abstração longe das estruturas de dados e das estratégias de implementação; na prática, as linguagens de ontologias são mais próximas, em poder de expressividade, à

---

<sup>10</sup> Salienta-se que o termo ontologia é interpretado por diferentes formas na Filosofia, na Linguística, na Ciência da Computação e na Ciência da Informação. Essa característica, muitas vezes, torna difícil o aprendizado do tema nessas áreas do conhecimento. A definição de ontologia apresentada por Gruber em 1993 tem sido frequentemente utilizada em estudos em Ciência da Informação. Porém, neste estudo, adota-se a definição de Gruber de 2009, pois consiste em uma versão atualizada e revisada da definição anterior. Para outras interpretações do termo, recomenda-se a leitura do artigo de Moreira, Alvarenga e Oliveira (2004) e da dissertação e tese de Ramalho (2006, 2010).

lógica de primeira ordem do que as linguagens utilizadas para modelar bancos de dados. Por essa razão, ontologias são consideradas no nível “semântico”, enquanto que esquemas de banco de dados são modelos de dados considerados no nível “lógico” ou “físico”. Devido a sua independência dos modelos de dados de nível inferior, ontologias são utilizadas para integrar banco de dados heterogêneos, possibilitando a interoperabilidade entre sistemas diferentes e especificando interfaces para serviços independentes baseados em conhecimento. No conjunto de tecnologias dos padrões da Web Semântica, ontologias estão na camada explícita. Há atualmente linguagens padrão disponíveis e uma variedade de ferramentas *open source* e comerciais para criar e trabalhar com ontologias (GRUBER, 2009, tradução nossa).

Ramalho (2006, p. 97) define ontologia como sendo

Um artefato tecnológico que descreve um modelo conceitual de um determinado domínio em uma linguagem lógica e formal, a partir da descrição dos aspectos semânticos de conteúdos informacionais, possibilitando a realização de inferências automáticas por programas computacionais.

No sentido apresentado por essas definições de ontologia, utiliza-se aqui o termo para significar certo tipo de artefato computacional, geralmente apresentado como um documento, contendo um conjunto de declarações descritivas precisas sobre alguma parte do mundo, normalmente referida como o domínio de interesse ou o domínio do discurso da ontologia.

Para descrever precisamente um domínio de interesse, é necessário ter um conjunto de termos centrais – comumente chamado de vocabulário – e determinar seus significados. Além de uma definição concisa em linguagem natural, o significado de um termo pode ser definido por meio de declarações que expressem como esse termo está inter-relacionado com outros termos<sup>11</sup>. Além desse conhecimento terminológico, uma ontologia também pode conter o conhecimento assertivo que lida com os objetos concretos do domínio considerado, ao invés de apenas noções gerais<sup>12</sup> (W3C, 2012b).

Tal como a RDFS, a OWL é uma linguagem de representação do conhecimento destinada a formular, intercambiar e raciocinar com o conhecimento sobre um domínio de interesse. Uma ontologia expressa em OWL busca capturar o conhecimento de um domínio de interesse e comunicar esse conhecimento para outras aplicações, porém, esse tipo de

<sup>11</sup> Ver seção 2.4.4, *RDF Vocabulary Description Language: RDF Schema (RDFS)*, para entender sobre as interdependências semânticas existentes entre termos de um domínio de interesse.

<sup>12</sup> Quando uma ontologia também inclui o conhecimento assertivo (ABox), além do conhecimento terminológico (TBox), esse artefato tecnológico é chamado de base de conhecimento (*knowledge base*) (BRESLIN; PASSANT; DECKER, 2009, p. 58).

conhecimento que é expresso pela OWL não reflete, obviamente, todos os aspectos do conhecimento humano. As declarações realizadas por meio da linguagem OWL são chamadas de axiomas. Todos os constituintes atômicos das declarações, sejam eles objetos (João, Maria), categorias (Macho, Fêmea) ou relações (casado com, divorciado de), são chamados de entidades. Em OWL denotam-se objetos como indivíduos, categorias como classes e relações como propriedades. Propriedades são ainda subdivididas em propriedades de objeto (*object properties*) e propriedade de tipos de dados (*datatype properties*) (W3C, 2012b).

Com classes, propriedades e indivíduos, a modelagem básica feita com a OWL permite realizar diversas declarações, tais como as mostradas no Quadro 2.

Quadro 2 – Exemplos de modelagens básicas realizadas com a OWL.

<b>Modelagem</b>	<b>Exemplo de utilização</b>	<b>Propriedades, classes e tipos de dados necessários</b>
Uma determinada entidade é um indivíduo (não uma classe, propriedade ou tipo de dado)	João é um indivíduo	<i>rdf:type</i> e <i>owl:NamedIndividual</i>
Uma determinada entidade é uma classe	Homem é uma classe	<i>rdf:type</i> e <i>owl:Class</i>
Uma determinada entidade é uma propriedade que relaciona indivíduos	a propriedade casado com relaciona dois indivíduos	<i>rdf:type</i> e <i>owl:ObjectProperty</i>
Uma determinada entidade é uma propriedade que relaciona indivíduos a valores concretos, ou seja, valores de tipos de dados	a propriedade nome relaciona um indivíduo a um valor concreto	<i>rdf:type</i> e <i>owl:DatatypeProperty</i>
Uma determinada entidade faz parte de uma classe	João é um Homem	<i>rdf:type</i>
Uma determinada classe é subclasse de outra	Homem é uma Pessoa	<i>rdfs:subClassOf</i>
Uma determinada classe equivale à outra	Pessoa equivale a Humano	<i>owl:equivalentClass</i>
Uma determinada propriedade equivale à outra	casa publicadora equivale a editora	<i>owl:equivalentProperty</i>
Uma determinada entidade que pertence a uma classe não pode pertencer à outra(s)	Maria ou é Mulher ou é Homem	<i>owl:disjointWith</i> ou <i>owl:AllDisjointClasses</i> e <i>owl:members</i>
Uma determinada entidade está relacionada à outra	Cibercultura possui como criador Pierre Lévy	<i>dc:creator</i> por exemplo
Um determinado indivíduo não está relacionado a outro determinado indivíduo por uma determinada propriedade	João não tem como esposa Joana	<i>owl:NegativePropertyAssertion</i> , <i>owl:sourceIndividual</i> , <i>owl:assertionProperty</i> e <i>owl:targetIndividual</i>

<b>Modelagem</b>	<b>Exemplo de utilização</b>	<b>Propriedades, classes e tipos de dados necessários</b>
Uma determinada propriedade é subpropriedade de outra	esposa é subpropriedade de cônjuge	<i>rdfs:subPropertyOf</i>
Uma determinada propriedade relaciona um tipo de indivíduo a outro	a propriedade esposa relaciona um Homem a uma Mulher	<i>rdfs:domain</i> e <i>rdfs:range</i>
Um determinado indivíduo é diferente de outro	João não é José	<i>owl:differentFrom</i> ou <i>owl:AllDifferent</i> e <i>owl:distinctMembers</i>
Um determinado indivíduo é o mesmo que outro	João S. é o mesmo indivíduo que João da Silva	<i>owl:sameAs</i>
Um determinado indivíduo não tem um determinado valor concreto em uma determinada propriedade	João não tem 53 anos	<i>owl:NegativePropertyAssertion</i> , <i>owl:sourceIndividual</i> , <i>owl:assertionProperty</i> e <i>owl:targetValue</i>
Indivíduos pertencentes a uma determinada classe precisam ter um determinado tipo de dado definido em uma determinada propriedade	Pessoas só podem ter valores inteiros e não negativos na propriedade idade	<i>rdfs:domain</i> , <i>rdfs:range</i> e <i>xsd:nonNegativeInteger</i> por exemplo

Fonte: Elaborado pelo autor.

Com relacionamentos complexos entre classes e usos avançados de propriedades e tipos de dados, é possível realizar declarações mais complexas como, por exemplo, as mostradas no Quadro 3.

Quadro 3 – Exemplos de modelagens complexas realizadas com a OWL.

<b>Modelagem</b>	<b>Exemplo de utilização</b>	<b>Propriedades, classes e tipos de dados necessários</b>
Uma classe equivale à intersecção de duas ou mais classes	Esposas são indivíduos que são Mulheres e que são Casadas	<i>owl:equivalentClass</i> , <i>rdf:type</i> , <i>owl:Class</i> e <i>owl:intersectionOf</i>
Uma classe equivale à união de duas ou mais classes	Cônjuges são aqueles indivíduos que são Esposos ou Esposas.	<i>owl:equivalentClass</i> , <i>rdf:type</i> , <i>owl:Class</i> e <i>owl:unionOf</i>
Uma classe equivale à intersecção de uma determinada classe com os não membros de outra determinada classe	Pessoas Sem Filhos são aqueles indivíduos que são Pessoas, mas que não são Pais	<i>owl:equivalentClass</i> , <i>rdf:type</i> , <i>owl:Class</i> , <i>owl:intersectionOf</i> e <i>owl:complementOf</i>
Uma classe precisa ser subclasse da intersecção de duas ou mais classes	Para ser um Avô um indivíduo precisa ser Homem e Pai	<i>rdfs:subClassOf</i> , <i>rdf:type</i> , <i>owl:Class</i> e <i>owl:intersectionOf</i>

<b>Modelagem</b>	<b>Exemplo de utilização</b>	<b>Propriedades, classes e tipos de dados necessários</b>
Uma entidade pertence a uma classe equivalente a intersecção de uma determinada classe com a classe de não membros de outra determinada classe	João é uma Pessoa que Não É Pai	<i>rdf:type, owl:Class, owl:intersectionOf e owl:complementOf</i>
Uma classe equivale àquela em que seus indivíduos (cada um) têm pelo menos algum valor em uma determinada propriedade	Pais são aqueles indivíduos que possuem pelo menos um filho	<i>owl:equivalentClass, rdf:type, owl:Restriction, owl:onProperty e owl:someValuesFrom</i>
Uma classe equivale àquela em que seus indivíduos possuem uma relação com outros indivíduos que estão, todos eles, em outra determinada classe	Pessoa Felizes são aqueles indivíduos que possuem todos os seus filhos Felizes	<i>rdf:type, owl:Class, owl:equivalentClass, owl:Restriction, owl:onProperty e owl:allValuesFrom</i>
Uma determinada propriedade é o inverso de outra	tem pai é o inverso de tem filho	<i>owl:inverseOf</i>
Uma classe equivale àquela em que seus indivíduos têm um determinado valor em uma determinada propriedade inversa	Órfãos são aqueles indivíduos que tem o inverso de "tem filho" (isto é, "tem pais") com todos os valores pertencentes à classe Mortos	<i>owl:equivalentClass, rdf:type, owl:Restriction, owl:onProperty, owl:inverseOf e owl:allValuesFrom</i>
Uma determinada propriedade é simétrica: a relação entre A e B é a mesma que entre B e A	João é irmão de José do mesmo modo que José é irmão de João	<i>rdf:type e owl:SymmetricProperty</i>
Uma determinada propriedade é assimétrica: a relação entre A e B não pode ser a mesma que entre B e A	João tem como filha Joana, mas Joana jamais pode ter como filho seu próprio pai, João	<i>rdf:type e owl:AsymmetricProperty</i>
Um indivíduo que possui uma determinada propriedade com um determinado valor de propriedade não pode possuir outra determinada propriedade com o mesmo valor da primeira propriedade	Maria é esposa de João ou é mãe de João, não as duas coisas	<i>owl:propertyDisjointWith ou owl:AllDisjointProperties e owl:members</i>
Uma determinada propriedade é reflexiva: relaciona tudo a ela mesma	Todo mundo é parente de si mesmo	<i>rdf:type e owl:ReflexiveProperty</i>
Uma propriedade é irreflexiva: nenhum indivíduo pode ser relacionado a si mesmo com tal propriedade	Ninguém é pai de si mesmo	<i>rdf:type e owl:IrreflexiveProperty</i>

Modelagem	Exemplo de utilização	Propriedades, classes e tipos de dados necessários
Uma determinada propriedade é funcional: A só pode ter zero ou um valor em uma determinada propriedade. Assim, se A está relacionado a B e também a C, B e C são o mesmo indivíduo porque A só pode estar relacionado a zero ou um indivíduo com essa propriedade	Maria esta casada com João da Silva, se ela estiver casada também com João S., então João da Silva e João S. são o mesmo indivíduo	<i>rdf:type</i> e <i>owl:FunctionalProperty</i>
Uma determinada propriedade é funcional inversa: B só pode ser o valor de uma determinada propriedade para zero ou um indivíduo. Assim, Se B é o valor de uma determinada propriedade tanto para A quanto para C, então A e C são o mesmo indivíduo	Maria é esposa de João da Silva, se Maria é também esposa de João S., então, João da Silva e João S. são o mesmo indivíduo	<i>rdf:type</i> e <i>owl:InverseFunctionalProperty</i>
Uma determinada propriedade é transitiva: se o indivíduo A está relacionado a B e B está relacionado a C com essa mesma propriedade, então, A também está relacionado a C com essa mesma propriedade	João Pai é ancestral de João Filho, que por sua vez, é ancestral de João Neto, logo, João Pai é ancestral de João Neto	<i>rdf:type</i> e <i>owl:TransitiveProperty</i>
Uma classe é subclasse daquela em que os indivíduos possuem um determinado valor mínimo, máximo ou exato em uma determinada propriedade	Uma Defesa de Mestrado precisa ter no mínimo (ou no máximo, ou exatamente) três professores	<i>owl:Class</i> , <i>rdfs:subClassOf</i> , <i>owl:Restriction</i> , <i>owl:onProperty</i> , <i>owl:minCardinality</i> ou <i>owl:maxCardinality</i> ou <i>owl:cardinality</i> e <i>xsd:nonNegativeInteger</i> por exemplo

Fonte: Elaborado pelo autor.

Tais combinações de mecanismos e termos fornecidos pelas tecnologias XML, RDF, RDFS e OWL permitem o compartilhamento de conhecimentos complexos acerca de um domínio de interesse.

Para Breslin, Passant e Decker (2009, p. 61, tradução nossa), a “OWL estende a noção de classes e propriedades definidas na *RDF Schema* e fornece novos axiomas para definir características avançadas e restrições em relação a classes e propriedades”.

Segundo Hitzler, Krötzsch e Rudolph (2009), encontrar um equilíbrio entre a expressividade da linguagem de um lado e o raciocínio lógico eficiente (escalabilidade) do

outro, foi central para o desenvolvimento da OWL. Isso ocorreu porque, dadas observações gerais, construções complexas de linguagens para a representação de conhecimentos implícitos produzem altas complexidades computacionais. A fim de fornecer ao desenvolvedor uma escolha entre diferentes graus de expressividade, três sub-linguagens da OWL foram definidas: OWL Lite, OWL DL e OWL Full.

OWL Lite estende a RDFS provendo a possibilidade de expressar novos axiomas como simetria e cardinalidade (embora a cardinalidade em OWL Lite seja restrita a zero e um) e é a menos expressiva das três sub-linguagens.

OWL DL (*Description Logics*) contém a OWL Lite, adiciona a possibilidade de expressar mais axiomas como intersecção, união e disjunção entre classes, bem como amplia as restrições de cardinalidade oferecidas na OWL Lite, é largamente suportada pela maioria das ferramentas de software e, segundo Santarem Segundo (2010, p. 128),

dá suporte aos usuários que desejam o máximo de expressividade sem perder a completude computacional (todas as conclusões são garantidas de serem computadas) e capacidade de decisão (todas as computações em um tempo finito) dos sistemas de raciocínio. O OWL DL inclui todos os construtores da linguagem OWL, com restrições, como separação entre tipos [de entidades] (uma classe não pode ser ao mesmo tempo um indivíduo ou tipo [de dados], e uma propriedade não pode ser ao mesmo tempo um indivíduo ou uma classe). OWL DL tem expressividade menor que OWL Full, mas conta com melhor eficiência, computacionalmente falando, pois garante que todas as conclusões sejam computáveis (implementadas em máquinas que contenham processador) e que todas as computações sejam resolvidas num tempo finito.

OWL Full contém a OWL DL (e, por conseguinte, a OWL Lite), é a única sub-linguagem que contém todos os elementos da RDFS, é bastante expressiva mas é semanticamente difícil de entender e de se trabalhar, é dificilmente suportada por alguma ferramenta de software e, segundo Santarem Segundo (2010, p. 129), foi desenvolvida para os usuários que desejam o máximo de expressividade, porém, sem garantia computacional.

#### **2.4.6 SPARQL Protocol and RDF Query Language (SPARQL)**

Como visto nas seções anteriores, RDF, RDFS e OWL são úteis para a descrição de recursos por meio de metadados e de ontologias. No entanto, uma vez que essas descrições foram construídas, linguagens de consulta são necessárias para que os dados contidos nelas possam ser utilizados. *SPARQL Protocol and RDF Query Language* (SPARQL) é um protocolo e



uma linguagem construída exatamente para resolver essa necessidade: consultar dados modelados em RDF.

Segundo o W3C (2013a, tradução nossa), a SPARQL é “um conjunto de especificações que fornecem linguagens e protocolos para consultar e manipular o conteúdo de grafos RDF disponíveis na Web ou em um banco de dados RDF”.

Mais especificamente,

A SPARQL pode ser utilizada para expressar consultas entre fontes de dados diversas, dados esses armazenados nativamente como RDF ou visualizados como RDF via *middleware*. A SPARQL tem a capacidade para realizar consultas por meio de padrões de grafos obrigatórios e opcionais juntamente com conjunções e disjunções. A SPARQL também suporta agregação, subconsulta, negação, criação de valores por expressões, teste de valor extensível e consultas de restrição tendo como fonte um grafo RDF. Os resultados das consultas SPARQL podem ser apresentados em conjuntos de resultados ou em grafos RDF (W3C, 2012e, tradução nossa).

A SPARQL teve a sua primeira especificação publicada pelo W3C em 15 de janeiro de 2008 (SPARQL 1.0) (W3C, 2008b) e uma atualização (SPARQL 1.1) publicada em 21 de março de 2013 (W3C, 2013b).

Segundo Breslin, Passant e Decker (2009), a SPARQL é compreendida como a SQL<sup>13</sup> da Web Semântica, ela oferece um meio poderoso para consultar grafos e triplas RDF.

Para Berners-Lee (JACOBS; FORGUE; HIRAKAWA, 2008) “tentar utilizar a Web Semântica sem SPARQL é como tentar utilizar um banco de dados relacional sem SQL. [...] A SPARQL torna possível consultar informações a partir de bancos de dados e de outras diversas fontes na Web”.

Devido aos dados em RDF estarem representados como grafos, a SPARQL é, então, uma linguagem de consultas baseada em grafos, o que significa que a abordagem é diferente da SQL, que lida com tabelas e linhas (tuplas). Dizendo de outra maneira, a SQL e a SPARQL operam em estruturas de dados bastante diferentes e, além dessas diferenças, a SPARQL fornece certa extensibilidade dentro dos padrões de consulta, possibilitando a realização de consultas avançadas baseadas em uma representação de grafos. Com a SPARQL é possível realizar, por exemplo, a seguinte consulta: encontre seis coisas que possuem como autor o autor Z, desde que o título e o número de páginas de cada coisa sejam conhecidos, além

---

<sup>13</sup> *Structured Query Language* (SQL) (Linguagem de Consulta Estruturada), linguagem padrão para banco de dados relacionais.

disso, ordene alfabeticamente os resultados pelo título (BRESLIN; PASSANT; DECKER, 2009; HITZLER; KRÖTZSCH; RUDOLPH, 2009). O exemplo da Figura 8 ilustra essa consulta.

Figura 8 – Exemplo de uma consulta utilizando a linguagem SPARQL.

---

```

PREFIX vocab: <http://padraodemetadados-y.org/uri#>
SELECT ?titulo ?numeroDePaginas
WHERE {
    ?livro vocab:possuiAutor <http://biblioteca-x.org/uri#autor-Z> .
    ?livro vocab:titulo ?titulo .
    ?livro vocab:paginas ?numeroDePaginas
}
ORDER BY ?titulo
LIMIT 6

```

---

Fonte: Elaborada pelo autor.

Assumindo que o autor Z represente o autor Pierre Lévy, a consulta SPARQL apresentada na Figura 8 poderia trazer o resultado apresentado no Quadro 4:

Quadro 4 – Exemplo de um resultado de uma consulta SPARQL.

titulo	numeroDePaginas
"A inteligência coletiva"	"212"
"A máquina universo"	"173"
"As tecnologias da inteligência"	"208"
"Cibercultura"	"264"
"Ciberdemocracia"	"249"
"O que é o virtual?"	"157"

Fonte: Elaborado pelo autor.

Faz-se necessário salientar que os resultados de consultas SPARQL, tal como apresentado no Quadro 4, não visam à apresentação imediata aos agentes humanos, pelo contrário, são construídos para a posterior formatação e apresentação ou processamento por outras aplicações, tal como ocorre com os resultados de consultas SQL.

Breslin, Passant e Decker (2009) afirmam que a linguagem SPARQL (1.0) oferece quatro formas de consulta que são utilizadas para fazer diferentes tipos de consultas:

- SELECT - utilizada para recuperar informações baseadas em um padrão de tripla(s);
- CONSTRUCT - utilizadas para criar um documento RDF baseado em um documento RDF de entrada e que pode ser usado como um serviço de tradução para os dados em RDF (entre ontologias diferentes);
- ASK - utilizado para identificar se um padrão particular de consulta pode ser correspondido no grafo RDF consultado;
- DESCRIBE - utilizado para identificar todas as triplas relacionadas a um objeto particular que precisa ser descrito.

Estudos aprofundados sobre a linguagem SPARQL não foram encontrados em Ciência da Informação, provavelmente por ser uma tecnologia criada recentemente e por ter uma forte ligação com a área da Ciência da Computação, mais especificamente, com as disciplinas sobre bancos de dados.

Algumas das tecnologias abordadas neste capítulo foram intensamente pesquisadas e apresentadas em estudos nacionais e internacionais da Ciência da Informação. Alguns desses estudos estão relacionados a seguir.

XML: Bryan (1998), Bax (2001), Almeida (2002b), Siqueira (2003), Miller e Clarke (2004), Santarem Segundo (2004), Alves (2005), Flamino (2006), Furgeri (2006), Ramalho (2006), Castro (2008), Eito Brun (2008), Gama (2011) e Assumpção (2013).

RDF: Miller (1998), Dziekaniak e Kirinus (2004), Santarem Segundo (2004), Alves (2005), Ramalho (2006), Castro (2008), Santarem Segundo (2010), Dunsire e Willer (2011), Gama (2011) e Catarino e Souza (2012).

RDFS: Santarem Segundo (2004), Alves (2005), Ramalho (2006), Castro (2008), Santarem Segundo (2010) e Gama (2011).

OWL: Ramalho (2006), Santarem Segundo (2010), Gama (2011) e Dunsire, Hillmann e Phipps (2012).

Este capítulo teve por objetivo fornecer um breve histórico das origens da Web Semântica, apresentar seus conceitos mais pertinentes a esta pesquisa e os padrões de codificação, representação, descrição, relação e consulta de dados estruturados (URI, XML, RDF, RDFS, OWL e SPARQL), realçando seus aspectos principais para a compreensão dos wikis semânticos, tema abordado no próximo capítulo.

### 3 Wikis semânticos: histórico, definições e tecnologias

Este capítulo apresenta sucintamente o que é um wiki e quais são as suas características comuns, e, de modo mais aprofundado, o que são wikis semânticos e quais as suas funções. Aborda também a origem dos wikis semânticos, seus principais conceitos e características, seus tipos e suas classificações e algumas de suas principais implementações.

#### 3.1 Os wikis e suas características

O termo wiki é a forma curta para o termo *WikiWikiWeb*, da mesma maneira que o termo Web é a forma curta para o termo *WorldWideWeb*. O termo wiki é derivado do termo havaiano “*wiki wiki*” que significa rápido. Os sistemas wiki surgiram em 1995 e, seguindo seu criador, Ward Cunningham, um wiki é

uma coleção livremente expansível de “páginas” Web interligadas, um sistema de hipertexto para armazenar e modificar informações – um banco de dados, no qual cada página é facilmente editável por qualquer usuário com um navegador que aceite formulários (LEUF; CUNNINGHAM, 2001, p. 14, tradução nossa).

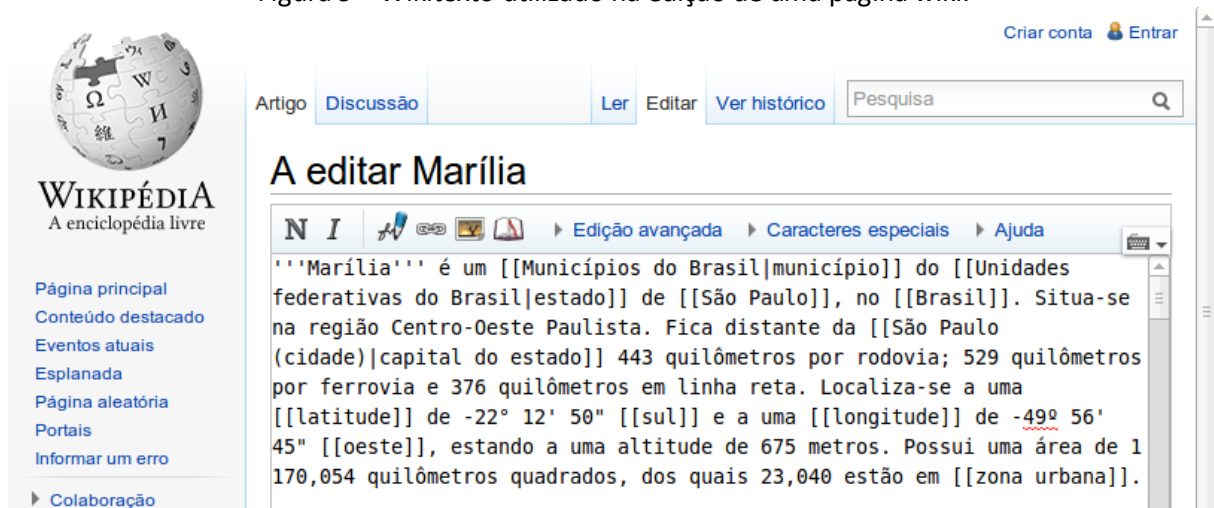
Essencialmente, um wiki é uma coleção de páginas Web que estão conectadas entre si por hiperlinks e que são editadas colaborativamente de modo rápido e fácil. Embora haja uma gama de sistemas wiki disponíveis (*MediaWiki*, *Pyki*, *Swiki*, etc.) com propósitos e públicos-alvo diferentes, a maior parte deles compartilha as características apresentadas a seguir por Schaffert (2006b):

**A edição de páginas é realizada via navegador:** o conteúdo de um wiki é editado com apenas o auxílio de um navegador comum, não sendo necessárias quaisquer instalações de programas adicionais. Isso torna a edição de páginas uma tarefa simples de ser realizada e permite que as modificações dessas páginas possam ser realizadas em qualquer parte do mundo e com o mínimo de requisitos técnicos. Como consequência, os criadores de conteúdo podem acessar e atualizar o site onde quer que estejam.

**Sintaxe wiki (ou linguagem de marcação wiki) simplificada:** o conteúdo das páginas de um wiki é normalmente expressado por meio de uma linguagem de marcação, chamada de sintaxe wiki, código wiki, linguagem wiki ou *wikitexto*. Por ser simples, tal linguagem é mais fácil de aprender e de usar que a linguagem de marcação HTML, por exemplo. A linguagem de marcação wiki é exibida aos usuários apenas durante o processo de edição do

conteúdo de uma página. Após a edição, o conteúdo é armazenado em um banco de dados. Para a exibição da página em um navegador, o sistema wiki realiza uma consulta ao banco de dados e converte o conteúdo expresso em linguagem wiki em conteúdo expresso em HTML. A Figura 9 exemplifica o *wikitexto* utilizado na edição de uma página wiki.

Figura 9 – *Wikitexto* utilizado na edição de uma página wiki.



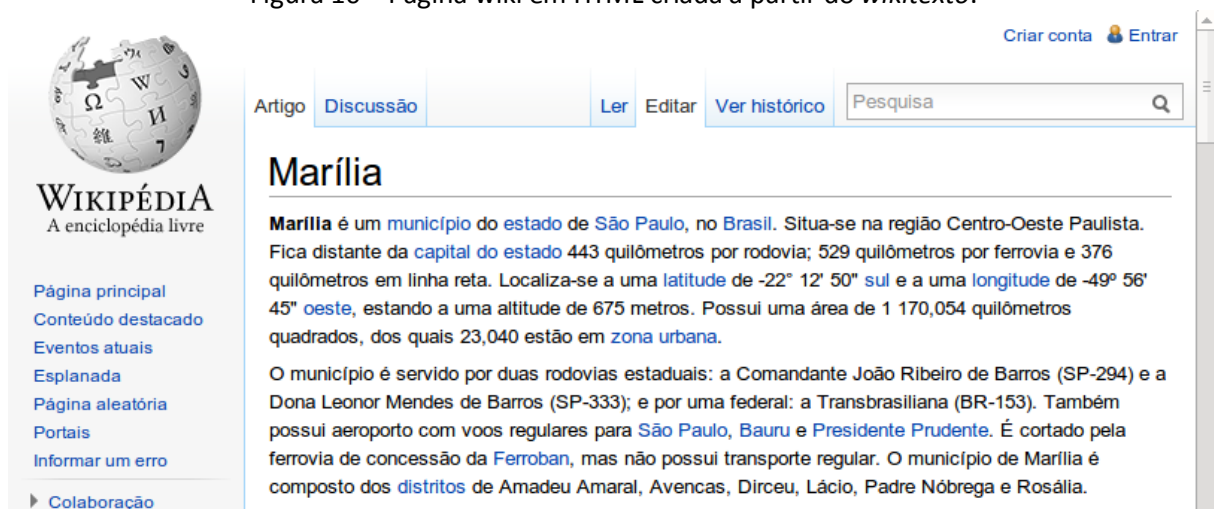
Fonte: Disponível em:

<<http://pt.wikipedia.org/w/index.php?title=Mar%C3%ADlia&oldid=35863131>>.

Acesso em: 24 jun. 2013.

A Figura 10, por sua vez, apresenta uma página wiki em HTML criada a partir do *wikitexto*.

Figura 10 – Página wiki em HTML criada a partir do *wikitexto*.



Fonte: Disponível em:

<<http://pt.wikipedia.org/w/index.php?title=Mar%C3%ADlia&oldid=35863131>>.

Acesso em: 24 jun. 2013.

**Mecanismo de reversão de edições:** as alterações no conteúdo de um wiki são armazenadas e “versionadas”, ou seja, as versões anteriores das páginas são mantidas em um histórico de edições. Isso permite fácil e rapidamente a reversão de uma página atual para uma versão anteriormente salva, caso alguma parte importante do conteúdo seja acidentalmente apagada ou indesejavelmente modificada por outro usuário. Além disso, muitos dos sistemas wiki permitem realizar a comparação de duas versões de uma página, tornando possível a identificação das mudanças ocorridas entre essas edições.

**Acesso aberto:** na maioria dos sistemas wiki o acesso é irrestrito, isto é, qualquer pessoa pode corrigir, modificar, completar, complementar ou mesmo apagar qualquer informação. Embora isso possa parecer estranho ou mesmo perigoso de uma perspectiva tradicional, a prática mostra que o sistema funciona: por um lado, usuários malcomportados são raros, por outro, todas as mudanças são facilmente desfeitas utilizando o mecanismo de reversão mencionado anteriormente. Embora essa liberdade de edição seja encontrada em muitos wikis, ela não é obrigatória. Existem opções para aplicar restrições de acesso a determinados usuários e/ou a grupos de usuários, funcionalidade comumente encontrada em sistemas de gerenciamento de conteúdo tradicionais.

**Edição colaborativa:** com as características anteriormente mencionadas, os wikis se tornam ferramentas bastante propícias para a edição colaborativa. Assim que uma pessoa cria um conteúdo, outras podem contribuir com mais informações para esse conteúdo, estendendo-o, corrigindo-o, etc. Muitos sistemas wiki fornecem suportes adicionais à edição colaborativa por meio de fóruns de discussão, resumos de edições e listas que exibem as últimas alterações realizadas no site, de modo geral ou em alguma página específica.

**Forte interconexão entre páginas:** as páginas de um wiki são bastante interconectadas entre si por meio de links. A razão para isso é que a sintaxe simplificada dos wikis torna fácil a tarefa de criação de links entre páginas. Links que apontam para páginas inexistentes normalmente são exibidos com uma cor diferente. Quando um usuário clica em um desses links, o sistema direciona-o para uma página na qual é possível criar a página inexistente. Os links em um wiki são a ferramenta de navegação mais importante e, além disso, vários desses sistemas permitem aos usuários não somente seguir os links na direção em que eles foram definidos, mas também na direção reversa (*back-links* ou páginas afluentes).

**Função de busca:** como ferramenta complementar de navegação, quase todos os sistemas wiki possuem um mecanismo de busca de texto completo que atua sobre o conteúdo de todas as páginas. Alguns wikis mais sofisticados permitem também a busca por campos, propriedades, atributos, categorias ou outros modos de seleção e filtragem de conteúdo.

**Carregamento de outros tipos de conteúdo:** além da criação de conteúdo textual, os wikis mais sofisticados permitem o carregamento (*uploading*) de conteúdos arbitrários como arquivos de texto, imagens, vídeos, sons, aplicativos, etc.

Diferentemente de outras ferramentas de gerenciamento de conteúdo, um sistema wiki fornece aos usuários, administradores ou não, a liberdade quase total sobre o processo de desenvolvimento e de organização do conteúdo, sem fluxos de trabalho rígidos, sem restrições de acesso ou estruturas predefinidas. Desse modo, os usuários não precisam adaptar as suas práticas à “ditadura do sistema”, pelo contrário, as práticas dos usuários é que podem definir a estrutura do wiki. Isso é importante porque diferentes domínios de interesse frequentemente possuem ou mesmo requerem diferentes tipos de fluxos de trabalho (SCHAFFERT, 2006b).

Assim, os sistemas wiki permitem uma fácil e rápida colaboração na criação e na edição de textos na Web. Eles mantêm um histórico das versões de suas páginas possibilitando a recuperação de versões anteriores a qualquer momento, permitem a visualização de mudanças recentes realizadas em todo o site por editores registrados ou não, dispõem de recursos que alertam os editores sobre as mudanças ocorridas em páginas de interesse, permitem a realização de buscas pelos títulos e pelos conteúdos das páginas e funcionam como ferramentas de desenvolvimento de conteúdos por comunidades de colaboradores.

De acordo com Reitz (2010), um wiki é uma aplicação Web que permite aos usuários adicionarem conteúdo a um recurso Web de hipertexto colaborativo, além de permitir que outros usuários editem aquele conteúdo. Wikis podem ter políticas de edição e procedimentos para o tratamento de conflitos de editores. As atividades dentro do wiki podem ser vistas e revisadas por qualquer visitante do site. Wiki também se refere ao software de servidor utilizado para criar coletivamente um site desse tipo, permitindo que páginas Web, armazenadas em um banco de dados, sejam facilmente criadas e atualizadas.

A fim de elencar as características de web sites tradicionais em comparação às características comuns presentes em sites wiki, é apresentado o Quadro 5.

Quadro 5 – Comparação entre sites web tradicionais e sites wiki.

Sites web tradicionais	Sites wiki
Objetiva principalmente prover informações	Além de prover informação, incentiva a prática de colaboração, exploração
A edição frequentemente requer conhecimento da linguagem HTML	A edição é simples utilizando uma linguagem de marcação simples ou um sistema WYSIWYG <sup>14</sup>
As informações são dadas, majoritariamente, em texto contínuo	As informações são dadas, majoritariamente, em texto contínuo
Uma página é criada como um novo arquivo	Uma página pode ser criada a partir de um link para uma página inexistente no wiki
Páginas se tornam públicas em um momento decidido pelo administrador	Páginas são atualizadas em tempo real a menos que uma função de aprovação de conteúdo esteja ativada
Cada página tem poucos autores	Cada página pode ter muitos autores
Geralmente os autores são conhecidos	Os autores podem ser “anônimos” <sup>15</sup>
A edição normalmente está limitada a pessoas autorizadas	As edições podem ser abertas a todos os leitores sem necessidade de se registrarem
Sistemas de controle de versão precisam ser desenvolvidos ou implementados para manter versões antigas	As versões anteriores das páginas são normalmente guardadas automaticamente
Pode não ser possível restaurar uma versão anterior da página	É possível restaurar versões anteriores da página
A reversão requer algum esforço	A reversão é rápida e simples, requer apenas um ou dois cliques
Geralmente não há a possibilidade de observar, de modo automático, as mudanças recentes ocorridas no site	As mudanças recentes das páginas são rastreadas facilmente
A pesquisa pode ser um pouco sofisticada	A maioria dos wikis oferecem pesquisas rudimentares dentro do wiki
A estrutura de permissões é hierárquica	A estrutura de permissões normalmente é simples
Os processos de aprovação de conteúdo são comuns	Os processos de aprovação de conteúdo são raramente implantados
O conteúdo é definido pelo autor ou pelo patrocinador da página	O conteúdo é definido pela comunidade de autores
A estrutura do site é predefinida	A estrutura do wiki se desenvolve conforme as páginas são criadas pelos usuários
Podem suportar uma comunidade existente, mas não são uma ferramenta de desenvolvimento de comunidades	Podem suportar uma comunidade existente e também desenvolver uma comunidade de leitores e contribuidores
As páginas normalmente são consideradas finalizadas	As páginas normalmente são consideradas um processo sempre em desenvolvimento

<sup>14</sup> WYSIWYG: *What You See Is What You Get* (O que você vê é o que você obtém).

<sup>15</sup> O “anonimato” aqui é encarado de forma branda, já que endereços IP podem identificar pessoas.



Sites web tradicionais	Sites wiki
Pode ou não ocorrer a exportação de dados em diversos formatos	Como as informações são estruturadas em forma de texto, são disponibilizados poucos formatos de exportação (por exemplo PDF e XML)
A representação e a descrição das informações são diversificadas, depende do site	A representação e a descrição das informações ocorre, mas de forma pouco estruturada

Fonte: Adaptado de Klobas (2006, p. 215-216).

Os wikis, de modo geral, são ambientes abertos, descentralizados, dinâmicos, flexíveis, de fácil utilização, com foco na colaboração entre pessoas e de baixo custo (a maioria dos sistemas wiki é gratuita e de código aberto).

Além dos pontos fortes e positivos possibilitados pelos wikis, apresentados por este estudo e por Ferreira (2011), os wikis também possuem pontos fracos.

Os pontos fracos dos wikis, que são relevantes para esta pesquisa, são oriundos da natureza aberta e flexível desses sistemas. Um site wiki geralmente cresce muito rápido e, com a quantidade sempre crescente de usuários colaboradores, a criação de páginas é realizada, muitas vezes, de forma exponencial. Com isso, o aumento do banco de dados do wiki pode trazer consequências negativas, tais como dificuldades na navegação, na orientação e na busca por conteúdos (GONÇALVES, 2008).

As características do wiki [...] estimulam a colaboração e, muitas vezes, promovem o crescimento da base de dados de um ambiente e [desse modo] alguns problemas são enfrentados. Quando a base de dados de um ambiente wiki fica muito grande, pode ser necessária a presença de recursos para a organização dos conteúdos, sem, contudo, comprometer as características básicas do wiki (GONÇALVES, 2008, p. 64).

De acordo com Gonçalves (2008), recursos como a categorização de páginas, as predefinições (*templates*) e as notificações ocorridas quando novos conteúdos são disponibilizados e/ou alterados podem ser úteis para a organização do conteúdo e também para que administradores do site possam revisar esses conteúdos e organizar as páginas no site.

Uma vez que o conteúdo de um wiki tradicional é apenas sintático, não é possível a interpretação das informações pelas máquinas, o que seria de grande relevância para a organização das informações ali contidas, para a obtenção automática de informações e para a realização de buscas mais complexas no banco de dados (GONÇALVES, 2008).

Segundo Schaffert (2006a) e Krötzsch et al. (2007), os sistemas wiki são ferramentas que estão se tornando cada vez mais populares na gestão de conteúdos e na gestão do conhecimento em ambientes digitais. Atualmente, uma quantidade considerável de conhecimento está disponível em sistemas como a Wikipédia, mas, infelizmente, esse conhecimento não está acessível às máquinas.

Para solucionar esses problemas foram propostos sistemas wiki melhorados com tecnologias sofisticadas para o tratamento de informações. Essas tecnologias, conhecidas como tecnologias da Web Semântica, permitiram, por exemplo, a incorporação de metadados nos wikis. Combinando as tecnologias dos wikis (sintáticos) com as tecnologias da Web Semântica surgiram os wikis semânticos, tema abordado nas seções seguintes.

### 3.2 Definições de wiki semântico

O termo wiki semântico foi usado pela primeira vez em uma postagem na rede Usenet no dia 21 de janeiro de 2001 por Andy Dingley (DINGLEY, 2001). Porém, somente em 2003 o termo voltou a aparecer, dessa vez, em uma dissertação de mestrado realizada pelo cientista da informação austríaco Leopold Sauer mann (SAUERMAN, 2003). Uma das propostas presentes na dissertação de Sauer mann incluía a adaptação de um sistema wiki para suportar a utilização de tecnologias da Web Semântica (URI, XML, RDF, RDFS, OWL) pelos usuários desse wiki.

Nos anos seguintes, propostas e sistemas de wikis semânticos começaram a aparecer. A maioria deles surgiu por volta de 2004, 2005 e 2006.

De 2006 em diante foram realizados vários eventos internacionais sobre o tema, sendo que os mais importantes foram:

- *First Workshop on Semantic Wikis: From Wiki to Semantics* (2006), evento que ocorreu conjuntamente com a *3rd Annual European Semantic Web Conference* em Budva, Montenegro;
- *Second Workshop on Semantic Wikis: Wiki-based Knowledge-Engineering* (2006), evento que ocorreu conjuntamente com o *International Symposium on Wikis and Open Collaboration* (WikiSym) em Odense, Dinamarca;
- *Third Workshop on Semantic Wikis: The Wiki Way of Semantics* (2008), evento que ocorreu conjuntamente com a *5th Annual European Semantic Web Conference* em Tenerife, Espanha;

- *Fourth Workshop on Semantic Wikis: The Semantic Wiki Web* (2009), evento que ocorreu conjuntamente com a *6th Annual European Semantic Web Conference* em Creta, Grécia; e
- *Fifth Workshop on Semantic Wikis: Linking Data and People* (2010), evento que ocorreu conjuntamente com a *7th Annual Extended Semantic Web Conference* em Heraklion, Grécia.

A partir de 2005 o assunto wikis semânticos evoluiu de um tema obscuro para um de interesse de uma ampla comunidade (MILLARD et al., 2008; VÖLKELE; SCHAFFERT; DECKER, 2006). Segundo Völkel, Schaffert e Decker (2006), o campo de pesquisa sobre wikis semânticos é bastante diversificado: há estudos sobre sistemas de wikis semânticos (*engines*), sobre modos de combinar os wikis com as ideias da Web Semântica e sobre as aplicações de wikis semânticos, por exemplo, nas áreas de biociência, matemática, *e-learning* e multimídia.

Völkel, Schaffert e Decker (2006) apontam que os wikis semânticos são explorados principalmente de duas formas: wikis que são melhorados com tecnologias da Web Semântica e aplicações da Web Semântica que estão sendo “wikificadas”.

Para Vrandečić e Krötzsch (2006) inúmeras abordagens têm sido desenvolvidas com o objetivo de combinar os wikis com as tecnologias semânticas. Enquanto os objetivos e métodos desses sistemas wiki variam bastante, muitos deles focam na capacidade de possibilitar aos usuários especificar propriedades e relacionamentos entre elementos individuais.

Um wiki semântico pode ser definido como

um software wiki que se utiliza de tecnologias da Web Semântica para incorporar conhecimento formalizado, conteúdo, estruturas e links nas páginas wiki. O conhecimento formalizado é representado utilizando padrões da Web Semântica e é, portanto, acessível e reutilizável por aplicações Web (BUFFA et al., 2007, p. 85, tradução nossa).

Para Schaffert et al. (2006, p. 112), um wiki semântico estende o wiki tradicional (ou sintático) pelas suas tecnologias semânticas como a RDFS e a OWL, por exemplo. Segundo os autores, o principal objetivo de um wiki semântico é – dado a estrutura inerente de vinculação entre páginas nos wikis – possibilitar que informações estruturadas possam ser “entendidas” por máquinas (agentes, serviços), possibilitando ir além da mera navegação por links.

Berners-Lee (2006, tradução nossa), em uma entrevista a IBM, considera que “Os wikis semânticos são muito interessantes. Estes são wikis em que as pessoas podem adicionar dados e, em seguida, esses dados podem ser navegados e divididos utilizando várias ferramentas da Web Semântica [...]”.

Devido aos wikis semânticos estenderem os wikis tradicionais com tecnologias da Web Semântica, as informações contidas no wiki podem ser anotadas (ou marcadas) semanticamente com metadados para facilitar a realização de buscas automatizadas, de consultas complexas, possibilitar o intercâmbio de dados e a realização de inferências sobre esses dados (BAO; DING; HENDLER, 2008).

Buffa et al. (2007, p. 85, tradução nossa) apontam que em um wiki semântico o conhecimento formalizado é utilizado para adicionar funcionalidades avançadas, tais como “uma melhor busca por documentos, sugestões de links novos, identificação de redes de conhecimento, atualização de conteúdo dinâmico, checagem e notificação de informações, etc.”.

Para Boulos (2009), as páginas de um wiki tradicional possuem textos estruturados e formatados (destinados à leitura e ao entendimento por humanos) e links tradicionais para outras páginas relacionadas no wiki (links também destinados a humanos). Já os wikis semânticos oferecem a habilidade de capturar (por humanos), armazenar e posteriormente identificar (por máquinas) mais meta-informações ou metadados sobre essas páginas e links, bem como as suas relações.

Para Zaidan e Bax (2011b, p. 1379),

Um wiki semântico é um wiki que tem subjacente um modelo de conhecimento descrito em suas páginas. Wikis clássicos, ou sintáticos, constituem-se de textos e hiperlinks não tipados. Wikis semânticos, por outro lado, permitem seus usuários identificar informações sobre os dados descritos nas páginas, e as relações entre as páginas, de maneira que podem ser consultados ou exportados como um banco de dados. [...] Existe uma grande variedade de cenários de aplicação de wikis semânticos. Para citar apenas alguns: a engenharia de ontologias, a gestão do conhecimento, e os ambientes educacionais.

Segundo os autores (ZAIDAN; BAX, 2011b, p. 1385), os wikis semânticos permitem a marcação semântica em links, a apresentação dinâmica dos conteúdos, uma navegação mais rica, a atribuição de metadados, a exportação dos dados em triplas (RDF), uma busca semântica e a inclusão de consultas em páginas do wiki (consultas embutidas).

Kiesel (2006, p. 31, tradução nossa) afirma que os wikis semânticos tentam explicitar as informações contidas no wiki utilizando tecnologias da Web Semântica para facilitar o reuso da informação ou, de modo geral, facilitar o acesso à informação por meios automatizados. Ao mesmo tempo, para o autor, “o conhecimento da estrutura interna da informação contida nas páginas wiki pode ser utilizado para melhorar a navegação e a busca dentro do wiki”. Kiesel (2006, p. 31) afirma também que uma tarefa bastante comum nos wikis semânticos é a de *anotar* estrutura no wiki, isto é, adicionar metadados aos recursos já existentes como links e páginas.

Nessa mesma linha de raciocínio, Oren et al. (2006a) e Oren, Breslin e Decker (2006) afirmam que os wikis semânticos permitem aos usuários realizar descrições formais de recursos por meio de marcações semânticas (também chamadas de anotações semânticas ou simplesmente de anotações) nas páginas que representam aqueles recursos. Enquanto um wiki tradicional possibilita aos usuários descrever recursos em linguagem natural, um wiki semântico possibilita, adicionalmente, descrever recursos em uma linguagem formal. Pela adição de metadados ao conteúdo tradicional, os usuários adquirem benefícios tais como melhor acesso e recuperação, intercâmbio de informações e reuso do conhecimento.

Para Di Iorio et al. (2006, p. 212) os wikis semânticos combinam o sucesso do modelo wiki com a poder das tecnologias da Web Semântica. Segundo esses autores, um wiki semântico é um wiki melhorado com o objetivo de codificar mais conhecimento do que apenas texto estruturado e hiperlinks e tornar esse conhecimento também legível por máquinas.

Segundo Al-Husain e El-Masri (2012), os avanços recentes nos wikis semânticos enriquecem os wikis tradicionais pela adição de marcações semânticas estruturadas ao texto tradicional dos wikis. Com essa prática, o conteúdo do wiki se torna mais organizado e estruturado e, desse modo, suporta uma melhor busca e uma navegação facilitada. Segundo os autores, esses wikis adicionam recursos semânticos como esquemas de dados, categorização e marcações semânticas. Os autores ressaltam que, apesar de esses novos recursos melhorarem a organização do conteúdo, eles não afetam negativamente as características de abertura e flexibilidade dos wikis.

Buffa et al. (2006, p. 84), analisando o estado-da-arte dos projetos de wiki semântico, os distinguem com relação à abordagens que consideram o uso de wikis para ontologias e abordagens que consideram o uso de ontologias para wikis, enquanto que poucos projetos

consideram as duas abordagens ao mesmo tempo. A maioria dos projetos se enquadra na primeira categoria, isto é, projetos que consideram as páginas wiki como conceitos e os links no conteúdo das páginas como relações ou atributos.

Segundo Schaffert et al. (2006, p. 112), alguns sistemas de wiki semântico ainda estão focados principalmente no conteúdo das páginas e encaram as marcações semânticas como algo opcional. Para outros sistemas, as marcações semânticas estão em primeiro plano. De acordo com os autores, sistemas diferentes servem a diferentes propósitos, por exemplo, estendendo o conteúdo com marcações semânticas para permitir uma melhor navegação ou permitindo a construção colaborativa de ontologias, etc. Schaffert et al. (2006) declaram que as características comumente encontradas em sistemas de wiki semântico são:

- **Classificação ou marcação semântica em links (*typed links*)**. Praticamente todos os wikis semânticos permitem a marcação semântica nos links, desse modo, dá-se aos links um tipo específico de informação. A ideia subjacente a isso é que as ligações entre as páginas não existem apenas por mera navegação. Geralmente, cada link carrega algum relacionamento ou significado entre as páginas envolvidas. Os modos pelos quais os links são marcados semanticamente diferem de sistema para sistema. Alguns wikis semânticos incluem a marcação semântica como parte da sintaxe wiki, enquanto outros fornecem um editor separado para a adição dessas marcações.
- **Apresentação consistente ao contexto**. Alguns wikis semânticos podem modificar o modo como o conteúdo é apresentado baseando-se em marcações semânticas. Isso pode incluir o enriquecimento das páginas pela exibição de páginas semanticamente relacionadas, pela exibição de informações que são derivadas a partir de uma base de conhecimento subjacente, ou mesmo pela apresentação do conteúdo de uma página de uma maneira diferente que é mais adequada para um dado contexto.
- **Navegação semântica**. Enquanto um wiki tradicional somente permite ao usuário seguir um link, um wiki semântico fornece informações adicionais sobre a relação que o link descreve. Tal informação pode ser utilizada para oferecer uma navegação adicional ou mais sofisticada. Por exemplo, os links

são mais independentes do contexto textual e podem ser exibidos, por exemplo, em uma caixa separada para informações relacionadas.

- **Busca semântica.** A maioria dos wikis permite uma busca semântica na base de conhecimento subjacente. Em alguns wikis as consultas são expressas na linguagem SPARQL. Utilizando a busca semântica os usuários podem realizar consultas como “recupere todas as músicas compostas por Mozart” ou “recupere todos os documentos em que a licença permita obras derivadas”.
- **Suporte ao raciocínio/inferência.** Raciocínio ou inferência, na Web Semântica, significa a derivação adicional e implícita de informações a partir de fatos cadastrados no sistema utilizando regras predefinidas na base de conhecimento. Por exemplo, a partir do fato de que Mozart compôs A flauta mágica, um sistema capaz de raciocínio lógico poderia deduzir (ou inferir) que Mozart é um compositor. Embora o raciocínio lógico de fatos seja um recurso importante, ele é suportado por poucos sistemas. Como dito anteriormente, construções complexas de linguagens para a representação de conhecimentos implícitos produzem altas complexidades computacionais, o que leva os sistemas a ficarem mais lentos. Além disso, esse tipo de funcionalidade requer uma memória intensiva e pode produzir resultados que não são esperados ou mesmo rastreáveis para os usuários.

Segundo Oren et al. (2006b) um wiki semântico é composto por, pelo menos, uma interface de usuário, um analisador sintático (*parser*), um servidor de páginas, um analisador de dados e um banco de dados.

Al-Husain e El-Masri (2012) comentam que praticamente todos os sistemas wiki clássicos oferecem suporte a marcação simples nas páginas na forma de *tagging* ou de categorização do conteúdo. Os wikis semânticos estendem essa capacidade pela adição de marcações semânticas sobre as páginas wiki ou sobre as relações (links) entre essas páginas. Essas marcações semânticas são representadas em RDFS e OWL e, uma vez que essas marcações foram criadas, elas estão, então, disponíveis para consultas complexas.

Os wikis semânticos utilizam quase o mesmo formato de marcação que os wikis tradicionais, eles só modificam as estruturas de links nas páginas para fornecer dados semânticos. Nos wikis semânticos os links são marcados com propriedades (atributos) e valores. Há dois tipos de marcações semânticas: marcações de links que expressam

relacionamentos entre recursos e marcações contendo propriedades e valores de dados (metadados com valores literais).

Deixando de lado os aspectos relacionados às tecnologias da Web Semântica, Wallace (2012) acredita que para que um wiki seja considerado um wiki semântico, ele precisa fornecer, no mínimo:

1. a habilidade para classificar as páginas: em uma página wiki chamada Cibercultura, por exemplo, ter a possibilidade de declarar que a página representa um livro (ontologicamente, que o recurso descrito é um livro);
2. a habilidade para atribuir significado aos links entre as páginas: ter a possibilidade de declarar que a página Cibercultura tem uma relação (link) com a página Pierre Lévy, sendo que essa relação expressa a autoria do primeiro (Cibercultura tem como autor Pierre Lévy);
3. a habilidade para atribuir significado aos valores de dados (literais) em uma página: ter a possibilidade de declarar, por exemplo, que o recurso Cibercultura possui o ISBN número 85-7326-126-9; e
4. a habilidade para consultar esses dados possibilitando a geração de conteúdo dinâmico.

A fim de elencar as características dos wikis tradicionais em comparação às características comuns presentes em wikis semânticos, é apresentado o Quadro 6, com a primeira coluna extraída do Quadro 5.

Quadro 6 – Comparação entre características de wikis tradicionais e de wikis semânticos.

<b>Wikis tradicionais</b>	<b>Wikis semânticos</b>
Além de prover informação, incentiva a prática de colaboração, exploração	Além de prover informação legível por humanos, almeja prover informação estruturada para máquinas
A edição é simples utilizando uma linguagem de marcação simples ou um sistema WYSIWYG <sup>16</sup>	A edição normalmente é baseada em formulários para a estruturação dos dados
As informações são dadas, majoritariamente, em texto contínuo	As informações geralmente são apresentadas de modo mais estruturado (tabelas, quadros, gráficos, listas, etc.)
Uma página pode ser criada a partir de um link para uma página inexistente no wiki	Uma nova página pode ser criada a partir de um link para uma página inexistente no wiki, no entanto, essa nova página é criada com o auxílio de um formulário

<sup>16</sup> WYSIWYG: *What You See Is What You Get* (O que você vê é o que você obtém).



<b>Wikis tradicionais</b>	<b>Wikis semânticos</b>
Páginas são atualizadas em tempo real a menos que uma função de aprovação de conteúdo esteja ativada	Páginas são atualizadas em tempo real a menos que uma função de aprovação de conteúdo esteja ativada
Cada página pode ter muitos autores	A quantidade de autores depende de quão especializado é o assunto e de quais restrições à edição foram impostas
Os autores podem ser “anônimos” <sup>17</sup>	Os autores podem ser “anônimos” <sup>17</sup>
As edições podem ser abertas a todos os leitores sem necessidade de se registrarem	Normalmente os wikis semânticos possuem mais restrições à edição que os wikis tradicionais
As versões anteriores das páginas são normalmente guardadas automaticamente	As versões anteriores das páginas são normalmente guardadas automaticamente
É possível restaurar versões anteriores da página	É possível restaurar versões anteriores da página
A reversão é rápida e simples, requer apenas um ou dois cliques	A reversão é rápida e simples, requer apenas um ou dois cliques
As mudanças recentes das páginas são rastreadas facilmente	As mudanças recentes das páginas são rastreadas facilmente
A maioria dos wikis oferecem pesquisas rudimentares dentro do wiki	A pesquisa em um wiki semântico pode ser sofisticada (por texto completo, com operadores booleanos, por busca facetada, por propriedade ou em classes específicas, etc.)
A estrutura de permissões normalmente é simples	A estrutura de permissões é mais complexa que em wikis tradicionais
Os processos de aprovação de conteúdo são raramente implantados	Os processos de aprovação de conteúdo são implantados com mais frequência que em wikis tradicionais
O conteúdo é definido pela comunidade de autores	O conteúdo é definido pela comunidade de autores
A estrutura do wiki se desenvolve conforme as páginas são criadas pelos usuários	A estrutura do site é predefinida (ocorre a criação de formulários, predefinições, propriedades, classes, etc. anteriormente a aceitação de dados)
Podem suportar uma comunidade existente e também desenvolver uma comunidade de leitores e contribuidores	Podem suportar uma comunidade existente e também desenvolver uma comunidade de leitores e contribuidores
As páginas normalmente são consideradas um processo sempre em desenvolvimento	As páginas normalmente são consideradas um processo sempre em desenvolvimento
Como as informações são estruturadas em forma de texto, são disponibilizados poucos formatos de exportação (por exemplo PDF e XML)	Como as informações são mais estruturadas que em wikis tradicionais, uma gama de formatos de exportação são disponibilizados (XML, RDF, OWL, CSV, JSON, etc.)
A representação e a descrição das informações ocorre, mas de forma pouco estruturada	A representação e a descrição das informações ocorre de forma estruturada
<b>Exemplos:</b> Wikipedia (de.wikipedia.org) Wikitravel (wikitravel.org)	<b>Exemplos:</b> NYCpedia (nyc.pediacities.com) NYCFacets (nycfacets.com)

<sup>17</sup> O “anonimato” aqui é encarado de forma branda, já que endereços IP podem identificar pessoas.

Wikis tradicionais	Wikis semânticos
wikiHow (wikihow.com)	SNPedia (snpedia.com)
Recipes Wiki (recipes.wikia.com)	LawMantic (lawmantic.com)
Wikimapia (wikimapia.org)	Familypedia (familypedia.wikia.com)
AboutUs (aboutus.org)	The Free Software Directory (directory.fsf.org)
LyricWikia (lyrics.wikia.com)	saveMLAK (savemlak.jp)
Wikibooks (pt.wikibooks.org)	Wikiparques (wikiparques.com)

Fonte: Elaborado pelo autor, com base em Klobas (2006, p. 215-216).

Tendo definido e caracterizado o que são os wikis semânticos, convém classificá-los de alguma maneira. Um modo de classificar os wikis semânticos é apresentado por Bry (2012). Segundo o autor, os wikis semânticos podem ser classificados em pertencentes a uma primeira geração ou a uma segunda geração.

Os wikis semânticos da **primeira geração** focam na marcação semântica aplicada aos links para expressar relacionamentos específicos entre recursos. Nessa geração, os wikis semânticos expressam o conteúdo semântico dos recursos por meio da descrição em triplas (sujeito-predicado-objeto). Os wikis dessa geração possuem uma simples interface de usuário tipicamente com três painéis: um painel para a edição do texto da página, um painel para a visualização e/ou edição dos metadados da página e um painel para a navegação, utilizado para a listagem de páginas relacionadas de acordo com os metadados atribuídos. Os wikis dessa geração são subdivididos em dois tipos.

O primeiro tipo compreende os wikis em que os links com marcações semânticas encontram-se em meio ao texto. Nesses wikis os metadados dos recursos descritos na página wiki são especificados como marcações semânticas atribuídas aos links em meio ao texto. Além disso, os metadados não podem ser editados por outras vias senão pela própria interface de edição do texto e o painel de metadados só permite a visualização dos mesmos.

Nos wikis do segundo tipo os links com marcações semânticas encontram-se separados do texto. Nesses wikis o texto e os metadados são definidos separadamente e o painel de metadados é utilizado tanto para a leitura quanto para a edição dos metadados.

Os wikis semânticos da **segunda geração**, por sua vez, focam na utilização de ontologias. Esses wikis expressam o conteúdo semântico de forma mais rica, isto é, na forma

de ontologias. Além disso, pequenas inferências podem se realizadas com as ontologias armazenadas por esses wikis. Uma das motivações para o surgimento dessa segunda geração de wikis semânticos foi a Wikipédia. As *infoboxes*<sup>18</sup> da Wikipédia podem ser vistas como expressões textuais de conceitos semanticamente ricos e relações entre conceitos. Uma vez que as ontologias ricas semanticamente dificilmente são expressas em linguagem natural, a maioria dos wikis semânticos da segunda geração possuem painéis para a edição de metadados. Segundo Bry (2012), o desenvolvimento dos wikis semânticos da segunda geração tem sido impulsionado pela comunidade de pesquisa em Web Semântica que os consideram como “wikis para a Web Semântica” ou “a Web Semântica em miniatura”. Assim como os wikis da primeira geração, os wikis da segunda geração também estão divididos em dois tipos.

O primeiro tipo compreende os wikis com a capacidade para o desenvolvimento de ontologias. Esses wikis foram desenvolvidos com o objetivo de serem a “Web semântica em miniatura”. Os sistemas se esforçam para serem capazes de importar ontologias complexas, utilizar essas ontologias para estruturar semanticamente o conteúdo dos wikis e suportar o desenvolvimento dessas ontologias. Com essa configuração, os wikis semânticos podem ser considerados como “interfaces de ontologias legíveis por humanos” ou sistemas de gerenciamento de ontologias mais ou menos sofisticados. Esses wikis semânticos se aproximam dos sistemas editores de ontologias.

Já o segundo tipo compreende os wikis sem a capacidade para o desenvolvimento de ontologias. Esses wikis foram desenvolvidos com a ideia de que não deveriam fornecer ferramentas para o desenvolvimento ou o gerenciamento de ontologias complexas. Para Bry (2012), os desenvolvedores desses wikis acreditam que ontologias complexas são melhor especificadas fora de um wiki semântico.

Bry (2012) destaca também os chamados wikis semânticos não convencionais. Tais wikis não se encaixam em nenhuma das duas gerações.

Para sumarizar os sistemas de wikis semânticos, classificados de acordo com a proposta de Bry (2012), apresenta-se o Quadro 7, que contém, além das divisões por geração e suas subdivisões, o nome do sistema de wiki semântico e o ano de sua divulgação.

---

<sup>18</sup> Uma *Infobox* é uma tabela de formato fixo planejada para ser adicionada ao lado direito das páginas da Wikipédia para apresentar de modo consistente o resumo de alguns aspectos que um determinado grupo de páginas compartilha e, às vezes, para melhorar a navegação para outras páginas relacionadas.

Quadro 7 – Sistemas de wikis semânticos classificados de acordo com seus focos.

<b>Primeira geração:</b> foco na marcação semântica aplicada aos links	
<b>Primeiro tipo:</b> links com marcações semânticas em meio ao texto	<b>Segundo tipo:</b> links com marcações semânticas fora do texto
Kawawiki (2005) Makna e MultiMakna (2006) MewKISS (2005) Platypus (2004) Rhizome (2005) TaOPis (2008) WikSAR (2006)	Kawawiki (2006) Platypus (2004) Rhizome (2005) SemperWiki (2005) SHAWN (2005)
<b>Segunda geração:</b> foco na utilização de ontologias	
<b>Primeiro tipo:</b> wikis com a capacidade para o desenvolvimento de ontologias	<b>Segundo tipo:</b> wikis sem a capacidade para o desenvolvimento de ontologias
Freebase (2005) <sup>19</sup> Graphingwiki (2006) Gnowsis Wiki (2006) Makna e MultiMakna (2006) Semantic MediaWiki (2005) OntoWiki (2006) Swooki (2008) Topincs Wiki (2006) Ylvi (2006)	BOWiki (2008) COW (2006) IkeWiki (2006) Kaukolu (2006)
<b>Wiki semânticos não convencionais</b>	
AceWiki (2008) BOWiki (2008) Freebase (2005) <sup>19</sup> Graphingwiki (2006) KnowWE (2010) Makna (2006) OntoWiki (2006) OpenRecord (2006)	ShyWiki (2011) SweetWiki (2006) SWiki (2005) Swooki (2008) TaOPis (2008) Topincs Wiki (2006) wikiCalc (2007)

Fonte: Bry (2012).

Apesar de dividir os wikis semânticos em primeira e segunda gerações com seus respectivos tipos, Bry (2012) reconhece que determinados wikis semânticos podem

<sup>19</sup> Bry (2012) atenta para o fato de que o Freebase é um caso especial: ele é uma aplicação com características wiki, não um sistema wiki propriamente dito.

pertencer a mais de um tipo ou mesmo a mais de uma geração devido as características diversas apresentadas por eles.

Durante a revisão de literatura, a categorização dos wikis semânticos realizada por Bry (2012) foi a única encontrada, não sendo objetivo da presente pesquisa desenvolver uma categorização para esses sistemas.

Para finalizar, concorda-se com Millard et al. (2008) ao concluírem que os wikis semânticos oferecem um mecanismo poderoso para a criação e a publicação de dados semânticos de um modo simples. Para os autores, assim como o wiki tradicional possibilitou a autoria hipertextual na Web para os usuários não especialistas, os wikis semânticos poderiam tornar a Web Semântica acessível também a esses usuários não especialistas. Pela utilização de ontologias que se expandem com os wikis e de uma linguagem de marcação significativa, os usuários podem construir grafos semânticos e desenvolver novas ontologias sem sequer construir uma declaração RDF ou um axioma na linguagem OWL.

Para um maior entendimento sobre alguns dos wikis semânticos listados anteriormente, são apresentados, na próxima seção, os principais wikis semânticos e suas características de destaque.

### 3.3 Principais iniciativas e tecnologias

Assim como os sistemas wiki tradicionais, não existe um sistema de wiki semântico padrão. Cada sistema possui seu próprio foco, seus pontos fortes e suas fraquezas.

A lista de sistemas de wiki semântico apresentada no Quadro 7 da seção 3.2 não almeja englobar todos os sistemas existentes. Salienta-se, também, que o presente estudo não pretende comparar tais sistemas. Estudos que realizam tais comparações podem ser obtidos em Bry (2012), Buffa et al. (2007), Millard et al. (2008), Panagiotou e Mentzas (2007) e Semantic Wiki Projects (2013). As breves descrições sobre os wikis semânticos listados a seguir pretendem fornecer apenas uma visão geral de cada sistema. A escolha dos sistemas descritos a seguir foi baseada na popularidade que esses sistemas tiveram nas etapas de levantamento bibliográfico e de análise da literatura realizada durante a elaboração desta pesquisa.

- **AceWiki** (KUHN, 2008). Um wiki semântico que permite a representação de conhecimento por meio de sentenças escritas em um inglês controlado

(*Attempto Controlled English* - ACE<sup>20</sup>). Textos em inglês, escritos nesses termos controlados, podem ser traduzidos em uma variante sintática da lógica de primeira ordem e em OWL DL. No *AceWiki*, entidades ontológicas como indivíduos, conceitos e propriedades são mapeadas para entidades linguísticas como nomes próprios, substantivos, preposições e verbos.

- **IkeWiki** (SCHAFFERT; GRUBER; WESTENTHALER, 2005). O *IkeWiki* e sua derivação, *KiWi* (SCHAFFERT et. el, 2009), utilizam o RDF e a OWL para o armazenamento e o raciocínio lógico com o conhecimento formal. Segundo os autores, o *IkeWiki* é fácil de usar, possui uma interface interativa, é compatível com o software *MediaWiki* e com os padrões da Web Semântica, utiliza as anotações semânticas de modo imediato, suporta diferentes níveis de experiência e diferentes níveis de formalização e deduções na base de conhecimento.
- **OntoWiki** (AUER; DIETZOLD; RIECHERT, 2006). O *OntoWiki* é um wiki semântico baseado em formulários. Assim como o *IkeWiki*, seu foco está na construção colaborativa de ontologias e bases de conhecimento. Ele também utiliza a RDF como estrutura de modelagem e a SPARQL como linguagem de consulta. Seu desenvolvimento foi financiado em grande parte pelo Ministério Federal de Educação e Pesquisa da Alemanha.
- **Platypus Wiki** (TAZZOLI; CASTAGNA; CAMPANINI, 2004). Desenvolvido inicialmente em 23 de dezembro de 2003, o *Platypus Wiki* foi um dos primeiros wikis semânticos a aparecerem. Segundo os autores, o *Platypus Wiki* é um wiki melhorado que utiliza “o modelo de dados RDF e o vocabulário OWL para representar metadados e relações entre páginas”. Gonçalves (2008, p. 68) salienta que o *Platypus Wiki*, “não possui mecanismo para verificar a consistência da ontologia, não tem suporte a inferências, e suporta apenas consultas básicas em sua base de dados”.
- **Rhizome** (SOUZIS, 2005). O wiki semântico *Rhizome* suporta a ZML, uma linguagem de script para a criação de conteúdo semântico, e a RxML, uma

---

<sup>20</sup> *Attempto Controlled English* (ACE) é uma linguagem natural controlada. ACE é um subconjunto do inglês com uma gramática restrita e uma semântica formal.

estrutura alternativa ao RDF para a exibição de conteúdo. A ZML pode ser traduzida em RDF. O Rhizome não possui uma linguagem de consulta própria.

- **Semantic MediaWiki (SMW)** (KRÖTZSCH et al., 2007). Iniciado em 2005, esse sistema de wiki semântico se concentra no cenário enciclopédico da Wikipédia e, desse modo, dá ênfase na escalabilidade e na compatibilidade com a Enciclopédia. Nenhuma ontologia predefinida é requerida para a realização das anotações semânticas. O *Semantic MediaWiki* é o sistema de wiki semântico mais conhecido dentre todos os outros e o único com um número significativo de usuários na Web (AL-HUSAIN; EL-MASRI, 2012; KOREN, 2012; KRÖTZSCH et al., 2007; MILLARD et al., 2008; ZAIDAN; BAX, 2011a). Frequentemente escolhido como o representante dos wikis semânticos, o *Semantic MediaWiki* será abordado com maior profundidade na seção 3.4.
- **SweetWiki** (BUFFA et al., 2007). O *SweetWiki* combina indexação social (*tagging*) com ontologias formais, desse modo, usuários podem atribuir facilmente termos de indexação arbitrários às páginas e tais termos serão, posteriormente, associados a conceitos de ontologias subjacentes. O *SweetWiki* pode utilizar a RDFa, a RDFS ou a OWL Lite para a representação do conhecimento e utiliza a SPARQL como linguagem de consulta.

A seguir é apresentado com mais detalhes o principal sistema de wiki semântico, desenvolvido e utilizado por uma grande comunidade de usuários: o *Semantic MediaWiki*.

### 3.4 *Semantic MediaWiki*

Para a maioria dos wikis, públicos ou privados, os objetivos principais de sua existência consistem na organização da informação coletada em uma determinada área de interesse e em seu compartilhamento. Como mencionado anteriormente, wikis são considerados ferramentas fáceis e rápidas para o gerenciamento de conteúdo online. Esse gerenciamento ocorre, principalmente, pela edição de um conteúdo textual, conhecido como *wikitexto*, escrito em uma linguagem de marcação relativamente simples. Esse *wikitexto* está presente em praticamente todas as páginas dos wikis.

A Wikipédia, site que fornece a maior fonte de conhecimento enciclopédico editado colaborativamente, é o exemplo mais conhecido de um site wiki. Porém, as informações

contidas nas páginas da Wikipédia são de difícil utilização por ferramentas externas: **utilizar** a Wikipédia significa, atualmente, apenas **ler** artigos. Não há qualquer modo de acessar informações que estão distribuídas entre múltiplos artigos enciclopédicos ao mesmo tempo. Por exemplo, ainda não é possível consultar a Wikipédia solicitando uma lista contendo todos os filmes da década de 1960 dirigidos por italianos, mesmo que esses dados estejam presentes e distribuídos por entre as várias páginas da Enciclopédia. Embora os dados estejam um tanto estruturados (cada filme tem seu artigo enciclopédico, existem links para atores e diretores, etc.), seus significados são desconhecidos para os computadores, pois não estão representados e descritos de um modo processável por máquinas, ou seja, representados e descritos de maneira formal.

Para solucionar esse problema, Markus Krötzsch e Denny Vrandečić criaram em 2005 o *Semantic MediaWiki* (SMW), “um sistema wiki melhorado com tecnologias da Web Semântica que possibilita aos usuários anotar semanticamente [isto é, marcar semanticamente] o conteúdo do wiki com informações explícitas legíveis por máquinas” (KRÖTZSCH et al., 2007, p. 251, tradução nossa). O *Semantic MediaWiki* é um *software* livre, de código aberto, disponível como uma extensão do *software MediaWiki*.

A integração entre o *Semantic MediaWiki* e o *MediaWiki* está baseada no mecanismo de extensão do último: o *Semantic MediaWiki* se registra para certos eventos ou requisições do *software* e o *MediaWiki* trata de chamar as funções do *Semantic MediaWiki* quando são requeridas. Desse modo, o *Semantic MediaWiki* não sobrescreve qualquer parte do *MediaWiki* e, por isso, pode ser adicionado, sem custos de migração, aos sites que já utilizam o *MediaWiki*.

Embora o *Semantic MediaWiki* tenha sido criado em 2005 por Markus Krötzsch e Denny Vrandečić no *Institut für Angewandte Informatik und Formale Beschreibungsverfahren* (AIFB) (Instituto de Informática da Faculdade de Economia e Engenharia de Negócios) do *Karlsruhe Institut für Technologie* (KIT) (Instituto de Tecnologia de Karlsruhe), hoje ele é mantido por, principalmente, Markus Krötzsch (atualmente pesquisador na Universidade de Oxford) e Jeroen De Dauw (desenvolvedor contratado pela *Wikimedia Foundation*), além de vários outros contribuidores (SMW Project, 2013)<sup>21</sup>.

---

<sup>21</sup> Site oficial do *Semantic Mediawiki*: <http://semantic-mediawiki.org>.



Krötzsch et al. (2007, p. 251) afirmam que, utilizando-se dos dados marcados semanticamente, o *Semantic MediaWiki* soluciona os principais problemas dos wikis atuais:

- **Consistência do conteúdo:** a mesma informação frequentemente ocorre em muitas páginas em um wiki. Como assegurar que tais informações em diferentes páginas estejam consistentes, especialmente se elas podem ser modificadas de modo distribuído?
- **Acesso ao conhecimento:** grandes wikis possuem milhares de páginas. Encontrar e comparar informações de diferentes páginas wiki é um grande desafio e uma atividade demasiadamente demorada para ser concluída.
- **Reuso do conhecimento:** muitos wikis são gerenciados com o desejo de tornar acessíveis informações para uma grande quantidade de pessoas e dispositivos. Mas, nos wikis tradicionais, o conteúdo baseado em texto somente é utilizado para a leitura por humanos por meio de navegadores.

Nesse contexto, é possível observar que as tecnologias da Web Semântica aplicadas nos wikis semânticos podem facilitar o reuso das informações neles contidas. No entanto, para que isso ocorra, as informações precisam estar formalmente representadas e descritas de modo que possam ser processadas por máquinas nas requisições de consultas, permitindo, assim, o acesso, a recuperação e o uso dessas informações.

Visando a apresentar a aplicação das tecnologias da Web Semântica nos wikis semânticos, o capítulo seguinte descreve as atividades de representação, de descrição, de recuperação e de intercâmbio de informações no principal sistema de wiki semântico atual, o *Semantic MediaWiki*.

## **4 Descrição, recuperação e intercâmbio de informações em wikis semânticos**

Este capítulo explora, apresenta e analisa as características do *Semantic MediaWiki* nas atividades de representação, de recuperação e de intercâmbio de informações, tendo como objetivo principal favorecer o entendimento, a discussão e o uso das tecnologias da Web Semântica em ambientes informacionais digitais.

Segundo Ferneda (2012, p. 153), “É interessante observar que no projeto da Web Semântica estão inseridos conceitos e ideias que há muito tempo são utilizados pela Ciência da Informação no tratamento documental”. A Ciência da Informação, por desenvolver e estudar técnicas e teorias para a representação, a descrição, a recuperação e o intercâmbio de informações, provê subsídios teórico-metodológicos que podem ser aproveitados nas atividades de representação e de descrição em wikis semânticos utilizando as tecnologias da Web semântica. Alguns desses subsídios teórico-metodológicos estão apresentados no decorrer deste capítulo.

### **4.1 A descrição de informações em wikis semânticos**

A representação e a descrição de recursos são atividades há muito tempo conhecidas pela Biblioteconomia, uma das áreas de aplicação da Ciência da Informação.

Na Biblioteconomia, mas especificamente na disciplina e na atividade de Catalogação, os profissionais dessa área constroem formas de representação que são inseridas em catálogos. Essas formas de representação devem ser abrangentes, do ponto de vista cognitivo, e sintéticas para a tomada de decisão dos usuários sobre os recursos buscados. Surgindo a partir de uma descrição padronizada dos recursos, essas formas de representação objetivam tornar o recurso informacional representado único e não confundível com nenhum outro recurso presente no acervo. Nesse sentido, ao descrever o recurso de forma abrangente, é possível criar e/ou multiplicar seus pontos de acesso com vistas a sua identificação, localização e recuperação (SANTOS, 2008, p. 155-171).

Nesse contexto, como afirma Catarino e Souza (2012, p. 78), a Catalogação está intimamente associada com o *Resource Description Framework* (RDF), um “modelo de descrição dos recursos” que, segundo as autoras, é “o alicerce da Web Semântica”.

Como os wikis semânticos se utilizam de tecnologias da Web Semântica, eles fazem uso do modelo RDF para a descrição de recursos. Mas antes de possuírem capacidades de

descrição de recursos, os wikis possuem mecanismos para a estruturação e a organização das informações neles contidas. A seção 4.1.1, Estruturação e organização do conteúdo em um wiki, abordará brevemente esses mecanismos que serão utilizados para a realização das descrições baseadas nas tecnologias semânticas.

#### 4.1.1 Estruturação e organização do conteúdo em um wiki

Nos wikis baseados no *MediaWiki* (um dos mais populares sistemas wiki), existem seis principais mecanismos que servem para a organização e/ou a estruturação do conteúdo, a saber: páginas (e subpáginas), *namespaces*, categorias, redirecionamentos, páginas de desambiguação e predefinições.

O mecanismo estrutural da maioria dos wikis é a organização do conteúdo distribuída por entre páginas wiki.

O principal método de entrada de informação em um wiki é a utilização do *wikitexto*, uma linguagem de marcação simplificada que é transformada em HTML para a leitura do texto contido nas páginas wiki. Conseqüentemente, o *wikitexto* já fornece várias facilidades para a formatação e, até mesmo, algumas facilidades para a estruturação do conteúdo. Para definir a inter-relação de páginas no wiki, a utilização de hiperlinks é a ferramenta mais importante.

No *MediaWiki*, essas páginas são ainda agrupadas e separadas pelos *namespaces*<sup>22</sup> (também chamados de domínios, espaços de domínio, espaços de nomes ou espaços nominais), que distinguem diferentes tipos de páginas de acordo com as suas funções.

O *namespace* de uma página é identificado por um prefixo específico, como “Usuário:” para páginas destinadas a identificação de usuários, “Ajuda:” para páginas de documentação, tutoriais, etc., “Projeto:” para páginas relacionadas às políticas de edição, normas, comunicados oficiais, etc. Os *namespaces* não são criados por usuários, mas sim definidos por administradores do site.

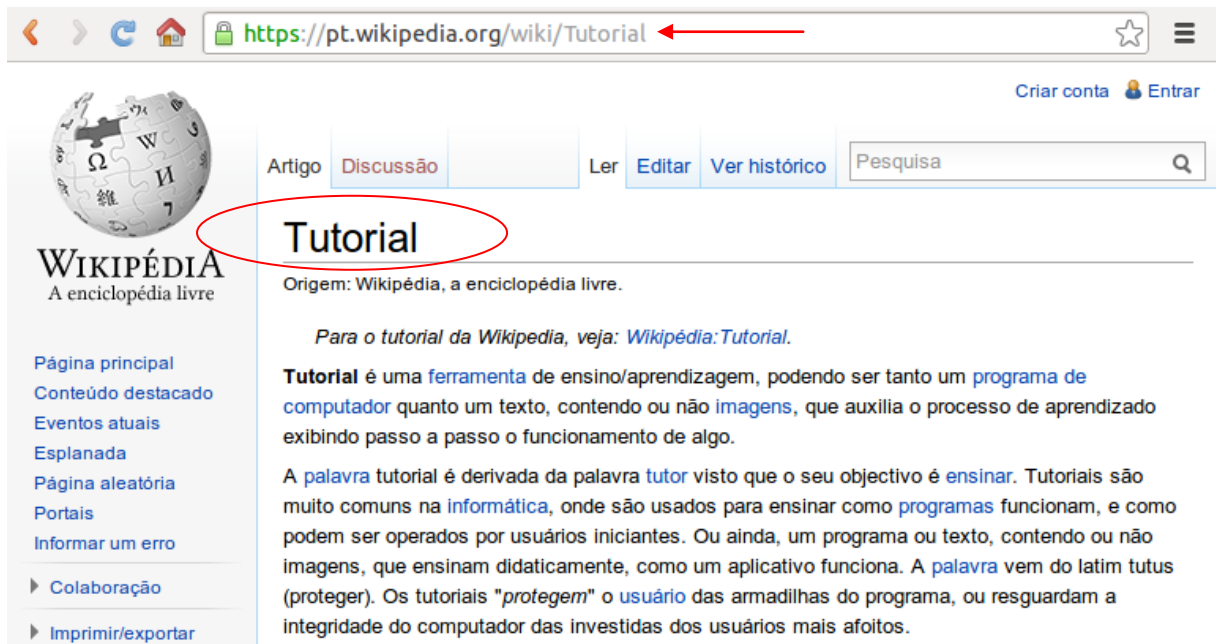
Por padrão, quando uma página é criada, ela está associada ao *namespace* principal (que não possui prefixo). No entanto, se o usuário deseja que a página por ele criada esteja associada a um *namespace* que não seja o principal, por exemplo, o *namespace* “Ajuda:”, é

---

<sup>22</sup> Apesar de possuir funções semelhantes o mecanismo de *namespace* do *MediaWiki* não deve ser confundido com o mecanismo de *namespace* da XML, apresentado na seção 2.4.2, *Extensible Markup Language* (XML).

necessário que o título da página seja precedido pelo prefixo do *namespace* escolhido (NAMESPACES, 2012). Para ilustrar esse exemplo, considere a Figura 11 e a Figura 12, que mostram, respectivamente, as páginas “Tutorial” e “Ajuda:Tutorial” na Wikipédia.

Figura 11 – Página no *namespace* principal (sem prefixo).



Fonte: Disponível em:

<<https://pt.wikipedia.org/w/index.php?title=Tutorial&oldid=35646427>>.

Acesso em: 24 jun. 2013.

Figura 12 – Página no *namespace* Ajuda (com o prefixo Ajuda:).



Fonte: Disponível em:

<<https://pt.wikipedia.org/w/index.php?title=Ajuda:Tutorial&oldid=36189891>>.

Acesso em: 24 jun. 2013.

*Namespaces* diferentes são utilizados para distinguir as funções semânticas que as páginas wiki podem representar: o *namespace* principal é utilizado para agrupar as páginas que descrevem os elementos (instâncias) do domínio de interesse; o *namespace* “Categoria:” é utilizado para agrupar as páginas que representam as classes, que por sua vez são utilizadas para classificar as instâncias e também para classificar outras classes (subclasses); já o *namespace* “Propriedade:” é utilizado para agrupar as páginas que representam relações entre dois elementos ou entre um elemento e um valor de dado literal.

Vários sistemas wiki se utilizam de links para a classificação/categorização de páginas. Por exemplo, buscando por todas as páginas que contenham um link cujo destino aponta para a página “Brasil” é um modo de encontrar informações sobre o país em questão. No *MediaWiki*, porém, esse uso tem sido substituído por um sistema mais elaborado de categorias. Cada página pode ser atribuída a uma ou mais categorias e cada categoria possui uma página no *namespace* “Categoria:”. Páginas de categorias, isto é, páginas que representam categorias e que listam todas as páginas categorizadas naquela determinada categoria, são utilizadas para navegar pelas páginas categorizadas e também para organizar as categorias de um modo hierárquico. As páginas de categorias e suas hierarquias podem ser editadas pelos usuários por meio de uma sintaxe específica do *wikitexto*. Um exemplo de página de categoria é apresentado na Figura 13.

Figura 13 – Página de categoria.

The screenshot shows the Wikipedia page for the category "Espécies descritas em 1927". The page layout includes a sidebar on the left with navigation links, a main content area with a title, a description, a year navigation bar, and a list of species organized by letter. The species listed are:

- A**
  - Aerangis collum-cygni
  - Alethe choloensis
  - Amietia vertebralis
  - Aotus miconax
  - Assimineea palauensis
- B**
  - Bonobo
  - Brugia malayi
  - Buchema hadromeres
- C**
  - Caecorhanda brasiliensis
  - Calconiscellus gotscheensis
  - Callicebus purinus
  - Callicebus regulus
  - Chirixalus hansenae
  - Chirixalus nongkhorensis
- H**
  - Hamptophryne boliviana
  - Hapalomys delacouri
  - Helminthoglypta mohaveana
  - Hesperilla flavescens
  - Hylodes meridionalis
- L**
  - Labeotropheus fueleborni
  - Lactarius clarkeae
  - Lactarius russuliformis
  - Ladigesocypris ghigii
  - Lagarto-da-ilha-de-ratas
  - Lampsilis rafinesqueana
  - Lampsilis streckeri
  - Leopardo-persa
  - Lepus comus
  - Limnonectes arathooni
  - Limnonectes dammermani
- P**
  - Paladilhia hungarica
  - Paludinella semperi
  - Paludinella vitrea
  - Peixe-mandarim
  - Philautus andersoni
  - Philautus fergusonianus
  - Philautus malcolmsmithi
  - Philautus rugatus
  - Philautus zimmeri
  - Phrynobatrachus liberiensis
  - Pleurobema marshalli
  - Pleurodema tucumana
  - Polypedates longinasus
  - Pseudopaludicola boliviana
  - Pseudotropheus macrophthalmus
- R**

Fonte: Disponível em:

<[http://pt.wikipedia.org/w/index.php?title=Categoria:Esp%C3%A9cies\\_descritas\\_em\\_1927&oldid=34835483](http://pt.wikipedia.org/w/index.php?title=Categoria:Esp%C3%A9cies_descritas_em_1927&oldid=34835483)>. Acesso em: 24 jun. 2013.

Outros problemas de estruturação encontrados em grandes wikis são os títulos sinônimos e os títulos homônimos. Em um processo de edição descentralizado podem surgir várias páginas diferentes para o mesmo objeto de interesse, gerando, desse modo, páginas com títulos sinônimos. Para esses casos, o *MediaWiki* possui um mecanismo de redirecionamento pelo qual páginas wiki são utilizadas para enviar todas as requisições de uma determinada página diretamente para outra determinada página. A Figura 14 e a Figura 15 apresentam duas páginas de redirecionamento que levam o usuário a uma mesma página, de modo semelhante à remissiva “ver” dos catálogos, dicionários, enciclopédias, etc.

Figura 14 – Redirecionamento de MASP para Museu de Arte de São Paulo.



Fonte: Disponível em:  
 <<http://pt.wikipedia.org/w/index.php?title=MASP&oldid=21075347>>.  
 Acesso em: 24 jun. 2013.

Figura 15 – Redirecionamento de Museu de Arte de São Paulo Assis Chateaubriand para Museu de Arte de São Paulo.



Fonte: Disponível em:  
 <[http://pt.wikipedia.org/w/index.php?title=Museu\\_de\\_Arte\\_de\\_S%C3%A3o\\_Paulo\\_Assis\\_Chateaubriand&oldid=14905582](http://pt.wikipedia.org/w/index.php?title=Museu_de_Arte_de_S%C3%A3o_Paulo_Assis_Chateaubriand&oldid=14905582)>.  
 Acesso em: 24 jun. 2013.

Homônimos, por sua vez, ocorrem quando o título de uma página é ambíguo e pode se referir a diferentes objetos de interesse dependendo do contexto. Esse problema é resolvido pelas chamadas páginas de desambiguação que sucintamente listam os diferentes significados possíveis de um título. As páginas referidas pela página de desambiguação geralmente são, então, sufixadas com um qualificador entre parênteses para distingui-las,

por exemplo, “1984 (livro)”, “1984 (filme de 1956)”, “1984 (álbum de Van Halen)”, etc. A Figura 16 apresenta uma página de desambiguação.

Figura 16 – Página de desambiguação.

The screenshot shows the Portuguese Wikipedia page for 'Manga'. The browser address bar displays 'https://pt.wikipedia.org/wiki/Manga'. The page features the Wikipedia logo and navigation links on the left. The main content area includes a search bar, tabs for 'Artigo', 'Discussão', 'Ler', 'Editar', and 'Ver histórico', and a search input field. The title 'Manga' is prominently displayed, followed by the text 'Origem: Wikipédia, a enciclopédia livre.' Below this, a blue banner indicates that the page is a disambiguation page. A list of related articles is provided, including 'Biologia', 'Localidades', 'Pessoas', 'Artes', and 'Vestuário'. The 'Biologia' section lists 'Manga (fruta)', 'Manga-papaia', 'Manga (Arctiidae)', and 'Cajá-manga'. The 'Localidades' section lists 'Manga (Minas Gerais)', 'Manga (Burkina Faso)', and 'Ribeira da Manga'. The 'Pessoas' section lists 'Agenor Gomes', 'Carlos Manga', 'Luciano Manga', and 'Halilton Corrêa de Aruda'. The 'Artes' section lists 'Amarelo Manga' and 'Manga (banda)'. The 'Cinema' section lists 'Amarelo Manga' and the 'Música' section lists 'Manga (banda)'.

Fonte: Disponível em:

<<http://pt.wikipedia.org/w/index.php?title=Manga&oldid=34456103>>.

Acesso em: 24 jun. 2013.



Outro recurso significativo utilizado para a estruturação do conteúdo no *MediaWiki* é o mecanismo de predefinições (*template system*). O analisador sintático do *MediaWiki* substitui marcações de predefinições com o texto contido nas próprias páginas dessas predefinições. O texto predefinido pode ser parametrizado e, desse modo, utilizado para alcançar uma alta consistência de conteúdo, uma vez que, uma tabela, por exemplo, pode ser definida uma única vez (na página da predefinição), mas utilizada em várias páginas wiki (TEMPLATE, 2013).

Segundo Krötzsch et al. (2007), aderindo aos princípios básicos do *MediaWiki*, os dados semânticos no *Semantic MediaWiki* também são estruturados por meio de páginas, de modo que todo o conteúdo semântico está explicitamente atrelado a páginas wiki. Semanticamente falando, cada página corresponde a uma entidade ontológica (incluindo classes e propriedades) que pode ser descrita por marcações presentes em tal página. Segundo os autores, essa localização é crucial para a manutenção das informações: se a informação é reutilizada em vários locais, os usuários precisam saber onde a informação se originou.

Além dos mecanismos citados, o *MediaWiki* possui ainda outros modos de estruturar o conteúdo textual dentro das páginas como, por exemplo, por seções e tabelas. O *Semantic MediaWiki*, porém, objetiva coletar e armazenar informações sobre o conceito, isto é, algo abstrato representado por uma página, não sobre o texto associado a ela.

Para Krötzsch et al. (2007), o principal pré-requisito para a exploração de tecnologias semânticas é a disponibilidade de dados semânticos adequadamente estruturados. Para esse objetivo, o *Semantic MediaWiki* introduz modos para adicionar mais estrutura ao *MediaWiki* por meio da marcação (ou anotação) do conteúdo textual do wiki.

De acordo com Krötzsch et al. (2007), o *Semantic MediaWiki* coleta os dados semânticos por meio das anotações semânticas que os usuários, cientes ou não, adicionam ao *wikitexto* das páginas wiki. Desse modo, as anotações (ou marcações) semânticas são utilizadas para explicitar, por exemplo, relacionamentos entre entidades e instanciamentos de classes.

Embora a sintaxe de marcação seja relevante para os editores do wiki, ela é apenas uma pequena parte de todo o sistema do *Semantic MediaWiki*. Mais relevante do que a sintaxe de marcação é a estrutura conceitual subjacente do *Semantic MediaWiki*, que é baseada, principalmente, em propriedades e em tipos de dados.

Nesse contexto, é possível observar a possibilidade da descrição de recursos (representados, descritos e apresentados por meio de páginas wiki) utilizando-se de propriedades, valores e classes. O relato da atividade de descrição de recursos é apresentado nas seções 4.1.2, Propriedades e sua utilização, e 4.1.3, Mapeamento para a linguagem OWL.

#### 4.1.2 Propriedades e sua utilização

Cada comunidade de descrição está interessada em descrever diferentes relacionamentos entre entidades dependendo de sua área de interesse. Por exemplo, uma biblioteca descreve um livro de modo diferente de uma livraria, que por sua vez descreve um livro de modo diferente de uma editora, e assim por diante. Partindo desse princípio, o *Semantic MediaWiki* permite que os usuários (gestores, administradores, etc.) controlem o conjunto de propriedades disponíveis para a realização dessas descrições.

Segundo Krötzsch et al. (2007), o mecanismo de propriedades presente no *Semantic MediaWiki* segue o padrão de formalismo da Web Semântica: propriedades são um dos mecanismos expressivos centrais de todo o funcionamento do sistema.

As propriedades no *Semantic MediaWiki* são utilizadas para expressar os relacionamentos entre uma entidade semântica (representada por uma página wiki) e alguma outra entidade semântica ou entre uma entidade semântica e um valor de dado (valor literal).<sup>23</sup>

O *MediaWiki* não oferece, por padrão, qualquer mecanismo para a atribuição de propriedades e seus valores às páginas wiki. Quando o *Semantic MediaWiki* é instalado com o *MediaWiki*, a atribuição de propriedades e valores passa a ser possível, podendo, assim, tornar-se disponível uma grande quantidade de dados devido aos relacionamentos entre as páginas e entre as páginas e seus valores de dados.

O tipo de relacionamento mais aparente nos wikis tradicionais são os links. Cada link estabelece *algum* relacionamento entre duas páginas do wiki, sem especificar, no entanto, de que tipo esse relacionamento é ou se esse relacionamento é significativo para algum propósito além de mera navegação. Krötzsch et al. (2007) afirmam que o *Semantic MediaWiki* permite que os links sejam caracterizados por propriedades, de maneira que o

---

<sup>23</sup> Propriedades e valores de dados (valores literais) foram abordados na seção 2.4.3, *Resource Description Framework (RDF)*.

alvo do link se torne o valor de uma determinada propriedade. Os autores também atentam que nem todas as propriedades requerem outras páginas wiki como seus valores: quantidades numéricas, datas e coordenadas geográficas são exemplos de outros tipos de dados disponíveis como valores para as propriedades.

Para um exemplo concreto, considere o primeiro parágrafo da Figura 17. Os elementos de marcação são fáceis de ler: `'''texto'''` é utilizado para formatar o texto em negrito; `[[texto]]` é utilizado para transformar um termo em um link para a página com aquele nome. Portanto, `[[capital]]` e `[[Brasil]]` são marcações que transformam os termos em links que apontam para as páginas com o nome “Capital” e “Brasil”. Esses links não carregam qualquer significado. Para declarar que Brasília (objeto de interesse da página em questão) é a capital do Brasil, um usuário do *Semantic MediaWiki* apenas precisaria estender a sintaxe padrão do link `[[Brasil]]` para `[[capital do::Brasil]]`. Essa marcação declara que o indivíduo chamado “Brasília” tem uma propriedade chamada “capital do” com o valor “Brasil”.

Figura 17 – *Wikitexto* de uma página wiki sobre Brasília no *MediaWiki* (primeiro parágrafo) e no *Semantic MediaWiki* (segundo parágrafo).

```
'''Brasília''' é a [[capital]] federal do [[Brasil]]. Em 2011, a
população de Brasília era estimada em 2.609.997 habitantes. Brasília
cobre uma área de 5.801.937 km².
```

```
[[Categoria:Cidade]]
```

```
'''Brasília''' é a [[capital]] federal do [[capital do::Brasil]]. Em
2011, a população de Brasília era estimada em
[[população::2.609.997]] habitantes. Brasília cobre uma área de
[[área::5.801.937 km²]].
```

```
[[Categoria:Cidade]]
```

Fonte: Adaptado de Krötzsch et al. (2007, p. 253).

O primeiro parágrafo da Figura 17 evidencia também valores de dados que não correspondem a links, mas que são importantes para a descrição de um indivíduo, como o número de habitantes e a área coberta. Para declarar esses dados literais, a sintaxe utilizada é a mesma utilizada para declarar relações entre páginas: `[[propriedade::valor]]`. Como mostrado no segundo parágrafo da Figura 17, para declarar que “Brasília” possui uma população de 2.609.997 habitantes, bastaria utilizar a marcação `[[população::2.609.997]]`. Observa-se que, nesse caso, 2.609.997 não se refere a uma página wiki cujo nome é “2.609.997”, mas sim a um valor de dado absoluto (literal). O

*wikitexto* contendo essas marcações (segundo parágrafo da Figura 17) é exibido com o *Semantic MediaWiki* como mostrado na Figura 18.

Figura 18 – Uma página em que foram marcadas propriedades e valores.

The screenshot shows the Semantic MediaWiki page for Brasília. The page title is "Brasília". The main text states: "Brasília é a capital federal do Brasil. Em 2011, a população de Brasília era estimada em 2.609.997 habitantes. Brasília cobre uma área de 5.801.937 km²." Below the text, there is a table of properties and values:

Dados sobre Brasília	
Capital do	Brasil +
População	2.609.997 +
Área	5.801.937 km² +

Red boxes and arrows highlight specific elements: "Recurso" points to the "Cidade" category; "Valores" points to the numerical values in the table; and "Propriedades" points to the property names in the table.

Fonte: Elaborada pelo autor.

Para especificar que os valores de uma propriedade devem ser valores literais, um usuário precisa criar a página da referida propriedade no *namespace* "Propriedade:" e declarar o tipo de dado (*datatype*) que essa propriedade deverá aceitar (texto, número, URL, e-mail, coordenada geográfica, temperatura, data, etc.).

Portanto, assim como categorias possuem páginas no wiki (Figura 13) que servem para propósitos de organização e de navegação do conteúdo, as propriedades também possuem as suas páginas para propósitos similares.

Um wiki pode conter a página "Propriedade:População" na qual "Propriedade:" é o prefixo de *namespace* que funciona para indicar que a página é uma propriedade (não um indivíduo, uma classe, um tipo de dado, um usuário, uma página de ajuda, etc.). A página da propriedade pode conter uma descrição textual sobre a propriedade em questão visando a auxiliar os usuários a empregá-la consistentemente no wiki. Além disso, é comum utilizar a página da propriedade para especificar características semânticas da própria propriedade. Uma dessas características semânticas é a especificação do tipo de dado, mencionado anteriormente.

No caso da propriedade "População" um usuário administrador do wiki faria a marcação "[[tem tipo::número]]" para descrever que a propriedade espera somente

valores numéricos (desse modo não gerando links na exibição)<sup>24</sup>. A Figura 19 mostra uma página que descreve uma propriedade tendo seu tipo de dados especificado como numérico.

Figura 19 – Página de uma propriedade.

The screenshot shows a Semantic MediaWiki page for the property "População". The page layout includes a sidebar on the left with navigation links like "Página principal", "Mudanças recentes", and "Ajuda". The main content area features a search bar, tabs for "Property", "Discussão", "Ler", "Editar", and "Ver histórico", and a search input field. The title is "Propriedade:População". Below the title, it states "Esta propriedade tem valores do tipo número." and "Páginas que utilizam a propriedade 'População'". It lists four pages using this property, grouped by the first letter of the city name: B (Brasília), F (Florianópolis), R (Rio de Janeiro), and S (São Paulo). Each entry shows the city name, a plus icon, a down arrow, the population value, and another plus icon. At the bottom, there is a section "Dados sobre População" with a "Feed RDF" link and a "Has type" label pointing to "Number". A footer note states "Esta página foi modificada pela última vez às 14h45min de 24 de junho de 2013."

Fonte: Elaborada pelo autor.

Os tipos de dados que podem ser atribuídos a uma propriedade no *Semantic MediaWiki* estão baseados nos principais tipos de dados apresentados pela especificação da *XML Schema*. Cada tipo de dado possui seu próprio método de processamento, armazenamento e exibição padrão no wiki. Os tipos de dados mais comuns presentes no *Semantic MediaWiki* são: página wiki, texto, número, data, temperatura, coordenada geográfica, numero de telefone, valor booleano, URL, URI, e-mail, código, quantidade e registro.

<sup>24</sup> A propriedade "tem tipo" é uma propriedade pré-definida no *Semantic MediaWiki*, com um significado também já definido por padrão.

Salienta-se que, apesar da possibilidade da realização das marcações semânticas diretamente no *wikitexto* de uma página, tal opção não é obrigatória para o usuário final. Na verdade, na maioria dos wikis semânticos são utilizados formulários<sup>25</sup> para a entrada dos dados. Isso facilita a padronização das informações, já que os formulários dispõem de recursos tais como o “auto completar” para valores cadastrados anteriormente, a validação de dados por meio de padrões de caracteres, o uso de botões, menus, abas, recursos de ajuda, entre outros. A Figura 20 ilustra uma página em que os dados são preenchidos por uma interface de formulário.

---

<sup>25</sup> Para a utilização de formulários em wikis semânticos do tipo *Semantic MediaWiki*, veja a extensão *Semantic Forms*: [http://mediawiki.org/wiki/Extension:Semantic\\_Forms](http://mediawiki.org/wiki/Extension:Semantic_Forms).

Figura 20 – Página de edição de dados via formulário.

Wikincat

Página inicial  
Pesquisa  
Catalogação  
Aquisição  
Circulação  
Registro de usuário  
Conversor de ISBN  
Mudanças recentes

► Ferramentas  
► Donate

português do Brasil Jaider Discussão Links administrativos Preferências Páginas vigiadas Contribuições Sair

Título Ler Editar Editar fonte Ver histórico Pesquisa

## Editar manifestação: A ciência da informação (2004)

Dados de manifestação e de expressão **Dados de item**

Obra: La science de l'information

Título preferido: A ciência da informação

Subtítulo:

Responsabilidade: Yves-François Le Coadic; tradução de Maria Yêda F. S. de Filgueiras Gomes

Edição: 2. ed.

Local de publicação: Brasília

Editora: Bnquet de Lemos/Livros

Ano de publicação: 2004  
[Caso não tenha os dados da publicação, informe os dados da distribuição ou da manufatura]

Ano de *copyright*: 2004  
[Se for uma publicação seriada, informe o título, o responsável, o ISSN e a numeração dentro da série]

Tipo de conteúdo: text

Idiomas: pt

Tipo de mídia: unmediated

Tipo de suporte: volume

Modo de emissão: single unit

Extensão do texto: 124 páginas

Ilustrações: Exemplo: ilustrado

Dimensões: Exemplo: 23 cm

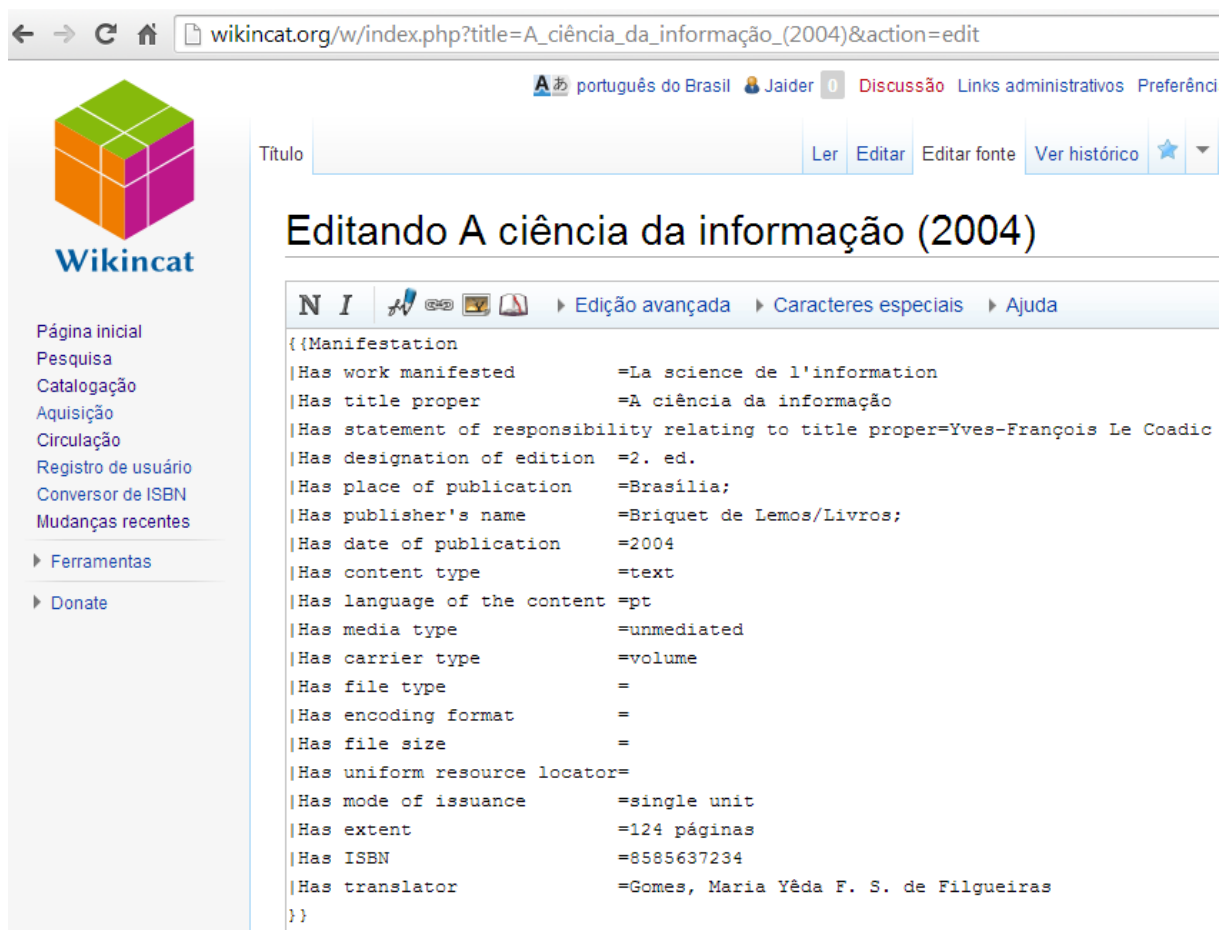
Nota:

Sumarização do conteúdo: Resumo ou sinopse

Fonte: Elaborada pelo autor.

Os dados inseridos via formulários fazem parte da marcação semântica armazenada em forma de *wikitexto* (utilizando predefinições). A Figura 21 apresenta um *wikitexto* contendo dados inseridos via formulário.

Figura 21 – Página contendo dados inseridos via formulário.



The screenshot shows the Wikicat website interface. The browser address bar displays the URL: `wikicat.org/w/index.php?title=A_ciência_da_informação_(2004)&action=edit`. The page title is "Editando A ciência da informação (2004)". The main content area shows a table of metadata for a manifestation, with the following data:

Property	Value
Has work manifested	=La science de l'information
Has title proper	=A ciência da informação
Has statement of responsibility relating to title proper	=Yves-François Le Coadic
Has designation of edition	=2. ed.
Has place of publication	=Brasília;
Has publisher's name	=Briquet de Lemos/Livros;
Has date of publication	=2004
Has content type	=text
Has language of the content	=pt
Has media type	=unmediated
Has carrier type	=volume
Has file type	=
Has encoding format	=
Has file size	=
Has uniform resource locator	=
Has mode of issuance	=single unit
Has extent	=124 páginas
Has ISBN	=8585637234
Has translator	=Gomes, Maria Yêda F. S. de Filgueiras

Fonte: Elaborada pelo autor.

Esta seção teve por objetivo apresentar as propriedades e sua utilização para a descrição de recursos de maneira simples em um wiki semântico, isto é, sem levar em consideração aspectos relacionados às ontologias. A seção a seguir apresenta alguns meios para conseguir realizar descrições mais rigorosas com relação aos aspectos semânticos providos por ontologias e/ou por vocabulários (internos ou externos ao wiki).

#### 4.1.3 Mapeamento para a linguagem OWL

Com o *Semantic MediaWiki* é possível adicionar vários tipos de informações ontológicas ao wiki (VRANDEČIĆ; KRÖTZSCH, 2006). Conforme apresentado na seção 2.4.5, *Web Ontology Language* (OWL), é possível dividir as principais entidades ontológicas em **instâncias** que representam os indivíduos de um domínio descrito, em **classes** que representam conjuntos de indivíduos, e em **propriedades** que representam relações entre indivíduos. A maneira com que o *Semantic MediaWiki* representa o conhecimento foi



parcialmente inspirado pela OWL DL e, desse modo, pode-se mapear os elementos do wiki, isto é, as páginas individuais de conteúdo, para os elementos da linguagem OWL DL. Nesse sentido,

- **Indivíduos** OWL são representados por páginas normais do wiki. Essas páginas geralmente constituem a maior parte do conteúdo do site e estão localizadas no *namespace* principal. No entanto, podem existir páginas em *namespaces* adicionais, tais como “Imagem:” ou “Usuário:”, que, adicionalmente, podem ser interpretadas como indivíduos OWL.
- **Classes** OWL são representadas por páginas de categorias. Conforme visto na seção 4.1.1, o mecanismo de categorias é um dos mecanismos de organização de páginas mais utilizados pelos usuários de um wiki. Esse mecanismo possibilita que as páginas sejam classificadas de acordo com algum critério escolhido previamente. Essas categorias são representadas como páginas pertencentes ao *namespace* “Categoria:”. As categorias podem ser organizadas de um modo multi-hierárquico, tal como acontece com classes OWL.
- Ao contrário dos indivíduos (páginas comuns) e das classes (páginas de categorias) diretamente fornecidos pelo *MediaWiki*, **propriedades** OWL fazem parte especificamente do *Semantic MediaWiki*. Tanto as propriedades de objeto (*object properties*) quanto as propriedades de dados (*datatype properties*) são declaradas no *namespace* “Propriedade:”.

Assim, baseado no mapeamento para a linguagem OWL, o SMW permite aos usuários realizar declarações ontológicas em páginas wiki. Para demonstrar esse pequeno mapeamento entre os elementos ontológicos da OWL e os elementos de descrição do SMW, é apresentado o Quadro 8.

Quadro 8 – Representação dos principais elementos da linguagem OWL no *Semantic MediaWiki*.

OWL	<i>Semantic MediaWiki</i>
<i>owl:NamedIndividual</i>	Página normal do wiki
<i>owl:Class</i>	Página no <i>namespace</i> Categoria
<i>owl:ObjectProperty</i>	Página no <i>namespace</i> Propriedade (com tipo de dados <i>Page</i> )
<i>owl:DatatypeProperty</i>	Página no <i>namespace</i> Propriedade (com algum outro tipo de dados)
Declaração sobre uma entidade	Marcação wiki ( <i>wikitexto</i> ) utilizado
Relação entre entidades – <i>owl:ObjectProperty</i>	[[ <i>Nome da propriedade::Objeto</i> ]]
Relação entre uma entidade e um valor de dado literal – <i>owl:DatatypeProperty</i>	[[ <i>Nome da propriedade::Valor literal</i> ]]
Instanciamento de classe – <i>rdf:type</i>	[[Categoria: <i>Nome da classe</i> ]] (em páginas normais)
Instanciamento de classe (subclasse) – <i>rdf:subClassOf</i>	[[Categoria: <i>Nome da classe</i> ]] (em páginas de categorias)

Fonte: Adaptado de Vrandečić e Krötzsch (2006, p. 21).

Assim, conforme visto em 4.1.2, para realizar declarações sobre indivíduos, classes e propriedades representadas no wiki, são realizadas marcações semânticas utilizando-se do *wikitexto* nas páginas que representam tais elementos ontológicos.

O *Semantic MediaWiki* inclui uma função de exportação que gera documentos OWL/RDF de acordo com o mapeamento do apresentado no Quadro 8. A função de exportação também associa URIs a todas as páginas wiki. Esses URIs correspondem aos conceitos descritos (de modo formal, legíveis por máquinas, em OWL/RDF) e suas respectivas páginas de descrição (apresentação do conteúdo, legíveis por humanos, em HTML).

O Quadro 8 indica que, embora algumas descrições simples sobre o domínio sejam explicitadas, um wiki semântico é principalmente utilizado para fornecer descrições concretas de indivíduos e seus relacionamentos. Como visto na seção 2.4.4, *RDF Vocabulary Description Language: RDF Schema* (RDFS), essa parte do conhecimento descrito é conhecido como conhecimento assertivo. Por padrão, o *Semantic MediaWiki* se limita a explicitar esse tipo de conhecimento, embora também seja possível explicitar o conhecimento de ordem mais complexa, isto é, o conhecimento terminológico. Essa característica padrão do sistema permite que as marcações semânticas exportadas em OWL/RDF sejam logicamente consistentes em relação à especificação da OWL DL.

Para realizar o mapeamento entre um indivíduo, uma classe ou uma propriedade presente no wiki com um indivíduo, uma classe ou uma propriedade definida de maneira formal e complexa em uma ontologia externa ao wiki, utiliza-se da propriedade especial “*Equivalent URI*”, já presente no *Semantic MediaWiki*.

O Quadro 9 apresenta o mapeamento entre as propriedades OWL que indicam equivalências e a propriedade especial “*Equivalent URI*”.

Quadro 9 – Mapeamento para conceitos externos ao wiki por meio de URIs.

OWL	Sintaxe da linguagem de marcação wiki
<i>owl:sameAs</i>	[[Equivalent URI::URI]] em páginas normais
<i>owl:equivalentClass</i>	[[Equivalent URI::URI]] em páginas de categorias
<i>owl:equivalentProperty</i>	[[Equivalent URI::URI]] em páginas de propriedades

Fonte: Adaptado de Vrandečić e Krötzsch (2006, p. 21).

A Figura 22 ilustra a utilização da propriedade especial “*Equivalent URI*”.

Figura 22 – Exemplo de um mapeamento entre uma propriedade local do wiki e uma propriedade de um vocabulário externo por meio de seu URI.

The screenshot shows a Semantic MediaWiki page for the property "Criador". The page title is "Propriedade:Criador". The main content area displays the property name and its type, "texto". It also shows that the property is equivalent to the "creator" property of the Dublin Core ontology, with the URI "http://purl.org/dc/elements/1.1/creator". The page includes a sidebar with navigation links such as "Página principal", "Mudanças recentes", "Ajuda", "Ferramentas", and "Semantic Tools". There is also a search bar and a "Feed RDF" link.

Fonte: Elaborada pelo autor.

Vrandečić e Krötzsch (2006) alertam que é necessário cuidado quando se faz uso desse mapeamento para conceitos de ontologias externas, pois essas declarações, que fogem ao controle do wiki, podem levar a ontologia gerada a um estado lógico inconsistente e, devido a complexidade e a quebra de regras da linguagem OWL DL, ao uso inesperado da linguagem OWL Full.

Esta seção abordou a descrição de informações no *Semantic MediaWiki*, mais especificamente, destacou como as informações podem ser estruturadas, organizadas e descritas utilizando páginas wiki, *namespaces*, categorias (classes), redirecionamentos, páginas de desambiguação, predefinições, propriedades e tipos de valores de dados. As próximas seções deste capítulo abordarão, respectivamente, a recuperação e o intercâmbio de informações no *Semantic MediaWiki*.

#### 4.2 A recuperação de informações em wikis semânticos

Embora o processo de marcação semântica (descrição de recursos) seja uma tarefa relativamente simples de realizar nos wikis semânticos, a maioria dos usuários iriam justificadamente negligenciar esse processo se ele não trouxesse benefícios imediatos. Nesse sentido, pelo menos dois são os benefícios imediatos trazidos pela marcação semântica nos wikis: a navegação por meio dos dados e a consulta por esses dados.

Como visto na Figura 18, a página wiki que possui marcações semânticas pode exibir os dados de forma sumarizada em uma área específica abaixo do conteúdo textual. Essa área é chamada de *Factbox* (caixa de fatos, em uma tradução livre). Nessa *Factbox* são informados: o recurso descrito (sujeito), suas propriedades (predicados) e os valores dessas propriedades (objetos). Além disso, a *Factbox* pode fornecer avisos quando há erros nos dados (quando um dado, por exemplo, não corresponde ao tipo de dado especificado pela propriedade) e também oferecer links para funções relacionadas aos dados (dados geográficos, por exemplo, podem conter links para serviços de mapas, externos ao wiki).

Esses links são utilizados para a navegação pelo wiki por meio do conteúdo semântico (KRÖTZSCH et al., 2007). O título da página presente no cabeçalho da *Factbox* leva o usuário a uma interface de navegação semântica que exibe não apenas as marcações presentes na dada página, mas também todas as marcações em que a dada página é utilizada como valor de uma propriedade. A figura de uma lupa, próxima a cada valor, leva o usuário a uma busca inversa por todas as páginas com marcações similares. Portanto, ao clicar em uma lupa próxima ao valor “2004” de uma propriedade “data de publicação” o wiki buscará por todos os recursos (sujeitos) que possuem a propriedade em questão com o valor “2004” e exibirá esses resultados. Além dessas funcionalidades, a *Factbox* exibe links para as páginas de propriedades, as quais, por sua vez, listam todas as marcações para a dada propriedade, como exemplificada pela Figura 19, mencionada anteriormente.

Segundo Krötzsch et al. (2007), todas essas funcionalidades de navegação são interconectadas por links apropriados, desse modo, permitindo aos usuários navegar por entre a base de conhecimento semântico do wiki.

Para além da mera navegação, o *Semantic MediaWiki* inclui uma linguagem de consulta (*query language*) chamada *Ask*, que possibilita o acesso ao conhecimento formalizado por meio da marcação semântica presente no wiki.

Krötzsch et al. (2007) declaram que a linguagem de consulta pode ser utilizada de dois modos: para consultar diretamente os dados ou para adicionar (embutir) os resultados de uma consulta em páginas wiki. Esse último modo possibilita aos editores adicionar a uma página listas ou tabelas criadas dinamicamente e, conseqüentemente, tornar disponíveis os resultados atualizados de uma consulta aos leitores do wiki, os quais não precisam sequer saber sobre consultas semânticas.

A sintaxe da linguagem de consulta *Ask* é semelhante ao *wikitexto* embora sua semântica corresponda a certas expressões da OWL DL (KRÖTZSCH et al., 2007). Segundo Krötzsch et al. (2007), condições fundamentais são codificadas como valores atômicos de consultas, os quais possuem sintaxe similar àquela das marcações semânticas. Por exemplo, `[[localizado em::São Paulo]]` é o valor atômico da consulta para todas as páginas com essa marcação semântica. Consultas com outros tipos de propriedades e pertencimentos a categorias são construídas seguindo o mesmo princípio. Ao invés de valores fixos únicos, é possível também especificar sequências ou conjuntos de valores ou, até mesmo, expressões de consultas aninhadas, isto é, consultas dentro de consultas (ou subconsultas).

Os símbolos principais de controle para estruturar as consultas são:

- `| |` ou `OR` como operador de disjunção;
- `<q>` e `</q>` como delimitador de subconsultas;
- `+` como condição vazia que corresponde a todos os valores;
- `<`, `<<`, `>`, `>>`, `!` para expressar os operadores de comparação  $\leq$ ,  $<$ ,  $\geq$ ,  $>$ ,  $\neq$ ;
- `~` (*like*) e `!~` (*not like*) para comparação de *strings*<sup>26</sup>.

---

<sup>26</sup> Para mais informações acerca da expressividade da linguagem *Ask*, veja: BAO, Jie; DING, Li; HENDLER, James. *Knowledge Representation and Query in Semantic MediaWiki: A Formal Study*, 2009. Disponível em: <http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.187.4263&rep=rep1&type=pdf>. Acesso em: 10 nov. 2013.

Para exemplificar a sintaxe de uma consulta, o seguinte trecho ilustra uma requisição por todas as cidades que estão localizadas em países membros da União Europeia ou que tenham mais de 500.000 habitantes:

```
[[Categoria:Cidade]] <q> [[localizada em:: <q> [[Categoria:País]]
[[membro da::União Europeia]] </q> ]] || [[população::>500.000]]
</q>
```

Segundo Krötzsch et al. (2007), a semântica formal de tais consultas é dada pelo mapeamento entre a sintaxe wiki e expressões em OWL DL, isto é, a consulta recupera todos os membros inferidos de acordo com uma classe OWL (por exemplo, se a consulta busca por “pessoas” e “João” é uma pessoa, logo, “João” será recuperado).

As consultas no *Semantic MediaWiki*, como introduzidas acima, meramente definem um conjunto de páginas como resultado. Para recuperar mais informações sobre esses resultados, o *software* permite que sejam realizadas as chamadas declarações de exibição (*printout statements*) como partes das consultas. Por exemplo, escrevendo `?população` em uma consulta por cidades, fará com que a cada cidade listada, seja exibido o número de habitantes (valores da propriedade *população*) correspondente. A Figura 23 demonstra essa utilização:

Figura 23 – O resultado de uma consulta semântica e a sua respectiva requisição.



The screenshot shows a Semantic MediaWiki page titled "Editando Cidades por habitantes". The page is in edit mode, showing a table of city data and a SPARQL query in the edit mode.

**Table Data:**

Cidade	População	Unidade Federativa
São Paulo	11.376.685	São Paulo
Rio de Janeiro	6.323.037	Rio de Janeiro
Brasília	2.609.997	Distrito Federal
Florianópolis	538.084	Santa Catarina

**SPARQL Query:**

```

{{#ask: [[categoria:cidade]]
|?População
|?Unidade Federativa
|format=table
|mainlabel=Cidade
|sort=população
|order=descending
}}

```

Fonte: Elaborada pelo autor.

Ao utilizar mais parâmetros em uma requisição de consulta, o resultado pode ser controlado para uma grande escala de formatos de saída e/ou de exibição<sup>27</sup> (listas ordenadas, listas não ordenadas, tabelas, mapas, calendários, linhas de tempo, cronogramas, gráficos, galerias, nuvens de *tags*, *slides*, funções matemáticas, XML, RDF, JSON, CSV, DSV, RSS, Atom, etc.).

Formulários, como descritos na seção 4.1.2, também podem ser utilizados no auxílio a consultas nos wikis semânticos do tipo *Semantic MediaWiki*. Para exemplificar esse uso, a Figura 24 ilustra um formulário utilizado em uma consulta por publicações cadastradas em um wiki semântico.

<sup>27</sup> Para mais informações sobre os formatos de resultado disponíveis, veja a extensão *Semantic Result Formats*: [http://www.semantic-mediawiki.org/wiki/Semantic\\_Result\\_Formats](http://www.semantic-mediawiki.org/wiki/Semantic_Result_Formats).

Figura 24 – Página de busca que se utiliza de um formulário.

The screenshot shows the Wikicat search interface. The browser address bar displays 'wikicat.org/wiki/Especial:RunQuery/AdvancedSearch'. The page title is 'Pesquisar'. The search form includes a title dropdown set to 'contém ao menos uma palavra', a search input field containing 'Wiki Wikipedia', and filters for 'Ano de publicação' (set to '>2001' and '<<2010') and 'Idioma' (set to 'English'). A 'Pesquisar' button is present. To the right, a help box titled 'Ajuda: ano de publicação' provides instructions on date search syntax. Below the form, a list of search results is displayed, including titles like 'Blogs, Wikipedia, Second Life, and beyond: from production to produsage' and 'How Wikipedia works: and how you can be a part of it'.

**Pesquisar**

Título ▼ contém ao menos uma palavra ▼

Wiki Wikipedia

Ano de publicação: >2001 <<2010

Idioma: English ▼

Pesquisar

Pesquisa simples | Lista de assuntos | Lista de autores | Lista de livros

**Ajuda: ano de publicação**

- Busca por uma **única data**: insira a data desejada no primeiro campo.  
Ex. 2001
- Busca por **mais de uma data**: insira as datas desejadas no primeiro campo separando-as por "ou".  
Ex. 2001 ou 2005 ou 2010
- Busca por um **intervalo de datas**: insira a data inicial no primeiro campo, acrescente antes dela > (maior ou igual) ou >> (maior). Insira a data final no segundo campo, acrescente antes dela < (menor ou igual) ou << (menor).  
Ex. >2001 e <<2010 (busca por data de publicação maior ou igual a 2001 e menor que 2010).

- Blogs, Wikipedia, Second Life, and beyond: from production to produsage**  
Axel Bruns  
New York, Washington, D.C./Baltimore, Bern Frankfurt am Main, Berlin, Brussels, Vienna, Oxford: Peter Lang, 2009.
- How Wikipedia works: and how you can be a part of it**  
Phoebe Ayers, Charles Matthews, Ben Yates  
San Francisco: No Starch Press, 2008.
- MediaWiki**  
Daniel J. Barrett  
Sebastapol, Calif.: O'Reilly Media, 2009.
- The Wiki way: quick collaboration on the Web**  
Bo Leuf, Ward Cunningham  
New York: Addison-Wesley, 2001.
- The Wikipedia revolution: how a bunch of nobodies created the world's greatest encyclopedia**  
Andrew Lih; prefácio de Jimmy Wales  
New York: Hyperion, 2009.
- Wikis: tools for information work and collaboration**  
Jane Klobas  
Oxford: Chandos, 2006.

Fonte: Elaborada pelo autor.

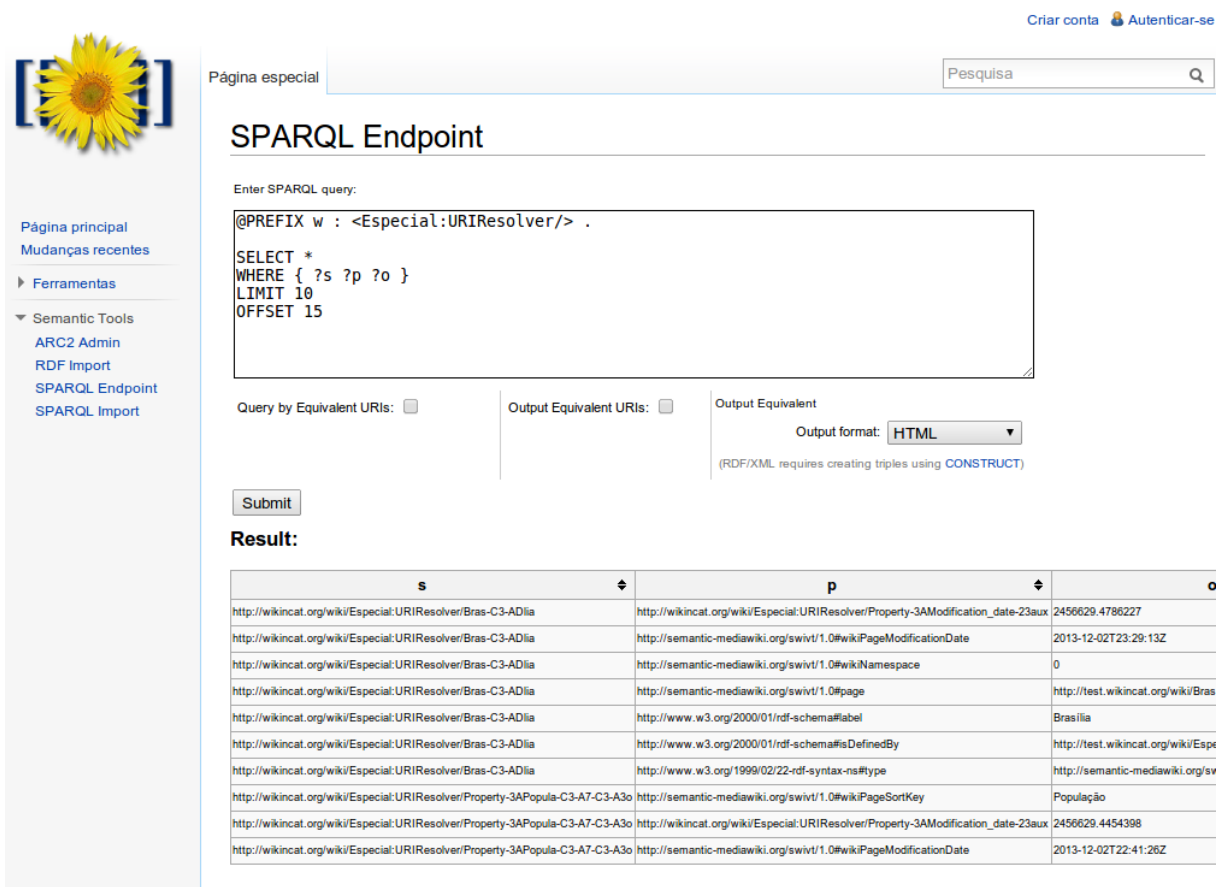
Embora não disponível na instalação padrão do Semantic MediaWiki, existem extensões<sup>28</sup> que permitem realizar consultas SPARQL nos dados registrados no wiki. A Figura 25 ilustra essa opção.

<sup>28</sup> Para a realização de consultas SPARQL em wikis, veja as extensões:

<http://www.mediawiki.org/wiki/Extension:RDFIO>,  
<http://www.mediawiki.org/wiki/Extension:LinkedWiki>,  
<http://www.mediawiki.org/wiki/Extension:SparqlExtension> e  
<http://www.mediawiki.org/wiki/Extension:Spark>.



Figura 25 – Uma consulta SPARQL por todas as triplas (sujeito, predicado, objeto) presentes no wiki, começando pelo 15º resultado e limitando-se a 10 resultados.



The screenshot shows the SPARQL Endpoint interface. On the left is a navigation menu with links like 'Página principal', 'Mudanças recentes', 'Ferramentas', and 'Semantic Tools'. The main area is titled 'SPARQL Endpoint' and contains a text input for the query. The query entered is:
 

```
@PREFIX w : <Especial:URIResolver/> .
SELECT *
WHERE { ?s ?p ?o }
LIMIT 10
OFFSET 15
```

 Below the query input are checkboxes for 'Query by Equivalent URIs' and 'Output Equivalent URIs', and a dropdown for 'Output format' set to 'HTML'. A 'Submit' button is at the bottom left. The results are displayed in a table with three columns: 's', 'p', and 'o'.
 

s	p	o
http://wikincat.org/wiki/Especial:URIResolver/Bras-C3-ADlia	http://wikincat.org/wiki/Especial:URIResolver/Property-3AModification_date-23aux	2456629.4786227
http://wikincat.org/wiki/Especial:URIResolver/Bras-C3-ADlia	http://semantic-mediawiki.org/swivt/1.0#wikiPageModificationDate	2013-12-02T23:29:13Z
http://wikincat.org/wiki/Especial:URIResolver/Bras-C3-ADlia	http://semantic-mediawiki.org/swivt/1.0#wikiNamespace	0
http://wikincat.org/wiki/Especial:URIResolver/Bras-C3-ADlia	http://semantic-mediawiki.org/swivt/1.0#page	http://test.wikincat.org/wiki/Bras
http://wikincat.org/wiki/Especial:URIResolver/Bras-C3-ADlia	http://www.w3.org/2000/01/rdf-schema#label	Brasilia
http://wikincat.org/wiki/Especial:URIResolver/Bras-C3-ADlia	http://www.w3.org/2000/01/rdf-schema#isDefinedBy	http://test.wikincat.org/wiki/Espe
http://wikincat.org/wiki/Especial:URIResolver/Bras-C3-ADlia	http://www.w3.org/1999/02/22-rdf-syntax-ns#type	http://semantic-mediawiki.org/sv
http://wikincat.org/wiki/Especial:URIResolver/Property-3APopula-C3-A7-C3-A3o	http://semantic-mediawiki.org/swivt/1.0#wikiPageSortKey	População
http://wikincat.org/wiki/Especial:URIResolver/Property-3APopula-C3-A7-C3-A3o	http://wikincat.org/wiki/Especial:URIResolver/Property-3AModification_date-23aux	2456629.4454398
http://wikincat.org/wiki/Especial:URIResolver/Property-3APopula-C3-A7-C3-A3o	http://semantic-mediawiki.org/swivt/1.0#wikiPageModificationDate	2013-12-02T22:41:26Z

Fonte: Elaborada pelo autor.

Nesse contexto, percebe-se que são várias as formas de navegação pelos dados e de recuperação desses dados no wiki, seja por uma navegação exploratória, por uma consulta já predefinida e embutida em uma página ou por uma consulta construída no momento requerido utilizando a linguagem Ask ou a linguagem SPARQL.

#### 4.3 O intercâmbio de informações em wikis semânticos

A Web Semântica requer a troca e a reutilização de informações, facilitadas pelos formatos padrão que possibilitam o intercâmbio de informações estruturadas entre produtores e consumidores dessas informações. Os wikis semânticos, por serem produtores e, às vezes, consumidores dessas informações estruturadas, não ficam de fora desse aspecto. Os dados do *Semantic MediaWiki*, por exemplo, são fornecidos na codificação OWL/RDF e utilizam URIs apropriados como identificadores, prevenindo a confusão com URLs das páginas HTML do wiki (KRÖTZSCH et al., 2007). O documento OWL/RDF gerado

pelo *software* é navegável e, nesse sentido, seus URIs são utilizados para localizar mais recursos: todos os URIs apontam para um *web service* do wiki que utiliza a negociação de conteúdo (*content negotiation*) para direcionar o requerente para o serviço de exportação OWL/RDF (caso o requerente seja um agente não-humano buscando por mais informações) ou para a página wiki correspondente (caso o requerente seja um agente humano). A negociação de conteúdo juntamente a exportação em OWL/RDF permite o máximo reuso dos dados presentes no wiki.

A Figura 26 exemplifica a exportação em OWL/RDF de uma página sobre a cidade do Rio de Janeiro, contendo apenas duas propriedades: População e Unidade Federativa.

Figura 26 – Exemplo de exportação de dados em OWL/RDF a partir do *Semantic MediaWiki*.

---

```
<?xml version="1.0" encoding="UTF-8"?>
<rdf:RDF xmlns:rdf="http://www.w3.org/1999/02/22-rdf-syntax-ns#"
  xmlns:rdfs="http://www.w3.org/2000/01/rdf-schema#"
  xmlns:owl="http://www.w3.org/2002/07/owl#"
  xmlns:swikt="http://semantic-mediawiki.org/swikt/1.0#"
  xmlns:wiki="http://xyz.org/wiki/Special:URIResolver/"
  xmlns:property="http://xyz.org/wiki/Special:URIResolver/Property-3A">
  <owl:Ontology rdf:about="http://xyz.org/wiki/Special:ExportRDF/Rio de Janeiro">
    <swikt:creationDate rdf:datatype="http://www.w3.org/2001/XMLSchema#dateTime">2013-12-04T21:27:02-
02:00</swikt:creationDate>
    <owl:imports rdf:resource="http://semantic-mediawiki.org/swikt/1.0"/>
  </owl:Ontology>
  <swikt:Subject rdf:about="http://xyz.org/wiki/Special:URIResolver/Rio de Janeiro">
    <rdf:type rdf:resource="http://xyz.org/wiki/Special:URIResolver/Categoria-3ACidade"/>
    <rdfs:label>Rio de Janeiro</rdfs:label>
    <swikt:page rdf:resource="http://xyz.org/wiki/Rio de Janeiro"/>
    <rdfs:isDefinedBy rdf:resource="http://xyz.org/wiki/Special:ExportRDF/Rio de Janeiro"/>
    <property:Populacao rdf:datatype="http://www.w3.org/2001/XMLSchema#double">6323037</property:Populacao>
    <property:Unidade_Federativa rdf:datatype="http://www.w3.org/2001/XMLSchema#string">Rio de
Janeiro</property:Unidade_Federativa>
    <swikt:wikiPageModificationDate rdf:datatype="http://www.w3.org/2001/XMLSchema#dateTime">2013-12-
02T23:28:53Z</swikt:wikiPageModificationDate>
    <swikt:wikiPageSortKey rdf:datatype="http://www.w3.org/2001/XMLSchema#string">Rio de
Janeiro</swikt:wikiPageSortKey>
  </swikt:Subject>
  <owl:DatatypeProperty rdf:about="http://semantic-mediawiki.org/swikt/1.0#creationDate"/>
  <owl:Class rdf:about="http://xyz.org/wiki/Special:URIResolver/Categoria-3ACidade"/>
  <owl:ObjectProperty rdf:about="http://semantic-mediawiki.org/swikt/1.0#page"/>
  <owl:DatatypeProperty rdf:about="http://semantic-mediawiki.org/swikt/1.0#wikiNamespace"/>
  <owl:DatatypeProperty rdf:about="http://xyz.org/wiki/Special:URIResolver/Property-3APopulacao"/>
  <owl:DatatypeProperty rdf:about="http://xyz.org/wiki/Special:URIResolver/Property-3AUnidade_Federativa"/>
  <owl:DatatypeProperty rdf:about="http://semantic-mediawiki.org/swikt/1.0#wikiPageModificationDate"/>
  <owl:DatatypeProperty rdf:about="http://semantic-mediawiki.org/swikt/1.0#wikiPageSortKey"/>
</rdf:RDF>
```

---

Fonte: Elaborada pelo autor.

Além de fornecer a exportação de uma ou mais páginas do wiki em OWL/RDF, o *Semantic MediaWiki* fornece scripts que geram a exportação completa de todos os dados do

wiki, o que é útil para diversas tarefas como operações *off-line* com os dados (análise/estudo, tratamento dos dados, conversões de formatos, etc.), disponibilização dos dados em repositórios de dados abertos como o CKAN, envio para serviços de indexação de conteúdo semântico, *backup* dos dados, entre outras operações.

Outro meio de exportação e/ou de utilização externa dos dados presentes no *Semantic MediaWiki* é pelo uso da API (*Application Programming Interface*) do *MediaWiki*. Com a API, é possível consultar dados com a linguagem *Ask*, por exemplo, e receber esses dados em XML ou JSON. A Figura 27 ilustra um documento XML obtido como resultado de uma consulta feita com a linguagem *Ask* na API do *MediaWiki*.

Figura 27 – Documento XML obtido pela API do *MediaWiki*.

---

```
<?xml version="1.0" encoding="UTF-8"?>
<api query-continue-offset="3">
  <query serializer="SMW\Serializers\QueryResultSerializer" version="0.5">
    <printrequests>
      <printrequest label="" typeid="_wpg" mode="2"/>
      <printrequest label="Has title proper" typeid="_txt" mode="1" format=""/>
      <printrequest label="Has place of publication" typeid="_txt" mode="1" format=""/>
    </printrequests>
    <results>
      <subject fulltext="11 de setembro (2002)" fullurl="http://wikincat.org/wiki/11_de_setembro_(2002)">
        <printouts>
          <property label="Has title proper">
            <value>11 de setembro</value>
          </property>
          <property label="Has place of publication">
            <value>Rio de Janeiro</value>
          </property>
        </printouts>
      </subject>
      <subject fulltext="Administração de bibliotecas (1975)"
fullurl="http://wikincat.org/wiki/Administra%C3%A7%C3%A3o_de_bibliotecas_(1975)">
        <printouts>
          <property label="Has title proper">
            <value>Administração de bibliotecas</value>
          </property>
          <property label="Has place of publication">
            <value>São Paulo</value>
          </property>
        </printouts>
      </subject>
      <subject fulltext="Admirável mundo novo (1980)"
fullurl="http://wikincat.org/wiki/Admir%C3%A1vel_mundo_novo_(1980)">
        <printouts>
          <property label="Has title proper">
            <value>Admirável mundo novo</value>
          </property>
          <property label="Has place of publication">
            <value>São Paulo</value>
          </property>
        </printouts>
      </subject>
    </results>
  </query>
</api>
```

---

Fonte: Elaborada pelo autor.

Este capítulo apresentou e analisou as características do *Semantic MediaWiki* nas atividades de representação, de recuperação e de intercâmbio de informações.

A seção 4.1 e suas subseções abordaram as atividades de estruturação, de organização e de descrição de informações por meio de páginas, *namespaces*, categorias (classes), redirecionamentos, páginas de desambiguação, predefinições, propriedades e tipos de dados em um wiki semântico.

A seção 4.2 abordou a recuperação de informações em wikis semânticos, seja por meio de navegação exploratória pelo wiki, seja por meio de consultas aos dados.

Finalmente, a seção 4.3 abordou o intercâmbio de informações em wikis semânticos com ênfase na questão da exportação dos dados em formatos e meios já consagrados.

A seguir, seguem as considerações finais desta pesquisa.

## 5 Considerações finais

Esta pesquisa, que surgiu a partir do questionamento “como se configuram os processos de representação, recuperação e intercâmbio de informações em wikis semânticos?”, teve como objetivos específicos:

- apresentar brevemente os conceitos de Web Semântica, de wiki e de wiki semântico (objetivo alcançado nas seções 2.3, 3.1 e 3.2);
- compreender os padrões de codificação, de representação, de relação e de consulta de dados estruturados relacionados à Web Semântica (URI, XML, RDF, RDFS, OWL e SPARQL) (alcançado nas seções 2.4.1 a 2.4.6);
- investigar as principais iniciativas que englobam as tecnologias wiki com as tecnologias da Web Semântica (*AceWiki*, *IkeWiki*, *OntoWiki*, *Semantic MediaWiki*, etc.) (alcançado nas seções 3.3 e 3.4); e
- descrever as principais contribuições das tecnologias da Web Semântica aplicadas aos wikis no que se refere à descrição, à recuperação e ao intercâmbio de informações (marcações semânticas em links, atribuições de metadados aos recursos, consultas complexas aos dados, navegação melhorada, exportação de dados, etc.) (alcançado no capítulo 4).

Sobre a realização desses objetivos, são apresentadas nessa seção as considerações finais desta pesquisa.

O estudo realizado sobre as raízes históricas da Web Semântica contextualizou essa temática possibilitando uma melhor compreensão acerca de suas definições e das principais tecnologias que fazem parte de sua construção. Tendo as definições e características básicas da Web Semântica em mente, pôde-se expandir o conhecimento geral acerca dessa importante iniciativa do W3C.

Uma vez que as tecnologias e os padrões da Web Semântica são condições indispensáveis para a consolidação de uma Web de dados, cada uma de suas tecnologias deve ser estudada cuidadosamente. No escopo desta pesquisa, foram destacados apenas os aspectos relevantes dessas tecnologias, necessários à compreensão dos wikis semânticos.

Por meio da revisão de literatura acerca do modelo para a descrição de recursos *Resource Description Framework* (RDF), sua origem e suas características, observa-se sua importância para a descrição de quaisquer recursos, sejam eles digitais ou não.

A partir dos apontamentos apresentados, considera-se que a área de Catalogação descritiva está intimamente relacionada a esse modelo, tido como um dos alicerces da Web Semântica.

Utilizando-se da *Extensible Markup Language* (XML) como sintaxe comum para o intercâmbio e o processamento de metadados, o RDF colabora para a interoperabilidade entre os vários sistemas de informação e de descrição existentes, contribuindo, desse modo, para a construção de mecanismos de busca mais integrados que permitirão a oferta de serviços mais especializados aos seus usuários.

A *RDF Vocabulary Description Language* (*RDF Schema* ou RDFS) especifica propriedades e classes que são utilizadas para nomear e descrever propriedades e classes que descreverão recursos, dando a ela um aspecto de linguagem de ontologias. Porém, aplicações mais sofisticadas requerem linguagens de representação mais expressivas. Esse é o caso da linguagem OWL que permite o compartilhamento de conhecimentos complexos acerca de um domínio de interesse na forma de ontologias processáveis por máquinas.

Tendo descrições construídas com base no modelo RDF, linguagens de consulta são necessárias para que os dados contidos nessas descrições possam ser utilizados. No contexto da Web Semântica, essa é a função da linguagem SPARQL, construída para consultar quaisquer dados modelados em RDF.

Assim, foram estudados e apresentados padrões de representação, codificação, descrição, relação e consulta de dados estruturados (URI, XML, RDF, RDFS, OWL e SPARQL) que, junto a outras tecnologias, formam a base da Web Semântica e apoiam o funcionamento dos wikis semânticos.

Wikis são sistemas que possibilitam a edição de páginas Web por um navegador, utilizam uma sintaxe própria para a edição de textos, possibilitam a edição colaborativa, possuem mecanismos de reversão de edições e de busca. Porém, apesar de todos os seus pontos fortes e positivos, os wikis possuem fraquezas em alguns pontos. Devido a sua natureza aberta e flexível, wikis geralmente crescem de modo muito rápido e, com a quantidade sempre crescente de usuários colaboradores, a criação de páginas em sites desse tipo é realizada, muitas vezes, de forma exponencial. Isso aumenta a base de dados do wiki trazendo consequências negativas, tais como dificuldades para a navegação, orientação e busca por conteúdos.

Para solucionar esses problemas, wikis semânticos foram propostos. Esses sistemas são wikis melhorados devido à utilização de tecnologias sofisticadas para o tratamento de informações. Essas tecnologias são as tecnologias da Web Semântica que permitem a representação, a codificação, a descrição, a relação e a consulta de dados estruturados dentro do wiki. Com isso, podem ser realizadas marcações semânticas em links, atribuições de metadados aos recursos e consultas complexas aos dados. Os resultados dessas consultas, por sua vez, podem ser embutidos em páginas wiki (possibilitando uma apresentação dinâmica dos conteúdos e uma navegação mais sofisticada no site) e exportados em OWL/RDF.

Porém, embora os esforços teóricos e práticos envidados pelos desenvolvedores de wikis semânticos com o objetivo de tornar os seus sistemas úteis e utilizáveis, a realidade evidencia que poucos são os sistemas que ganharam popularidade para serem utilizados de forma significativa na Web. Durante o estudo, observou-se que o principal sistema de wiki semântico utilizado é o *Semantic MediaWiki*, um sistema desenvolvido e utilizado por uma expressiva comunidade de desenvolvedores e de usuários<sup>29</sup>.

Visando a apresentar a aplicação das tecnologias da Web Semântica nos wikis semânticos, foram descritas as principais atividades de descrição, de recuperação e de intercâmbio de informações no *Semantic MediaWiki*. Nessa etapa da pesquisa, foi possível observar que as tecnologias da Web Semântica aplicadas aos wikis semânticos podem facilitar o reuso das informações neles contidas. No entanto, para que isso ocorra, as informações precisam estar formalmente representadas e descritas de modo que possam ser processadas por máquinas.

A partir do momento em que as informações estão formalmente representadas e descritas no wiki semântico, são vários os modos de navegação por essas informações e de recuperação destas. Existe a opção por realizar uma navegação exploratória, seguindo os links gerados pelo wiki, por uma consulta já predefinida e embutida em uma página wiki ou por uma consulta construída no momento requerido utilizando as linguagens *Ask* ou *SPARQL*.

---

<sup>29</sup> De acordo com o WikiApiary, que coleta informações sobre sites wiki que utilizam o software MediaWiki, o Semantic MediaWiki é utilizado em 1.284 sites wiki. Disponível: <[https://wikiapiary.com/wiki/Semantic\\_statistics](https://wikiapiary.com/wiki/Semantic_statistics)>. Acesso em: 02 ago. 2008.

No que se refere ao intercâmbio de dados, os dados presentes em um *Semantic MediaWiki*, por exemplo, são fornecidos na codificação OWL/RDF, além de vários outros formatos de exportação (XML, JSON, CSV, DSV, RSS, etc.) e em uma API, que pode ser utilizada tanto para uso interno quanto para uso externo ao wiki.

Apesar de haver uma vasta literatura científica, inclusive nacional, sobre as tecnologias da Web Semântica na área da Ciência da Informação, é notável a falta da aplicação dessas tecnologias em projetos nacionais existentes. No campo da Biblioteconomia, por exemplo, até mesmo a XML, que possui mais de 15 anos, parece não ser completamente compreendida e utilizada nos padrões bibliográficos usados nacionalmente, haja vista a utilização do Formato MARC 21 ainda com a codificação da ISO 2709 e de sistemas de catalogação do tipo CDS ISIS, padrões esses originados na década de 1970. Desse modo, uma vez que os sistemas de bibliotecas, por exemplo, não fazem uso de tecnologias já consagradas internacionalmente por outras áreas, faz-se necessária, pelo menos, uma aproximação informal entre os profissionais (e futuros profissionais) da Ciência da Informação e as tecnologias da Web Semântica que já estão sendo incorporadas em software não especializado, como sistemas de gerenciamento de conteúdo, wikis, blogs, redes sociais, etc. É nesse sentido que os wikis semânticos, por trazerem conceitos da Web Semântica, facilitam o aprendizado e o uso dessas tecnologias ainda pouco utilizadas nacionalmente na Ciência da Informação para a representação, a descrição, a recuperação e o intercâmbio de informações.

Portanto, tendo apresentado os aspectos de representação, de descrição, de recuperação e de intercâmbio de informações em wikis semânticos, espera-se ter contribuído para o entendimento, a discussão e, porque não, para o uso de tecnologias da Web Semântica em ambientes informacionais digitais gerenciados e propostos, de preferência, por profissionais da informação.

Considera-se que, na utilização de wikis semânticos os gestores desses wikis são incentivados constantemente a refletir sobre a melhor maneira de

- a) representar as entidades a serem descritas;
- b) descrever as entidades representadas;
- c) utilizar ou não vocabulários e/ou ontologias externas existentes para a modelagem e a descrição dos dados;
- d) apresentar os dados armazenados aos agentes humanos e não humanos;



- e) seguir padrões internacionais de descrição e de intercâmbio de dados; e
- f) disseminar os dados a todos os interessados.

Com essas considerações acerca da Web Semântica, das tecnologias a ela relacionadas e dos wikis semânticos, entende-se que o objetivo geral proposto foi alcançado: explorar, apresentar e descrever as características dos wikis semânticos com enfoque nas atividades de representação, de recuperação e de intercâmbio de informações apoiadas em tecnologias da Web Semântica para, assim, favorecer o entendimento, a discussão e o uso dessas tecnologias em ambientes informacionais digitais.

Devido à constatação da falta de projetos reais em Ciência da Informação que se utilizam de wikis semânticos no dia a dia de um sistema de informação, aponta-se como sugestões para estudos futuros, pesquisas que versem sobre o planejamento, o desenvolvimento e o uso efetivo de tais wikis semânticos em ambientes informacionais, levando em consideração os objetivos do sistema de informação, a missão da instituição, o público a que se destina, os recursos (humanos, informacionais, financeiros) envolvidos, as necessidades de informação dos usuários reais e potenciais, a integração e/ou interoperabilidade com outros sistemas, e tantos outros aspectos que envolvem a implantação e a manutenção de tais sistemas em curto, médio e longo prazos.

## Referências

- AL-HUSAIN, Luluah; EL-MASRI, Samir. An automatic approach for generating tables in semantic wikis. *Journal of Theoretical and Applied Information Technology*, v. 40, n. 1, p. 89-97, Jun. 2012. Disponível em: <<http://www.jatit.org/volumes/Vol40No1/13Vol40No1.pdf>>. Acesso em: 10 jan. 2013.
- ALMEIDA, Maurício Barcellos. *Inter-operabilidade entre fontes heterogêneas: um meta-modelo baseado em ontologias*. Belo Horizonte, 2002. 146 f. Dissertação (Mestrado em Ciência da Informação). Escola de Ciência da Informação – Universidade Federal de Minas Gerais, 2002a. Disponível em: <<http://hdl.handle.net/1843/LHLS-6AZQHU>>. Acesso em: 10 ago. 2012.
- ALMEIDA, Maurício Barcellos. Uma introdução ao XML, sua utilização na Internet e alguns conceitos complementares. *Ciência da Informação*, Brasília, v. 31, n. 2, p. 5-13, maio/ago. 2002b. Disponível em: <<http://www.scielo.br/pdf/ci/v31n2/12903>>. Acesso em: 13 ago. 2012.
- ALVES, Rachel Cristina Vesú et al. Ciência da Informação, Ciência da Computação e Recuperação da Informação: algumas considerações sobre os métodos e tecnologias da informação utilizados ao longo do tempo. *Revista Eletrônica Informação e Cognição*, v. 6, n. 1, p. 28-40, 2007.
- ALVES, Rachel Cristina Vesú. *Web Semântica: uma análise focada no uso de metadados*. 2005. 180 f. Dissertação (Mestrado em Ciência da Informação) – Faculdade de Filosofia Ciências, Universidade Estadual Paulista, Marília, 2005. Disponível em: <[http://www.marilia.unesp.br/Home/Pos-Graduacao/CienciadaInformacao/Dissertacoes/alves\\_rcv\\_me\\_mar.pdf](http://www.marilia.unesp.br/Home/Pos-Graduacao/CienciadaInformacao/Dissertacoes/alves_rcv_me_mar.pdf)>. Acesso em: 28 ago. 2012.
- AQUINO, Maria Clara. A folksonomia como hipertexto potencializador de memória coletiva: um estudo dos links e das tags no del.icio.us e no Flickr. *Liinc em Revista*, v. 4, n. 2, 2008. Disponível em: <<http://revista.ibict.br/liinc/index.php/liinc/article/viewFile/263/175>>. Acesso em: 15 jun. 2011.
- ASSUMPÇÃO, Fabrício Silva. *Conversão de registros em XML para MARC 21: um modelo baseado em folhas de estilo XSLT*. 2013. Dissertação (Mestrado em Ciência da Informação) – Faculdade de Filosofia Ciências, Universidade Estadual Paulista, Marília, 2013. Disponível em: <[http://www.marilia.unesp.br/Home/Pos-Graduacao/CienciadaInformacao/Dissertacoes/assumpcao\\_fs\\_me\\_mar.pdf](http://www.marilia.unesp.br/Home/Pos-Graduacao/CienciadaInformacao/Dissertacoes/assumpcao_fs_me_mar.pdf)>. Acesso em: 22 fev. 2014.
- AUER, Sören; DIETZOLD, Sebastian Tramp; RIECHERT, Thomas. OntoWiki - A Tool for Social, Semantic Collaboration. In: THE SEMANTIC WEB - ISWC 2006, INTERNATIONAL SEMANTIC WEB CONFERENCE, ISWC 2006, 5th, Athens, GA, USA, *Proceedings...*, 2006. Disponível em: <[http://dx.doi.org/10.1007%2F11926078\\_53](http://dx.doi.org/10.1007%2F11926078_53)>. Acesso em: 23 jan. 2013.
- BAO, Jie; DING, Li; HENDLER, James. Knowledge Representation and Query in Semantic MediaWiki: A Formal Study. In: TETHERLESS WORLD CONSTELLATION TECH REPORT, *Proceedings...*, 2008. Disponível em: <<http://citeseerx.ist.psu.edu/viewdoc/summary?doi=10.1.1.187.4263>>. Acesso em: 15 dez. 2012.
- BAX, Marcello Peixoto. Introdução às linguagens de marcas. *Ciência da Informação*, Brasília, v. 30, n. 1, p. 32-38, jan./abr. 2001. Disponível em: <<http://www.scielo.br/pdf/ci/v30n1/a05v30n1.pdf>>. Acesso em: 21 nov. 2012.
- BECKETT, Dave; MILLER, Eric; BRICKLEY, Dan. *Expressing Simple Dublin Core in RDF/XML*. Ohio: Dublin Core Metadata Initiative, 2002. Disponível em: <<http://www.dublincore.org/documents/dcmes-xml>>. Acesso em: 1 out. 2012.
- BERNERS-LEE, Tim. *A Short History of "Resource" in web architecture*. [S.l.: s.n.], 2009. Disponível em: <<http://www.w3.org/DesignIssues/TermResource.html>>. Acesso em: 02 fev. 2013.

BERNERS-LEE, Tim. *DeveloperWorks Interviews: Tim Berners-Lee*. Entrevistador: Scott Laningham, 2006. Disponível em: <<http://www.ibm.com/developerworks/podcast/dwi/cm-int082206txt.html>>. Acesso em: 28 ago. 2012.

BERNERS-LEE, Tim. *Information management: a proposal*. Genebra: CERN, 1989. Disponível em: <<http://www.w3.org/History/1989/proposal.html>>. Acesso em: 9 jul. 2012.

BERNERS-LEE, Tim. *Semantic Web Road map*. [S.l.: s.n.], 1998. Disponível em: <<http://www.w3.org/DesignIssues/Semantic.html>>. Acesso em: 20 ago. 2012.

BERNERS-LEE, Tim. *The Semantic Toolbox: Building Semantics on top of XML-RDF*. [S.l.: s.n.], 1999. Disponível em: <<http://www.w3.org/DesignIssues/Toolbox.html>>. Acesso em: 20 ago. 2012.

BERNERS-LEE, Tim. *Universal Resource Identifiers -- Axioms of Web Architecture*. [S.l.: s.n.], 1996. Disponível em: <<http://www.w3.org/DesignIssues/Axioms.html>>. Acesso em: 20 ago. 2012.

BERNERS-LEE, Tim; FISCHETTI, Mark. *Weaving the Web: the original design and ultimate destiny of the World Wide Web by its inventor*. New York: Harper Collins, 2000.

BERNERS-LEE, Tim; HENDLER, James; LASSILA, Ora. The semantic Web: a new form of Web content that is meaningful to computers will unleash a revolution of new possibilities. *Scientific American*, London, Nature, May 2001. Disponível em: <<http://www.med.nyu.edu/research/pdf/mainim01-1484312.pdf>>. Acesso em: 28 ago. 2010.

BORKO, Harold. Information Science: what is it? *American Documentation*, Washington, D.C., v. 19, n. 1, p. 3-5, 1968. Disponível em: <<http://dx.doi.org/10.1002/asi.5090190103>>. Acesso em: 4 mar. 2013.

BOULOS, Maged N. Kamel. Semantic Wikis: A Comprehensible Introduction with Examples from the Health Sciences. *Journal of Emerging Technologies in Web Intelligence*, v. 1, n. 1, Aug. 2009. Disponível em: <<http://ojs.academypublisher.com/index.php/jetwi/article/download/01019496/11>>. Acesso em: 1 fev. 2013.

BRATT, Steve. *Semantic Web, and Other Technologies to Watch*. Cambridge: W3C, 2007. Disponível em: <<http://www.w3.org/2007/Talks/0130-sb-W3CTechSemWeb>>. Acesso em: 24 set. 2012.

BRESLIN, John G.; PASSANT, Alexandre; DECKER, Stefan. *The social semantic Web*. New York: Springer, 2009.

BRY, François. Semantic Wikis: Approaches, Applications, and Perspectives. In: REASONING WEB SUMMER SCHOOL, 8th, *Proceedings...* Vienna, 2012. Disponível em: <Erro! A referência de hiperlink não é válida.>. Acesso em: 30 jan. 2013.

BRYAN, Martin. An Introduction to the Extensible Markup Language (XML). *Bulletin of the American Society for Information Science and Technology*, v. 25, n. 1, p. 11-14, Oct./Nov. 1998. Disponível em: <<http://dx.doi.org/10.1002/bult.104>>. Acesso em: 25 jun. 2012.

BUFFA, Michel et al. SweetWiki: A semantic wiki. *Journal of Web Semantics: Science, Services and Agents on the World Wide Web*, [S.l.], Elsevier, v. 6, n. 1, p. 84-97, 2007. Disponível em: <<http://dx.doi.org/10.1016/j.websem.2007.11.003>>. Acesso em: 25 jun. 2011.

BUFFA, Michel et al. SweetWiki: Semantic Web Enabled Technologies in Wiki. In: WORKSHOP ON SEMANTIC WIKIS, 1st, 2006, Budva. *Proceedings of the First Workshop on Semantic Wikis: from Wiki to Semantics*. Budva: Karlsruhe Institut für Technologie, 2006. p. 74-88. Disponível em: <<http://hnsp.inf-bb.uni-jena.de/proceedings/eswc2006/proceedings-workshops/eswc2006-workshop-semantic-wikis.pdf>>. Acesso em: 25 jun. 2011.

BURKE, Peter. *Uma história social do conhecimento: de Gutenberg a Diderot*. Rio de Janeiro: Jorge Zahar, 2003.

BUSH, Vannevar. *As we may think*. The Atlantic Monthly, Boston, v. 176, n. 1, p. 101-108, 1945. Disponível em: <<http://www.theatlantic.com/doc/194507/bush>>. Acesso em: 18 nov. 2012

CARVALHO, Rodrigo Aquino de. *Perspectivas na Web Semântica para a Ciência da Informação*. Campinas, 2009, 186 f. Dissertação (Mestrado em Ciência da Informação). Centro de Ciências Humanas e Sociais Aplicadas, Programa de Pós-Graduação em Ciência da Informação – Pontifícia Universidade Católica de Campinas, 2009. Disponível em: <[http://rabci.org/rabci/sites/default/files/Dissertacao\\_Mestrado\\_-\\_Rodrigo\\_A.\\_de\\_Carvalho\\_-\\_versao\\_final.pdf](http://rabci.org/rabci/sites/default/files/Dissertacao_Mestrado_-_Rodrigo_A._de_Carvalho_-_versao_final.pdf)>. Acesso em: 15 abr. 2012.

CASTELLS, Manuel. *A sociedade em rede*. 10. ed. São Paulo: Paz e Terra, 2007.

CASTRO, Fabiano Ferreira de. *Elementos de interoperabilidade na catalogação descritiva: configurações contemporâneas para a modelagem de ambientes informacionais digitais*. 202 f. Tese (Doutorado em Ciência da Informação) - Faculdade de Filosofia e Ciências, Universidade Estadual Paulista, Marília, 2012. Disponível em: <[http://www.marilia.unesp.br/Home/Pos-Graduacao/CienciadaInformacao/Dissertacoes/Castro,%20F.F.\\_doutorado\\_CI\\_2012.pdf](http://www.marilia.unesp.br/Home/Pos-Graduacao/CienciadaInformacao/Dissertacoes/Castro,%20F.F._doutorado_CI_2012.pdf)>. Acesso em: 15 dez. 2012.

CASTRO, Fabiano Ferreira de. *Padrões de representação e descrição de recursos informacionais em bibliotecas digitais na perspectiva da ciência da informação: uma abordagem do MarcOnt initiative na era da web semântica*. 2008. 201 f. Dissertação (Mestrado em Ciência da Informação) - Faculdade de Filosofia e Ciências, Universidade Estadual Paulista, Marília, 2008. Disponível em: <[http://www.marilia.unesp.br/Home/Pos-Graduacao/CienciadaInformacao/Dissertacoes/castro\\_ff\\_me\\_mar.pdf](http://www.marilia.unesp.br/Home/Pos-Graduacao/CienciadaInformacao/Dissertacoes/castro_ff_me_mar.pdf)>. Acesso em: 15 set. 2012.

CATARINO, Maria Elisabete; SOUZA, Terezinha Batista de. A representação descritiva no contexto da web semântica. *TransInformação*, Campinas, v. 24, n. 2, p. 77-90, maio/ago. 2012. Disponível em: <<http://revistas.puc-campinas.edu.br/transinfo/viewarticle.php?id=472>>. Acesso em: 2 out. 2012.

DI IORIO et al. Automatic Deployment of Semantic Wikis: a Prototype. In: WORKSHOP ON SEMANTIC WIKIS, 1st, 2006, Budva. *Proceedings of the First Workshop on Semantic Wikis: from Wiki to Semantics*. Budva: Karlsruhe Institut für Technologie, 2006. p. 212-221. Disponível em: <<http://hnsp.inf-bb.uni-jena.de/proceedings/eswc2006/proceedings-workshops/eswc2006-workshop-semantic-wikis.pdf>>. Acesso em: 25 jun. 2012.

DINGLEY, Andy. Theory: "opportunistic hypertext". Usenet, 2001. Disponível em: <<https://groups.google.com/forum/?hl=en#!msg/comp.infosystems.www.authoring.site-design/KqBkAgpFLs/yvrVEClbRhYJ>>. Acesso em: 12 fev. 2014.

DUNSIRE, Gordon; HILLMANN, Diane; PHIPPS, Jon. Reconsidering Universal Bibliographic Control in Light of the Semantic Web. *Journal of Library Metadata*, v. 12, n. 2-3, p. 164-176, 2012. Disponível em: <<http://dx.doi.org/10.1080/19386389.2012.699831>>. Acesso em: 6 out. 2012.

DUNSIRE, Gordon; WILLER, Mirna. Standard library metadata models and structures for the Semantic Web. *Library Hi Tech News*, v. 28, n. 3, p. 1-12, Apr. 2011. Disponível em: <<http://dx.doi.org/10.1108/07419051111145118>>. Acesso em: 6 out. 2012.

DZIEKANIAK, Gisele Vasconcelos; KIRINUS, Josiane Boeira. Web Semântica. *Encontros Bibli*, Florianópolis, n. 18, 2º sem. 2004. Disponível em: <<http://www.periodicos.ufsc.br/index.php/eb/article/view/1518-2924.2004v9n18p20>>. Acesso em: 1 out. 2012.

EITO BRUN, Ricardo. *Lenguajes de marcas para la gestión de recursos digitales: aproximación técnica, especificaciones y referencia*. Gijón: Trea, 2008.

FERNEDA, Edberto. *Introdução aos modelos computacionais de recuperação de informação*. Rio de Janeiro: Ciência Moderna, 2012.

FERREIRA, Jaider Andrade. *A Wikipedia, a inteligência coletiva e o conceito Web 2.0*. 2011. 102 f. Trabalho de conclusão de curso (Graduação em Biblioteconomia) - Faculdade de Filosofia e Ciências, Universidade Estadual Paulista, Marília, 2011.

FLAMINO, Adriana Nascimento. *MARXML: um padrão de descrição para recursos informacionais em Open Archives*. 2006. 164 f. Dissertação (Mestrado em Ciência da Informação) - Faculdade de Filosofia e Ciências, Universidade Estadual Paulista - UNESP, Marília. 2006. Disponível em: <[http://www.marilia.unesp.br/Home/Pos-Graduacao/CienciadaInformacao/Dissertacoes/flamino\\_an\\_me\\_mar.pdf](http://www.marilia.unesp.br/Home/Pos-Graduacao/CienciadaInformacao/Dissertacoes/flamino_an_me_mar.pdf)>. Acesso em: 20 set. 2012.

FUMERO, Antônio; ROCA, Genís. *Web 2.0*. Madrid: Fundación Orange, 2007. Disponível em: <[http://www.fundacionorange.es/areas/25\\_publicaciones/WEB\\_DEF\\_COMPLETO.pdf](http://www.fundacionorange.es/areas/25_publicaciones/WEB_DEF_COMPLETO.pdf)>. Acesso em: 26 ago. 2011.

FURGERI, Sérgio. O papel das linguagens de marcação para a Ciência da Informação. *TransInformação*, Campinas, v. 18, n. 3, p. 225-239, set./dez. 2006. Disponível em: <<http://revistas.puc-campinas.edu.br/transinfo/viewarticle.php?id=184>>. Acesso em: 1 out. 2012.

GAMA, Fernando Alves da. *As contribuições das linguagens de marcação para a gestão da informação arquivística digital*. 2011. 165 f. Dissertação (Mestrado em Ciência da Informação) – Faculdade de Filosofia e Ciências, Universidade Estadual Paulista, Marília, 2011. Disponível em: <[http://www.marilia.unesp.br/Home/Pos-Graduacao/CienciadaInformacao/Dissertacoes/gama\\_fa\\_me\\_mar.pdf](http://www.marilia.unesp.br/Home/Pos-Graduacao/CienciadaInformacao/Dissertacoes/gama_fa_me_mar.pdf)>. Acesso em: 30 jun. 2012.

GONÇALVES, José Jesse. *Um repositório de experiência docente integrando estrutura a priori e estrutura emergente a partir da abordagem de wiki semântico*. 2008. 173 f. Dissertação (Mestrado em Gestão do Conhecimento e Tecnologia da Informação) – Universidade Católica de Brasília, Brasília, 2008. Disponível em: <[http://www.bdt.d.uceb.br/tede/tde\\_arquivos/3/TDE-2008-11-27T072103Z-693/Publico/Texto%20completo%20Jesse.pdf](http://www.bdt.d.uceb.br/tede/tde_arquivos/3/TDE-2008-11-27T072103Z-693/Publico/Texto%20completo%20Jesse.pdf)>. Acesso em: 28 ago. 2012.

GRUBER, Thomas Robert. Ontology. In: LIU, Ling; ÖZSU, M. Tamer. *Encyclopedia of database systems*. [S.l.]: Springer-Verlag, 2009. Disponível em: <<http://tomgruber.org/writing/ontology-definition-2007.htm>>. Acesso em: 5 jun. 2012.

GRUBER, Thomas Robert. Toward principles for the design of ontologies used for knowledge sharing. In: INTERNATIONAL WORKSHOP ON FORMAL ONTOLOGY, 1993, Padova, Italy. *Technical Report KSL 93-04*. Stanford: Knowledge Systems Laboratory, Stanford University, 1993. Disponível em: <<http://tomgruber.org/writing/onto-design.pdf>>. Acesso em: 5 jun. 2012.

HITZLER, Pascal; KRÖTZSCH, Markus; RUDOLPH, Sebastian. *Foundations of Semantic Web technologies*. Boca Raton: CRC Press, 2010.

JACOBS, Ian; FORGUE, Marie-Claire; HIRAKAWA, Yasuyuki. *W3C Opens Data on the Web with SPARQL: Powerful Technology for Querying Distributed and Diverse Data*. Cambridge: W3C, 2008. Disponível em: <<http://www.w3.org/2007/12/sparql-pressrelease>>. Acesso em: 15 jan. 2013.

JORENTE, Maria José Vicentini; SANTOS, Plácida Leopoldina Ventura Amorim da Costa. Web social e intersemioses digitais: apontamentos para inovação e inclusão no compartilhamento de conhecimento. In: ENCONTRO NACIONAL DE PESQUISA EM CIÊNCIA DA INFORMAÇÃO, XI, 2010, Rio de Janeiro. *Anais...* Rio de Janeiro: UNIRIO, 2010. Disponível em: <<http://enancib.ibict.br/index.php/xi/enancibXI/paper/view/109/154>>. Acesso em: 5 ago. 2012.

- JORENTE, Maria José Vicentini; SANTOS, Plácida Leopoldina Ventura Amorim da Costa; VIDOTTI, Silvana Aparecida Borsetti Gregorio. Quando as Webs se encontram: social e semântica - promessa de uma visão realizada? *Informação & Informação*, Londrina, v. 14, n. esp, p. 1-24, 2009. Disponível em: <<http://dx.doi.org/10.5433/1981-8920.2009v14nesp.p1>>. Acesso em: 7 ago. 2012.
- KLOBAS, Jane Elizabeth. *Wikis: tools for information work and collaboration*. Oxford: Chandos, 2006. 229 p.
- KOIVUNEN, Marja-Riitta; MILLER, Eric. *W3C Semantic Web Activity*. Cambridge: W3C, 2001. Disponível em: <<http://www.w3.org/2001/12/semweb-fin/w3csw>>. Acesso em: 21 ago. 2012.
- KOREN, Yaron. *Working with MediaWiki*. New York: WikiWorks Press, 2012.
- KRÖTZSCH, Markus et al. Semantic Wikipedia. *Journal of Web Semantics: Science, Services and Agents on the World Wide Web*, v. 5, n. 4, p. 251-261, Dec. 2007. Disponível em: <<http://dx.doi.org/10.1016/j.websem.2007.09.001>>. Acesso em: 15 mar. 2013.
- KUHN, Tobias. AceWiki: A Natural and Expressive Semantic Wiki. In: SEMANTIC WEB USER INTERACTION AT CHI 2008: Exploring HCI Challenges, *Proceedings...*, 2008. Disponível em: <<http://attempto.ifi.uzh.ch/site/pubs/papers/kuhn08acewiki.pdf>>. Acesso em: 10 dez. 2012.
- LEÃO, Lucia. *O labirinto da hipermídia: arquitetura e navegação no ciberespaço*. São Paulo: FAPESP, Iluminuras. 2005.
- LE COADIC, Yves-François. *A ciência da informação*. 2. ed. Brasília: Briquet de Lemos/Livros, 2004.
- LEUF, Bo; CUNNINGHAM, Ward. *The Wiki way: quick collaboration on the Web*. Boston: Addison-Wesley, 2001.
- LÉVY, Pierre. *Cibercultura*. 2. ed. Rio de Janeiro: Ed. 34, 2000.
- LIBRARY OF CONGRESS. *A Bibliographic Framework for the Digital Age*. Washington, D.C., 2011. Disponível em: <<http://www.loc.gov/marc/transition/news/framework-103111.html>>. Acesso em: 20 set. 2012.
- LIBRARY OF CONGRESS. *Bibliographic Framework as a Web of Data: Linked Data Model and Supporting Services*. Washington, D.C., 2012. Disponível em: <[www.loc.gov/marc/transition/pdf/marclid-report-11-21-2012.pdf](http://www.loc.gov/marc/transition/pdf/marclid-report-11-21-2012.pdf)>. Acesso em: 22 jan. 2013.
- MILLARD, David E. et al. Semantics on demand: Can a Semantic Wiki replace a knowledge base? *New Review of Hypermedia and Multimedia*, v. 14, n. 1, p. 95-120, Jul. 2008. Disponível em: <<http://dx.doi.org/10.1080/13614560802316111>>. Acesso em: 20 set. 2012.
- MILLER, Dick R.; CLARKE, Kevin S. *Putting XML to work in the library: tools for improving access and management*. Chicago: American Library Association, 2004.
- MILLER, Eric. An Introduction to the Resource Description Framework. *D-Lib Magazine*, v. 4, n. 5, May 1998. Disponível em: <<http://www.dlib.org/dlib/may98/miller/05miller.html><http://www.w3.org/2004/Talks/0120-semweb-umich>>. Acesso em: 16 set. 2012.
- MILLER, Eric. *The semantic Web*. Cambridge: W3C, 2004. Disponível em: <<http://www.w3.org/2004/Talks/0120-semweb-umich>>. Acesso em: 16 set. 2012.
- MOREIRA, Alexandra; ALVARENGA, Lídia; OLIVEIRA, Alcione de Paiva. O nível do conhecimento e os instrumentos de representação: tesouros e ontologias. *DataGramaZero*, v. 5, n. 6, dez. 2004. Disponível em: <[http://www.dgz.org.br/dez04/Art\\_01.htm](http://www.dgz.org.br/dez04/Art_01.htm)>. Acesso em: 25 mar. 2012.

- MORESI, Eduardo (Org.). *Manual de metodologia da pesquisa*. Brasília: Universidade Católica de Brasília, 2003. Disponível em: <<http://www.inf.ufes.br/~falbo/files/MetodologiaPesquisa-Moresi2003.pdf>>. Acesso em: 23 jun. 2012.
- NAMESPACES. In: MediaWiki. Florida: Wikimedia Foundation, 2012. Disponível em: <<http://www.mediawiki.org/w/index.php?title=Namespaces&oldid=512843>>. Acesso em: 12 jun. 2013.
- NELSON, Theodor Holm. Complex information processing: a file structure for the complex, the changing and the indeterminate. In: ACM/CSC-ER NATIONAL CONFERENCE, 20., 1965. *Proceedings...* 1965. p. 84-100. Disponível em: <<http://dx.doi.org/10.1145/800197.806036>>. Acesso em: 23 jun. 2012.
- NILSSON, Mikael et al. *Expressing Dublin Core metadata using the Resource Description Framework (RDF)*. Ohio: Dublin Core Metadata Initiative, 2008. Disponível em: <<http://www.dublincore.org/documents/dc-rdf>>. Acesso em: 1 out. 2012.
- NISO. *Understanding Metadata*. Bethesda: NISO Press, 2004. Disponível em: <<http://www.niso.org/publications/press/UnderstandingMetadata.pdf>>. Acesso em: 23 set. 2012.
- O'REILLY, Tim. *Web 2.0 compact definition: trying again*. Sebastopol, C.A.: O'Reilly Media, 2006. Disponível em: <<http://radar.oreilly.com/2006/12/web-20-compact-definition-tryi.html>>. Acesso em: 02 set. 2010.
- O'REILLY, Tim. *What is web 2.0: design patterns and business models for the next generation of software*. Sebastopol, C.A.: O'Reilly Media, 2005. Disponível em: <<http://oreilly.com/web2/archive/what-is-web-20.html>>. Acesso em: 28 ago. 2010.
- OBITKO, Marek. *Semantic Web Architecture*. Prague: Czech Technical University, 2007. Disponível em: <<http://obitko.com/tutorials/ontologies-semantic-web/semantic-web-architecture.html>>. Acesso em: 21 ago. 2012.
- OREN, Eyal et al. Annotation and Navigation in Semantic Wikis. In: WORKSHOP ON SEMANTIC WIKIS, 1st, 2006a, Budva. *Proceedings of the First Workshop on Semantic Wikis: from Wiki to Semantics*. Budva: Karlsruher Institut für Technologie, 2006a. p. 16-30. Disponível em: <<http://hnsf.inf-bb.uni-jena.de/proceedings/eswc2006/proceedings-workshops/eswc2006-workshop-semantic-wikis.pdf>>. Acesso em: 25 jun. 2012.
- OREN, Eyal et al. Semantic Wikis for Personal Knowledge Management. In: INTERNATIONAL CONFERENCE ON DATABASE AND EXPERT SYSTEMS APPLICATIONS (DEXA), *Proceedings...*, p. 509-518. Sep. 2006b. Disponível em: <<http://hdl.handle.net/10379/428>>. Acesso em: 13 jan. 2013.
- OREN, Eyal; BRESLIN, John; DECKER, Stefan. How Semantics Make Better Wikis. In: INTERNATIONAL WORLD WIDE WEB CONFERENCE, 15th, 2006, Edinburgh. *Proceedings...*, New York: ACM Press, 2006. Disponível em: <<http://www2006.org/programme/files/pdf/p171.pdf>>. Acesso em: 25 jun. 2012.
- ORTEGA, Cristina Dotta. Do princípio monográfico à unidade documentária: exploração dos fundamentos da catalogação. *Liinc em Revista*, v. 7, n. 1, p. 43-60, 2011. Disponível em: <<http://revista.ibict.br/liinc/index.php/liinc/article/viewFile/402/263>>. Acesso em: 15 set. 2012.
- PANAGIOTOU, Dimitris; MENTZAS, Gregoris. A comparison of Semantic Wiki Engines. In: EUROPEAN CONFERENCE ON OPERATIONAL RESEARCH, 22nd, *Proceedings...* Prague, 2007. Disponível em: <<http://www.docstoc.com/docs/135937925>>. Acesso em: 7 jan. 2013.
- POMBO, Olga. Da classificação dos seres à classificação dos saberes, leituras. *Revista da Biblioteca Nacional de Lisboa*, n. 2, p. 19-33, Primavera 1998. Disponível em:

<<http://www.educ.fc.ul.pt/docentes/opombo/investigacao/opombo-classificacao.pdf>>. Acesso em: 5 jun. 2012.

POOLE, David; MACKWORTH, Alan; GOEBEL, Randy. *Computational Intelligence: a logical approach*. New York: Oxford University Press, 1998.

RAMALHO, Rogério Aparecido Sá. *Desenvolvimento e utilização de ontologias em bibliotecas digitais: uma proposta de aplicação*. Marília, 2010. 145 f. Tese (Doutorado em Ciência da Informação). Programa de Pós-Graduação em Ciência da Informação – Faculdade de Filosofia e Ciências, Universidade Estadual Paulista, Campus de Marília, 2010. Disponível em: <[http://www.marilia.unesp.br/Home/Pos-Graduacao/CienciadaInformacao/Dissertacoes/ramalho\\_ras\\_do\\_mar.pdf](http://www.marilia.unesp.br/Home/Pos-Graduacao/CienciadaInformacao/Dissertacoes/ramalho_ras_do_mar.pdf)>. Acesso em: 10 abr. 2012.

RAMALHO, Rogério Aparecido Sá. *Web Semântica: aspectos interdisciplinares da gestão de recursos informacionais no âmbito da Ciência da Informação*. 2006. 120 f. Dissertação (Mestrado em Ciência da Informação). Faculdade de Filosofia e Ciências – Universidade Estadual Paulista, Marília, 2006. Disponível em: <[http://www.marilia.unesp.br/Home/Pos-Graduacao/CienciadaInformacao/Dissertacoes/ramalho\\_ras\\_me\\_mar.pdf](http://www.marilia.unesp.br/Home/Pos-Graduacao/CienciadaInformacao/Dissertacoes/ramalho_ras_me_mar.pdf)>. Acesso em: 10 abr. 2012.

RAMALHO, Rogério Aparecido Sá; VIDOTTI, Silvana Aparecida Borsetti Gregorio; FUJITA, Mariângela Spotti Lopes. Web semântica: uma investigação sob o olhar da Ciência da Informação. *DataGramZero*, Rio de Janeiro, v. 8 n. 6, dez. 2007. Disponível em: <[http://www.datagramazero.org.br/dez07/Art\\_04.htm](http://www.datagramazero.org.br/dez07/Art_04.htm)>. Acesso em: 25 jun. 2011.

RAY, Erik. *Aprendendo XML*. Rio de Janeiro: Campus; O'Reilly, 2001.

REITZ, Joan M. *Online Dictionary for Library and Information Science*. Westport: Libraries Unlimited, 2010. Disponível em: <[http://www.abc-clio.com/ODLIS/odlis\\_W.aspx?#wiki](http://www.abc-clio.com/ODLIS/odlis_W.aspx?#wiki)>. Acesso em: 27 set. 2010.

RUSSELL, Stuart; NORVIG, Peter. *Artificial Intelligence: a modern approach*. 3th ed. New Jersey: Prentice Hall, 2009.

SÁEZ VACAS, Fernando. Contextualización sociotécnica de la web 2.0: vida y sociedad en el Nuevo Entorno Tecnosocial. In: FUMERO, Antônio; ROCA, Genís. *Web 2.0*. Madrid: Fundación Orange, 2007. 131 p. Disponível em: <[http://www.fundacionorange.es/areas/25\\_publicaciones/WEB\\_DEF\\_COMPLETO.pdf](http://www.fundacionorange.es/areas/25_publicaciones/WEB_DEF_COMPLETO.pdf)>. Acesso em: 26 abr. 2012.

SANTAREM SEGUNDO, José Eduardo. *Recursos tecno-metodológicos para descrição e recuperação de informações na Web*. Marília, 2004. 157 f. Dissertação (Mestrado em Ciência da Informação) - Faculdade de Filosofia e Ciências, Universidade Estadual Paulista. 2004. Disponível em: <[http://www.marilia.unesp.br/Home/Pos-Graduacao/CienciadaInformacao/Dissertacoes/santaremsegundo\\_je\\_me\\_mar.pdf](http://www.marilia.unesp.br/Home/Pos-Graduacao/CienciadaInformacao/Dissertacoes/santaremsegundo_je_me_mar.pdf)>. Acesso em: 20 ago. 2011.

SANTAREM SEGUNDO, José Eduardo. *Representação Iterativa: um modelo para Repositórios Digitais*. Marília, 2010. 244 f. Tese (Doutorado em Ciência da Informação) - Faculdade de Filosofia e Ciências, Universidade Estadual Paulista. 2004. Disponível em: <[http://www.marilia.unesp.br/Home/Pos-Graduacao/CienciadaInformacao/Dissertacoes/santaremsegundo\\_je\\_do\\_mar.pdf](http://www.marilia.unesp.br/Home/Pos-Graduacao/CienciadaInformacao/Dissertacoes/santaremsegundo_je_do_mar.pdf)>. Acesso em: 5 set. 2011.

SANTAREM SEGUNDO, José Eduardo; VIDOTTI, Silvana Aparecida Borsetti Gregorio. Organização da informação na web: a busca na qualidade do armazenamento e da recuperação com a utilização de XML e RDF. In: SIMPÓSIO EM FILOSOFIA E CIÊNCIAS, 5., 2003, Marília. *Anais...* Marília: Unesp Marília Publicações, 2003.

SANTOS, Plácida Leopoldina Ventura Amorim da Costa. Redes informacionais como ambientes colaborativos e de empoderamento: a catalogação em foco. In: GUIMARÃES, José Augusto



Chaves; FUJITA, Mariângela Spotti Lopes (Orgs.). Ensino e pesquisa em biblioteconomia no Brasil: a emergência de um novo olhar. Marília: Cultura acadêmica, 2008, p. 155-171.

SANTOS, Plácida Leopoldina Ventura Amorim da Costa; ALVES, Rachel Cristina Vesú. Metadados e Web Semântica para estruturação da Web 2.0 e Web 3.0. *DataGramaZero*, Rio de Janeiro, v. 10, n. 6, dez. 2009. Disponível em: <[http://www.dgz.org.br/dez09/Art\\_04.htm](http://www.dgz.org.br/dez09/Art_04.htm)>. Acesso em: 25 jun. 2011.

SANTOS, Plácida Leopoldina Ventura Amorim da Costa; VIDOTTI, Silvana Aparecida Borsetti Gregório. Perspectivismo e Tecnologias de Informação e Comunicação: acréscimos à Ciência da Informação? *DataGramaZero: revista de Ciência da Informação*, Rio de Janeiro, v. 10, n. 3, jun. 2009. Disponível em: <[http://dgz.org.br/jun09/Art\\_02.htm](http://dgz.org.br/jun09/Art_02.htm)>. Acesso em: 10 jul. 2010.

SARACEVIC, Tefko. Ciência da informação: origem, evolução e relações. *Perspectivas em Ciência da Informação*, Belo Horizonte, v. 1, n. 1, p. 41-62, jan./jun. 1996. Disponível em: <<http://portaldeperiodicos.eci.ufmg.br/index.php/pci/article/viewFile/235/22>>. Acesso em: 17 ago. 2012.

SAUERMAN, Leopold. *The Gnowsisis: Using Semantic Web Technologies to build a Semantic Desktop*. Vienna, 2003. 93 f. Dissertação (Mestrado em Ciência da Informação) - Information Systems Institute, Vienna University of Technology. 2003. Disponível em: <<http://www.dfki.uni-kl.de/~sauermann/papers/sauermann2003.pdf>>. Acesso em: 20 jan. 2013.

SCHAFFERT, Sebastian et al. *KiWi: A Platform for Semantic Social Software*. In: EUROPEAN SEMANTIC WEB CONFERENCE (ESWC09), 6th, Heraklion, Greece, 2009. Disponível em: <<http://www.schaffert.eu/wp-content/uploads/kiwi-demo.pdf>>. Acesso em: 25 jan. 2013.

SCHAFFERT, Sebastian et al. Learning with semantic wikis. In: WORKSHOP ON SEMANTIC WIKIS, 1st, 2006, Budva. *Proceedings of the First Workshop on Semantic Wikis: from Wiki to Semantics*. Budva: Karlsruher Institut für Technologie, 2006a. p. 109-123. Disponível em: <<http://hnsf.inf-bb.uni-jena.de/proceedings/eswc2006/proceedings-workshops/eswc2006-workshop-semantic-wikis.pdf>>. Acesso em: 25 jun. 2012.

SCHAFFERT, Sebastian. *IkeWiki: A Semantic Wiki for Collaborative Knowledge Management*. In: INTERNATIONAL WORKSHOP ON SEMANTIC TECHNOLOGIES IN COLLABORATIVE APPLICATIONS (STICA 06), 1st, Manchester, UK, June 2006. *Proceedings...*, 2006b. Disponível em: <[http://www.schaffert.eu/wp-content/uploads/schaffert06\\_ikewiki.pdf](http://www.schaffert.eu/wp-content/uploads/schaffert06_ikewiki.pdf)>. Acesso em: 25 jan. 2013.

SCHAFFERT, Sebastian; GRUBER, Andreas; WESTENTHALER, Rupert. A Semantic Wiki for Collaborative Knowledge Formation. In: SEMANTICS, 2005, Vienna, Austria. *Proceedings...*, Vienna: 2005. Disponível em: <<http://citeseerx.ist.psu.edu/viewdoc/summary?doi=10.1.1.94.6443>>. Acesso em: 25 jan. 2013.

SCHIESSL, Marcelo; BRÄSCHER, Marisa. Do texto às ontologias: uma perspectiva para a ciência da informação. *Ciência da Informação*, Brasília, v. 40 n. 2, p. 301-311, maio/ago. 2011. Disponível em: <<http://revista.ibict.br/ciinf/index.php/ciinf/article/view/1947/1413>>. Acesso em: 11 set. 2012.

SCHOROW, Stephanie. Berners-Lee receives Draper Prize. *MIT TechTalk: serving the MIT community*, v. 51, n. 14, p. 1-6, Jan. 2007. Disponível em: <<http://web.mit.edu/newsoffice/2007/techtalk51-14.pdf>>. Acesso em: 25 jun. 2012.

SEMANTIC wiki projects. In: SEMANTICWEB.ORG, 2013. Disponível em: <<http://semanticweb.org/index.php?oldid=54685>>. Acesso em: 10 jan. 2013.

SIQUEIRA, Marcos Antonio. *XML na Ciência da Informação: uma análise do MARC 21*. Marília: 2003. 134 f. Dissertação (Mestrado em Ciência da Informação). Faculdade de Filosofia e Ciências

- Universidade Estadual Paulista, Marília, 2003. Disponível em: <[http://www.marilia.unesp.br/Home/Pos-Graduacao/CienciadaInformacao/Dissertacoes/siqueira\\_ma\\_me\\_mar.pdf](http://www.marilia.unesp.br/Home/Pos-Graduacao/CienciadaInformacao/Dissertacoes/siqueira_ma_me_mar.pdf)>. Acesso em: 25 ago. 2012.

SMW Project. In: SMW: Semantic MediaWiki. [S.l.: s.n], 2013. Disponível em: <[http://semantic-mediawiki.org/w/index.php?title=Help:SMW\\_Project&oldid=25644](http://semantic-mediawiki.org/w/index.php?title=Help:SMW_Project&oldid=25644)>. Acesso em: 23 abr. 2013.

SOUZA, Renato Rocha; ALVARENGA, Lídia. A Web Semântica e suas contribuições para a Ciência da Informação. *Ciência da Informação*, Brasília, v. 33, n. 1, p. 132-141, jan./abr. 2004. Disponível em: <<http://dx.doi.org/10.1590/S0100-19652004000100016>>. Acesso em: 25 jun. 2012.

SOUZIS, Adam. Building a Semantic Wiki. *IEEE Intelligent Systems*, v. 20, n. 5, p. 87-91, Sep./Oct. 2005. Disponível em: <[http://www.liminalzone.org/static/IEEE\\_IS\\_Souzis\\_v20n5.pdf](http://www.liminalzone.org/static/IEEE_IS_Souzis_v20n5.pdf)>. Acesso em: 10 jan. 2013.

TAZZOLI, Roberto; CASTAGNA, Paolo; CAMPANINI, Stefano Emilio. Towards a Semantic Wiki Wiki Web. In: INTERNATIONAL SEMANTIC WEB CONFERENCE (ISWC2004), 3rd. *Proceedings...*, 2004. Disponível em: <<http://platypuswiki.sourceforge.net/whatis/documents/platypuswiki.pdf>>. Acesso em: 10 jan. 2013.

TEMPLATE. In: Meta-Wiki. Florida: Wikimedia Foundation, 2013. Disponível em: <<http://meta.wikimedia.org/w/index.php?title=Help:Template&oldid=5552594>>. Acesso em: 28 mar. 2013.

VRANDEČIĆ, Denny; KRÖTZSCH, Markus. Reusing Ontological Background Knowledge in Semantic Wikis. In: WORKSHOP ON SEMANTIC WIKIS, 1st, 2006, Budva. *Proceedings of the First Workshop on Semantic Wikis: from Wiki to Semantics*. Budva: Karlsruher Institut für Technologie, 2006. p. 16-30. Disponível em: <<http://hnsf.inf-bb.uni-jena.de/proceedings/eswc2006/proceedings-workshops/eswc2006-workshop-semantic-wikis.pdf>>. Acesso em: 25 jun. 2012.

W3C. *Extensible Markup Language (XML) 1.0 (Fifth Edition)*. Cambridge: W3C, 2008a. Disponível em: <<http://www.w3.org/TR/REC-xml>>. Acesso em: 15 jun. 2012.

W3C. *OWL 2 Web Ontology Language: Document Overview (Second Edition)*. Cambridge: W3C, 2012a. Disponível em: <<http://www.w3.org/TR/owl2-overview>>. Acesso em: 30 out. 2012.

W3C. *OWL 2 Web Ontology Language: Primer (Second Edition)*. Cambridge: W3C, 2012b. Disponível em: <<http://www.w3.org/TR/2012/REC-owl2-primer-20121211>>. Acesso em: 30 out. 2012.

W3C. *OWL Web Ontology Language: Overview*. Cambridge: W3C, 2004a. Disponível em: <<http://www.w3.org/TR/owl-features>> Acesso em: 30 out. 2012.

W3C. *RDF Primer*. Cambridge: W3C, 2004b. Disponível em: <<http://www.w3.org/TR/rdf-primer>>. Acesso em: 15 jun. 2012.

W3C. *RDF Vocabulary Description Language 1.0: RDF Schema*. Cambridge: W3C, 2004c. Disponível em: <<http://www.w3.org/TR/rdf-schema>>. Acesso em: 15 jun. 2012.

W3C. *Resource Description Framework (RDF) Model and Syntax Specification*. Cambridge: W3C, 1999. Disponível em: <<http://www.w3.org/TR/1999/REC-rdf-syntax-19990222>>. Acesso em: 15 jun. 2012.

W3C. *Semantic Web*. Cambridge: W3C, 2012c. Disponível em: <<http://www.w3.org/standards/semanticweb>>. Acesso em: 21 ago. 2012.

W3C. *SPARQL 1.1 Overview*. Cambridge: W3C, 2013a. Disponível em: <<http://www.w3.org/TR/sparql11-overview>>. Acesso em: 28 jul. 2013.

W3C. *SPARQL 1.1 Query Language*. Cambridge: W3C, 2013b. Disponível em: <<http://www.w3.org/TR/sparql11-query>>. Acesso em: 28 jul. 2013.

W3C. *SPARQL Query Language for RDF*. Cambridge: W3C, 2008b. Disponível em: <<http://www.w3.org/TR/rdf-sparql-query>>. Acesso em: 22 dez. 2012.

W3C. *Vocabularies*. Cambridge: W3C, 2012f. Disponível em: <<http://www.w3.org/standards/semanticweb/ontology>>. Acesso em: 1 out. 2012.

W3C. *W3C Semantic Web Activity*. Cambridge: W3C, 2013c. Disponível em: <<http://www.w3.org/2001/sw>>. Acesso em: 25 ago. 2013.

WALLACE, Mark. What Makes a Wiki Semantic? In: SEMAPPS: Creating Semantic Applications, 2012. Disponível em: <<http://semapps.blogspot.com.br/2012/02/what-makes-wiki-semantic.html>>. Acesso em: 20 jan. 2013.

ZAIDAN, Fernando Hadad; BAX, Marcello Peixoto. Semantic Wikis and the Collaborative Construction of Ontologies: case study. *Journal of Information Systems and Technology Management (JISTEM)*, São Paulo, v. 8, n. 3, p. 539-554, Sep./Dec. 2011a. Disponível em: <<http://dx.doi.org/10.4301/S1807-17752011000300002>>. Acesso em: 25 jun. 2012.

ZAIDAN, Fernando Hadad; BAX, Marcello Peixoto. Wikis semânticos e a construção colaborativa de ontologias: um estudo de caso. In: CONTECSI INTERNATIONAL CONFERENCE ON INFORMATION SYSTEMS AND TECHNOLOGY MANAGEMENT, 8th, 2011b, São Paulo. *Anais...* São Paulo: USP, 2011b. Disponível em: <<http://eprints.rclis.org/handle/10760/15989>>. Acesso em: 25 jun. 2012.