

Web semântica: uma investigação sob o olhar da Ciência da Informação*Semantic Web: an inquiry under the view of the Information Science*

por Rogério Aparecido Sá Ramalho e Silvana Aparecida Borsetti Gregorio Vidotti e Mariângela Spotti Lopes Fujita

Resumo: No âmbito da gestão de recursos informacionais os modelos e métodos de organização e recuperação de informações sempre estiveram condicionados às tecnologias utilizadas, de modo que a partir do desenvolvimento e intensificação da utilização dos meios digitais uma nova gama de possibilidades vem sendo incorporada, atingindo um estágio em que os modelos clássicos de organização e recuperação de informações precisam ser (re)pensados sob diferentes perspectivas, evidenciando a necessidade de novos métodos que possibilitem otimizar a recuperação de informações. Este artigo comprehende uma discussão de literatura, de caráter interdisciplinar, com o objetivo de favorecer a “desmistificação” dos conceitos e tecnologias subjacentes ao projeto Web Semântica, avaliando em que medida a área de Ciência da Informação pode contribuir para sua concretização e ressaltando os reflexos das novas abordagens tecnológicas de representação e recuperação de recursos informacionais no corpus teórico da área de Ciência da Informação.

Palavras-chave: Web semântica; Recuperação de informação; Ontologia; Sistemas de informação; Gestão de recursos informacionais.

Abstract: In the field of information resources management, the models and methods of information organization and retrieval have always been linked to the current technologies in a way that a new range of possibilities have appeared from the development and improvement on the use of digital means, reaching a stage in which the classical models of information organization and retrieval must be (re)thought under different perspectives, highlighting the need of new methods that allow to optimize the information retrieval. This article brings out an interdisciplinary literature discussion, aiming to clarify the concepts and technologies related to the Semantic Web project, evaluating how the Information Science area can contribute to its concretization, as well as bringing out the consequences of new technological approaches of information resources representation and retrieval in the Information Science theoretical corpus.

Keywords: Semantic web; Information retrieval; Ontology; Information system; Information resources management.

Introdução

É inegável o fato de que o ambiente Web constitui-se como uma das mais ricas fontes de informações da atualidade, apresentando-se também como um ambiente interativo que possibilita a troca de informações em escala global. Tal fato, que à primeira vista apresenta ser o seu maior apelo, é ao mesmo tempo um de seus fatores críticos.

Nos últimos anos, o avanço exponencial na quantidade de recursos informacionais disponíveis no ambiente Web vêm conduzindo a um estágio em que os modelos clássicos de representação e recuperação de informações precisam ser (re)pensados sob diferentes perspectivas, pois considerando a representação como elemento fundamental, para a garantia de qualidade na recuperação, apresenta-se como desafio a necessidade de singularização contextual na reconstrução do conhecimento, a partir da determinação de requisitos de qualidade e relevância das informações, que permitam categorizar e organizar, de maneira eficiente, o “oceano” de dados disponíveis, favorecendo a identificação de informações que realmente interessam ao usuário.

Tradicionalmente a Internet, uma infra-estrutura de redes, servidores e canais de comunicação, popularizada principalmente a partir do desenvolvimento da *World Wide Web*, ou simplesmente Web, tem como uma de suas principais funcionalidades a tarefa de disponibilizar conteúdos informacionais de modo que estes possam ser visualizados e interpretados por usuários humanos.

Contudo, a partir do final da década de 1990, começaram a formalizar-se pesquisas relacionadas ao desenvolvimento de uma nova geração da Web, com o objetivo de possibilitar a incorporação de ligações semânticas aos recursos informacionais, de modo que os próprios computadores possam “compreendê-las” de forma automatizada. *Machine understandable information*, com esta sucinta

expressão [Berners-Lee](#) (1998) impulsionou os primeiros estudos em direção ao projeto da Web Semântica.

Esta investigação comprehende uma discussão de literatura acerca o projeto Web Semântica, buscando favorecer a “desmistificação” dos conceitos e tecnologias subjacentes e avaliar em que medida a área de Ciência da Informação pode contribuir para sua concretização, ressaltando os reflexos das novas abordagens tecnológicas de representação e recuperação de recursos informacionais no corpus teórico da área de Ciência da Informação.

Na próxima seção busca-se identificar a fundamentação teórica inerente ao projeto Web Semântica, descrevendo as principais tecnologias necessárias para sua concretização e como estas se relacionam em sua arquitetura de camadas. Na seção 3 são apresentadas e discutidas novas abordagens tecnológicas de representação e recuperação de recursos informacionais, contextualizando-as a partir do enfoque da área de Ciência da Informação. A seção 4 apresenta os desafios e responsabilidades dos profissionais da informação no âmbito do projeto Web Semântica, a partir de um “*Espectro Funcional*” de sua arquitetura de camadas, proposto com o intuito de omitir detalhes técnicos e facilitar a compreensão de suas principais características. Ainda neste trabalho são apresentadas as conclusões da análise realizada.

Arquitetura da Web Semântica

Desde seus primórdios o projeto Web Semântica foi concebido como um conjunto de tecnologias relacionadas, de modo que no ano de 2000 o World Wide Web Consortium (W3C) [L], tendo como seu maior expoente [Tim Berners-Lee](#), divulgou publicamente a primeira proposta de arquitetura da [Web Semântica](#), com base em uma série de camadas sobrepostas, onde cada camada ou tecnologia deveria obrigatoriamente ser complementar e compatível com as camadas inferiores, ao mesmo tempo em que não deveria depender das camadas superiores, possibilitando assim uma estrutura idealmente escalonável, que indicasse os passos e as tecnologias necessários para a concretização do projeto Web Semântica.

Segundo Bernes-Lee (1999), o primeiro passo para o desenvolvimento da Web Semântica seria a inclusão de dados em um formato que os sistemas computacionais pudessem naturalmente compreender de forma direta ou indireta. Após a publicação desta definição, em seu livro Weaving the Web, a expressão “Web Semântica” passou a ser disseminada como um título genérico que representa uma série de pesquisas que têm como objetivo principal possibilitar um melhor aproveitamento das potencialidades do ambiente Web, onde por meio do uso intensivo de linguagens computacionais e instrumentos de metadados espera-se obter o acesso automatizado às informações de maneira mais precisa, a partir da utilização de processamentos semânticos e heurísticas automáticas.

Conforme afirmam [Berners-Lee](#) et al. (2001, p.2, tradução nossa): “A Web Semântica é uma extensão da Web atual, onde a informação possui um significado claro e bem definido, possibilitando uma melhor interação entre computadores e pessoas”. Assim, observa-se que comparando com as abordagens tradicionalmente desenvolvidas, o projeto Web Semântica constitui-se como uma tentativa inversa de solução que tem como objetivo desenvolver meios para que as máquinas possam servir aos humanos de maneira mais eficiente, mas para isso torna-se necessário construir instrumentos que forneçam sentido lógico e semântico aos computadores.

Com o intuito de ilustrar as dificuldades encontradas nos processos de recuperação de informações na Web atual pode-se analisar os resultados de uma busca realizada por meio dos tradicionais “motores de busca” (search engines), como por exemplo, a busca de textos científicos de um determinado autor. Utilizando-se, por exemplo, “Arlindo Machado” como expressão de busca ter-se-ia como resultado todos os tipos de documentos que contenham “Arlindo Machado” em alguma parte de seu

conteúdo.

Caso esta mesma busca seja realizada utilizando-se apenas o sobrenome deste autor o problema se agravaría ainda mais, pois “Machado” pode igualmente se referir ao sobrenome de uma pessoa, com também a um instrumento cortante utilizado para rachar madeira e a Web atual não fornece condições que possibilitem distinguir entre os vários significados semânticos que um termo pode comportar, o que favorece a recuperação de uma grande quantidade de documentos irrelevantes ou não relacionados com a busca realizada, tornando algumas vezes inexequível a tarefa de localizar informações específicas no ambiente Web.

Segundo Koivunen e Miller (2001), um dos princípios fundamentais do projeto Web Semântica é o fato de que “*tudo*” pode ser identificado por um Uniform Resource Identifier (URI), de modo que pessoas, lugares e elementos do mundo físico possam ser referenciados a partir de tais identificadores.

Possibilitando assim identificar uma instituição a partir do URI de sua página Web, por exemplo, ou uma pessoa por meio do URI de sua caixa de e-mail.

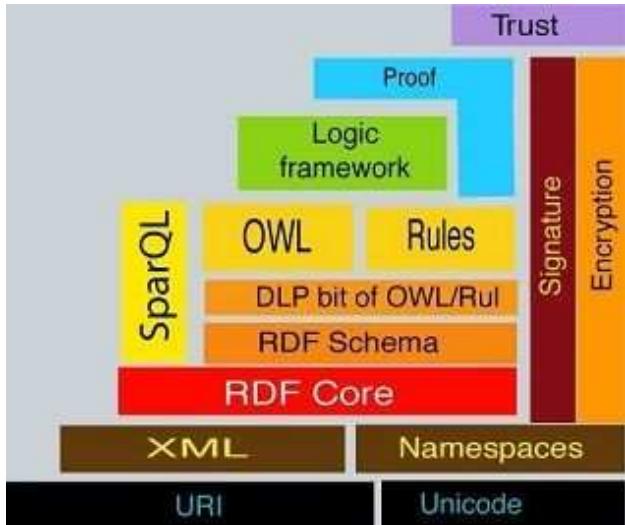
Conforme afirmam Berners-Lee et al. (1994) e Fielding (1995), um URI é um padrão conjunto que abrange os conceitos de *Uniform Resource Locator (URL)* e do *Uniform Resource Name (URN)*, de modo que pode ser representado por qualquer um destes, ou por ambos. Fazendo uma analogia de um recurso disponível no ambiente Web com um livro armazenado em uma biblioteca, pode-se considerar o *URN* de um recurso como o número *ISBN* de um livro, os quais fornecem uma identificação exclusiva, porém não oferecendo informações a respeito de onde o livro/recurso pode ser obtido.

Do mesmo modo, pode-se considerar que o código que identifica onde um livro está localizado, em meio ao acervo de uma biblioteca, desempenha a mesma função do *URL* de um recurso no ambiente Web, indicando o local onde o livro/recurso pode ser obtido. Segundo Krishnamurthy e Rexford (2001, p.183), o modo mais popular de apresentação de um *URI* é utilizando um *URL*, o qual pode ser considerado como uma cadeia de caracteres formada por componentes padronizados.

Outra característica importante do projeto Web Semântica é que os links podem possuir diferentes tipos, possibilitando a definição de conceitos úteis para as máquinas, como por exemplo, indicando que um recurso é uma versão de outro recurso ou que contém informações a respeito de uma determinada pessoa. A Web atual também consiste de recursos e links, porém estes links são criados apenas para o uso de seres humanos, de modo que é relativamente simples para um ser humano identificar se um link, contido em um determinado recurso, referencia uma fatura, um romance ou um trabalho científico, contudo tais informações não estão acessíveis para as máquinas, pois os links na Web atual não indicam formalmente quais são os tipos de relações existentes entre os recursos referenciados.

Devido ao grande número de pesquisadores envolvidos no projeto Web Semântica, não tardou para que a arquitetura proposta inicialmente pelo The World Wide Web Consortium, W3C sofresse modificações, sendo que em 2005 foi publicada uma nova proposta de arquitetura incorporando novas tecnologias, com o intuito de possibilitar uma maior integração entre as camadas e facilitar a realização de consultas semânticas. Como pode ser verificado na figura 1.

FIGURA 1 – Arquitetura da web semântica proposta em 2005



Fonte: [Berners-Lee, 2005.](#) p.17

Para fins deste estudo, baseando-se na proposta de arquitetura apresentada acima, pode-se descrever as principais tecnologias e camadas inerentes ao projeto Web Semântica, sucintamente, da seguinte maneira:

- **URI:** Conforme apresentado anteriormente, tal componente consiste de um Identificador Único de Recursos que possibilita a definição e adoção, de maneira precisa, de nomes aos recursos e seus respectivos endereços na Internet.
- **UNICODE:** Esquema padronizado de codificação dos caracteres, que diminui consideravelmente a possibilidade de redundâncias dos dados, pois funciona independentemente da plataforma utilizada.
- **Signature:** Conjunto de tecnologias desenvolvidas com o intuito de substituir em ambiente computacional a função exercida pela assinatura formal de uma pessoa em um suporte físico. Segundo [Pfützenreuter \(2004\)](#), a assinatura digital garante a integridade dos dados e a comprovação da procedência dos recursos.
- **Encryption:** Consiste de um processo em que as informações são cifradas de modo que não possam ser interpretadas por qualquer pessoa ou sistema computacional, garantindo assim a confidencialidade das informações. Segundo [Nakamura e Geus \(2003, p.287\)](#), “[...] encryption é o processo de disfarçar a mensagem original, [...] , de tal modo que sua substância é escondida em uma mensagem com texto cifrado”.
- **XML:** É uma linguagem computacional que possibilita a estruturação dos dados por meio da definição de elementos e atributos, e que permite a descrição de regras sintáticas para a análise e validação dos recursos.
- **Namespace:** Coleção de nomes, identificados por um URI, que são utilizados em documentos XML para validar elementos e atributos.
- **RDF Core:** Núcleo que comprehende as especificações do modelo e a sintaxe da Resource Description Framework (Estrutura de Descrição de Recursos), possibilitando a descrição dos recursos por meio de suas propriedades e valores. Segundo Daum e Merten (2002), a RDF pode ser vista como uma tecnologia de capacitação para a modelagem semântica, sobre a qual podem ser criadas linguagens computacionais específicas.
- **RDF Schema:** Utilizada para a descrição do vocabulário RDF, possibilitando a

definição de taxonomias de recursos em termos de uma hierarquia de classes. Segundo Brickley (2004), a RDF Schema é uma extensão semântica do código RDF, fornecendo mecanismos para descrever grupos de recursos e os relacionamentos existentes entre eles.

[FO] *SparQL: Segundo alguns pesquisadores pertencentes ao W3C, ([Prud'hommeaux e Seaborne](#), 2005; [Clark](#), 2005), SparQL é uma linguagem computacional utilizada para realizar consultas a partir de estruturas RDF, favorecendo a recuperação de informações de maneira mais eficaz. Tal linguagem ainda não se encontra completamente padronizada, motivo pelo qual não é recomendada oficialmente pelo W3C, sendo denominada como uma tecnologia candidata à recomendação.*

[FO] *DLP: A DLP também é uma tecnologia candidata à recomendação e constitui a intersecção entre os dois principais paradigmas utilizados atualmente para desenvolver sistemas computacionais baseados em representação do conhecimento, Lógica Descritiva (OWL DL) e Programação Lógica (F-Logic), fornecendo uma estrutura extremamente flexível. ([Grosof](#), (2003) e [Vrandecic](#). (2005)). Deste modo, a DLP ainda não é considerada uma linguagem de representação do conhecimento, mas sim como uma “ponte” que possibilita a união entre os dois principais paradigmas utilizados.*

[FO] *OWL: Linguagem computacional recomendada pelo W3C para o desenvolvimento de ontologias. Segundo [McGuinness e Harmelen](#) (2004), a linguagem OWL permite descrever formalmente, de modo mais eficiente, os aspectos semânticos dos termos utilizados e seus respectivos relacionamentos, possibilitando representações mais abrangentes das linguagens RDF e RDF Schema e favorecendo uma maior interoperabilidade.*

[FO] *Rules: Permite a definição de regras lógicas relacionadas aos recursos informacionais. Segundo [Daconta, Obrst e Smith](#) (2003), esta camada possibilita uma espécie de “Introdução Lógica”, enquanto que a camada superior, Logic Framework, possibilita a incorporação de “Lógicas Avançadas”.*

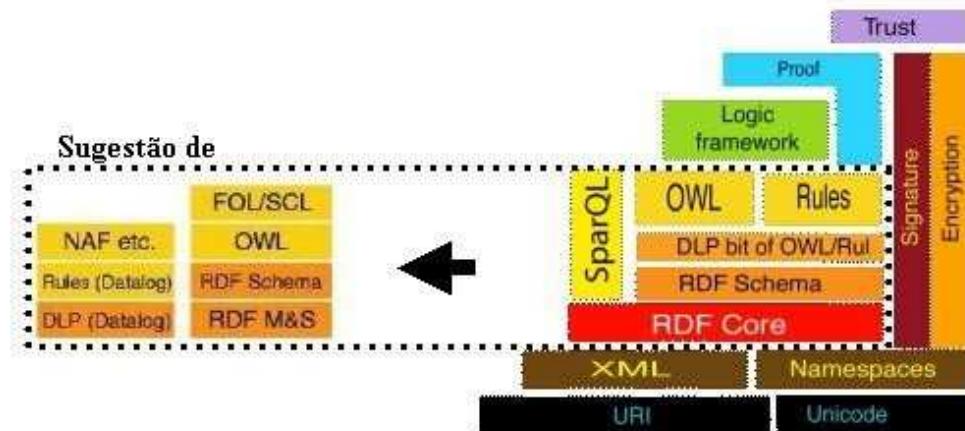
[FO] *Logic Framework: Camada para a definição de regras mais abrangentes, utilizadas no tratamento das informações descritas nos níveis inferiores, possibilitando que agentes computacionais possam realizar inferências automáticas a partir das relações existentes entre os recursos informacionais, podendo inclusive inferir novas informações.*

[FO] *Proof: Espera-se que esta camada possibilite a verificação/comprovação da coerência lógica dos recursos, de modo que os aspectos semânticos das informações estejam descritos de maneira consideravelmente adequada, atendendo a todos os requisitos das camadas inferiores.*

[FO] *Trust: Camada de Confiança, a partir da qual se espera garantir que as informações estejam representadas de modo correto, possibilitando um certo grau de confiabilidade.*

A partir de tais definições observa-se que o projeto Web Semântica encontra-se em constante desenvolvimento, de modo que muitas das tecnologias propostas em sua arquitetura ainda estão em fase de avaliação e verificação de seus resultados. Assim sendo, poucos meses após a publicação da arquitetura descrita acima foram apresentadas críticas e sugestões de alterações para a mesma. Segundo [Horrocks](#) (2005), algumas das linguagens apresentadas nesta arquitetura não são semanticamente compatíveis de forma direta, como a [Description Logic Programs](#) (DLP) e a linguagem de desenvolvimento de ontologias [Web Ontology Language](#), OWL, de modo que tais tecnologias deveriam ser apresentadas paralelamente ao invés de sobrepostas, formando duas torres em uma parte da arquitetura, conforme figura 2, a seguir.

FIGURA 2 – Sugestão de alteração na arquitetura proposta para web semântica



adaptado de: [Berners-Lee](#), 2005, p.17 e Horrocks et al., 2005, p. 3.

Por se tratar de um projeto em fase de desenvolvimento verifica-se que provavelmente a arquitetura da Web Semântica ainda sofrerá modificações, para comprovar esta tendência basta observar que as camadas “*Logic Framework*” e “*Prof*” ainda não possuem sequer tecnologias recomendadas para suas implementações, pois apesar da padronização da linguagem [Web Ontology Language](#), OWL, como recomendação para o desenvolvimento de ontologias, atualmente ainda não está muito claro como esta camada irá relacionar-se com as demais camadas da arquitetura, informação esta imprescindível para o desenvolvimento das camadas superiores.

É importante ressaltar que apesar das novas tecnologias incorporadas na proposta de arquitetura publicada em 2005 pelo [Wide Web Consortium](#), W3C e das críticas e sugestões de alterações de tal arquitetura, é possível identificar aspectos que não devem ser alterados no projeto Web Semântica, afinal novas tecnologias estarão sempre sendo desenvolvidas, mas os conceitos básicos que norteiam o desenvolvimento do projeto Web Semântica tendem a permanecer estáveis.

Novas abordagens de representação e recuperação de recursos informacionais

Atualmente, verifica-se uma tendência de aproximação entre as áreas de Ciência da Informação e Ciência da Computação, principalmente no que tange ao desenvolvimento de novos instrumentos de representação e recuperação de recursos informacionais. Segundo [Saracevic](#) (1996), a recuperação de informação pode ser considerada como a vertente tecnológica da Ciência da Informação.

No entanto, apesar de relacionadas, é possível identificar um distanciamento teórico entre tais Ciências. [Ferneda](#) (2003, p.1-2) ao comentar sobre as relações existentes entre a Ciência da Informação e a Ciência da Computação, ressalta que: “[...] a informação, objeto de comum interesse de ambas as ciências, é paradoxalmente o que mais as distancia”.

Tal afirmação sustenta-se no fato de que ambas as Ciências utilizam diferentes conceitos ao se referir ao termo informação. Segundo [Le Coadic](#) (1996, p.5), para a área de Ciência da Informação, a informação “[...] comporta um elemento de sentido. É um significado transmitido a um ser consciente por meio de uma mensagem inscrita em um suporte espaço-temporal: impresso, sinal-eletrico, onda sonora, etc.”. Em contra partida, para a área de Ciência da Computação o conceito de informação tradicionalmente restringe-se aos conceitos relacionados com a Teoria Matemática da

Informação, mais especificamente com os fundamentos estabelecidos por [Shannon e Weaver](#) (1949), presentes no artigo intitulado *The Mathematical Theory of Communication*, onde é apresentada uma teoria matemática para a transmissão de mensagens e troca de sinais, a qual não se preocupa com a semântica dos dados, porém adequada para a construção de sistemas computacionais, onde a informação possa ser quantificada, processada e transmitida por máquinas. Conforme afirma [Shannon](#) (1948, p.3): “*the theory of computing machines*”.

A identificação de tal enfoque na área de Ciência da Computação pode ser comprovada a partir da análise de sua literatura e dos resultados práticos identificados no seu campo de atuação. No contexto nacional também é possível verificar tal abordagem nas diretrizes curriculares de cursos na área de Computação e Informática elaboradas pela *Comissão de Especialistas de Ensino em Computação e Informática* ([CEEInf](#), 1999) e disponibilizadas pelo *Ministério da Educação* (MEC), assim como no currículo de referência para cursos de graduação em Computação, proposto pela Sociedade Brasileira de Computação ([SBC](#), 1999, p.14), no qual figura como parte do currículo básico para cursos de computação, o estudo de:

Princípios da teoria da informação: codificação da informação e sua medida, entropia de código. Transmissão da informação e modelagem do sistema de transmissão, maximização do fluxo de informação por um canal. Processamento digital de sinais, análise espectral. Transmissão analógica e digital.

Segundo [Capurro](#) (2003), embora a área de Ciência da Informação também tenha sido influenciada em seu campo teórico pela assim chamada “*information theory*” de [Shannon e Weaver](#), a mesma não se limita a esta visão puramente [físicista](#), abarcando também outros aspectos: contextuais, semânticos, sociais e culturais.

Contextualizando tais definições, verifica-se que a Ciência da Computação tradicionalmente preocupa-se com o desenvolvimento de sistemas computacionais que possibilitem “*manusear*” informações, sem necessariamente preocupar-se com os aspectos semânticos subjacentes, enquanto que a Ciência da Informação preocupa-se com a natureza das informações, assim como sua comunicação e uso.

Utilizando como exemplo um processo de recuperação de informação no ambiente Web, pode-se verificar que no âmbito da área de Ciência da Informação é necessário levar em consideração os componentes semânticos inerentes a tal processo, no entanto, de acordo com o enfoque da área de Ciência da Computação, observa-se que os tradicionais “*motores de busca*”, *search engines*, baseiam-se exclusivamente na recuperação de dados, não levando em consideração as semânticas contidas nas páginas da Web, recuperando apenas seqüências de caracteres que satisfaçam determinadas condições de busca.

Um dos principais “slogans” do projeto Web Semântica baseia-se na possibilidade de permitir a classificação de recursos informacionais disponíveis no ambiente Web, “rotulando-os” a partir de categorias que possam ser “interpretadas” automaticamente pelos computadores. Analisando tal definição sob o prisma da área de Ciência da Informação pode-se fazer uma analogia à definição apresentada por [Lancaster](#) (2004, p. 21):

No campo do armazenamento e recuperação de informação a classificação de documentos refere-se à formação de classes de itens com base no conteúdo temático. Tesauros, cabeçalhos de assuntos e esquemas de classificação bibliográfica são essencialmente listas de rótulos com os quais se identificam e, por ventura, se organizam essas classes.

Para tornar possível a classificação e categorização dos recursos informacionais disponíveis no

ambiente Web, a ponto de permitir a realização de inferências automáticas, como esperam os idealizadores e defensores do projeto Web Semântica, é necessário primeiramente que os computadores sejam capazes de captar as informações descritivas e temáticas inerentes a tais recursos. Tal fato justifica o motivo pelo qual o projeto Web Semântica tem alavancado uma forte demanda de desenvolvimento de modelos/instrumentos de representação no âmbito computacional.

Conforme apresenta Campos (2004), no contexto da área de Ciência da Computação, espera-se que os instrumentos de representação auxiliem na implementação de estruturas computáveis, que possibilitem aos computadores a realização de tarefas mais sofisticadas de forma automatizada, enquanto que no âmbito da área de Ciência da Informação, os modelos de representação são utilizados há muito tempo na elaboração de linguagens documentárias verbais e notacionais, visando à recuperação de informação e à organização dos conteúdos informacionais dos documentos.

Nos últimos anos, diferentes abordagens computacionais têm sido propostas com o objetivo de favorecer o aperfeiçoamento dos processos de representação e classificação de recursos informacionais, contudo pode-se destacar o desenvolvimento de ontologias como um dos principais temas que vem despertando o interesse de pesquisadores, das mais variadas áreas do conhecimento, empenhados no estudo desta nova e instigante categoria de instrumentos de representação. Segundo Ramalho (2006, p.97), no âmbito do projeto Web Semântica pode-se definir ontologia como:

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.

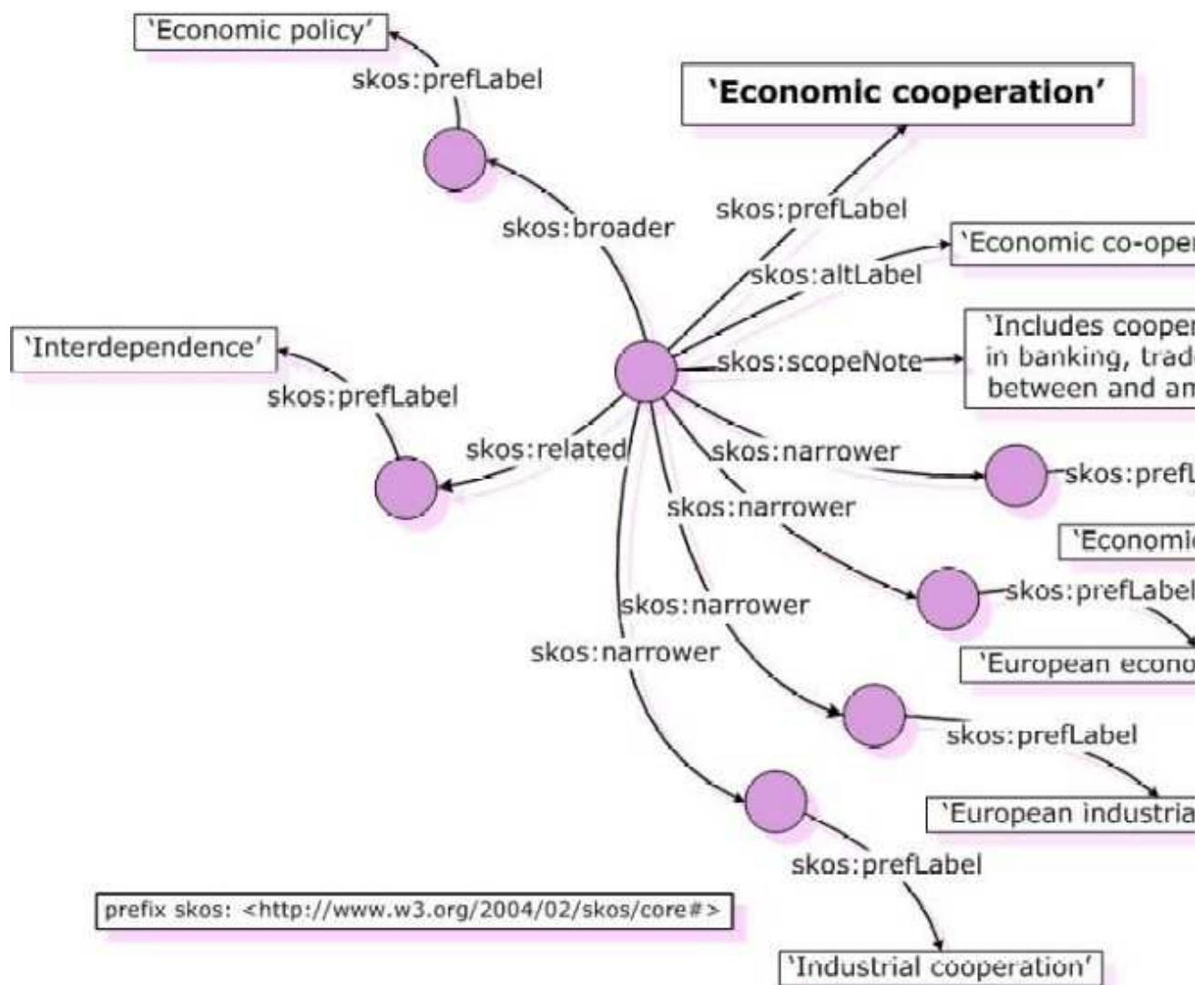
Entre as inúmeras iniciativas relacionadas ao desenvolvimento de instrumentos de representação e recuperação de recursos informacionais pode-se destacar os estudos desenvolvidos no *Semantic Web Advanced Development for Europe2 (SWAD-E)*[2], que têm demonstrado um forte interesse de pesquisadores ligados ao W3C nos instrumentos de representação tradicionalmente utilizados no âmbito da área de Ciência da Informação, conforme pode ser identificado na *SWAD-E Thesaurus Activity* [3], uma iniciativa que tem como objetivo principal desenvolver tecnologias que permitam expressar, de maneira formal, a estrutura básica e o conteúdo de tesouros e vocabulários controlados, possibilitando a sua utilização de forma automatizada em *Knowledge Organization Systems (KOS)*, Sistemas de Organização do Conhecimento.

Segundo Binding e Tudhope (2004), é possível identificar nos últimos anos uma crescente demanda por pesquisas relacionadas a Knowledge Organization Systems, KOS, devido principalmente ao rápido crescimento de comunidades científicas empenhadas no desenvolvimento de projetos relacionados com: Web Semântica, *Semantic Grid* e Bibliotecas Digitais.

Várias iniciativas têm sido propostas com o intuito de atualizar os padrões de tesouros internacionais para que considerem esses desenvolvimentos on-line, paralelamente o W3C tem se empenhado no desenvolvimento de padrões que dêem suporte ao uso de Knowledge Organization Systems, KOS a partir da estrutura da Web Semântica, entre os quais pode-se destacar o *Simple Knowledge Organisation System (SKOS)*. Segundo trabalhos apresentados por pesquisadores ligados ao *Wide Web Consortium*, W3C (Miles e Brickley, 2005), o Simple Knowledge Organisation System, SKOS representa desde estruturas simples utilizadas para expressar sistemas de organização do conhecimento, até estruturas mais complexas e abrangentes, fornecendo um modelo para expressar a estrutura básica e o conteúdo de tesouros, esquemas de classificação, lista de cabeçalho de assunto, taxonomias e também outros esquemas conceituais.

O Simple Knowledge Organisation System, SKOS é um modelo, ainda em fase de desenvolvimento, que compreende um conjunto de propriedades descritas em *Resource Description Framework*, RDF, a partir de classes RDFS, que podem ser utilizadas para expressar o conteúdo e a estrutura de um esquema de conceito como um gráfico RDF, possibilitando descrever formalmente os termos e relacionamentos existentes em um tesauro, conforme apresentado na seguir na figura 3.

FIGURA 3 – Parte de um tesauro representado a partir do SKOS



Fonte: [Miles e Brickley \(2005\)](#).

Segundo [Garcia Jiménez \(2004, p.90,\)](#): “As relações entre ontologias e tesouros parecem demonstrar uma tendência evidente: a intenção por parte de diversos especialistas em elaborar determinadas ontologias a partir de um tesauro”. Seguindo esta tendência, verifica-se também o crescente desenvolvimento de projetos no âmbito de bibliotecas digitais que utilizam-se das tecnologias relacionadas ao projeto Web Semântica, tais como o [JeromeDL](#) e o [MarcOnt](#).

Conforme [Kruk](#), (2005) apresentam no artigo sugestivamente intitulado como *JeromeDL - Reconnecting Digital Libraries and the Semantic Web*, o projeto [JeromeDL](#) consiste de uma biblioteca digital de código aberto baseada nas principais tecnologias presentes no projeto Web Semântica, permitindo a descrição de recursos a partir da linguagem computacional *Resource Description Framework*, RDF e a realização de buscas semânticas baseadas em ontologias,

possibilitando uma melhora considerável na precisão das buscas e um maior nível de interoperabilidade.

Quanto ao projeto MarcOnt, segundo Synak e Kruk (2005), seu principal objetivo é o desenvolvimento de uma ontologia capaz de tornar-se um padrão de representação de informações para bibliotecas digitais, possibilitando a descrição dos aspectos semânticos dos conteúdos e favorecendo a integração de bibliotecas. Estando ainda em fase de implementação e avaliação, a ontologia *MarcOnt* vem sendo desenvolvida a partir da linguagem *Web Ontology Language*, OWL, de modo que se espera que tal ontologia seja compatível com o formato MARC 21, permitindo que as descrições semânticas possam ser convertidas para outros formatos, possibilitando grande interoperabilidade a partir do reaproveitamento das bases de conhecimento e por meio da incorporação de outras ontologias que sigam os mesmos critérios de desenvolvimento.

Assim, verifica-se que muitos dos conceitos inerentes aos métodos convencionais de representação podem ser aproveitados no desenvolvimento de novas abordagens voltadas para o ambiente Web, do mesmo modo que muitas tecnologias desenvolvidas a partir do desenvolvimento do projeto Web Semântica não limitam seu escopo de aplicabilidade ao ambiente Web, como por exemplo, a linguagem computacional XML e mais recentemente o desenvolvimento de ontologias que têm despertado o interesse de inúmeros pesquisadores da área de Ciência da Informação, conforme pode ser observado nos trabalhos de Qin e Paling (2001), Alvarenga (2001), Ding e Foo (2002), Soergel (2002), Almeida e Bax (2003), Ferneda (2003), Golbeck (2003), Pincemin (2003), Arano e Codina (2004), Campos (2004), Garcia Gimenez (2004), Pérez Agüera (2004), Souza e Alvarenga (2004), Moreira e Oliveira (2005), Ramalho, Vidotti e Fujita (2005), entre outros.

Conforme afirmam Souza e Alvarenga (2004, p. 139), no artigo intitulado: A Web Semântica e suas contribuições para a ciência da informação: “*Tudo indica que os padrões que estão sendo desenhados para esta nova Web também sejam adotados na arquitetura de bibliotecas digitais e de novos sistemas de informação*”.

Web Semântica e os profissionais da informação

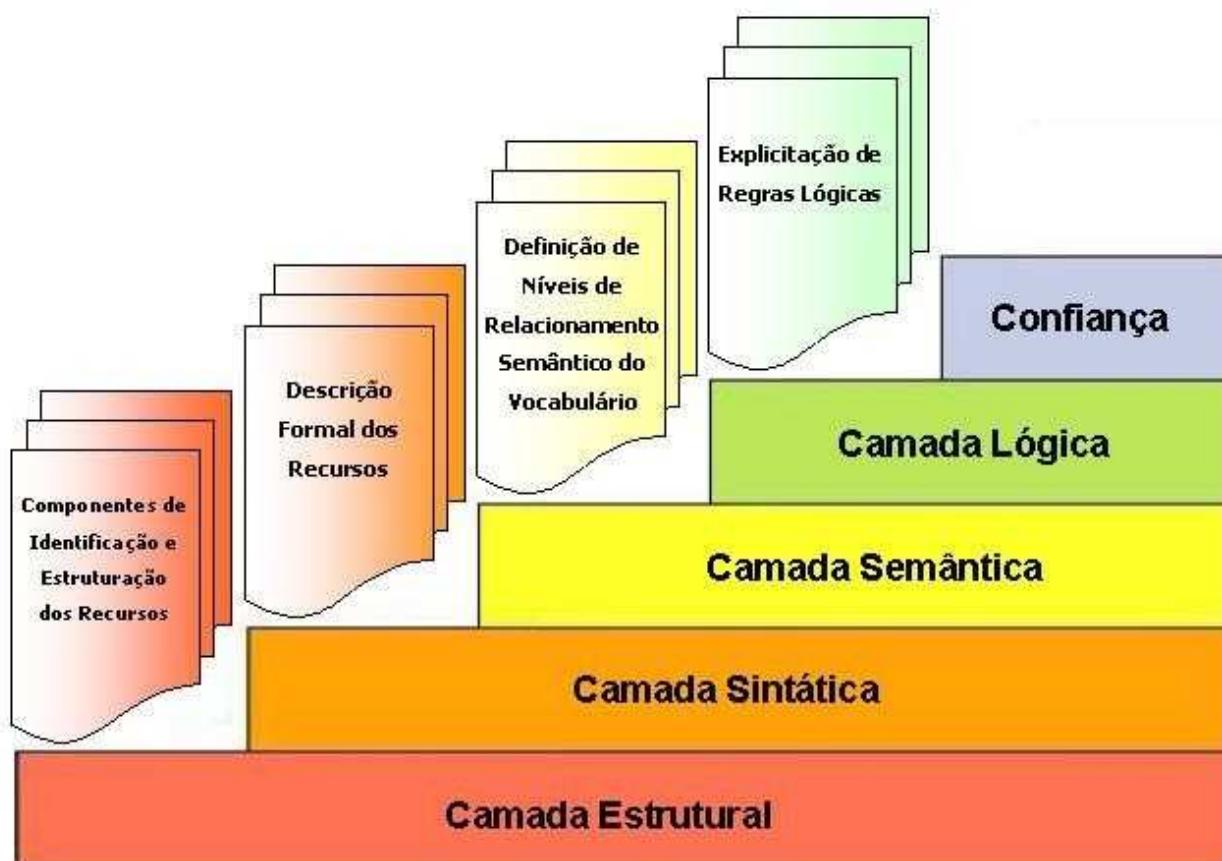
Segundo Alvarenga (2001), novos desafios vêm sendo apresentados aos profissionais da informação, a partir do aumento das possibilidades de disponibilização e processamento de informações no meio digital, e muitas vezes a formação convencional no campo da biblioteconomia garante pouca fundamentação matemática e computacional a tais profissionais, sendo necessário o estabelecimento de parcerias e a formação de equipes interdisciplinares.

Conforme afirma Garcia-Gimenez (2004), deve-se lembrar que as novas tecnologias informacionais relacionadas à organização e recuperação de informações nasceram imersas nos ambientes computacionais, de modo que tal imbricação e dependência tecnológica exigem de seus criadores determinados conhecimentos e habilidades próprias em um contexto técnico e qualificado, para que possam representar o conhecimento e permitir sua posterior recuperação a partir de métodos que superem os tradicionais, melhorando sua eficácia.

No âmbito do projeto Web Semântica também observa-se a necessidade de uma maior familiarização dos profissionais da informação com as novas tecnologias, para que as mesmas também possam ser desenvolvidas a partir de princípios éticos sociais e não baseadas única e exclusivamente em conhecimentos e processos puramente técnicos. Afinal as tecnologias subjacentes ao projeto Web Semântica e os instrumentos de representação de informações desenvolvidos no âmbito da área de Ciência da Informação possuem como objetivo comum propiciar meios mais adequados de representar e organizar conteúdos informacionais, possibilitando responder de maneira mais eficiente às buscas realizadas diretamente pelos usuários finais.

É importante observar que muitos dos conceitos apresentados na arquitetura do projeto Web Semântica constituem uma nova “roupagem tecnológica” para métodos e técnicas que já são utilizados há décadas na área de Ciência da Informação. Assim, propõe-se um “*Espectro Funcional*” da arquitetura da Web Semântica, figura 4, desenvolvido a partir do enfoque da área de Ciência da Informação e baseado nos conceitos básicos das principais tecnologias presentes na arquitetura da Web Semântica, apresentados na figura 1 - pág. 4, com o intuito de omitir detalhes técnicos e facilitar a compreensão de tal arquitetura.

FIGURA 4 – Espectro Funcional da Arquitetura da Web Semântica



De acordo com a figura acima pode-se observar que para a concretização do projeto Web Semântica torna-se necessária primeiramente a identificação dos recursos a partir da “*Camada Estrutural*”, a qual permite identificar cada recurso de forma única e padronizada e possibilita meios seguros de transmissão e armazenamento das informações.

A “*Camada Sintática*” possibilita a descrição dos recursos, por meio da definição e validação de regras sintáticas formalmente descritas, possibilitando a estruturação dos recursos informacionais. Analisando esta camada pode-se destacar o forte embasamento da área de Ciência da Informação no que se refere a práticas de catalogação e indexação que podem ser de grande valia para o desenvolvimento do projeto Web Semântica.

Quanto à “*Camada Semântica*”, é nesta camada que se espera que sejam desenvolvidos vocabulários que permitam descrever os aspectos semânticos inerentes aos recursos informacionais e sistemas de conceitos que definam formalmente as relações existentes, de modo que a partir de tais vocabulários possam ser definidas, na “*Camada Lógica*”, as regras a serem interpretadas computacionalmente, possibilitando a realização de inferências automáticas e a verificação do nível de coerência lógica dos recursos.

Cabe ressaltar que a partir dos instrumentos de representação utilizados tradicionalmente no âmbito da área de Ciência da Informação (*tesauros, lista de cabeçalhos de assunto, taxonomias, etc*), é possível o desenvolvimento de ontologias, permitindo a representação formal dos relacionamentos existentes entre os termos e conceitos.

Observa-se então que para a concretização do projeto Web Semântica é necessária a utilização intensiva de lógicas computacionais, para que assim possa ser realizada, na “*Camada de Confiança*”, a comprovação de que os aspectos semânticos dos recursos estão descritos de modo consideravelmente adequado, atendendo a todos os requisitos das camadas anteriores e possibilitando um certo grau de confiança das informações.

É importante ressaltar que muitas vezes devido às pressões para o desenvolvimento em curto prazo de novas tecnologias, buscando atender demandas de mercado, as metodologias propostas a partir da área de Ciência da Computação podem não abranger as devidas preocupações quanto às possíveis consequências que tais tecnologias possam acarretar, subestimando uma das principais preocupações da área de Ciência da Informação, quanto à indexação de informações.

Conforme relata [Fujita](#) (2003, p.180):

Uma das principais preocupações dos pesquisadores em indexação é a rápida evolução das técnicas de recuperação automática, acarretando o aumento da responsabilidade do indexador na determinação do assunto do documento. Novas formas de recuperação exigem maior aprofundamento teórico do indexador para que se evite o risco de uma prática descompromissada com a representação do contexto do documento e do sistema de recuperação de informação.

Os profissionais da informação têm grandes responsabilidades dentro deste contexto, desempenhando um importante papel de agentes sociais, de modo que um de seus principais desafios é favorecer a diminuição das desigualdades no acesso à informação, a partir da utilização das novas tecnologias de acordo com princípios éticos que respeitem as especificidades, subjetividades e os valores de cada indivíduo ou comunidade.

Segundo [Fernández-Molina e Guimarães](#) (2002), quando tratamos das questões éticas que envolvem os profissionais da informação e as novas tecnologias, em geral levamos os conceitos e aplicações dentro de um nível restrito, deixando desta forma de gerar análises mais abrangentes sobre seu uso em uma esfera global. Assim, é evidente que o projeto Web Semântica, a partir da criação de categorias para a classificação dos recursos informacionais disponíveis no ambiente Web, trará no bojo de seu desenvolvimento novos dilemas éticos em uma escala global.

[Buchanan](#) (1999) afirma que os profissionais da informação têm a obrigação moral de responder aos novos dilemas éticos emergentes e de se esforçar para balancear as limitações práticas e tecnológicas de fornecimento e uso da informação.

Conclusões

A partir da análise da literatura apresentada, pode-se constatar que o projeto Web Semântica constitui uma evolução no modo como as informações são organizadas no ambiente Web, projetado com o intuito de possibilitar a incorporação de aspectos semânticos aos dados, favorecendo a contextualização das informações de forma automatizada, de acordo com o “contexto” no qual os dados estão inseridos e os critérios da busca realizados.

Observou-se que considerando os pressupostos da área de Ciência da Computação, é possível justificar a origem da expressão “*Web Semântica*”, devido ao fato que os “*motores de busca*” e as páginas Web tradicionais não levam em consideração os aspectos semânticos inerentes aos recursos informacionais. Fato que não ocorre caso um processo de recuperação de informação no ambiente Web seja analisado sob o prisma de um usuário que inegavelmente irá utilizar-se de aspectos semânticos para localizar as informações que procura.

Verificou-se que os conceitos e tecnologias subjacentes ao projeto Web Semântica podem ser considerados como uma renovação ou desdobramento dos tradicionais métodos representação, organização e recuperação de informações, apontando a possibilidade de contribuições da área de Ciência de Informação, devido ao seu embasamento teórico referente a formas de representação e as práticas profissionais identificadas em seu campo de atuação.

Ao mesmo tempo em que devido ao fato do projeto Web Semântica ter despertado o interesse de profissionais das mais variadas áreas do conhecimento, novos conceitos vêm sendo incorporados ou reformulados na área de Ciência da Computação, constituindo um novo paradigma computacional, segundo o qual os aspectos semânticos estão intrínsecos ao conceito de informação.

Portanto, torna-se evidente o caráter interdisciplinar que delinea o corpus teórico do projeto Web Semântica, englobando essencialmente áreas como a Ciência da Informação e a Ciência da Computação, entre outras, apresentando-se como um campo fértil para pesquisas. Assim, esforços interdisciplinares são necessários para possibilitar o desenvolvimento de soluções multidisciplinares, respeitando as especificidades de cada área do conhecimento e tendo como objetivo comum auxiliar na evolução do conhecimento humano de forma integral.

Verificou-se também que a partir dos estudos relacionados ao projeto Web Semântica, torna-se possível uma aproximação do objeto de estudo da Ciência da Informação com a área da Ciência da Computação, pois a partir do momento que os pesquisadores da área de Ciência da Computação empenham-se em desenvolver mecanismos que possibilitem descrever os aspectos semânticos inerentes aos recursos informacionais, pode-se considerar que estão trabalhando com o mesmo conceito de informação utilizado pelos pesquisadores da área de Ciência da Informação. **Barreto** (2002, p.73) já sinalizava para tal fato, quando em seu artigo “*A condição da Informação*” utilizou a expressão “*tecnologista da informação*” ao se referir a Tim Berners-Lee, pesquisador diretamente ligado à área de Ciência da Computação e justamente o grande precursor do projeto Web Semântica.

Deste modo, é possível considerar que o principal objetivo do projeto Web Semântica pode ser identificado pelo próprio slogan do Wide Web Consortium, W3C “*Leading the Web to Its Full Potential*”, Conduzir a Web para o Seu Potencial Máximo, entretanto, verifica-se poucas possibilidades de sua implantação de modo integral, ao menos em curto prazo, constituindo-se assim como um ideal abstrato muito mais do que uma possibilidade real a ser concretizada. Tal afirmação baseia-se no fato de que a concretização de uma Web Semântica global depende de uma série outros fatores sociais, econômicos, políticos e culturais, que vão além do desenvolvimento de novas tecnologias e padrões de representação.

Conforme afirma **Codina** (2003), atualmente a expressão Web Semântica representa o rótulo de uma aspiração, um projeto de como seria idealmente o ambiente Web. É inegável o fato de que as novas tecnologias têm provocado avanços consideráveis e verdadeiras revoluções em vários setores da sociedade, sendo evidente que o projeto Web Semântica traz no bojo de suas inovações tecnológicas uma série de avanços que podem possibilitar melhorias significativas nos processos de organização e recuperação de informações em ambiente digital.

No entanto, observa-se também que tais tecnologias criam/agravam uma série de problemas sociais,

econômicos, políticos e culturais, que em sua maioria ainda não foram enfrentados, ou muitas vezes sequer identificados, tanto no campo teórico quanto prático. Constituindo assim, um dos grandes desafios a ser superado por uma sociedade que almeja ostentar o título de “*Sociedade da Informação*”.

Notas

[1] O W3C consiste de um consórcio mundial liderado por Tim Berners-Lee que reúne empresas, instituições acadêmicas, profissionais e cientistas, com o intuito de padronizar novas tecnologias que possibilitem estender gradativamente as funcionalidades do ambiente Web.

[2] <http://www.w3.org/2001/sw/Europe>

[3] <http://www.w3.org/2001/sw/Europe/reports/thes>

Referências Bibliográficas

ALMEIDA, M. B.; BAX, M. P. Uma visão geral sobre ontologias: pesquisa sobre definição, tipos, aplicações, métodos de avaliação e de construção. *Ciência da Informação*, Brasília, v. 32, n. 3, p. 7-20, set/dez. 2003.

ALVARENGA, L. A teoria do conceito revisada em conexão com ontologias e metadados no contexto das Bibliotecas tradicionais e digitais. *DataGramZero*, Rio de Janeiro, v. 2, n. 6, dez. 2001. Disponível em: <http://www.dgz.org.br/dez01/Art_05.htm> Acesso em: 15 set. 2005.

ARANO, S.; CODINA, L. La estructura conceptual de los tesauros en el entorno digital: nuevas posibilidades en la representación de la información y el control terminológico. In: *JORNADES CATALANES D'INFORMACIÓ I DOCUMENTACIÓ*, 9. Barcelona, 2004.

BARRETO, A. de A. A condição da informação. *São Paulo em Perspectiva*, v.16, n.3, p.67-74, jul/set., 2002.

BERNERS-LEE, T. *Semantic Web road map*. 1998. Disponível em: <<http://www.w3.org/DesignIssues/Semantic.html>>. Acesso em: 18 set. 2005.

BERNERS-LEE, T. ;HENDER, J. ;LASSILA, O. The semantic Web: a new form of Web content that is meaningful to computers will unleash a revolution of new possibilities. *Scientific American*, New York, may. 2001.

BERNERS-LEE, T. *Information Management: a proposal*. CERN, Genebra, mar. 1989. Disponível em: <<http://www.w3.org/History/1989/proposal.html>>. Acesso em: 20 Set. 2005.

BERNERS-LEE, T. *Semantic Web - XML2000*. 2000. Disponível em: <<http://www.w3.org/2000/Talks/1206-xml2k-tbl>>. Acesso em: 18 set. 2005.

BERNERS-LEE, T. *Semantic Web Concepts*. 2005. Disponível em: Disponível em: <<http://www.w3.org/2005/Talks/0517-boit-tbl>>. Acesso em: 18 out. 2005.

BERNERS-LEE, T. *The Semantic Web*. 2002. Disponível em: <<http://www.w3.org/2002/Talks/04-sweb-sloan/Overview.html>>. Acesso em: 20 set. 2005.

BERNERS-LEE, T. *Weaving the Web*. San Francisco: Harper, 1999.

BERNERS-LEE, T.; MASINTER, L.; MCCAHILL, M. Uniform Resource Locators (URL). *RFC Sourcebook*, n.1738, mar. 1994. Disponível em: <<http://www.networksorcery.com/enp/rfc/rfc1738.txt>>. Acesso em: 20 Set. 2005.

BINDING, C.; TUDHOPE, D. KOS at your Service: Programmatic Access to Knowledge Organisation Systems. *Journal of Digital Information*, Texas, v. 4, no. 4, 2004.

BRICKLEY, D. et al. RDF vocabulary description language 1.0: RDF schema. . *W3C recommendation*, 10 Fev. 2004. Disponível em: <<http://www.w3.org/TR/2004/REC-rdf-schema-20040210>>. Acesso em: 16 set. 2005.

BUCHANAN, E.A. An overview of information ethics issues in a world-wide context. *Ethics and Information Technology*, Berlin, v.1, no. 3, p. 193-201, sep. 1999.

CAMPOS, M. L. de A. *Modelização de domínios de conhecimento: uma investigação de princípios fundamentais*. Ciência da informação, Brasília, v. 33, n. 1, p. 22-32, jan./abr. 2004.

CAPURRO, R. Epistemologia e ciência da informação. IN: ENCONTRO NACIONAL DE PESQUISA EM CIÊNCIA DA INFORMAÇÃO - ENANCIB, 5., 2003. Belo Horizonte. *Anais...* Belo Horizonte: UFMG, 2003.

CEEInf - Comissão de Especialistas de Ensino de Computação e Informática. *Diretrizes Curriculares de cursos da área de Computação e Informática*. Brasília: CEEInf/MEC, 1999. Disponível em: <<http://www.inf.ufrgs.br/mec/ceeinf/diretrizes.html>>. Acesso em: 10 Jan 2006.

CLARK, K. G. SPARQL Protocol for RDF. *W3C Working Draft*, 14 set.2005. Disponível em: <<http://www.w3.org/TR/2005/WD-rdf-sparql-protocol-20050914>>. Acesso em: 18 set. 2005.

CODINA, L. La web semántica : una visión crítica. *El profesional de la información*, Barcelona, v. 12, n. 2, p. 149-152. 2003.

DACONTA, M. C.; OBRST, L. J.; SMITH, K. T. *The Semantic Web: a guide to the Future of XML, Web Services, and Knowledge Management*. Indianápolis: Wiley Publishing, 2003. 279 p.

DAUM, B.; MERTEN, U. *Arquitetura de sistemas com XML*. Rio de Janeiro: Campus, 2002.

DING, Y.; FOO, S. Ontology research and development, Part 1 - A review of ontology generation. *Journal of Information Science*, v. 28, no. 2, p. 123-136. 2002.

FERNÁNDEZ MOLINA, J.C.; GUIMARÃES, J. A Ethical aspects of knowledge organization and representation in the digital environment: their articulation in professional codes of ethics. IN: López-Huertas, M.M., (Ed.). *Challenges in knowledge representation and organization for the 21st century: integration of knowledge across boundaries*. Wurzburg: ERGON-Verlag, 2002. p. 487-492.

FERNEDA, E. *Recuperação de Informação: análise sobre a contribuição da Ciência da Computação para a Ciência de Informação*. 2003. 137 f. Tese (Doutorado em Ciência da Informação) – Escola de Comunicações e Artes, Universidade de São Paulo, São Paulo, 2003.

FERNEDA, E. *Recuperação de Informação: análise sobre a contribuição da Ciência da Computação para a Ciência de Informação*. 2003. 137 f. Tese (Doutorado em Ciências da Comunicação) – Escola de Comunicações e Artes, Universidade de São Paulo, São Paulo, 2003.

FIELDING, R. Relative Uniform Resource Locators. *RFC Sourcebook*, no.1808, jun. 1995. Disponível em: <<http://www.networksorcery.com/enp/rfc/rfc1808.txt>>. Acesso em: 20 Set. 2005.

FUJITA, M. S. L. *A Leitura Documentária do indexador: aspectos cognitivos e lingüísticos influentes na formação do leitor profissional*. 2003. 321f. Tese (Livre-Docência em Análise Documentária e Linguagens Documentárias Alfabéticas) – Faculdade de Filosofia e Ciências,

Universidade Estadual Paulista, Marília.

GARCIA JIMÉNEZ, A. Instrumentos de representación del conocimiento: tesauros versus ontologías. *Anales de documentacion*, Mucia, n.7, p.79-95, 2004. Disponível em: <<http://www.um.es/fccd/anales/ad07/ad0706.pdf>>. Acesso em: 10 jan 2006.

GOLBECK et al. The National Cancer Institute's Thesaurus and Ontology. *Journal of Web Semantics*, v.1, no.1, p. 75-80, 2003.

GROSOF, B. et al. Description logic programs: Combining logic programs with description logics. In: *Proc. of WWW 2003*, Budapest, mai 2003, p. 48-57. Disponível em: <<http://www.cs.man.ac.uk/~horrocks/Publications/download/2003/p117-grosop.pdf>>. Acesso em: 16 set. 2005.

HORROCKS, I., et al. Semantic web architecture: stack or two towers? In: FAGES, F.; SOLIMAN, S. (Ed.). *Principles and Practice of Semantic Web Reasoning*, (PPSWR 2005), no. 3703 in LNCS, 2005. p. 37-41.

KOIVUNEN, M. R.; MILLER, E. *W3C Semantic Web Activity*. 2001. Disponível em: <<http://www.w3.org/2001/12/semweb-fin/w3csw>>. Acesso em: 10 out. 2005.

KRISHNAMURTHY, B.; REXFORD, J. *Redes para a Web*. Rio de Janeiro: Campus, 2001.

KRUK, S. R. et al. *JeromeDL reconnecting digital libraries and the semantic Web*. Disponível em: <http://www.marcont.org/marcont/pdf/www2005_jeromedl.pdf>. Acesso em: 10 out. 2005.

LANCASTER, F.W. *Indexação e resumo: teoria e prática*. Brasília: Briquet de Lemos, 2004. 452 p.

LE COADIC, Y-F. *A Ciência da informação*. Brasília: Briquet de Lemos, 1996.

MCGUINNESS, D. L.; HARMELEN, F. OWL Web Ontology Language Overview. *W3C Recommendation*, 10 Feb. 2004. Disponível em: <<http://www.w3.org/TR/2004/REC-owl-features-20040210>>. Acesso em: 16 set. 2005.

MEY, E. S. A. Biblioteca Alexandrina. *Revista Digital de Biblioteconomia e Ciência da Informação*, Campinas, v .1, n.2, p. 71-91. 2004. Disponível em: <<http://server01.bc.unicamp.br/seer/ojs/index.php>> Acesso em: 10 de out. 2005.

MILES, A.; BRICKLEY, D. SKOS Core Guide. *W3C Working Draft*, 02 Nov. 2005. Disponível em: <<http://www.w3.org/TR/swbp-skos-core-guide>>. Acesso em: 15 jan. 2006.

MOREIRA, A.; OLIVEIRA, A. P. Contribuição da terminologia na modelagem de sistemas computacionais. *DataGramZero*, Rio de Janeiro, v. 6, n. 5, out. 2005. Disponível em: <http://www.dgz.org.br/out05/Art_01.htm>. Acesso em: 20 set. 2005.

NAKAMURA, E. T.; GEUS, P. L. *Segurança de redes em ambientes corporativos*. São Paulo: Futura, 2003.

PÉREZ AGÜERA, J. R. Automatización de tesauros y su utilización en la Web semántica. *textos universitarios de biblioteconomía i documentació*, n.13, 2004. Disponível em: <<http://www.ub.es/bid/13perez2.htm>>. Acesso em: 15 jan 2006.

PFÜTZENREUTER, E. *Aplicabilidade e desempenho do protocolo de transporte SCTP*. 2004. 119 f. Dissertação (Mestrado em Ciência da Computação) - Universidade Federal de Santa Catarina, Florianópolis, 2004.

PINCEMIN, B. Thésaurus documentaires et ontologies: divergences et ressemblances. *Journée d'étude Web Sémantique*. Paris, 2003. Disponível em: <http://www-lli.univ-paris13.fr/membres/biblio/1195_pincemin_ws_0410.pdf>. Acesso em: 10 jan. 2006.

PRUD'HOMMEAUX, E.; SEABORNE, A. SPARQL query language for RDF. *W3C working draft*, 21 July 2005. Disponível em: <<http://www.w3.org/TR/2005/WD-rdf-sparql-query-20050721>>. Acesso em: 16 set. 2005.

QIN, J.; PALING, S. Converting a controlled vocabulary into an ontology: the case of GEM. *Information Research*, v. 6, no. 2, 2001. Disponível em: <<http://informationr.net/ir/6-2/paper94.html>>. Acesso em: 10 jan 2006.

RAMALHO, R. A. S. *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.

RAMALHO, R. A. S.; VIDOTTI, S. A. B. G. ; FUJITA, M. S. L. Bibliotecas Digitais na era da Web Semântica: reflexões no âmbito da gestão de conteúdos informacionais. In: FUJITA, M. S. L. (Org.). *A dimensão social da biblioteca digital na organização e acesso ao conhecimento: aspectos teóricos e aplicados*. São Paulo:USP/IBICT, 2005, v. 2, p. 91-113.

RAYWARD, W. B. The origins of information science and the International Institute of Bibliography/International Federation for Information and Documentation (FID). *Journal of the American Society for Information Science*, no. 48, p. 289-300, Apr. 1997. Disponível em: <<http://alexia.lis.uiuc.edu/~wrayward/otlet/OriginsofInfoSci.htm>>. Acesso em: 8 jan. 2006.

SARACEVIC, T. Ciência da Informação: origem, evolução e relações. *Perspectivas em Ciência da Informação*, Belo Horizonte, v. 1, n. 1, p. 41-62, jan/jun. 1996.

SBC, Sociedade Brasileira de Computação. *Currículo de referência da SBC para cursos de graduação em Computação*. Campinas, 2003. Disponível em: <www.sbc.org.br/educacao>. Acesso em: 10 jan. 2006.

SHANNON, C ; WEAVER, W. *The Mathematical theory of communication*. Urbana: University of Illinois Press, 1949. Disponível em: <<http://cm.bell-labs.com/cm/ms/what/shannonday/shannon1948.pdf>>. Acesso em: 18 ago. 2005.

SHANNON, C. E. A Mathematical theory of communication. *The Bell System Technical Journal*, v. 27, p.379-423, 623-656, July/Oct. 1948.

SOERGEL, D. Thesauri and ontologies in digital libraries: tutorial. In: *EUROPEAN CONFERENCE ON DIGITAL LIBRARIES*, (ECDL), 2002, Roma. Disponível em: <http://www.dsoergel.com/cv/B63_rome.pdf>. Acesso em: 10 jan. 2006.

SOUZA, R. R.; ALVARENGA, L. 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.

SYNAK, M.; KRUK, S. R. MarcOnt initiative the ontology for the librarian world. In: *EUROPEAN SEMANTIC WEB CONFERENCE ESWC*, 2., 2005, Heraklion, Grécia. *Proceedings...* Heraklion: 2005.

VRANDEČIĆ, D. et al. *DLP: an introduction*. 2005. Disponível em: <<http://logic.aifb.uni-karlsruhe.de/download/dlpintro.pdf>>. Acesso em: 18 set. 2005.

Sobre os autores / About the Author:

Rogério Aparecido de Sá Ramalho

ramalho@marilia.unesp.br

Doutorando do Programa de Pós-Graduação em Ciência da Informação, UNESP, Campus de Marília.

Silvana Aparecida Borsetti Gregorio Vidotti

vidotti@marilia.unesp.br

Professora do Programa de Pós-Graduação em Ciência da Informação, UNESP, Campus de Marília.

Mariângela Spotti Lopes Fujita

fujita@marilia.unesp.br

Professora do Programa de Pós-Graduação em Ciência da Informação, UNESP, Campus de Marília.