

AS POTENCIALIDADES DA WEB SEMÂNTICA PARA A CIÊNCIA DA INFORMAÇÃO

Resumo

Este artigo aborda a Web Semântica em termos conceituais e explora as suas potencialidades para a Ciência da Informação. Através de uma revisão de literatura, referenciada nos principais autores da área, analisa a interface entre a Tecnologia da Informação e as técnicas utilizadas em Biblioteconomia e Documentação. Avalia os resultados e a relevância da interação entre estas áreas diversas do conhecimento, em particular para a Ciência da Informação. A investigação permite concluir que a Web Semântica consiste em uma rede altamente estruturada de informação que tem como objetivo processar os conteúdos publicados e provenientes de diversas fontes. A semântica propiciada pela Web favorece o entendimento dos conteúdos pelos sistemas digitais, apresentando resultados mais condizentes com as necessidades informacionais dos usuários. Assim, a Web Semântica pode favorecer o avanço significativo das técnicas de indexação, armazenamento e recuperação da informação em rede, contribuindo notadamente para a evolução da Ciência da Informação.

Palavras-chave: Ciência da Informação. Web Semântica. Sistemas Digitais de Informação.

Ana Paula de Oliveira Villalobos

Doutora em Educação
(FACED/UFBA), Professora da
ICI/UFBA.
villalob@ufba.br

Daniel Cerqueira Silva

Bacharel em Biblioteconomia
(ICI/UFBA). Universidade Estadual
de Feira de Santana.
falecomdaniel@yahoo.com.br

THE POTENCIAL OF SEMANTIC WEB FOR THE INFORMATION SCIENCE

Abstract

This article discusses the concept of Semantic Web and explores its potential for Information Science. Through a literature review, referenced in the main authors in the field, analyzes the interface between information technology and the techniques used in Librarianship and Documentation. Evaluates the results and relevance of the interaction between these different areas of knowledge, particularly for Information Science. The research concludes that the Semantic Web consists of a highly structured network of information that would process the contents published by different sources. The semantics provided by the Web facilitates the understanding of content by digital systems, with results more consistent with the information needs of users. Thus, the Semantic Web may encourage the improvement of techniques of indexing, storage and retrieval of networked information, contributing remarkably to the development of information science.

Key-words: Information Science. Semantic Web. Digital Systems of Information.

INTRODUÇÃO

Recentemente tem se percebido um notável aumento do acesso à internet. Além do aumento quantitativo, há também um crescente aumento qualitativo, já que as conexões de alta velocidade estão, cada vez, mais acessíveis. Diferentemente de alguns tempos atrás, agora os usuários não estão apenas sendo espectadores, mas, acima de tudo, produtores. Esta condição, aliada à necessidade de obtenção ágil de informações, faz da internet um repositório de dados que nem sempre condiz com as necessidades de seus usuários. Uma pesquisa na internet pode ser uma tarefa exaustiva, a depender de como e do que se pretende pesquisar. Diante deste cenário emerge a ideia de criar uma nova Web, com uma padronização das informações disponibilizadas, onde os computadores possam processar informações provenientes de diversas fontes, propiciando uma rede de informações realmente estruturada.

A denominação para este novo estágio da internet é Web Semântica e tem como objetivo não apenas exibir as páginas na internet, mas também capturar e processar todos os conteúdos publicados. Ou seja, a página criada pela Web Semântica tem a capacidade de propiciar o entendimento pelas máquinas dos conteúdos exibidos.

Esta nova tendência certamente trará impactos para a forma de utilizar a internet, repercutindo em vários setores profissionais e na sociedade. Mas, em particular, existirá uma atenção especial para os profissionais da Ciência da Informação que não simplesmente lidam com a informação, mas com a maneira de processá-la e potencializá-la. Entretanto, a tecnologia por si só não será suficiente para um avanço nas técnicas de pesquisa na internet, é necessária uma representação adequada dos documentos, e para que isso ocorra é imprescindível criar sistemas de indexação eficazes, de forma que a recuperação das informações neles contidas seja a mais significativa possível.

A Web Semântica possibilitará transformações para a Ciência da Informação, havendo um processo de simbiose entre estas áreas, representando a junção real da Tecnologia da Informação com as técnicas utilizadas em Biblioteconomia e Documentação. Este artigo

focaliza justamente este ponto, permitindo a avaliação dos resultados e da relevância deste processo para cada uma destas áreas do conhecimento, em particular para a Ciência da Informação.

1 AS TECNOLOGIAS ASSOCIADAS À WEB 2.0

O termo Web 2.0 foi criado em 2004 pela empresa O'Reilly Media¹ para se referir à uma segunda geração de comunidades e serviços, tendo a Web como plataforma e incorporando [wikis](#), aplicativos baseados em [folksonomia](#), [redes sociais](#) e [Tecnologia da Informação](#). Conforme O'Reilly (2004),

a Web 2.0 é a mudança para uma internet como plataforma, e um entendimento das regras para obter sucesso nesta nova plataforma. Entre outras, a regra mais importante é desenvolver aplicativos que aproveitem os efeitos de rede para se tornarem melhores quanto mais são usados pelas pessoas, aproveitando a inteligência coletiva.

Na Web 2.0 os programas são abertos, ou livres, uma parte do programa pode ser modificada ou utilizada para se criar outro programa. As aplicações que permitem esta flexibilidade são conhecidas como Application Programming Interface (API) – Interface de Programação de Aplicativos. É uma lista de [rotinas](#) e padrões de um [software](#) para a utilização das suas funcionalidades por programas aplicativos. Ainda, na Web 2.0 cada usuário é um servidor de arquivos e os arquivos são trocados diretamente entre eles, caracterizando uma arquitetura P2P, *Peer to Peer*, ou entre pares. Neste tipo de arquitetura os computadores conectados em rede podem atuar tanto como clientes, solicitando aplicações, quanto como servidores.

A criação das páginas na internet, baseadas nas linguagens de programação como XML e RDF, possibilitou um impulso altamente significativo para que as tecnologias da

chamada Web 2.0 pudessem, realmente, participar do aperfeiçoamento dos processos de organização e transferência da informação. Na tentativa de melhor compreender as relações entre a tecnologia da Web 2.0 e a Ciência da Informação, é pertinente ressaltar que a maioria dos elementos que compõe esta tecnologia está sendo amplamente aplicada na concepção de bibliotecas digitais. Prova disso é que atualmente já é possível utilizar alguns destes recursos, inclusive muitos deles já se encontram disponíveis em alguns sites. São eles:

Blog: é um diário ou jornal disponível na internet (em geral, é de acesso público). Por ser atualizado diariamente, apresenta a vantagem de manter os seus usuários permanentemente atualizados sobre determinado assunto. Tem duas grandes vantagens em relação aos sites: o fato de que o autor do *blog* não necessita saber construir páginas para a internet e a possibilidade de permitir que se faça comentários sobre os conteúdos publicados. Os *blogs* podem, ainda, serem criados e geridos por professores (individualmente ou em grupo), por alunos (individualmente, por grupos de trabalho, ou por turmas) e até simultaneamente por professores e os seus alunos;

Wiki: é um sistema de hipertexto, cujo conteúdo pode ser não apenas lido pelo usuário, mas também modificado *online*. Um dos principais exemplos desta ferramenta é a Wikipédia (www.wikipedia.com), onde as informações do site são postadas pelos próprios usuários, propiciando ampliar a interatividade e a qualidade da informação. Conforme Lèvy (2000), a existência de uma internet colaborativa possibilita a disseminação da inteligência coletiva;

RSS: é um serviço na web, semelhante às chamadas de notícias, que contém o título, um pequeno texto e o *link* para a página principal, ou seja, permite que o usuário receba as últimas atualizações de vários sites, sem precisar se dirigir a cada um deles,

¹ A O'Reilly Media (antes chamada de O'Reilly & Associates) é uma companhia de [mídia](#) (editora) americana criada por [Tim O'Reilly](#), que publica livros e *websites* e organiza conferências sobre temas de [informática](#).

possibilitando também que vários sites, ou pelo menos as atualizações deles, possam ser vistos em uma única página;

Tagging: indexação coletiva é uma forma de indexação livre, na qual o cliente ordena os conteúdos através de descritores (palavras-chave) com o auxílio de diferentes *softwares*;

Mash-up: web sites ou aplicações que combinam conteúdo de múltiplas fontes, mas que aparecem integradas para o usuário.

Entre as aplicações associadas à Web 2.0, destaca-se a folksonomia. Segundo Catarino e Baptista (2007), folksonomia é um neologismo criado a partir da combinação entre *folk* (pessoa) e *taxonomy* (taxonomia), resultante da codificação de informações pessoais realizada em ambientes sociais. Assim, esta ferramenta possibilita ao usuário indexar o conteúdo de seu interesse para posterior recuperação. Neste processo, o usuário utiliza as *Tags*, etiquetas ou mesmo palavras-chave empregadas para a representação da informação.

A utilização destes recursos possibilita vantagens imediatas, tais como: respostas mais rápidas nas pesquisas; aumento da especificidade das respostas; disponibilização gratuita dos serviços na Web; possibilidade de incrementar o número de usuários; favorece a inteligência coletiva e permite a renovação mais rápida dos conteúdos.

É visível a mudança de paradigma estabelecida entre os serviços oferecidos e atividades desenvolvidas pelas bibliotecas tradicionais e o novo conceito advindo da Web 2.0. Através das tecnologias baseadas na interatividade e na participação, a exploração da inteligência coletiva transforma a busca em localização, o acesso em compartilhamento e a exposição em debate. Maness (2006) conclui que a Web 2.0 não está voltada para buscar e sim para localizar, não em acessar e sim em compartilhar informações.

3 DA WEB 2.0 PARA A WEB SEMÂNTICA

A internet é um dos grandes acontecimentos de nosso século, tornando-se mais popular a partir dos anos 90 do século passado. Diante da sua capacidade de evolução e da ampla utilização das suas possibilidades, a internet tem vivenciado um crescimento acentuado nos anos recentes. O desenvolvimento de infraestrutura de redes e servidores aliado ao uso maciço sem critérios fez da internet uma avalanche sem controle, onde diversos tipos de informações de relevâncias variadas estão sendo disponibilizadas e consultadas a todo o momento. Monteiro (2006) afirma que no ciberespaço não há centro de significância estruturado, hierarquizado, linear ou instrumentos de organização do conhecimento que reproduzam o modelo de significância, sentido único e referência fixa.

Estas consequências da utilização sem critérios da web, acrescidas à necessidade de se ter informações mais precisas em espaço de tempo mais curto, despertaram a necessidade da elaboração de uma nova filosofia. Neste enfoque os dados deixam de ser combinados aleatoriamente e se tornam informações confiáveis e estruturadas. A partir daí foi criado o termo Web Semântica para descrever este novo estágio da internet, onde agora haverá uma preocupação mais evidente em relação aos critérios e a metodologia de recuperação da informação.

A Web Semântica é uma extensão da atual Web, na qual a informação apresenta um significado bem definido, permitindo melhor interação entre os computadores e as pessoas. Com estas palavras, Berners-Lee (2001) define os projetos de seu grupo de trabalho no World Wide Web Consortium (W3C) para modificar a *Web*. O W3C representa o Consórcio de empresas, profissionais, cientistas e instituições acadêmicas responsável pela criação de padrões tecnológicos que regulam a World Wide Web.

Para a viabilização da Web Semântica, os dados devem ser acompanhados de descrições baseadas nos padrões W3C. Em vista da presença desta camada adicional de dados será possível a ação de programas especiais, os agentes, para processar as informações. Os agentes são considerados assistentes de tarefa, programas que empregam

técnicas de inteligência artificial para auxiliar o usuário na realização de uma dada tarefa (SOUZA; ALVARENGA, 2004).

A Web 3.0, assim chamada inicialmente pelo jornalista John Markoff em 2006, representa uma terceira geração de serviços baseados na Internet. Estes serviços se apoiam em um tipo de Web que se baseia na maior capacidade do *software* em interpretar os conteúdos em rede, permitindo resultados mais objetivos e personalizados de pesquisa. Esta é uma das características da Web Semântica, uma proposta de Tim Berners-Lee identificável com a terceira geração da Web, em que os conteúdos podem ser registrados de maneira a serem compreendidos, interpretados e processados por determinados agentes de *software*, os quais passam a pesquisar, partilhar e integrar a informação de uma forma mais eficiente.

A Web Semântica é associada à [Web 3.0](#), como um próximo movimento da Internet depois da Web 2.0 que já inicia seu crescimento. A ideia da Web Semântica surgiu em 2001, quando [Tim Berners-Lee](#), [James Hendler](#) e [Ora Lassila](#) publicaram um artigo na revista *Scientific American*, intitulado: *Web Semântica: um novo formato de conteúdo para a Web que tem significado para computadores e vai iniciar uma revolução de novas possibilidades*.

O objetivo principal da Web semântica não é, pelo menos para já, treinar as máquinas para que se comportem como pessoas, mas sim desenvolver tecnologias e linguagens que tornem a informação legível para as máquinas. A finalidade passa pelo desenvolvimento de um modelo tecnológico que permita a partilha global de conhecimento assistido por máquinas (W3C, 2001). A integração das linguagens ou tecnologias [eXtensible Markup Language \(XML\)](#), [Resource Description Framework \(RDF\)](#), arquitecturas de metadados, [ontologias](#), agentes computacionais, entre outras, favorecerá o aparecimento de serviços Web que garantam a interoperabilidade e cooperação.

A Web Semântica consiste em uma rede onde os documentos podem ser representados de forma estruturada, possibilitando uma indexação mais eficaz e apresentando resultados mais condizentes com as necessidades informacionais dos usuários. Conforme Berners-Lee, Hendler e Lassila (2001), os computadores necessitam ter acesso a

coleções estruturadas de informações (dados e metadados) e de conjuntos de regras de inferência que ajudem no processo de dedução automática para que seja administrado o raciocínio automatizado, ou seja, a representação do conhecimento.

Os mecanismos de busca utilizados atualmente agem de forma genérica e apresentam uma baixa precisão na recuperação da informação. Eles aproveitam a palavra-chave escolhida e realizam a busca através de *search engine* (motores de busca) que apenas pesquisam o termo da mesma forma que ele foi sugerido sem contextualização alguma. Consequentemente, isso faz com que estes resultados apresentem baixas taxas de precisão, relevância e revocação, o que torna imprescindível a filtragem das informações realmente apropriadas.

Atualmente as linguagens utilizadas para a criação de páginas na internet são baseadas na representação dos caracteres existentes em cada um dos documentos, sendo possível apenas fazer referência aos mesmos considerando a forma como eles são exibidos. Na afirmação de Pickler (2007) para recuperar informações disponíveis na Web, os mecanismos de busca valem-se da indexação de palavras em linguagem natural presentes nos próprios documentos, embora cada ferramenta opere de acordo com princípios, por vezes, divergentes.

A Web Semântica baseia-se na utilização da linguagem XML² que possibilita a utilização de categorias semânticas nos dados que serão exibidos na Web. Isso permitirá pesquisar levando em consideração o conteúdo semântico e os significados contextuais dos documentos. Conforme Oliveira (2002, p. 5),

Ao contrário da HTML que através das marcas pré-definidas gerenciam os textos marcados e controlam sua representação estabelecendo ligações entre os documentos, a linguagem XML marca semanticamente um documento. XML consiste em padrão utilizado para marcação de documentos que contém informações estruturadas, ou seja, documentos

² Versão mais simples do padrão SGML, padrão universal para a escrita de documentos no formato de hipertexto adotado nos sites da Internet.

que contém uma estrutura clara e precisa da informação armazenada e obtida com XML. Esta estruturação define e separa claramente conteúdo, significado e apresentação. Assim os documentos em XML podem ser indexados com maior precisão que as páginas planas escritas em HTML.

A Web Semântica tem uma finalidade clara, que é simplesmente melhorar a sistemática da recuperação da informação através de técnicas avançadas de representação da informação. Enquanto a Web atual utiliza um mecanismo de representação de caracteres voltado para o entendimento dos usuários, a Web Semântica promove a compreensão da informação através das máquinas, quer dizer, a Web Semântica terá a vantagem de estruturar os dados encontrados nos sites para que eles, antes de serem entendidos pelos usuários sejam entendidos pelo próprio sistema. Desta forma, ao se pesquisar termos polissêmicos não existirá o risco de se recuperar significados diferentes daquele que seja de interesse do usuário.

Podemos esperar que a *Web* tenha grande melhoria dos índices de revocação e precisão no atendimento das necessidades de informação, porque a semântica embutida nos documentos permitirá aos dispositivos de recuperação evitar os problemas comuns de polissemia e sinonímia, além de considerar as informações em seus contextos de significado (SOUZA; ALVARENGA, 2004, p.13).

No entanto, não basta apenas a infraestrutura tecnológica. É necessária também uma rede estruturada de conhecimento humano para dar suporte complementar a este processo desenvolvido pelo computador. Ou seja, a vantagem semântica só existirá se os computadores tiverem acesso a coleções padronizadas de informações, as quais possibilitarão acesso ao conteúdo semântico dos dados.

Para atingir tal propósito, é necessária uma padronização de tecnologias, de linguagens e de metadados descritivos, de forma que 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. (SOUZA; ALVARENGA, 2004, p.5).

E é justamente neste contexto que surge a proposta de se combinar as técnicas de Ciência da Informação com as técnicas de busca automática pela internet.

A web pode oferecer ferramentas automatizadas para busca da informação sem tratamento e a Biblioteconomia pode oferecer sua experiência e teoria na organização da informação. Estas comunidades, aliando suas técnicas e tecnologias, podem atrelar recursos poderosos para enfrentar problemas de acesso e manutenção à informação eletrônica de qualidade. (DZIEKANIAK; KIRINUS, 2004, p.18).

Espera-se com isso classificar o conhecimento da Web de forma hierárquica, o que possibilitará a existência de um padrão de compartilhamento, armazenamento e exibição das informações para que a informação seja explorada de forma exclusiva e não ambígua.

4 A RELEVÂNCIA DA WEB SEMÂNTICA PARA A CIÊNCIA DA INFORMAÇÃO

Com a conseqüente necessidade de melhoria da organização da informação na internet, a Web Semântica trará consigo algumas mudanças extremamente significativas. Acompanhando este paradigma, a Ciência da informação não fica alheia a esta transformação. Várias vantagens poderão ser aproveitadas para os procedimentos ligados às ciências documentais, principalmente para a pesquisa, a indexação e a representação eletrônica de documentos. Assim, é possível elencar alguns pontos relevantes que poderão reforçar os benefícios que a tecnologia semântica trará para a Ciência da Informação:

Recuperação da informação: o entendimento da linguagem humana pelas máquinas é um fator primordial que facilitará as técnicas de pesquisa, tanto o entendimento, quanto a forma de aplicar os mecanismos de recuperação da informação não necessitarão de estratégias de busca tão refinadas e específicas para a obtenção de resultados precisos, já que os termos pesquisados poderão ser entendidos sem ambigüidades pelos computadores. A utilização de linguagens de marcação como o XML

e o RDF permitirão que os usuários criem tags³ personalizadas com características singulares, impedindo a duplicação de descritores;

Indexação: a utilização de metadados aumentará a qualidade da informação na internet, possibilitando uma representação mais exata dos dados. Através do significado semântico será possível elaborar *thesaurus*, termos controlados e cabeçalhos de assuntos mais precisos com capacidade de revocação mais elevada. A competência interativa proporcionada por esta nova arquitetura, na qual o usuário pode contribuir com os conteúdos disponibilizados (folksonomia), permite a criação de novos termos ou de uma taxonomia mais simples, ampliando a compreensão e propiciando qualidade aos termos descritores. Este fato já acontece em alguns sites como a *Wikipédia*, *Amazon*, ou *You Tube*, o que comprova que a participação coletiva dos usuários pode também contribuir para a qualidade da indexação;

Processo de referência: através dos mais variados recursos, a Web Semântica possibilitará que as informações sejam disponibilizadas com uma dinâmica mais ágil e precisa do que a atual. Mecanismos *feeds*, páginas em XML serão utilizadas para agilizar a disponibilização da informação e possibilitar que os usuários possam se atualizar de acordo com as áreas do conhecimento da sua preferência. Este processo é simples, os sites disponibilizam os *feeds* e estes por sua vez são rastreados pelos leitores ou agregadores, documentos XML que realizam essa rotina de convergir os temas de interesse do usuário e procurá-los periodicamente em um novo local na internet. Assim como nos serviços de alerta, os *feeds* têm a função de manter os seus usuários atualizados sobre um determinado tema e, ao mesmo tempo, já é realizada a Disseminação Seletiva da Informação;

³ Estruturas de linguagem de marcação que consistem em breves instruções, tendo uma marca de início e outra de fim.

Uso dos metadados: a proposta da adoção de um vocabulário controlado para a internet é um projeto extremamente desafiador. Mas, através da utilização de alguns padrões como o *Dublin Core* essa empreitada torna-se muito mais viável. O uso de metadados sempre foi largamente aplicado para a organização da informação. Conforme Milstead e Feldman (1999), “O termo metadados antecede a Web tendo, aparentemente, sido cunhado por Jack Myers nos anos 60 para descrever arquivos eletrônicos”. A adoção de metadados se intensificou nos anos 80 e atualmente já pode ser considerada uma premissa básica para o ordenamento das informações em meio eletrônico.

A definição de metadados é complexa, vários autores definem características particulares para esse termo. Conforme reforça Ikematu (2001), a definição de metadados tem sido uma tarefa difícil, pois as várias interpretações sobre o assunto estão relacionadas ao estágio da organização dentro da hierarquia evolucionária de gestão do conhecimento. No entanto, os autores convergem ao afirmar que os metadados são dados sobre dados, descritores de atributos de um determinado material, tais como: autoria, data de criação, publicação, características físicas, controle de acesso, palavras-chave, entre outros descritores, os quais objetivam ampliar a possibilidade de recuperação da informação. O formato MARC – Machine Readable Cataloging, ou catalogação legível por computador –, por exemplo, é um tipo de metadados que pode servir como modelo para a criação de novos modelos de metadados, ou ser aperfeiçoado baseado em algum outro modelo de metadados.

A nova tendência já aponta que o uso dos metadados será indispensável para o gerenciamento das informações em meio digital. No caso em particular do padrão *Dublin Core*, trata-se de um conjunto de metadados que tem uma relação muito estreita com a Ciência da informação. O *Dublin Core* foi criado em 1994 por bibliotecários da Online Computer Library Center (OCLC) com a finalidade de fazer a descrição de materiais tradicionais como livros, documentos em papel, iconografias etc.

Com o surgimento da internet o padrão *Dublin Core* foi adaptado quase que exclusivamente, para a representação dos registros eletrônicos. Weibel (1997) descreve este conjunto de metadados, que consiste em 15 (quinze) elementos de metadados equivalentes a uma ficha catalográfica para descrição de recursos de informação. O padrão *Dublin Core* possui as seguintes funcionalidades de acordo com Rocha e outros (2003): simplicidade na descrição dos recursos – pode ser usado por não-catalogadores, autores ou *websiters* sem conhecimento prévio de todas as regras de catalogação; interoperabilidade semântica – promove o entendimento comum dos descritores, ajudando a unificar padrões de conteúdo e a aumentar a possibilidade de interoperabilidade semântica entre as disciplinas; consenso internacional – reconhecimento da cobertura internacional do escopo do recurso; extensibilidade – constitui uma alternativa aos modelos de descrição mais elaborados e caros. Possui flexibilidade e extensibilidade para codificar semânticas mais elaboradas em padrões mais sofisticados.

Uso das ontologias: na literatura atual a terminologia ontologia vem sendo fortemente utilizada em áreas cada vez mais diversas e para diferentes finalidades. Entre as diversas explicações sobre a definição deste termo, uma das mais unânimes é aquela apresentada por Gruber (1993), conforme este autor uma ontologia é uma especificação formal explícita de uma conceitualização compartilhada. Considerando o uso da ontologia na Ciência da Informação, é interessante destacar o grande potencial que esta tecnologia tem na interoperabilidade entre sistemas. Na prática, interoperar significa ter a capacidade de proporcionar interação lógica, uma troca de dados entre os sistemas e os seus usuários de forma mais inteligível, permitindo que uma estrutura uniforme e integrada possibilite que vários usuários compartilhem o mesmo banco de dados. Um dos impasses mais comuns, relativo à gestão da informação, refere-se a falta de padronização ou a falta de uma linguagem uniforme para a indexação e a pesquisa de termos nas bases de dados (PACHECO; KERN, 2001);

Mudança de perfil profissional: com a consolidação da Web Semântica o profissional da informação terá que buscar conhecimentos mais aprofundados relacionados às novas ferramentas de informação e comunicação. Além do domínio técnico sobre representação da informação e de conhecimentos através das formas tradicionais como AACR2, CDD e CDU, este profissional terá que se adaptar a uma nova realidade marcada por um novo ambiente digital que transformará as formas de produção e recuperação dos documentos. Breending (2007) reforça o uso das tecnologias da Web 2.0 nas bibliotecas chamando-as de Library 2.0 – Biblioteca 2.0 – e aponta a necessidade de ir além dos serviços tradicionais, estáticos, assíncronos e observar a interação e colaboração dos usuários. A utilização da tecnologia semântica obrigará que muitos conceitos ligados à Ciência da informação passem por uma readaptação.

5 CONCLUSÃO

Este artigo procurou abordar a relevância da utilização das tecnologias associadas à Web Semântica para a Ciência da informação. A evolução da Internet desde os estágios da Web 1.0 para 2.0 até a Web 3.0 ou Web Semântica, aqui tratada com mais destaque, evidenciam as imbricações concretas entre a Ciência da Informação e as demais áreas do conhecimento, especialmente a Ciência da Computação.

A Web Semântica traz benefícios não apenas operacionais, em termos de ferramentas, que propiciam aprimorar os processos de indexação, armazenamento e recuperação da informação, mas também possibilita uma grande transformação paradigmática e contextual, que envolve toda a sociedade, já que esta tem a necessidade de se apropriar das tecnologias que promovam acesso mais rápido e qualificado à informação.

É possível considerar que esta mudança ainda irá muito além, já que o nível de transformação provocado por esta nova realidade será altamente impactante para todos os

usuários da internet, permitindo que possam ser desenvolvidos novos estudos mais aprofundados que relacionem a Ciência da Informação a esta nova modalidade tecnológica.

As ferramentas e as linguagens aqui citadas (Blog, Wiki, RDF, XML etc.) são interfaces da área da Ciência da Computação, que têm sido amplamente aproveitadas em processos de gerenciamento da informação em sistemas digitais, tanto em linguagem natural como para a criação de vocabulários controlados.

As ontologias e os metadados apresentam-se como parte fundamental para o conhecimento compartilhado e padronizado, servindo como ferramentas interativas, propiciando aos usuários o desfrute da informação semântica. A evolução destas ferramentas deverá estabelecer novos modelos de taxonomias, vocabulários e terminologia, favorecendo o desenvolvimento de novos parâmetros apropriados para uma nova forma de cognição relacionada ao processamento da informação *online*.

Vale a pena destacar também que a ampliação deste universo de pesquisa, cada vez mais ligado à tecnologia, permitirá aproximar mais ainda o já existente diálogo entre a Ciência da Informação e a Ciência da Computação. Pode-se mencionar que a Web Semântica encontra-se na interface entre estes dois campos do conhecimento, proporcionando à Ciência da Informação uma grande oportunidade de receber contribuições significativas na construção e desenvolvimento de métodos orientados à pesquisa em rede.

Como arcabouço de apoio à representação da informação, o binômio Web Semântica e Ciência da Informação já apresenta grande potencialidade para a produção científica, possibilitando uma melhor organização, classificação, transmissão, utilização e descrição da informação através da sua estruturação semântica com condições de atender às necessidades da comunicação humana contemporânea.

Existe, ainda, um aspecto importante a ser ressaltado, relacionado ao papel que o cientista da informação exercerá neste novo ambiente. Este deve procurar proceder a organização da informação na internet, visando associar a práxis tradicional às novas tecnologias semânticas.

O incremento e a melhoria das características dos serviços de informação na internet são responsáveis diretos pela qualidade de vida dos cidadãos. O aproveitamento da tecnologia semântica por parte da Ciência da Informação permitirá que esta fique bem mais próxima de alcançar um de seus grandes objetivos: proporcionar a democratização e a autonomia no acesso à informação via tecnologias digitais. Essa nova forma de explorar o conhecimento na internet colaborará de forma crucial para que a Ciência da Informação consiga ofertar serviços de informação de alta qualidade para as comunidades, contribuindo para a produção de conteúdos inteligentes, que por sua vez refletirão na formação de uma sociedade mais desenvolvida.

Artigo submetido em 10/13/2009 e aceito para publicação em 08/06/2010

REFERÊNCIAS

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

BERNERS-LEE, T., LASSILA, O.; HENDLER, J. The semanticweb. **Scientific American**, May 2001.

Disponível em:

<<http://www.sciam.com/article.cfm?articleID=00048144-10D2-1C70-84A9809EC588EF21>>. Acesso em: abr. 2008.

BLATTMANN, U.; SILVA, F. C. C. da. Colaboração e interação na Web 2.0 e Biblioteca 2.0. **Revista ACB: Biblioteconomia em Santa Catarina, Florianópolis**, v. 12, n. 2, p. 191-215, jul./dez. 2007.

BREEDING, M. We need to go beyond Web 2.0. **Computers in Libraries**, v. 27, n. 5, p. 22-25, May 2007. Disponível em: <<http://www.infotoday.com/cilmag/may07/index.shtml>> Acesso em: abr. 2009.

CATARINO, M. E.; BAPTISTA, A. A. Folksonomia: um novo conceito para a organização dos recursos digitais na Web. **DatagramaZero: Revista de Ciência da Informação**, São Paulo, v. 8, n. 3, jun. 2007.

DZIEKANIAK, G. V.; KIRINUS, J. B. Web semântica. **Encontros Bibli: Revista Eletrônica de Biblioteconomia e Documentação**. Florianópolis, n.18, p.20-39, jul./dez. 2004. Disponível em: <<http://redalyc.uaemex.mx/redalyc/pdf/147/14701803.pdf>>. Acesso em: 20 maio 2009.

73

GÓMEZ-PÉREZ, A.; BENJAMINS, V. R. Overview of knowledge sharing and reusable components: ontologies and problem-solving methods. In: INTERNATIONAL JOINTCONFERENCE ON ARTIFICIAL INTELLIGENCE, 1., 1999, Sweden. **Proceedings...** Sweden: VR Editors. 1999.

GRUBER, T. R. A translation approach to portable ontology specifications. **Knowledge Acquisition**, London. v. 5, i.2, p.199-220, jun. 1993.

GUIMARÃES, F. J. Z. **Utilização de ontologias no domínio B2C**. 2002, 194 f. Dissertação (Mestrado em Informática) - Departamento de informática. Pontifícia Universidade Católica do Rio de Janeiro, Rio de Janeiro.

IANNELA, R.; WAUGH, A. **Metadata**: enabling the Internet. Disponível em: <<http://archive.dstc.edu.au/RDU/reports/CAUSE97>>. Acesso em: maio 2009.

IKEMATU, R. S. Gestão de Metadados: sua evolução na Tecnologia da Informação. **DataGramZero – Revista de Ciência da Informação**, v. 2, n. 6, dez. 2001

LÉVY, P. **A inteligência coletiva**: por uma antropologia do ciberespaço. São Paulo: Loyola, 2000.

MANESS, J. Library 2.0 theory: Web 2.0 and Its Implications for Libraries. **Webology**, v. 3, n. 2, 2006. Disponível em: <<http://www.webology.ir/2006/v3n2/a25.html>>. Acesso em: maio 2009.

MILSTEAD, J.; FELDMAN, S. Metadata: project and standards. **Online**, v. 23, p. 32-40, jan./feb. 1999.

MONTEIRO, S. D. O ciberespaço e os mecanismos de busca: novas máquinas semióticas. **Ciência da Informação**, Brasília, v.1, n.1, p. 31-38, jan./abr. 2006.

NOVELLO, T. C. Ontologias, sistemas baseados em conhecimento e modelos de banco de dados. Disponível em: <http://www.inf.ufrgs.br/~clesio/cmp151/cmp15120021/artigo_taisa.pdf>. Acesso em: 15 maio 2009.

NASCIMENTO, M. S. O. do et.al. A ontologia na ciência da informação. **Revista Digital de Biblioteconomia e Ciência da Informação**, v.5, n.1, 2007. Disponível em: <<http://server01.bc.unicamp.br/seer/ojs/viewarticle.php?id=105>>. Acesso em: 23 maio 2009.

OLIVEIRA, R. M. V. B. Web Semântica: novo desafio para os profissionais da informação. In: SEMINÁRIO NACIONAL DE BIBLIOTECAS UNIVERSTÁRIAS. Disponível em <<http://www.sibi.ufrj.br/snbu/snbu2002/oralpdf/124.a.pdf>>. Acesso em: 19 maio 2009.

O'REILLY, TIM. **O que é Web 2.0**: padrões de design e modelos de negócios para a nova geração de software. Disponível em: <<http://www.cipedya.com/doc/102010>>. Acesso em: abr. 2010.

PACHECO, R. C. dos S.; KERN, V. M. A common ontology for information and knowledge integration in science and technology. **Ciência da Informação**, Brasília, v. 30, n. 3, p. 56-63, set./dez. 2001.

PICKLER, M. E. V. Web Semântica: ontologias como ferramentas de representação do conhecimento. **Perspectivas em Ciência da Informação**, v. 12, p. 65-83, 2007.

SOUZA, M. I. F.; VENDRUSCULO, L. G.; MELO, G. C. Metadados para a descrição de recursos de informação eletrônica: utilização do padrão Dublin Core. **Ciência da Informação**, Brasília, v. 29, n. 1, p. 93-102, jan./abr. 2000.

_____. **Informação para internet: uso de metadados e o padrão Dublin Core para catalogação de recursos eletrônicos na Embrapa**. Disponível em :
<<http://dici.ibict.br/archive/00000702/01/T042.pdf>> . Acesso em: 23 maio 2009.

SOUZA, R. R.; ALVARENGA; L. A Web Semântica e suas contribuições para a ciência da informação. **Ciência da Informação**. Brasília, v. 33, n.1, jan./abr. 2004.

THOMAZ, K. P.; SANTOS V. M. dos. Metadados para gerenciamento eletrônico de documentos de caráter arquivístico - GED/A: estudo comparativo de modelos e formulação de uma proposta preliminar. **DataGramZero** - Revista de Ciência da Informação, Rio de Janeiro, v. 4,n. 4, p. 1-25, ago., 2003.