

A Contribuição do Controle de Autoridade para a Recuperação da Informação no Ambiente Digital: o caso VIAF

The contributions of the authority control in information retrieval in a digital environment: the VIAF *case*

Denise Mancera Salgado (Programa de Pós-Graduação em Ciência da Informação/Escola de Comunicações e Artes/Universidade de São Paulo) – dsalgado@usp.br

Resumo

Discorre sobre as contribuições das áreas da representação e organização da informação, em especial do controle de autoridade, na recuperação da informação na Web. Aborda-seos desafios para organização da informação no ambiente digital e como a interação e integração dos conceitos e processos advindos da Ciência da Informação e da Computação pode auxiliar na recuperação da informação na Web. Enfatiza a necessidade de interoperabilidade e estruturação da informação no ambiente digital por meio dos conceitos da web semântica em colaboração com as ferramentas e processos de estruturação desenvolvidos pela Ciência da Informação. Apresenta o controle de autoridade como mecanismo auxiliarna estruturação da informação. Aborda a utilização do controle de autoridade na recuperação da informação e seus benefícios para bibliotecários e usuários finais. Destaca o desenvolvimento do Projeto *Virtual International Authority File (VIAF)* e sua interação com a web semântica como recurso de estruturação da informação em ambiente digital, permitindo uma recuperação confiável e precisa da informação.

Palavras-chave: controle de autoridade, ambiente digital, recuperação da informação, controle bibliográfico, *VIAF*

Abstract

It discusses the contributions of the areas of representation and organization of information, in particular the authority control, in information retrieval on the Web. Analyze the challenges for the organization of information in the digital environment and how the interaction and integration of the concepts and processes arising of Information Science and Computer can help in information retrieval on the Web. It emphasizes the need for interoperability and information structure in the digital environment through semantic web concepts in collaboration with the tools and structuring processes developed by information Science. Displays the authority control as a mechanism assist in structuring information. It covers the use of authority control in information retrieval and its benefits for librarians and end users. It highlights the development of the Virtual International Authority File Project (VIAF) and its interaction with the semantic web as information structuring feature in the digital environment, enabling a reliable and accurate retrieval of information.

Key-word: authority control, digital environment, retrieval of information, bibliographic control, *VIAF*

1. Introdução

A popularização da Internet a partir da década de 90 permitiu aos profissionais da informação entrever o sonho de Otlet e Vanevar Bush do cérebro mundial ou *Brain World*. Porém, isso é apenas um vislumbre. Tanto Otlet quanto Bush imaginaram um grande compêndio do conhecimento organizado, onde qualquer informação, através de seu tratamento, poderia ser facilmente acessada e localizada em um grande dispositivo de armazenamento. A Web permite que muitas informações possam estar ao alcance de um clique, mas ao mesmo tempo ela é caótica e desordenada.

A massa de informações armazenadas na Web aumenta de forma exponencial. A Pesquisa Nacional por Amostra de Domicílios (PNAD), realizada pelo Instituto Brasileiro de Geografia e Estatística (IBGE), demonstra que somente no Brasil aproximadamente 86 milhões de pessoas eram usuárias da Internet em 2013. Estatísticas constataam que até o ano de 2013 foram registrados por volta de 3 milhões de domínios brasileiros na internet. O uso constante de redes sociais, blogs e outros recursos para a troca de informações, sejam elas imagens, sons ou textos, principalmente uma variedade de todas essas formas, promove uma profusão de fontes de informação para qualquer pesquisador. Mas como acessar essas informações?

Mecanismos de busca populares como o Google se propõem a executar essa tarefa. Porém percebe-se uma busca entremeadada de resultados muitas vezes não pertinentes à questão de pesquisa. Novos mecanismos para filtrar e permitir uma busca de resultados mais consistente estão em desenvolvimento, mas o caos ainda impera. Os problemas enfrentados pelos mecanismos na recuperação da informação também estão relacionados, entre outros fatores, a mudanças de endereços (URLs) e recuperação de informações fora do contexto, provocadas por problemas semânticos e de ambiguidade.

Este artigo tem por objetivo discorrer sobre como a Ciência da Informação, por meio do controle de autoridade, e em conjunto com as tecnologias desenvolvidas pela área da computação, é capaz de auxiliar e fornecer recursos que permitam uma recuperação da informação mais consistente na Web.

Metodologicamente, caracteriza-se como estudo descritivo baseado em artigos selecionados sobre o tema da recuperação da informação na Web, e do controle de autoridade. Num primeiro momento são abordados os desafios para a organização da informação no ambiente Web e a necessidade de interoperabilidade no ambiente digital, seguido da busca por uma estruturação dessa informação com a utilização dos conceitos de Web semântica. Posteriormente, enfoca-se a utilização do controle de autoridade na recuperação da informação, apresentando o desenvolvimento do Projeto VIAF (*Virtual International Authority File*).

2.A Web - os desafios da organização da informação e a interoperabilidade

O movimento caótico e (supostamente) anárquico da Web, considerado benéfico ao permitir a inserção de uma grande quantidade de recursos, tais como: documentos científicos, bancos de dados, programas, arquivos multimídias; é influenciado pela criação de tais recursos de forma autônoma, sem a preocupação com regras de estruturação dos

dados, catalogação e descrições de propriedades. Esses aspectos impactam diretamente na qualidade das respostas fornecidas pelos mecanismos de busca existentes. Alia-se a isso a baixa qualidade da indexação, que resulta em quantidades de informações recuperadas e sem relevância; a cobertura parcial da Internet, pois os buscadores não são capazes de acessar registros de banco de dados que possuem interfaces que pressupõe uma interação entre um usuário humano com a base de dados, estando inacessíveis aos programas robôs; e as ferramentas de busca não especializadas (Marcondes, Sayão, 2001; Dias, Santos, 2003). Esforços vêm sendo despendidos na minimização desses problemas. A integração entre a área de Ciência da Informação e da Computação poderia ser benéfica, aliando conhecimentos técnicos e de organização da informação.

Há 30 anos, ainda no início da sociedade global em rede, a abordagem documental parecia ultrapassada, mas agora se torna mais relevante do que nunca. As tecnologias da informação e da comunicação que deveriam permitir uma maior ordem na documentação têm sido desafiadas pelo ambiente de dados não estruturados da *World Wide Web*. Apesar de ser um ambiente de comunicação pública, que permite a qualquer pessoa publicar qualquer coisa, a Web também se tornou um espaço mais difícil de navegar e encontrar material relevante e confiável (Lund, 2009).

Durante séculos, bibliotecas, museus e arquivos procuram organizar a informação armazenada. Sua expertise no processo de armazenamento, organização e recuperação da informação é vasta. Esses processos foram pensados para um armazenamento analógico e linear, entretanto, a era digital demanda novas formas de pensar a organização e o armazenamento de informações. Para Marcondes e Sayão (2001), mais que recursos informacionais a Internet apresenta novas ferramentas cognitivas, que permitem interações mentais que suplantam aquelas oferecidas por documentos impressos de leitura. O cenário tradicional de disseminação da informação, onde sujeitos autônomos se comunicam ou trocam informações uns com os outros, é posto em xeque com a documentação digital. Sua intencionalidade característica está ausente na geração de um vasto conjunto de enunciados digitais (Frohmann, 2008).

Não apenas a disseminação, mas a busca pela informação é alterada. Segundo pesquisa realizada pela *British Library* e o *Joint Information Systems Committee* (JISC) com jovens estudantes, o comportamento na busca de informações é dinâmico, horizontal, controlado e visual por natureza. As ferramentas usadas por eles para a busca exigem pouca habilidade ou conhecimento de técnicas de pesquisa. Ficam satisfeitos com uma pesquisa básica ou simples. Esses jovens fazem pouco uso de avançados recursos de pesquisa, assumindo que os mecanismos de busca "entendem" suas perguntas. Ademais, há poucos julgamentos de relevância quando se considera a presença ou ausência de palavras que descrevem exatamente o termo de pesquisa. Como resultado, perdem-se documentos relevantes e as pesquisas acabam por ser repetidas (British Library, JISC, 2008).

As bibliotecas e outras instituições tradicionais na organização da informação estão migrando para o mundo digital. Mas essa migração ainda apresenta vícios do ambiente físico. Bibliotecas digitais e portais temáticos são criados em número crescente. Seu armazenamento é feito em servidores diferentes e isolados; operados por interfaces de busca distintas que ainda são versões eletrônicas dos catálogos em fichas. Essas interfaces apresentam telas lineares, obrigando usuários a uma dispendiosa busca, site a site para encontrar informações relevantes (Marcondes, Sayão, 2001; Tillett, 2007).

Esse panorama pode ser alterado ao aplicar-se o princípio da interoperabilidade. As bibliotecas, ao colocarem seus catálogos acessíveis na Web, perceberam que a interoperabilidade beneficia seus usuários, e permitiu reduzir custos de processamento e tratamento da informação, além de promover uma melhor padronização dos dados processados.

Sayão e Marcondes (2008) destacam que o conceito de interoperabilidade não é uma novidade no domínio das bibliotecas. Desde meados do século XX as bibliotecas se organizam em serviços cooperativos, propiciando um ordenamento universal dessas informações. Toda uma estrutura global foi montada em torno da ideia do compartilhamento e da cooperação entre bibliotecas. Dessa forma, a interoperabilidade pode ser definida como “a possibilidade de um usuário realizar buscas a recursos informacionais heterogêneos, armazenados em diferentes servidores na rede, utilizando-se de uma interface única sem tomar conhecimento de onde nem como estes recursos estão armazenados” (Marcondes, Sayão, 2001 :27).

É preciso extrapolar esse conceito para além dos catálogos e atingir outros serviços como repositórios e bibliotecas digitais. A interação exerce influência sobre os sistemas de informação, crescendo o interesse tanto na interoperabilidade quanto no intercâmbio de informações documentárias. Os obstáculos tecnológicos e de organização da informação podem e devem ser superados, sendo necessária a construção de condições para que os usuários possam acessar múltiplas bases de dados através de uma interface definida e uma linguagem de busca normalizada (Moreira, Lara, 2012).

Como a capacidade de compartilhar dados e possibilitar o intercâmbio de informações, a interoperabilidade deve abranger três aspectos: o técnico, o semântico e o sintático. A reutilização é uma consequência direta do compartilhamento e da recuperação inequívoca e pertinente da informação, sendo realizada através da melhora da interoperabilidade, possibilitando que diferentes recursos e aplicações Web troquem informações através de conceitualizações compartilhadas e de uma codificação comum (Moreiro González, Sánchez Cuadrado, Morato Lara, 2012). Isso facilita o fluxo e acesso à informação. Contudo, são necessárias discussões que abranjam formatos, padrões e também o desenvolvimento técnico colaborativo, auxiliar na disseminação desses formatos (Mucheroni, Silva, 2011).

O alinhamento semântico promove uma melhor interoperabilidade entre unidades de informação, como arquivos, museus e bibliotecas. Isso ocorre através da ligação entre entidades equivalentes, similares ou semanticamente relacionadas em bases de autoridades, vocabulários de metadados ou dados de diferentes instituições, como por exemplo, o mesmo nome em bases de dados de autoridades (Marcondes, 2012).

3.A web semântica e a estruturação da informação

Em um ambiente caótico como a Web, onde o crescimento informacional pode ser verificado segundo a segunda lei de Moore, a busca de informações não processadas apresenta-se como desafio a ser suplantado. Pesquisas demonstram que o usuário em geral utiliza-se da linguagem natural para suas estratégias de busca, como já citado, inferindo que

os mecanismos de busca “entendem” a questão formulada. Entretanto, o que se observa como resultados dessas buscas é o conjunto de páginas de resposta com pouca ou nenhuma relevância. Isso ocorre porque a Web atual apresenta conteúdos pouco estruturados, sem tratamento semântico. Os buscadores por sua vez, operam analisando a coincidência de palavras, um mero encadeamento de bits e bytes iguais.

A Web semântica, pensada por Tim Berners-Lee, tem por objetivo tratar os conteúdos da Web de forma semântica, onde uma estrutura poderosa de conhecimento possa ligar dados, documentos e informações. A diferença entre a Web atual e a Web semântica é que a primeira tem suas buscas baseadas nas coincidências de caracteres ligando páginas ou documentos em HTML, e está baseada em padrões de forma. Já, a Web semântica liga dados estruturados, e está fundamentada em uma estrutura de conteúdo. Para isso utiliza um padrão que modela os dados, permitindo sua identificação e relacionando-os uns aos outros segundo propriedades e classes de propriedades, promovendo a construção de declarações (Lara, 2012).

Lund (2009) apresenta a seguinte equação para que os documentos digitais estejam disponíveis para uso: Documento eletrônico = texto informado + conhecimento. O texto deve ser estruturado de forma personalizado, permitindo a adição de conhecimento na forma de metadados que se adaptem a um leitor / usuário específico. Isso se traduz na Web semântica que, mediante o uso de ontologias e outras tecnologias, permite a análise dos elementos do documento digital por um sistema com o objetivo de sua exploração por um leitor competente. O avanço mais imediato é a utilização de estruturas de marcadores para modular as análises de texto, preocupando-se com a integração de elementos de formatação de material como fonte.

A Web semântica, ao tratar da estruturação de conteúdos, trabalha com o conceito de metadados, um ambiente há muito conhecido da Ciência da Informação. É nesse cenário que a área da Representação e Organização da Informação pode auxiliar a Web a ser mais eficiente e estruturada. Recursos existentes em bibliotecas, como vocabulários controlados por assuntos, nomes e títulos podem ser incorporados a Web, principalmente no que concerne a construção de ontologias para a Web semântica.

O desenvolvimento da Web semântica envolve a criação de uma infraestrutura de conexões de recursos e o uso controlado de vocabulários, que são chamados de ontologias. Uma das definições mais utilizadas para o termo “ontologia” na área da computação e informação é a fornecida por Gruber (1995, 2009), que explicita conceituações e relações existentes em um determinado domínio do conhecimento. Nesse contexto, pode-se descrever a ontologia de um programa definindo um conjunto de termos de representação. Em tal ontologia, definições associam nomes de entidades no universo do discurso (por exemplo, classes, relações, funções ou outros objetos) com texto legível descrevendo o que os nomes significam e os axiomas formais que restringem a interpretação e uso desses termos. No âmbito dos sistemas de banco de dados, a ontologia pode ser vista como um nível de abstração de modelos de dados, de forma análoga aos modelos hierárquicos e relacionais, mas se destina a modelagem do conhecimento sobre os indivíduos, seus atributos e suas relações com outros indivíduos. São usadas para integrar bancos de dados heterogêneos, permitindo a interoperabilidade entre sistemas distintos, e especificando as interfaces para serviços baseados no conhecimento.

Essas ontologias poderiam ser usadas em formatos de exibição mais fáceis para os usuários e na sua própria língua e grafia. Essa é a oportunidade para que bibliotecas e

outros serviços de informação contribuam para a infraestrutura da Web semântica, pois suas bases de dados trabalham com vocabulários controlados. Esses vocabulários podem conectar-se com outros vocabulários controlados e sistemas de classificação, tais como dicionários biográficos e fontes de referência, com a finalidade de implantar a precisão em pesquisas e auxiliar os usuários em suas buscas (Tillett, 2007).

4. O controle de autoridade e a recuperação da informação

O controle de autoridade como o processo de unificar os pontos de acesso em catálogos, mediante a utilização de uma forma normalizada, mostra as relações existentes entre os elementos de um registro bibliográfico e busca facilitar a identificação e a recuperação dos documentos armazenados (Herrero Pascual, 1999). Objetiva evitar ambiguidades ou confusões que possam ser causadas por homônimos, sinônimos ou diversidades de nomes pelos quais as pessoas, entidades, obras, temáticas ou conceitos possam ser denominados. Catálogos desenvolvidos por serviços de informação auxiliam no intercâmbio de informações, sendo benéficos na redução dos custos globais para a manutenção e elaboração de bases de dados bibliográficas e de autoridades.

Com isso, o controle de autoridade permite precisão e recuperação, algo que falta, atualmente, em pesquisas realizadas na Web. A precisão é evidenciada ao recuperar apenas os registros ou itens de interesse, e a sua rede de referências cruzadas, mostrando os relacionamentos entre os pontos de acesso ou descritores, garante uma recuperação relevante dos dados, bem como a navegação entre informações e dados relacionados (Tillett, 2000).

A capacidade de estabelecer relações nos catálogos de autoridade, por meio do uso das conexões entre as formas autorizadas de nomes, assuntos e títulos, podem ser usadas em variados recursos disponíveis na Web como diretórios, biografias, repositórios, serviços de indexação, facilitando a navegação e promovendo explicações sobre variações e inconsistências de termos.

A catalogação inexiste sem a padronização dos pontos de acesso, e o controle de autoridade é o mecanismo pelo qual podemos alcançar o grau de padronização necessária. A padronização beneficia os catalogadores ao capacitá-los a identificar e distinguir os pontos de acesso controlados dentro de um catálogo. Mais importante, o controle de autoridade beneficia os usuários finais ao facilitar a pesquisa por qualquer forma controlada de nome de um autor, ou de um título, para recuperar os recursos bibliográficos em catálogos (Gorman, 2004; FRANAR, 2009).

A adoção do controle de autoridade para nomes pessoais e de instituições pode fornecer uma pesquisa mais adequada para os diferentes usuários em seu próprio contexto. Payette, Blanchi, Lagoze e Overly (1999) destacam que os esforços de interoperabilidade e estrutura semântica permitem que diferentes comunidades, com diferentes tipos de informação e usando diferentes tecnologias, consigam um nível geral de compartilhamento de informação criando novos e poderosos serviços informacionais.

Essa afirmação é reforçada pela cultura da pós-modernidade apontada por García Canclini e Cruces (2012) ao destacarem a ocorrência de uma mudança na relação com a territorialidade, onde as tendências são internacionais e transfronteiriças. A

interculturalidade deve propiciar a continuidade de pertencimentos étnicos, grupais e nacionais, junto com o acesso fluido a repertórios transnacionais difundidos pelos meios de comunicação urbanos e de massas (García Canclini, 2007). O desenvolvimento do ambiente digital e da tecnocultura empreende um exercício de pluralismo ético na criação de ferramentas que não apenas auxiliem na organização dos conhecimentos dominantes, mas, especialmente, que promovam a resistência dos conhecimentos considerados subalternos (García Gutiérrez, 2006).

É preciso que os serviços de informação, sejam eles digitais ou analógicos, reflitam sobre as necessidades nacionais e culturais de seus usuários. Os catálogos de autoridade devem considerar a indicação da estrutura sindética das referências cruzadas e formas autorizadas dos vocabulários para serem usados de forma a atender audiências específicas, considerando-se, também, a inclusão de variantes em idiomas e grafias alternativas (Tillett, 2007).

Os recursos existentes podem ser explorados de maneira crítica para sua reutilização, buscando por conceitos adequados às necessidades de diferentes comunidades de usuário e centrando esforços para formalizar sistemas de organização do conhecimento não são suportados por grandes instituições, sendo formados por sistemas de organização do conhecimento que expressam culturas locais, pontos de vista originais, e que respondam às necessidades específicas de informação (Lara, 2012).

5. O Virtual International Authority File

O esforço em tornar a Web mais adequada a usuários multilocais e com necessidades específicas de informação, em sua grafia e idioma pode ser exemplificado pelo projeto *Virtual International Authority File* (VIAF) - Catálogo Virtual Internacional de Autoridade. Desenvolvido em parceria pela *Library of Congress*, a *Deutsche Nationalbibliothek*, a *Bibliothèque Nationale de France* e a *Online Computer Library Center* (OCLC), tem por objetivo facilitar o acesso aos maiores catálogos de autoridade de nomes do mundo. Foi idealizado como uma base para a Web semântica, de forma a permitir a alteração do modo de exibição dos nomes de pessoas para que possam ser apresentados no idioma e grafia escolhido pelo usuário da Web. O início de funcionamento ocorreu em 2012, contando com 20 agências de 16 países. Atualmente participam do projeto 35 agências de 28 países. Relatório da OCLC, de 2014, totaliza a base com 44.885,637 (milhões) de registros (OCLC, 2014).

O VIAF é um esforço de cooperação internacional entre bibliotecas e outras organizações que contribuem com arquivos de autoridade nome, aumentando o acesso à informação. Seu uso auxilia na redução de custo na manutenção de catálogos, tornando-os úteis para as bibliotecas e usuários em geral. Para sua execução o VIAF compara e associa os catálogos de autoridades das bibliotecas e instituições participantes, agrupando todos os registros de autoridade em um registro único que contém os diferentes nomes fornecidos para essa autoridade.

Como expansão prática do conceito de controle bibliográfico universal, o VIAF é embasado no trabalho realizado por cada agência bibliográfica. Permite que variações

nacionais ou regionais da forma autorizada do nome coexistam, apoiando assim as necessidades dos usuários em todo o mundo para as variações de idioma preferido, grafia e ortografia. As atuais propostas para o futuro da Web descrevem o uso de ontologias para torná-la mais compreensível para a máquina e o processamento automático. O VIAF pode ser um dos elementos básicos para a Web semântica quando combinado com outros vocabulários controlados e arquivos de autoridade de fontes tais como serviços de resumos e indexação, arquivos, museus, editoras, etc. (Bennett, Hengel-Dittrich, O'Neil, Tillett, 2006).

O primeiro objetivo do VIAF era criar links entre catálogos de autoridades de Bibliotecas Nacionais. Atualmente, outras instituições fazem parte do projeto, como o *GettyResearchInstitute*. Como um catálogo de autoridade, o VIAF baseia-se nas relações existentes entre um determinado arquivo bibliográfico e arquivo de autoridade que o dirige. Um registro VIAF possui:

- identificador VIAF;
- todas as formas preferidas de um nome de pessoa ou de entidade coletiva, estabelecida pelos parceiros e a bandeira do país relacionado à agência. Um gráfico também mostra dinamicamente como as correspondências foram feitas;
- todas as formas variantes do nome utilizadas por pelo menos um dos parceiros;
- todas as formas afins usadas por pelo menos um dos parceiros;
- os países onde as obras da pessoa ou da entidade foram publicadas. Um planisfério mostra os resultados globais;
- estatísticas mostrando o histórico de publicações de uma autoridade;
- as principais editoras que lançaram as obras da autoridade cadastrada;
- o sexo, a nacionalidade, o idioma mais usado para a autoridade;
- links externos para outros recursos, tais como *WorldCatIdentities* ou *Wikipédia*;
- links para a representação do registro VIAF em UNIMARC, MARC21 XML e em RDF (Bourdon,Boulet, 2011).

Um usuário na Alemanha que tenha interesse em localizar uma obra de Jorge Amado, pode assim, solicitar o termo de busca em seu próprio idioma e grafia, o mesmo ocorrendo com um usuário na China. Além de instituições ligadas a serviços de informações, a *Wikipedia* também se utiliza do VIAF para a padronização e vinculação de referências em seu conteúdo. Para tanto, desenvolveu-se um projeto em parceria entre wikipedistas e a *British Library*, para integrar dados de autoridade do VIAF com artigos biográficos da *Wikipédia*. Esse projeto estabelece correspondência de identificadores VIAF para centenas de milhares de artigos biográficos da *Wikipédia*, usando um algoritmo e o *VIAFbot*, uma conta automatizada *Wikipedia*. (Klein,Kyrois, 2013).

O VIAF permite uma melhor desambiguação de nomes, evitando o uso incorreto de nomes e referências. O potencial do VIAF na Web semântica já é uma realidade, pois possui incorporado o uso do *ResourceDescription Framework* (RDF) e do *UniformResourceIdentifier* (URI), ferramentas da Web semântica para a estruturação de dados.

Considerações finais

Com a aplicação dos modelos conceituais *Functional Requirements for Bibliographic Records* (FRBR) e do *Functional Requirements for Authority Data* (FRAD) na catalogação de recursos em bases de dados, é possível construir ligações entre obras e expressões - como obras originais e suas traduções ou romances que formam a base para roteiros cinematográficos - e navegar por essas relações. Algo diverso de confiar em notas textuais que não são legíveis por máquina.

Dessa maneira, os serviços de informação podem tornar seus dados mais acessíveis na web. O uso desses conceitos amplia as possibilidades de atuação de profissionais da informação. Os usuários Web precisam dispor de catálogos online, repositórios e bibliotecas digitais como recursos que permitam a navegabilidade ao longo de suas páginas, estabelecendo relações com outros recursos disponíveis na Web, sejam eles virtuais ou textuais.

Além de pensarmos em uma nova forma de representar a informação, é preciso adequar os instrumentos de busca às necessidades atuais dos usuários. Segundo o estudo da British Library e JISC (2008), citado anteriormente, os serviços de informação estão em constante competição pela atenção dos grupos de usuários, especialmente os jovens, que exigem conteúdo personalizado, dinâmico e interativo. Os usuários da biblioteca demandam acesso durante 24 horas por dia, 7 dias por semana, além de satisfação imediata em um clique. Procuram cada vez mais "a resposta" em vários formatos. Os sistemas utilizados por bibliotecas e outras instituições para a disponibilização de seus produtos na Web são menos intuitivos do que o buscador comum, fazendo com que os usuários preferiram uma pesquisa global pelo Google a fornecida pela biblioteca, mais sofisticada e demorada. Não precisamos apenas de uma compreensão ampla de como os sistemas de recuperação trabalham, e como a informação é representada dentro de bases de dados bibliográficas e de texto completo, mas, também, alguma apreciação da natureza do espaço informacional e como contribui a ortografia, a gramática e a estrutura da frase para a eficácia da pesquisa.

O momento é de reflexão e de ação. O papel social e educativo dos profissionais da informação pode ser fundamental para a realização da tão almejada Enciclopédia Universal. A Web é capaz de absorver, criar, produzir e comunicar milhares de informações, dados e recursos, mas a capacidade de organização e recuperação desses é frágil. Os serviços de informação possuem ferramentas necessárias e adequadas para a maximização dos recursos da Web, mas é preciso deixar o pensamento linear que perdurou por séculos na área da Ciência da Informação e dedicarmos esforços para a adequação de nossas ferramentas ao pensamento digital, dinâmico e horizontal da atual geração.

Bibliotecários, documentalistas, arquivistas e museólogos podem contribuir significativamente na construção da Web semântica, promovendo uma maior interação humano-computador, por meio da adoção de ferramentas que possibilitem aos programas inferirem sobre os dados e informações e processá-los de forma inteligente. Mais que ontologias para a Web semântica, esses profissionais são hábeis em estabelecer relações entre dados, documentos e informações, propiciando uma navegação e recuperação da informação satisfatória para aos usuários da Web.

Referências

- Bennett, R., Hengel-Dittrich, C., O'Neil, E.T., Tillett, B.B. (2006, August). VIAF (Virtual International Authority File): linking Die Deutsche Bibliothek and Library of Congress name authority files. *Proceedings of the World Library and Information Congress: IFLA General Conference and Council*. Seoul, KR, 72.
- Bourdon, F., Boulet, V. (2011, August) VIAF: a hub for a multilingual access to varied collections. *Proceedings of the World Library and Information Congress: IFLA General Conference and Assembly*. San Juan, PRI, 77.
- British Library, JISC. (2008). Comportamiento informacional del investigador del futuro: informe CIBER. *Anales de Documentación*, 11, 235-258. Recuperado em 24 julho, 2013, de <http://revistas.um.es/analesdoc/issue/view/2651>.
- Dias, T.D., Santos, N. (2003). Web semântica: conceitos básicos e tecnologias associadas. *Cadernos do IME*, 14, 25-28. Recuperado em 14 outubro, 2013, de <http://www.e-publicacoes.uerj.br/index.php/cadinf/article/view/6619/4734>.
- FRANAR. (2009). *Functional Requirements for Authority Data: a conceptual model*. Recuperado em 16 agosto, 2012 de <https://netfiles.uiuc.edu/trainor1/errtpapers/FRAD.pdf>.
- Frohmann, B. (2008). O caráter social, material e público da informação. In FUJITA, M.S.L., MARTELETO, R.M., LARA, M.L.G. (Org). *A dimensão epistemológica da Ciência da Informação e suas interfaces técnicas, políticas e institucionais nos processos de produção, acesso e disseminação da informação*(p.13-34). São Paulo: Cultura Acadêmica/Marília: FUNDEPE.
- García Canclini, N. (2007). *Diferentes, desiguais e desconectados*. Rio de Janeiro: Editora da UFRJ.
- García Canclini, N., Cruces, F.(2012) Conversación a modo de prólogo. In GARCÍA Canclini, N., Cruces, F. (Org). *Jóvenes, culturas urbanas y redes digitales: prácticas emergentes en las artes, las editoriales y la música*. Fundación Telefónica. Recuperado em 06 abril, 2013 de http://www.articaonline.com/wp-content/uploads/2011/07/jovenes_culturas_urbanas_completo.pdf.
- Garcia Gutiérrez, A. (2006). Cientificamente favelados: uma visão crítica do conhecimento a partir da epistemografia. *Transinformação* 18(2), 103-112. Recuperado em 05 junho, 2013 de <http://revistas.puc-campinas.edu.br/transinfo/viewarticle.php?id=151>.
- Gorman, Michael. (2004). Authority control in the context of bibliographic control in the electronic environment. *Cataloguing & Classification Quarterly*, 3-4(38), 11-22.
- Gruber, T. R. (1995). Toward principles for the design of ontologies used for knowledge sharing. *International Journal of Human Computer Studies*, 43(5-6), 907-928. Recuperado em 09 novembro, 2013 de <http://tomgruber.org/writing/onto-design.pdf>.
- Gruber, T. R. (2009). Ontology. In LIU, L., ÖZSU, M.T. (Ed.). *Encyclopedia of Database Systems*. Recuperado em 09 novembro, 2013 de <http://tomgruber.org/writing/ontology-definition-2007.htm>.

Herrero Pascual, C. (1999). El control de autoridades. *Anales de documentación*, (2), 121-136. Recuperado em 30 janeiro, 2013 de <http://redalyc.uaemex.mx/pdf/635/63500208.pdf>.

Klein, M., Kyrios, A. (2013). VIAFbot and the integration of library data on wikipedia. *The Code4Lib Journal*, (22). Recuperado em 17 outubro, 2013 de <http://journal.code4lib.org/articles/8964>.

Lara, M. L. G. (2013). Novas tecnologias nos sistemas de organização do conhecimento: possibilidades de 'escutar' outros discursos?. Anais da Jornada Científica Internacional da Rede Mussi 'Redes e processos info-comunicacionais: mediações, memórias, apropriações. *Anais da 2a. jornada Científica Internacional Redes e Processos Info-Comunicacionais: mediações, memórias, apropriações*. Rio de Janeiro, RJ, BRA, 2.

Lund, N.W. (2009). Documenttheory. *Annual review of Information Science and Technology*, 43(1), 1-55. Recuperado em 04 maio, 2014 de <http://onlinelibrary.wiley.com/doi/10.1002/aris.2009.1440430116/full>.

Marcondes, C.H. (2012). "Linked data" – dados interligados – e interoperabilidade entre arquivos, bibliotecas e museus na Web. *Encontros Bibli*, 17(34), 171-192.

Marcondes, C.H., Sayão, L.F. (2001). Integração e interoperabilidade no acesso a recursos informacionais eletrônicos em C&T: a proposta da Biblioteca Digital Brasileira. *Ciência da Informação*, 30(3), 24-33. Recuperado em 12 julho, 2013 de <http://revista.ibict.br/ciinf/index.php/ciinf/article/view/190/167>.

Moreira, W., Lara, M.L.G. (2012). Ontologias, categorias e interoperabilidade semântica. *DataGramaZero*, 13(4). Recuperado em 17 agosto, 2013 de http://www.dgz.org.br/ago12/Art_03.htm.

Moreiro González, J.A., Sánchez Cuadrado, S., Morato Lara, J. (2012). Mejora de la interoperabilidad semántica para la reutilización de contenidos mediante sistemas de organización del conocimiento. *Encontros Bibli*, 17(33), 46-58. Recuperado em 09 novembro, 2013 de <https://periodicos.ufsc.br/index.php/eb/article/view/1518-2924.2012v17n33p46/21710>.

Mucheroni, M.L., Silva, J.F.M. (2011). A interoperabilidade dos sistemas de informação sob o enfoque da análise sintática e semântica de dados na Web. *Ponto de Acesso*, 5(1), 3-18. Recuperado em 28 outubro, 2013 de <http://www.pontodeacesso.ici.ufba.br>.

OCLC. 2014 Annual Report to VIAF Council. Recuperado em 10 de janeiro, 2015 de <http://goo.gl/j5Jy0F>.

PAYETTE, S., BLANCHI, C., LAGOZE, C., OVERLY, E.A. (1999). Interoperability for digital objects and repositories: the Cornell/CNRI Experiments. *D-Lib Magazine*, 5(5). Recuperado em 12 julho, 2013 de <http://webdoc.sub.gwdg.de/edoc/aw/d-lib/dlib/may99/payette/05payette.html>.

Sayão, L.F., Marcondes, C.H. (2008). O desafio da interoperabilidade e as novas perspectivas para as bibliotecas digitais. *TransInformação*, 20(2), 133-148. Recuperado em 12 julho, 2013 de <http://periodicos.puc-campinas.edu.br/seer/index.php/transinfo/article/view/530/510>.

Tillett, B. (2000, November). Authority Control on the Web. *Proceedings of the Bicentennial Conference on Bibliographic Control for the New Millennium: Confronting the Challenges of Networked Resources and the Web*, Washington, D.C. Recuperado em 15 maio, 2013 de http://www.loc.gov/catdir/bibcontrol/tillett_paper.html.

Tillett, B.(2007). Base Virtual Internacional de Autoridade. Proceedings for the International Meeting of Experts for an International Cataloguing Code, Pretória, ZA, 5. Recuperado em 25 outubro, 2010 de http://www.imeicc5.com/download/portuguese/Presentations3_VIAF_IME%20ICC_Pretoria_Port.pdf.