

OS METADADOS COMO INSTRUMENTOS TECNOLÓGICOS NA PADRONIZAÇÃO E POTENCIALIZAÇÃO DOS RECURSOS INFORMACIONAIS NO ÂMBITO DAS BIBLIOTECAS DIGITAIS NA ERA DA WEB SEMÂNTICA

artigo de revisão

Fabiano Ferreira de Castro*
Plácida Leopoldina V. A. da Costa Santos**

RESUMO

As bibliotecas digitais, como ambientes propícios para a recuperação de informações, têm na utilização de metadados a padronização das formas de representação e a possibilidade de garantia de interoperabilidade entre sistemas, favorecendo a integridade e a acessibilidade dos recursos informacionais de forma eficiente pelo usuário final. Com base na literatura disponível, aponta-se o uso de metadados na construção de bibliotecas digitais como elementos-chave na consolidação da Web Semântica. Destacam-se a representação e a descrição de forma e de conteúdo dos recursos informacionais como foco na busca por iniciativas para a construção de padrões de representação e descrição de recursos que se utilizam de tecnologias subjacentes à Web Semântica.

Palavras-chave

BIBLIOTECAS DIGITAIS
METADADOS
WEB SEMÂNTICA
PADRONIZAÇÃO E REPRESENTAÇÃO DA INFORMAÇÃO
INTEROPERABILIDADE.

* Programa de Pós-Graduação em Ciência da Informação da Universidade Estadual Paulista-UNESP/Campus de Marília.
E-mail: fabianoffc@marilia.unesp.br

** Departamento de Ciência da Informação – Programa de Pós-Graduação em Ciência da Informação da Universidade Estadual Paulista-UNESP/Campus de Marília.
E-mail: placida@marilia.unesp.br

I INTRODUÇÃO

No cenário atual, marcado principalmente pelo uso intensivo de tecnologias, em especial, as Tecnologias da Informação e Comunicação - TICs -, é possível verificar uma mudança de paradigmas nos ambientes informacionais, tais como as bibliotecas, caracterizadas como espaços de preservação do conhecimento acumulado pela humanidade. Nesse contexto, as bibliotecas digitais se caracterizam como ambientes facilitadores de acesso às informações sem a limitação de espaço e tempo.

Entretanto, a padronização da descrição das informações continua sendo processo essencial

para a caracterização dos documentos, a fim de garantir com eficiência o acesso e a recuperação dos conteúdos digitais.

Nesse sentido, este estudo tem como epicentro o destaque da relevância dos metadados no âmbito de bibliotecas digitais na construção Web Semântica, como instrumentos fundamentais nos ambientes de manipulação de dados bibliográficos e que potencializam os resultados de busca, garantindo ao usuário melhores condições de acesso, recuperação e reutilização dos recursos informacionais digitais.

Por meio de uma metodologia de revisão de literatura sobre as bibliotecas digitais e da aplicação de metadados a este contexto, foi possível identificar o papel primordial dos metadados no cenário de

bibliotecas digitais, devendo tal assunto ser repensado e refletido constantemente pela comunidade científica, em especial pelos profissionais da informação, uma vez que tem aumentado a quantidade de padrões de metadados que buscam contemplar as necessidades dos usuários nos processos de recuperação e busca dos recursos informacionais digitais.

2 BIBLIOTECAS DIGITAIS: ALGUMAS CONSIDERAÇÕES

A evolução das bibliotecas tem merecido destaque ao longo destes anos, no que se refere ao desenvolvimento e ao uso de tecnologias, principalmente as da informação e comunicação que potencializam e geram recursos de acesso e formatos de intercâmbio para o tratamento das informações bibliográficas.

Vale destacar que as bibliotecas, em diferentes épocas, sempre se utilizaram das tecnologias disponíveis para a otimização de seus processos e serviços com o propósito de agilizar as tradicionais rotinas de tratamento das informações, a fim de garantir a integridade e a qualidade na recuperação das informações.

A história das bibliotecas digitais está intimamente ligada à realidade dos últimos anos do século XX, período em que o uso corrente das tecnologias da informação e comunicação permitiram a ampliação da produção de informações em formato digital e o êxito social das redes de comunicação impulsionaram o intercâmbio de informações em diferentes formatos e suportes.

A diferença entre a biblioteca tradicional e a biblioteca digital está na potencialização da disponibilização e do acesso ao recurso digital que, por meio do uso das ferramentas tecnológicas, tem ampliadas as possibilidades de descrição, de armazenamento, de recuperação, de preservação e de uso.

Uma variedade de estudos para a definição de bibliotecas digitais está disponível, alguns centrados nos objetivos que elas devem contemplar, outros voltados para o conteúdo informacional, no sentido de que tipo de informação deve ser armazenado e disponibilizado, e outros direcionados ao público a que elas devem servir.

Segundo Borgman (1996, p. 6, tradução nossa),

Bibliotecas digitais são conjuntos de fontes eletrônicas e serviços técnicos associados

para a criação, pesquisa e uso da informação, que possibilitam uma extensão e um aumento do armazenamento da informação e dos sistemas de recuperação de informação, manipulando dados digitais em qualquer meio (texto, imagens, sons, imagens dinâmicas e estáticas) em redes distribuídas de trabalho. O conteúdo das bibliotecas digitais inclui dados e metadados que descrevem vários aspectos do dado (representação, criador, dono, direitos de reprodução) e metadados que consistem em ligações ou relacionamentos com outros dados ou metadados, sejam esses externos ou internos à biblioteca digital.

Parece consensual para vários autores que uma característica das bibliotecas digitais é o fácil compartilhamento instantâneo dos recursos informacionais através de redes de computadores, de forma que o acesso aconteça, local ou remotamente, em qualquer lugar e a qualquer tempo, com custo proporcionalmente baixo.

Neste sentido, Machado, Novaes e Santos (1999, p.216), reforçando tal assertiva, dizem que

[...] a biblioteca digital tem como característica uma coleção de documentos eminentemente digitais, independentemente se foram criados na forma digital ou digitalizados a partir de documentos impressos, e permite, por meio do uso de redes de computadores, compartilhar a informação instantânea e facilmente.

Em uma das obras mais recentes no Brasil sobre a temática biblioteca digital, Toutain (2006, p. 16), assim a define:

Biblioteca que tem como base informacional conteúdos em texto completo em formatos digitais - livros, periódicos, teses, imagens, vídeos e outros -, que estão armazenados e disponíveis para acesso, segundo processos padronizados, em servidores próprios ou distribuídos e acessados via rede de computadores em outras bibliotecas ou redes de bibliotecas da mesma natureza.

É possível identificar, na literatura científica, que biblioteca digital é a terminologia mais empregada no contexto de um cenário marcado pelos aparatos tecnológicos, o que por si implica uma mudança de paradigma, especialmente no que

diz respeito ao fazer profissional dos bibliotecários, uma vez que nas bibliotecas digitais o tratamento dado ao recurso informacional requer uma descrição de forma e de conteúdo legível por máquinas, com resultados compreensíveis aos humanos.

Desse modo, destaca-se a necessidade de um tratamento adequado de forma e conteúdo para a representação e a apresentação de informações para uma recuperação mais eficiente.

Segundo Cardoso e Oliveira (2000, p.1),

As bibliotecas digitais vêm ganhando espaço na rede, como uma referência a informações de qualidade e um novo campo de possibilidade à cultura. Um dos desafios, neste momento, é estimular os usuários a explorar suas potencialidades e assim tirar maior proveito destes sistemas, realizando consultas de maneira mais eficaz e criando um novo ambiente de leitura.

Destaca-se, entretanto, que o planejamento e o desenvolvimento estrutural estarão pautados no padrão de organização das bibliotecas tradicionais, ou seja, a seleção, o armazenamento, a disseminação, a recuperação, a preservação, o serviço de referência, devem estar contemplados de acordo com uma política institucional que norteie a funcionalidade da biblioteca digital.

Nesta perspectiva, conforme pode ser observado por Wainwright (1996, p. 2, tradução nossa),

[...] as bibliotecas digitais ainda permanecem sendo bibliotecas, [sustentadas] com os mesmos propósitos, funções e metas das bibliotecas tradicionais [convencionais]. A parte digital do termo indica meramente que os materiais são armazenados e acessados digitalmente.

O armazenamento e a preservação dos recursos informacionais e de seu conteúdo têm sido objetos de estudo de pesquisas atuais na área, uma vez que a massa de informações em meio digital cresce exponencialmente, dificultando muitas vezes o controle, o acesso e a integridade das informações.

Ainda nessa perspectiva, Lesk (1997 *apud* ARELLANO, 2004, p.17) afirma que “as bibliotecas digitais são meios mais dinâmicos para preservação digital do que as bibliotecas tradicionais, no sentido da sua adaptação às freqüentes mudanças tecnológicas”. Tal contexto requer uma preocupação

da comunidade científica com a identificação e o desenvolvimento de normas e padrões para a representação das informações que facilitem a identificação de sua descrição e sua localização, e ainda proporcionem a interoperabilidade entre os sistemas informacionais. A definição de normas que facilitem a melhor localização e recuperação das informações na ambiência de bibliotecas digitais possibilita maior aproveitamento dos seus recursos, facilitando o acesso e a assimilação dos conteúdos digitais disponibilizados ao usuário final. (MODESTO, 2005).

Nesse sentido, os metadados têm papel fundamental, pois estão estruturados em uma ambiência padronizada e facilitarão os processos de busca e recuperação dos recursos informacionais em ambientes informacionais digitais.

3 BIBLIOTECAS DIGITAIS: UMA ABORDAGEM SOBRE OS METADADOS

Os bibliotecários produzem e padronizam metadados há séculos, desde as primeiras tentativas de organização da informação a partir da descrição de documentos. O que vem acontecendo ultimamente é que profissionais de diversas áreas estão buscando criar métodos de descrição da informação, mas seu desconhecimento dos métodos e processos da Biblioteconomia tem gerado uma variedade de padrões que muitas vezes não atende satisfatoriamente às exigências de uma lógica descritiva estabelecida na Biblioteconomia. (MILSTEAD; FELDMAN, 1999; ALVES, 2005).

Para definir metadados, recorremos a Alves (2005), pois há uma pluralidade discursiva a respeito do termo nas diferentes áreas de conhecimento. A autora relata que

[...] os metadados são conjuntos de atributos, mais especificamente dados referenciais, que representam o conteúdo informacional de um recurso que pode estar em meio eletrônico ou não. Já os formatos de metadados, também chamados de padrões de metadados, são estruturas padronizadas para a representação do conteúdo informacional que será representado pelo conjunto de dados-atributos (metadados) (ALVES, 2005, p.115).

Os metadados são elementos que compõem estruturas para dar tratamento às informações, tendo em vista o gerenciamento informacional. (ALVES, 2005).

Para os profissionais da Ciência da Informação, o termo metadados está relacionado com o tratamento da informação, mais especificamente com as formas de representação de um recurso informacional para fins de descrição, identificação, localização, busca e recuperação, ou seja, dados bibliográficos e catalográficos que servem para organizar, representar e tornar a informação identificável, localizável e acessível. (GILLILAND-SWETLAND, 1999).

Formas diferenciadas de representação para um tratamento mais acurado sobre tais informações com a finalidade de obter uma recuperação eficiente dos recursos informacionais em ambientes digitais e que atendam aos requisitos da interoperabilidade entre os sistemas informacionais, apresentam-se como uma necessidade para a otimização de ambientes informacionais digitais.

Para tanto, Alves (2005) aponta os metadados como a ferramenta capaz de proporcionar uma forma de representação dos recursos informacionais em meio digital e, conseqüentemente, promover a mediação entre a informação em ambiente digital e o usuário.

Vale destacar que a interoperabilidade é uma vertente importante a ser discutida, pois o compartilhamento de informações e registros entre bibliotecas digitais com diferentes arquiteturas, formatos e tecnologias é uma das metas no uso das TICs.

Ao buscar ações que proporcionarão o alcance dessa meta, voltamos aos metadados que, quando aplicados ao contexto de bibliotecas digitais, conforme apontam Souza, Catarino e Santos (1997, p. 4), apresentam-se como facilitadores da

interoperabilidade entre as diversas fontes de dados; na definição da linguagem de consulta; na agilidade e no acesso com qualidade na recuperação da informação; e no intercâmbio informacional.

Um dos objetivos do uso de metadados no contexto da *Web* e das bibliotecas digitais é permitir não só descrever documentos eletrônicos e informações em geral, possibilitando sua avaliação de relevância por usuários humanos, mas também permitir agenciar computadores e programas especiais, robôs e agentes de *software*, para que eles compreendam os metadados associados a documentos e possam então recuperá-los, avaliar sua relevância e manipulá-los com mais eficiência.

Diante disso, podemos dizer que, sem o uso de metadados, não haverá o estabelecimento de uma biblioteca digital de acordo com um conjunto de padrões e normas que garantam o acesso, a recuperação, o uso e a reutilização dos recursos informacionais e que atendam às necessidades dos usuários apresentadas nas solicitações de busca.

Neste contexto, as tecnologias atreladas à *Web Semântica* caminham em direção aos objetivos das bibliotecas digitais, destacando como ponto primordial proporcionar meios mais adequados de representar e organizar os conteúdos digitais, tomando como base os aspectos semânticos inerentes a tais conteúdos, no que diz respeito à sua recuperação pelo usuário final.

A *Web Semântica* propõe-se estruturar e dar semântica aos conteúdos representados com o intuito de diminuir problemas de busca e recuperação de informações, apresentando uma estrutura que possibilitará a compreensão e o gerenciamento do conteúdo dos recursos informacionais, por meio da valorização da semântica destes recursos e de agentes capazes de processar informações e trocá-las com outros programas (BERNERS-LEE; HENDLER; LASSILA, 2001).

Souza e Alvarenga (2004) alertam: para que a *Web Semântica* seja implementada, é necessário o trabalho conjunto de várias ferramentas tecnológicas que propiciem uma melhor estruturação e representação dos conteúdos.

Nesse sentido, para que haja uma recuperação mais eficiente dos conteúdos, principalmente no âmbito das bibliotecas digitais, é necessário investir em ferramentas subjacentes à *Web Semântica*, tais como padrões de metadados e ontologias.

Conforme *Semantic Web* (2005, tradução nossa), a definição de ontologia está relacionada a uma especificação de uma conceituação de um domínio de conhecimento, ou seja, um vocabulário controlado que descreve objetos e as relações entre eles; as ontologias podem incluir glossários, taxonomias e tesouros para o estabelecimento de um vocabulário comum.

Podemos dizer, então, que “as ontologias funcionam como qualificadores dos metadados e não haverá estabelecimento da semântica sem a representação por metadados” (ALVES, 2005, p. 110).

Para melhor visualização da arquitetura da *Web Semântica*, especificamente a camada mais baixa, onde são descritos e representados os metadados e outras ferramentas tecnológicas que compõem a camada de estrutura, vejamos a Figura 1.

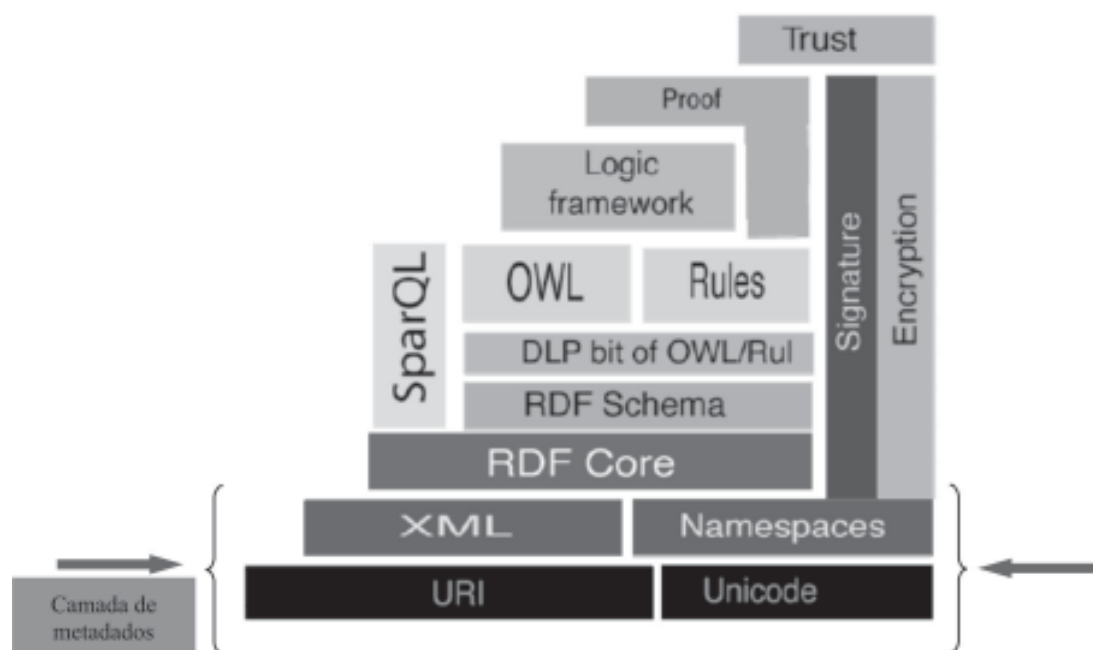


Figura 1: Arquitetura da Web Semântica
 Fonte: Berners-Lee (2005, p. 27)

O estabelecimento pleno da *Web Semântica* acontecerá através de um conjunto de recursos informacionais que deverão estar representados com um mínimo de exigência para assegurar a eficácia na recuperação. Tal exigência requer que os recursos informacionais estejam devidamente representados por meio da utilização de metadados ou padrões/formatos de metadados padronizados para a construção-base a fim de que a *Web Semântica* se desenvolva.

Outro ponto bastante peculiar para a construção *Web Semântica* é a forma como as bibliotecas digitais podem contribuir para o seu estabelecimento e concretização. As bibliotecas digitais, que são consideradas ambientes específicos e padronizados, e distribuídas de acordo com áreas do conhecimento, já se constituem como segmentos na Internet e proporcionarão o estabelecimento da semântica de forma mais precisa, uma vez que os metadados atribuídos nas bibliotecas digitais garantirão a representação dos recursos informacionais e contribuirão para aplicações mais efetivas na *Web Semântica*.

Souza e Alvarenga (2004) apontam que a *Web Semântica* faz uso, ou seja, adota padrões de metadados e compartilha estes padrões para melhor utilizar o imenso repositório de informação disponibilizado no ambiente *Web*.

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 (SOUZA; ALVARENGA, 2004, p.139).

Nessa perspectiva, identificamos na comunidade científica algumas iniciativas que dizem respeito ao desenvolvimento de bibliotecas digitais, bem como à construção de padrões de representação e descrição dos conteúdos, e que se utilizam de tecnologias subjacentes à *Web Semântica*, tais como o *JeromeDL* e o *MarcOnt Initiative*, sendo o último objeto de estudo de um dos autores deste artigo.

O *JeromeDL* é um projeto integrado entre a *National University of Ireland - NUI* - e *Digital Enterprise Research Institute - DERI* - e pode ser localizado no site: <http://www.jeromedl.org/>.

A iniciativa *MarcOnt* é o resultado dos estudos de Sebastian Ryszard Kruk, que apresenta a proposta de um novo padrão para representação e descrição de recursos informacionais em bibliotecas digitais no âmbito da *Web Semântica*. As informações sobre o projeto estão disponíveis no Portal *MarcOnt*: <http://portal.marcont.org>. Vale destacar que os estudos referentes ao projeto *MarcOnt Initiative* contam com a parceria da *National University of Ireland/Galway, Leopold Franzens*

University/Innsbruck, Gdansk University of Technology/Gdansk e Digital Enterprise Research Institute/Galway.

O esforço da *Web Semântica*, parcialmente originado a partir da comunidade de bibliotecas digitais, está provendo tecnologias como as ontologias que podem ser potencialmente aplicadas para o problema de desenvolvimento e descrição dos recursos informacionais. (KRUK; SYNAK; ZIMMERMANN, 2005, tradução nossa).

Cada tecnologia aplicada nas bibliotecas digitais e na *Web Semântica* realiza tarefas específicas. Os metadados, por exemplo, garantem as formas de representação dos recursos informacionais, e as ontologias, neste contexto, determinam semanticamente o conceito dos dados dos recursos (dos próprios metadados).

Dessa maneira, observa-se que os metadados são instrumentos fundamentais na padronização da representação dos recursos informacionais aplicados em distintas ambiências e que através destes, serão otimizados o acesso e as possibilidades de uso das informações.

4 CONSIDERAÇÕES FINAIS

Os metadados assumem importância vital para as bibliotecas digitais e para o ambiente *Web* de um modo geral, pois são estes que determinarão

a padronização para a descrição dos recursos, garantirão a integridade da apresentação dos recursos informacionais e facilitarão a interoperabilidade entre os sistemas informacionais, o que garantirá a acessibilidade de forma eficiente para diferentes tipos de usuário.

Como pode ser observado na literatura, a temática bibliotecas digitais abordada no texto apresenta-se como um desafio e traz inúmeras possibilidades para a exploração da ambiência digital por parte de uma diversidade de profissionais, com destaque para os da área da Ciência da Informação, no tocante às questões de representação e descrição de conteúdo e forma dos recursos informacionais, bem como na mediação destes por profissionais da informação.

Com a *Web Semântica*, verifica-se que os metadados aplicados nesse contexto são considerados elementos-chave na padronização da representação, na descrição e na apresentação dos recursos informacionais, no propósito de disponibilizar os conteúdos digitais pautados em normas e padrões, de modo a garantir a localização, o acesso e a recuperação dos recursos informacionais de maneira eficiente, potencializando o uso e o reuso, bem como a preservação das informações, como instrumentos tecnológicos de importância capital na estrutura das bibliotecas digitais.

THE METADATA AS TECHNOLOGICAL TOOLS IN STANDARDIZATION AND POTENTIALIZATION OF INFORMATIONAL RESOURCES IN THE SCOPE OF THE DIGITAL LIBRARIES IN THE SEMANTIC WEB ERA

ABSTRACT

The digital libraries, as a promising environment for the retrieval of information, with the metadata use, make possible the standardization of the representational forms and the guarantee of interoperability among systems, favoring the integrity and accessibility to the informational resources in an efficient way by the final user. Based on the literature available, it is pointed out the use of metadata in the building of digital libraries as key elements in the consolidation of the Semantic Web. It is highlighted the representation and description of the shape and content of the informational resources as the focus in the search for initiatives for the building of representational patterns and description of resources which use technologies that underlie the Semantic Web.

Keywords

DIGITAL LIBRARIES
METADATA
SEMANTIC WEB
STANDARDIZATION AND REPRESENTATION OF INFORMATION
INTEROPERABILITY

Artigo recebido em 13.04.2007 e aceito para publicação em 02.05.2007

REFERÊNCIAS

- ALVES, R. C. V. *Web semântica: uma análise focada no uso de metadados*. 180 f. 2005. Dissertação (Mestrado em Ciência da Informação) - Faculdade de Filosofia e Ciências, Universidade Estadual Paulista, Marília, 2005.
- ARELLANO, M. A. Preservação de documentos digitais. *Ciência da Informação*, Brasília, v. 33, n. 2, p. 15-27, maio/ago. 2004.
- BERNERS-LEE, T. *Semantic Web Concepts*. 2005. Disponível em: <<http://www.w3.org/2005/Talks/0517-boit-tbl>>. Acesso em: 13 nov. 2006.
- BERNERS-LEE, T. ; HENDLER, 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. Disponível em: <<http://www.sciam.com/2001/0501issue/0501berners-lee.html>>. Acesso em: 28 jun. 2005.
- BORGMAN, C. L. *Social aspects of digital libraries*. In: WORKSHOP ON SOCIAL ASPECTS OF DIGITAL LIBRARIES, 1996, Los Angeles. *Final Report...* Los Angeles: UCLA/NSF, 1996. Disponível em: <<http://www.lis.gseis.ucla.edu/DL>>. Acesso em: 27 jul. 2006.
- CARDOSO, J. C.; OLIVEIRA, J. B. *Uma proposta de interface para bibliotecas digitais configuráveis ao perfil do usuário*. Porto Alegre, 2000. Plano de Estudo e Pesquisa.
- KRUK, S. R.; SYNAK, M.; ZIMMERMANN, K. *MarcOnt initiative: mediation services for digital libraries*. 2005. Disponível em: <<http://www.marcOnt.org/marcOnt/pdf/marcontecdl2005.pdf>>. Acesso em: 16 jan. 2006.
- LESK, M. *Practical digital libraries: books, bytes and bucks*. San Francisco, CA: Morgan Kaufmann, 1997.
- MACHADO, R. N.; NOVAES, M. S. F.; SANTOS, A. H. Biblioteca do futuro na percepção de profissionais da informação. *Transinformação*, Campinas, v. 11, n. 3, p. 215-222, set./dez. 1999.
- MILSTEAD, J.; FELDMAN, S. *Metadata: cataloging by any other name*. Online, [S. l.], January 1999. Disponível em: <<http://www.online.com/online/ol1999/milstead1.html>>. Acesso em: 22 mar. 2007.
- MODESTO, F. *Metadados: introdução básica*. São Paulo, 2005. Apostila de aula da disciplina representação descritiva II. Disponível em: <<http://www.eca.usp.br/prof/fmodesto/textos/metadado2005f.pdf>>. Acesso em: 25 mar. 2007.
- SEMANTIC WEB. Disponível em: <<http://www.semanticweb.org/about.html>>. Acesso em: 25 ago. 2005.
- SOUZA, R. R.; ALVARENGA, L. A web semântica e suas contribuições para a ciência da informação. *Ciência da Informação*, v. 33, n. 1, p. 132-141, jan./abr. 2004.
- SOUZA, T. B. de; CATARINO, M. E.; SANTOS, P. C. dos. Metadados: catalogando dados na Internet. *Transinformação*, v. 9, n. 2, maio/ago. 1997. Disponível em: <<http://www.biblioestudantes.hpg.ig.com.br/146.htm>>. Acesso em: 05 mar. 2007.
- TOUTAIN, L. M. B. B. Biblioteca digital: definição de termos. In: MARCONDES, C. H. et al. (Org.). *Bibliotecas digitais: saberes e práticas*. 2. ed. Salvador: EDUFBA; Brasília: IBICT, 2006. p. 15-24.
- WAINWRIGHT, E. *Digital libraries: some implications for government and education from the Australian development experience*. Canberra, Austrália: National Library of Australia, 1996. Disponível em: <<http://www.nla.gov.au/nla/staffpaper/ew6.html>>. Acesso em: 12 mar. 2007.

AGRADECIMENTOS

À CAPES, pelo financiamento dessa pesquisa.