

REVENDO AS PRÁTICAS DE ORGANIZAÇÃO DA INFORMAÇÃO NO INFOBCI: PRIMEIROS ESTUDOS SOBRE O VOCABULÁRIO CONTROLADO

REVIEWING THE PRACTICES OF INFORMATION ORGANIZATION IN INFOBCI: FIRST STUDIES ON CONTROLLED VOCABULARY

Laura Inafuko linafuko@hotmail.com
Universidade Estadual Paulista

Pedro Andretta andretta_pedro@yahoo.com.br
Universidade Federal de São Carlos

Renan Ramos renan@nit.ufscar.br
Universidade Federal de São Carlos

Resumo

Tendo em vista a capacidade da web para a disseminação e compartilhamento da informação, foi criado um ambiente virtual dedicado à comunidade interessada em Ciência da Informação. Neste espaço, denominado InfoBCI (disponível em: <<http://www.infobci.wordpress.com>>), os artigos publicados nas revistas nacionais são indexados usando um vocabulário controlado sugerido para organização de acervos de Biblioteconomia e Ciência da Informação e aberto um espaço para comentários sobre as publicações. Entretanto, percebeu-se que os termos sumarizados não vem sendo eficazes na representação dos conteúdos. Neste contexto, parte de um estudo para a determinação de um vocabulário controlado mais adequado à organização e compartilhamento das informações científicas. Explorou-se como temática o uso e importância desse instrumento de representação, levantando algumas questões sobre a criação ou adoção de índices já prontos na organização da informação, comparando ainda a lista de assuntos adotados no blog com outros voltados a Ciência da Informação.

Palavras-chave

Organização da Informação. Blog InfoBCI. Vocabulário Controlado.

INTRODUÇÃO

A crescente utilização da Internet como recurso de acesso à informação tem contribuído para mudanças nos quadros de distribuição e disseminação do conhecimento na comunidade científica, acarretando na reestruturação do sistema de comunicação da ciência. Fixada antes pelas vias orais e impressas, é na Internet que se firma o terceiro canal de comunicação científica: a comunicação científica eletrônica.

A comunicação científica ocorre quando há compartilhamento de conhecimento entre membros de uma comunidade científica, neste processo há pelo menos duas atividades: a produção e a disseminação do conhecimento.

No que se refere à disseminação do conhecimento, percebe-se que vem sendo muito comum a disponibilização livre e gratuita de publicações científicas na internet. Sobre isso, Mueller (2006, p.27) afirma que “o movimento para acesso livre ao conhecimento científico pode ser considerado como o fato mais interessante e talvez importante de nossa época no que se refere à comunicação científica”, entretanto, a produção científica, sempre crescente, encontra-se espalhada na *web*.

Com o propósito de complementar o suporte oferecido por bases de dados de acesso público que indexam as publicações brasileiras relacionados à área de Ciência da Informação (como BRAPCI¹, Diálogo Científico², Holmes³, Catálogo da Biblioteca do IBICT⁴, Peri⁵ e SciELO⁶), foi criado o *blog* InfoBCI⁷, para acesso de pesquisadores e interessados. No *blog* são divulgadas e organizadas as publicações de cada periódico científico eletrônico brasileiro, de forma a possibilitar a interação de pesquisadores por meio da disponibilização de um espaço para comentários, permitindo o compartilhamento de experiências, além ainda de oferecer inúmeros outros recursos, tais como *links* para artigos segundo sua temática, para cada periódico indexado, para *blogs* da área e fontes de informação em Ciência da Informação.

Segundo Barreto (2007, s/p), a Ciência da Informação “[...] se preocupa e se ocupa com os princípios teóricos e as práticas da criação, organização e distribuição da

¹ Disponível em: <<http://brapci.ufpr.br/>>

² Disponível em: <<http://dici.ibict.br/view/subjects/6.07.00.00.html/>>

³ Disponível em: <<http://holmes.feudo.org/>>

⁴ Disponível em: <<http://biblioteca.ibict.br/>>

⁵ Disponível em: <<http://base.eci.ufmg.br/>>

⁶ Disponível em: <<http://scielo.br/>>

⁷ Disponível em: <<http://infobci.wordpress.com/>>

informação”, isto é, a função da Ciência da Informação é procurar soluções para os problemas relacionadas à informação, sendo responsável

[...] pelo aperfeiçoamento técnico da fluidez da informação, através de procedimentos, técnicas e tecnologias que buscam promover a interação entre usuários e informação, considerando a infra-estrutura digital e seus impactos na disseminação informacional. (MOURA, 2009, p. 60)

Com o objetivo de cobrir seu objeto de estudo e com as atividades relacionadas a ela, a ciência em questão trabalha como descreve Le Coadic (c1994, p.22) “[...] com conhecimentos advindos principalmente, a psicologia, a lingüística, a sociologia, a informática, a matemática, a lógica, a estatística, a eletrônica, a economia, o direito, a filosofia, a política e as telecomunicações”.

Para dar conta da organização da informação e permitir a fácil recuperação dos conteúdos informacionais, a Ciência da Informação usa de representações semânticas para descrever abreviadamente objetos e ideias. Essas substituem aquilo que representam, entretanto por melhores que pareçam, as representações não abarcam a ideia total do conteúdo informacional propriamente dito, além de estarem sujeitas a um determinado ponto de vista, dando foco a determinados conceitos em detrimento de outros.

Considerando que “[...] uma recuperação eficaz da informação constante dos documentos, por parte dos usuários tem sido considerada a meta suprema dos profissionais dos serviços de informação.” (ALVARENGA, 2003, p. 29), e esta só ocorre quando são trabalhadas as práticas de organização da informação. a presente pesquisa recorre à teoria da representação temática da informação, com ênfase nos vocabulários controlados para levantar as questões sobre sua utilização e importância, assim como entender a ação de seu funcionamento tanto do lado do indexador como do utilizador do sistema de informação.

Este trabalho parte dos estudos relacionados à adaptação da linguagem documentária adotada no blog InfoBCI. Pretende-se, ainda, extrair as primeiras conclusões quanto à abrangência da atual lista de assuntos e como podem ser percebidos os índices de recuperação de precisão e revocação comparando-os com outros vocabulários controlados na área de Ciência da Informação.

Com o desenvolvimento da *Web 2.0*, surgiram possibilidades de elaboração de espaços interativos entre usuários, nos quais é possível criar, avaliar e colaborar conteúdos em documentos digitais. Assim, o *blog* InfoBCI foi criado com o objetivo de se tornar um canal de comunicação entre estudantes, profissionais, pesquisadores e a comunidade em geral, interessados na área de Ciência da Informação.

Este espaço de comunicação vincula-se a outro canal já formalizado pela comunidade científica. Em verdade há muitas formas de comunicar a ciência como em congressos, palestras e visitas, mas pode-se dizer que:

Dentro do ciclo produtivo da ciência, o periódico científico mantém-se como o principal veículo de comunicação formal dos resultados de pesquisas originais. Instrumental qualitativo, reproduz as sanções e exigências próprias do campo científico, aponta o grau de evolução de cada área de conhecimento, estabelece a propriedade intelectual, legitima novos campos de estudos, é um índice nos sistemas de julgamento e distribuição de verbas para a produção científica. (GRUSZYNSKI; GOLIN, 2007)

Dessa forma, o *blog* divulga as publicações de artigos de 19 periódicos científicos eletrônicos brasileiros, criando um espaço para análise e comentários do público leitor. Este ambiente caracteriza-se pela possibilidade de interação entre pesquisadores e a comunidade em geral por disponibilizar um espaço para publicações de comentários, o que permite o compartilhamento de experiências, sugestões e críticas.

As visitas ao *blog* têm crescido progressivamente, sendo que no primeiro mês da criação, em janeiro de 2009, o *blog* registrou 19 visitas, e nos seguintes 243, 400, 901 e 1413, nos meses respectivos. Atualmente, o *blog* conta com cerca de 45 visitas por dia, sendo que no último mês (julho/2009) o *blog* recebeu cerca de 1550 visitas.

Como pode ser visto na figura 1, o *blog* está dividido em 3 colunas, sendo elas: coluna de postagem, barra lateral 1 e barra lateral 2. Na coluna de postagem se encontra os textos indexados, em que cada artigo é publicado individualmente com: sua referência, resumo e o link para acesso ao texto completo, e uma representação do assunto, de forma a possibilitar sua posterior recuperação.

InfoBCI
Sempre Informando!

Representação descritiva e temática de recursos de informação no sistema agência Embrapa: uso do padrão dublin core

06/08/2009 in Revista Digital de Biblioteconomia e Ciência da Informação with Leave a Comment
Tags: Metadados, Representação da Informação

SOUZA, Marcia Izabel Fujisawa; ALVES, Maria das Dores Rosa. Representação descritiva e temática de recursos de informação no sistema agência Embrapa: uso do padrão dublin core. **Revista Digital de Biblioteconomia e Ciência da Informação**, Campinas, v.7, n. 1, p. 208-223, jul./dez. 2009. Disponível em: <<http://polaris.bc.unicamp.br/seer/ojs/include/getdoc.php?id=6978&article=2098&mode=pdf>> Acesso em: 6 ago. 2009.

Resumo
O artigo relata a experiência da Embrapa na utilização do padrão de metadados Dublin Core (DC) na representação descritiva e temática de recursos de informação eletrônicos. É apresentada uma descrição sucinta de cada elemento metadado, incluindo sua definição, qualificadores e valor. A adoção do padrão Dublin Core teve sua origem no âmbito de um projeto de pesquisa dedicado à organização e tratamento da

PÁGINAS
Apresentação
Autores
Lista de Assuntos
Lista dos periódicos e Qualis

CATEGORIAS
Ciência da Informação (9)
CRB-8 Digital (6)
Cultura em recorte (2)
DataGramaZero (19)
Encontros Bibli (20)
Informação & Informação (7)
Informação & Sociedade: Estudos (9)
LIINC em Revista (9)
Perspectivas em Ciência da Informação (14)
Ponto de Acesso (5)
Revista ACB (20)
Revista Brasileira de Biblioteconomia e Documentação (9)

search here

FEEDS
Assine seu feed!
Assine seu feed por e-mail
QUE DIA É HOJE?
Agosto 2009
D S T Q J S S
1
2 3 4 5 6 7 8
9 10 11 12 13 14 15
16 17 18 19 20 21 22
23 24 25 26 27 28 29
30 31
< Jul
ARQUIVOS
Agosto 2009
Julho 2009
Junho 2009

Figura 1 – InfoBCI: Sempre informando!
Fonte: <http://infobci.wordpress.com/>.

Nas barras laterais estão as opções de navegação do *blog*, como a organização de textos pelo nome dos periódicos em que são publicados, ferramenta de pesquisa, links para os periódicos eletrônicos, links para *blogs* da área e bases de dados para pesquisa, opção para assinatura do *feed* do *blog* e opção para recuperar os artigos por seus temas, entre outros.

Para representar os artigos segundo o assunto, a equipe do *blog* InfoBCI utiliza a proposta de classificação para Trabalhos de Conclusão de Curso da área de Biblioteconomia e Ciência da Informação da Universidade Federal de São Carlos, elaborada por Cerantola et al (2007), sendo 48 termos distribuídos em 6 grandes categorias.

No entanto, a partir das indexações dos artigos, percebeu-se que a proposta adotada é insuficiente para abarcar os assuntos referentes aos artigos, de forma que alguns destes foram publicados no *blog* sem as *tags* de assunto, podendo prejudicar as possíveis buscas do usuário por um artigo específico.

Dessa forma, pretende-se com este trabalho estudar e reformular o vocabulário controlado do *blog*, de modo que apresente melhorias para a busca e recuperação dos artigos indexados, assim como facilitar a tarefa de indexação para a equipe do *blog* InfoBCI.

LINGUAGENS DE INDEXAÇÃO E VOCABULÁRIOS CONTROLADOS

Para entender o que são vocabulários controlados e as implicações de seu uso é necessário para um bom entendimento localizar essa região temática dentro dos estudos da Ciência da Informação. Para isso recorreremos ao ciclo documentário conforme figura a seguir.

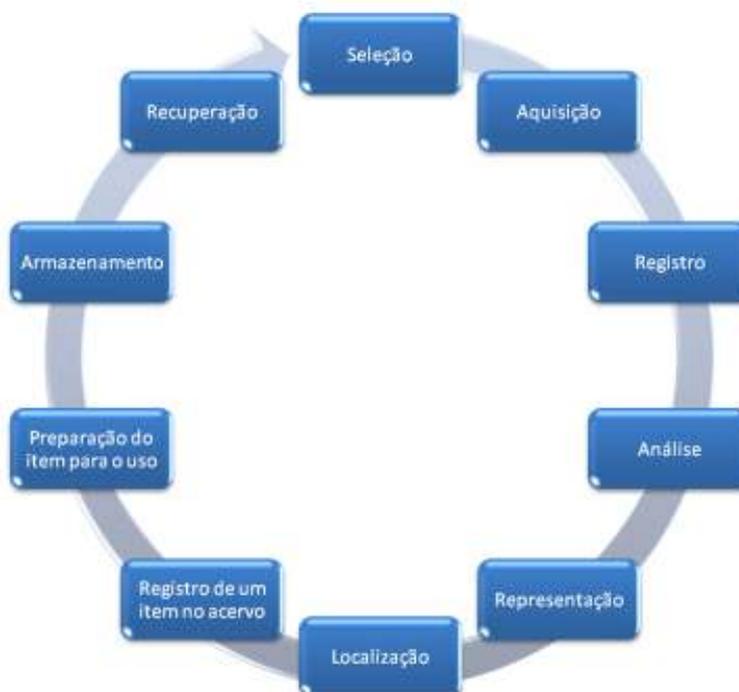


Figura 2 – Ciclo documentário
Fonte: Própria

Para a Ciência da Informação há um processo denominado representação da informação, que se subdivide em pelos menos dois segmentos: o de representar a informação por meio da descrição do suporte (ênfase na localização da informação) e por meio da representação temática do conteúdo (ênfase no acesso a informação). Nesse contexto, pode-se distinguir a análise documentária como parte dos procedimentos para representação temática da informação, uma vez que oferece suporte para a representação dos conceitos, que se fará por meio de uma linguagem, seja ela natural ou controlada.

Segundo Novellino (1996, p. 38) o processo de representação temática do conteúdo envolve principalmente dois passos:

- a) Análise de assunto de um documento e a colocação do resultado desta análise numa expressão linguística.
- b) Atribuição de conceitos ao documento analisado

Para atribuir conceitos aos documentos é necessário o uso de uma linguagem, que represente com fidedignidade o conteúdo informacional registrado no suporte, substituindo-o para fins de recuperação. Em verdade Campos (2004) afirma:

[...] Fidelidade perfeita é, em geral, impossível, tanto na prática quanto em princípio. A única representação completamente precisa de um objeto é o objeto em si. Qualquer outra representação é imprecisa e, inevitavelmente, contém simplificações. (CAMPOS, 2004, p. 24)

Por conta disso, tem-se buscado técnicas que permitam uma boa representação, neste intento, estabeleceram-se duas formas de trabalhar as linguagens de indexação, uma natural e outra controlada.

A utilização de linguagem natural ocorre quando um termo é extraído do próprio texto para que figure como meio de recuperação do documento; já a linguagem controlada ocorre quando este termo é eleito segundo um conjunto de termos autorizados e limitado.

As linguagens documentárias são instrumentos que tem por função representar sinteticamente o conhecimento apresentado no documento, padronizando as entradas de assunto e tornando possível uma melhor recuperação da informação, haja vista, que o vocabulário utilizado torna-se comum tanto para usuário como indexadores.

Lopes (2002) aponta as vantagens e desvantagens do uso de vocabulário controlado, que podem ser vistos no seguinte quadro:

VANTAGENS	DESVANTAGENS
<ol style="list-style-type: none"> 1. Controle total do vocabulário de indexação, minimizando os problemas de comunicação entre indexadores e usuários. 2. Com o uso de um <i>thesaurus</i> e suas respectivas notas de escopo, os indexadores podem assinalar mais corretamente os conceitos dos documentos. 3. Se bem constituído, o vocabulário controlado poderá oferecer alta recuperação e relevância e, também, ampliar a confiança do usuário diante de um possível resultado negativo. 4. As relações hierárquicas e as remissivas do vocabulário controlado auxiliam tanto o indexador, quanto o usuário na identificação de conceitos relacionados. 5. Redução no tempo de consulta à base, pois a estratégia de busca será mais bem elaborada com o uso do <i>thesaurus</i>. 	<p>Custos: a produção e manutenção da base de dados terá despesas maiores com a equipe de indexadores. Será necessário ainda manter pessoal especializado na atualização do <i>thesaurus</i>.</p> <p>O vocabulário controlado poderá não refletir adequadamente os objetivos do produtor da base, caso esteja desatualizado.</p> <p>Um vocabulário controlado poderá se distanciar dos conceitos adequados para a representação das necessidades de informação dos usuários.</p> <p>Necessidade de treinamento no uso dos vocabulários controlados tanto para os intermediários, quanto para os usuários finais.</p> <p>Desatualização do vocabulário controlado poderá conduzir a falsos resultados.</p>

Quadro 1 – Vantagens e desvantagens do uso de vocabulário controlado

Fonte: Lopes (2002)

Para Vale (1987, p. 12) as linguagens documentárias podem ser divididas em dois grupos:

- Pré-coordenadas, que combinam ou coordenam os termos no momento da indexação. São utilizados principalmente em sistemas manuais: catálogos de bibliotecas, bibliografias e índices impressos.
- Pós-coordenados, que combinam ou coordenam os termos no momento da busca. São utilizados principalmente em sistemas automatizados: bases/banco de dados, bibliotecas ou centros de documentação que operam com computador.

As primeiras linguagens documentárias foram as listas de cabeçalhos de assunto que adotavam um sistema pré-coordenado, passando-se depois para sistemas pós-coordenados sem um controle de vocabulário e retomando-se adoção de termos controlados em sistemas pós-coordenados com listas de termos autorizados (para se evitar falsas coordenações).

Os estudos em linguagens documentárias tiveram seu desenvolvimento intensificado na década de 60 devido ao aumento explosivo da produção do conhecimento em âmbito científico e tecnológico. Nesse contexto, têm hoje se mostrado essencial para representação sistemática e unívoca da informação e do conhecimento disponível em quaisquer suportes de informação como afirma Lara (2004, p. 232):

[...] no momento em que as tecnologias digitais têm sido utilizadas na comunicação, as metodologias de construção da linguagem documentária são referidas como meios para tratar também universos não exclusivamente bibliográficos, como sites na WEB, conteúdos de manuais técnicos, de cd-roms de distintos assuntos, etc. Essas metodologias são essenciais à arquitetura da informação.

Diante desse contexto, a utilização e o conhecimento da base teórica do campo de estudo da linguagem documentária é essencial para o desenvolvimento de instrumentos que auxiliem a representação da informação e do conhecimento e posteriormente sua recuperação.

COMPARAÇÃO ENTRE VOCABULÁRIOS CONTROLADOS DISPONÍVEIS EM CIÊNCIA DA INFORMAÇÃO

Com o propósito de analisar o atual vocabulário controlado adotado no blog InfoBCI e extrair as primeiras conclusões quanto sua adaptação, recorreu-se a três

esquemas temáticos e classificatórios em Ciência da Informação (entre eles o esquema atualmente em uso), comparando-os quanto aos índices de precisão e revocação.

Quanto aos índices, Le Coadic (c1994, p.63) afirma que “a revocação é a proporção de documentos relevantes recuperados, enquanto a precisão mede a proporção de documentos recuperados que são relevantes”. Isso pode ser entendido pensando no quadro e fórmulas descritas abaixo:

Documentos	Relevantes	Não Relevantes
Recuperados	A	B
Não recuperados	C	D

Quadro 2: Índices de revocação e precisão
Fonte: Própria

Em que temos:

$$\text{Precisão} = A / (A + B)$$

$$\text{Revocação} = A / (A + C)$$

A recuperação ideal seria aquela na qual todos os documentos recuperados são relevantes (boa precisão) ou ainda o inverso, todos os documentos relevantes são recuperados (boa revocação).

Os esquemas temáticos em Ciência da Informação foram analisados a partir do assunto mais frequente nas indexações do blog InfoBCI, sendo a comparação feita a partir do trabalho de Odonne e Gomes (2004), Vocabulário Controlado do SIBi/USP, além ainda do trabalho de Cerantola et al (2007), sendo este último, utilizado atualmente para indexação dos registros.

Cerantola et al (2007) apropria-se do estudo desenvolvido por Pinheiro (2002), adaptando a sistemática oferecida no trabalho para uso na organização do acervo dos trabalhos de conclusão de curso em Biblioteconomia e Ciência da Informação da Universidade Federal de São Carlos. Pinheiro (2002) trabalha os desdobramentos disciplinares, interdisciplinares e transdisciplinares da Ciência da Informação, refazendo os passos empregados em sua tese de doutorado, analisando as publicações do *Annual Review of Information Science and Technology – ARIST* entre os anos de 1979 a 2004, além de verificar as transformações nos temas de pesquisa e montando um esquema para os temas mais recorrentes segundo seis grandes eixos disciplinares.

Segundo Oddone e Gomes (2004), a ausência de uma taxonomia em Ciência da Informação acaba por dificultar o processo de transmissão do conhecimento e dessa forma a sua ampla difusão, acarretando em uma barreira para a consolidação da identidade da disciplina e sua visibilidade no cenário acadêmico-científico nacional. Nesse âmbito, as autoras procuram construir um sistema taxonômico adaptado aos assuntos quem vem sendo estudados e publicados no Brasil.

Para a construção desse sistema, Oddone e Gomes (2004) recorrem à tabela utilizada por Neusa Dias de Macedo na criação do Catálogo de Teses e Dissertações da Ciência da Informação, os esquemas desenvolvidos por Kalervo Järvelin e Pertti Vakkari nos anos de 1990 e 1993, as tabelas desenvolvidas por Leilah Santiago Bufrem e Sonia Kasuko Sakai Teixeira, além do sistema adotado no repositório eletrônico E-LIS e o *Thesaurus of Information Science*, da American Society for Information Science & Technology. Quando concluído o esquema, as autoras tomam conhecimento da proposta dos autores Hawkins, Larson e Caton (2003), cujos objetivos convergiam para o mesmo fim, e reveem a proposta, consolidando por fim a estrutura.

A criação do Vocabulário Controlado do SIBi/USP se deu junto a criação do catálogo geral informatizado da Universidade de São Paulo (USP) no ano de 1985, com o objetivo de atender a todas as unidades de informação da instituição. Esse catálogo foi planejado para o acesso on-line e foi chamado de Banco de Dados Bibliográficos da USP - DEDALUS.

A Lista de Assuntos USP foi desenvolvida com o objetivo de criar uma linguagem que pudesse representar a temática do acervo de forma comum para todas as bibliotecas da USP. A princípio essa Lista era composta por oito mil entradas, atualmente, esta Lista passou a ser chamada Vocabulário Controlado do SIBi/USP sendo dividido em 7 grandes áreas do conhecimento, e estas se subdividem em outras 37 áreas.

MÉTODOS E RESULTADOS

Para iniciar as comparações entre os vocabulários, foi elaborada uma lista com os assuntos (*tags*) do *blog* InfoBCI e a quantidade de vezes que estes apareceram, conforme pode ser observado no quadro de uso das *tags* – Marcadores de assunto, que segue abaixo.

Aplicações de informação	7	Informação científica	2
Arquitetura de informação	1	Informação em arquivos	12
Automação de bibliotecas	2	Informação em Arte	1
Ética na informação	1	Informação em bibliotecas	29
Bases de dados	1	Informação em museus	3
Bibliometria / Informetria/ Cientometria / Webmetria	4	Informação industrial	0
Bibliotecas Virtuais e Digitais	1	Informação tecnológica	3
Catologação	4	Inteligência Competitiva	6
Classificação	4	Internet/Web	7
Competência informacional	7	Marketing de informação	1
Comunicação científica	6	Metadados	1
Comunicação mediada por computador	4	Metodologias da Ciência da Informação	2
Disseminação da informação	2	Mineração de dados	0
Divulgação científica	0	Ontologia	2
Economia da informação	2	Organização do Conhecimento	4
Educação à distância	6	Organização e Processamento da Informação	4
Epistemologia da Ciência da Informação	1	Políticas de informação	7
Estudos de necessidades e usos de informação	4	Preservação e Segurança digital	0
Estudos de usuários	5	Processamento automático de linguagem	1
Estudos Interdisciplinares	27	Redes e Sistemas de informação	2
Formação Profissional	16	Representação da Informação	6
Fundamentos da Ciência da Informação	2	Sistemas de recuperação da informação	7
Gestão da Informação	11	Tecnologias da Informação	19
Gestão de Qualidade de Informação	2	Teoria da Informação	1
Gestão do Conhecimento	4	Tesauros	2
História da Ciência da Informação	1	Transferência de Informação	2
Inclusão digital	13	Vocabulários controlados	1
Indexação	2		

Quadro 3: Uso das tags

Fonte: Própria

É necessário lembrar que cada artigo é indexado o número de termos que forem necessários, segundo a exposição no resumo e palavras-chave. No entanto, quando não se identifica a representação precisa do conteúdo esta é expressa pelo termo mais

aproximado por conceito. A utilização do resumo e palavras-chave não implica em descrédito na indexação quando pretende-se apenas descrição do assunto geral do documento, sem entrar em detalhes quanto: ação, objeto, agente, métodos do agente, local e causa-efeito.

Depois de identificados os assuntos mais frequentes, foram selecionados os três mais recorrentes e quantificados os grupos temáticos e as representações para aquele assunto nas diversas listas como visto no quadro abaixo:

Nome	Quantidade de grandes grupos temáticos	Representações para		
		Bibliotecas	Estudos Interdisciplinares	Tecnologias da informação
PINHEIRO, 2002	6	1	1	9
ODDONE, GOMES, 2004	11	5	28	46
Vocabulário Controlado do SIBi/USP	7	23	0	0

Quadro 4: Assuntos mais frequentes
Fonte: Própria

Observando a variedade numérica em relação aos grandes grupos temáticos em Ciência da Informação (que pode ser visto em Apêndice A), pode-se dizer que isto se deve as diversas visões da ciência, reflexo inclusive da própria metodologia empregada para seccioná-la. Exemplo disso é o fato da Lista 3 do Apêndice A não possuir representações em sua estrutura taxonômica para conceitos como “Estudos interdisciplinares” e “Tecnologias da informação” na qual certamente devem estar expressos em outras Ciências.

O fato de uma lista (vocabulário controlado) possuir mais ou menos representações para um objeto ou idéia implicará respectivamente em uma maior precisão ou revocação na recuperação da informação.

CONSIDERAÇÕES PARCIAIS

Este estudo percorreu a teoria das linguagens documentárias atendo-se à prática de uso de vocabulários controlados para a representação e recuperação da informação, comparando quantitativamente como são feitas as representações de três assuntos em

três vocabulários controlados brasileiros construídos para indexar publicações em Ciência da Informação.

Apesar da limitação metodológica (que se fez a partir de uma pequena avaliação quantitativa dos termos para a representação da informação), esses primeiros estudos permitiram verificar algumas das consequências proveniente da adaptação do vocabulário do blog InfoBCI, além de levantar algumas afirmações tais como:

- O uso das linguagens documentárias permite a comunicação entre o sistema e o usuário;
- Os vocabulários controlados são instrumentos úteis na padronização das indexações e, por conseguinte, na sua recuperação;
- É possível que haja inúmeras taxonomias ou modelos para qualquer ciência, sendo que uma pode ser mais adequada a um determinado propósito que outra;
- A metodologia de construção de uma taxonomia influencia em sua estrutura e conseqüentemente na quantidade de entradas (termos);
- A quantidade de entradas no vocabulário controlado irá influenciar diretamente os índices de precisão e revocação no momento de recuperação das informações;
- Um sistema no qual possibilite alta precisão não é necessariamente melhor que outro que ofereça alta revocação, pois isso está ligado aos objetivos de construção do sistema e seu público alvo.

Assim, se obtém como conclusão preliminar que antes da adoção ou criação de um vocabulário para uso no *blog* InfoBCI na organização temática dos artigos científicos publicados nos periódicos nacionais é necessário antes que a equipe reveja o objetivo do *blog*, estude a movimentação dos usuários nas buscas e a infra-estrutura de sua plataforma.

Por fim acredita-se que o uso da inteligência coletiva na seleção dos termos-chave para cada publicação poderá ser uma alternativa para proporcionar um maior compartilhamento das informações e difusão da ciência. Entretanto, ainda sim, os estudos de representação e construção de vocabulários não são descartados haja vista que convenções regionais e ideológicas acabam por fragmentar e dispersar o conhecimento científico.

REFERÊNCIAS

ALVARENGA, L. Representação do conhecimento na perspectiva da Ciência da Informação em tempo e espaço digitais. *Encontros Bibli*, v.1, n.15, jan/mar. 2003.

BARRETO, A. Glossário sobre a Ciência da Informação. *DataGramaZero*, v.8 n.1 fev. 2007. Disponível em: <http://dgz.org.br/fev07/Ind_com.htm> Acesso em: 08 ago. 2009.

CAMPOS, M. L. de A. Modelização de domínios de conhecimento: uma investigação de princípios fundamentais. *Ciência da Informação*, v.33, n.1, p.22-32, jan./abril 2004. Disponível em: <<http://revista.ibict.br/ciinf/index.php/ciinf/article/view/56/53>> Acesso em: 04 set. 2009.

CERANTOLA, A. et al. O acervo de tccs do curso de biblioteconomia e ciência da informação da UFSCar: uma nova perspectiva de classificação. In: ENCONTRO NACIONAL DE ESTUDANTES DE BIBLIOTECONOMIA, CIÊNCIA E GESTÃO DA INFORMACAO E DOCUMENTACAO, 30., 2007, São Carlos. *Anais...* São Carlos: s.n., 2007.

GRUSZYNSKI, A. C.; GOLIN, C. Periódicos científicos e eletrônicos e a visibilidade da ciência na web: estudo de caso na UFRGS. *DataGramaZero*. v.8, n.3, jun., 2007. Disponível em: <http://www.dgz.org.br/jun07/Art_02.htm> Acesso em 22 dez., 2008.

HAWKINS, D. T., LARSON, S. E., CATON, B. Q. Information science abstracts : tracking the literature of information science. Part 2 : a new taxonomy for information science. *Journal of the American Society for Information Science and Technology*, v. 54, n. 8, p. 771-781, 2003.

LARA, Marilda Lopes Ginez de. Linguagem documentária e terminologia. *Transinformação*, Campinas, v 16, n. 3, 2004. Disponível em: <<http://revistas.puccampinas.edu.br/transinfo/viewissue.php?id=8>>. Acesso em: 23 jun. 2009.

LE COADIC, Y. *A ciência da informação*. Tradução de Maria Yeda F.S. de Filgueiras Gomes. Brasília: Briquet de Lemos, c1994.

LOPES, I. L. Uso das linguagens controlada e natural em bases de dados: revisão da literatura. *Ciência da Informação*, v. 31, n. 1, p. 41-52, jan./abr. 2002. Disponível em: <<http://revista.ibict.br/ciinf/index.php/ciinf/article/view/172/151>> Acesso em: 04 ago. 2009.

MOURA, M. A. Informação, ferramentas ontológicas e redes sociais ad hoc: a interoperabilidade na construção de tesouros e ontologias. *Informação & Sociedade: Estudos*, v.19, n.1, p. 59-73, jan./abr. 2009. Disponível em: <<http://www.ies.ufpb.br/ojs2/index.php/ies/article/view/2396/2688>>. Acesso em: 30 abr. 2009.

MUELLER, S. P. M. A comunicação científica e o movimento de acesso livre ao conhecimento. *Ciência da Informação*, Brasília, v. 35, n. 2, p. 27-38, maio/ago. 2006. Disponível em: <<http://revista.ibict.br/index.php/ciinf/article/view/826/668>> Acesso em: 08 ago. 2009.

NOVELLINO, M. S. F. Instrumentos e metodologias de representação da informação. *Informação & Informação*, Londrina, v. 1, n. 2, p.37-45, jul./dez. 1996. Disponível em: <<http://www.uel.br/revistas/uel/index.php/informacao/article/viewPDFInterstitial/1603/1358>> Acesso em: 09 ago. 2009.

ODDONE, N. E.; GOMES, M. Y. F. S. F. Os temas de pesquisa em Ciência da Informação e suas implicações político-epistemológicas. In CIFORM - ENCONTRO NACIONAL DE CIÊNCIA DA INFORMAÇÃO, 5, 2004, Salvador. *Anais...* Salvador, 2004. Disponível em: <http://dici.ibict.br/archive/00000558/01/temas_de_pesquisa.pdf> Acesso em: 08 ago, 2009

PINHEIRO, L. V. R. *Ciência da informação: desdobramentos disciplinares, interdisciplinaridade e transdisciplinaridade*. Rio de Janeiro: IBICT/MCT, 2002. Disponível em: <<http://www.uff.br/ppgci/editais/lenavanialeituras.pdf>>. Acesso em: 08 ago. 2009.

UNIVERSIDADE DE SÃO PAULO. *Vocabulário Controlado do SIBi/USP*. Disponível em: <<http://143.107.73.99/Vocab/Sibix652.dll/ARV?HIER=CH751>> Acesso em: 09 ago. 2009

VALE, E. A. Linguagens de Indexação. In: SMIT, J. W. (Coord.). *A análise documentária: a análise da síntese*. Brasília: IBICT, 1987. p.11 - 26.

Title

Reviewing the practices of information organization in InfoBCI: first studies on controlled vocabulary

Abstract

Viewing the Web capacity for disseminate and share information, was created a virtual environment dedicated to the community interested in Information Science. In this space, called InfoBCI (available at: <<http://www.infobci.wordpress.com>>), articles published in national journals are indexed using a controlled vocabulary suggested for organization library and information science holdings and open a space for comments about the publications, however, we noticed that the terms summarized not being effective in representation of the contents. In this context, part of a study for the determination of a controlled vocabulary appropriate organization and sharing of scientific information, explored as issue the use and importance of this instrument of representation, introducing some relevant issues about the establishment or adoption of ready-made indices in the organization of information, even comparing it to the list of issues adopted in the blog with others focused on Information Science.

Keywords

Organization of information. InfoBCI Blog. Controlled vocabulary.

Apêndice A - Listas dos grandes grupos temáticos em Ciência da Informação

LISTA 1

Fundamentos da Ciência da Informação
Organização e processamento da informação
Gestão da Informação
Tecnologias da Informação
Transferência de Informação
Aplicações de informação

PINHEIRO, L. V. R. *Ciência da informação: desdobramentos disciplinares, interdisciplinaridade e transdisciplinaridade*. Rio de Janeiro: IBICT/MCT, 2002. Disponível em:
<<http://www.uff.br/ppqci/editais/lenavanialeiturass.pdf>>. Acesso em: 08 ago. 2009.

LISTA 2

Pesquisa em ciência da informação
Organização do conhecimento
Profissão da informação
Questões sociais
Indústria da informação
Produção editorial e distribuição
Tecnologias da informação
Sistemas e serviços de informação eletrônica
Fontes e aplicações para assuntos específicos
Bibliotecas e serviços bibliotecários
Informação e questões governamentais e legais

ODDONE, N. E.; GOMES, M. Y. F. S. F. Os temas de pesquisa em Ciência da Informação e suas implicações político-epistemológicas. In *CINFORM - ENCONTRO NACIONAL DE CIÊNCIA DA INFORMAÇÃO*, 5, 2004, Salvador Proceedings ... Salvador, 2004. Disponível em:
<http://dici.ibict.br/archive/00000558/01/temas_de_pesquisa.pdf> Acesso em: 08 ago, 2009

LISTA 3

ARQUIVÍSTICA
BIBLIOTECONOMIA
DOCUMENTAÇÃO
DOCUMENTOS
INFORMAÇÃO
BIBLIOGRAFIA
USUÁRIOS DA INFORMAÇÃO

UNIVERSIDADE DE SÃO PAULO. *Vocabulário Controlado do SIBi/USP*. Disponível em:
<<http://143.107.73.99/Vocab/Sibix652.dii/ARV?HIER=CH751>> Acesso em: 09 ago. 2009