

# INDEXAÇÃO SOCIAL: ABORDAGEM CONCEITUAL

**Roger de Miranda Guedes**  
**Eduardo José Wense Dias**

**Resumo:** Diante da discussão sobre nomenclaturas dada à indexação orientada pelo usuário em ambientes virtuais o presente trabalho tem por objetivo situar o conceito “indexação social” no campo do tratamento da informação, baseando-se nas diferentes tipologias de linguagens e orientações de indexação. A partir da literatura especializada e de observações no ambiente da *World Wide Web* pretende-se contribuir para a discussão e o entendimento desta modalidade de indexação.

**Palavras-Chave:** Representação da informação; Linguagens de indexação; Indexação social; *World Wide Web*.

## 1 INTRODUÇÃO

O advento de novos suportes e contextos de criação, disponibilização, acesso e preservação de informação gerou, e continua gerando, uma série de transformações nas ciências que tem a informação como objeto de estudo. Observa-se na literatura em ciência da informação certa preocupação que diz respeito às mudanças na terminologia utilizada neste campo de estudo. Lancaster (2004), em seu texto de apresentação da segunda edição publicada no Brasil do livro *Indexação e resumos: teoria e prática*, chama a atenção para os princípios (oriundos, sobretudo da biblioteconomia) que estão sendo “reinventados”. Abordagens e conceitos estudados a mais de meio século por pesquisadores de biblioteconomia e ciência da informação parecem estar sendo negligenciados por outras áreas e, muitas vezes, rebatizados.

A terminologia científica surge da necessidade de denominar sistemas de conceitos para contextos específicos, com o objetivo de permitir uma comunicação eficiente entre especialistas. Na prática, significa estabelecer a linguagem de uma determinada área (KOBASHI; SMIT; TÁLAMO, 2001).

Sabe-se que é recorrente, e muitas vezes necessário, a criação ou adaptação de termos científicos para atender as necessidades e evolução científica, o que preocupa é a falta de um “alicerce sólido sobre o qual construir” (LANCASTER, 2004, p. X). Segundo o autor o problema é causado pela negligência dos que trabalham com os temas relacionados à biblioteconomia e ciência da informação, pesquisadores que não tomam conhecimento das contribuições que já haviam sido feitas, que parecem desconhecer o que é uma revisão de literatura.

Frequentemente no meio científico ocorre que objetos e processos de mesma natureza geralmente podem receber denominações distintas devido a uma série de fatores, como o contexto ao qual estão inseridos, a época em que foram conceituados e a função daquele objeto ou processo para determinada ciência. Na área de biblioteconomia e ciência da informação isto se revela algo comum, por exemplo, muitas vezes quando falamos em “metadados” ou “pontos de acesso de um catálogo” estamos nos referindo a mesmos objetos em contextos diferentes. Mas Alvarenga (2006, p. 97) nos alerta, “mudam-se os meios, sofisticam-se os instrumentos e surgem nomes novos para designar coisas novas e velhas. Entretanto, a essência das coisas permanece”.

É notável a quantidade exagerada de novos termos para descrever processos e ações, que são há muito tempo estudados pelos bibliotecários e cientistas da informação, sem sequer fazer algum tipo de referência ou contextualização. Um bom exemplo é o termo “etiquetagem” (*tagging*), usado para denominar a indexação orientada pelo usuário na *Word Wide Web* (*www*). Neste contexto ainda existem outras designações, umas mais fundamentadas e outras

pouco expressivas, a saber: etiquetagem colaborativa, classificação social, indexação social e folksonomia (VOSS, 2007, p. 2).

Dentre os conceitos apresentados adotou-se o termo “indexação social” para se referir, neste trabalho, a ação de etiquetagem desempenhada por usuários de ferramentas sociais em ambientes *Web*.

Nesse sentido, o trabalho visa o enquadramento conceitual do termo “indexação social” e aponta onde o mesmo se insere nos estudos sobre tratamento da informação na grande área da biblioteconomia e ciência da informação.

## **2 TRATAMENTO DA INFORMAÇÃO**

No contexto da ciência da informação a representação do conhecimento é considerada um processo organizacional aliado a processos intelectuais aplicados a *objeto-representando-conhecimentos* (ALVARENGA, 2003), ou seja, processos capazes, através de uma linguagem de símbolos, de descreverem documentos. Em outras palavras, uma representação de conhecimento é um meio de expressão, uma linguagem na qual pode-se dizer coisas sobre o mundo.

A representação pode ser classificada em dois níveis: representação primária e representação secundária. A representação primária feita pelos autores, sendo o produto de seus pensamentos a partir da observação de fenômenos naturais e sociais. A representação secundária ocorre quando os objetos (documentos) produzidos pelo homem se tornam parte de um conjunto documental, como o acervo de uma biblioteca, sendo mais uma vez representados para que possam ser manipulados facilmente.

O tratamento e a organização da informação são os processos motores, presentes em um ambiente informacional, que têm como

finalidade principal substituir as representações primárias por simulacros dos documentos – representações secundárias.

Ao aprofundar nos estudos referentes ao processo de representação documentária encontram-se ações comuns em todas as formas de representação, a indexação é uma delas. Ainda que receba diferentes nomenclaturas (indexação; catalogação de assunto; *tagging*; etiquetagem; etc.) o ato de representar um documento através de um conceito é de mesma natureza nos diferentes contextos – físico e digital.

Analisando o tratamento da informação em contexto digital e, especificando a *Web* como o ambiente a ser avaliado, percebe-se que algo já está sendo feito quanto a preocupação com o tratamento da informação disponível na Rede. Devido ao seu caráter dinâmico parece inviável tratar a informação na *Web* com os modelos de tratamento da informação tradicionais. Isso por uma série de fatores, como, por exemplo, o aumento do volume de dados que cresce exponencialmente chegando a ser impossível precisar a quantidade de informação veiculada neste ambiente. A magnitude de informações na *Web* tomou proporções tão gigantescas que tanto o gerenciamento como o tratamento de informação na Internet só pode ser feito com o auxílio de ferramentas e *softwares* especializados.

### 3 INDEXAÇÃO

De modo geral a indexação é definida como um conjunto de procedimentos com objetivo de expressar/representar o conteúdo temático de documentos através de linguagens de indexação ou documentárias visando a recuperação posterior (LANCASTER, 2004; LANGRIDGE, 2006; ROBREDO; CUNHA, 1986). O produto da representação temática é denominado termos de indexação. Segundo Lancaster (2004, p. 6) os termos atribuídos ao documento,

no processo de indexação, servem como ponto de acesso mediante os quais um item é localizado e recuperado. O autor apresenta as duas etapas principais da indexação de assuntos: a análise conceitual e a tradução.

Em relação ao nível de abrangência da indexação Robredo e Cunha (1994) classificam a descrição temática em: categorização, que seria o reconhecimento dos aspectos dominantes, segundo a divisão por assuntos preexistentes; indexação superficial, que permite obter os conceitos principais tratados no documento; indexação profunda, que consiste em obter todos os conceitos considerados fundamentais. Ainda, segundo os autores, em relação às partes analisadas dos documentos, a indexação pode ser feita com base: no título; no resumo; no título e no resumo; em partes determinadas (sumários, introdução, conclusão); no documento completo.

Outra distinção que se faz pertinente nesse assunto são as ações de indexação humana e indexação automática. A primeira é realizada por pessoas, seja por profissionais especialistas seja pelo usuário do documento, a segunda é realizada por máquinas programadas a reconhecer características em um documento que poderão ser úteis em sua recuperação. A indexação humana e a automática oferecem resultados igualmente válidos e que se complementam. Entretanto, a automática possui vantagens frente à humana devido a redução de custo e tempo de processamento.

### **3.1 Linguagens de Indexação**

As linguagens de indexação são instrumentos convencionais de uso das unidades de informação para a descrição dos conteúdos temáticos dos documentos (GUINCHAT; MENO, 1994), podendo

ser classificadas como linguagens controladas; linguagens naturais e linguagens livres (ROWLEY, 2002, p. 169).

- Linguagens controladas de indexação – definidas como um conjunto de termos autorizados para uso na indexação do assunto de documentos. É subdividida em dois tipos: as linguagens alfabéticas de indexação, como os tesouros e listas de cabeçalhos de assunto; e os sistemas de classificação, representado por código ou notação.
- Linguagens naturais de indexação – referem-se a quaisquer expressões que ocorram em alguma parte do documento. Todos os termos no corpo do documento são candidatos a serem termos de indexação.
- Linguagens livres de indexação – para esta linguagem não existem limitações quanto aos termos a serem utilizados no processo de indexação.

A primeira linguagem é considerada a mais sistemática e eficiente proporcionando mais qualidade e facilidade de utilização, é usada em sistemas de recuperação da informação onde o processamento técnico ocorre manualmente e, portanto, com um custo elevado. A segunda linguagem é a mais utilizada em sistemas de recuperação da informação automatizados onde os documentos se apresentam em formato digital, pode ser caracterizado pelo seu baixo custo e exaustividade presente nas buscas. A terceira linguagem é usada por um indexador humano e sua qualidade irá depender do conhecimento do indexador sobre o(s) assunto(s) tratado no documento.

### **3.2 Orientação da Indexação**

Para Hjørland (2001) a indexação deve ser moldada para se ajustar às necessidades de determinada clientela, indo além da escolha de linguagem de indexação adequada. A indexação também poder ser analisada tomando como perspectiva o agente executor do processo. Segundo Rafferty e Hilderley (2007, p. 399) existem três grupos de candidatos a atores no processo de indexação, são eles:

- Indexação orientada por especialistas – baseia-se no tratamento da informação através da intervenção de intermediários (bibliotecários, indexadores, editores voluntários), é a indexação feita por especialistas sendo dispendiosa e cara.
- Indexação orientada pelo autor – esta abordagem pressupõe que o autor irá utilizar termos que são comumente compreendidos e geralmente aceitos. Um problema que essa abordagem enfrenta é o fato do autor não ser necessariamente um gestor de informação com os conhecimentos profissionais de um especialista.
- Indexação orientada pelo usuário – esse tipo de indexação possibilita um elevado nível de interação com a comunidade que, provavelmente, não seria possível se tivesse decisões a serem feitas sobre códigos, convenções e regras que regem qualquer taxonomia controlada.

## **4 ORGANIZAÇÃO DA INFORMAÇÃO NA WEB**

Desde o seu surgimento, a *Web* passou por diversas transformações, não só no âmbito tecnológico, mas também por mudanças histórico-social centradas no usuário. De uma rede explicitamente corporativa e comercial a *Web* passa a ser conhecida pelo seu poder de interação. Esta abordagem frente a *www* teve sua

ascensão nos últimos anos e foi nomeada *Web 2.0* “surgindo um nível de interação onde as pessoas poderiam colaborar para a qualidade do conteúdo disponível, produzindo, classificando e reformulando o que já está disponível” (BLATTMANN; SILVA, 2007, p. 197).

Com isso surgem diversos questionamentos a serem tratados por profissionais e pesquisadores que lidam com informação, como as alternativas para tratamento e organização da informação em contextos digitais.

Os primeiros passos no sentido de minimizar o problema da falta de organização da informação na *Web* foram dados no desenvolvimento de estudos sobre bibliotecas digitais – sendo o conceito mais utilizado hoje para definir o conjunto de artefatos, conhecimentos, práticas referente a uma comunidade presente no contexto digital (DIAS, 2001). Tão importante quanto o tratamento da informação no contexto digital é o processo de seleção, deve-se ter em mente o que é relevante e merece ser processado dentro do desordenado espaço que a *www* se tornou.

As bibliotecas digitais presentes na *Web* são caracterizadas por reunir uma coleção de itens relacionados a um mesmo tipo de material ou mesmo assunto, e até obras de/sobre um determinado autor. São repositórios onde a informação recebe tratamento adequado podendo ser recuperada com mais facilidade. Este modelo de organização da informação revela um sistema caracterizado pelo movimento *top-down*, ou seja, uma organização que segue um fluxo do centro para a periferia.

Uma recente alternativa adotada no ambiente *Web* para auxiliar a organização da informação são as ferramentas de gerenciamento de informação voltadas ao usuário. São funcionalidades disponíveis em *softwares* sociais – espaços virtuais propícios à interação mediada por computador e ao compartilhamento de conteúdo – que deixa nas mãos do usuário todo o processo de gerenciamento da informação. Moura (2009) classifica estas ferramentas pertencentes a um grupo

de instrumentos de organização da informação mais abrangente, denominado ferramentas ontológicas. De acordo com a autora, as ferramentas ontológicas seriam

[...] estruturas informacionais contextualizadas, derivadas de esquemas intelectuais mais complexos e desenvolvidos sob um ponto de vista e com um propósito específico. Tais ferramentas têm por objetivo orientar os sujeitos no entendimento acerca do conhecimento em áreas específicas bem como na adoção consciente desses esquemas representacionais em sistemas de organização e recuperação da informação. (MOURA, 2009, p. 62)

Entre as principais ações dessas ferramentas estão a atribuição de etiquetas (*tagging*) pelos usuários aos itens compartilhados e, como consequência, a geração de índice de descritores que possibilite posteriormente a recuperação dos itens etiquetados.

Uma etiqueta (*tag*) criada por um usuário pode ser utilizada por todos os outros, assim como um item classificado por um usuário também é livre para ser classificado por quem quiser, criando um ambiente favorável ao desenvolvimento do conhecimento coletivo. Esse conhecimento emerge de um sistema governado por forças *bottom-up*, isto é, da periferia para o centro.

## 5 INDEXAÇÃO SOCIAL

Em tempos onde o principal caráter da informação em contexto virtual é a ubiquidade percebe-se a mudança de posicionamento dos atores que integram o contexto informacional. Um mesmo documento é acessado, manipulado, organizado e recuperado por uma infinidade de usuários em diversas partes do mundo e ao mesmo tempo. A emergência do ciberespaço aliada à potencialidade de

ferramentas sociais na *Web* propiciam um ambiente ideal para que haja o movimento de descentralização da organização da informação.

No centro deste arranjo estão as folksonomias, que podem ser definidas como sistemas orgânicos baseados na atribuição livre e pessoal de etiquetas à informações ou objetos visando à organização e recuperação (WAL, 2006). O neologismo folksonomia – formado pelas palavras, em inglês, *folks* (pessoas) e *taxonomy* (taxonomia) – foi cunhado por Thomas Vander Wal como forma de expressar contraposição às classificações do conhecimento tradicionais, elaboradas por especialistas e construídas baseando-se em arranjos hierárquicos (MATHES, 2004).

A ação mais explícita na folksonomia enquanto processo é a indexação, mais conhecida nos ambientes digitais por *tagging*, ou seja, a ação de etiquetar. Antes mesmo da popularização desta modalidade de representação da informação na *Web* autores que trabalham com formas de indexação alternativas já estudavam o assunto.

Lancaster (2004) já dizia que para certos tipos de materiais, a indexação orientada pelo usuário pode até ser mais importante do que o é no caso de artigos de periódicos, livros, ou relatórios técnicos. Em seu trabalho, Brown et al. (1996) apud Lancaster (2004), chamam a atenção para a necessidade de um tratamento democrático da indexação, em que os usuários acrescentariam aos registros termos de sua própria escolha, quando isso fosse necessário e apropriado. E, posteriormente Rafferty e Hilderley (1997), que já trabalhavam com um método de tratamento democrático da indexação.

A indexação social é definida por Hassan-Montero (2006, *online*) como

[...] um novo modelo de indexação, em que são os próprios usuários ou consumidores dos recursos os que levam a cabo sua descrição [...] A descrição de cada

recurso se obteria por agregação, ou seja, um mesmo recurso seria indexado por inúmeros usuários, dando como resultado uma descrição intersubjetiva e portanto mais fiel que a realizada pelo autor do recurso (Tradução do autor).

O autor ressalta que seria válido se referir a essa modalidade somente em sistemas que permitem uma indexação agregada, ou seja, nos quais vários usuários indexam um mesmo recurso. A atribuição de etiquetas, quando feita apenas pelo criador do recurso, não pode ser considerada um modelo inovador.

Ressalta-se que a indexação social não deve ser vista como um contraponto às outras modalidades de indexação – indexação realizada por profissionais, indexação automática, etc., ela é apenas mais uma alternativa de organização, surgida da necessidade de tratar conteúdos no ambiente virtual.

## **6 ALGUMAS CONSIDERAÇÕES**

O objetivo deste ensaio centrou-se na contextualização do conceito “indexação social” dentro do campo do tratamento da informação, a partir de suas características essenciais enquanto processo de representação, procurando evidenciar a possibilidade de sedimentação do termo dentro da área da Ciência da Informação.

A partir das classificações apresentadas nas sessões anteriores sobre os tipos de indexação buscou-se situar a indexação social dentro daquelas estruturas.

Quanto ao tipo de linguagem a indexação social é considerada uma linguagem livre – por não haver nenhum tipo de controle, o usuário atribui etiquetas ao item sem se preocupar com regras e normas, mas tendo em mente a futura recuperação daquele item por ele ou outro usuário. Em relação a sua orientação é baseada no

usuário – ainda que também possa ser considerada orientada pelo autor em *softwares* de compartilhamento onde o usuário é o criador do item.

Essas características são os traços mais evidentes dessa forma de indexação contribuindo na compreensão do conceito adotado para denominar este modelo descentralizado de representação da informação em ambientes virtuais. Assim, a expressão “indexação social”, se justifica não apenas pelo fato da ação ser concretizada por indivíduos, mas também por ser um ato colaborativo e democrático, onde o papel de todos os indivíduos tem o mesmo valor e peso dentro do sistema.

## REFERÊNCIAS

ALVARENGA, Lídia. Representação do conhecimento na perspectiva da Ciência da Informação em tempos e espaços digitais. *Enc. Bibli.*, n. 15, v. 8, 2003. Disponível em: <<http://www.periodicos.ufsc.br/index.php/eb/article/view/97>>. Acesso em: 15 out. 2008.

ALVARENGA, Lídia. Organização da informação nas bibliotecas digitais. In: NAVES, Madalena M. L.; KURAMOTO, Hélio (Org.). *Organização da informação: princípios e tendências*. Brasília: Briquet de Lemos/Livros, 2006. p. 79-98

BLATTMANN, Ursula; SILVA, Fabiano Couto Corrêa da. Colaboração e interação na Web 2.0 e biblioteca 2.0. *Revista ACB*, Florianópolis, v. 12, n. 2, p. 191-215, jul./dez. 2007. Disponível em: <<http://revista.acb.org.br/index.php/racb/article/view/530/664>>. Acesso em: 18 abr. 2009.

BROWN, P. et al. The democratic indexing of images. *New Review of Hypermedia and Multimedia*, 2, 1996, p. 107-120 apud LANCASTER, F. W. *Indexação e resumos: teoria e prática*. Brasília: Briquet de Lemos, 2004. 452 p.

DIAS, Eduardo José Wense. Contexto digital e tratamento da informação. *Data Grama Zero*, Rio de Janeiro, v. 2, n. 5, 2001. Disponível em: <[http://dgz.org.br/out01/F\\_I\\_art.htm](http://dgz.org.br/out01/F_I_art.htm)>. Acesso em: 17 maio 2009.

GUINCHAT, C.; MENO, M. *Introdução geral às ciências técnicas da informação e documentação*. Brasília: IBICT, 1994. 540 p.

HASSAN-MONTERO, Yusef. Indización Social y Recuperación de Información. *No Solo Usabilidad Journal*, Granada, n. 5 nov. 2006. Disponível em: <[http://www.nosolousabilidad.com/articulos/indizacion\\_social.htm](http://www.nosolousabilidad.com/articulos/indizacion_social.htm)>. Acesso em: 10 nov. 2008.

HJORLAND, B. Toward a theory of aboutness, subject, topicality, theme, domain, field, content... and relevance. *Journal of the American Society for Information Science and Technology*, 52, p. 249-298, 2001.

KOBASHI, N. Y.; SMIT, J. W.; TÁLAMO, M. F. G. M. A função da terminologia na construção do objeto da Ciência da Informação. *Data Grama Zero*, Rio de Janeiro, v. 2, n. 2, abr. 2001. Disponível em: <[http://www.dgzero.org/abr01/Art\\_03.htm](http://www.dgzero.org/abr01/Art_03.htm)>. Acesso em: 20 jun. 2009.

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

LANGRIDGE, Derek. *Classificação: abordagem para estudantes de biblioteconomia*. Rio de Janeiro: Interciência, 2006. 126 p.

MATHES, Adam. Folksonomies - cooperative classification and communication through shared metadata. *Computer Mediated Communication – LIS590CMC*, Urbana : University of Illinois, 2004. Disponível em: <<http://www.adammathes.com/academic/computer-mediated-communication/folksonomies.html>>. Acesso em: 15 Apr. 2009.

MOURA, Maria Aparecida. Informação, ferramentas ontológicas e redes sociais *ad hoc*: a interoperabilidade na construção de tesouros e ontologias. *Informação e Sociedade*, João Pessoa, 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: 25 jun. 2009.

RAFFERTY, Pauline; HIDDENLEY, Rob. Flickr and democratic indexing: dialogic approaches to indexing. *Aslib Proceedings*, v. 59, Issue 4/5, 2007. p. 397-410. Disponível em: <<http://www.emeraldinsight.com/Insight/viewPDF.jsp?Filename=html/Output/Published/EmeraldFullTextArticle/Pdf/2760590407.pdf>>. Acesso em: 1º maio 2008.

ROBREDO, Jaime; CUNHA, Murilo Bastos da. *Documentação de hoje e de amanhã: uma abordagem informatizada da biblioteconomia e dos sistemas de informação*. São Paulo: Global, 1994. 400 p.

ROWLEY, Jennifer. *A biblioteca eletrônica*. Brasília: Briquet de Lemos, 2002. 399 p.

VOSS, Jakob. *Tagging, folksonomy & Co – renaissance of manual indexing?* 2007. Disponível em:  
<[http://arxiv.org/PS\\_cache/cs/pdf/0701/0701072v2.pdf](http://arxiv.org/PS_cache/cs/pdf/0701/0701072v2.pdf) > Acesso em:  
21 Apr. 2009.

WAL, Thomas Vander. *Folksonomy definition and wikipedia.* 2005.  
Disponível em:  
<<http://www.vanderwal.net/random/entrysel.php?blog=1750>>.  
Acesso em: 22 may. 2009.

---

## **SOCIAL INDEXING: CONCEPTUAL APPROACH**

**Abstract:** *Given the discussion of nomenclature given to user-orientated indexing in virtual context the present study aims to situate the concept Indexing Social on field of information processing, based on the different types of indexing languages and guidance. From a literature and careful observations in the World Wide Web it is intended to contribute to the discussion and understanding of this model of subject indexing.*

**Keywords:** *Representation of information; Indexing languages; Social indexing; World Wide Web.*

---

### **Roger de Miranda Guedes**

Mestrando em Ciência da Informação (PPGCI/UFMG), Universidade Federal de Minas Gerais

Contato: rogerotoni@gmail.com

### **Eduardo José Wense Dias**

Doutor em Ciência da Informação. Professor aposentado da Escola de Ciência da Informação ECI/UFMG

Contato: edias@eci.ufmg.br

Artigo: Recebido em: 03/08/2009 Aceito em: 08/09/2009
---