

As folksonomias entre os conceitos e os pontos de acesso: as funções de descritores, citações e marcadores nos sistemas de recuperação da informação

Leticia Strehl

**Bibliotecária da Biblioteca Central da
Universidade Federal do Rio Grande dos Sul.
Doutoranda do Programa de Pós-graduação em
Comunicação e Informação da UFRGS**

O artigo discute as potencialidades das folksonomias, a partir da proposição e análise de esquemas de representação de conceitos, que constituem os diferentes recursos de indexação de assuntos. Aborda o desenvolvimento das fontes de informação especializadas e revisa os fundamentos técnicos das linguagens natural e documentária. Trata dos sistemas de recuperação, que exploram as redes de colaboração inerentes à produção e ao uso da informação, ilustrando o funcionamento desse tipo de recurso, ao tratar dos índices de citação e dos favoritos socializados. Por fim, menciona a distinção necessária entre o papel das folksonomias, como instrumentos de representação conceitual e como pontos de acesso.

Palavras-chave: Folksonomias; Linguagens de indexação; Citações.

The folksonomies between the concepts and the access points: the roles of descriptors, citations and tags in the information retrieval systems

The paper discusses the potentialities of folksonomies through the proposition and analysis of concept representation schemas that constitute the different resources of subject indexing. The development of specialized information sources is approached, and, the technical basis for natural and documental languages are

reviewed. It deals with retrieval systems that make use of collaboration networks inherent to the information production and use; this is done to illustrate the functioning of this kind of resources when citation indexes and social bookmarking tools are analyzed. Finally, it makes a necessary distinction between the role of the folksonomies as instruments of conceptual representation and access points.

Keywords: Folksonomies; Indexing languages; Citations.

Recebido em 05.09.2009 Aceito em 01.06.2011

1 Introdução

O advento da Internet ou, mais especificamente, da Web 2.0, revolucionou a forma de produzir e disseminar informações. Antes, os recursos para veiculação ampla de informações reduziam-se, basicamente, aos meios de comunicação de massa. Estes, por sua vez, estavam disponíveis a um limitado número de pessoas, que tinha acesso a uma estrutura técnica e complexa de produção.

Hoje, qualquer indivíduo, com um treino mínimo, pode tornar público o que pensa e produz. O desenvolvimento das tecnologias de informação e o aperfeiçoamento das interfaces permitem que uma pessoa escreva, divulgue sua música, fotografia ou vídeo em um blog, no YouTube, no Orkut, na Wikipédia, entre tantas outras centenas de recursos, gratuitamente, sem necessariamente dominar HTML ou disputar espaço em um servidor de Internet.

Além da veiculação de informações, os usuários da web 2.0 podem localizar pessoas com interesses comuns, utilizando um computador conectado na Internet, para frequentar um espaço público. Eles podem entrar em uma sala de bate-papo, formada em torno de um tema específico, participar de uma comunidade que lhes interesse em um sítio de relacionamento, escrever um comentário em um blog, ampliando, sobremaneira, os espaços sociais de convivência.

O surgimento dessas novas formas de sociabilidade se estende, também, para os ambientes acadêmicos, onde, na realidade, tudo começou, desde os idos dos primeiros usos dados à Arpanet. A Internet revolucionou profundamente os modos de comunicação da ciência, dando novas formas aos tradicionais veículos de informação científica e criando diferentes espaços para a comunicação informal entre pesquisadores.

Nesse contexto, o presente artigo tem por objetivo discutir as recentes possibilidades de representação e recuperação da informação com a web 2.0, tratando, especificamente, do uso das folksonomias nas ferramentas de favoritos socializados. Para realização das análises, propomos esquemas que ilustram a forma de representação dos conceitos inerentes aos tradicionais recursos de indexação por linguagens e citações. Explorando as questões de representação conceitual, pretendemos chegar a uma compreensão das potencialidades das

ferramentas baseadas em tecnologia de web 2.0 para os sistemas de recuperação da informação, como desenvolvemos nas seções a seguir.

2 O uso da linguagem para constituição dos pontos de acesso em sistemas de recuperação da informação

As bases de dados especializadas e seus antecessores, os índices e resumos, vêm desempenhando, historicamente, um papel de extrema importância no processo de comunicação da ciência. A partir da representação detalhada das partes e dos conteúdos das publicações, estes instrumentos permitem que o pesquisador, em meio a uma imensa gama de informações, identifique trabalhos sobre determinado assunto de seu interesse.

Estas fontes surgiram logo após a invenção da imprensa e eram inicialmente produzidas com recursos técnicos bastante rudimentares. Somente com o avanço da automação, nos anos 1960 e 1970, as grandes empresas comerciais começaram a publicar índices e resumos mais robustos, capazes de manipular uma quantidade maior de dados e de permitir que os itens fossem indexados por um número maior de pontos de acesso.

O ponto de acesso pode ser definido como o meio pelo qual um item bibliográfico é recuperado, no momento da realização de uma busca (LANCASTER, 1993). Na realidade, esse conceito é indispensável para que avaliemos qualquer sistema de recuperação de informação, pois dependemos da qualidade dos pontos de acesso para maximizar, tanto a identificação dos itens úteis (revocação), quanto a omissão dos itens inúteis (precisão).

O ponto de acesso que cumpre essa função é representado por um termo que identifica inequivocamente um conceito e, no contexto de um sistema específico, deve ser expresso vocabularmente de modo coerente. Como forma de alcançar este ideal, no decorrer da história dos sistemas de recuperação da informação, foram desenvolvidas as técnicas de exame de documentos, de identificação conceitual e de tradução de termos para linguagens de indexação.

O exame dos documentos tem como objetivo verificar os assuntos tratados em seu conteúdo. Esta etapa possibilita a realização da tarefa seguinte: a identificação dos conceitos inerentes às temáticas tratadas. A análise conceitual de um documento é um procedimento complexo que envolve, segundo a Associação Brasileira de Normas Técnicas (ABNT) (1992, p. 2), a formulação e a resposta a questões do tipo:

- a) Qual o assunto de que trata os documentos?
- b) Como se define o assunto em termos de teorias, hipóteses, etc.?
- c) O documento contém uma ação, uma operação, um processo?

- d) O documento trata do agente dessa ação, operação, processo, etc.?
- e) O documento se refere a métodos, técnicas e instrumentos especiais?
- f) Esses aspectos foram considerados no contexto de um local ou ambiente especial?
- g) Foram identificadas variáveis dependentes ou independentes? e
- h) o assunto foi considerado sob um ponto de vista interdisciplinar?

Depois dessa etapa, os conceitos identificados devem ser traduzidos para uma linguagem de indexação. Esse procedimento pode ser realizado de duas formas (LANCASTER, 1993). Primeiro, por extração, quando as palavras ou expressões escritas pelos próprios autores são utilizadas para representar o conteúdo temático dos documentos. Dizemos, neste caso, que foi adotada a linguagem natural para a indexação.

Como mostra a FIG. 1, com o uso da linguagem natural, um único conceito pode ser expresso por diferentes descritores em um sistema de recuperação, dependendo da forma como os textos são originalmente escritos por seus autores. Partindo dos elementos expostos no esquema a seguir, por exemplo, se um usuário estiver interessado em artigos que tratem dos conceitos A e B e utilizar, em sua expressão de busca, apenas os descritores 1A e 1B, deixará de recuperar o Artigo 2, indexado apenas com os descritores 2A e 2B, não contemplados na expressão de busca.

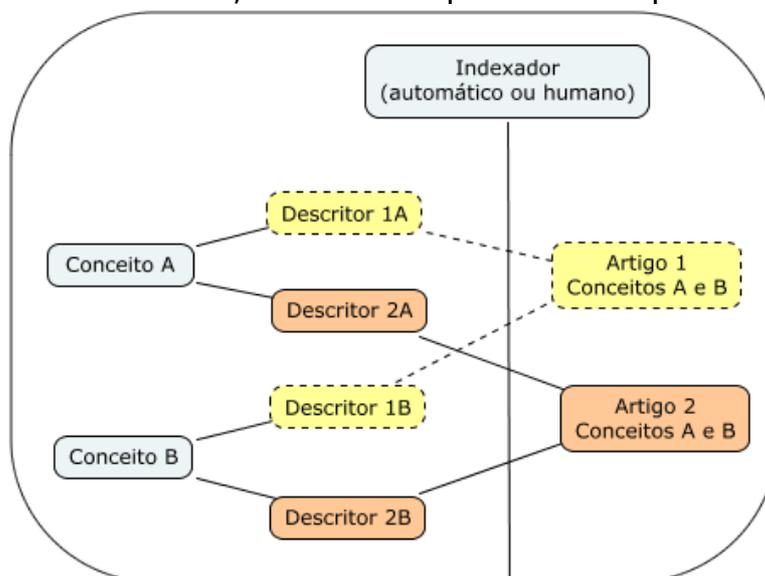


FIGURA 1 – Esquema da representação de conceitos com uso de linguagem natural

Fonte: Dados da pesquisa.

Em uma tentativa de resolução desse problema, uma outra forma de tradução de conceitos foi desenvolvida, a indexação por atribuição. Por atribuição, os termos que representam os assuntos tratados são oriundos de uma fonte que não é o próprio documento, mas, sim, de uma

linguagem documentária, frequentemente, de algum tipo de vocabulário controlado ou tesauro.

De acordo com Cintra e outros (2002, p. 42):

As linguagens documentárias mais consistentes para representação documentária dispõem de um vocabulário que integra elementos, de um lado, da linguagem de especialidade e das terminologias e, de outro, da linguagem natural que é a linguagem dos usuários. Essas unidades, acompanhadas ou não de uma noção, constituem o 'léxico' das linguagens documentárias [...].

A FIG. 2 exemplifica que, em um sistema que utiliza recursos para controle de vocabulário, um conceito apresenta uma única forma de representação. Neste caso, os dois artigos, por tratarem de um conceito em comum, são representados igualmente pelo Descritor A.

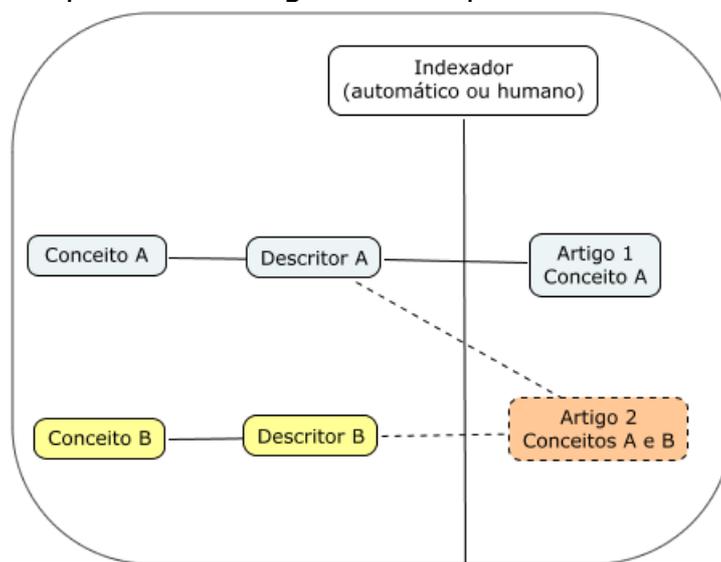


FIGURA 2 – Esquema da representação de conceitos com uso de linguagem documentária

Fonte: Dados da pesquisa.

Esse tipo de recurso só é passível de ser implementado quando as variações de estilo e até mesmo as imprecisões naturais da linguagem humana são trabalhadas no âmbito de uma linguagem documentária por técnicas que, em sua forma mais sofisticada, resultam em instrumentos imbuídos de uma rede lógico-semântica, que se destinam a: controlar sinônimos, diferenciar homógrafos e agrupar termos cujos significados apresentem relações mais estreitas entre si (LANCASTER, 1993).

Por sua vez, as linguagens documentárias também não são, por si só, capazes de resolver todos os problemas de recuperação dos sistemas de informação. Além do seu alto custo de desenvolvimento, de implementação e de manutenção, percebemos que, muitas vezes, as bases de dados não disponibilizam ferramentas para que os usuários

consigam aproveitar em suas buscas, as possibilidades de recuperação oferecidas pela linguagem documentária.

Por exemplo, se um usuário utilizar em uma base de dados a expressão “implantação de íons” e o descritor adotado para este assunto for “implantação iônica”, ele acabará por não recuperar qualquer item em sua busca se o sistema não dispuser de instrumentos remissivos que o redirecionem para a forma como o sistema localizaria a informação. Assim, mesmo que o conceito esteja representado de forma coerente e padronizada, a recuperação não ocorrerá se o usuário e a interface não estiverem aptos a reconhecer a convenção.

A era digital modificou radicalmente os instrumentos informacionais e, concomitantemente, as discussões sobre pontos de acesso ganharam uma nova dimensão. Nos primórdios, a quantidade de pontos de acesso confinava-se aos limites possíveis de uma obra impressa. Hoje, cada palavra de um texto disponível em meio eletrônico e indexado por um motor de busca qualquer constitui um ponto de acesso.

Desse modo, todas as discussões sobre vocabulários controlados e formação de indexadores parecem ter ficado perdidas em um tempo que, da capacidade de representação sucinta e exata de um conceito, dependia o acesso do usuário à informação. Todavia, mesmo sendo esse o senso comum, as questões de representação da informação não estão completamente solucionadas com a indexação do texto integral dos documentos.

A realização de uma busca rápida no Google, ou mesmo em sua versão especializada, o Google Acadêmico, dá a dimensão dos desafios impostos ao usuário, quando o sistema retorna muitos resultados. Por mais que tenham evoluído, os algoritmos de avaliação automática da relevância dos documentos, o aumento incomparável no fornecimento de pontos de acesso acaba, também, por ampliar as possibilidades de ocorrência de falsas associações ou de relações incorretas entre os termos.

Nessas circunstâncias, os resultados produzidos por sistemas desse tipo tendem a carecer de precisão: são eficazes na identificação de itens úteis, mas pouco capaz de omitir itens irrelevantes para satisfação de uma necessidade específica de informação. Em um primeiro momento, este parece ser um problema menor, mas não o é. Não podemos esquecer que a omissão é quase tão fundamental quanto a identificação de itens em um sistema, dado o desafio constante imposto pela explosão documental à limitada capacidade humana para o processamento de informações.

Tendo em vista a complexidade do uso da linguagem como recurso para representação de conceitos, as citações feitas nos trabalhos científicos passaram a ser um importante mecanismo para recuperação de informações, agrupando documentos semelhantes pela identificação de referências comuns. As citações podem ser vistas como marcas da estrutura colaborativa da produção do conhecimento e são precursoras, mesmo que não sejam as inspiradoras, dos sistemas atuais de Web 2.0, que utilizam as folksonomias para favorecer a reunião de pessoas com

interesses comuns. Na próxima seção aprofundaremos os argumentos favoráveis a esta analogia.

3 Redes de colaboração, citações, URLs e Folksonomia: recursos de recuperação da informação para além da linguagem

A avaliação por pares e as citações são os principais elementos que caracterizam o que se reconhece, hoje, como o moderno artigo científico. Estes dois elementos, presentes na realidade de produção do discurso acadêmico, surgiram em meados do século XIX, e definem que um artigo poderá compor o universo das publicações científicas, quando receber um parecer que declare seu mérito e incluir referências claras e precisas aos autores que apresentaram contribuições para a temática tratada (ELLMAN, 1998).

O ato de citar explicita a natureza colaborativa da ciência, é a evidência de que a produção do conhecimento científico se dá de forma progressiva (não necessariamente cumulativa), a partir da reunião dos esforços de diferentes pesquisadores que trabalham, de forma articulada ou não, sobre um determinado tema. Em suas publicações, os cientistas referenciam os trabalhos anteriores, revelando a rede de colaborações que tornou possível a obtenção dos resultados da pesquisa.

Tendo em vista a importância dessas informações para o processo de produção do conhecimento, no início da década de 1950, como membro do Projeto *John Hopkins University Medical Indexing*, Eugene Garfield identificou nas citações bibliográficas, um recurso para representar os assuntos dos documentos, por meio de procedimentos completamente automáticos (CRONIN; ATKINS, 2000). Depois de desligar-se do Projeto, Garfield aproveitou esta experiência e iniciou uma série de outros estudos, que culminaram na fundação do *Institute for Scientific Information* (ISI, atual *Thomson-Reuters*) e, em 1963, na publicação da primeira edição do *Science Citation Index* (SCI).

Operacionalizando esta lógica, os índices de citações foram desenvolvidos a partir do princípio de que as referências citadas por um autor identificam de maneira mais precisa o relacionamento entre documentos que tratam do mesmo assunto (GARFIELD, 1955). Korfhage (1997) destaca que as pesquisas que analisam a ocorrência de co-citações, verificam que a similaridade de dois documentos pode ser medido pelo número de artigos que ambos citam.

O uso das citações, como recurso de identificação de documentos semelhantes alterou o enfoque utilizado nos sistemas de recuperação da informação, ao investir na possibilidade de que os autores são mais precisos quando identificam a contribuição anterior de seus colegas para o tema de pesquisa, do que quando empregam a linguagem para representar os conceitos de que tratam.

Do mesmo modo que trabalhamos o uso das linguagens de indexação para representação dos assuntos em sistemas de informação,

descrevemos na FIG. 3 um esquema que ilustra o uso das referências, como forma de identificação de documentos tematicamente afins.

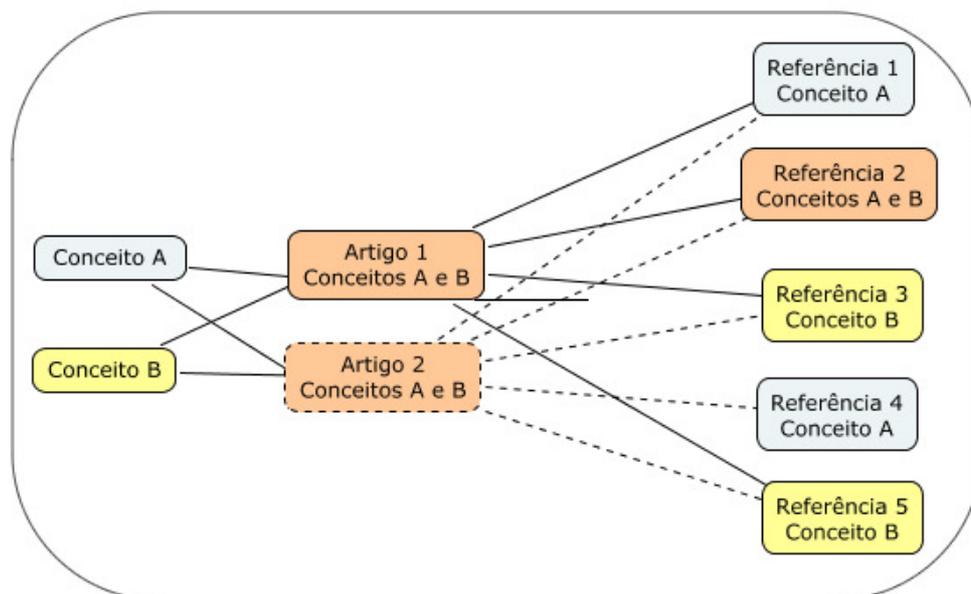


FIGURA 3 – Esquema de representação de conceitos a partir das referências citadas nos artigos

Fonte: Dados da pesquisa.

De acordo com o exemplo, observamos que artigos semelhantes referem trabalhos comuns. Desse modo, a partir de um índice de citações, se buscássemos os documentos que citaram a Referência 5, recuperaríamos os artigos 1 e 2, que tratam exatamente dos mesmos conceitos. Entretanto, como os autores possuem estilo próprio e preferências pessoais, a relação assunto-citação não é exata. Essa realidade também é descrita em nossa ilustração, quando indica que um usuário do índice que pesquisasse pela Referência 4, não chegaria aos 2 trabalhos, mas apenas ao Artigo 2.

Todavia, não sendo exata a correspondência entre assunto e referências, em que medida os índices de citações apresentam-se como um recurso diferenciado de recuperação?

Na realidade, este tipo de base de dados passou a ser um instrumento potente de recuperação, após a informatização das fontes de referência e da implementação de estruturas hipertextuais de navegação. Com esta interface, o usuário passou a ser capaz de localizar facilmente informações a partir de um único registro bibliográfico, ampliando o potencial de identificação de documentos relacionados à sua necessidade específica, ao percorrer, com o auxílio dos links, a rede de pesquisas que se entrelaça pelas citações, como mostra a FIG. 4.

organização e o compartilhamento de favoritos, os chamados favoritos socializados. Delicious¹, Connotea² e CiteULike³, exemplos de ferramentas desta natureza, possibilitam que seus usuários colem e indexem os documentos eletrônicos de seu interesse, disponíveis na Internet. Estes procedimentos acabam por formar inúmeras coleções individuais, quando vistas no nível de cada um dos usuários, e um imenso acervo, quando examinados no âmbito do sistema como um todo.

Farooq e outros (2007) definem que a unidade básica de informação em um serviço de favoritos socializados compreende três elementos: usuário, documento e marcador. Estes elementos são processados pelos sistemas e compilados naquilo que se denomina uma folksonomia (GORDON-MURNANE, 2006).

O termo folksonomia é derivado de taxonomia. Taxonomias são geralmente controladas por especialistas e são estáticas, tendendo para o uso de terminologia oficial, em vez de frases vernaculares. Ao contrário, as folksonomias são sistemas de classificação distribuídos, criados por usuários individuais (GUY; TONKIN, 2006). Assim, a folksonomia, como seu nome sugere, é uma taxonomia elaborada pelo povo.

Com esse recurso, qualquer usuário pode navegar ou pesquisar nas folksonomias em busca de documentos do seu interesse. Para Aquino (2007, p. 99), "ao adicionar tags [marcadores] a uma informação na web estamos ativando uma rede de associações, contribuindo para a formação de uma memória coletiva [...]."

De acordo com Morrison (2008), uma definição bem estrita de folksonomia deve incluir apenas os marcadores e as relações entre eles, excluindo qualquer sistema que apresente títulos e descrições de documentos, sistemas de classificação e recomendação etc. Enquanto determinada folksonomia pode assemelhar-se a uma rede de conexões sociais entre usuários, outra pode ignorá-la. A definição preferida pelo autor, e adotada no presente trabalho, inclui um espectro mais amplo de aspectos da seguinte forma:

- a) a coleção do sistema é formada com contribuições de usuários;
- b) os usuários participam na classificação ou avaliação; e
- c) a adição, classificação ou avaliação de itens se realiza por intermédio de uma rede social.

Retornando aos esquemas que utilizamos anteriormente para compreender a representação de conceitos pelas linguagens natural e documentária e a recuperação de informações, a partir das referências comuns citadas por diversos artigos científicos, ilustramos, na FIG. 5, os elementos que compõem o uso de marcadores em serviços de socialização de favoritos.

¹ Disponível em: <<http://delicious.com>>. Acesso em: 01 ago. 2008.

² Disponível em: <<http://www.connotea.org/>>. Acesso em: 01 ago. 2008.

³ Disponível em: <<http://www.citeulike.org/>>. Acesso em: 01 ago. 2008.

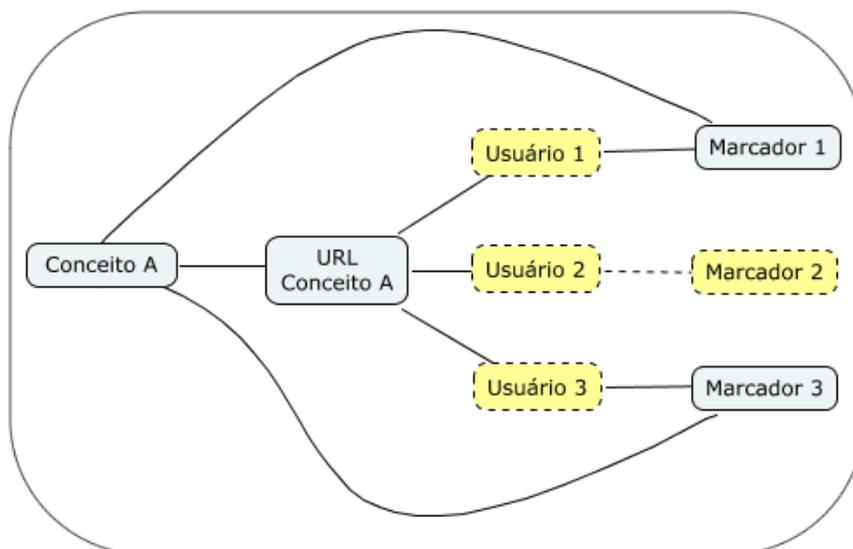


FIGURA 5 – Esquema de uso de marcadores em folksonomias

Fonte: Dados da pesquisa.

O esquema⁴ acima, apresenta várias diferenças das demais ilustrações feitas nas figuras de 1 a 3. Provavelmente, a divergência mais significativa é que em um sistema baseado em folksonomias não existe uma única diretriz que oriente a representação dos documentos, a tarefa está distribuída entre os diversos usuários do serviço, prescindindo de uma política global de indexação. Isto se deve ao fato de que nas folksonomias a atribuição de marcadores está relacionada com a identificação da significação particular dos documentos para cada um dos usuários do sistema, não sendo, necessariamente, representativas dos conceitos ali tratados. Em contraposição, nas bases de dados especializadas tradicionais, a representação pretende captar a dimensão dos conceitos abordados nos documentos, tendo em vista o público-alvo do serviço e os demais materiais cobertos pela fonte de informação.

Assim, mesmo que as folksonomias se assemelhem aos sistemas que utilizam a linguagem natural, quando autorizam a existência de formas variadas de representação para um único conceito, os recursos devem ser diferenciados por sua natureza: os marcadores resultam, não somente da diversidade de formas de expressão permitidas pela linguagem, mas, também, da imensidão de sentidos que podem ser atribuídos a um documento por diferentes indivíduos.

O exemplo da FIG. 5 pretende ilustrar esta questão. Os marcadores atribuídos pelos usuários 1 e 3 à URL⁵, que trata do conceito A, são diferentes e, apesar disso, ambos são variações de representação do conceito. O Usuário 2, por sua parte, que utilizou o Marcador 2, difere dos demais, não por uma diferença de estilo, de linguagem, mas porque o marcador é representativo de sua necessidade individual e não do conteúdo acessível pela URL.

Outra associação interessante pode ser feita entre as folksonomias e as referências citadas nas publicações científicas, considerando que ambas

⁴ Outros esquemas são propostos por Cattuto (2006) e Farooq e outros (2007).

⁵ Uniform Resource Locator (URL): endereço de um recurso disponível em uma rede (URL, 2008).

funcionam como elos entre pessoas com interesses comuns. Se nos índices de citação um pesquisador pode localizar outras publicações, que se utilizaram de um artigo reconhecidamente importante para sua investigação, nos favoritos socializados, por exemplo, o usuário pode identificar pessoas que possuem itens comuns aos seus em suas listas de preferências ou que utilizam os mesmos marcadores para indexar suas URLs (MACROBERTS; MACROBERTS, 1989).

Informações desta natureza eram completamente inacessíveis antes de sistemas deste tipo, mesmo no contexto da comunicação científica. Um pesquisador era capaz de identificar referências que foram relevantes para seus colegas de área, apenas depois de passada a última fase da produção do conhecimento, a publicação. E, mesmo neste estágio, o pesquisador ainda arriscava não ter ciência de documentos importantes que, por um motivo ou outro, seu colega utilizou, mas resolveu não citar.

O desenvolvimento de ferramentas que captam, não somente o conteúdo os documentos, mas o sentido que as pessoas lhes atribuem, ampliam, sobremaneira, as possibilidades de recuperação da informação. Deste modo, o recurso ultrapassa os limites da comunicação formal de conhecimento pelos autores, para alcançar um estágio anterior, ainda informal, de elaboração de ideias pelos leitores.

4 Considerações Finais

Ao falarmos do uso da linguagem natural e das linguagens documentárias, como formas de representação de conceitos e de todas as etapas envolvidas no processo de indexação (automática ou não), demonstramos que a complexidade dos recursos de representação da informação está intimamente ligada aos requisitos necessários para garantia da qualidade dos sistemas de recuperação.

Na realidade, algumas concepções oriundas do senso comum têm afirmado que as folksonomias são recursos mais modernos e igualmente eficazes de representação da informação. Neste ponto, afirmamos que somos contrários a esta posição. Particularmente, acreditamos que as folksonomias não foram desenvolvidas com o compromisso de representar os conceitos dos documentos no âmbito de um sistema de informação, mas, simplesmente, como recursos de organização de informações de que o usuário dispõe de acordo com suas conveniências individuais.

Essa característica não obscurece o potencial da folksonomia para criação de novos pontos de acesso, para produção de formas adicionais de recuperação relacionadas com o conteúdo dos documentos e para além dele. Estas observações também não menosprezam a capacidade do usuário de, por intermédio de marcadores, apreender o significado dos conceitos tratados. Todavia, sobre este último aspecto, apenas os estudos sistemáticos que estão sendo empreendidos, e os que futuramente ainda virão, revelarão a real capacidade das folksonomias para o cumprimento desta função (MORRISON, 2008).

Para concluir, gostaríamos de ressaltar que o fascínio despertado pelas folksonomias nos inspirou, especificamente, pelo fato de que não podíamos imaginar, até bem poucos anos atrás, que teríamos à nossa disposição as ferramentas que utilizamos para realizar tarefas que, hoje, nos parecem imprescindíveis, mas que, até então, não faziam parte de nossa realidade. Esses recursos são possíveis graças, não apenas à evolução tecnológica, que nos levou às ferramentas de Web 2.0, mas, sobretudo, ao espírito de colaboração que caracterizam algumas formas de sociabilidade em amplo desenvolvimento na Era da Informação.

Antes, precisávamos pagar muito caro por um serviço que se propunha a coletar, tratar e disseminar dados para uma comunidade específica. Atualmente, temos uma legião de usuários, que fazem estas mesmas tarefas de modo colaborativo, alimentando bancos de dados imensos, que tão melhores ficam, quanto mais pessoas os utilizam. Desse recente comportamento, temos muito a aprender sobre o usuário-bibliotecário, que compartilha suas coleções de informações eletrônicas individuais e as representa para si, mas com um benefício, em alguma medida, de recuperação para os outros usuários.

Referências

AQUINO, M. C. *O hipertexto como potencializador da memória coletiva: um estudo dos links na Web 2.0*. 2008. 175 f. Dissertação (Mestrado em Comunicação e Informação) – Programa de Pós-graduação em Comunicação e Informação, Universidade Federal do Rio Grande do Sul, 2008. Disponível em: <<http://www.lume.ufrgs.br/handle/10183/12761>>. Acesso em: 01 ago. 2008.

ASSOCIAÇÃO BRASILEIRA DE NORMAS TÉCNICAS (ABNT). *Métodos para análise de documentos: determinação de seus assuntos e seleção de termos de indexação: procedimento*. Rio de Janeiro: ABNT, 1992.

CATTUTO, C. Semiotic dynamics in online social communities. *European Physical Journal C*, v. 46, n. 2, p. 33-37, 2006.

CINTRA, A. M. M. *et al. Para entender as linguagens documentárias*. 2. ed. rev. e ampl. São Paulo: Polis, 2002.

CRONIN, B.; ATKINS, H. B. Introduction: the scholar's spoor. In: CRONIN, B. ; ATKINS, H.B. (Eds.). *The web of knowledge: a festschrift in honor of Eugene Garfield*. Medford: Information Today, 2000. (ASIS monograph series).

ELLMAN, H. *Grandes debates da ciência*. São Paulo: Unesp, 1998.

FAROOQ, U. *et al. Social bookmarking for scholarly digital libraries*. *IEEE Internet Computing*, v. 11, p. 29-35, 2007.

GARFIELD, E. Citation indexes for science: a new dimension in documentation through association of ideas. *Science*, Washington, v. 122, n. 3159, p. 108-111, 1955.

GORDON-MURNANE, L. *Social bookmarking, folksonomies and Web 2.0 tools*. 2006. Disponível em: <http://www.redorbit.com/news/technology/541192/social_bookmarking_folksonomies_and_web_20_tools/index.html>. Acesso em: 07 set. 2008.

GUY, M.; TONKIN, E. Folksonomies: tidying up tags? *D-Lib Magazine*, v. 12, n. 1, 2006. Disponível em: <<http://www.dlib.org/dlib/january06/guy/01guy.html>>. Acesso em: 18 ago. 2008.

KORFHAGE, R. R. *Information storage and retrieval*. New York: John Willy & Sons, 1997.

LANCASTER, F. W. *Indexação e resumos: teoria e prática*. Brasília: Briquet de Lemos, 1993.

MACROBERTS, M. H.; MACROBERTS, B. R. Problems of citation analysis: a critical review. *Journal of the American Society for Information Science*, Washington, v. 40, n. 5, p. 342-9, 1989.

MORRISON, P. J. Tagging and searching: search retrieval effectiveness of folksonomies on the World Wide Web. *Information Processing and Management*, v. 44, p. 1562-1579, 2008.

PRIMO, A. Fases do desenvolvimento tecnológico e suas implicações nas formas de ser, conhecer, comunicar e produzir em sociedade. In: PRETTO, N. L.; SILVEIRA, S. A. *Além das redes de colaboração: internet, diversidade cultural e tecnologias do poder*. Salvador: UFBA, 2008. Disponível em: <<http://rn.softwarelivre.org/alemdasredes/wp-content/uploads/2008/08/livroalemdasredes.pdf>>. Acesso em: 4 set. 2008.

URL. In: WIKIMEDIA FOUNDATION. *Wikipédia: a enciclopédia livre*. 2008. Disponível em: <<http://pt.wikipedia.org/wiki/URL>>. Acesso em: 4 set. 2008.