



A taxonomia e a folksonomia na representação da informação de fotografias

Thais Helen do Nascimento Santos

Doutora em Informação e Comunicação em Plataformas Digitais pela Universidade do Porto (UP). Mestra em Ciência da Informação pela Universidade Federal da Paraíba (UFPB). Graduada em Arquivologia pela Universidade Estadual da Paraíba (UEPB). Professora substituta do curso de Arquivologia da Universidade Estadual da Paraíba (UEPB).

<http://dx.doi.org/10.1590/1981-5344/2395>

A folksonomia se caracteriza como um recurso democrático e interativo para a representação de documentos em ambiente on-line. No tratamento das fotografias, a utilização da folksonomia se torna oportuna em virtude da exaustividade e especificidade em pontos de acesso, representação por meio da linguagem natural e o fomento à inteligência coletiva. Em contrapartida, as desvantagens da polissemia conceitual e o descontrole do vocabulário implicam a eficácia dos sistemas de recuperação da informação em uso da folksonomia, fundamentando, assim, a proeminência dos vocabulários controlados na representação da informação. Nesse contexto, o objetivo deste estudo é o de indicar e preconizar o ajustamento das vantagens da folksonomia e da taxonomia na representação do conteúdo de fotografias. A metodologia empregada se voltou à revisão de literatura nas áreas da Ciência da Informação, Arte, História e Tecnologia para a elaboração do alicerce teórico, bem como à exploração de casos práticos. A investigação deflagrou os resultados satisfatórios na convergência da folksonomia com a taxonomia, isto é, na adequação dos contributos dos usuários (por meio de etiquetas, descrições e/ou comentários) aos vocabulários controlados adotados para a representação e recuperação

da informação por documentos fotográficos, textuais ou de qualquer outro gênero.

Palavras-chave: *Representação da Informação. Fotografias. Taxonomia. Folksonomia.*

The taxonomy and folksonomy in the representation of photographs information

Folksonomy is characterized as a democratic and interactive feature as far as the representation of documents online is concerned. In what regards the treatment of photographs, the use of folksonomy becomes a current issue due to the accuracy and specificities in access points, the representation by means of natural language and the promotion of collective intelligence. Yet, the conceptual polysemy and the lack of control of vocabulary both imply disadvantages concerning the effectiveness of the information retrieval systems in use by folksonomy; thus substantiating the prominence of controlled vocabulary in the representation of information. In this context, the purpose of this study is to indicate and to bring forward the adjustment of the advantages of folksonomy and taxonomy in the representation of the content of photographs. The methodology employed has mainly to do with literature review in the areas of Information Science, Art, History and Technology for the elaboration of the theoretical basis for the present study. Furthermore, this contributed to the identification of case studies. The investigation provided satisfactory results regarding the convergence of folksonomy and taxonomy, i.e., in the adequacy of users' contributes (by means of tags, descriptions and/or comments) to the controlled vocabularies adopted in the representation and retrieval of information by means of photograph documents, textual ones and documents of any other gender.

Keywords: *Representation of Information. Photographs. Taxonomy. Folksonomy.*

Recebido em 14.09.2015 Aceito em 27.02.2018

1 Introdução

“Para onde quer que nos viremos, existe a imagem” (JOLY, 2008, p. 18). A imagem se constitui em um dos principais meios de expressão e comunicação humana que remonta desde o período pré-histórico (com os desenhos encontrados nas paredes das cavernas) e se estende na contemporaneidade (interação icônica por meio de fotografias, vídeos, mensagens midiáticas e publicitárias e outras formas de representação compartilhadas no ciberespaço).

Embora o seu reconhecimento date do período pré-histórico, ainda não foi possível estabelecer um conceito de imagem. O caráter heterogêneo, múltiplo e complexo, que configura a imagem, faz com que Joly (2003, p. 197) conceba a sua noção como “[...] um poço sem fundo, onde cabem tanto as artes visuais como a imagética midiática, científica, informática, as célebres auto-estradas da informação, e assim por diante”.

A imagem, enquanto um documento no campo de estudo da Ciência da Informação, se pauta na concepção apresentada por Silva (2012), quando versa sobre as duas polaridades que compõem o documento: a inscrição e o suporte. A inscrição diz respeito aos elementos simbólicos e intelectuais que expressam a mensagem com o conteúdo informacional da imagem. No que concerne ao suporte, envolve a materialidade para o registro da mensagem, que conforme exposto acima, teve início no período pré-histórico nas paredes das cavernas e, atualmente, se cristaliza, sobretudo, em um espaço significativo potencial por meio dos *bites* e *bytes* do ciberespaço.

O documento imagético assume funções diversificadas no decorrer da sua utilidade e finalidade. Aumont (1993) elenca três funções distintas da imagem, a saber: a função simbólica, a função epistemológica e a função estética. Em termos gerais, a função simbólica se operacionaliza nas representações religiosas e pelos rituais realizados por diferentes grupos sociais. A função epistemológica concebe a imagem como um instrumento de transmissão do conhecimento. A função estética diz respeito às sensações afetivas e emotivas oriundas das imagens artísticas.

Perante as múltiplas formas de materialidade da imagem, neste estudo, nossa ênfase incide nas fotografias. No escopo da compreensão etimológica, a palavra fotografia é definida como a arte de escrever com a luz (foto = luz, grafia = escrita) (RODRIGUES, 2007; BOCCATO; FUJITA, 2006), que representa o efeito da captação da realidade em um determinado tempo por meio de processo físico-químico ou tecnológico. A fotografia é uma das principais formas de materialização da imagem na representação visual, e é passível de assumir diferentes funções segundo a intencionalidade de produção e apropriação da mensagem expressa.

O registro fotográfico é intencionalmente construído. Esta intencionalidade permeia todo o processo fotográfico, desde a seleção do tema até as técnicas apreendidas para a captação da imagem e exposição. A subjetividade inerente à imagem ainda é associada aos recursos tecnológicos (empregados no antes, durante e depois da produção fotográfica) e à volatilidade na produção, disseminação e

armazenamento dos registros imagéticos (JOLY, 2008). Sendo assim, a fotografia não se restringe a uma réplica/reprodução da realidade, uma vez que é capaz de transformá-la e recriá-la. Para tanto, cabem a interpretação e a análise do observador as quais, geralmente, ocorrem por meio da representação atribuída pelo profissional da informação (indexador), onde a fotografia está acondicionada.

Nesse contexto, o objetivo deste estudo é o de indicar e preconizar o ajustamento das vantagens da folksonomia e da taxonomia na representação do conteúdo de fotografias. Apesar da proeminência da adoção dos vocabulários controlados no processo de representação da informação, uma vez que estes garantem o controle terminológico e a eficácia da recuperação da informação, a indexação social revela a sua potencialidade nos atributos de especificidade analítica e exaustividade de pontos de acesso. A metodologia utilizada se pautou na revisão de literatura nas áreas da Ciência da Informação, Arte, História e Tecnologia para a elaboração do embasamento teórico, bem como na exploração de casos práticos em diferentes contextos.

Além desta introdução, o artigo se encontra estruturado em duas seções teóricas acerca da análise de conteúdo de imagens e a representação da informação por meio dos vocabulários controlados organizados pela taxonomia. Posteriormente, remetemo-nos à reflexão dos recursos da folksonomia, enquanto indexação de cunho social, para a representação do conteúdo de fotografias. As considerações gerais se ocupam do compêndio do estudo e da indicação de outros temas para futuras investigações.

2 Análise de conteúdo para a representação de imagens

A análise de conteúdo informacional, também designada como análise documentária, se fundamenta em um conjunto de ações específicas “[...] no interior da Documentação, que trata da análise [conceitual], síntese e representação da informação, com o objetivo de recuperá-la e disseminá-la” (CINTRA *et al.*, 2002, p. 34). A análise de conteúdo é uma das principais atividades no gerenciamento da informação, uma vez que culmina na indexação do documento para inserção na base de dados e posterior recuperação, acesso e uso pelo usuário.

A análise conceitual, dentro do processo da análise do conteúdo, corresponde à leitura do documento a fim de identificar do que se trata, ou seja, o seu conteúdo. Esta primeira etapa viabiliza a elaboração dos resumos e/ou legendas (caso as imagens não as possuam), tendo como resultado a síntese, segunda fase da análise de conteúdo. Por fim, o indexador terá os elementos necessários para a representação da informação, também denominada de indexação, que implica na conversão do conteúdo do documento em termos indexadores. Os termos indexadores podem ser extraídos do próprio documento ou atribuídos por meio do vocabulário controlado (esquemas de classificação bibliográfica, como a Classificação Decimal de Dewey (CDD) e a Classificação Decimal

Universal (CDU); lista de cabeçalhos de assunto ou tesouros). Contudo, é válido ressaltar que não há um conjunto exato de termos indexadores para representar qualquer documento. Assim, a indexação é caracterizada como uma atividade complexa, que é estabelecida a partir da missão da instituição e função do documento, bem como pelas necessidades e comportamentos de busca da informação dos usuários.

Não obstante, a análise de conteúdo assume diferentes metodologias conforme as características do documento. Os documentos em linguagem textual têm o seu conteúdo exposto por signos e significantes socialmente compartilhados pelo código alfabético, os quais instituem uma zona de familiaridade ao indexador. Ademais, os documentos de gênero textual são acompanhados por elementos que facilitam a análise, a síntese e a representação do conteúdo tais como título, sumário e outros. O documento fotográfico, por seu turno, não possui os indicadores de conteúdo e ainda são constituídos por signos e significantes próprios que podem ser inteligíveis ou não. A análise conceitual do documento fotográfico envolve o conteúdo informacional e as técnicas apreendidas na produção, em virtude de ambos os elementos formarem o conjunto da mensagem expressa. Alves e Valerio (1998, p. 5) asseveram que a fotografia demanda “[...] processamento técnico, considerando as características específicas desses documentos, que requerem uma abordagem diferente daquela que é reservada aos documentos textuais”.

Joly (2008) recomenda que a análise de imagens deve estar pautada em objetivos preestabelecidos e no uso de metodologias adequadas para o alcance daqueles. A autora ainda destaca a necessidade de estabelecer limites e pontos de referência na análise. Os referenciais analíticos devem se concentrar no usuário/receptor da informação e as suas necessidades.

A revisão de literatura nas áreas da Ciência da Informação, Arte, História e Tecnologia desvelou diferentes metodologias para a análise de imagens, tais como a de Panofsky (1989), Shatford (1986), Cordeiro (2010), Rodrigues (2007), Smit (1996¹ e 1997² *apud* MANINI, 2004), Manini (2004), Kossoy (2011), dentre outros. Os elementos de análise alternam de acordo com a(s) função(ões) da imagem (simbólica, epistemológica e/ou estética) assim como diante das necessidades de recuperação dos usuários/receptores, detectadas em estudos empíricos.

A revisão dos modelos de análise de imagens permitiu apreciar e estabelecer um panorama geral das investigações sobre a indexação e a recuperação da informação de imagens. Identificamos que o principal objetivo das investigações com essa temática é o de propor modelos de análise mais adequados ao contexto institucional, social e/ou documental pesquisado, passíveis de aplicação em outras unidades de informação. Além disso, a análise das metodologias para a descrição de imagens aponta a escassez de investigações voltadas às demais etapas da análise

¹ SMIT, J. W. A representação da imagem. *Informare*, Rio de Janeiro, v. 2, n. 2, p. 28-36, 1996.

² SMIT, J. W. *Propostas para a indexação de informação iconográfica*. 1997. (Mimeo).

de conteúdo, como é o caso da síntese e da representação da informação para a indexação. Quando identificadas, as investigações se aplicam aos sistemas de recuperação de imagens em formato digital que empregam recursos da Ciência da Computação e/ou da Engenharia de *Software* para desenvolver sofisticados métodos de recuperação da informação por voz, semelhança, códigos visuais como a cor, textura, forma e outros (MARTINÉZ COMECHE, 2013).

Para a análise das fotografias, são recomendados e adotados métodos específicos. No entanto, a tradução do conteúdo deste documento em termos indexadores ocorre, essencialmente, do mesmo modo que os documentos textuais na atribuição de termos preestabelecidos pelos vocabulários controlados.

3 A taxonomia por meio dos vocabulários controlados

A taxonomia integra os sistemas de organização da informação que se constituem em uma estruturação e sistematização de conceitos para a ordenação, organização, recuperação e comunicação da informação. Destarte, os sistemas de organização da informação incorporam instrumentos que apresentam os conceitos organizados em uma estrutura lógica pelos vocabulários controlados.

Os vocabulários controlados correspondem a uma lista de termos estruturados por relações de equivalência, homografia, hierarquia e/ou associação. A norma ANSI/NISO Z39.19, desenvolvida pela *National Information Standards Organization* (NISO) integrante da *American National Standards Institute* (ANSI) a qual é responsável pelo estabelecimento de diretrizes para a construção, formatação e gestão de vocabulários controlados, define a taxonomia como “uma coleção de termos de vocabulário controlado organizada em uma estrutura hierárquica” (ANSI/NISO, 2005, p. 9, tradução nossa). O objetivo principal destes instrumentos é o de estabelecer consistência terminológica por meio da eliminação de ambiguidades, controle de sinônimos, estabelecimento de relações lógico-semânticas entre os termos a fim de melhorar a eficácia no armazenamento, facilitar e garantir a recuperação da informação (AGANETTE; ALVARENGA; SOUZA, 2010; BOCCATO, 2011; AQUINO; CARLAN; BRASCHER, 2009).

Aganette; Alvarenga e Souza (2010) e Novo (2010) apontam duas dimensões teórico-práticas de exploração da taxonomia: 1) a taxonomia na organização e na representação da informação; e 2) a taxonomia no gerenciamento da informação. Por um lado, a primeira perspectiva diz respeito ao uso da taxonomia como base hierárquica ou poli-hierárquica na elaboração de vocabulários controlados. Por outro lado, a taxonomia no gerenciamento da informação se dirige à sua utilização como instrumento operacional na comunicação corporativa.

Na organização e representação da informação, a taxonomia aspira à ordenação dos termos com vistas ao desenvolvimento de um vocabulário controlado. Os sistemas de classificação, listas de cabeçalhos e tesouros estão correlacionados à taxonomia por assumir características

e finalidades equivalentes (SILVA; SOUZA; ALMEIDA, 2008). Contudo, as taxonomias não correspondem a um instrumento concreto, mas a uma organização abstrata dos termos de forma lógica e hierarquizada que fornece subsídios para a elaboração dos vocabulários controlados. Assim, a taxonomia não seria uma tipologia dos vocabulários controlados, e sim, um elemento que os fundamenta.

A segunda dimensão de exploração da taxonomia é no gerenciamento da informação. As taxonomias corporativas são instrumentalizadas em vocabulários controlados que refletem os contextos, a cultura organizacional e a missão da organização para classificar documentos e representar informações produzidas e acessadas dentro desta. Além disso, as taxonomias podem admitir diferentes funções, tais como: a organização de conteúdos, a recuperação, a navegação, a filtragem de informações, dentre outros (AGANETTE; ALVARENA; SOUZA, 2010). Conway e Siglar (2002) identificaram três formas de apresentação das estruturas taxonômicas em ambientes corporativos: a taxonomia descritiva, a taxonomia de navegação e a taxonomia de gerenciamento de dados. Sumariamente, a taxonomia descritiva compreende a elaboração de um vocabulário controlado com a finalidade de garantir a eficácia na recuperação da informação. A taxonomia de navegação tem como base os modelos mentais dos trabalhadores/funcionários para o estabelecimento de agrupamentos apropriados para as informações. Por conseguinte, a taxonomia de gerenciamento de dados almeja o compartilhamento de informação por grupos organizacionais através de pequenos conjuntos de termos controlados.

Na prática das unidades de informação, seja em ambientes físicos, seja em ambientes digitais, a taxonomia é empregada na organização e representação categórica da informação. No caso das bibliotecas, Vignoli; Souto e Cervantes (2013, p. 67) afirmam que as taxonomias são aplicadas para a organização “[...] sistemática dos conceitos gerais para o específico, de forma a facilitar e estruturar a classificação da informação [...] em seu tratamento pelo bibliotecário e, conseqüentemente, para a recuperação da informação para o usuário da *web*”. As estruturas taxonômicas são adotadas nas bibliotecas para a organização física dos documentos (com os sistemas de classificação bibliográfica) nas estantes e nos catálogos a fim de guiar o usuário na busca e recuperação da informação. Nos arquivos, a taxonomia é subsídio na elaboração dos planos de classificação, que se constituem em esquemas de organização dos documentos em classes e orientam o arquivamento e a representação da informação.

A adoção da taxonomia como base para o sistema de representação da informação apresenta as vantagens da garantia de busca pela exclusão de termos mútuos, possibilidade de buscas genéricas, precisão semântica e capacidade de concepção de um instrumento de orientação ao indexador e ao usuário. No que concerne às desvantagens, são recorrentes o desconhecimento do usuário do vocabulário controlado utilizado pelo indexador, obsolescência das relações lógicas (em razão do tempo),

revisões terminológicas de alto custo e demoradas e a coibição das relações de busca (CAMPOS; GOMES, 2008).

No processo de análise e indexação, a taxonomia é um dos principais recursos para a operacionalização da última etapa: a representação da informação para a tradução em termos indexadores. Geralmente, as unidades de informação utilizam os vocabulários controlados para orientar o trabalho do indexador e facilitar a busca e a recuperação da informação pelo usuário. Na representação da informação de fotografias, Manini (2004, p. 3) confirma que

De forma semelhante ao que ocorre com o texto escrito, os termos verbais, linguísticos, empregados para indexar uma imagem fotográfica estão também sob a ação das regras da polissemia, da homonímia e da antonímia; por isso são adotados os vocabulários controlados.

Sobre isso, é relevante destacar que o processo de seleção de termos de busca para a localização e recuperação de um documento em uma base de dados é um desafio para o usuário, pois este não conhece e não tem acesso ao vocabulário controlado utilizado para a representação, indexação e entrada do documento no sistema (CAMPOS; GOMES, 2008). Este desafio é ainda maior quando se trata da recuperação de fotografias, visto que o próprio usuário não tem precisão do que deseja na sua busca.

4 A folksonomia como contributo à indexação de imagens

Os estudos e experimentos recentes apontam a folksonomia como um recurso potencialmente contributivo à indexação e posterior recuperação de imagens (MARTINÉZ COMECHE, 2013) para minimizar as lacunas/desvantagens pela adoção das taxonomias (com os vocabulários controlados) na organização e representação do conteúdo de fotografias.

A folksonomia, também conhecida como etiquetagem social, classificação social, indexação social e/ou marcação social, se mostra como uma ferramenta útil para a representação e recuperação da informação em base de dados *on-line*, pois se constitui pelo e para o usuário/receptor da informação. "O usuário é o ator principal e integrante ativo em que a sua percepção e a sua opinião são pressupostos referenciais para a efetivação desses processos" (BOCCATO, 2011, p. 185).

A denominação folksonomia é derivada da substituição do prefixo 'tax', da taxionomia, por 'folk' (do inglês: povo). O conceito contesta a regularização de termos indexadores, prevista com a taxonomia (por meio dos vocabulários controlados), para a atribuição livre e pessoal de etiquetas indexadoras.

Folksonomia é o resultado da etiquetagem livre e pessoal da informação e objetos (algo com uma URL) para sua própria recuperação. A etiquetagem é realizada em um ambiente social (usualmente compartilhado e aberto para outros). Folksonomia é

criada para a ação de etiquetagem por pessoas consumidoras de informação (WAL, 2007, *on-line*, tradução nossa).

A indexação elaborada pelo povo é não hierárquica e se estrutura a partir de correlações associativas. A atribuição de termos/etiquetas advém, especialmente, do interesse em recuperar o mesmo documento em uma nova e/ou outra necessidade de informação. A seleção das etiquetas é subjetiva e ocorre por interpretações e motivações diferentes. Destarte, o conjunto de termos atribuídos ao documento implica a diversidade terminológica com a exaustividade e a especificidade, assim como em um maior número de índices de recuperabilidade do documento (BRANDT; MEDEIROS, 2010; SANTOS, 2013; CATARINO; BAPTISTA, 2007; STREHL, 2011).

De acordo com Brandt e Medeiros (2010) e Sinha (2005), o processo cognitivo para a atribuição de etiquetas a um documento através da folksonomia envolve três diferentes estágios. O primeiro estágio consiste no desejo do usuário em recuperar, em outro momento, o documento. Esse interesse provoca a interpretação cognitiva (livre e pessoal) dos signos e símbolos apresentados naquela informação gerando a produção de termos conceituais, que confere o segundo estágio. O terceiro estágio culmina na seleção do(s) termo(s) e na sua atribuição ao documento. Portanto, a inserção de etiquetas por meio da folksonomia segue a mesma estrutura funcional utilizada pela taxonomia, desde a leitura e interpretação do documento textual e/ou imagético à seleção de termos que o representam (LANCASTER, 2004).

Os estudos de indexação pela folksonomia divergem entre a relevância das vantagens e das desvantagens dessa prática. Dentre as principais vantagens, estão: a colaboração que provoca a inteligência coletiva (LÉVY, 1997), formação de comunidades em torno de assuntos de interesse, inexistência de controle do vocabulário (liberdade sociocultural), etiquetas em rede (acessíveis em qualquer lugar por qualquer usuário). Em contrapartida, o descontrole do vocabulário, o grande índice de revocação e de baixa taxa de precisão e a polissemia terminológica se destacam como algumas das desvantagens (BRANDT; MEDEIROS, 2010; SANTOS, 2013; CATARINO; BAPTISTA, 2007; STREHL, 2011).

O ambiente colaborativo e participativo da *web 2.0* é onde se manifesta a folksonomia. Na transição de sistemas rígidos da *web 1.0*, a *web 2.0* se cristaliza na flexibilidade para a interação e a participação do usuário (O'REILLY, 2005). As plataformas *on-line* mais conhecidas que utilizam o recurso da folksonomia são o Del.icio.us e o Flickr, sendo que, respectivamente, uma concerne à coleta e descoberta de diversos conteúdos *web*, e a outra é destinada ao armazenamento, compartilhamento e recuperação de imagens.

Na representação do conteúdo de fotografias, a folksonomia tem se desvelado como recurso que traz significativas vantagens em relação aos vocabulários controlados de base taxonômica. Ménard (2010, p. 246, tradução nossa) escreve que "o uso de folksonomias, ou etiquetas, agora

também é considerado outro caminho possível". Sobre esta hipótese, a autora ainda complementa:

Usuários podem atribuir livremente palavras-chave aos documentos, incluindo recursos visuais, a fim de melhorar a recuperação. A atribuição destas palavras-chave livres conquistou popularidade durante os anos. Este tipo de indexação adiciona uma nova camada para prover um determinado vocabulário de indexação, às vezes muito diferente do que o especialista em indexação atribuiria no mesmo contexto (MÉNARD, 2010, p. 247, tradução nossa).

A defesa de Ménard (2010) pela etiquetagem social destaca a vantagem da representação exaustiva em linguagem natural. Grupos de usuários podem ver e interpretar as fotografias de modo diferente. Ora, a complexidade em relação à imagem é dualista, pois é inerente ao trabalho do indexador na análise conceitual do documento e do usuário, que não consegue precisar o documento que deseja recuperar em meio aos atributos extrínsecos (conteúdo informacional) e aos atributos intrínsecos (cor, forma e textura) (LANCASTER, 2004). Nesse sentido, a indexação livre e pessoal fomenta a exaustividade de pontos de acesso aos documentos dentro do sistema de recuperação da informação. Ademais, as diferentes formas de leitura e interpretação do mundo pelos usuários fornecem variadas percepções e modelos de representação da informação. Essas representações resultam em etiquetas, descrições e comentários compartilhados em linguagem natural que facilitam a recuperação do documento por diferentes grupos de usuários.

Martínez Comeche (2013) realizou uma densa revisão de literatura acerca dos sistemas de recuperação de imagens. Seu estudo elencou três etapas destes sistemas, a saber: 1) sistemas de recuperação de imagens baseados em representações textuais das características das imagens; 2) sistemas de recuperação de imagens pautados nos recursos visuais das imagens e; 3) uso simultâneo do código visual e do código textual para a representação e recuperação nos sistemas. Os efeitos da investigação destacam que um dos diferenciais na eficácia dos sistemas está na adoção de folksonomia para a representação das imagens.

Entre as conclusões mais destacadas obtidas durante estes poucos anos de investigação, destacam-se os melhores resultados obtidos empregando programas de aprendizagem de máquina que implementam classificadores a partir de descritores básicos e o abandono paulatino de ferramentas complementares como taxonomias e ontologias em favor do emprego de dados provenientes da etiquetagem social (MARTÍNEZ COMECHE, 2013, p. 429, tradução nossa).

Apesar de embrionários, os estudos sobre o uso da folksonomia na representação e recuperação de imagens têm mostrado resultados satisfatórios. Conquanto, reconhecemos que o descontrole do vocabulário, a alta taxa de revocação e baixa taxa de precisão, assim como a polissemia terminológica implicam a eficácia dos sistemas de recuperação da informação. Assim, nossa defesa se pauta na convergência das

vantagens da folksonomia à taxonomia, isto é, na adequação dos contributos dos usuários com etiquetas, descrições e comentários para os vocabulários controlados (sistemas de classificação bibliográfica, listas de cabeçalhos de assuntos, tesouros e outros) adotados pelas unidades de informação.

Algumas iniciativas são identificadas nesse sentido, como relata Gracioso (2010) em um projeto que visa o desenvolvimento de uma metodologia para um instrumento de recuperação da informação *web* através da linguagem natural cotidiana. O seu *corpus* de análise é a plataforma digital utilizada pelo Centro de Comunicação Social da Universidade Federal de São Carlos (UFSCar). Em andamento, o projeto tem como um de seus objetivos o de elaborar taxonomias por meio da folksonomia e categorizar uma lista de termos estabelecendo relações entre estes.

Outro estudo que explorou a relação entre as taxonomias (vocabulário controlado) e as folksonomias foi desenvolvido por Yi e Chan (2009). A finalidade deste projeto foi a de explorar o potencial da *Library of Congress Subject Headings* (LCSH) – a lista de cabeçalho de assuntos (tipo de vocabulário controlado) utilizada pela *Library of Congress* dos Estados Unidos – “[...] em mitigar a ambiguidade e complexidade causadas pelo descontrolo das etiquetas selecionadas pelos usuários (folksonomia) na área da descoberta e organização da informação colaborativa e o compartilhamento com outros sistemas de informação” (YI; CHAN, 2009, p. 874, tradução nossa). Para tanto, confrontaram as etiquetas que foram atribuídas em cerca de quatro mil e quinhentas páginas no Del.icio.us com as da LCSH.

A LCSH é estruturada por quatro diferentes tipos de relações hierárquicas, baseadas na taxonomia, que são: termos estabelecidos (termos autorizados), termos *lead-in* (termos não autorizados), termos genéricos e termos específicos. A análise empírica revelou que de todos os termos atribuídos pelos usuários, cerca de 61% estão presentes na LCSH. Os outros 39% dos termos passaram por um tratamento linguístico e semântico e, posteriormente, foram adequados ao LCSH por três padrões de distribuição: *skewedness* (gerenciamento da dispersão hierárquica entre os termos), *multifacet* (controle de sinônimos e homônimas) e *Zipfian-pattern* (organização e/ou reorganização das relações hierárquicas entre os termos em face da relevância (por recorrência) dos mesmos). A conclusão do projeto apontou benefícios na recuperação da informação com a integração da folksonomia à LCSH.

No caso das imagens, Guldogan e Gabbouj (2010) desenvolveram uma metodologia para a adequação da folksonomia em esquemas de classificação em dois conjuntos de imagens: um com trezentos e sessenta imagens armazenadas no Flickr, e o outro com mais de dez mil imagens arquivadas em um sistema particular. Em termos gerais, a metodologia proposta pelos autores se constitui em três fases principais: 1) alimentação do banco de dados de imagens com as etiquetas de atributos extrínsecos (conteúdo informacional); 2) desenvolvimento de recursos visuais para as imagens pertencentes ao mesmo grupo de categorização;

e 3) recategorização em toda a base de dados para adaptação dos novos termos aos termos já existentes. Igualmente, os resultados desse estudo apontaram aspectos satisfatórios na precisão da classificação com a integração da folksonomia.

5 Considerações gerais

A análise de conteúdo é uma atividade elementar à representação da informação. Os resultados desta atividade permitem a localização e a recuperação da informação desejada pelo usuário. Porém, em face da heterogeneidade, multiplicidade e complexidade que caracterizam as fotografias, o tratamento deste gênero documental demanda recursos e metodologias específicas de análise, representação e indexação pelo profissional da informação.

Ainda que haja o reconhecimento da necessidade de formas de tratamento diferenciadas às fotografias, a representação do seu conteúdo se dá pelos mesmos parâmetros utilizados na representação de documentos textuais com os vocabulários controlados, uma vez que estes conferem o controle terminológico, a precisão semântica e, logo, a eficácia na recuperação da informação.

No processo de análise e indexação de imagens é fundamental considerar o usuário/receptor e as suas dificuldades em recuperar a informação. Precisar a imagem desejada por seus atributos intrínsecos e/ou extrínsecos é uma tarefa difícil ao usuário (BOCCATO; FUJITA, 2006) e, por vezes, a sua necessidade é modelada conforme os documentos recuperados na busca. Nesse pensamento, fazem-se pertinentes a exaustividade e a especificidade na indexação de fotografias fomentadas pela folksonomia, em virtude da oferta de um maior índice de pontos de acesso.

Em suma, aquiescemos que o controle do vocabulário na indexação de fotografias torna a recuperação da informação mais eficiente. Todavia, reconhecemos que os vocabulários controlados desconsideram a colaboração do usuário – escopo das atividades desenvolvidas em qualquer unidade de informação – e a sua capacidade de análise, interpretação e sugestão de termos indexadores. Assim, a folksonomia se desvela como um recurso para a indexação de imagens capaz de obter contributos significativos para a atualização e adequação pertinentes aos vocabulários controlados adotados pelas unidades de informação, pelo que alcançamos o objetivo proposto por este estudo que foi o de indicar e preconizar o ajustamento das vantagens da folksonomia e da taxonomia na representação do conteúdo de fotografias.

A defesa da representação das fotografias pela convergência dos recursos da taxonomia com os da folksonomia não esgota o tema do tratamento das imagens pelas unidades de informação. Por isso, futuras investigações acerca desta temática podem estar voltadas à ênfase nas vantagens de adequação dos instrumentos da taxonomia com os recursos da folksonomia em outros gêneros documentais, tais como os sonoros, audiovisuais, cartográficos, dentre outros. Outrossim, em virtude da

carência de estudos práticos da adoção da indexação social pelas unidades e/ou serviços de informação, é oportuno explorar, descrever e analisar essas práticas tendo em vista a sua divulgação e análise crítica pelos pesquisadores em Ciência da Informação.

Referências

AGANETTE, E.; ALVARENGA, L.; SOUZA, R. R. Elementos constitutivos do conceito de taxonomia. *Informação & Sociedade: Estudos*, João Pessoa, v. 20, n. 3, p. 77-93, 2010.

ALVES, M. C.; VALERIO, S. A. *Manual para a indexação de documentos fotográficos*. Rio de Janeiro: Fundação Biblioteca Nacional, 1998.

NATIONAL INFORMATION STANDARDS ORGANIZATION (NISO). AMERICAN NATIONAL STANDARDS INSTITUTE (ANSI). *ANSI/NISO Z39.19: guidelines for the construction, format, and management of monolingual controlled vocabularies*. Baltimore, 2005.

AQUINO, I. J.; CARLAN, E.; BRASCHER, M. B. Princípios classificatórios para a construção de taxonomias. *PontodeAcesso*, Salvador, v. 3, n. 3, p. 196-215, 2009.

AUMONT, J. *A imagem*. Campinas: Papirus, 1993.

BOCCATO, V. R. C. Os sistemas de organização do conhecimento nas perspectivas atuais das normas internacionais de construção. *InCID*, Ribeirão Preto, n. 2, v. 1, p. 165-192, 2011.

BOCCATO, V. R. C.; FUJITA, M. S. L. Discutindo a análise documental de fotografias: uma síntese bibliográfica. *Cadernos BAD*, Lisboa, n. 2, p. 84-100, 2006.

BRANDT, M.; MEDEIROS, M. B. B. Folksonomia: esquema de representação do conhecimento? *Transinformação*, Campinas, v. 22, n. 2, p. 111-121, 2010.

CAMPOS, M. L. de A.; GOMES, H. E. Taxonomia e classificação: o princípio de categorização. *DataGramaZero*, Rio de Janeiro, v. 9, n. 4, 2008. Disponível em: <http://www.dgz.org.br/ago08/Art_01.htm>. Acesso em: 10 abr. 2015.

CATARINO, M. E.; BAPTISTA, A. A. Folksonomia: um novo conceito para a organização dos recursos digitais na Web. *DataGramaZero*, Rio de Janeiro, v. 8, n. 3, 2007. Disponível em: <http://www.dgz.org.br/jun07/Art_04.htm>. Acesso em: 10 abr. 2015.

CINTRA, A. M. M. *et al. Para entender as linguagens documentárias*. São Paulo: Polis, 2002.

CONWAY, S.; SLIGAR, C. *Unlocking knowledge assets*. Redmond: Microsoft Press, 2002.

CORDEIRO, R. I. N. Análise e representação do conteúdo de imagens para o acesso à informação. In: FREITAS, L. S. de; MARCONDES, C. H.;

RODRIGUES, A. C. (Org.). *Documento: gênese e contextos de uso*. Niterói: Editora da UFF, 2010. p. 235-246.

GRACIOSO, L. de S. Parâmetros teóricos para elaboração de instrumentos pragmáticos de representação e organização da informação na Web: considerações preliminares sobre uma possível proposta metodológica. *InCID*, Ribeirão Preto, v. 1, n. 1, p. 138-158, 2010.

GULDOGAN, E.; GABBOUJ, M. Adaptive image classification based on folksonomy. *In: INTERNATIONAL WORKSHOP ON IMAGE ANALYSIS FOR MULTIMEDIA INTERACTIVE SERVICES*, 11., 2010, Desenzano del Garda. *Anais...* Desenzano del Garda: IEEE, 2010. p. 1-4.

JOLY, M. *A imagem e a sua interpretação*. Lisboa: Edições 70, 2003.

JOLY, M. *Introdução à análise da imagem*. Lisboa: Edições 70, 2008.

KOSSOY, B. *Fotografia e história*. São Paulo: Ateliê, 2001.

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

LÉVY, P. *A inteligência coletiva: para uma antropologia do ciberespaço*. São Paulo: Loyola, 1997.

MANINI, M. P. Análise documentária de fotografias: leitura de imagens incluindo sua dimensão expressiva. *Cenário arquivístico*, Brasília, v. 3, n. 1, p. 16-28, 2004.

MARTÍNEZ COMECHE, J. A. La recuperación automatizada de imágenes: retos y soluciones. *Revista General de Información y Documentación*, Madrid, v. 23, n. 2, p. 423-436, 2013.

MÉNARD, E. Image indexing and retrieval: challenges and new perspectives. *Canadian Journal of Information and Library Science*, Montreal, v. 34, n. 3, p. 245-248, 2010.

NOVO, H. F. A taxonomia enquanto estrutura classificatória: uma aplicação em domínio de conhecimento interdisciplinar. *PontodeAcesso*, Salvador, v. 4, n. 2, p. 131-156, 2010.

O'REILLY, T. *What is Web 2.0: design patterns and business models for the next generation of software*. 2005. Disponível em: <<http://oreilly.com/pub/a/Web2/archive/what-is-Web-20.html?page=1>>. Acesso em: 2 abr. 2015.

PANOFSKY, E. *Significado das artes visuais*. Barcarena: Editorial Presença, 1989.

RODRIGUES, R. C. Análise e tematização da imagem fotográfica. *Ciência da Informação*, Brasília, v. 36, n. 3, p. 67-76, 2007.

SANTOS, H. P. Etiquetagem e folksonomia: o usuário e a sua motivação para organizar e compartilhar informação na Web 2.0. *Perspectivas em Ciência da Informação*, Belo Horizonte, v. 18, n. 2, p. 91-104, 2013.

SHATFORD, S. Analyzing the subject of a picture: a theoretical approach. *Cataloging and Classification Quarterly*, Chicago, v. 6, n. 3, p. 39-62, 1986.

SILVA, A. M. da. O impacto do uso generalizado das TIC (Tecnologias de Informação e Comunicação) no conceito de documento: ensaio crítico-analítico (I). *Prisma.com*, Porto, n. 16, p. 1-61, 2012.

SILVA, D. L. da; SOUZA, R. R.; ALMEIDA, M. B. Comparação de metodologias para construção de ontologias e vocabulários controlados. In: SEMINÁRIO DE PESQUISA EM ONTOLOGIA NO BRASIL, 1., 2008, Niterói. *Anais...* Niterói: UFF, 2008. p. 1-5.

SINHA, R. A cognitive analysis of tagging (or how the lower cognitive cost of tagging makes it popular). 2005. Disponível em: <<http://rashmisinha.com/2005/09/27/a-cognitive-analysis-of-tagging/>>. Acesso em: 2 abr. 2015.

STREHL, L. As folksonomias entre os conceitos e os pontos de acesso: as funções de descritores, citações e marcadores nos sistemas de recuperação da informação. *Perspectivas em Ciência da Informação*, Belo Horizonte, v. 16, n. 2, p. 101-114, 2011.

VIGNOLI, R. G.; SOUTO, D. V. B.; CERVANTES, B. M. N. Sistemas de organização do conhecimento com foco em ontologias e taxonomias. *Informação & Sociedade: Estudos*, João Pessoa, v. 23, n. 2, p. 59-72, 2013.

WAL, T. V. *Folksonomy*. 2007. Disponível em: <<http://vanderwal.net/folksonomy.html>>. Acesso em: 2 abr. 2015.

YI, K.; CHAN, L. M. Linking folksonomy to Library of Congress subject headings: na exploratory study. *Journal of Documentation*, Londres, v. 65, n. 6, p. 872-200, 2009.