



XI Encontro Nacional de Pesquisa em Ciência da Informação
Inovação e inclusão social: questões contemporâneas da informação
Rio de Janeiro, 25 a 28 de outubro de 2010

GT 2 :Organização do Conhecimento e Representação da Informação

Modalidade de apresentação: comunicação oral

FOLKSONOMIA: ANÁLISE DA ETIQUETAGEM DE IMAGENS NO FLICKR

Andre Abreu Rodrigues

Universidade Federal de Minas Gerais

Manoel Palhares Moreira

Pontifícia Universidade Católica de Minas Gerais

Resumo: A Folksonomia, através da utilização de etiquetas, representa uma mudança nos processos de organização e tratamento da informação na web. Se antes todo o conteúdo armazenado era indexado por heurísticas de máquinas de busca, agora este processo conta com a participação daqueles que disponibilizam os documentos e deles fazem uso: os usuários. A participação direta do usuário em um processo que outrora cabia somente a pessoal especializado levantou o debate sobre os efeitos da Folksonomia na organização da informação na internet tornando clara a necessidade de se estudar como os usuários realizam este processo. Tal inserção por si só traz grandes desafios que se ampliam em ambientes onde o usuário realiza sua etiquetagem para documentos do tipo imagem. Este trabalho objetivou conhecer as estratégias de indexação utilizadas pelos usuários na etiquetagem de imagens no Flickr. Apresenta-se uma revisão da literatura sobre o eixo temático Folksonomia, Indexação e Indexação de imagens, como base para a análise e a classificação da atividade de etiquetagem pelo usuário. Baseou-se em adaptações no método proposto por Manini (2002) aos níveis de interpretação de imagens de Panofsky (1979), propondo uma forma de análise e interpretação da etiquetagem do usuário. A Netnografia foi empregada para compor o percurso metodológico permitindo abordar diretamente o usuário através de entrevistas para verificar as estratégias de etiquetagem percebidas na análise da etiquetagem dos usuários, possibilitando também avaliar o método de análise proposto. Essa abordagem culminou numa forma de análise e classificação da atividade do usuário mais precisa e consistente, sendo possível através dela perceber as estratégias de etiquetagem de imagens utilizadas pelos usuários do Flickr envolvidos no trabalho.

Palavras-chave: Folksonomia. Indexação. Indexação de Imagens. Etiquetagem de Imagens. Netnografia.



1 INTRODUÇÃO

Resultado de dois anos de pesquisa em torno da temática da Folksonomia, o presente artigo representa, de forma resumida, a dissertação de mestrado homônima apresentada e aprovada ao Programa de Pós-graduação em Ciência da Informação da Universidade Federal de Minas Gerais em agosto de 2010. Para apresentar de forma resumida o referido trabalho, este artigo conta com os seguintes elementos: no capítulo 1 é apresentada sucinta introdução ao tema contando com as justificativas, o problema de pesquisa e os objetivos estabelecidos para o trabalho; no capítulo 2 é apresentado o referencial teórico utilizado na pesquisa; no capítulo 3 é descrito o percurso metodológico desenvolvido; em seguida, o capítulo 4 apresenta um exemplo da análise proposta realizada; no capítulo 5 estão expostos os principais resultados e contribuições da pesquisa; para por fim, no capítulo 6 são exibidas as considerações finais e indicações de trabalhos futuros.

Tidos como a grande sensação da Internet, atualmente, *sites* e programas de compartilhamento de arquivos surgiram na rede atraindo um número alto de usuários. Eles permitem que o usuário disponibilize livremente todo tipo de conteúdo, em variados formatos, para que estes sejam acessados pelos demais usuários, representando assim uma potente ferramenta de compartilhamento. Também possibilitam a decisão da forma como esse conteúdo será representado, através da inserção de etiquetas que possibilitem a recuperação. Este processo de livre escolha da representação de um conteúdo pelo usuário foi denominado Folksonomia, termo cunhado por Thomas Vander Wal em 2004.

Observa-se a necessidade de alargar a compreensão sobre esse instrumento, no sentido de conhecer como o usuário vem utilizando a Folksonomia e se aproveitando da oportunidade de indexar o conteúdo que disponibiliza, sendo possível assim responder perguntas como: qual a relação da etiqueta com o conteúdo representado? Qual a motivação que sustenta o processo de etiquetagem do usuário? Como pode ser classificada a etiquetagem do usuário? Quais estratégias de etiquetagem são utilizadas? Qual a influência de aspectos coletivos e de aspectos individuais na formação da Folksonomia do sistema e na Folksonomia do usuário?



Elegeram-se como objeto de estudo o Flickr (FLICKR, 2008), que é um aplicativo *on line* de compartilhamento e gerenciamento de fotos e vídeos onde é permitido ao usuário inserir, organizar e compartilhar o conteúdo que disponibiliza.

Este trabalho tem como objetivo geral observar as estratégias da utilização de etiquetas pelos usuários do Flickr, em língua portuguesa, e propor uma forma de análise e classificação das mesmas; e, como objetivos específicos: a) Conhecer os procedimentos e estratégias que envolvem o processo de etiquetagem de imagens no Flickr; b) Propor uma forma de análise e classificação da etiquetagem do usuário a partir do modelo de Manini (2002) e das categorias de Panofsky (1979); c) Observar as diferenças e similaridades da etiquetagem de usuários com diferentes níveis de experiência e participação no Flickr; e d) observar diferenças entre a etiquetagem mais antiga e mais recente de um mesmo usuário.

2 FUNDAMENTAÇÃO TEÓRICA

2.1 Indexação

A Indexação é uma prática bastante antiga que remete às origens da escrita e da biblioteca. No entanto, é a partir de meados do século XIX e durante o século XX que a atividade ganha importância e espaço, sendo conceituada e elaborada cientificamente como um processo capaz de tratar a explosão documental que estava em curso (FUJITA, 2004).

No final do século XIX e início do XX começam a aparecer as primeiras ferramentas de descrição de documentos com fins de pesquisa documental. Nas formulações que envolviam a fundação do *Office International de Bibliographie*¹, em Bruxelas, no ano de 1892 (MOREIRA, 2005), por Paul Otlet, há outro aspecto da evolução conceitual da Documentação, destacando que a atividade não deveria estar focada apenas no livro e na biblioteca, mas também na “(...) idéia de bibliografia como registro, memória do conhecimento científico, desvinculada dos organismos, como arquivos e

¹ Instituto Internacional de Bibliografia (tradução do autor).



XI Encontro Nacional de Pesquisa em Ciência da Informação
Inovação e inclusão social: questões contemporâneas da informação
Rio de Janeiro, 25 a 28 de outubro de 2010

bibliotecas, e de acervos.” (PINHEIRO, 1997). O avanço nos métodos e técnicas de construção de índices os transformou nos mais importantes instrumentos de recuperação da informação, a mais poderosa ferramenta de busca (ROBREDO, 2005), o que seria mantido até o surgimento dos modernos motores de busca, que, em última análise, funcionam com base na construção de índices.

Exibindo essa expansão e consolidação do conceito de Indexação, Lancaster (1993) conclui que a Indexação deve identificar, descrever ou caracterizar o conteúdo informacional de um documento para posterior exposição em linguagem natural ou em índices, sendo que isso pode ser feito na forma de uma indexação sintética (de acordo com seu assunto) ou analítica (de acordo com os conceitos), em que deve sempre ser observada também sua finalidade.

Viabilizado pelo surgimento do computador (AZEVEDO, 2009), nos primórdios da década de 1970, a utilização dos bancos de dados na organização da informação em bibliotecas iniciou um novo paradigma para a representação e recuperação da informação. O desenvolvimento da Internet, em 1991, por Tim Berners-Lee, revolucionou a velocidade e o fluxo de informações na sociedade de tal forma que seus desdobramentos são estudados até hoje. Essas contundentes transformações levaram à revisão de vários conceitos, como o conceito de documento, que teria sua compreensão ampliada e passaria a ser entendido como contendo três dimensões. Assim, o documento passou a ser percebido como forma, como signo e como meio (MOURA, 2009).

Apesar das diferenças nos processos que compõem a Folksonomia, ela cumpre os objetivos de um processo de indexação na representação de um documento, agora na visão do usuário. A questão que se coloca se refere à participação do usuário. Nos processos de indexação convencional, o usuário não é um participante ativo, mas o pressuposto da garantia de usuário faz com que ele não seja esquecido. Na Folksonomia, é ele próprio quem assegura esta garantia, além de ser o responsável pela representação do recurso e da facilidade ou não de sua recuperação.



2.2 Folksonomia

O termo Folksonomia é um neologismo criado por Thomas Vander Wal (2004) e consiste na junção das palavras taxonomia e *folk* (povo, em inglês). Trata-se de uma estrutura de indexação através da atribuição de palavras-chaves que representem o conteúdo, sendo que é arbitrado ao próprio usuário, que insere esse conteúdo, a escolha dessas palavras.

A Folksonomia é o processo de indexação adotado por muitos dos sistemas de compartilhamento de arquivos na Internet, atualmente. Neles a representação do conteúdo é feita em linguagem livre do usuário, sem qualquer tipo de intervenção direta de gestores ou mecanismos de organização especializados. De forma concisa, a Folksonomia pode ser entendida como "(...) uma forma de os próprios usuários disponibilizarem os rótulos adotados intuitivamente no compartilhamento, organização e na recuperação futura da informação." (MOURA, 2009).

Apesar das controvérsias e críticas feitas à Folksonomia quanto à sua efetividade enquanto mecanismo de organização e estratégia de busca, dada a falta de controle do vocabulário e de estrutura organizacional, "(...) o resultado para quem pesquisa é uma maior facilidade para encontrar termos que as demais linguagens de indexação não conseguem acompanhar em suas tabelas hierárquicas." (BLATTMANN e SILVA, 2007). Assim, este trabalho considera a Folksonomia como uma atividade, um processo de representação da informação, decorrente da evolução do uso da Internet.

2.3 Indexação de imagens: Análise Documentária de Imagens

O conceito de imagem refere-se a um vasto leque de documentos iconográficos, como pinturas, gravuras, pôsteres, cartões postais, fotografias, entre outros (SMIT, 1996). Os recursos encontrados no *Flickr* podem representar fotograficamente uma pintura, um cartão postal ou uma gravura, materiais que enquanto registros físicos necessitam de um tratamento diferenciado (SMIT, 1996). No entanto, em última análise, os recursos contidos no *Flickr* são imagens fotográficas desses e de outros tipos de registros.



XI Encontro Nacional de Pesquisa em Ciência da Informação
Inovação e inclusão social: questões contemporâneas da informação
Rio de Janeiro, 25 a 28 de outubro de 2010

Apesar de pouco numerosa, principalmente em relação ao volume de conteúdo publicado sobre a produção e realização da Indexação de livros, artigos, resumos, etc., a bibliografia sobre Análise Documentária de Imagens possui importantes referências como Shatford (1986), Shatford Layne (1994), Smit (1996) e Manini (1997, 2002 e 2009).

Um dos elementos mais básicos da representação de fotografias é a identificação de seu referente. De acordo com Barthes (1984), Dubois (1994), o conceito de referente diz respeito ao que está focado na imagem, ao objeto que está representado na imagem, sem que, em última análise, este represente na totalidade o objeto focado ou a imagem fotográfica enquanto documento. O referente adere à imagem fotográfica e representa o objeto ou conceito que está em foco sem, no entanto, confundir-se com ele; e, apesar de poder ser isolado, não serve de representação final ou total da imagem fotográfica (SMIT, 1996). No caso de uma imagem fotográfica cujo objeto em foco é uma ponte, o referente genérico é “ponte” e o referente específico é a identificação de qual ponte está representada na imagem. (SMIT, 1996)

A revisão bibliográfica sobre Análise Documentária de Imagens permitiu o encaixe dos níveis de análise de uma imagem prescritos por Panofsky (1979), com a análise da relação entre referente e etiqueta proposto neste trabalho. Apesar de suas análises não visarem à documentação de uma obra, Panofsky contribuiu para a evolução dos conceitos sobre o tema, ao distinguir em três níveis a análise de uma imagem, a saber: pré-iconográfico, iconográfico e iconológico.

Conforme Smit (1996), os níveis de análise podem ser resumidos da seguinte forma: o nível pré-iconográfico refere-se à descrição genérica dos objetos e ações representados em uma imagem; o nível iconográfico refere-se ao estabelecimento de um assunto secundário ou convencional, à determinação do significado abstrato, simbólico, da imagem, a partir dos componentes detectados na análise pré-iconográfica; o nível iconológico diz respeito à interpretação do significado intrínseco do conteúdo da imagem baseado nos níveis anteriores, porém com influência do conhecimento do analista sobre o ambiente cultural, artístico e social no qual a imagem foi gerada (PANOFSKY, 1979). Os níveis de análise de imagem de Panofsky (1979) oferecem categorias que nos permitem caracterizar a etiquetagem do usuário e perceber a movimentação realizada por este no momento da etiquetagem. Concluindo sobre os elementos da análise documentária de



imagem, destaca-se a evolução destas formulações. Para analisar pinturas sacras da época da Renascença e revelar seu significado, Panofsky (1979) elaborou os níveis de leitura de uma imagem que serviram de base para Shatford (1986) na formulação das categorias “DE genérico”, “DE específico” e “SOBRE” (Quadro 1).

QUADRO 1 Quadro representando adaptação de Shatford (1986) às categorias de Panofsky (1979)

PANÓFSKY	Exemplo	SHATFORD	Exemplo
Nível pré-iconográfico, significado fatural	Homem levanta o chapéu.	DE genérico	Ponte
Nível iconográfico, significado fatural	Sr. Andrade levanta o chapéu	DE específico	Ponte das Bandeiras
Níveis pré-iconográfico + iconográfico, significado expressivo.	Ato de cortesia, demonstração de educação etc.	SOBRE	Transporte urbano, São Paulo, Rio Tietê, arquitetura, urbanização, etc.

Fonte: Smit (1996).

Manini (2002) propõe a inclusão da dimensão expressiva da imagem na leitura documentária da mesma. Integrando e ampliando as considerações de Smit (1996), Panofsky (1979), Barthes (1984) e Shatford Layne (1994), a autora enfoca o fato de que a leitura documentária de uma fotografia não deve se ater somente ao conteúdo informacional da mesma, levando-se em conta também a forma como esta foi produzida (Dimensão Expressiva) (MANINI, 2002). O modo de leitura de uma imagem elaborado por Manini (2002) propõe as seguintes categorias de análise: a) Quem/O que = identificação do “objeto focado”: seres vivos, artefatos, construções, acidentes naturais, etc.; b) Onde = localização da imagem no espaço: espaço geográfico ou espaço da imagem; c) Quando = localização da imagem no tempo: tempo cronológico ou momento da imagem; e d) Como = descrição de atitudes ou detalhes relacionados ao “objeto focado” quando este é um ser vivo (MANINI, 2002). Essas categorias são associadas às categorias “DE genérico”, “DE específico” e “SOBRE” de Shatford Layne (1994), e à categoria “Dimensão Expressiva”, conforme Quadro 2.



QUADRO 2 Quadro que resume o método de análise proposto por Manini (2002)

	Conteúdo Informacional		Dimensão Expressiva
	DE	SOBRE	
Categoria	Genérico	Específico	
Quem/O Que			
Onde			
Quando			
Como			

Fonte: MANINI, 2002.

3 PROCEDIMENTOS METODOLÓGICOS

Ao se definir o tipo de pesquisa a ser realizada em um trabalho é importante considerar aspectos como os recursos materiais, temporais e das pessoas envolvidas, disponíveis para responder a uma determinada pergunta científica, bem como, e principalmente, o objetivo a ser alcançado com tal pesquisa. Faz-se necessário o levantamento das características de cada uma delas e, a partir daí, sua definição (Godoy, 1995a). A pesquisa qualitativa é eficaz para se entender a profundidade dos fenômenos, levando em conta toda a sua complexidade e particularidade com o objetivo de se alcançar o entendimento das singularidades. Diante do objetivo principal de se observar as estratégias de etiquetagem dos usuários e de observar seu comportamento em um determinado contexto, optou-se pela pesquisa Qualitativa, que, segundo Godoy (1995b), oferece ainda três diferentes possibilidades de trabalho: a pesquisa documental, o estudo de caso e a etnografia.

A Pesquisa Documental é realizada através do estudo de materiais que ainda não foram analisados ou que podem ser reexaminados para uma nova ou complementar interpretação. Tem a vantagem de permitir o estudo de eventos e pessoas distantes temporal e fisicamente (Godoy, 1995b). O Estudo de Caso permite uma análise aprofundada de pessoas ou eventos a serem estudados, permitindo se chegar a conclusões dos motivos pelos quais e das maneiras através das quais estes acontecem. É o método mais eficaz para a análise de eventos sobre os quais a possibilidade de controle é reduzida ou quando os fenômenos analisados são atuais e só fazem sentido



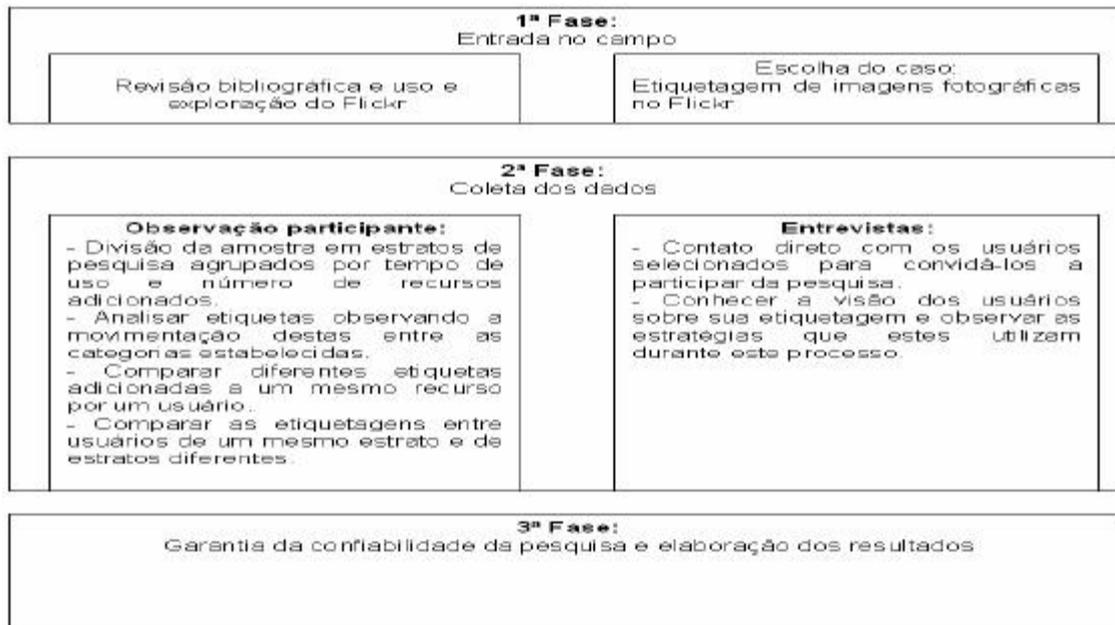
XI Encontro Nacional de Pesquisa em Ciência da Informação
Inovação e inclusão social: questões contemporâneas da informação
Rio de Janeiro, 25 a 28 de outubro de 2010

dentro de um contexto específico. A Etnografia é um método onde o pesquisador tem um maior envolvimento pessoal e se utiliza de técnicas de observação, contato direto e participação em atividades. É um método onde se exige uma dedicação maior de tempo. Como o objeto de pesquisa deste trabalho trata da atividade do usuário no ambiente da internet, elegeu-se o método da Netnografia, que é uma adaptação dos processos etnográficos de pesquisas para as comunidades virtuais. Esta movimentação em torno da formulação de uma metodologia que atenda aos objetivos e expectativas deste trabalho se fez necessária em vista do aspecto particular da pesquisa, ao analisar etiquetas feitas em torno de imagens e, conseqüentemente, da falta de um percurso metodológico estabelecido que a atenda, fato comum em pesquisas realizadas com adventos e processos recentes no ambiente da internet.

A associação de métodos e técnicas de pesquisa na forma de uma triangulação encontra respaldo nas postulações de Minayo (2000) e um excelente exemplo recente de utilização no trabalho de Araújo (2009). Em linhas gerais, pretendeu-se abordar a atividade do usuário e analisá-la como um processo de indexação, classificando-a através da análise de etiquetas. Essa análise foi realizada de forma estratificada enfocando os diferentes níveis de experiência dos usuários. Em uma visão geral (Figura 1), este trabalho é composto por uma pesquisa documental, como fonte de dados e base para outro tipo de estudo qualitativo, o estudo de caso. Essa pesquisa foi realizada por meio da análise da etiquetagem do usuário segundo os elementos prescritos por Manini (2002, 2009) e Smit (1996) para a indexação de imagens, permitindo caracterizar e classificar as etiquetas de acordo com as categorias elementares apresentadas por Panofsky (1979).



FIGURA 1 - Modelo de aplicação da Netnografia de Kozinets, 2002.



Fonte: Adaptado de Nascimento, 2008

Nesse sentido, as etiquetas foram classificadas como pré-iconográficas, iconográficas ou iconológicas da seguinte forma: a) foi considerada pré-iconográfica a etiqueta que descreve de forma genérica os objetos e ações representados na imagem; b) iconográfica a etiqueta que nomeia ou identifica a imagem, sendo que os elementos identificados estejam presentes na imagem; c) iconológica a etiqueta que estabelece um assunto, identifica um evento, nomeia ou identifica elementos da imagem, sem que estes elementos estejam representados na imagem. Isso significa dizer que estas ações foram feitas pelo usuário levando-se em consideração não aquilo que está representado na imagem, mas a relação destes elementos com outros não exibidos na mesma. Indica que uma imagem foi etiquetada levando-se em conta mais o conhecimento e as relações que o usuário fez dessa imagem com outros conteúdos, imagens ou eventos.

3.1 Seleção e identificação dos usuários

Em uma Pesquisa Qualitativa, a fonte de dados é o ambiente natural e o pesquisador é ativo em todo o processo, direcionando sua pesquisa durante seu



XI Encontro Nacional de Pesquisa em Ciência da Informação
Inovação e inclusão social: questões contemporâneas da informação
Rio de Janeiro, 25 a 28 de outubro de 2010

desenvolvimento. Para a análise dos dados se faz necessário considerar cada significado dos fenômenos fornecidos pelo ambiente em seu determinado contexto. Devido à preocupação com a compreensão dos fenômenos e não com a generalização dos dados, o critério para a definição da amostragem da pesquisa não é numérico, ou seja, não é calculado estatisticamente, mas selecionado dentre um grupo ou corte mais relevante para o estudo.

É necessário realizar um corte sobre o objeto de estudo, que define o campo e a dimensão do trabalho (MANNING, 1979), não de forma aleatória e sim escolhido dentre um grupo que melhor atenda aos objetivos específicos da pesquisa. No caso deste trabalho, o corte será feito em torno da língua portuguesa, nativos do Brasil, sendo que os usuários foram divididos em três estratos de pesquisa, a saber: a) *Treinee*, para usuários iniciantes, que possuam no máximo doze semanas de experiência na utilização do Flickr e no máximo cinquenta itens adicionados; b) *Júnior*, para usuários medianos, que possuam entre doze e vinte e quatro semanas de experiência e tenham adicionado entre cinquenta e cem itens ao sistema; c) *Sênior*, para usuários que possuam mais de vinte e quatro semanas de experiência e mais de cem itens adicionados ao sistema. Foi definido o número de três usuários por estrato, tendo em vista a duração média das entrevistas, verificada em testes prévios, e a profundidade das análises propostas.

Para alcançar os usuários que atendem ao corte e aos estratos de pesquisa foi realizada uma busca por grupos com a palavra “brasileiro”. Feita a busca, o pesquisador tornou-se membro do grupo e passou a observar e participar das atividades do mesmo no intuito de identificar os usuários que se enquadravam nos estratos da pesquisa, convidando-os a participar dela. Quando o usuário se recusou a participar foi selecionado outro usuário até que se atingiu o número estabelecido no estrato de pesquisa.

Selecionados os usuários, foram escolhidas para análise duas fotos de cada um deles, sendo estas escolhidas pelo tempo no sistema: a mais antiga adicionada e a mais recentemente adicionada, critérios estabelecidos previamente à escolha e análise. Para cada foto selecionada foram analisadas duas etiquetas, sendo estas a primeira e a última. Assim, foram selecionados e entrevistados três usuários por estrato, totalizando nove usuários; duas fotos por usuário (a mais antiga e a mais recente), totalizando dezoito



XI Encontro Nacional de Pesquisa em Ciência da Informação
 Inovação e inclusão social: questões contemporâneas da informação
 Rio de Janeiro, 25 a 28 de outubro de 2010

fotos; duas etiquetas por foto (a primeira e a última inseridas), totalizando trinta e quatro etiquetas analisadas (uma das fotos selecionadas não possuía etiquetas).

Desta forma, foram enviados convites para participação na pesquisa apenas para usuários que mantiveram atividade constante até a data da seleção de usuários. Esta seleção foi bastante criteriosa e demorada e ocorreu no período entre outubro de 2009 e abril de 2010. Apesar de exigir muito esforço e dedicação, esta fase foi importante também na própria definição e adequação dos valores definidos para elaboração dos estratos de pesquisa, pois foi preciso observar na prática se os valores estabelecidos correspondiam à atividade de fato dos usuários.

Foram analisados diversos perfis para se obter pelo menos cinco perfis de usuários ativos para cada estrato de pesquisa. Como era de se esperar, nem todos os usuários contatados tiveram interesse em participar da pesquisa, o que fez com que trinta usuários fossem contatados para que se atingissem os nove definidos na seleção da amostra.

Conforme os elementos importantes da Netnografia defendidos por Kozinets (2002), em especial os aspectos éticos, os usuários que participaram da pesquisa serão identificados por uma letra e por um número. A letra representa o estrato em que este usuário está inserido, ou seja, T para usuários *Treinee*, J para usuários Juniores e S para usuários Seniores. Já o número representa a ordem com que aquele usuário foi entrevistado. Visando a resguardar a identidade dos usuários, estes serão apresentados apenas com a nomenclatura estabelecida para identificação e com informações de sua atividade que foram observadas para sua seleção, tais como data de ingresso no Flickr, número de itens postados, título das imagens, etiquetas analisadas e as demais etiquetas atribuídas à imagem (ver Tabela 1).

TABELA 1 Tabela de identificação dos usuários

USUÁRIOS COM POUCA EXPERIÊNCIA – TREINEE				
Identificação do usuário	Data ingresso	Nº itens	Título Fotos/Demais etiquetas	Etiquetas
T1	Jan/2010	32	“01/365” <u>01/365</u> , january, pink, <u>white</u>	Projetc365 Yellow
			“Boa segunda-feira” Flor, Roxo	Abelha segunda-feira
T2	Mar/2010	50	“AOSP – Aaeragis Luteoalba”	Sem etiqueta
			“Cym Amarela 01” “Pombos e Pirâmide”	Cym Amarela
T3	Mar/2010	50		Paris



XI Encontro Nacional de Pesquisa em Ciência da Informação
 Inovação e inclusão social: questões contemporâneas da informação
 Rio de Janeiro, 25 a 28 de outubro de 2010

			<i>museum, Louvre, musée, birds, pyramid, fly</i>	Leo Porto
			<i>"Paulista" avenida paulista, buildings, architecture</i>	São Paulo Léo Porto
USUÁRIOS COM MÉDIA EXPERIÊNCIA - JÚNIOR				
Identificação do usuário	Data ingresso	Nº itens	Título Fotos/Demais etiquetas	Etiquetas
J1	Jan/2010	62	"DSC01056" margarida, flores	Rosa Abelha
			"DSC01046" vermelho	Lírio Flores
J2	Jan/2010	100	"Parque da Cidade - Natal/RN" <u>parque da cidade, Natal, Brasil, marcusu2, Niemeyer, Brazil.</u>	Parque Cidades Brasileiras
			"Arco do sol - Natal/RN" <u>rn, natal-rn, cidade do sol,</u>	Natal Cidades Brasileiras
J3	Jan/2010	52	"Aspone" <i>amateur</i>	Cameraphone K580i
			Sem título <u>cameraphone, wtf, dirty, Centro</u>	K850i Noite
USUÁRIOS MAIS EXPERIENTES - SÊNIOR				
Identificação do usuário	Data ingresso	Nº itens	Título Fotos/Demais etiquetas	Etiquetas
S1	Ago/2008	200	"57.365" Dias	365 Days
			"Raios no Ibura" <u>pernambuco, brasil, tiltshift, nikon</u>	Recife Luciana Santos
S2	Set/2007	519	"Just outside your window" <u>house, konica, minolta, dimage, Z2, brick, tijolo, reflection</u>	Janela Fotógrafo
			"A arte de viver da fé" <u>senhor, bonfim, Bahia, igreja fé, faith, religiao</u>	Wolfgang Salvador
S3	Mar/2008	112	"Corrosivo 8" Colagem	Arte Desenho
			"Oratório" hiena	Oratório Santo

Fonte: Dados de pesquisa

4. EXEMPLO DA ANÁLISE

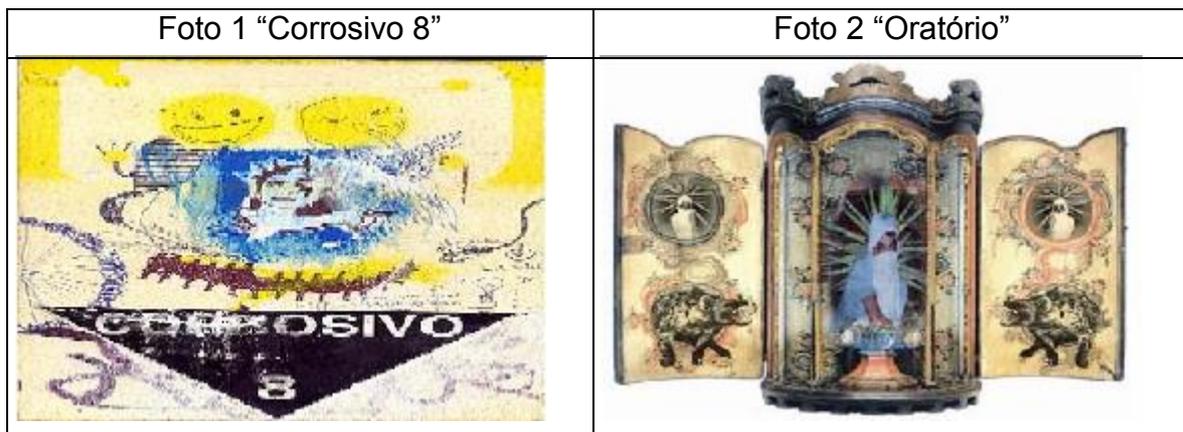
Devido às limitações de tamanho e formato prescritas para este artigo, será apresentada apenas a análise das etiquetas de um usuário para ilustrar como estas foram



XI Encontro Nacional de Pesquisa em Ciência da Informação
 Inovação e inclusão social: questões contemporâneas da informação
Rio de Janeiro, 25 a 28 de outubro de 2010

realizadas. Para tanto, foi selecionada a análise feita em torno do usuário S3, do qual foram selecionadas as imagens (Quadro 3) “Corrosivo 8” (Foto 1) e “Oratório” (Foto 2) e as etiquetas “Arte”, “Desenho”, “Oratório” e “Santo”, respectivamente, em par. As análises das etiquetas de S3 demandaram muita observação e pesquisa tendo vista tratarem da difícil indexação de obras de arte. As etiquetas “Arte” e “Desenho” estão associadas à “Foto 1” e as etiquetas “Oratório” e “Santo” à “Foto 2”. O caso da etiqueta “Arte” foi o mais difícil de estabelecer a classificação, pois, ao mesmo tempo em que descreve o referente de forma genérica, estabelece um assunto para a imagem. Analisando as comunicações do usuário e, principalmente, as outras etiquetas adicionadas à imagem, concluiu-se que a intenção do usuário foi estabelecer um assunto/categoria para a imagem, classificando a etiqueta como iconográfica.

QUADRO 3 Imagens coletadas de S3



Fonte: Dados de pesquisa.

A etiqueta “Desenho” foi considerada iconográfica por estabelecer um assunto secundário, observável no referente da imagem, que se refere à técnica utilizada na elaboração da obra de arte. Representa uma etiquetagem mais elaborada em relação à primeira etiqueta, pois acrescenta informação importante para a compreensão da imagem; indica a extração do conteúdo informacional da imagem. Da imagem “Oratório” (Foto 2) foram extraídas as etiquetas “oratório” e “santo”. A etiqueta “oratório” foi considerada iconográfica pois identifica um elemento do referente. Novamente fica evidenciada a atenção de S3 à extração do conteúdo informacional da imagem. A etiqueta “santo” foi considerada pré-iconográfica, pois descreve genericamente o referente da



imagem. É interessante observar a movimentação das etiquetas que culminam por representar uma boa extração do conteúdo informacional da imagem tanto específica quanto genericamente.

Ao contrário do que se esperava, a etiquetagem da “Foto 1” é mais elaborada que a da “Foto 2”, sendo que isso se justifica pelo fato da “Foto 1” ser mais complexa para se etiquetar, para se extrair e descrever o conteúdo informacional do seu referente, visto que é uma obra de arte. De modo geral, mesmo no que se refere à etiquetagem da primeira imagem de S3, a etiquetagem do usuário apresentou boas e variadas estratégias onde se destacam a extração genérica e específica do referente da imagem, o estabelecimento de assunto e da categoria.

Apesar de não utilizar a estratégia de adicionar a identificação do autor, comum nos outros dois usuários experientes da pesquisa e também nos usuários pró² do Flickr, S3 apresentou uma etiquetagem bem elaborada, levando-se em conta o maior grau de dificuldade da etiquetagem de suas imagens.

5. PRINCIPAIS RESULTADOS E CONTRIBUIÇÕES

De modo geral, a realização das análises permitiu o levantamento de um conjunto de estratégias (Tabela 5), bem como surgiram questões a serem esclarecidas nas entrevistas. Em outras palavras, as entrevistas permitiram observar as razões dos usuários na realização de suas etiquetas e tornar mais precisas e válidas as estratégias levantadas. A seleção da tipologia das estratégias de indexação de imagens foi realizado com base nos trabalhos de Smit (1996), Manini (2002).

² São chamados de usuários pró os usuários que remuneram o Flickr em troca de privilégios na utilização do site.



XI Encontro Nacional de Pesquisa em Ciência da Informação
 Inovação e inclusão social: questões contemporâneas da informação
 Rio de Janeiro, 25 a 28 de outubro de 2010

TABELA 5 Estratégias utilizadas pelos usuários

Estratégias	Usuários								
	T1	T2	T3	J1	J2	J3	S1	S2	S3
Extração genérica do conteúdo informacional	X	X		X		X		X	X
Extração específica do conteúdo informacional	X	X		X				X	X
Elaboração de assunto para imagem	X		X	X	X		X	X	X
Sugestões e normas estipuladas por grupos				X					
Utilização de etiquetas populistas				X	X			X	
Utilização de etiquetas egoístas	X				X	X	X		
Adição de informação da dimensão expressiva da imagem			X			X	X	X	X
Adição de informação do contexto do referente			X		X		X	X	

Fonte: Dados de pesquisa.

Na Tabela 5 estão representadas as estratégias levantadas e a indicação de quais estratégias cada usuário utilizou. Desta forma foi possível visualizar melhor as relações entre as estratégias levantadas, os usuários que as realizaram e os estratos. Ao contrário do que se esperava, a extração do conteúdo informacional da imagem, seja ela genérica ou específica, não foi utilizada por todos os usuários, sendo que nenhuma das estratégias levantadas foi utilizada por todos. Como estavam sendo analisadas pessoas sem qualquer tipo de conhecimento estruturado sobre métodos e processos de indexação e como a elaboração de assuntos para imagens é um procedimento que demanda muito esforço do usuário, tinha-se a expectativa de que a ação mais comum dos usuários fosse a extração do conteúdo informacional da imagem. Entretanto, a estratégia mais utilizada pelos usuários para etiquetagem de suas imagens é a elaboração de um assunto para a mesma, sendo que apenas T2 e J3 não a utilizaram. Percebeu-se, de forma clara, o aumento do número de estratégias utilizadas pelos usuários sendo que os usuários *Treinee* apresentaram o menor número de estratégias e os usuários Sênior o maior. Isso representa a evolução da etiquetagem dos usuários que, à medida que vão se relacionando com o sistema de etiquetagem, vão apreendendo e elaborando suas próprias estratégias. Apesar do estrato de usuários Sênior apresentar o maior número de estratégias utilizadas, o estrato de usuários Júnior foi o único que se utilizou de todas as estratégias levantadas nas análises. Na reflexão sobre este fato e na observação das entrevistas e comunicações dos usuários, conclui-se que isso se deve ao fato de, no



processo de amadurecimento de sua etiquetagem, os usuários Júnior experimentam uma variedade maior de estratégias.

De qualquer forma, causa grande surpresa o fato dos usuários Júnior terem rapidamente desenvolvido padrões de etiquetagem já repetidos na imagem mais recente e na mais antiga. Foi observado também o fato da elaboração de assunto para a imagem ter sido a estratégia mais utilizada entre todos os usuários – sendo utilizada por sete deles – e a influência dos grupos ter sido a estratégia menos utilizada, sendo comprovada sua efetiva utilização por apenas um usuário. Isso posto, deve-se ressaltar que, comparando apenas as primeiras etiquetagens dos usuários, a estratégia mais utilizada foi a extração do conteúdo informacional das imagens, ressaltando o maior esforço exigido para elaborar um assunto para uma imagem.

Contrariando a expectativa criada durante a execução do trabalho, de que a utilização da estratégia de inserir etiquetas populistas seria comum, esta foi utilizada por apenas três usuários, sendo que a utilização de etiquetas egoístas, sobre a qual havia expectativa de pouca utilização, foi utilizada por quatro usuários. Comparando diretamente a atividade dos usuários, temos a etiquetagem de S2, considerada a mais exaustiva, sendo a que contou com o maior número de estratégias – seis, no total – e a etiquetagem de T2, sendo a que menor número de estratégias apresentou – utilizando-se de apenas duas.

Concluindo a comparação entre os usuários, a etiquetagem de S3, considerada a mais precisa, exibiu poucas estratégias – quatro, ao todo – realizadas de maneira bem elaborada, contribuindo bastante para a precisão de uma busca, enquanto a etiquetagem de J3, que também contou com um número pequeno de estratégias – três, no total – foi considerada a de menor qualidade, sendo disseminada quase que exclusivamente para os grupos onde as fotos participam.

6. CONSIDERAÇÕES FINAIS E TRABALHOS FUTUROS

Em relação ao objetivo geral estabelecido na pesquisa – observar as estratégias da utilização de etiquetas pelos usuários do Flickr, em língua portuguesa, e propor uma



XI Encontro Nacional de Pesquisa em Ciência da Informação
Inovação e inclusão social: questões contemporâneas da informação
Rio de Janeiro, 25 a 28 de outubro de 2010

forma de análise e classificação das mesmas – foi considerado que a adaptação envolvendo os métodos de avaliação da indexação de imagens de Manini (2002) com a adaptação das categorias de Panofsky (1979) viabilizou de forma mais precisa o estabelecimento de uma classificação precisa, que traduz as movimentações e intenções dos usuários ao etiquetar imagens.

No que diz respeito aos objetivos específicos – propor uma forma de análise e classificação da etiquetagem do usuário a partir do modelo de Manini (2002) e das categorias de Panofsky (1979) e conhecer os procedimentos e estratégias que envolvem o processo de etiquetagem de imagens no Flickr – considera-se que o método de análise proposto permitiu de forma mais consistente estabelecer uma classificação mais precisa e rica para atividade do usuário, sendo possível perceber a relação entre o referente da imagem e a etiqueta, clareando, assim, as estratégias utilizadas pelos usuários na etiquetagem de imagens. Encerrando a retomada dos objetivos específicos – a) observar as diferenças e similaridades da etiquetagem de usuários com diferentes níveis de experiência e participação no Flickr; e b) observar diferenças entre a etiquetagem mais antiga e mais recente de um mesmo usuário – foram feitas comparações entre as etiquetagens de um mesmo usuário, bem como etiquetagens de usuários de estratos diferentes.

De modo geral, percebe-se como importante a movimentação de tratar da temática da Folksonomia dentro do contexto dos métodos e processos de indexação, conforme foi realizado neste trabalho. Observou-se que a maior parte dos trabalhos pesquisados na revisão bibliográfica mostrou este processo sendo mais abordado como um processo de comunicação pela área de Comunicação Social e, principalmente, como um instrumento de navegação pela Computação, onde a maioria das abordagens enfoca a Folksonomia do ponto de vista técnico e estatístico. Essa importância é corroborada por outros trabalhos, como o de Nascimento (2008), que contribuem para a evolução da Folksonomia, abordando de forma interdisciplinar, construindo percursos metodológicos que viabilizem o desenvolvimento da Folksonomia enquanto um processo social e humano. Por fim, no que diz respeito à análise deste trabalho, observa-se que o percurso metodológico percorrido por este culmina para contribuir para a precisão e classificação



XI Encontro Nacional de Pesquisa em Ciência da Informação
Inovação e inclusão social: questões contemporâneas da informação
Rio de Janeiro, 25 a 28 de outubro de 2010

da atividade dos usuários em ambientes que envolvem a Folksonomia realizada a partir de imagens.

Em relação a trabalhos futuros levantadas por esta pesquisa, ressalta-se a possibilidade de ampliação da precisão das análises aqui contidas com utilização de outras categorias, formulação de outros estratos de pesquisa e abordagem de comunidades mais específicas, bem como a utilização de mais etiquetas ou todas as etiquetas de uma imagem. Finalizando, pretende-se que este trabalho abra caminho para que, no futuro, seja possível traduzir a relação das etiquetas com os referentes das imagens em artefatos de organização da informação como as ontologias, no sentido de possibilitar ambientes de recuperação baseados em Folksonomia mais precisos e customizados.

ABSTRACT: The Folksonomy, by use of tags, represents a change in the processes of organization and information retrieval on the web. If before all content was stored heuristics indexed by search engines, this process now has the participation of those who provide the documents and make use of them: the users. The direct participation of the user in a process that once it was only the specialist staff raised the discussion about the effects of Folksonomy in information organizing on the Internet making clear the need to study how users carry out this process. This incorporation by itself brings great challenges, however, when performed in environments where the user takes their tags from images that challenge is amplified, because the greater complexity of indexing a polysemic feature as an image. Therefore, this study focused on the indexing strategies used by users in image tagging on Flickr. Thus, this study reviewed the literature on the thematic Folksonomy, indexing and indexing images, in order to draft a form of analysis and classification of user activity. This work made adjustments to the method of image reading proposed by Manini (2002) and the image interpretation levels of Panofsky (1979) in order to propose a form of analysis and interpretation of the user tagging. To accomplish this task Netnography was used to compose the methodological approach allowing to directly hit the users through interviews to check the tagging strategies perceived in the tag analysis, also allowing evaluate the proposed method of analysis. This approach reach a form of analysis and classification of user activity more accurately and consistently, and through her is possible to see the image tagging strategies used by Flickr users.

Keywords: Folksonomy. Indexing. Image Indexing. Image Tagging. Netnography.



REFERÊNCIAS:

ARAÚJO, Ronaldo F. **Apropriações de Bruno Latour pela Ciência da Informação no Brasil**: descrição, explicação e interpretação. Escola de Ciência da Informação/UFMG. Belo Horizonte, 2009.

AZEVEDO, Alexander W. A construção da Ciência da Informação na pós-modernidade: dialética histórica, in **Revista Digital de Biblioteconomia e Ciência da Informação**. Campinas, v. 6, n. 2, p. 71-82, jan./jun. 2009.

BARTHES, Roland. **A câmara clara**: nota sobre a fotografia. Rio de Janeiro: Nova Fronteira, 1984.

BLATTMANN, U.; SILVA, F. C. C. Colaboração e interação na Web 2.0 e Biblioteca 2.0. **Revista ACB**: Biblioteconomia em Santa Catarina, Florianópolis, v. 12, n. 2, p. 191-215, jul./dez., 2007.

DUBOIS, P. **O ato fotográfico e outros ensaios**. Campinas: Papirus, 1994.

FLICKR. Disponível em <www.flickr.com>, acesso em 10 ago. 2010

FUJITA, M. S. L.; SILVA, M. R. A prática da indexação: análise da evolução de tendências teóricas e metodológicas. **Transinformação**, Campinas, v. 2, n. 16, p. 133-161, mai./ago. 2004.

GODOY, Arilda S. Introdução à pesquisa qualitativa e suas possibilidades. **Revista de Administração de Empresas**, v. 35, n. 2, Mar./Abr. 1995a, p. 57-63.

GODOY, Arilda S. Pesquisa qualitativa – tipos fundamentais, in **Revista de Administração de Empresas**, v. 35, n. 3, Mai./Jun. 1995b, p. 20-29.

KOZINETS, Robert V. The field behind the screen: using netnography for marketing research in online communities. **Journal of Marketing Research**, Chicago, v.39, n.1, fev. 2002

LANCASTER, F. W. **Indexação e resumos**: teoria e prática. Trad. de Antonio Agenor Briquet de Lemos. Brasília: Briquet de Lemos/Livros, 1993.

MANNING, Peter K. Metaphors of the field: varieties of organizational discourse, in **Administrative Science Quarterly**, v. 24, n. 4, December 1979, p. 660-671.

MANINI, Miriam P. **Análise documentária de fotografias**: um referencial de leitura de imagens fotográficas para fins documentários. Tese (doutorado) – Escola de Comunicações e Artes, USP São Paulo, 2002.



XI Encontro Nacional de Pesquisa em Ciência da Informação
Inovação e inclusão social: questões contemporâneas da informação
Rio de Janeiro, 25 a 28 de outubro de 2010

MANINI, Miriam P. Análise documentária de imagens: a fotografia e seus textos. Campinas, 1997. In **III Encontro Nacional de Pesquisa em Ciência da Informação**, 1997, Rio de Janeiro.

MANINI, Miriam P. Aspectos informacionais do tratamento de documentos fotográficos tradicionais e digitais. In **X Encontro Nacional de Pesquisa em Ciência da Informação**, João Pessoa. 2009.

MINAYO, M. C. de S. **O desafio do conhecimento: pesquisa qualitativa em saúde**. 7ª ed. São Paulo: Hucitec, 2000.

MOREIRA, Manoel Palhares. **Ambiente para geração e manutenção semi-automática de tesouros**. Tese (Doutorado em Ciência da Informação) - Escola de Ciência da Informação, Universidade Federal de Minas Gerais, Belo Horizonte, 2005.

MOURA, Maria A. Informação, ferramentas ontológicas e redes sociais ADHOC: a interoperabilidade na construção de tesouros e ontologias. **Informação e Sociedade**, João Pessoa, v. 19, n. 1, p. 59-73, jan./abr. 2009.

NASCIMENTO, Geysa Flávia C. de L. **Folksonomia como estratégia de indexação dos bibliotecários no Del.icio.us**. João Pessoa: PPGCI, 2008.

PANOFSKY, E. **Significado nas artes visuais**. 2ª ed. São Paulo: Perspectiva, 1979.

PINHEIRO, L. V. R. **A Ciência da Informação entre sombra e luz: domínio epistemológico e campo interdisciplinar**. 1997. Tese (Doutorado em Comunicação e Cultura) - UFRJ/Escola de Comunicação, 1997.

ROBREDO, J. **Documentação de hoje e amanhã: uma abordagem informatizada de biblioteconomia e dos sistemas de informação**. 2ªed. rev. ampl. Brasília: Ed. do autor, 2005.

SHATFORD LAYNE, Sara. Some issues in the indexing of images. **Journal of the American Society for Information Science**, v. 45, n. 8, p. 583-588, 1994.

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

SMIT, Johanna, A representação da imagem. **INFORMARE** Rio de Janeiro, v. 2, n. 2, p. 28-36, jul./dez. 1996.