

Análise e indexação de imagens na rede Flickr

Rafael Alves de Oliveira

Doutorando; Universidade Federal da Paraíba;
raffael@gmail.com

Luciane Paula Vital

Doutoranda; Universidade Federal de Santa Catarina;
lucianepv@yahoo.com.br

Resumo: O artigo apresenta um breve panorama sobre a disseminação e democratização da internet colaborativa, e de como esse processo popularizou o uso da imagem em suporte digital. Faz uma breve análise sobre a rede social Flickr, identificando seus objetivos, suas ferramentas, e alguns dos estudos relativos à sua usabilidade no âmbito da Ciência da Informação. Delimita a abordagem para as imagens armazenadas na rede, propondo um estudo qualitativo em torno da indexação livre realizada por seus usuários. Aplica a metodologia estabelecida em estudo anterior como a mais apropriada para indexação de imagens fotográficas em ambiente web, comparando os resultados obtidos com a indexação livre realizada pelos usuários.

Palavras-chave: Imagem. Indexação de imagens. Flickr. Redes sociais. Indexação social.

1 Introdução

A década de 2001 a 2010 foi testemunha do surgimento de um novo conceito de uso da internet. A possibilidade de publicação e tratamento livre de informações por parte de seus usuários tornou-se a característica fundamental deste novo modelo, que, embora não seja um consenso, os estudiosos nomearam de Web 2.0. Em alguma medida, pode-se dizer que a criação das redes sociais foi uma consequência natural desse fato, e que a aceitação dessas por parte de uma grande parcela dos usuários da internet é um reflexo de como este modelo se estabeleceu, e tende a se aprimorar.

Essas novas ferramentas disseminaram o conceito e democratizaram o uso de um suporte informacional bastante peculiar: a imagem, que migrou dos

álbuns de fotografias e das molduras de parede para o suporte digital, podendo agora ser acessada a qualquer instante, de qualquer lugar do planeta.

[...] com as novas tecnologias de comunicação e informação uma nova fase inicia-se, e nesta, as imagens são disponibilizadas livremente pelos usuários na internet. Essa liberdade de publicação, naturalmente, impeliu os métodos de tratamento e recuperação para essa mesma ideia. Agora o usuário não só pode publicar as imagens como também categorizá-las. Esse *boom* de imagens disponibilizadas iniciou-se com o surgimento das redes sociais. A sociedade aderiu a esta nova forma de interagir, e as redes cresceram e popularizaram-se significativamente ao longo de poucos anos. À medida que as ferramentas destas redes se desenvolviam com o avanço dos seus servidores, as redes podiam armazenar cada vez mais informações. A publicação livre de imagens foi apenas um passo natural nesse desenvolvimento. (OLIVEIRA, 2011, p. 2).

Partindo desse pressuposto, fica evidente que uma problemática tende a surgir com esse avanço, e diz respeito especificamente à maneira como estas imagens serão tratadas e recuperadas. Na literatura científica de Ciência da Informação ainda não existe um acordo sobre como as imagens devem ser tratadas, pois diversos fatores implicam nessa análise, incluindo o tipo de imagem (pinturas, fotografias, etc.) e o suporte no qual se encontra. Entretanto, alguns modelos bem construídos e eficientes podem ser encontrados, como o de Manini (2002) e Rodrigues (2007).

No trabalho intitulado *Metodologias para indexação de imagens fotográficas em ambiente web* (OLIVEIRA, 2011) são apresentados e analisados os elementos propostos e o comportamento de três das metodologias para indexação de imagens mais citadas na literatura científica da área da Ciência da Informação, visando identificar aquela que poderia ser considerada mais adequada à aplicação em imagens fotográficas coletadas em ambiente web. Este estudo analisa, dentre outras, a metodologia proposta por Rodrigues (2007), na qual o autor revela que uma imagem não apenas mostra, mas também representa algo, que pode não necessariamente ter uma relação direta aos objetos apresentados. Devido a isso, entende-se que uma imagem terá dois níveis ou sentidos principais: o denotativo, que se refere “[...] àquilo que a imagem representa com ‘certa precisão’, no seu sentido real” (RODRIGUES, 2007, p.69), e o conotativo, referente àquilo “[...] que a imagem pode ‘interpretar’ em

um determinado contexto, em um sentido figurado e simbólico”.(RODRIGUES, 2007, p. 69). A partir dessa constatação, Rodrigues (2007) estabelece que devem ser considerados os seguintes elementos para uma análise apropriada:

- a) descrição física, ou o “[...] formato e tamanho da imagem fotográfica, tipo de suporte, autor, transformações ocorridas a partir do original etc”. (RODRIGUES, 2007, p.75);igo de periódico NBR 6022/03;
- b) composição, ao considerar “[...] tipo de luz, nível de nitidez dos assuntos, ponto de vista do fotógrafo, profundidade de campo e hierarquia das figuras, enquadramento etc”.(RODRIGUES, 2007, p. 75);
- c) contexto arquivístico, ou os locais e fatos, históricos ou não, correspondentes àquela fotografia;
- d) sentidos denotativos e conotativos, correspondendo ao que imagem mostra e ao que ela representa, respectivamente;
- e) tematização, ao atribuir a imagem a um contexto diferente do que é mostrado, mas que possui relação direta aos elementos denotativos apresentados.

Após traçar um quadro comparativo observou-se que, para uma análise de imagens coletadas em ambiente web, a metodologia proposta por Rodrigues (2007) pode ser considerada a mais apropriada, já que enfatiza os elementos conotativos identificados numa imagem, considerando-se que uma imagem coletada na internet tem grande possibilidade de não apresentar informações técnicas referentes a esta.

A partir das constatações obtidas neste estudo, buscaremos obter um breve panorama sobre a indexação realizada livremente pelos usuários de redes sociais por meio da aplicação prática da metodologia de Rodrigues (2007) e da interpretação dos dados obtidos. O banco de imagens escolhido para a pesquisa foi a rede *Flickr*, por se tratar de uma rede social amplamente utilizada e ter seu foco no compartilhamento de recursos visuais. Desde já delimitaremos o objeto de estudo apenas às imagens armazenadas pela rede *Flickr*, embora a rede atualmente permita também o compartilhamento de vídeos.

2 Flickr: uma breve análise

O *Flickr* é uma rede de compartilhamento de imagens hospedada na internet. O *Flickr*, segundo os próprios idealizadores, propõe novas formas para organizar fotografias, que poderão ser acessadas livremente por qualquer pessoa.

Parte da solução é tornar colaborativo o processo de organizar fotos [...] No *Flickr*, é possível permitir que seus amigos, família e outros contatos organizem suas coisas - não apenas adicionem comentários, mas também notas e tags. [...] E, à medida que essas informações crescem como metadados, você poderá encontrar as coisas facilmente mais tarde, uma vez que toda essa informação pode ser buscada. (*FLICKR*, 2014c, documento eletrônico sem paginação).

Embora seja considerado uma rede social, o *Flickr* permite que mesmo aqueles que não possuem cadastro na base tenham acesso às imagens disponibilizadas. No entanto, o usuário tem a liberdade de restringir o acesso a algumas de suas fotografias apenas àqueles usuários que lhes forem mais convenientes. O *Flickr* é um claro exemplo de ambiente que permite a utilização e o emprego daquilo que a literatura científica nomeou de *folksonomias*,

[...] que podem ser definidas como sistemas orgânicos baseados na atribuição livre e pessoal de [termos] à informações ou objetos visando à organização e recuperação. O neologismo folksonomia – formado pelas palavras, em inglês, *folks* (pessoas) e *taxonomy* (taxonomia) – foi cunhado [...] como forma de expressar contraposição às classificações do conhecimento tradicionais, elaboradas por especialistas e construídas baseando-se em arranjos hierárquicos. (GUEDES; DIAS, 2010, p.48).

No *Flickr* podemos identificar dois tipos de usuários: aqueles que possuem cadastro na base e, portanto, podem publicar fotos e se utilizar de todas as ferramentas que o *Flickr* oferece; e aqueles que não possuem qualquer vínculo com o sistema, apenas o utilizam para fazer pesquisas no conjunto de fotografias disponibilizadas para consulta pública. Ao primeiro tipo, chamaremos de usuário regular; ao segundo, usaremos a definição usuário-consultor.

Apesar de ser descrito como um ambiente que utiliza *folksonomias* (HIDDERLEY; RAFFERTY, 2007), na prática o *Flickr* não permite que a indexação seja feita livremente pelo usuário-consultor, como acontece na maioria desses ambientes. Apenas o usuário regular pode fazê-lo e, no máximo, outros usuários regulares que receberam do proprietário das imagens uma permissão para tanto. Esses poderão indexar as imagens utilizando-se do recurso *detagging* (do inglês, *etiquetagem*), que consiste na atribuição de termos que representem os conteúdos informacionais disponibilizados na web.

As tags são como palavras-chave ou títulos que você adiciona a uma foto para facilitar encontrá-la posteriormente. Você pode criar uma tag para uma foto com frases como "Catarina caminhada trilha montanha Yosemite." Posteriormente, se você procurar por imagens de Catarina, bastará clicar nessa tag e obter todas as fotos marcadas dessa maneira. (FLICKR, 2014a, documento eletrônico sem paginação).

As *tags* atribuídas a determinado conteúdo servirão como uma referência conceitual aos usuários que terão acesso posterior àquela informação. No caso do *Flickr*, o usuário regular disponibiliza a imagem na base e atribui, a partir da sua interpretação, termos que representem a imagem. Em seguida, o usuário-consultor, ou outros usuários regulares, poderão pesquisar as imagens disponibilizadas através de palavras-chave e a base naturalmente utilizará as *tags* atribuídas a cada imagem como ferramenta de busca. Alguns autores preferem o termo *indexação social* para identificar esse processo.

Cada imagem pode ter um máximo de 75 tags atribuídas a ela. No entanto, o usuário regular tem, além das *tags*, outra forma de descrever uma imagem: o texto. Este pode conter, como sugere o próprio *Flickr* (2014a), a história da fotografia e/ou notas explicativas sobre esta. É importante ressaltar que o *Flickr* não se utiliza exclusivamente das *tags* atribuídas às imagens para fazer a recuperação, mas também de termos livres contidos nas descrições. Na hora da busca, o usuário pode escolher entre pesquisar apenas nas *tags*, ou por meio de termos livres, incluindo assim os textos.

Tags, ao contrário de termos de indexação (descritores); geralmente utilizados em unidades de informação, não possuem um mediador

que as controle. O *Flickr* diferencia-se de outros ambientes que utilizam *folksonomias* por estabelecer que apenas o usuário regular possua tal liberdade (embora alguns ambientes que empregam *folksonomias* também utilizem esse recurso), provavelmente como uma forma de minimizar tags ‘inúteis’. O *Flickr* também propõe uma lista de “*hot tags*”, ou “etiquetas quentes”, que consistem numa lista das *tags* mais utilizadas pelos usuários regulares. Essas são, possivelmente, estratégias desenvolvidas pelos administradores da base para aprimorar as formas de indexação, ou *tagging*, das imagens.

Problemas no processo de indexação e as consequências na recuperação das imagens foram identificados por alguns autores. Rafferty e Hilderley (2007), numa pesquisa controlada, mostram como os índices de revocação podem variar imensamente de acordo com o termo atribuído, quando explicam que

A [...] tag [...] “casamento” recuperou 795.280 imagens, quando utilizada no *Flickr* em 14 de fevereiro de 2006, sugerindo que este seria um exemplo de uma tag talvez ampla demais. A [...] tag “Ijsselmeer” recuperou apenas uma imagem, e é um exemplo de uma tag que talvez seja específica demais para o propósito de pesquisa pública. (RAFFERTY; HIDERLEY, 2007, p. 403, tradução nossa)¹.

No exemplo mostrado, o termo *WEDDING* é considerado muito geral pelos autores, recuperando um número de imagens considerado absurdo. Já o termo *IJSELMEER*, segundo os autores, é específico demais para uma imagem disponibilizada numa base para uso público, recuperando apenas uma imagem. Analisar os resultados da recuperação, no entanto, não é suficiente para deduzir que a causa do problema seja a forma como os usuários indexam as imagens no *Flickr*. Percebemos, inclusive, que há um interesse por parte dos próprios administradores em educar os usuários a controlar os termos utilizados. É possível que, apesar de os índices de recuperação no *Flickr* não terem sido considerados satisfatórios pelos autores, os termos utilizados pelos usuários na atribuição de tags representem de forma efetiva os elementos apresentados na imagem.

Partindo desse impasse, coletamos algumas imagens no *Flickr* para analisá-las com base na metodologia proposta por Rodrigues (2007), anteriormente considerada a mais adequada para o ambiente web dentre as metodologias para representação imagética encontradas, e comparar os resultados obtidos com as tags atribuídas pelos usuários. Assim, é possível perceber se a indexação de imagens que está sendo realizada por meio das ferramentas da Web 2.0 condiz qualitativamente com aquela indexação proposta pelos estudiosos da área.

3 Procedimentos metodológicos

Entendemos este estudo como sendo de caráter qualitativo, já que não buscamos uma coleta precisa de dados estatísticos, mas uma análise interpretativa, baseada essencialmente no que a literatura científica diz a respeito do tema. Através deste estudo, buscamos compreender um pouco mais sobre as modalidades de indexação imagética.

O *Flickr* possui uma quantidade de fotos armazenadas que cresce de forma rápida e significativa. Dessa forma, percebemos de imediato que o *Flickr* constitui um universo de amplitude *infinita*.

Sendo esta pesquisa de caráter qualitativo, recorremos a Gil (2008) para obter alguns esclarecimentos a esse respeito. O autor explica, dentre outras formas de coletar amostras, que a amostragem por acessibilidade ou por conveniência

[...]constitui o menos rigoroso de todos os tipos de amostragem. Por isso mesmo é destituída de qualquer rigor estatístico. O pesquisador seleciona os elementos a que tem acesso, admitindo que estes possam, de alguma forma, representar o universo. Aplica-se este tipo de amostragem em estudos exploratórios ou **qualitativos**, onde não é requerido elevado nível de precisão. (GIL, 2008, p. 94, grifo nosso).

Dessa forma, buscamos tornar possível a análise da metodologia em ambiente web, mesmo admitindo a infinidade de imagens que podem ser encontradas nesse contexto. Devemos destacar, antes de explicar os

procedimentos, que, visando delimitar o campo de estudo, fica estabelecido que apenas as *tags* atribuídas na língua portuguesa serão consideradas na análise. *Tags* em quaisquer outras línguas serão desconsideradas, assim como aquelas que apresentem composição confusa devido a possíveis erros de grafia.

Em vista disso, foram realizados os seguintes procedimentos para a coleta das imagens utilizadas neste estudo:

- a) optamos por fazer a pesquisa como *usuário consultor*, uma vez que, desta forma, o cadastro na base não interfere significativamente na busca;
- b) na página inicial do *Flickr*, utilizamos os termos TAGS (para realizar a pesquisa inicial de imagens, por se tratar de um termo bastante amplo, de certa forma desprovido de qualquer conceituação, e, como já visto, muito utilizado na base); e BRASIL (para delimitar ao máximo as imagens recuperadas à imagens com *tags* em língua portuguesa). Dessa forma, as imagens foram recuperadas de forma aleatória no que diz respeito à sua temática;
- c) após a pesquisa inicial, a base dá a possibilidade de escolha entre a pesquisa em todos os textos atribuídos a uma imagem ou apenas em suas *tags*. Visando uma maior generalidade de imagens recuperadas, optamos por permanecer na busca realizada em todos os textos;
- d) tendo em vista uma maior aleatoriedade na coleta, foi selecionada uma imagem de cada página dos resultados de busca, sendo analisadas quatro imagens ao final;
- e) buscando uma conceituação diferenciada e equilibrada nas imagens coletadas, duas dessas foram selecionadas por conterem seres humanos, e as duas restantes por não conterem humanos representados.

Coletados os dados, aplicamos a cada imagem a metodologia proposta por Rodrigues (2007) e, em seguida, comparamos os resultados com as *tags* atribuídas, visando analisar qualitativamente a indexação realizada pelos usuários.

4 Resultados obtidos

Utilizando os procedimentos metodológicos estabelecidos para a coleta e análise das imagens no *Flickr*, obtivemos as seguintes imagens:

Figura 1 – Praia de Jurerê – Florianópolis, SC



Fonte: *Flickr*(2014b).

Na área reservada para o texto descritivo da imagem, o autor indica que esta imagem refere-se à Praia de Jurerê, localizada em Florianópolis, estado de Santa Catarina. As tags atribuídas a esta imagem foram, em ordem de aparição:

Quadro 1 – Tags atribuídas à Figura 1 por usuário da rede *Flickr*

Rainbow	top20travel
Arco-íris	A Plus Photo
Jurerê	The World Throgh a Lens
Florianópolis	Platinum Photo
Santa Catarina	Geo-tagged
Brasil	100v+10f
Brazil	Novas Jurerê

Fonte: elaborado pelos autores

Aplicando a metodologia de Rodrigues (2007), chegamos à seguinte análise:

Quadro 2 – Aplicação da metodologia de Rodrigues (2007) para a Figura 1

DESCRIÇÃO FÍSICA	COMPOSIÇÃO	CONTEXTO ARQUIVÍSTICO
---	---	Jurerê Florianópolis Santa Catarina
SENTIDO DENOTATIVO	SENTIDOS CONOTATIVOS	TEMATIZAÇÃO
Praia Arco-íris Chuva	Tranquilidade	Meteorologia

Fonte: elaborado pelos autores.

Através da metodologia de Rodrigues (2007), atribuímos oito termos que poderiam recuperar essa imagem. No Contexto Arquivístico, os termos JURERÊ, FLORIANÓPOLIS e SANTA CATARINA referem-se à procedência da imagem, ou seja, ao local apresentado na fotografia, e, conseqüentemente, onde esta foi produzida. No Sentido Denotativo, os termos PRAIA, ARCO-ÍRIS e CHUVA representam os objetos identificados pelos autores, enquanto nos Sentidos Conotativos o termo TRANQUILIDADE expressa a sensação transmitida aos autores no momento da observação. Entretanto, por se tratar de um elemento essencialmente subjetivo, outros termos podem ser atribuídos por observadores diferentes. Por fim, no campo Tematização, o termo METEOROLOGIA reproduz um enquadramento temático feito pelos autores diante dos elementos objetivos apresentados. Apesar desse enquadramento temático ser subjetivo, é mais fácil apontar os elementos que levaram a tal determinação, que, no caso da fotografia em questão, os elementos que levaram à atribuição das *tags* PRAIA e CHUVA. Entretanto, da mesma forma que os Sentidos Conotativos, a Tematização pode variar de um observador para outro.

Como ficou claro nas *tags* associadas a essa imagem, alguns termos não correspondem à língua portuguesa, enquanto outros apresentam composição confusa e, portanto, deverão ser desconsiderados. Traçando um comparativo entre as *tags* válidas e os termos adquiridos por meio da metodologia, temos:

Quadro 3 – Comparação entre as *tags* consideradas válidas e os resultados obtidos na análise

TAGS	TERMOS
Arco-íris	Jurerê
Jurerê	Florianópolis
Florianópolis	Santa Catarina
Santa Catarina	Praia
Brasil	Arco-íris
Novas Jurerê	Chuva
---	Tranquilidade
---	Meteorologia

Fonte: elaborado pelos autores

Dentre as *tags* consideradas válidas para esta pesquisa, apenas quatro coincidem com os termos adquiridos através da metodologia. Termos considerados indispensáveis na indexação desta imagem, como PRAIA, não aparecem nas *tags*. A *tag* BRASIL, por sua vez, apesar de corresponder a um dos critérios de busca para a presente pesquisa e ser relevante para situar a localização do ambiente apresentado na ilustração, é excessivamente genérica para descrever, por si só, essa imagem, enquanto NOVAS JURERÊ parece ser uma indicação individual de localização para o próprio autor da fotografia, não podendo ser considerado relevante para descrevê-la.

Avançando para a próxima página de busca, coletamos a seguinte imagem:

Figura 2 – Maragogi, AL



Fonte: *Flickr* (2014b).

Nesta imagem é possível identificar um grupo de jovens observando o mar. Segundo as informações do autor da fotografia, ela foi tirada na Praia de Peroba, em Maragogi, Estado de Alagoas. Diversas *tags* referenciam essa imagem, como as listadas a seguir:

Quadro 4 – Tags atribuídas à Figura 2 por usuário da rede Flickr

Colors	Miramar	Caribe
Color	Boa viagem	Brasileiro
Maragogi	Contraste	O Caribe brasileiro
Alagoas	Cores	Viagem
Peroba	Fortes	Dicas
Praia	Vibrantes	Ficar
Beach	Viva	Salinas
Red	Paraíso	Férias
Yellow	Paradise	Piscinas
Mar	Geotagged	Naturais
Céu	Beaches	Maceió
Azul	Ondas	Turismo
Nuvens	Rio	Brasilviagem
Pentax	Lago	Pousada
Ist	Vivid	Hotéis
D	Praias	Brasil azul
17mm	Maragogi fotos	Pousadas em Maragogi
Fisheye	Fotos de Maragogi	Miramar Maragogi
Brasil	Praia de Peroba	Praia de Maragogi
Brazil	Maragogi online	Porto de Galinhas
Postal	Aventura	Japaratinga
Tamandaré	Galés	Catamarã
Piscinas naturais de Maragogi, Alagoas	Piscinas naturais de Maragogi	Passeio
Hotel Salinas	Fotos Maragogi	Referência no Turismo

Fonte: elaborado pelos autores.

Um olhar não muito atento é o suficiente para perceber que muitas palavras se repetem em *tags* diferentes; ainda, percebemos que palavras que deveriam formar uma única *tag* acabam tornando-se *tags* distintas, como, por exemplo, CARIBE e BRASILEIRO, seguidas da *tag* O CARIBE BRASILEIRO.

Aplicando a metodologia de Rodrigues (2007) obtemos:

Quadro 5—Aplicação da metodologia de Rodrigues (2007) para a Figura 2

DESCRIÇÃO FÍSICA	COMPOSIÇÃO	CONTEXTO ARQUIVÍSTICO
---	Azul Azul turquesa	Peroba Maragogi Alagoas
SENTIDO DENOTATIVO	SENTIDOS CONOTATIVOS	TEMATIZAÇÃO
Praia Mar Céu azul Jovens	Curiosidade Aventura Férias	Turismo Pontos turísticos

Fonte: elaborado pelos autores.

Na tabela comparativa entre as *tags* consideradas válidas e os termos obtidos teremos:

Quadro 6—Comparação entre as *tags* consideradas válidas e os resultados obtidos na análise

<i>TAGS</i>	TERMOS
Maragogi	Azul
Alagoas	Azul turquesa
Peroba	Peroba
Praia	Maragogi
Mar	Alagoas
Céu	Praia
Azul	Mar
Nuvens	Céu azul
Contraste	Jovens
Cores	Curiosidade
Fortes	Aventura
Vibrantes	Férias
Viva	Turismo
Paraíso	Pontos turísticos
<i>TAGS</i>	TERMOS
Brasil	---
Postal	---
Ondas	---
Rio	---
Lago	---
Praias	---
Fotos	---

Maragogi	---
Fotos de Maragogi	---
Praia	---
Peroba	---
Caribe	---
Brasileiro	---
O Caribe brasileiro	---
Pousada	---
Hotéis	---
Viagem	---
Dicas	---
Ficar	---
Salinas	---
Boa viagem	---
Brasilviagem	---
Aventura	---
Férias	---
Piscinas	---
Naturais	---
Maceió	---
Turismo	---
Brasil azul	---
Pousadas em Maragogi	---
Praia de Maragogi	---
Porto de Galinhas	---
Japaratinga	---
TAGS	TERMOS
Tamandaré	---
Piscinas naturais de Maragogi, Alagoas	---
Piscinas naturais de Maragogi	---
Referência no Turismo	---
Hotel Salinas	---
Catamarã	---
Galés	---
Passeio	---
Fotos Maragogi	---

Fonte: elaborado pelos os autores.

Com a metodologia de Rodrigues (2007) atribuímos 14 termos a essa imagem, os quais abrangem desde a sua composição, com a cor que notavelmente se destaca, até a sua tematização. Acreditamos ser possível, com esses termos, descrever eficientemente a fotografia e realizar uma recuperação eficiente. As *tags* consideradas válidas somam 56, dentre as quais oito coincidem com os termos atribuídos. 56 poderia ser considerada uma quantidade de termos elevada, especialmente quando comparada com o resultado obtido através da metodologia de Rodrigues (2007). Entretanto, é preciso levar em conta os aspectos conotativos e toda a gama de possibilidades de interpretação que a imagem pode despertar ao observador, de modo que, dependendo do contexto em que o documento seja analisado e representado, um número grande de termos pode ser considerado ideal. No entanto, é necessário estar atento a equívocos na interpretação que poderiam levar a conceituações distantes e passíveis de confusão com outras imagens. Algumas *tags* atribuídas à imagem em questão podem exemplificar isso, como é o caso de PORTO DE GALINHAS ou TAMANDARÉ, que são praias localizadas no estado de Pernambuco. Ainda, a *tag* MACEIÓ refere-se à capital do estado de Alagoas, que não possui relação direta com a fotografia apresentada, o que poderia provocar confusão no momento da busca, visto que esse termo não corresponde corretamente à procedência da fotografia.

Na terceira página de busca encontramos a seguinte imagem:

Figura 3 – Ponte de madeira – Itapuã, RS



Fonte: *Flickr* (2014b).

Nesta imagem temos apenas um panorama, sem a presença de seres humanos. Segundo a descrição do autor, esta fotografia representa uma região da cidade de Itapuã, no Estado do Rio Grande do Sul. As *tags* atribuídas pelo autor foram:

Quadro 7 – *Tags* atribuídas à Figura 3 por usuário da rede *Flickr*

Itapuã	PB	Bem <i>Flickr</i> ... Bem Brasil!
RS	Noiretblanc	Gutemberg Ostemberg
Interior	Bianco	Gutemberg
Gaúcho	Nero	Porto Alegre
Guaíba	Blanco	Mywinners
Rio Grande do Sul	Negro	Bem <i>Flickr</i> ... Bem Brasil!
BW	Brazil	---

Fonte: elaborado pelos autores.

Mais uma vez percebemos, em INTERIOR e GAÚCHO, palavras que provavelmente deveriam formar uma única *tag*, mas acabam tornando-se *tags* distintas. Aplicando a metodologia de Rodrigues (2007) obtemos:

Quadro 8 – Aplicação da metodologia de Rodrigues (2007) para a Figura 3

DESCRIÇÃO FÍSICA	COMPOSIÇÃO	CONTEXTO ARQUIVÍSTICO
---	Preto e branco	Itapuã Guaíba Rio Grande do Sul
SENTIDO DENOTATIVO	SENTIDOS CONOTATIVOS	TEMATIZAÇÃO
Ponte de madeira Lago Dique	Imensidão Infinito Solidão	Preservação ambiental Ecossistema

Fonte: elaborado pelos autores.

Adotamos 12 termos para representar esta imagem, que buscam englobar tanto o que a imagem mostra, o que ela transmite aos autores, e os contextos em que pode estar inserida de acordo com a história do local. Itapuã é o nome de uma reserva florestal localizada no município de Guaíba, Rio Grande do Sul. Percebe-se, assim, que alguns destes termos só puderam ser atribuídos porque o autor informa através dos textos onde a fotografia foi tirada. A partir daí o indexador pode investigar o que mais pode ser relevante na sua conceituação, indo além do que a imagem revela. É importante destacar que Lancaster (2004)

prevê isso em seus estudos, quando aborda as questões relacionadas à indexação de imagens. No quadro comparativo teremos:

Quadro 9—Comparação entre as *tags* consideradas válidas e os resultados obtidos na análise
 Fonte: elaborado pelos autores.

<i>TAGS</i>	TERMOS
Itapuã	Preto e branco
Interior	Itapuã
Gaúcho	Rio Grande do Sul
Guaíba	Ponte de madeira
Rio Grande do Sul	Lago
Negro	Dique
Porto Alegre	Imensidão
---	Infinito
---	Solidão
---	Guaíba
---	Preservação ambiental
---	Ecossistema

Fonte: elaborado pelos autores.

Dentre as sete *tags* relacionadas, apenas três coincidem com os termos atribuídos pela metodologia de Rodrigues. Identificamos uma situação semelhante à encontrada na imagem analisada anteriormente, quando se adotou a *tag* PORTO ALEGRE, referente à capital do Estado do Rio Grande do Sul, a uma imagem que representa uma localidade situada no município de Guaíba, também no Rio Grande do Sul. No entanto, nesse caso é possível afirmar que, apesar de não haver uma relação direta entre a *tag* e a localidade apresentada, visto que são municípios diferentes, ainda assim seria admissível reconhecer uma relação relevante, já que o município de Guaíba fica localizado na Região Metropolitana de Porto Alegre.

Até agora analisamos três imagens, sendo que em duas delas não aparecem seres humanos. Cumprindo os procedimentos metodológicos estabelecidos, a última imagem a ser analisada deve necessariamente conter um ser humano. Na quarta página dos resultados de busca encontramos a seguinte imagem:

Figura 4 – Chapada dos Veadeiros, GO



Fonte: Flickr(2014b).

Na imagem há apenas um rapaz com a sua câmera fotográfica em meio a uma imensa paisagem verde. Segundo os textos relacionados à fotografia, ela foi tirada na Chapada dos Veadeiros, localizada no estado de Goiás. As *tags* atribuídas a essa imagem, diferentemente das outras analisadas, não foram muitas, como relacionado abaixo:

Quadro 10 – *Tags* atribuídas à Figura 4 por usuário da rede Flickr

Chapada dos Veadeiros	Eu
Vale da Lua	Borguetti
Alto Paraíso	Top-v1111
Me	---

Fonte: elaborada pelos autores.

Com a metodologia de Rodrigues (2007), obtemos os seguintes resultados:

Quadro 11—Aplicação da metodologia de Rodrigues (2007) para a Figura 4

DESCRIÇÃO FÍSICA	COMPOSIÇÃO	CONTEXTO ARQUIVÍSTICO
---	Verde	Chapada dos Veadeiros Goiás
SENTIDO DENOTATIVO	SENTIDOS CONOTATIVOS	TEMATIZAÇÃO
Floresta Homem	Aventura	Pontos turísticos Ecossistema

Fonte: os autores

Na tabela comparativa, considerando os critérios determinados anteriormente para a validade das *tags*, teremos:

Quadro 12—Comparação entre as *tags* consideradas válidas e os resultados obtidos na análise

TAGS	TERMOS
Vale da Lua	Verde
Alto Paraíso	Chapada dos Veadeiros
Eu	Goiás
Chapada dos Veadeiros	Floresta
---	Homem
---	Aventura
---	Liberdade
---	Pontos turísticos
---	Ecossistema

Fonte: elaborado pelos autores.

A Chapada dos Veadeiros é uma vasta região que abrange diversos municípios, dentre eles Cavalcante, Alto Paraíso de Goiás e Teresina de Goiás. Possivelmente por isso o autor atribuiu a *tag* ALTO PARAÍSO, referindo-se ao município em que, talvez, localizava-se no momento da fotografia. No entanto, sendo este termo específico demais, não foi considerado na sua análise. Da mesma forma, VALE DA LUA refere-se a uma região específica da Chapada dos Veadeiros, mas que não é possível identificar nem na descrição textual, nem na imagem em si. Assim, consideramos apenas o termo genérico CHAPADA DOS VEADEIROS. Nesse sentido, apenas uma *tag* coincidiu com os termos atribuídos na análise, que foram nove no total.

Sobre a análise qualitativa de dados, podemos fazer uma breve comparação entre os termos e as *tags*, e verificar como os dados que obtivemos

nas análises se comportam. Todas as tags atribuídas às quatro imagens analisadas e consideradas válidas somam o total de 83, enquanto os termos adquiridos por meio da metodologia de Rodrigues (2007) somam 43. Do total de *tags*, apenas 16 coincidem com o total de termos, o que resulta em aproximadamente 39,02% de compatibilidade entre as *tags* e os termos.

Além disso, analisando as tabelas comparativas podemos perceber que, em geral, os termos que coincidiram foram apenas os que dizem respeito à localização geográfica do ponto em que a fotografia foi tirada, o que não garante uma riqueza na conceituação por meio das *tags*.

Observamos neste estudo, também, que é essencial para o indexador identificar e considerar os diversos aspectos que uma imagem pode conter. Durante as análises, percebemos que as imagens coletadas possuem a maior parte das características propostas por Rodrigues (2007) como sendo as mais relevantes. No entanto, à exceção dos termos PASSEIO, FÉRIAS, PARAÍSO, TURISMO e REFERÊNCIAS NO TURISMO (Figura 2), que remetem à tematização da imagem, não encontramos nenhuma *tag* que correspondesse a elementos conotativos ou temáticos, e poucas as *tags* correspondentes a elementos denotativos. Entende-se, assim, que não há um equilíbrio entre os tipos de elementos descritos, sendo esses predominantemente relativos à sua composição e ao seu contexto arquivístico (RODRIGUES, 2007).

5 Considerações finais

Apesar da baixa qualidade identificada na indexação realizada por meio das *tags*, o valor da indexação social não se perde. O fato de existirem ferramentas capazes de permitir esse tipo de atividade representa um avanço significativo na forma como os novos profissionais da informação devem encarar a representação da informação, seja ela imagética ou não. Alguns estudos inclusive já apontam para propostas de metodologias para a atribuição das *tags* (HIDDERLEY; RAFFERTY, 2007). Isso remonta, de alguma forma, aos estudos iniciais de Panofsky (1979) sobre a conceituação de imagens. Os conceitos iniciais evoluíram de tal forma que hoje somos capazes de analisar

minuciosamente imagens disponibilizadas em ambiente web, algo impensável algumas décadas atrás. É possível que, em alguma medida, estejamos no mesmo marco inicial em que se encontrava Panofsky (1979) quando iniciou o estudo sobre essas questões.

Ao lado de grandes ferramentas da Web 2.0, como a *Wikipedia*, por exemplo, o *Flickr* faz cada vez mais parte do cotidiano de nossa sociedade. No entanto, as possibilidades advindas dessa evolução e a velocidade com que essa ocorreu e ainda ocorre requerem dos profissionais estudos frequentes para compreender a usabilidade e os potencializar os mecanismos dessas novas ferramentas.

A indexação de imagens é uma atividade complexa e altamente passível a interpretações e questionamentos diferenciados. Devido a isso, foi necessário fazer um estudo aprofundado sobre o tema através da literatura existente, para entender realmente a problemática e buscar um embasamento científico para analisar qualitativamente as imagens.

Buscamos, a partir desse estudo, criar uma reflexão em torno das problemáticas que estão surgindo no ambiente de tratamento, organização e disseminação da informação. Por mais que diversos estudos proponham formas de tratamento de imagens, é preciso estar atento à eficiência desses métodos em ambientes que não sejam restritos a bibliotecas e centros de documentação.

Referências

FLICKR. **Ajuda: tags**. 2014a. Disponível em:
<<http://www.flickr.com/help/tags/#37>>. Acesso em: 15 out. 2014.

FLICKR. **Busca de imagens**. 2014b. Disponível em:
<<http://www.flickr.com/search>>. Acesso em: 15 out. 2014.

FLICKR. **Sobre o Flickr**. 2014c. Disponível em:
<<http://www.flickr.com/about>>. Acesso em: 15 out. 2014.

GIL, Antônio Carlos. **Métodos e técnicas de pesquisa social**. 6. ed. São Paulo: Atlas, 2008.

GUEDES, Roger de Miranda; DIAS, Eduardo José Wense. Indexação social: abordagem conceitual. **Revista ACB: Biblioteconomia em Santa Catarina**, Florianópolis, v. 15, n. 1, p.39-53, jan./jun. 2010.

HIDDERLEY, R.; RAFFERTY, P. *Flickr* and democratic indexing: dialogic approaches to indexing. **Aslib Proceedings**, London, v. 59, n. 4/5, p. 397-410, 2007. Disponível em: <www.emeraldinsight.com/0001-253X.htm>. Acesso em: 11 jan. 2012.

LANCASTER, F. W. **Indexação e resumos: teoria e prática**. 2. ed. Brasília: Briquet de Lemos, 2004.

MANINI, Miriam Paula. **Análise documentária de fotografias: um referencial de leitura de imagens fotográficas Para fins documentários**. São Paulo, 2002. Tese (Doutorado em Ciências da Comunicação) - Departamento de Biblioteconomia e Documentação, Escola de Comunicações e Artes, Universidade de São Paulo, São Paulo, 2002.

OLIVEIRA, Rafael Alves de. Metodologias para indexação de imagens fotográficas em ambiente web. In: ENCONTRO DE ESTUDOS SOBRE TECNOLOGIA, CIENCIA E GESTÃO DA INFORMAÇÃO. **Anais...** Recife: –Universidade Federal de Pernambuco, 2011. 1 CD-ROM. .

PANOFSKY, E. **Significado nas Artes Visuais**. 2. ed. São Paulo: Perspectiva, 1979.

RODRIGUES, Ricardo Crisafulli. Análise e tematização da imagem fotográfica. **Ciência da Informação**, Brasília, v. 36, n. 3, p. 67-76, set./dez. 2007.

Analysis and index of images on Flickr network

Abstract: This article presents a brief overview on the dissemination and democratization of the collaborative web, and how this process popularized the use of the image in digital form. A brief analysis of the Flickr social network, identifying their goals, tools, and some studies on its usability in the context of information science. Sets out the approach to the images stored on the network, proposing a qualitative study about the free indexing performed by users. Applies the methodology established in a previous study as the most appropriate

for indexing images in a web environment, comparing the results obtained with free indexing performed by users.

Keywords: Image. Indexing of images. Flickr. Social networks. Social indexing.

Recebido: 16/10/2014
Aceito: 01/04/2015



¹No original: “The [...] tag [...] “wedding” retrieved 795.280 images, when it was used on Flickr on 14 February 2006, which suggests that this an example of a tag that is perhaps too broad. The [...] tag “Ijsselmeer” only retrieved one image and is an example of a tag that is perhaps too specific for public searching purposes”.