

# FOLKSONOMIA: análise da etiquetagem do usuário no Flickr

André Augusto de Abreu Rodrigues\*  
Manoel Palhares Moreira\*\*

## RESUMO

A consolidação do novo paradigma da Internet, a Web 2.0, coloca em evidência os sites de compartilhamento de arquivos e recursos que são atualmente a grande sensação da Internet atraindo um grande e crescente número de usuários. A Folksonomia é o processo de indexação utilizado pela maioria destes sites e consiste na oportunidade ao usuário de representar de forma livre o recurso que utiliza. Como se trata de um recurso novo há necessidade de alargar a compreensão sobre a Folksonomia e observar como os usuários vem se utilizando deste recurso. Por considerar o fenômeno da Folksonomia como um processo de indexação renovado pela massificação de seu uso na internet, a conceituação e o trajeto histórico da indexação, são apresentados como embasamento teórico da Folksonomia e servem também como prisma para observar e caracterizar a atividade do usuário. Este trabalho se propõe a investigar as estratégias do usuário no momento da etiquetagem, procurando estabelecer uma forma de classificar esta atividade de acordo com a possibilidade de compartilhamento de significado da etiquetagem do usuário. Desta forma foi realizado um pré-teste da em escala reduzida de 10% do previsto na metodologia com o intuito de observar a validade da classificação da etiquetagem do usuário proposta por Cañada. Este pré-teste serviu ainda para verificar hipóteses sobre a atividade do usuário e suscitar mudanças na metodologia e na forma de análise das etiquetas.

**Palavras-chave:** Folksonomia. Indexação. Representação da Informação.

## 1 INTRODUÇÃO

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.

---

\*Mestrando em Ciência da Informação na Universidade Federal de Minas Gerais ECI. E-mail: andré.augusto.rodrigues@gmail.com

\*\*Doutor em Ciência da Informação. Universidade Federal de Minas Gerais – ECI. Pontifícia Universidade Católica de Minas Gerais. E-mail: palhares@pucminas.br

Este processo de livre escolha da representação de um conteúdo pelo usuário foi denominado, em termo cunhado por Thomas Vander Wal em 2004, de Folksonomia.

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 motivação que sustenta o processo de etiquetagem do usuário? Como pode ser classificada a etiquetagem do usuário? Quais os tipos de etiquetagem mais utilizados?

Este estudo pretende favorecer uma abertura para que sistemas que utilizem a Folksonomia sejam utilizados em ambientes mais específicos através do estudo de propostas já existentes de classificações para o processo de etiquetagem do usuário, somando a essas novas alternativas. Espera-se assim contribuir no crescimento da compreensão sobre o processo de indexação do usuário, bem como observar a forma como o usuário realiza e se utiliza da Folksonomia, permitindo a criação de ambientes de uso mais intuitivos.

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

Este trabalho tem como objetivo geral analisar as estratégias de indexação dos usuários do *Flickr* e como objetivos específicos: a) levantar quais são as estratégias e os tipos de etiquetagem mais utilizados e seus impactos na recuperação de informação do sistema; b) verificar a adequação dos tipos de etiquetagem propostos por Cañada (2006) à etiquetagem do usuário do Flickr; c) levantar percentuais de não utilização de etiquetas entre usuários; d) levantar percentuais de utilização de etiquetas em língua diferente da língua portuguesa entre usuários que utilizam tal língua.

## 2. 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.

O final do século XIX e o início do século XX começam a aparecer as primeiras ferramentas de descrição de documentos com fins de pesquisa documentária. Nas formulações

que envolviam a fundação do *Office International de Bibliographie*<sup>1</sup>, em Bruxelas, no ano 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, sim, na “*idéia de bibliografia como registro, memória do conhecimento científico, desvinculada dos organismos, como arquivos e bibliotecas, e de acervos*” (PINHEIRO, 1997).

O avanço nos métodos e técnicas de construção de índices transformou estes nos mais importante instrumento de recuperação da informação, a mais poderosa ferramenta de busca (ROBREDO, 1994), 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 de uma indexação 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 Barners-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).

O tratamento interdisciplinar dado pela CI às atividades de representação da informação culminou no desenvolvimento de teorias como a Web semântica e artefatos como os tesauros e ontologias, que apresentam formas de representação mais sofisticadas que os índices.

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 refere-se à participação do usuário. Nos processos de indexação convencional, o usuário não é um participante ativo, mas, o pressuposto da garantia

---

<sup>1</sup> Instituto Internacional de Bibliografia (tradução do autor).

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

### **3 FOLKSONOMIA**

O termo Folksonomia é um neologismo criado por 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 nenhum 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 a 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.

#### **3.1 Individual versus Coletivo**

Avançando sobre as teorias que basearão as análises deste trabalho, Cañada (2006) oferece uma proposta de classificação em que associa a motivação do usuário com o benefício social de sua etiquetagem para o sistema.

Nessa movimentação, percebe-se que a proximidade ou distância de uma etiqueta escolhida pelo usuário aos conceitos de categoria e palavra-chave está associada à motivação do usuário de organizar para uso próprio ou para compartilhar com um grupo restrito ou um grupo aberto. Cañada (2006) tomou por base a observação de etiquetagens realizadas nos sites *Flickr* e *Delicious* para propor quatro estilos de etiquetagem sendo estes: Egoísta,

Amigável, Altruísta e Populista.

A etiquetagem egoísta é considerada como o comportamento mais natural do usuário (MONTERO, 2006) e consiste em etiquetar algo com o intuito de apenas recuperá-lo posteriormente. A etiquetagem amigável retrata um processo com intuito de compartilhar o conteúdo com um grupo restrito de usuários. No caso da etiquetagem altruísta fica expressa a vontade sincera de compartilhar o conteúdo com a comunidade, com o coletivo geral de usuários. Já a etiquetagem populista consiste em algo para tornar o conteúdo mais atrativo, visando apenas obter visitas. Nesse sentido, interessa promover análises que verifiquem a ocorrência e a validade desses tipos através de um percurso metodológico científico que possibilite aprofundar o conhecimento sobre a etiquetagem do usuário e a responder perguntas colocadas neste trabalho, bem como verificar hipóteses estabelecidas nas teorias que nortearão a análise.

#### **4 PERCURSO METODOLÓGICO**

O trabalho constitui-se em uma pesquisa documental que servirá de fonte de dados e base para o estudo de caso a ser realizado empregando técnicas qualitativas. Esta pesquisa será realizada por meio da análise de discurso que, segundo Godoy (1995), parte do pressuposto que, por trás do discurso aparente, simbólico e polissêmico, esconde-se um sentido que convém desvendar.

A adoção de um estudo de caso como método na realização das análises acompanha o aspecto exploratório que visa conhecer com mais profundidade o fenômeno da etiquetagem do usuário. A pesquisa documental permite o acesso a pessoas e eventos distantes física e temporalmente, o que ocorre ao buscarmos etiquetas em um sistema virtual e disseminado mundialmente e, no qual, não teremos contato direto com os usuários que postaram e classificaram os materiais disponíveis.

##### **4.1 Seleção das Amostras**

a) Serão consultadas cem pessoas, cinquenta de cada sexo, de acordo com a faixa etária, com a seguinte pergunta: Cite cinco temas de fotos que você gostaria de ver; b) os quinhentos temas levantados serão organizados em ordem alfabética não importando a repetição entre eles; c) desta ordem alfabética serão selecionados cinquenta temas escolhidos por sorteio de lista aleatória de intervalo igual a dez, a partir de número aleatório sorteado

entre um (1) e dez (10); d) os cinquenta temas selecionados serão pesquisados em um mesmo dia; e) para cada tema serão selecionados os três primeiros usuários sem repetição de usuário; f) dos usuários selecionados serão coletadas as etiquetas retiradas da primeira e segunda imagem respectivamente, a mais recentemente adicionada e a segunda mais recente.

#### 4.2 Método das Análises

Segundo Montero (2006), os termos Categoria e Palavra-chave serão extrapolados de seu conceito original. A primeira para representar a movimentação entre uma etiqueta que produz significado apenas para quem a insere. A segunda, uma etiqueta que agrega possibilidade de significação dentro de um grupo restrito ou para o coletivo de usuários dentro do recorte selecionado para esta pesquisa. Para dar mais precisão a esta classificação será observada a relação da etiqueta com contexto em que está inserida (a imagem que classifica, o perfil, os contatos, os grupos e outras informações do sistema).

O benefício social de uma etiqueta será extraído da relação desta com o eixo Categoria - Palavra-chave e da verificação desta relação no contexto em que esta etiqueta foi inserida. A tabela 1 resume as relações entre as categorias que fundamentaram a análise das etiquetas.

<b>Classificação</b>	<b>Compartilhamento de significado</b>	<b>Benefício social</b>	<b>Relação com o referente da imagem</b>
Altruísta	Grande	Alto	Direta e forte.
Amigável	Restrito ao contexto	Médio a alto	Indireta ou fraca.
Egoísta	Pouco ou nenhum.	Pouco ou nenhum.	Indireta, genérica, nenhuma.
Populista	Grande	Nenhum. Muitas das vezes agrega ruído.	Indireta, fraca ou nenhuma.

Tabela 1: Descrição das características e relações observadas na análise.

## 5 PRÉ-TESTE, RESULTADOS e CONCLUSÃO

Para verificação de hipóteses e avaliação da metodologia, foi realizado um pré-teste seguindo integralmente as prescrições relatadas na metodologia com a seleção da amostra sendo esta reduzida em escala de 10% do previsto para o trabalho. Assim sendo, foram selecionados dez entrevistados mantidas as condições de faixa etária e sexo. Foram sorteados cinco temas de uma lista de cinquenta coletados nas entrevistas, selecionando cinco usuários (um por tema), uma foto de cada usuário e a coleta e análise de todas as etiquetas atribuídas à imagem selecionada, totalizando 27 etiquetas. Não foi encontrada nenhuma foto sem etiqueta,

mas quatro etiquetas em língua estrangeira, representando 15% do total de etiquetas selecionadas.

Realizada a análise para enquadrar as etiquetas selecionadas nos tipos descritos no capítulo 4.2, os resultados mostraram 33% de etiquetas altruístas, 26% de etiquetas amigáveis, 26% de etiquetas populistas e 15% de etiquetas egoístas, contrariando a hipótese de que as etiquetagens amigável e altruísta exigem mais esforço e, portanto, teriam percentuais menores.

A realização do pré-teste apontou acertos no caminho e indicou possibilidade de pesquisa pela metodologia criada. Criaram-se também possibilidades para que novas categorias de análise pudessem estar presentes, levantando críticas e acertos à metodologia e ao método de análise. Validou-se a existência das categorias de Cañada (2006), suscitando, entretanto, mudanças na forma de análise formulada em torno destas categorias, com inclusão de novos elementos que contribuirão para um melhor desenho das categorias e maior precisão das análises.

## REFERÊNCIAS

- AZEVEDO, Alexander W. A construção da Ciência da Informação na pós-modernidade: dialética histórica. **Revista Digital de Biblioteconomia e Ciência da Informação**, Campinas, v6, n. 2, p. 71-82, jan./jun. 2009.
- 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.
- CAÑADA, J. TIPOLOGÍAS y estilos en el etiquetado social. Terremoto.net: Diseño de interacción desde el año 2000. Disponível em: <http://www.terremoto.net/tipologias-y-estilos-en-el-etiquetado-social/>.2006
- FLICKR. Disponível em <[www.flickr.br](http://www.flickr.br)>. Acesso em 10 jan 2008.
- FUJITA, M. S. L. SILVA, M. R. A prática da indexação: análise da evolução de tendências teóricas e metodológicas. **Transinformação**, Campinas, n.16 (2), p.133-161, mai./ago. 2004.
- GODOY, Arilda S Pesquisa qualitativa – tipos fundamentais. **Revista de Administração de Empresas**, v.35, n.3, Mai./Jun. 1995b, p.20-29.
- LANCASTER, F. W. **Indexação e resumos**: teoria e prática. Trad. de Antonio Agenor Briquet de Lemos. Brasília: Briquet de Lemos/Livros, 1993.
- MONTERO, Yusef H. Indización Social y Recuperación de Información. **No solo usabilidad journal**, n. 5. 16 de novembro de 2006.

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, Informação, Ferramentas Ontológicas e Redes Sociais ADHOC: a interoperabilidade na construção de tesouros e ontologias. **Informação & Sociedade: Estudos**, João Pessoa, v.19, n.1, p. 59-73, jan./abr. 2009.

ROBREDO, J. **Documentação de hoje e amanhã**: uma abordagem informatizada de biblioteconomia e dos sistemas de informação. 2. ed. ver. ampl. Brasília: Edição do autor, 1994.