

UNIVERSIDADE FEDERAL DO RIO GRANDE DO SUL
FACULDADE DE BIBLIOTECONOMIA E COMUNICAÇÃO
DEPARTAMENTO DE CIÊNCIAS DA INFORMAÇÃO

PRISCILA DA SILVA LOPES

**Uma abordagem sobre a indexação utilizada no *Tumblr*: a
folksonomia em foco**

Porto Alegre

2012

Priscila da Silva Lopes

Uma abordagem sobre a indexação utilizada no *Tumblr*: a folksonomia em foco

Trabalho de Conclusão de Curso, submetido ao curso de Biblioteconomia da Faculdade de Biblioteconomia e Comunicação, da Universidade Federal do Rio Grande do Sul, como requisito parcial para obtenção do título de Bacharel em Biblioteconomia.

Orientador: Me. Jackson da Silva Medeiros

Porto Alegre
2012

UNIVERSIDADE FEDERAL DO RIO GRANDE DO SUL

Reitor: Prof. Dr. Carlos Alexandre Netto

Vice-Reitor: Prof. Dr. Rui Vicente Oppermann

FACULDADE DE BIBLIOTECONOMIA E COMUNICAÇÃO

Diretor: Prof. Ricardo Schneiders da Silva

Vice-Diretora: Prof. Dra. Regina Helena van der Laan

DEPARTAMENTO DE CIÊNCIAS DA INFORMAÇÃO

Chefe: Prof. Dra. Ana Maria Mielniczuk de Moura

Chefe Substituta: Prof. Dra. Sônia Elisa Caregnato

COMISSÃO DE GRADUAÇÃO DO CURSO DE BIBLIOTECONOMIA

Coordenadora: Prof. Me. Glória Isabel Sattamini Ferreira

Coordenadora Substituta: Prof. Dra. Samile Andréa De Souza Vanz

L864a Lopes, Priscila da Silva

Uma abordagem sobre a indexação utilizada no Tumblr: a folksonomia em foco / Priscila da Silva Lopes ; orientador Jackson da Silva Medeiros. – 2012.

69 f.

Trabalho de Conclusão de Curso (graduação) – Universidade Federal do Rio Grande do Sul, Faculdade de Biblioteconomia e Comunicação, Departamento de Ciências da Informação, Porto Alegre, 2012.

1. Organização e representação da informação. 2. Folksonomia. 3. Sistema de recuperação da informação. 4. Tags. 5. Tumblr. I. Medeiros, Jackson da Silva. II. Título.

CDU 025.4

Departamento de Ciências da Informação
Rua Ramiro Barcelos, 2705 Campus Saúde
Bairro Santana
Porto Alegre - RS
Cep: 90035-007
Telefone: (51) 3308-5146
E-mail: fabico@ufrgs.br

Priscila da Silva Lopes

Uma abordagem sobre a indexação utilizada no *Tumblr*: a folksonomia em foco

Trabalho de Conclusão de Curso, submetido ao curso de Biblioteconomia da Faculdade de Biblioteconomia e Comunicação, da Universidade Federal do Rio Grande do Sul, como requisito parcial para obtenção do título de Bacharel em Biblioteconomia.

Data de aprovação: 25 de junho de 2012.

BANCA EXAMINADORA

Me. Jackson da Silva Medeiros (orientador)

Prof. Dr. Rafael Port da Rocha - UFRGS

Prof. Dr. Rodrigo da Silva Caxias de Sousa - UFRGS

AGRADECIMENTOS

Primeiramente agradeço à *UFRGS* e à *Fabico* pela oportunidade de poder conviver em um ambiente que carrega tanta história e pelas pessoas competentes que pude conhecer e aprender com elas.

Ao meu orientador, *Jackson Medeiros*, pela paciência, tranquilidade e pelo seu conhecimento que foi fundamental para a realização deste estudo.

Aos professores que compõem a banca examinadora, pela disposição de tempo e atenção para com o este trabalho.

Aos melhores pais desse mundo, *Luis Nei* e *Marlene*, pelo apoio incondicional, carinho, educação, alento e confiança que depositaram em mim durante toda minha vida.

Aos meus queridos irmãos: *Tita*, pela paciência, dedicação e apoio incondicional que recebi durante toda esta trajetória, certamente sem ele esse sonho não seria possível; *Gordo* pelos momentos de alegria e por todo seu carinho; e *Manê* que, mesmo distante, sempre permaneceu ao meu lado dando conselhos, “puxões de orelha” e muito amor.

Às pessoas maravilhosas que este curso possibilitou eu conhecer. *Ana, Ane, Annie, Augusto, Carluxa, Dani, Fabri, Julinho, Karin, Liriens, Leandro K., Leandro O., Lu, Marcia, Nalines, Nata, Rose e Zanza*. Aos dias e noites felizes e infinitos, às trocas de figurinhas, ao apoio emocional e profissional e, principalmente, aos laços de amizade construídos durante toda essa jornada.

À *Biblioteca Lucília Minssen* pela oportunidade do primeiro contato com a profissão, especialmente à *Naida* pelo apoio, ensinamentos, amizade e pela inspiração de querer ser uma profissional competente igual à ela.

Ao *CDI Zero Hora* pelo privilégio de ver um outro lado da Biblioteconomia em que estou satisfeita com o trabalho que empenho. À equipe animada e responsável que divide o mesmo setor comigo: *Cla, Leti, Sarrafo, Fabi, Bibs, Rose, Julio e Stéfano*. Aos colegas que tive o prazer de trabalhar: *Anne, Cris, Leandro, Thiago e Zanza*. Todos acrescentaram valores e conhecimentos para a minha vida profissional.

A todos os *amigos* que pude conhecer em Porto Alegre e aos antigos de Alegrete, que sempre estão ali para me ajudar, seja com palavras, abraços ou cervejas. À *Bárbara* pela fidelidade e companheirismo durante todos estes anos.

À *Fabiane* por aguentar o mau-humor de finais de semestre e pelo suporte em dias difíceis. À *Bárbara Fabres*, pela amizade e também por ajudar com suas noções de inglês muito valiosas para este trabalho.

E a *todos* que, de alguma forma, fizeram parte destes anos de aprendizados pessoais e acadêmicos, meus sinceros agradecimentos.

*das coisas
que eu fiz a metro
todos saberão
quantos quilômetros
são*

*aquelas
em centímetros
sentimentos mínimos
ímpetos infinitos
não?*

PAULO LEMINSKI

RESUMO

Analisa a interface de recuperação da informação do site Tumblr e categoriza as *tags* indexadas pelos usuários de acordo com os estilos de etiquetagem propostos por Cañada (2006). O estudo teve abordagem quanti-qualitativa com a finalidade de obtenção e interpretação dos dados coletados a partir da observação assistemática do site Tumblr para a análise da interface de recuperação da informação e também para a classificação de estilos de etiquetagem. Escolheu-se três categorias aleatórias em que o próprio site classifica, que foram *Animals*, *Landscape* e *Fashion*. A partir das categorias, escolheu-se aleatoriamente duas *tags* para cada categoria. De acordo com as *tags*, foram selecionadas fotografias aleatórias e observou-se as outras *tags* relacionadas pelos usuários indexadores. Após, as *tags* foram classificadas nos estilos de etiquetagem propostos por Cañada (2006), a saber: egoísta, amigável, altruísta e populista. Conclui-se que os usuários do site *Tumblr*, em geral, classificam suas *tags* de acordo com o estilo altruísta, que facilita a recuperação de informações, e que o sistema do site *Tumblr*, apesar de confuso, é satisfatório dentro de sua proposta.

Palavras-chave: Organização e representação da informação. Folksonomia. Tags. Tumblr.

ABSTRACT

This paper analyzes the site Tumblr and its interface of information retrieval, categorizing the tags indexed by users from site using the styles of tagging proposed by Cañada (2006). It's a quantitative and qualitative study, with the purpose of obtain and interpret the data collected from the systematic observation of Tumblr to do the analysis of information retrieval interface and also for the classification of styles tagging. Three random categories classified by Tumblr were chosen: Animals, Landscape and Fashion. From this categories were chosen two randomly tags, for each class. By the tags, random photographs were selected to observe the other tags associated by indexers users. After the observations the tags were classified in the styles of tagging proposed by Cañada (2006): egoistic, friendly, altruistic and populist. The study concludes that users of Tumblr, generally, classify their tags by the altruistic style, that facilitates retrieval of information. Also the Tumblr's system, although confuse, is satisfactory in its proposal.

Keywords: Organization and representation of information. Folksonomia. Tags. Tumblr.

SUMÁRIO

1 INTRODUÇÃO	11
2 OBJETIVOS	13
2.1 OBJETIVOS GERAL.....	13
2.2 OBJETIVOS ESPECÍFICOS.....	13
3 REFERENCIAL TEÓRICO.....	14
3.1 REPRESENTAÇÃO DA INFORMAÇÃO.....	14
3.2 WEB 2.0.....	16
3.3 FOLKSONOMIA.....	19
3.3.1 Estilos de etiquetagem	23
3.4 TUMBLR	25
4 METODOLOGIA.....	32
4.1 ABORDAGEM E TIPO DE PESQUISA.....	32
4.2 OBJETOS DA PESQUISA.....	33
4.3 PLANO DE ANÁLISE DE DADOS	33
5 ANÁLISE DE DADOS	35
5.1 INTERFACE DE RECUPERAÇÃO DA INFORMAÇÃO NO <i>TUMBLR</i>	35
5.2 CLASSIFICAÇÃO DE TAGS.....	40
5.2.1 Animals.....	40
5.2.2 Landscape.....	50
5.2.3 Fashion.....	57
6 CONSIDERAÇÕES FINAIS	65
REFERÊNCIAS.....	67

1 INTRODUÇÃO

Em meio a tanta informação adquirida em tempos atuais, seja pelos meios tradicionais de comunicação, como televisão, jornal e rádio, a internet tornou-se um dos mais importantes produtos midiáticos destes novos tempos. Junto com ela, foi acrescida uma boa dose de imediatismo e ansiedade informacional, que é resultado de tanta informação que nos é disponibilizada. Sites de compartilhamentos e redes sociais, advindos da web 2.0, são um dos meios mais utilizados para manter-se informado. Para organizar esse enorme campo de informações e conhecimento, a folksonomia tornou-se uma das representações de informação mais necessárias para este tipo de mídia.

A web 2.0, surgida no início dos anos 2000, tem como conceito a web como plataforma, ou seja, uma web que englobe diversas ferramentas, aplicativos e tecnologias da informação. Os softwares funcionam com a internet, e não apenas em um computador local, facilitando assim a ação colaborativa. Focos da web 2.0 são o compartilhamento, os usuários e uma web colaborativa.

Um dos modos de participação e colaboração do usuário se dá através da folksonomia, que tornou-se elemento fundamental para obter-se a criação de um conhecimento coletivo. A folksonomia pode ser definida como:

Uma classificação livre de documentos, feita por etiquetagem, as chamadas *'tags'*, que utiliza a linguagem natural, nem sempre a utilizada nos documentos, e sim a linguagem comum do usuário, presente na memória coletiva” (SILVA; SANTOS, 2011, p. 94).

Neste contexto, as redes sociais utilizam a folksonomia para organizar o conteúdo alimentado diariamente em seus sistemas. O *Tumblr*, uma ferramenta relativamente nova, é um microblog e o seu foco é o compartilhamento de conteúdo (fotografias, animações, vídeo, áudio, texto e ilustrações). O usuário, além de poder publicar seu próprio conteúdo, também pode compartilhar o de seus amigos podendo etiquetar de acordo com a sua concepção.

A importância deste estudo está em analisar como o site Tumblr aborda a folksonomia e se ela realmente é eficaz para a recuperação da informação. Também

é relevante a importância de analisar um site que, apesar de recente, tem grande visibilidade mundial.

Por ser uma ferramenta nova, não existem publicações acerca da interface de recuperação de informação do *Tumblr*. Para isso, este trabalho preocupou-se em analisar este sistema e buscar compreender como ele funciona, suas vantagens e desvantagens.

Deste modo, o foco principal desta pesquisa está em analisar qual o estilo de etiquetagem mais utilizado no site e como os usuários indexam seu conteúdo, através da seguinte questão: *como está categorizada, de acordo com Cañada (2006), a etiquetagem de fotografias no Tumblr?*

Utilizou-se dos conceitos e categorias de Cañada (2006) para tomar partido das análises de *tags* indexadas no *Tumblr*. O autor explica que “há estilos muito marcados de etiquetar e há ambientes em que funcionam melhor ou pior, precisamente porque neles predominam um ou outro estilo de etiquetar” (CAÑADA, 2006). O autor quis dizer que nem em todos os sistemas predominam o mesmo modo de etiquetar por parte dos usuários, porém, dependendo do sistema, pode ser que algum outro estilo predomine sobre os demais.

O presente trabalho abordou como referencial teórico: a representação da informação e algumas de suas linguagens de representação; aspectos relativos à web 2.0; a folksonomia e os estilos de etiquetagem dados por Cañada (2006); por fim, detalhes e características do *site Tumblr*. Alguns autores utilizados foram: Bräscher e Café (2008), Van de Laan (2002), Rowley (1988), O’Reilly (2005), Primo (2007), Wal (2006), Catarino e Baptista (2007), Noruzi (2007), Cañada (2006) entre outros.

A metodologia utilizada foi de métodos mistos (quanti-qualitativa), com a utilização da observação assistemática como base para as análises de como funciona a interface de recuperação do *Tumblr* e a classificação de fotografias por parte dos usuários do site.

Pretende-se com este trabalho conhecer melhor o sistema utilizado pelo *site Tumblr*, além de verificar a consistência da folksonomia como linguagem de classificação.

2 OBJETIVOS

Os objetivos que nortearão o trabalho seguem abaixo e estão divididos em objetivo geral e objetivos específicos.

2.1 OBJETIVO GERAL

Analisar a classificação de fotografias realizada pelos usuários do sistema *Tumblr*.

2.2 OBJETIVOS ESPECÍFICOS

- a) Identificar as formas de recuperação de fotografias no *Tumblr*;
- b) identificar os problemas e as vantagens da interface de recuperação da informação no *Tumblr*;
- c) classificar, de acordo com a classificação de *tags* proposta por Cañada (2006), as etiquetas das fotografias pesquisadas;
- d) analisar o resultado da classificação obtida e identificar o(s) tipo(s) de etiquetagem(ns) predominante(s) do *Tumblr*, segundo Cañada (2006).

3 REFERENCIAL TEÓRICO

Neste capítulo faz-se uma revisão de literatura sobre representação da informação, web 2.0, folksonomia e o *Tumblr*.

3.1 REPRESENTAÇÃO DA INFORMAÇÃO

A representação da informação está intimamente ligada à organização da informação. Burke (2003 *apud* BRASCHER; CAFÉ, 2008, p. 3) atribui à informação as seguintes características: o que é relativamente 'cru', específico e prático. Ou seja, a informação seria apenas um modo de significar algum objeto.

Já para Fogl (1979, p.22 *apud* BRASCHER; CAFÉ, 2008, p. 3), "não há conexão direta entre informação e objeto, uma vez que a única fonte de origem da informação é o conhecimento, a consciência humana e não o próprio objeto que está sendo conhecido, avaliado ou transformado". Sendo assim, o conceito de informação pode estar além da transmissão e do armazenamento físico. O autor também explica que informação pode ser "um item definitivo do conhecimento expresso por meio da linguagem natural ou outros sistemas de signos percebidos pelos órgãos e sentidos" (FOGL, 1979 *apud* BRASCHER; CAFÉ, 2008, p. 4).

Neste contexto, Taylor (2004 *apud* BRASCHER; CAFÉ, 2008, p. 5) define o objeto como sendo uma "unidade de informação organizável - a informação registrada, que inclui, dentre outros, textos, imagem, registros sonoros, representações cartográficas e páginas web". A organização da informação é atingida quando é feita a descrição física e de conteúdo destes objetos, sendo a representação da informação o meio em que esse objetivo é concluído. A representação da informação é "entendida como um conjunto de elementos descritivos que representam os atributos de um objeto informacional específico" (BRASCHER; CAFÉ, 2008, p. 5). Para isso, existem linguagens específicas para a construção da representação da informação, descrevendo a informação e também o suporte destes objetos.

O processo de indexação é uma forma de representação da informação, permitindo que através da análise de um documento sejam atribuídos termos

representativos do conteúdo deste para serem utilizados no processo de recuperação da informação.

A finalidade da indexação, segundo Robredo e Cunha (1994), é que a recuperação da informação seja permitida atendendo sempre às necessidades dos usuários, classificando-os em dois tipos: (a) os que precisam de informações sobre certo assunto e (b) aqueles que identificam os documentos requeridos, ou seja, já têm certo conhecimento da informação que necessitam.

Neste contexto, a indexação é totalmente voltada a quem procura essa informação, ou seja, o usuário. De um ponto de vista profissional, é necessário indexar viabilizando uma gama de opções para que os mais diversos tipos de usuários possam recuperar tal informação.

Desta maneira, Van der Laan (2002, p. 11) afirma que a indexação é o: “[...] processo de identificação e organização dos itens necessários a posterior recuperação das informações contidas em um documento, observando-se que os principais pontos de acesso para a recuperação de informações são autor, título e assunto”.

As linguagens de indexação, segundo Araújo Junior (2007, p. 33), permitem representar o conteúdo dos documentos de forma analítica, ou seja, é necessário que a representação destes documentos seja minuciosa para que a recuperação da informação seja precisa. Austin e Dale (1991 *apud* VAN DER LAAN, 2002, p. 29) destacam que a linguagem de indexação é “um conjunto controlado de termos selecionados da linguagem natural, usado para apresentar, de forma abreviada, os assuntos do documento”.

Rowley (1988) afirma que a linguagem de indexação pode ser definida em três categorias, sendo elas: (a) linguagens de indexação controladas, (b) linguagens de indexação natural e (c) linguagem de indexação livre.

A autora define as três categorias da seguinte forma:

- a) Linguagens de indexação controladas são aquelas que utilizam cabeçalhos de assunto e/ou tesauros. São termos a partir de um vocabulário controlado;
- b) linguagens de indexação natural se apropriam da linguagem do próprio documento;

- c) linguagem de indexação livre, em que qualquer palavra ou termo que convenha para expressar o assunto pode ser designado como um termo de indexação.

Assim como existem as linguagens documentárias controladas para organizar a informação em unidades de informação, existe um outro tipo de indexação com linguagem natural e livre e que independe da pessoa ser profissional ou não, tendo como base a web. É a indexação social, que segundo Guedes e Dias (2010), trata de “sistemas orgânicos baseados na atribuição livre e pessoal de etiquetas a informações ou objetos visando a organização e recuperação da informação”. O foco da indexação social é o usuário, pois é ele quem indexa o conteúdo na rede e quem irá buscar pode ser ele ou qualquer outro usuário, já que esse processo explora o lado pessoal do usuário indexador.

A indexação social, também conhecida como folksonomia, é o foco deste trabalho. Ela propõe a liberdade de construção de conhecimento coletivo e aborda uma indexação social em que várias pessoas podem construir uma gama de palavras-chave para algum elemento da internet como vídeos, imagens, textos etc.

No contexto deste estudo, é importante entender como funciona a folksonomia e também o meio em que ela está inserida, a *web 2.0*.

3.2 WEB 2.0

Hoje o termo *web 2.0* já não gera a estranheza de quando surgiu. Conhecida também como *web social*, a *web 2.0* advém de uma nova era cultural em que o compartilhamento de informações é o foco da tecnologia. Segundo Primo (2007, p. 1):

O termo *web 2.0* faz um trocadilho com o tipo de notação em informática que indica a versão de um software, foi popularizado pela O'Reilly Media e pela MediaLive International como denominação de uma série de conferências que tiveram início em outubro de 2004. (PRIMO, 2007, p. 1).

Tim O'Reilly, um dos precursores deste estudo, define a *web 2.0* como:

A mudança para uma internet como plataforma, e um entendimento das regras para obter sucesso nesta nova plataforma. Entre outras, a regra mais importante é desenvolver aplicativos que aproveitem os efeitos de rede para se tornarem melhores quanto mais são usados pelas pessoas, aproveitando a inteligência coletiva. (O'REILLY, 2005, p. 3).

Primo (2007) colabora dizendo que a *web 2.0*:

[...] tem repercussões sociais importantes, que potencializam processos de trabalho coletivo, de troca afetiva, de produção e circulação de informações, de construção social de conhecimento apoiada pela informática. (PRIMO, 2007, p. 2).

Desta forma, a *web 2.0* funciona em ambientes propícios à criação e manutenção de redes sociais (abertas ou fechadas, públicas ou privadas), em que há ligações entre vários sistemas em um só site, como por exemplo, os aplicativos de outras redes sociais dentro do Facebook. O objetivo é de proporcionar funcionalidades adicionais aos usuários destas comunidades.

Na *web 1.0* os sites trabalhavam como sistemas isolados. Já na *web 2.0* a estrutura é integrada de funcionalidades e conteúdo. Segundo O'Reilly (2005), alguns dos princípios fundamentais são “a simplicidade, o foco no conteúdo, a colaboração, o compartilhamento e a *web* como plataforma”.

A simplicidade é por conta da apresentação do conteúdo do sistema, que deve ser de fácil manuseio para o usuário. São destacados cinco pontos importantes para atingir este princípio, baseados em O'Reilly (2005):

- a) objetividade: o site deve ir ao encontro do que o usuário deseja;
- b) escopo: deve delimitar a atuação do usuário no sistema, para que evite erros futuros por parte do usuário;
- c) minimalismo: quanto menos etapas o usuário precisar percorrer para alcançar o seu objetivo, melhor o sistema.
- d) usabilidade: o sistema deve ser de fácil uso e aprendizado;
- e) acessibilidade: a informação deve estar acessível e disponível a todo e qualquer tipo de usuário.

O conteúdo é um dos grandes focos da *web 2.0*, sejam textos, áudio, vídeo,

imagem etc. Segundo Primo (2007, p. 3):

A internet foi celebrada por sua tecnologia *pull* (o conteúdo é “puxado” pela audiência), que se opunha ao modelo *push* (o conteúdo é “empurrado” até a audiência) da mídia massiva. Enquanto esta se caracteriza por um sistema centralizado de distribuição dos mesmos conteúdos para toda a massa em horários e/ou espaços determinados. (PRIMO, 2007, p. 3).

A internet começou inteiramente como uma tecnologia em que o conteúdo é posto no sistema para ser requerido quando o receptor o necessitar. A web 2.0 e seu novo sistema possuem um conteúdo participativo e democrático, onde há sempre uma alimentação contínua de informação.

A colaboração dos usuários é, neste sentido, essencial para que isso funcione. O usuário acaba deixando de ser apenas consumidor de informação passando a ser o construtor da informação, formando uma comunidade, o que leva a garantir disseminação de conhecimento e credibilidade de suas publicações. De acordo com Primo (2007, p. 4), “[...] nestes casos importa menos a formação especializada de membros individuais. A credibilidade e relevância dos materiais publicados são reconhecidas a partir da constante dinâmica de construção e atualização coletiva”.

O compartilhamento é resultado do conteúdo e da colaboração dos usuários do sistema, em que há uma exposição de conteúdo produzido, opiniões, gostos, informações pessoais etc.

Por fim, a *web* como plataforma, encerrando os princípios básicos da *web* 2.0, funciona como “arquitetura de participação”. Segundo O’Reilly (2005), o sistema *web* viabiliza ações que antes funcionariam somente com algum software instalado no computador. Hoje elas ficam armazenadas em rede, podendo ser acessadas de qualquer lugar. Primo exemplifica a arquitetura de participação utilizando um tipo de rede, a *peer-to-peer* (par a par):

[...] nas redes *peer-to-peer* (P2P), voltadas para a troca de arquivos digitais, cada computador conectado à rede torna-se tanto “cliente” (que pode fazer *download* de arquivos disponíveis na rede) quanto um “servidor” (oferta seus próprios arquivos para que outros possam “baixá-lo”). Dessa forma, quanto mais pessoas na rede, mais arquivos se tornam disponíveis. (PRIMO, 2007, p.2).

Este exemplo ajuda a compreender um dos princípios chave da *web 2.0* segundo O'Reilly (2005), em que os serviços tornam-se melhores quanto mais pessoas os utilizarem.

Com tanto conteúdo produzido, este novo sistema precisa de uma organização que recupere as informações contidas na *web*. A *web 2.0* apresenta um processo coletivo para organização e recuperação de documentos eletrônicos. E aqui entra a folksonomia, que permite a facilidade de indexar e compartilhar conteúdos.

3.3 FOLKSONOMIA

Toda a facilidade que a *web 2.0* trouxe para os usuários da rede, como poder manusear um *site* sem entender sobre *HTML*¹ e outras linguagens de programação, fez com que as publicações de conteúdo na *web* fossem multiplicadas em números jamais imaginados. Para organizar todo esse material, nada mais justo que o próprio usuário/indexador faça este trabalho.

Pensando desta maneira, novos conceitos surgem para organizar e classificar os recursos digitais disponíveis na internet. A folksonomia é um destes novos conceitos, também conhecida como indexação social, etiquetagem colaborativa, etiquetagem social, classificação social etc.

O termo folksonomia foi criado pelo arquiteto da informação Thomas Vander Wal, em que *folk* (pessoas) mais taxonomia corresponde a uma indexação feita por pessoas (PRIMO, 2007, p.3).

Wal (2006) conceitua folksonomia como:

O resultado da marcação livre e pessoal de informações e objetos (qualquer coisa com uma *URL*) para uma recuperação própria. A marcação é feita em um ambiente social (compartilhado e aberto aos outros). O ato de marcação é feito pela pessoa que consome a informação. [...] É uma classificação social, fazendo com que outras pessoas que usem o mesmo vocabulário sejam capazes de encontrar o mesmo objeto. (WAL, 2006).

Wal (2006) explica que o valor desta marcação é derivado de pessoas que usam de seu próprio vocabulário e acrescentam significados que podem interferir na

¹ HTML = HyperText Markup Language (Linguagem de Marcação de Hipertexto)

² *Bookmark* é uma ferramenta presente em todos os *browsers* atuais que serve como um

informação e/ou objeto como ele é. Esta marcação é denominada *tag* (etiqueta), que é a classificação do conteúdo da informação. *Tagging* (etiquetar) é o ato de nomear, identificar alguma coisa.

Taxonomia é uma organização de informações através de classes “[...] que limita o pesquisador na medida em que o resultado das buscas é definido em função de uma determinada listagem de palavras relacionadas com determinado assunto” (AQUINO, 2007, p.4). Ao contrário disso, a folksonomia é um tipo de indexação livre com linguagem natural, em que o usuário detém os direitos de classificar ou não o conteúdo do documento em que está inserindo em meio à rede. Não existe uma hierarquia entre as tags e pode não haver relação entre elas. Um exemplo que pode ser dado é a indexação de um vídeo de uma corrida de carros, em que o usuário vai definir suas tags como: “carro”, “azul”, “vermelho”, “chuva”, “velocidade”. As tags não se relacionam e se referem a coisas distintas.

Este tipo de indexação gera muitas controvérsias pelo fato de não existir uma estrutura organizacional. Entretanto, para o pesquisador pode ser mais fácil encontrar termos que não são abrangidos nas demais linguagens de indexação que utilizam tabelas de hierarquia. Além disso, na folksonomia não é necessário ser um indexador profissional para classificar as informações de determinado objeto, como explica Mathes:

A folksonomia representa uma mudança fundamental pois é derivada não de profissionais ou criadores de conteúdo, mas de usuários de informações e documentos. Desta forma, ela diretamente reflete as escolhas de enunciação, terminologia e precisão. (MATHES, 2004, p.7 *apud* PRIMO, 2007).

Existe um outro tipo de funcionalidade de recuperação de informação em alguns sistemas que utilizam a folksonomia, que é o *tag cloud* (nuvem de etiquetas). Ela funciona como uma exibição das tags mais utilizadas pelos usuários, seja em seu perfil pessoal ou as tags mais utilizadas no site, em determinado tempo. As tags mais utilizadas são postas com fontes maiores e destacadas, facilitando o acesso aos termos mais utilizados.

Catarino e Baptista (2007), destacam algumas vantagens deste tipo de indexação.

A possibilidade de formar, automaticamente, comunidades em torno de assuntos de interesse na medida em que, ao utilizar serviços de folksonomia, o usuário tem acesso aos outros usuários que têm os mesmos interesses identificados através das etiquetas. Uma outra característica que se destaca é a de que não há uma regra preestabelecida de controle dos vocabulários. Esta característica pode ser vista como uma vantagem na medida em que os usuários dos recursos expressam, ao etiquetar estes conteúdos, a sua estrutura mental em relação àquela informação: há uma liberdade de expressão que possibilita abarcar todas as formas de ver um mesmo conteúdo respeitando as diferenças culturais, interpretativas, etc. (CATARINO E BATISTA, 2007).

Contudo, toda essa liberdade também acarreta alguns ruídos na recuperação da informação. Alguns autores, como Cañada (2006) e Noruzi (2007), listam uma série de questões que prejudicam a recuperação da informação e que estabelecem lacunas neste modo de indexar, mais precisamente, na recuperação de informação do conteúdo *online*.

Noruzi (2007) define quatro problemas de etiquetagem, que são: pluralidade, polissemia, sinonímia e especificidade da marcação. O autor explica-se dizendo que:

Plurais: plurais e partes do discurso ortográfico podem prejudicar um sistema de etiquetagem. Por exemplo, se as etiquetas “gato” e “gatos” são distintos, em seguida, uma consulta para uma das tags não irá recuperar muito, a menos que o sistema de busca inteligente tem a capacidade de executar tais substituições incorporadas.

Polissemia: polissemia refere-se a uma palavra que tem dois ou mais significados semelhantes. “Poli” significa “muitos”, e “semia” significa “significados”.

Sinonímia: palavras diferentes com significados semelhantes ou idênticos, apresentam um problema maior para os sistemas de marcação, porque a inconsistência entre os termos utilizados na marcação pode tornar muito difícil para um pesquisador ter certeza de que todos os itens relevantes foram encontrados.

Profundidade (especificidade) de marcação: como o usuário (classificador) deve ser específico para traduzir um conceito em tag? Recursos web podem ser marcados em diferentes níveis de especificidade, a partir de temas muito amplos ou tomados apenas a partir do título e resumo. A profundidade das tags refere-se a quantas tags existem em relação a um recurso web no sistema. (NORUZI, 2007).

Os plurais são um problema para sistemas menos inteligentes. Assim como o autor explica, a recuperação pode ser afetada caso algum recurso web for indexado no plural e o pesquisador procurar pelas palavras no singular e vice-versa. Dependendo do caso, o sistema pode até recuperar a informação, como, por

exemplo, plurais que não alteram tanto a palavra como *faca/facas*, onde só a letra “s” é acrescida. Agora, a palavra *coleção/coleções*, em que a palavra sofre alteração no final, dependendo de como foi indexada, pode não ser recuperada se apenas uma delas for colocada na indexação.

A polissemia a que o autor se refere seria utilizar da mesma palavra em situações diferentes. Como exemplo, utilizamos a palavra *colher*. Pode ser a colher, como *talher*, ou *colher*, do verbo *colher* (*colheita*). Este problema é comum acontecer, seja em sistemas de classificação hierárquicos ou sistemas de indexação livre, pois se refere a um problema gramatical das línguas em geral.

A sinonímia provoca insegurança no pesquisador, pois havendo este problema, o pesquisador acaba insatisfeito com os resultados de sua pesquisa, caso procure por apenas uma das palavras que contenham o mesmo significado. Podemos usar como exemplo, as palavras “*menina*”, “*garota*” e “*guria*”. São palavras distintas, mas utilizadas para a mesma coisa, aqui todas correspondem e nos remetem à figura de uma jovem.

A especificidade da marcação é atribuída ao usuário, de onde ele irá extrair etiquetas que possam descrever o objeto indexado. Como exemplo, um usuário que irá classificar o link de um artigo em seu *bookmark*² poderá indexá-lo apenas pelo título, autor, tema do artigo, como também poderá utilizar-se de palavras retiradas deste artigo, que ele julgar pertinente para uma posterior recuperação deste link. A profundidade da marcação é relacionada à exaustividade, quantas tags serão utilizadas para descrever o conteúdo de um documento digital.

Um dos sites pioneiros a incorporar a folksonomia foi o *Del.icio.us*.³, *bookmark* pessoal em que o usuário pode relacionar links de seu interesse e classificá-los por assunto. Também surgiram o *Youtube*⁴, que permite aos usuários compartilhar vídeos e a busca é feita através das tags utilizadas no momento do *upload*, e o *Flickr*⁵, site de compartilhamento de imagens em que os usuários determinam tags que remetam à pesquisa de outros usuários. Blogs também utilizam este tipo de

² *Bookmark* é uma ferramenta presente em todos os *browsers* atuais que serve como um bloco virtual de anotações. Nele, o usuário guarda os endereços que mais lhe interessam para poder acessá-los quando quiser. Também conhecidos como favoritos (AQUINO, 2007)

³ Disponível em: <<http://delicious.com/>>. Acesso em: 24 abr. 2012.

⁴ Disponível em: <<http://youtube.com/>>. Acesso em: 24 abr. 2012.

⁵ Disponível em: <<http://flickr.com/>>. Acesso em: 24 abr. 2012.

indexação, onde os autores classificam suas entradas ou artigos em várias “tags” (categorias) ao mesmo tempo.

Ultimamente, as redes sociais também adaptaram esta forma de indexar, como o *Twitter*⁶, microblog em que as informações publicadas são em forma de “tweets”, que são mensagens em até 140 caracteres. Inicialmente, a recuperação da informação era por meio das *hashtags*, que são palavras antecedidas por “#” como, por exemplo, #feriado, #copadomundo, #show. Atualmente, palavras soltas também são recuperadas. A recuperação é dada pelos usuários que utilizaram as palavras em seus tweets.

Neste trabalho, iremos trabalhar com a rede social *Tumblr* e como seus usuários indexam o conteúdo fotográfico postado em seus perfis. Para isso, será utilizado o estudo do autor Cañada (2006), sendo assim, as tags dos usuários do Tumblr serão categorizadas de acordo com estilos propostos pelo autor.

3.3.1 Estilos de etiquetagem

Neste trabalho serão utilizadas as constatações do autor Cañada (2006) sobre os estilos de etiquetagem na internet. Cañada utilizou-se da observação de etiquetagens utilizadas em sites como o *Flickr* e o *del.icio.us*, e propôs quatro estilos de etiquetagem sendo estes: egoísta, amigável, altruísta e populista. Neles, o autor classifica o benefício social ao grande grupo de usuários e a motivação que leva o usuário a etiquetar com cada estilo. Também encontra-se na literatura pesquisada trabalho que utiliza a abordagem de Cañada (2006) para classificação de fotografias do Flickr (CALDAS; MOREIRA, 2009).

O etiquetado egoísta consiste em classificar um documento eletrônico com etiquetas em que apenas o indexador possa recuperar. Cañada (2006) explica que: “Usamos etiquetas muito significativas para nós mesmos, mas carentes de significado fora de seu contexto pessoal. Exemplos: ‘trabalho’, ‘férias’, ‘leitura pendente’, ‘para Juan’, ‘oficina’”. O benefício social é de médio a baixo, segundo o autor, pois com o aumento do número de etiquetas utilizadas pelo usuário, surge a necessidade de organizá-las da melhor maneira. A motivação do etiquetado egoísta,

⁶ Disponível em: <<http://twitter.com/about>>. Acesso em: 24 abr. 2012.

segundo Cañada (2006), é alta, uma vez que não precisa de nenhum esforço intelectual do usuário e a ideia é de utilização apenas para benefício próprio.

O estilo de etiquetar amigável é baseado no processo de compartilhar o conteúdo do documento eletrônico a um grupo restrito de usuários, sejam eles amigos, família, colegas de trabalho, grupos de estudo, etc. Cañada (2006) constata que quando “[...] etiquetamos algo para compartilhá-lo com outras pessoas de nosso círculo (amigos, colegas, companheiros) e usamos termos consensuais e desconhecidos para o resto. Exemplos: ‘desconferência’, ‘cadius’, ‘asobine’”. Para o autor, o benefício social é alto, pois é muito importante para o fluxo da informação dentro do grupo. Em alguns casos, o tema usado pelo grupo pode tornar-se interessante para o público em geral, fazendo com que haja relevância do termo. A motivação é alta, pois reforça o sentimento de comunidade entre um grupo.

O estilo altruísta é o que mais agrega o sentimento de compartilhar o conteúdo com o grande grupo geral de usuários. O autor revela que é quando:

Etiquetamos algo com o desejo sincero de compartilhá-lo com o resto do mundo, elegendo as etiquetas que acreditamos que sejam mais descritivas, conhecidas e geralmente aceitas. Exemplos: ‘música’, ‘anos 60’, ‘clássicos’, ‘pop’, ‘the beatles’. (CAÑADA, 2006).

O benefício social desta etiquetagem é muito alto e é a forma que mais contribui para a recuperação da informação do conjunto total de usuários. Porém, segundo o autor, a motivação é baixa, pois é mais trabalhoso e exige esforço intelectual do usuário. Geralmente, esta motivação altruísta surge da alta relação do usuário com o conteúdo que compartilha.

E, por fim, a etiquetagem populista consiste em tornar o conteúdo indexado mais interessante, com a finalidade de obter mais acessos. Explica que isso ocorra quando “etiquetamos algo para fazê-lo mais atrativo e que obtenha visitas. Para isso usamos etiquetas que sejam populares e chamativas para a maioria. Exemplos: ‘muito bom’, ‘sexo’, ‘super interessante’, ‘top10’” (CAÑADA, 2006). O benefício social, segundo o autor, é nulo, sendo considerado lixo eletrônico. Acaba tornando-se inútil para a recuperação da informação tanto para o usuário que a utiliza, quanto para o usuário que realiza a busca. A motivação é alta, pois quem recorre a esse estilo de etiquetagem busca um benefício direto e evidente, como, por exemplo, ter muitos acessos a um conteúdo que não contenha nada do contexto de suas *tags*.

Cañada (2006) ressalta que não existe um indexador pré-disposto a ser egoísta ou altruísta, dependendo do que cada um irá indexar em determinado momento. Esclarecendo isso, Cañada (2006) diz que:

Não que haja “etiquetadores egoístas”, pois às vezes etiquetamos as coisas de modo egoísta. Depende do momento e, sobre tudo, do documento que vamos etiquetar. O estilo que usamos pode ser fruto de um esforço consciente ou voluntário (etiquetado altruísta, etiquetado populista) ou vem a surgir de forma mais natural (etiquetado egoísta, etiquetado amiguista). (CAÑADA, 2006).

Entretanto é difícil encontrarmos a utilização de apenas um estilo de etiquetagem. O usuário indexador determinará consciente ou inconscientemente suas *tags* de acordo com o documento que está disponibilizando. Cañada ressalta que: “[...] se em um sistema vão predominar as formas egoístas e amiguistas, pode ser que tenhamos muitos recursos ‘mal etiquetados’. Quem sabe seja necessário estabelecer mecanismos de equilíbrio para corrigir essa anomalia ‘antisocial’” (CAÑADA, 2006).

3.4 TUMBLR

Antes de entendermos como funciona a rede social Tumblr, é preciso esclarecer o que são as redes sociais. As redes sociais podem funcionar como formas de representação dos relacionamentos afetivos ou profissionais dos seres humanos entre si ou entre seus agrupamentos de interesses mútuos. Na internet, as redes sociais podem ser constituídas e mantidas através dos sites de redes sociais.

Estes sites de redes sociais são espaços que dão suporte ao surgimento de redes sociais, mas que vão depender de como são efetivamente apropriados para que se verifique ou não a existência de redes sociais. (BOYD; ELLISON, 2007).

A rede tem como característica o compartilhamento de ideias entre pessoas que possuem interesses e objetivos em comum. Assim, um grupo de discussão é composto por indivíduos que possuem identidades semelhantes.

Segundo Boyd e Ellison, os sites de redes sociais são ferramentas na internet que permitem aos indivíduos:

- a) construir um perfil público ou semi-público que fica restrito a um sistema limitado;
- b) articular uma lista de outros usuários com quem estes sujeitos vão partilhar uma conexão;
- c) permitir que sejam visualizadas as listas de conexões desenvolvidas por todos os usuários do sistema. (BOYD; ELLISON, 2007)

O *Tumblr*⁷ (lê-se *tãmlr*) possui todas essas características. Denominado também como um microblog, o Tumblr é uma rede social de compartilhamento de links, vídeos, textos, imagens e áudio. Criado em 2007 por David Karpo, o *Tumblr* é uma das redes sociais que mais crescem na atualidade. Na página sobre o site⁸ existe uma contagem de quantas visitas o site recebeu no mês, quantas postagens foram feitas no dia etc. Os números são exorbitantes, sempre na base dos milhões. Recentemente o site está disponível na língua portuguesa brasileira.

De fácil manuseio e atrativo em sua estética, o *Tumblr* ganha seus usuários com a simplicidade. O cadastro é rápido, simples e fácil. Basta o usuário digitar o e-mail pessoal, senha e *username* (Figura 1). O sistema irá pedir que o usuário coloque sua idade e dispõe os termos de uso do site, conforme Figura 2. Logo após, o sistema induz o usuário a customizar o seu perfil com temas já definidos, ou *HTMLs* criadas por outros usuários e também disponíveis em outros sites, como também tem a opção de criar o seu próprio *layout* (Figura 3).

Figura 1 - *Tumblr* login.

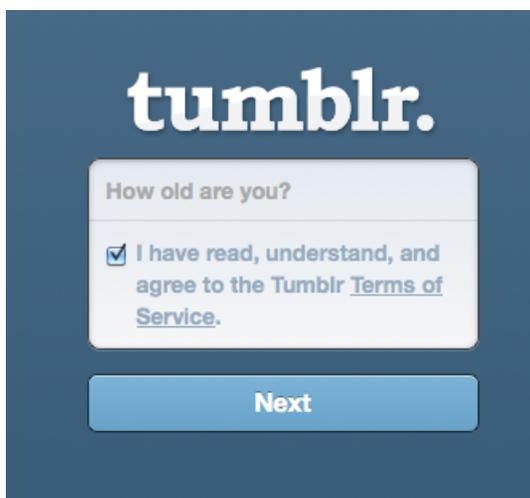


⁷ Disponível em: <<http://www.tumblr.com>>. Acesso em: 14 abr. 2012.

⁸ Disponível em: <<http://tumblr.com/about>>. Acesso em: 14 abr. 2012.

Fonte: *Tumblr*, 2012.

Figura 2 - Termos de uso



tumblr.

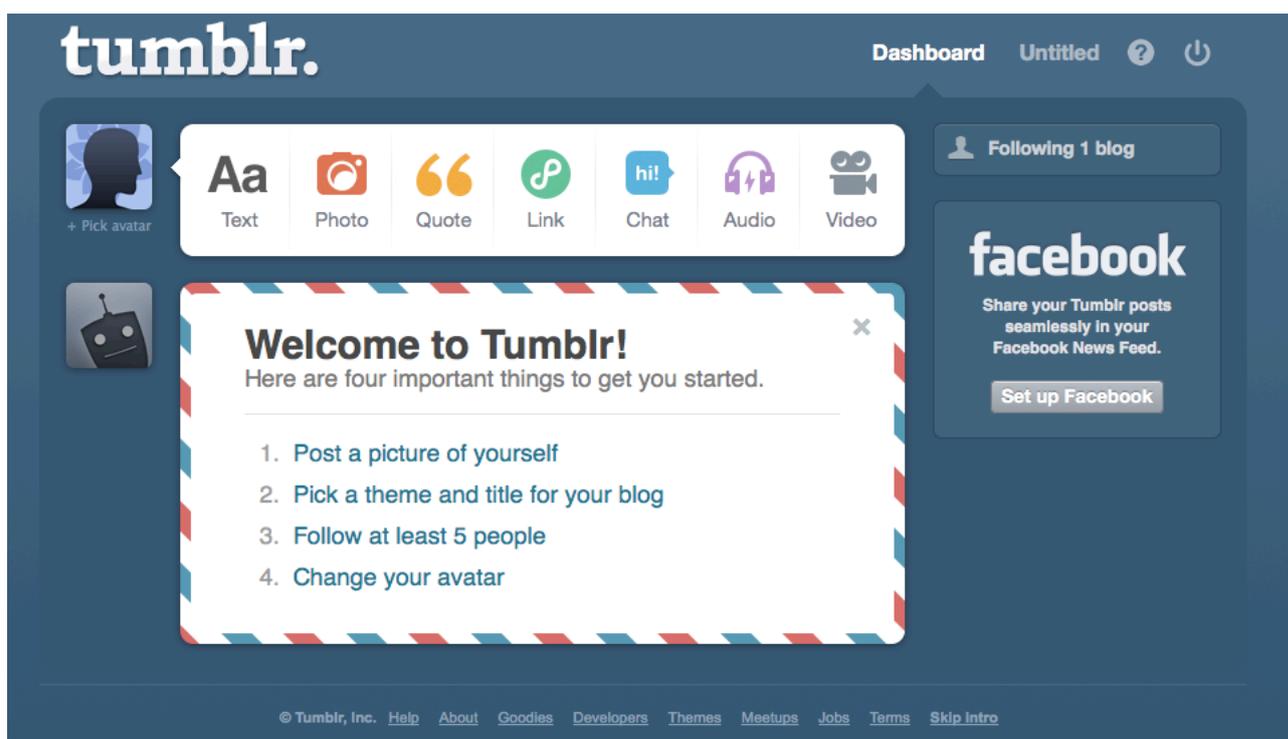
How old are you?

I have read, understand, and agree to the Tumblr [Terms of Service](#).

Next

Fonte: *Tumblr*, 2012.

Figura 3 – Bem-vindo ao *Tumblr*.



Fonte: *Tumblr*, 2012.

Como rede social, o sistema do site é parecido com o *Twitter*, em que o usuário pode “seguir” (*follow*) outros usuários, acompanhando pelo painel de controle

(*dashboard*) todo o conteúdo que é postado. O usuário não precisa atualizar a página para ver se há *posts* novos: um balão vermelho aparece no topo da página dando os números de novas postagens.

O compartilhamento de conteúdo do site é feito através da função *reblog*, ou seja, possibilita que o usuário poste novamente o que outro usuário publicou. A *reblogagem* pode ser feita por qualquer usuário do site, não precisa que um siga o outro. O crédito das postagens é feito automaticamente no ato da *reblogagem*, juntamente com uma lista de *notes* (notas) que mostram todos os outros que *reblogaram* aquilo também.

Há também a possibilidade de favoritar (Figura 4), que é através da figura de um coração, também nomeado de *like* (gostar, curtir), sendo para isso criada uma lista com todas as publicações que o usuário “curtiu”.

Figura 4 - Favoritar

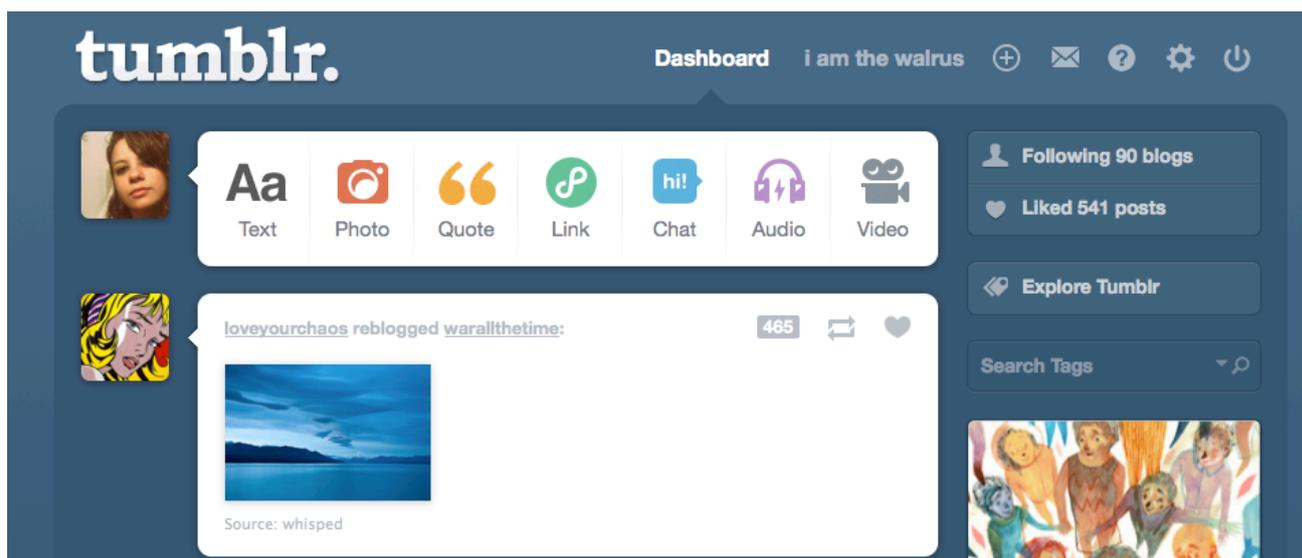


Fonte: Tumblr, 2012.

O *Tumblr* utiliza um sistema de *tags* para recuperar as informações que seus usuários postam. O sistema de *tags* é bastante aprofundado no site, tendo muitas dicas e suporte para os usuários. Entretanto, este sistema passa muitas vezes despercebido pelo usuário.

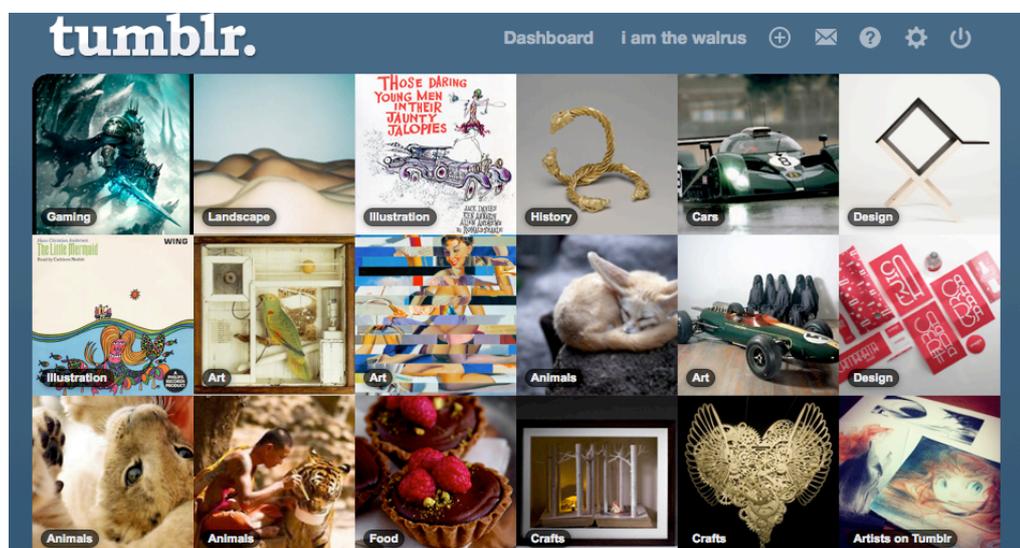
O site dispõe de uma sessão em que mostra as *tags* mais populares do momento. Para chegar nesta sessão, basta clicar em “*Explore Tumblr*” (Figura 5). A sessão (Figura 6) mostra de maneira simples todos os assuntos mais populares do momento como se fosse uma *tag cloud*, porém, formatado de outra maneira.

Figura 5 - Perfil e opções



Fonte: Tumblr, 2012.

Figura 6 - Explorando



Fonte: Tumblr, 2012.

O site⁹ também mostra em tempo real as *tags* mais populares (Figura 7), postagens com muita atividade (*reblogagens*) e opiniões. Também mostra o editor principal, que é um usuário escolhido por um “robô” do sistema *Tumblr* associado às *tags* que o sujeito incorpora às suas postagens.

⁹ Disponível em: <http://www.tumblr.com/docs/es/editor_faq>. Acesso em: 14 maio 2012.

Figura 7 - Etiquetas populares

ETIQUETAS POPULARES	OPINIÃO ▾	ACTIVIDAD	EDITOR PRINCIPAL
LOL	●●●●●		drinkyourjuice
Fashion	●●●●●		wetheurban
GIF	●●●●●		zbags
Vintage	●●●●●		hoodoothatvoodoo
Animals	●●●●●		samaralex
Landscape	●●●●●		kari-shma
Art	●●●●●		jacobvanloon
Illustration	●●●●●		theartofanimation
Film	●●●●●		thegoodfilms
Black and White	●●●●●		minusmanhattan
Makeup	●●●●●		makeuptips-
Tattoos	●●●●●		delacroix

Fonte: Tumblr, 2012.

A criação de *tags* (Figura 8) é feita na hora da postagem na parte direita da tela. Não é necessário colocar as tags entre aspas ou antecedida por um sustenido (*hashtag*). O usuário pode escrever a palavra que quiser colocando um ponto final ao fim de cada *tag* ou apenas teclando *enter*.

Figura 8 - Criação de tags

The screenshot shows the Tumblr 'Upload a Photo' interface. The main area contains a file upload section with 'Choose File' and 'P1140903.jpeg', a caption field with the text 'Zack looking especially dapper.', and a 'Set a click-through link' option. On the right side, there is a sidebar for post settings, including 'Post to' (my blog), 'publish now', 'now' (post date), 'http://' (content source), a 'tags' input field, 'Set a custom post URL' (with the example '/post/123456/'), and checkboxes for 'Send to Facebook' and 'Let people photo reply'.

Fonte: Tumblr, 2012

As *tags* não aparecem em qualquer tema (*layout* de cada microblog), mas o usuário pode escolher temas que as mostrem. Independente disso, o usuário pode buscar as *tags* de qualquer publicação mesmo que o tema não as mostre. Na busca de *tags* do *site* as etiquetas aparecem antecedidas de um sustenido.

Também é possível que o sujeito utilize *links* nas publicações ou na descrição do microblog. Tais *links* levarão a todas as publicações que contenham alguma *tag* a qual identifique todos os posts, como, por exemplo, utilizar a *tag* “usuárioxyz” em todos os posts. O *link* ficaria: “<http://usuário.tumblr.com/tagged/usuárioxyz>”. Basta utilizar o endereço do blog e as *tags* específicas que o usuário quiser fazer referência.

As *tags* do *Tumblr* somente não funcionam com blogs protegidos com senha. Também existe a possibilidade de “seguir” uma *tag*. Esta forma facilita para que os usuários possam manter-se atualizados das últimas postagens da referida *tag*.

Visto os detalhes do site, o funcionamento do sistema de etiquetagem e suas peculiaridades, isto irá facilitar o entendimento da atual pesquisa.

4 METODOLOGIA

Este capítulo descreve as etapas e processos que foram seguidos a fim de atingir os objetivos propostos.

4.1 ABORDAGEM E TIPO DE PESQUISA

A abordagem deste estudo foi mista, tanto qualitativa como quantitativa. O estudo qualitativo costuma ser direcionado ao longo do seu desenvolvimento e não busca enumerar ou medir eventos, além de não utilizar ferramentas estatísticas para análise dos dados. Conforme Neves (1996), faz parte da pesquisa qualitativa a obtenção de dados descritivos mediante contato direto e interativo do pesquisador com a situação objeto de estudo.

O estudo quantitativo busca a explicação para as causas da pesquisa. Segundo Figueiredo (2004, p. 107-108), o estudo quantitativo “[...] associa análise estatística à investigação dos significados das relações humanas, privilegiando a melhor compreensão do tema a ser estudado, facilitando desta forma a interpretação dos dados obtidos”.

Enquanto o método quantitativo confirma ou generaliza os resultados e conclusões, o método qualitativo tende a interpretar a apresentação de resultados. Unir os dois métodos é um modo de atingir resultados com embasamentos mais consistentes em teorias objetivas e subjetivas. Minayo (2000, p. 22), explica que “[...] o conjunto de dados quantitativo e qualitativo, porém, não se opõem. Ao contrário, se complementam, pois a realidade abrangida por eles interage dinamicamente, excluindo qualquer dicotomia.”

O tipo de pesquisa será estudo de caso, visto que este estudo pretende: “[...] analisar em profundidade um programa, um fato, uma atividade, um processo ou uma ou mais pessoas.” (CRESWELL, 2007, p. 32).

Por isso, os dados coletados são descritivos e sua análise é baseada na teoria que compõe o referencial teórico, principalmente no que se refere a Cañada (2006).

4.2 OBJETOS DA PESQUISA

Os objetos da pesquisa foram fotografias aleatórias do site Tumblr escolhidos pela interface de recuperação de *tags* do site. Foram selecionados cinco usuários que etiquetaram fotografias semelhantes, de acordo com o termo pesquisado das categorias pré-estabelecidas pela autora, como é explicitado na próxima seção.

4.3 PLANO DE ANÁLISE DE DADOS

Segundo Lakatos e Marconi (2009, p. 167) a coleta de dados é a “[...] etapa da pesquisa em que se inicia a aplicação dos instrumentos elaborados e das técnicas selecionadas, a fim de se efetuar a coleta dos dados previstos.”

A técnica de coleta de dados utilizada foi a observação assistemática. A observação é considerada uma coleta de dados para conseguir informações sob determinados aspectos da realidade. Ajuda o pesquisador a: “[...] identificar e obter provas a respeito de objetivos sobre os quais os indivíduos não têm consciência, mas que orientam seu comportamento” (LAKATOS; MARCONI, 1996, p. 79). Na observação assistemática, o pesquisador pretende “recolher e registrar os fatos da realidade sem a utilização de meios técnicos especiais, ou seja, sem planejamento ou controle” (BONI; QUARESMA, 2005, p. 71).

Foi analisado como a interface de recuperação da informação do *site Tumblr* funciona, além de suas vantagens e desvantagens na visão da autora. Para a interface de recuperação, buscou-se analisar qualitativamente as características da busca, a precisão e a relação entre *tag* e conteúdo, de acordo com a observação e interpretação da autora.

Optou-se por utilizar apenas o conteúdo fotográfico do site, já que o mesmo dispõe de áudio, vídeos, textos e animações. Selecionou-se três categorias de tipos de fotografias, com base nas *tags* mais populares do site, na seção “*Explore*” (Figura 8). Buscou-se selecionar as três categorias que geralmente estão entre as dez mais populares, que foram: *animals*, *landscape* e *fashion*.

A partir dessas três categorias, buscou-se dois termos em inglês (pois é a língua que abrange o site) de acordo com cada categoria: na categoria *animals*, foram buscadas as *tags* “*dog*” e “*cat*”; na categoria *landscape*, foram buscadas as

tags “*sea*” e “*mountain*”; e, por fim, na categoria *fashion* foram buscadas as *tags* “*shoe*” e “*dress*”. Foram analisados cinco usuários que publicaram fotos semelhantes e colocaram *tags* em seus *posts*.

Por fim, as *tags* foram categorizadas, de acordo com a tipologia de etiquetagem dada por Cañada (2006), que são egoísta, amigável, altruísta e populista.

Após a análise e classificação das *tags* dadas pelos usuários do site, os dados são analisados quantitativamente, para definir como as *tags* são utilizadas de acordo com a teoria e com os sujeitos analisados, e qualitativamente, para entendimento dos aspectos relacionados à folksonomia.

O estudo quantitativo é apresentado a partir de três tabelas elaboradas pela autora: uma para classificação das *tags* utilizadas pelos usuários estudados de acordo com as categorias de Cañada (2006); outra para quantificar as *tags* por categoria; e por fim, uma tabela para o resultado final do número de *tags* por categoria, que envolve todas as *tags* pesquisadas.

O estudo qualitativo foi embasado nas categorias estipuladas por Cañada (2006), onde a autora analisou e interpretou cada *tag* utilizada pelos usuários estudados neste presente trabalho.

Devido ao pouco tempo de pesquisa, o período de coleta e análise de dados deu-se em uma semana, de 21/05/2012 a 28/05/2012.

5 ANÁLISE DE DADOS

A análise de dados prevista neste trabalho busca compreender a interface de recuperação da informação e a classificação de *tags* feita por usuários no site *Tumblr*. Neste capítulo serão apresentados os resultados das análises, bem como detalhes da observação e interpretação desses dados, baseadas no referencial teórico deste trabalho.

5.1 INTERFACE DE RECUPERAÇÃO DA INFORMAÇÃO NO TUMBLR

O site *Tumblr* utiliza como classificação e organização de seu conteúdo as *tags* inseridas pelos usuários no momento de uma postagem, o que é uma característica da folksonomia. Para analisar a interface de recuperação da informação, utilizou-se apenas do modo de recuperação de *tags*, em “Buscar marcadores” (Figura 10) na página principal do usuário. Algumas peculiaridades foram percebidas no momento da coleta das *tags* pré-definidas (ver seção 4.3), como as características da busca, a precisão e a relação entre *tag* e conteúdo.

Figura 9 – Buscar marcadores



Fonte: Tumblr, 2012.

As características da busca consistem em: (i) quais os tipos de documentos são recuperados (vídeos, imagens, animações, textos etc); (ii) a classificação da

busca (métodos de pesquisa, relevância, mais acessos); e (iii) o sistema utilizado pelo site para a recuperação da informação. Ressalta-se novamente que foi levada em consideração apenas a busca por *tags* específicas e não as categorias estabelecidas pelo site.

Percebeu-se que não existe um modo com que o usuário possa escolher que tipo de documento que ele quer recuperar. A busca consiste em colocar a *tag* a ser recuperada e analisar todos os objetos indexados que se utilizam desta mesma *tag*.

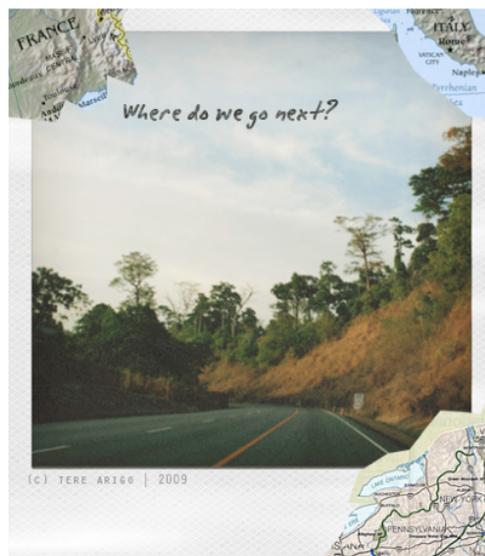
A classificação dos objetos recuperados a partir da busca é feita em ordem decrescente de data de postagem. Também não há outras opções para pesquisar com outra classificação, como por exemplo, por relevância, mais acessos da *tag* em questão, somente por categorias estabelecidas pelo site de acordo com acessos e reblogagens na seção “Explorar o Tumblr” (Figura 9).

A interface utilizada pelo site na recuperação da informação é um pouco confuso. Não existe no site uma seção¹⁰ explicando ao usuário como funcionam as *tags*, como etiquetar e algumas outras peculiaridades. Também não informa como funciona essa interface de recuperação dessas *tags*. Entrou-se em contato com o site para saber mais detalhes sobre como é feito este interface/sistema de recuperação, porém, não obteve-se respostas concretas.

Foi constatado que muitas vezes a busca recupera informações a partir das reblogagens dos usuários. Por exemplo: o usuário X publica uma fotografia de um gato e coloca as seguintes *tags*: “gato, amarelo, gatinho, sono”. O usuário Y *rebloga* a mesma publicação, porém não coloca *tags*. O usuário Z quer fazer uma pesquisa sobre gatos no site, e procura como “gatinho”. O sistema busca as duas publicações, já que no momento em que o usuário Y *reblogou* a postagem do usuário X, o sistema importa automaticamente as *tags* da primeira publicação, embora elas não sejam visíveis. Assim como o usuário Y não usou as *tags*, outros usuários que forem *reblogar* a mesma postagem podem inserir *tags*, provocando uma grande exaustividade na indexação.

Para melhor entendimento, foi pesquisada uma *tag* aleatória, “*travel*”. O exemplo aqui ilustrado (Figura 10) não obteve *tags* a partir do último usuário que reblogou (Figura 11), porém, foi recuperado pois o primeiro usuário a postar a foto utilizou a *tag* “*travel*” (Figura 12).

¹⁰ Disponível em: <http://www.tumblr.com/docs/br/using_tags>. Acesso em: 23 maio 2012.

Figura 10 – *Travel*

Fonte: Tumblr, 2012

Figura 11 – Última postagem

no-winter-for-a-year:

travelingcolors:

Where do we go next? (by thresca)

everywhere

Posted on May 13th (7:23pm), 1 week ago

247 notes

reblogged from spheres

originally posted by travelingcolors

 leoisdrunkonlife liked this

 getoutthereandexplore reblogged this from travelingcolors

 dina-ngeles reblogged this from himehikari

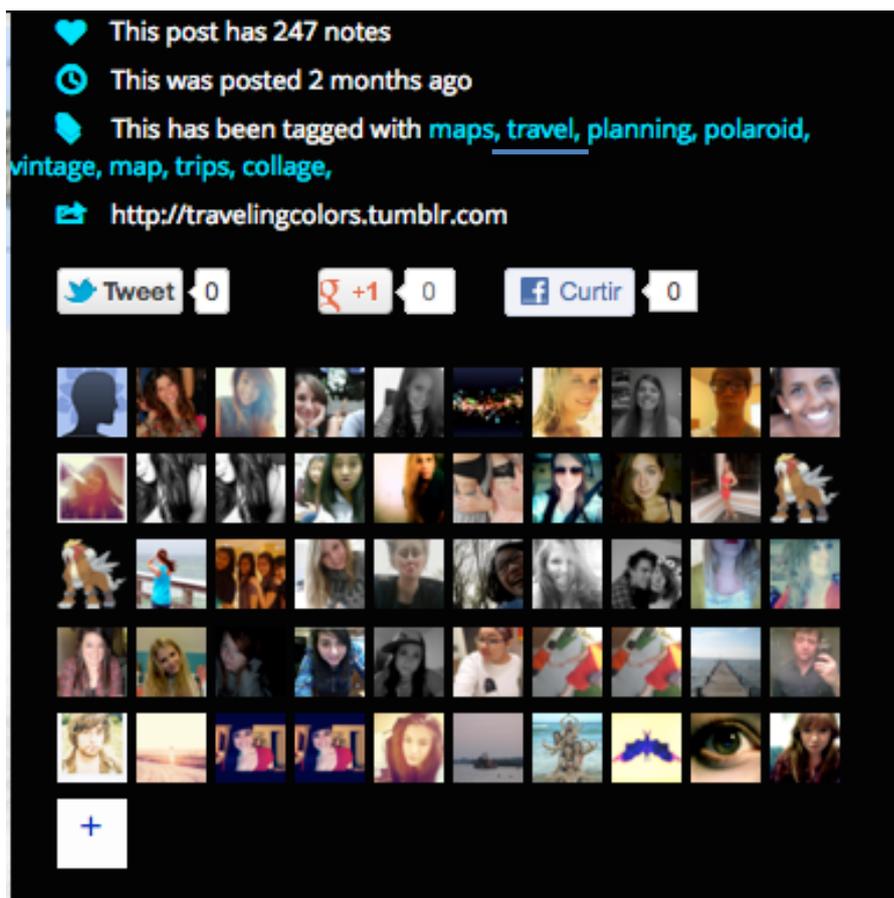
 stefgard liked this

 haileygoes reblogged this from maps-and-globes

 dog-master liked this

Fonte: Tumblr, 2012

Figura 12 – Primeira postagem



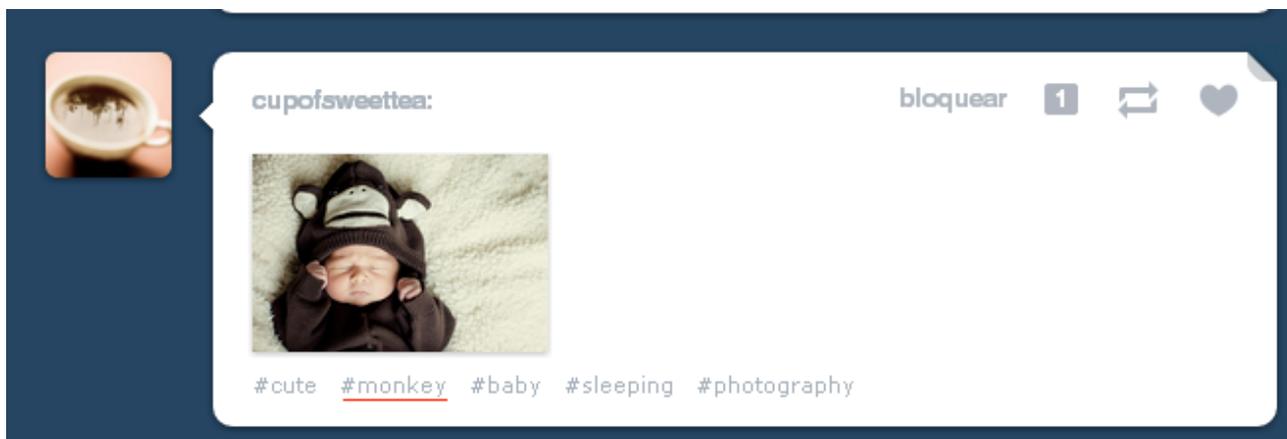
Fonte: Tumblr, 2012.

Durante a pesquisa, percebeu-se a ausência de *tags* em algumas fotografias que eram recuperadas por uma determinada etiqueta, muitas vezes de fotografias sendo publicadas pela primeira vez. Isso causou estranheza e curiosidade em entender como seria recuperada a informação daquela imagem. Num primeiro momento, criou-se a hipótese de existir um sistema de reconhecimento de fotografias. Entretanto, não houve como confirmar com o suporte técnico do *site*. Fazendo algumas buscas, foi afastada esta hipótese, pois algumas fotos de determinados objetos, como por exemplo, roupas, dependendo do ângulo, o sistema de reconhecimento de fotografias seria inviável.

Contudo, descobriu-se que clicando na *tag* abaixo da foto, o sistema busca todas as imagens que utilizaram da mesma etiqueta, retirando-a das demais *tags* indexadas. Como no exemplo abaixo, foi pesquisada a tag “*adorable*” (adorável). Foi selecionada aleatoriamente uma fotografia (Figura 13) que continha a *tag* “*monkey*” (macaco). Clicando nesta mesma *tag*, veio a mesma fotografia (Figura 14), porém

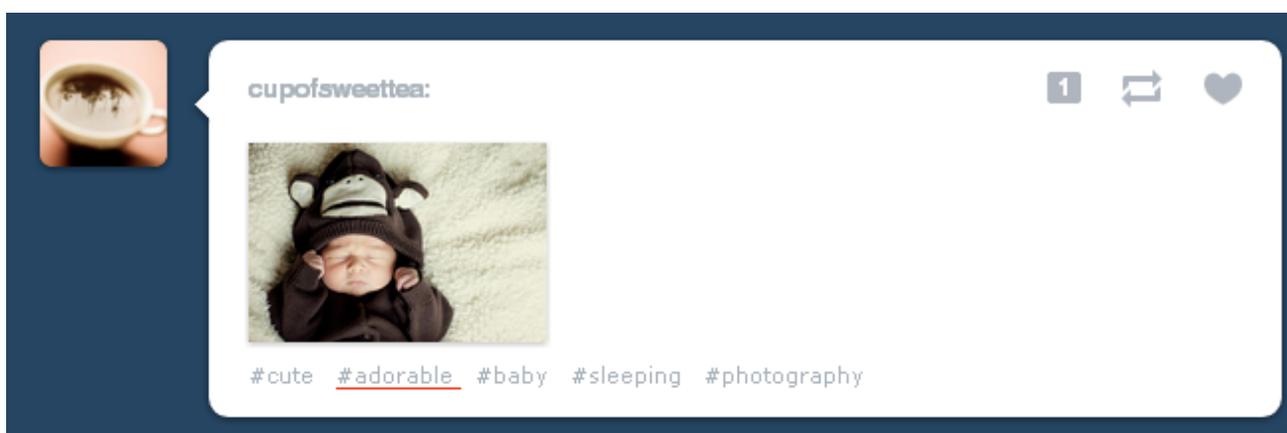
sem a *tag* “*monkey*”, e aparecendo a *tag* “*adorable*” que na primeira busca não aparecia. Ou seja, o sistema oculta a *tag* pesquisada.

Figura 13 - *Monkey*



Fonte: Tumblr, 2012.

Figura 14 - *Adorable*



Fonte: Tumblr, 2012.

Outra característica da interface de recuperação da informação é que não é possível procurar por mais de uma palavra na busca de etiquetas, como por exemplo “*monkey, adorable*”. Também não é possível intercalar entre duas *tags*, como exemplo “seguir” (ver seção 3.4) uma das *tags* e inserir outra na busca após. O sistema busca apenas a última *tag* pesquisada. Palavras compostas ou nomes próprios podem ser pesquisados, como “*Andy Warhol, Frida Kahlo*”.

A precisão da recuperação da informação no site deixa a desejar pelo motivo de que nem todo usuário etiqueta o conteúdo que está postando, o que faz compreender que nem todo conteúdo é recuperado já que com o teste acima

percebeu-se que apenas as *tags* são recuperáveis, a descrição de qualquer postagem não. O fato também da pesquisa ser feita apenas a partir de uma *tag* por vez dificulta a precisão da busca.

A relação entre *tag* e conteúdo a partir das análises obtidas mostrou que, em sua maioria, as *tags* correspondiam ao seu conteúdo. O que acontece, algumas vezes, são os casos considerados lixo eletrônico ou *tags* populistas, segundo Cañada (2006), que são aquelas etiquetas em que se aproveitam de um determinado assunto para que haja mais acessos ao conteúdo indexado.

5.2 CLASSIFICAÇÃO DE TAGS

A classificação de *tags* partiu do pressuposto de que todas as *tags* pesquisadas foram utilizadas pelos usuários em questão, pois como mostrado no capítulo anterior, o sistema do site oculta a etiqueta pesquisada. Serão apresentadas em seções as grandes categorias em que as *tags* pesquisadas estão incluídas, representadas pelas tabelas elaboradas pela autora. As *tags* escolhidas foram: “*cat*” e “*dog*” na categoria *Animals*; “*mountain*” e “*sea*” na categoria *Landscape*; e “*dress*” e “*shoe*” na categoria *Fashion* (ver seção 4.3). As *tags* que foram pesquisadas pela autora para a análise foram classificadas como altruístas, que segundo Cañada (2006) são etiquetas utilizadas para compartilhar o conteúdo com o grande público. Ou seja, as *tags* utilizadas na busca de imagens foi satisfatória, pois a autora encontrou o que queria.

Foram analisadas as imagens em que cada usuário utilizou as *tags* estabelecidas e as demais indexadas no momento da publicação de seu *post* no *Tumblr*. As *tags* serão precedidas de sustenido, pois é como aparecem nas imagens.

5.2.1 Animals

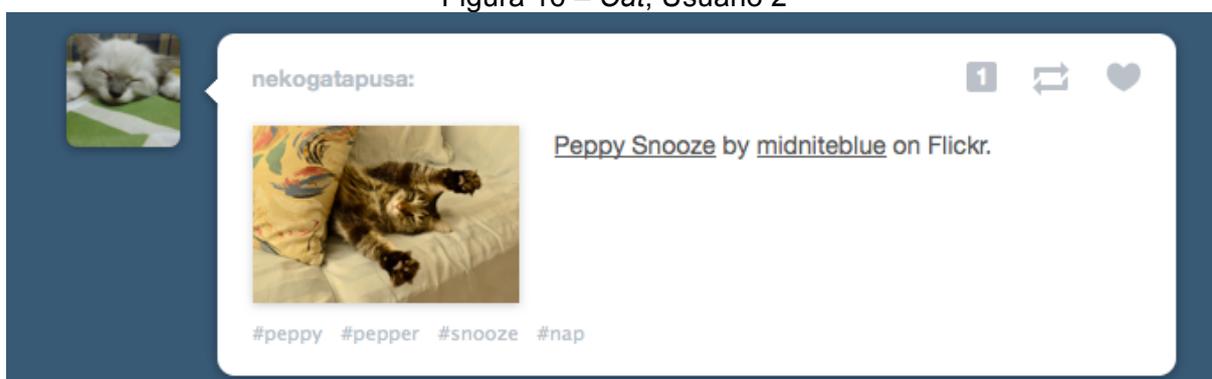
Abaixo são apresentadas as tabelas com as classificações de *tags* “*cat*” e “*dog*” para a categoria “*Animals*”.

Figura 15 – Cat, Usuário 1



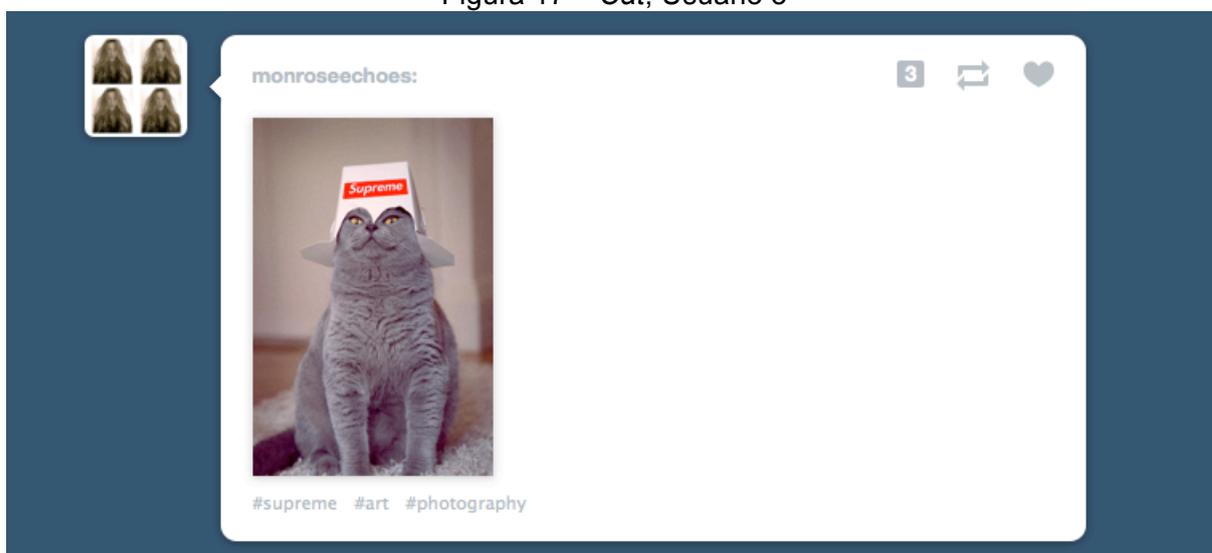
Fonte: Tumblr, 2012

Figura 16 – Cat, Usuário 2



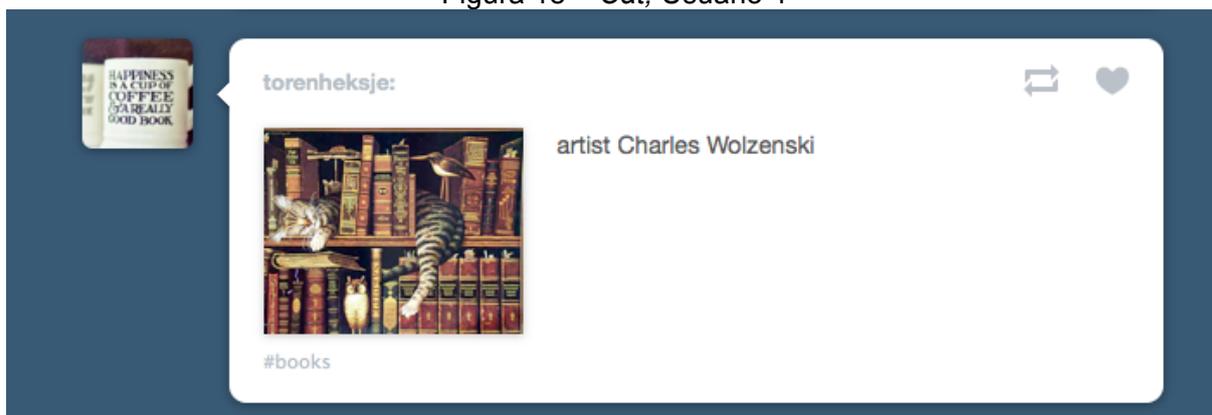
Fonte: Tumblr, 2012

Figura 17 – Cat, Usuário 3



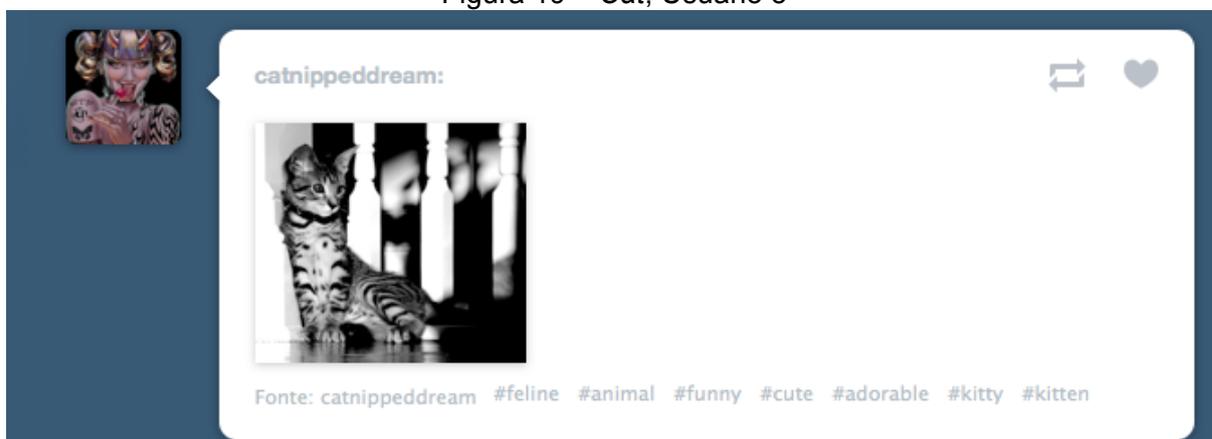
Fonte: Tumblr, 2012

Figura 18 – Cat, Usuário 4



Fonte: Tumblr, 2012

Figura 19 – Cat, Usuário 5



Fonte: Tumblr, 2012

Tabela 1: Classificação da imagem a partir da tag “cat”

	Egoísta	Amigável	Altruísta	Populista
Usuário 1	#light of my life; #boyfriend; #personal		#cat	
Usuário 2	#pepper; #peppy		#cat; #nap; #snooze	
Usuário 3	#art; #photography	#supreme;	#cat;	
Usuário 4			#cat; #books	
Usuário 5	#funny; #adorable		#cat; #feline; #kitty; #kitten	

Fonte: Autora

Tabela 2: Quantidade de tags atribuídas à imagem a partir da tag “cat”

	Egoísta	Amigável	Altruísta	Populista
Usuário 1	3	-	1	-
Usuário 2	2	-	3	-
Usuário 3	2	1	1	-
Usuário 4	-	-	2	-
Usuário 5	2	-	4	--
Resultado	9	1	11	0

Fonte: Autora

Foram classificadas como egoístas as tags:

- *#light of my life* (luz da minha vida), *#boyfriend* (namorado/amigo) e *#personal* (pessoal) utilizadas pelo Usuário 1;
- *#pepper* e *#peppy*, que pela interpretação da autora são nomes do animal, indexado pelo Usuário 2;
- *#art* (arte) e *#photography* (fotografia) utilizadas pelo usuário 3;
- e as tags *#funny* (engraçado) e *#adorable* (adorável) usadas pelo Usuário 5.

As *tags* mencionadas, na interpretação da autora, são significativas apenas para quem está indexando e também são julgamentos do próprio etiquetador para a imagem. Outra pessoa pode não achar, por exemplo, o animal da foto “engraçado”. Cañada (2006) explica que com o tempo, os usuários tendem a sistematizar o seu estilo com etiquetas temáticas mais gerais, no caso aqui as *tags* “*art*”, “*photography*”, “*funny*” e “*adorable*”. As *tags* “*pepper*” e “*peppy*” são nomes próprios no caso desta fotografia, pois são os nomes/apelidos do gato, indicando que somente o etiquetador saiba o significado. “*Pepper*” também significa pimenta em inglês, o que pode causar algum ruído na recuperação da informação para quem está buscando conteúdo relacionado à pimenta. Sendo considerada a primeira possibilidade, optou-se por classificá-las como o nome ou apelido do animal.

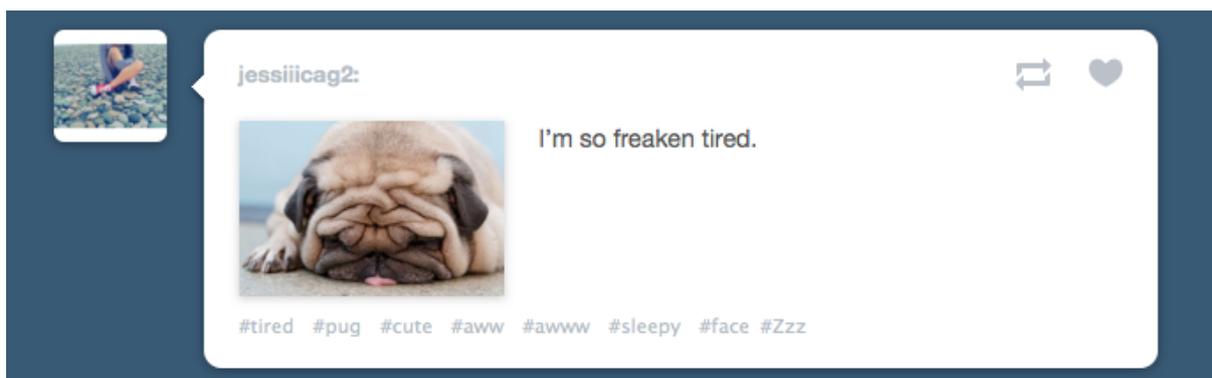
A *tag* *#supreme* (no português supremo, porém no contexto é uma marca de roupas) utilizada pelo Usuário 3, foi classificada como amigável, pois um determinado grupo pode procurar pela marca de roupas. Segundo Cañada (2006), uma *tag* amigável é para um grupo restrito de usuários, no caso, um grupo restrito de pessoas que conhecem a marca *Supreme* e que possam querer pesquisá-la.

Como altruístas, foram classificadas as *tags*:

- *#cat* (gato) utilizada por todos usuários;
- *#nap* (sesta), *#snooze* (soneca) pelo Usuário 2;
- *#books* (livros) utilizada pelo Usuário 4;
- e *#feline* (felino), *#kitty* (gatinho) e *#kitten* (gatinho) usadas pelo Usuário 5.

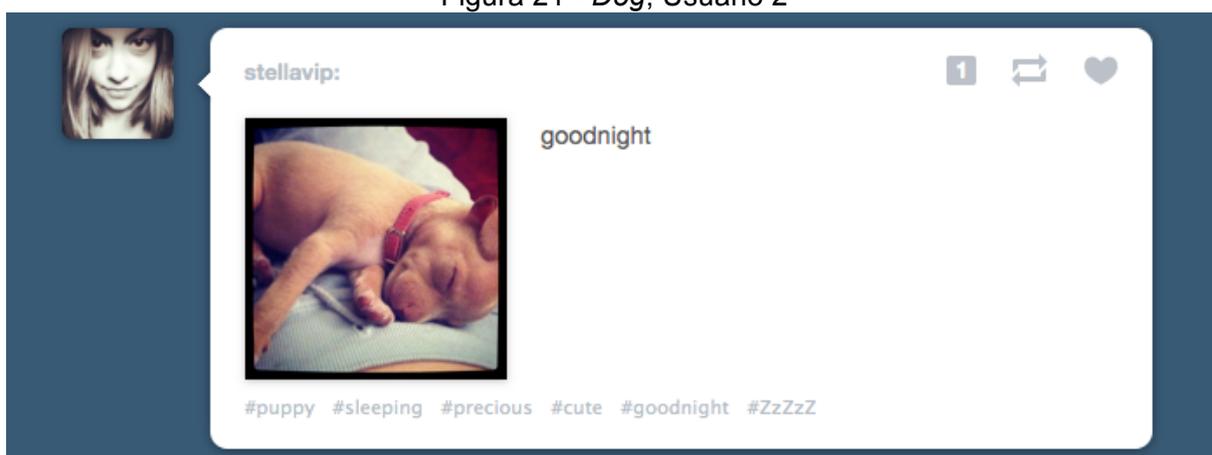
Estas *tags* têm alto benefício social, pois segundo Cañada (2006), são *tags* que contribuem para facilitar a recuperação da informação. Todas as etiquetas correspondem às imagens indexadas, ou seja, o usuário que pesquisar por qualquer uma das *tags* encontrará o que está buscando.

Percebeu-se que as *tags* altruístas e egoístas foram mais utilizadas para as imagens indexadas com a *tag* “*cat*”.



Fonte: Tumblr, 2012

Figura 21 - Dog, Usuário 2



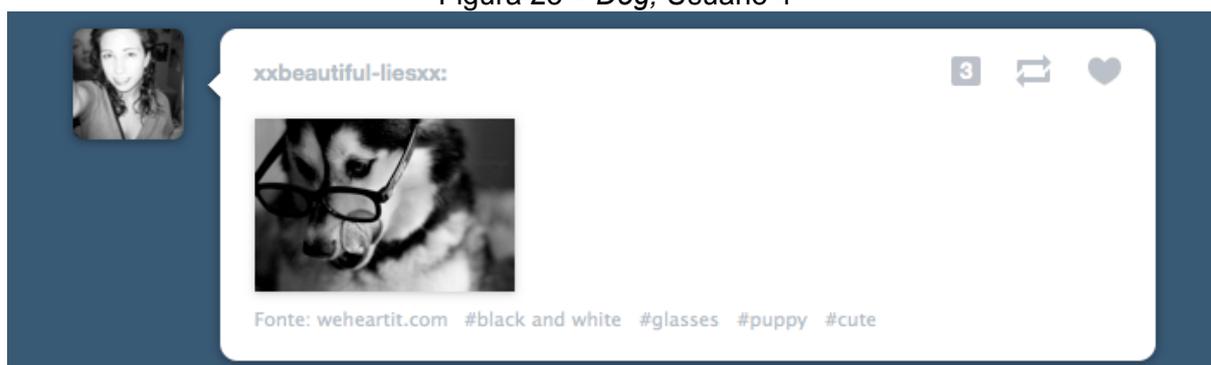
Fonte: Tumblr, 2012

Figura 22 – Dog, Usuário 3



Fonte: Tumblr, 2012

Figura 23 – Dog, Usuário 4



Fonte: Tumblr, 2012

Figura 24 - Dog, Usuário 5



Fonte: Tumblr, 2012

Tabela 3: Classificação da imagem a partir da tag “dog”

	Egoísta	Amigável	Altruísta	Populista
Usuário 1	#aww; #awww; #Zzz; #cute		#dog; #pug; #sleepy; #tired	#face
Usuário 2	#ZzZzZ; #cute; #precious;		#dog; #puppy; #sleeping	#goodnight;
Usuário 3			#dog; #bulldog; #english bulldog; #perro	#animal; #pet
Usuário 4	#cute		#dog; #glasses; #puppy; #black and white	
Usuário 5	#My Photo		#dog; #puppy; #Golden Retriever	#summer; #sun

Fonte: Autora

Tabela 4: Quantidade de tags atribuídas à imagem a partir da tag “dog”

	Egoísta	Amigável	Altruísta	Populista
Usuário 1	4	-	4	1
Usuário 2	3	-	3	1
Usuário 3	-	-	4	2
Usuário 4	1	-	4	-
Usuário 5	1	-	3	2
Resultado	9	0	18	6

Fonte: Autora

Para as imagens indexadas com a tag “dog” foram classificadas como egoístas as tags:

- *#aww*, *#awww* e *#zzz* pelo Usuário 1 e *#ZzZzZ* pelo Usuário 2, todas onomatopeias (as duas primeiras desconhecidas, as outras duas que indicam sono, dormir);
- *#cute* (fofo) usada pelos usuários 1, 2 e 4;
- *#precious* (precioso) usada pelo Usuário 2;
- e a tag *#My Photo* (minha foto) usada pelo usuário 5.

As tags remetem a algo que talvez nem o próprio usuário irá conseguir recuperar, como *#aww* e *#awww*, pois nota-se que são praticamente iguais, causando confusão pelo usuário indexador. Ainda assim são para benefício próprio, não se adequando a um contexto social. As tags “cute” e “precious” são subjetivas, pois são opiniões do próprio indexador.

As tags altruístas são:

- *#dog* (cachorro, cão) utilizada por todos os usuários;
- *#puppy* (filhote, cachorrinho) utilizada pelos usuários 2, 4 e 5;

- *#pug* (raça de cachorro), *#sleepy* (sonolento) e *#tired* (cansado) usadas pelo Usuário 1;
- *#sleeping* (dormindo) utilizadas pelo Usuário 2;
- *#bulldog* (raça de cachorro), *#english bulldog* (bulldog inglês, raça de cachorro), *#perro* (cachorro, em espanhol) usadas pelo usuário 3;
- *#glasses* (óculos), *#black and white* (preto e branco) utilizadas pelo Usuário 4;
- e *#Golden Retriever* (raça de cachorro) usadas pelo Usuário 5.

Cañada (2006) esclarece que as *tags* altruístas são mais descritivas. As *tags* referidas descrevem as imagens e detalham alguns dados como a raça dos animais, que facilita muito na recuperação da informação. “*Black and white*” apesar de ser uma *tag* bastante abrangente, foi classificada como altruísta pois a fotografia se enquadra nessa característica. “*Sleepy*” e “*tired*” são etiquetas que descrevem a situação do animal na foto. O usuário que buscar por essas palavras encontrará diversas situações, inclusive cachorros cansados e sonolentos.

Foram classificadas como populistas as etiquetas:

- *#face* (face, rosto) utilizada pelo Usuário 1;
- *#goodnight* (boa noite) usada pelo usuário 2;
- *#animal* e *#pet* (animal de estimação) usadas pelo Usuário 3;
- e *#summer* (verão) e *#sun* (sol) publicada pelo Usuário 5.

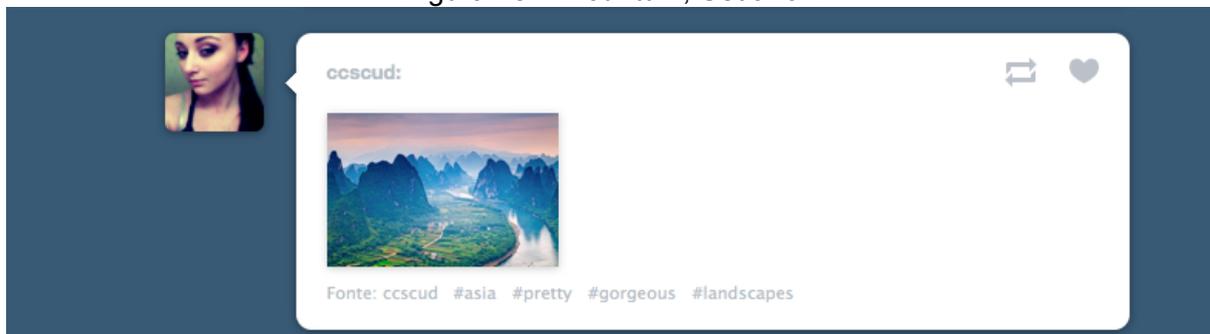
Cañada (2006) afirma que as *tags* populistas são indexadas para tornar o conteúdo mais interessante. As *tags* em questão têm baixo benefício social, pois são etiquetas que não descrevem as principais características da foto, “*summer*” e “*sun*” são *tags* que apenas acrescentam algumas informações de tempo à foto. Interpreta-se como *tags* que servem para buscar mais visualizações à imagem. “*Face*” e “*good night*” são *tags* que não atingem o conteúdo das fotos.

Observou-se que as *tags* egoístas e altruístas foram mais utilizadas nas imagens em que a *tag* “*dog*” foi atribuída.

5.2.2 Landscape

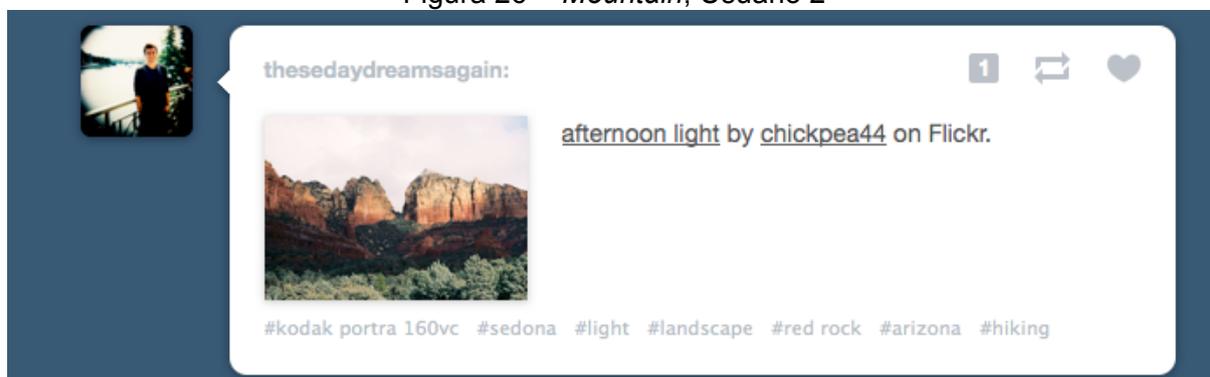
Abaixo são apresentadas as tabelas com as classificações de *tags* “*mountain*” e “*sea*” para a categoria “*Landscape*”.

Figura 25 – *Mountain*, Usuário 1



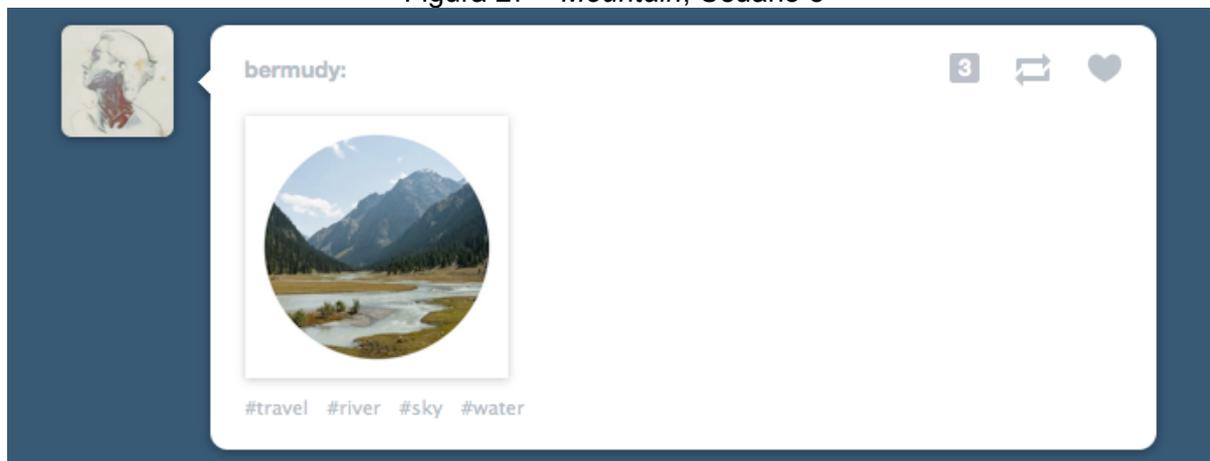
Fonte: Tumblr, 2012

Figura 26 - *Mountain*, Usuário 2



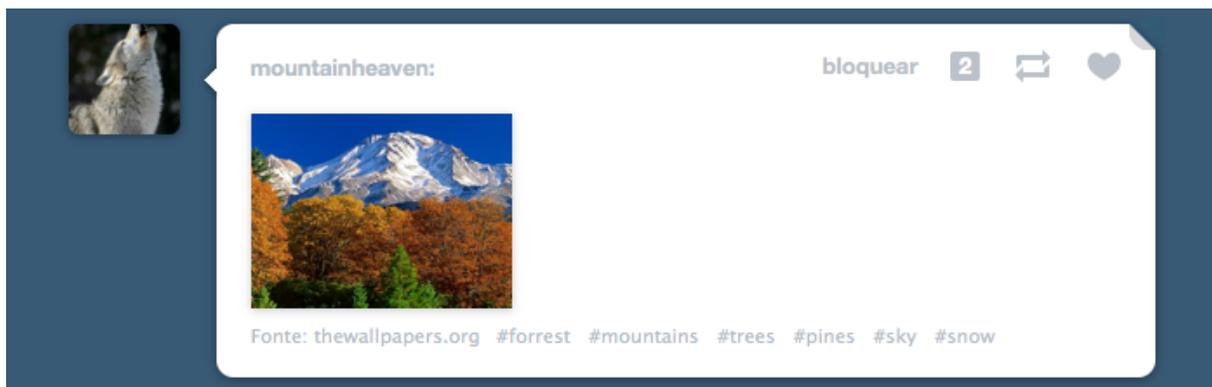
Fonte: Tumblr, 2012

Figura 27 – *Mountain*, Usuário 3



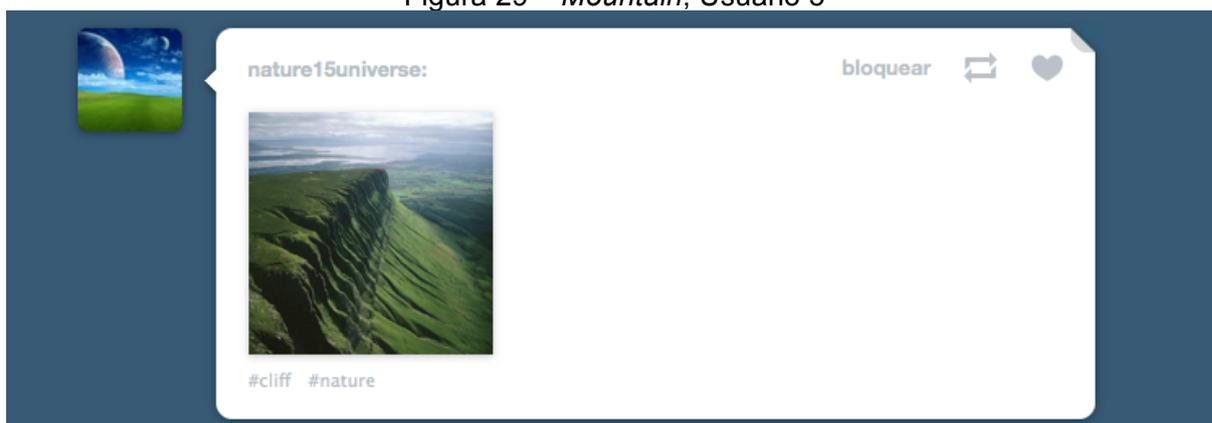
Fonte: Tumblr, 2012

Figura 28 – *Mountain*, Usuário 4



Fonte: Tumblr, 2012

Figura 29 – *Mountain*, Usuário 5



Fonte: Tumblr, 2012

Tabela 5: Classificação da imagem a partir da tag “*mountain*”

	Egoísta	Amigável	Altruísta	Populista
Usuário 1	#gorgeous; #pretty		#mountain; #asia; #landscapes	
Usuário 2		#kodak portra 160vc	#mountain; #sedona; #landscape; #red rock; #arizona	#hiking
Usuário 3			#mountain; #river; #sky	#water; #travel
Usuário 4			#mountain; #mountains; #forrest; #trees; #snow; #pines; #sky	
Usuário 5			#mountain; #cliff; #nature	

Fonte: Autora

Tabela 6: Quantidade de *tags* atribuídas à imagem a partir da *tag* “*mountain*”

	Egoísta	Amigável	Altruísta	Populista
Usuário 1	2	-	3	-
Usuário 2	-	1	5	1
Usuário 3	-	-	3	2
Usuário 4	-	-	7	-
Usuário 5	-	-	3	-
Resultado	2	1	21	3

Fonte: Autora

Para as imagens em que a etiqueta “*mountain*” foi atribuída, as *tags* classificadas como egoístas foram: *#gorgeous* (lindo, maravilhoso) e *#pretty* (bonito) utilizadas pelo usuário 1. As *tags* são concepções do próprio indexador, sendo assim opiniões próprias. Também são *tags* temáticas e reconhecidas pelo grande público, considerando-as então como *tags* egoístas, segundo Cañada (2006).

A *tag* *#kodak portra 160vc* (modelo de câmera fotográfica) foi considerada amigável, pois é reconhecida por um grupo restrito, que são fotógrafos ou apreciadores de fotografia. Segundo Cañada (2006), as *tags* amigáveis são classificadas de acordo com um determinado círculo e são de consenso com o meio e desconhecidas para o resto do público.

As *tags* classificadas como altruístas foram:

- *#mountain(s)* (montanha) usadas por todos os Usuários;
- *#landscape(s)* (paisagem) utilizadas pelos Usuários 1 e 2;
- *#sedona* (cidade do Arizona, nos EUA), *#red rock* (pedra vermelha); *#arizona* (estado dos EUA) utilizadas pelo Usuário 2;
- *#river* (rio) e *#sky* (céu) utilizadas pelo Usuário 3;
- *#forrest* (floresta), *#trees* (árvores), *#snow* (neve), *#pines* (pinheiros) e *#sky* (céu) utilizadas pelo Usuário 4;
- e *#cliff* (penhasco) e *#nature* (natureza) usadas pelo Usuário 5.

As *tags* em questão descrevem a maior parte de características das fotografias analisadas nesta categoria, por isto são altruístas, pois facilitam o trabalho do usuário pesquisador.

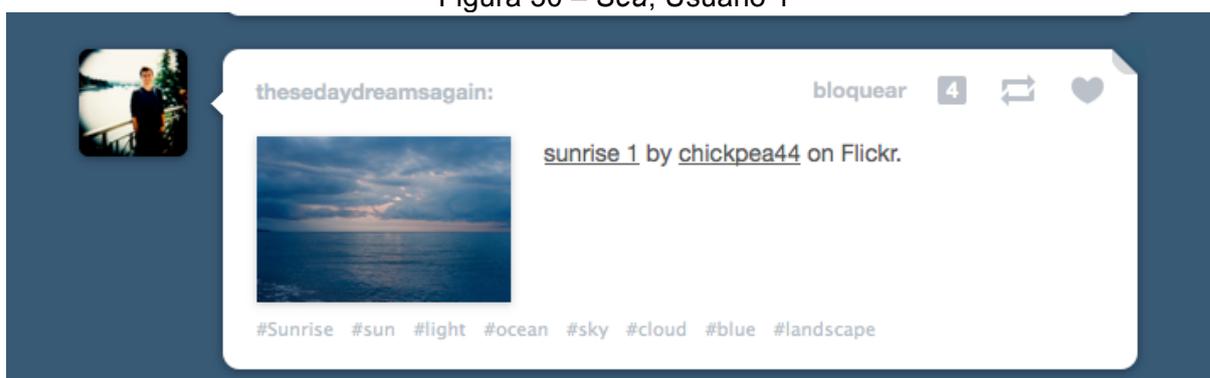
As *tags* populistas são:

- *#hiking* (caminhada) utilizada pelo usuário 2;
- e *#water* (água) e *#travel* (viagem) utilizadas pelo usuário 3.

Cañada (2006) observa que as *tags* populistas são etiquetas populares que são chamativas à maioria dos usuários. As *tags* acima citadas não fazem referência exclusivamente às imagens, elas podem ser subjetivas. Observou-se que as etiquetas apenas acrescentam informações desnecessárias e que podem causar ruído na recuperação da informação de outras pesquisas.

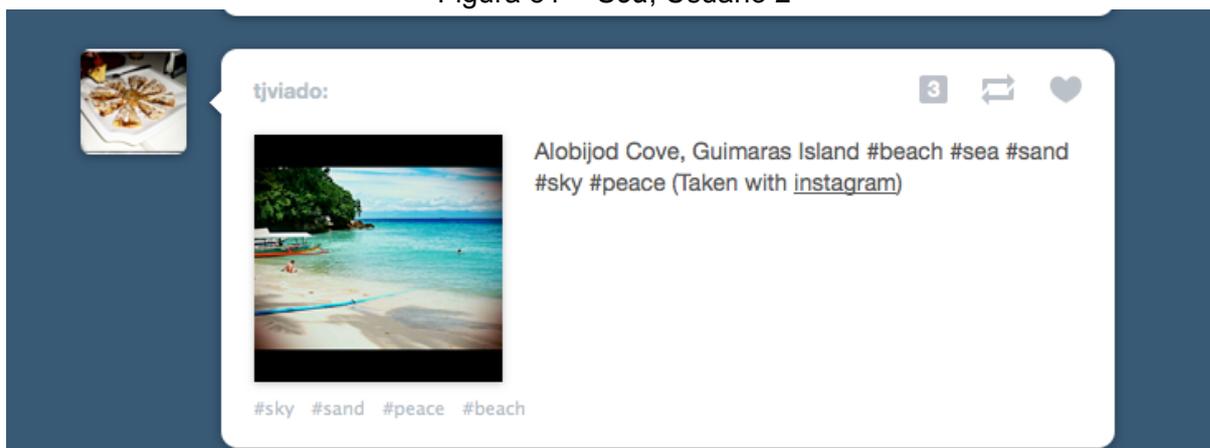
Para imagens que foram atribuídas a *tag* “*mountain*”, as *tags* altruístas predominaram perante as outras.

Figura 30 – Sea, Usuário 1



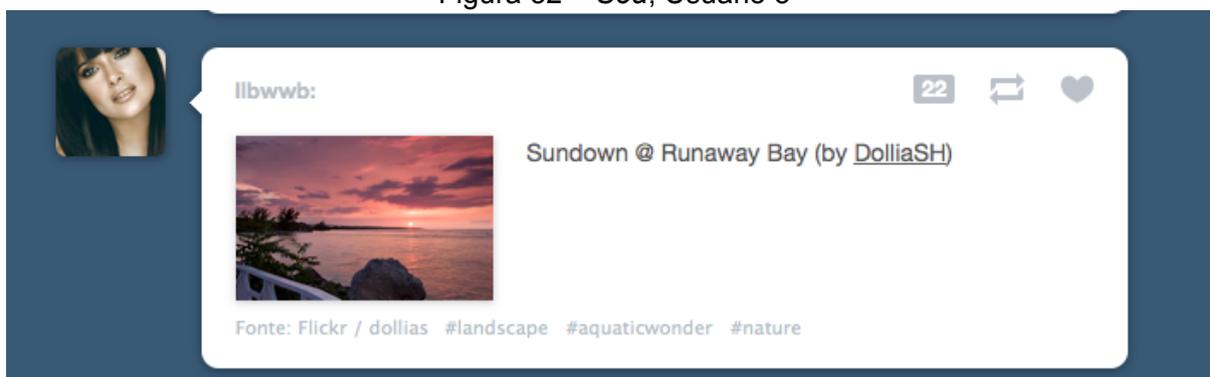
Fonte: Tumblr, 2012

Figura 31 - Sea, Usuário 2



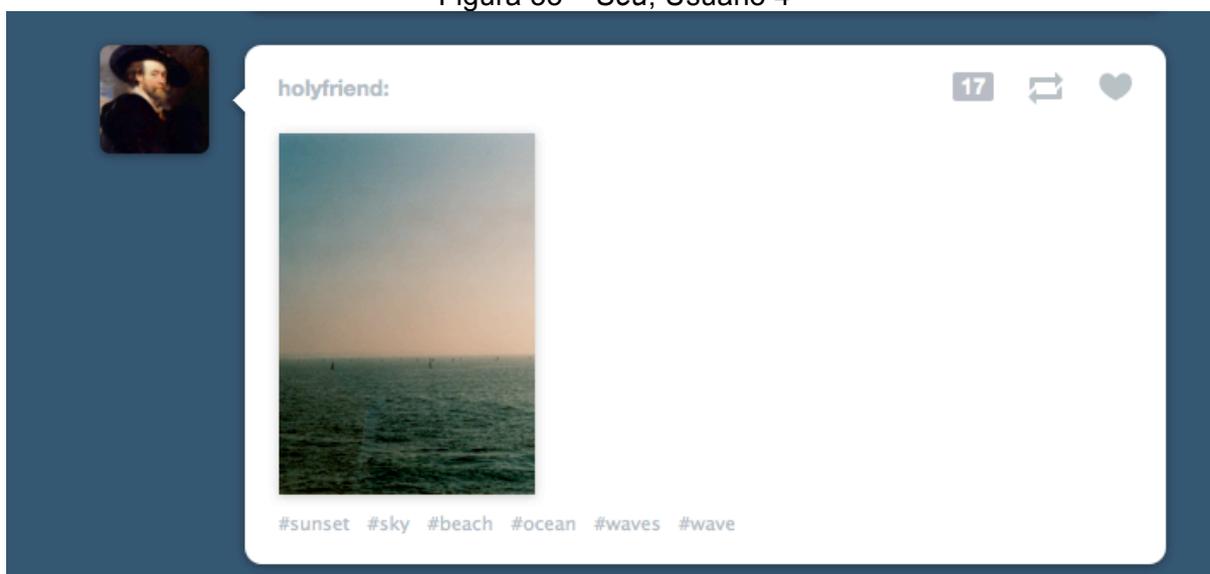
Fonte: Tumblr, 2012

Figura 32 – Sea, Usuário 3



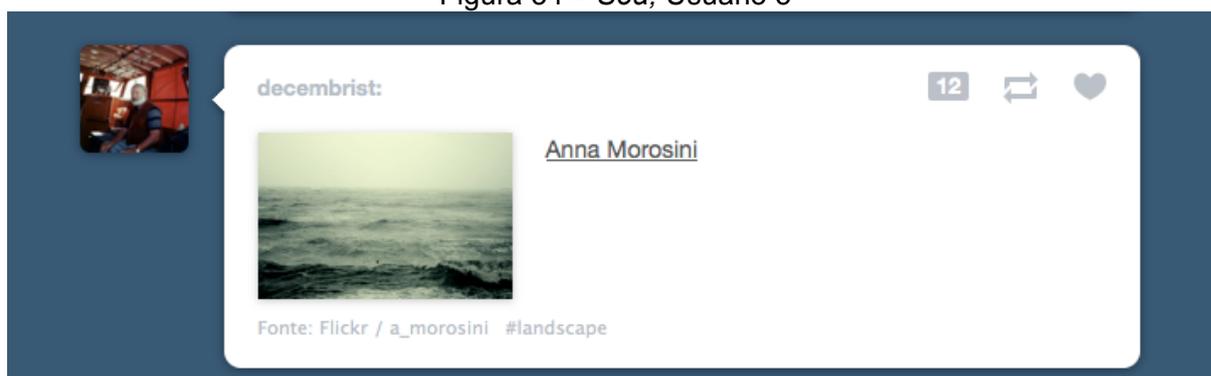
Fonte: Tumblr, 2012

Figura 33 – Sea, Usuário 4



Fonte: Tumblr, 2012

Figura 34 – Sea, Usuário 5



Fonte: Tumblr, 2012

Tabela 7: Classificação da imagem a partir da tag “sea”

	Egoísta	Amigável	Altruísta	Populista
Usuário 1			#sea; #ocean; #sunrise; #sun; #cloud; #sky; #landscape	#light; #blue
Usuário 2			#sea; #sand; #beach; #sky	#peace
Usuário 3			#sea; #landscape; #nature; #aquaticwonder	
Usuário 4			#sea; #sunset; #ocean; #sky	#beach; #wave; #waves
Usuário 5			#sea; #landscape	

Fonte: Autora

Tabela 8: Quantidade de tags atribuídas à imagem a partir da tag “sea”

	Egoísta	Amigável	Altruísta	Populista
Usuário 1	-	-	7	2
Usuário 2	-	-	4	1
Usuário 3	-	-	4	
Usuário 4	-	-	4	3
Usuário 5	-	-	2	
Resultado	0	0	21	6

Fonte: Autora

As imagens em que a *tag* “*sea*” foi atribuída foram classificadas como altruístas as seguintes etiquetas:

- *#sea* (mar) utilizada por todos os usuários;
- *#landscape* (paisagem) usada pelos usuários 1, 3 e 5;
- *#ocean* (oceano), *#sunrise* (nascer do sol) *#sun* (sol) *#cloud* (nuvem) *#sky* (céu) utilizadas pelo Usuário 2;
- *#nature* (natureza) e *#aquaticwonder* (maravilha aquática) utilizadas pelo Usuário 3;
- *#sand* (areia), *#beach* (praia) e *#sky* (céu) usadas pelo Usuário 4;
- e *#sunset* (pôr-do-sol), *#ocean* (oceano) e *#sky* (céu) utilizadas pelo Usuário 5.

As *tags* acima detalham as imagens de maneira abrangente, o que é uma característica das *tags* altruístas. Cañada (2006) afirma que o estilo altruísta é o que mais facilita na recuperação da informação. As *tags* mencionadas retratam de maneira descritiva as imagens postadas pelos usuários, acarretando em um benefício social muito alto.

Como *tags* populistas, foram classificadas as etiquetas:

- *#light* (luz) e *#blue* (azul) utilizadas pelo Usuário 1;
- *#peace* (paz) utilizada pelo Usuário 2;
- e *#beach* (praia) e *#wave(s)* (onda) utilizadas pelo Usuário 4.

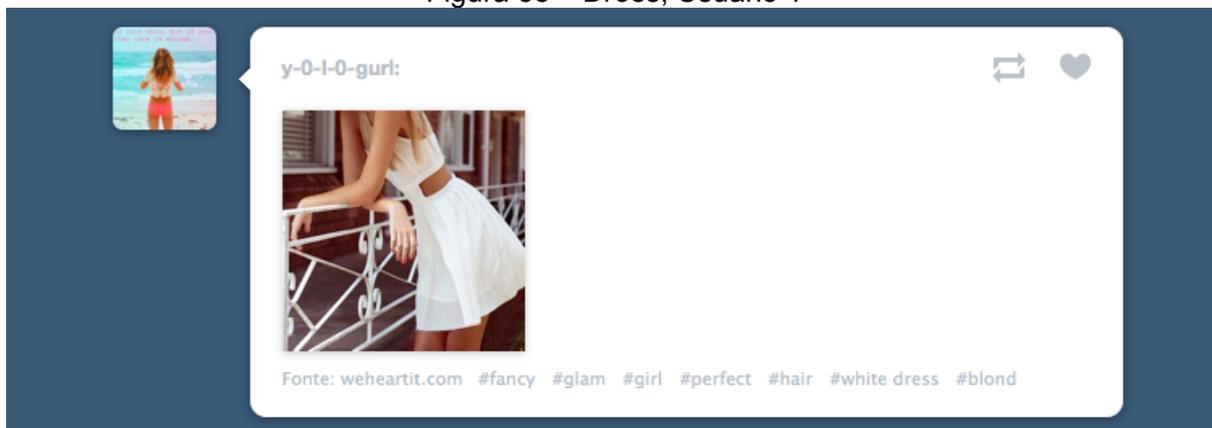
As *tags* populistas podem ser relacionadas às demais etiquetas relacionadas, porém, neste caso, as *tags* relacionam ao tema em geral, e não respectivamente às imagens. Como exemplo, “*blue*” se refere aos tons da imagem do Usuário 1, mas se imaginarmos um usuário pesquisando pela palavra “*blue*”, talvez a imagem deste caso não sirva para ele, sendo mais uma foto entre várias de sua pesquisa. Acaba por causar ruído na recuperação.

Na categoria *Landscape*, percebeu-se um grande número de *tags* altruístas indexadas, sendo um dos estilos mais utilizados pelos usuários desta pesquisa.

5.2.3 Fashion

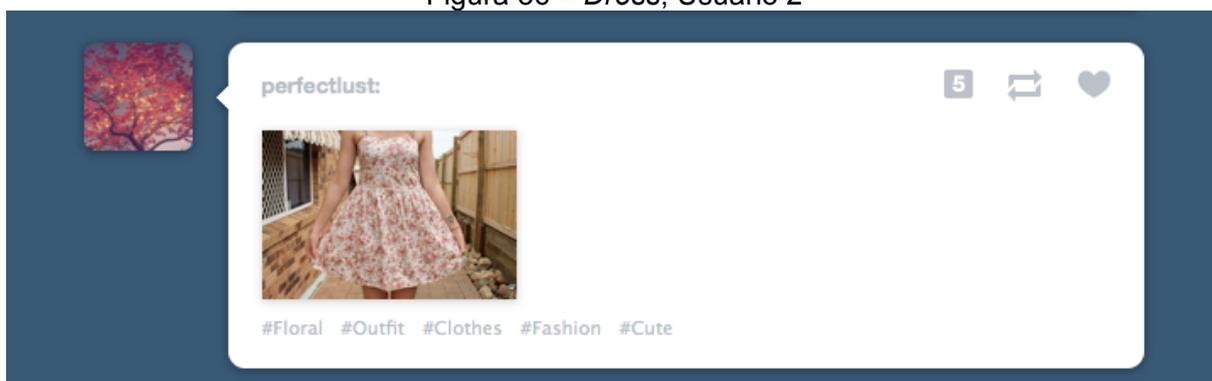
Abaixo são apresentadas as imagens e as tabelas com as classificações de tags “dress” e “shoe” para a categoria “Fashion”.

Figura 35 – Dress, Usuário 1



Fonte: Tumblr, 2012

Figura 36 – Dress, Usuário 2



Fonte: Tumblr, 2012

Figura 37 – Dress, Usuário 3



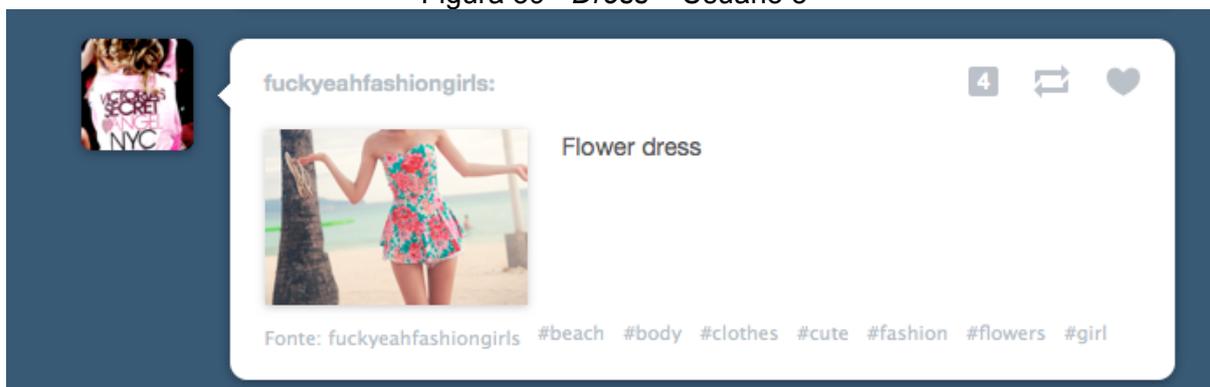
Fonte: Tumblr, 2012

Figura 38 – Dress, Usuário 4



Fonte: Tumblr, 2012

Figura 39 - Dress – Usuário 5



Fonte: Tumblr, 2012

Tabela 10: Classificação da imagem a partir da tag “dress”

	Egoísta	Amigável	Altruísta	Populista
Usuário 1	#glam; #perfect;		#dress; #white dress; #blond	#fancy; #girl; #hair
Usuário 2	#cute		#dress; #outfit; #floral; #clothes; #fashion	
Usuário 3	#pretty; #photography		#dress; #black and white; #fashion	
Usuário 4	#style; #summer		#dress; #lace; #ribbon; #fashion; #clothes; #outfit	
Usuário 5	#cute;		#dress; #beach; #body; #clothes; #fashion; #flowers	#girl

Fonte: Autora

Tabela 10: Quantidade de *tags* atribuídas à imagem a partir da *tag* “dress”

	Egoísta	Amigável	Altruísta	Populista
Usuário 1	2	-	3	3
Usuário 2	1	-	5	-
Usuário 3	2	-	3	-
Usuário 4	2	-	6	-
Usuário 5	1	-	6	1
Resultado	8	0	23	4

Fonte: Autora

Como estilo egoísta, foram classificadas as *tags*:

- *#glam* (glamour) e *#perfect* (perfeição) utilizadas pelo Usuário 1;
- *#cute* (fofo) utilizada pelos usuários 2 e 5;
- *#pretty* (bonito) e *#photography* (fotografia) usadas pelo Usuário 3;
- e *#style* (estilo) e *#summer* (verão) usadas pelo Usuário 4.

As *tags* seguem os mesmos critérios das demais classificações ao estilo egoísta. São *tags* em que ou refletem características pessoais do usuário indexador ou são etiquetas em que se referem a um grupo muito abrangente, aqui neste caso, “fotografia” e “estilo”.

No estilo altruísta as *tags* classificadas foram:

- *#dress* (vestido) utilizada por todos os usuários;
- *#white dress* (vestido branco) e *#blond* (loira) utilizadas pelo Usuário 1;
- *#outfit* (roupa), *#floral* (estampa floral), *#clothes* (roupas) e *#fashion* (moda) utilizadas pelo usuário2;
- *#black and white* (preto e branco) e *#fashion* (moda)
- *#lace* (renda), *#ribbon* (fita, faixa), *#fashion* (moda), *#clothes* (roupas) e *#outfit* (roupas) usadas pelo usuário 4;
- *#beach* (praia), *#body* (corpo), *#clothes* (roupas), *#fashion* (moda) e *#flowers* (flores – estampa) usadas pelo usuário 5.

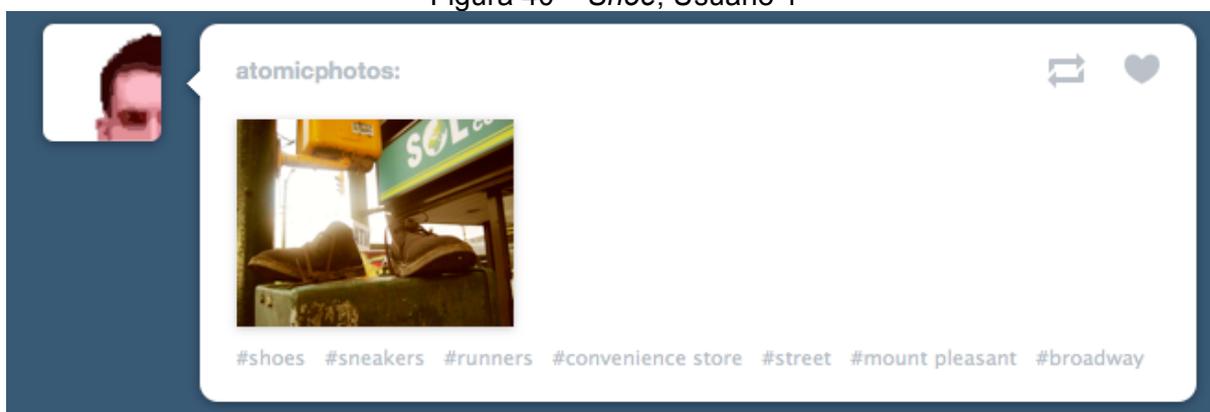
As *tags* são consideradas altruístas por descreverem as imagens postadas pelos usuários, segundo Cañada (2006). Seguindo a mesma linha de pensamento das análises acima descritas, o detalhamento de uma imagem em *tags* é muito importante para que a recuperação da informação seja satisfatória.

As *tags* populistas são:

- *#fancy* (fantasia, imaginação), *#girl* (garota, menina) e *#hair* (cabelo) utilizadas pelo usuário 1;
- *#girl* (garota, menina) utilizadas pelo Usuário 5.

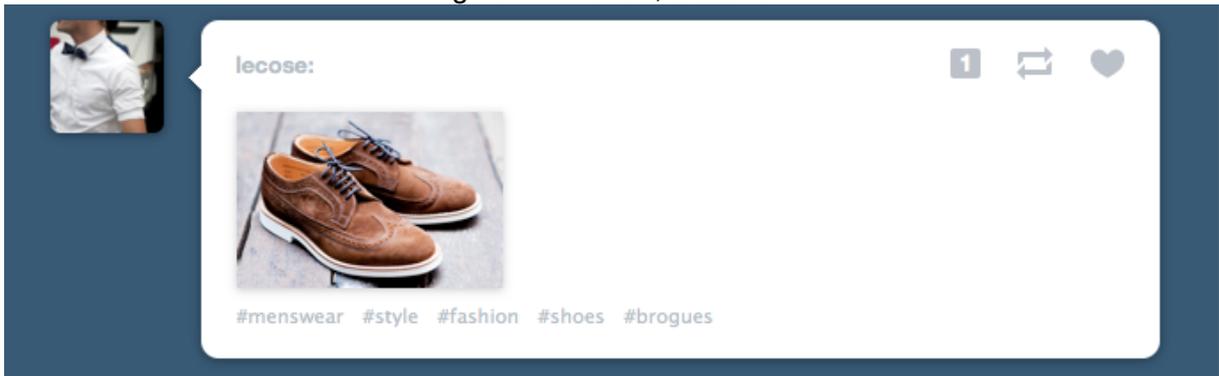
A *tag* “*hair*”, analisada isoladamente, não corresponde à imagem. É uma *tag* relacionada à *tag* “*blond*” (altruísta). As duas diferem, pois se analisarmos as duas separadamente comparada à imagem correspondente, nota-se que “*blond*” pode remeter à imagem pois aparece a cor do cabelo da menina. Já a *tag* “*hair*” não se enquadra neste exemplo, pois o cabelo da modelo não é o foco principal da imagem, mas sim o vestido. “*Girl*” e “*fancy*” são *tags* muito abrangentes relacionadas à imagem.

Figura 40 – *Shoe*, Usuário 1



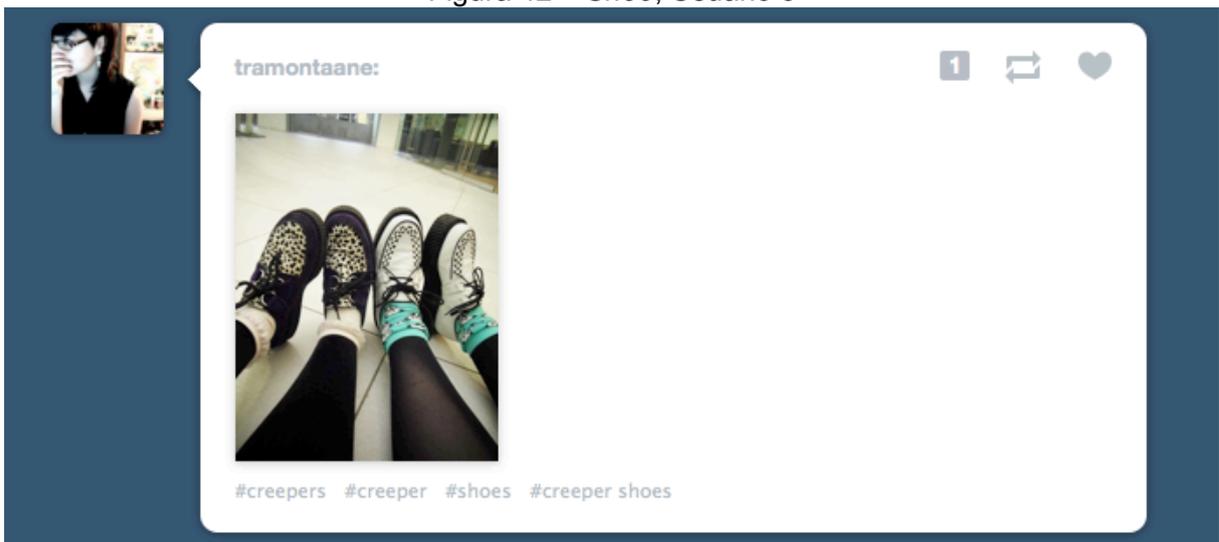
Fonte: Tumblr, 2012

Figura 41 – Shoe, Usuário 2



Fonte: Tumblr, 2012

Figura 42 – Shoe, Usuário 3



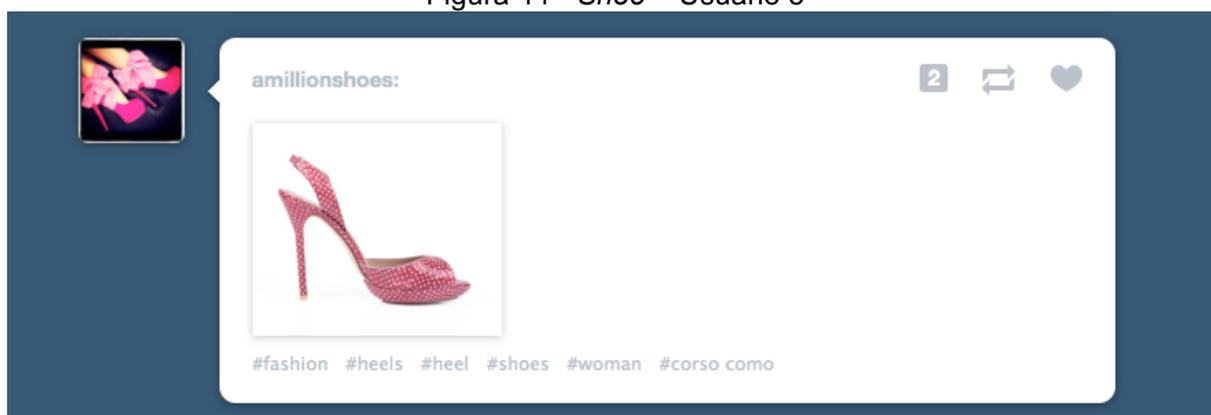
Fonte: Tumblr, 2012

Figura 43 – Shoe, Usuário 4



Fonte: Tumblr, 2012

Figura 44 - Shoe – Usuário 5



Fonte: Tumblr, 2012

Tabela 11: Classificação da imagem a partir da tag “shoe”

	Egoísta	Amigável	Altruísta	Populista
Usuário 1			#shoe; #shoes; #sneakers	#broadway; #street; #mount pleasant; #runners; #convenience store
Usuário 2	#style		#shoe; #shoes; #menswear; #fashion; #brogues	
Usuário 3			#shoe; #creeper; #shoes; #creeper #shoes	
Usuário 4	#photography		#shoe	
Usuário 5		#corso como	#shoe; #heels; #heel; #fashion; #shoes	#woman

Fonte: Autora

Tabela 12: Quantidade de tags atribuídas à imagem a partir da tag “shoe”

	Egoísta	Amigável	Altruísta	Populista
Usuário 1			3	5
Usuário 2	1		5	
Usuário 3			5	
Usuário 4	1		1	
Usuário 5		1	5	1
Resultado	2	1	19	6

Fonte: Autora

Por fim, nas imagens relacionadas à tag “shoe”, são consideradas egoístas as tags *#style* (estilo) utilizada pelo Usuário 2 e a tag *#photography* (fotografia) usada pelo Usuário 4, já vistas em exemplos acima. Foram classificadas como egoístas por serem temáticas e aceitas pelo grande público.

Como tag amigável, o estilo em que se adequa a um restrito grupo de pessoas, a tag *#corso como* foi classificada. *Corso Como* é uma marca de sapatos, e um determinado grupo que conhece esta marca poderá pesquisá-lo.

As tags altruístas foram classificadas como:

- *#shoe(s)* por todos os usuários;
- *#sneakers* (tênis) pelo usuário 1;
- *#menswear* (roupa masculina); *#fashion* (moda) e *#brogues* (o modelo do sapato) utilizadas pelo usuário 2;
- *#creeper* (modelo de sapato) utilizada pelo usuário 4;
- *#heel(s)* (salto) e *#fashion* (moda) utilizada pelo Usuário 5.

Os usuários acima descrevem as imagens de maneira precisa e também com muitas tags no singular e no plural, o que facilita muito na recuperação da informação. O benefício social é alto e enriquece a pesquisa.

As tags populistas são:

- *#broadway* (rua, via, nome de rua), *#street* (rua) *#mount pleasant* (nome de rua, cidade e vilarejos americanos), *#runners* (corredores) e *#convenience store* (loja de conveniência) utilizadas pelo usuário 1;
- *#woman* (mulher) pelo usuário 5.

As tags analisadas separadamente nada têm a ver com as imagens analisadas. Os usuários utilizaram-se de tags que descrevem uma esfera além do que realmente importa que é o conteúdo da imagem.

A partir das tags analisadas e classificadas em seus determinados estilos, criou-se uma tabela para fazer a comparação entre os estilos mais utilizados no sistema *Tumblr* pelos seus usuários.

Tabela 13 – Quadro geral da quantidade de *tags* por categoria, atribuídas às imagens a partir das *tags* analisadas.

	Egoísta	Amigável	Altruísta	Populista
Cat	9	1	11	0
Dog	9	0	18	6
Mountain	2	1	21	3
Sea	0	0	21	6
Dress	8	0	23	4
Shoe	2	1	19	6
Resultado	30	3	113	25

Fonte: Autora

De acordo com a tabela acima, observa-se que as etiquetas de estilo altruísta se destacam dentre as demais, com 113 *tags* ao todo. Isso demonstra que grande parte dos usuários do sistema *Tumblr*, ao publicar fotografias, estão preocupados com que seu conteúdo seja acessado. Consta-se que este tipo de *tag* é uma vantagem para o site, pois, além de facilitar a recuperação da informação, torna-a pesquisa satisfatória para o usuário pesquisador.

O estilo altruísta tem baixa motivação por parte do etiquetador pois é muito trabalhoso etiquetar desta maneira. Cañada (2006) explica que “[...] é difícil alguém motivar-se a fazer um trabalho para os outros quando não há benefício direto associado”. Fortalece então a ideia de vantagem, pois acredita-se que o benefício social é o mais indicado neste tipo de *site*.

O segundo estilo mais utilizado foi o egoísta, com 30 *tags* no total. Este estilo não chega a ser desvantajoso para o *site*, pois determinadas *tags* são abrangentes ao público como *tags* em que descrevem um sentimento ou uma constatação do usuário indexador. Cañada (2006) afirma que a motivação deste tipo de etiqueta é alta, pois “[...] fazemos para o nosso próprio benefício”. Nesta análise, constatou-se poucas *tags* em benefício próprio para o usuário indexador, pois não haviam tantas etiquetas que apenas o indexador pudesse recuperar posteriormente.

Com 25 *tags* ao todo, o estilo populista é o terceiro mais utilizado dentre as etiquetas analisadas. Causou surpresa este estilo não ser tão utilizado em relação

aos demais, pelo tipo do site analisado. O *Tumblr* sendo uma rede social onde os usuários cultuam o número de acessos, número de “seguidores” etc., o estilo populista adequaria-se a estas características. Constatou-se como uma vantagem - ainda que a interface de recuperação da informação seja problemática -, pois este estilo de *tags* somente prejudica os processos de recuperação da informação. Considerando que Cañada (2006) afirma que a motivação do usuário ao utilizar este tipo de etiqueta é alta, já que o benefício direto é evidente, entende-se que a ausência deste tipo de etiquetagem é uma vantagem para o sistema *Tumblr*.

E, por fim, com apenas 3 *tags* no total, o estilo amigável mostra-se como o menos utilizado pelos usuários do *Tumblr*. Cañada (2006), explica que o etiquetado amigável tem motivação muito alta por parte do usuário, pois fortalece o sentimento de comunidade entre um grupo. Por ser um estilo onde apenas um pequeno grupo irá ter benefício, o estilo amigável não é muito utilizado em decorrência das reblogagens. Como o *site* compartilha muito conteúdo de diversos usuários, a *tag* amigável não é muito utilizada devido ao fato de abranger um público maior e menos grupos. Constatou-se que apenas usuários em que postaram fotografias de autoria própria utilizaram de *tags* amigáveis.

6 CONSIDERAÇÕES FINAIS

Decorrente de *sites* - em sua maioria redes sociais - que utilizam da folksonomia como modo de organizar o conteúdo indexado por seus usuários, esta pesquisa preocupou-se em entender este fenômeno e classificar as *tags* utilizadas por usuários do site *Tumblr*. Através dos objetivos propostos, a pesquisa baseou-se principalmente em aspectos teóricos de Cañada (2006), o qual propõe estilos para as etiquetas utilizadas por usuários de sites que abordam a folksonomia em seus sistemas de recuperação da informação.

A folksonomia corrobora para a organização de *sites*, principalmente *sites* de compartilhamento. Uma de suas vantagens é não deixar a classificação de conteúdo de um site apenas com um administrador, tornando seus usuários classificadores de seu próprio conteúdo. O imediatismo das mídias atuais acaba tornando inviável deixar este trabalho só para uma pessoa. Os inúmeros usuários da rede podem compartilhar e recuperar informações a partir das *tags* que julgam ser pertinentes para determinado conteúdo.

O *Tumblr*, apesar de ser um *site* de compartilhamento relativamente novo, tem mais de 50 milhões de páginas/usuários. Seria inviável organizar tanta informação postada por dia, tanto pelo aspecto financeiro como também pela mão-de-obra necessária para isso. A organização folksonômica, ao permitir a participação das pessoas, facilita o acesso à informação deixando-a mais popular e também menos formal.

A limitação da pesquisa foi lidar com o sistema do *site Tumblr*. Não foram encontrados aspectos teóricos ou explicações no próprio *site* a respeito da forma como a recuperação da informação funciona. Sendo assim, as conclusões sobre a interface de recuperação da informação foram feitas a partir da interpretação e observação do site pelas formas viáveis.

Constatou-se um sistema problemático no *Tumblr*. O modo de recuperar conteúdo no site é restrito a apenas uma *tag* por vez. Também foi constatado que as reblogagens têm papel importante na recuperação da informação, já que parte dos usuários que reblogam algum conteúdo, publicam suas *tags*. O sistema recupera as *tags* utilizadas por todos os usuários que reblogaram uma mesma imagem, por exemplo.

Outra deficiência do *site* é não dispor de opções de pesquisa como relevância, datas, tipo do conteúdo a ser buscado (fotografia, animações, vídeos etc), mais acessados, entre outros métodos de pesquisa. O *site* apenas remete a *tag* ao conteúdo postado recentemente, em ordem decrescente.

Para classificar as *tags* de acordo com Cañada (2006), procurou-se pesquisar uma pequena amostra de usuários que etiquetaram imagens fotográficas. A busca baseou-se em categorias pré-estabelecidas pelo *site*. A partir destas, foram escolhidas as imagens de forma aleatória.

A partir de análises, interpretações e dados obtidos, observou-se uma maior utilização de *tags* no estilo altruísta, que é um estilo em que está ligado ao fato de compartilhar com um grande grupo o conteúdo de qualquer plataforma. Constata-se que, desta maneira, a recuperação da informação é eficaz e satisfatória, pois em um *site* como o *Tumblr* o conteúdo indexado é de interesse abrangente e não de determinados grupos.

Portanto, por mais que a interface de recuperação da informação do *site* seja limitado, os usuários indexadores facilitam a busca por etiquetarem de modo altruísta. O benefício social é alto, reforçando a ideia de que a folksonomia é uma das melhores alternativas para a qualidade de uma interface de recuperação da informação na *web*.

REFERÊNCIAS

AQUINO, Maria Clara. Hipertexto 2.0, folksonomia e memória coletiva: um estudo das tags na organização web. **Revista E-Compós**, v.18, n.16, ago. 2007. Disponível em: < http://encipecom.metodista.br/mediawiki/images/1/16/GT2-_07-_Hipertexto_2-0-_maria.pdf >. Acesso em: 16 maio 2012.

BONI, Valdete; QUARESMA, Sílvia Jurema. Aprendendo a entrevistar: como fazer entrevistas em ciências sociais. **Revista Eletrônica dos Pós-Graduandos em Sociologia Política da UFSC**, v. 2, n. 1, p. 68-80, jan./jul., 2010. Disponível em: <<http://www.journal.ufsc.br/index.php/emtese/article/view/18027>>. Acesso em: 03 jun. 2012.

BOYD, Danah. Social network sites: public, private, or what?. **Knowledge Tree: an e-Journal of Learning Innovation**. n.13, May 2007. Disponível em: <<http://tinyurl.com/2hqbu6>>. Acesso em: 23 abr. 2012.

BRÄSCHER, M.; CAFÉ, L. Organização da Informação ou Organização do Conhecimento? In: ENCONTRO NACIONAL DE PESQUISA EM CIÊNCIA DA INFORMAÇÃO, 9, 2008, São Paulo. **Anais**, São Paulo: ANCIB, 2008. Disponível em: < <http://www.enancib2008.com.br> >. Acesso em: 3 jun. 2012.

CALDAS, Wagner Ferreira; MOREIRA, Manoel Palhares. Folksonomia e classificação de etiquetas: estudo de caso Flickr. In: ENCONTRO NACIONAL DE PESQUISA EM CIÊNCIA DA INFORMAÇÃO, 10., 2009, João Pessoa. **Anais...** João Pessoa, 2009. Disponível em: <<http://jefersontcc.files.wordpress.com/2011/05/folksonomia-e-classificac3a7c3a3o-de-etiquetas-estudo-de-caso-flickr.pdf>>. Acesso em: jun. 2012.

CAÑADA, J. **Tipologías y estilos en el etiquetado social**. 2006. Disponível em: <<http://www.terremoto.net/tipologias-y-estilos-en-el-etiquetado-social/>>. Acesso em: 24 abr. 2012.

CATARINO, Maria Elisabete; BAPTISTA, Ana Alice. Folksonomia: um novo conceito para a organização dos recursos digitais na Web. **Revista Data Grama Zero**, v.8, n.3, jun. 2007. Disponível em: <http://www.dgz.org.br/jun07/F_I_art.htm>. Acesso em: 24 abr. 2012.

FIGUEIREDO, Nélia Maria Almeida de. **Método e metodologia na pesquisa científica**. Editora: Difusão, 2004.

GUEDES, Roger; DIAS, Eduardo José. Indexação Social: Abordagem Conceitual. **Revista ACB: Biblioteconomia em Santa Catarina, Florianópolis**, v.15, n.1, p. 39-53 jan./jun., 2010. Disponível em: <revista.acbsc.org.br/index.php/racb/article/download/686/754>. Acesso em: 27 abr. 2012.

LAKATOS, Eva Maria; MARCONI, Marina de Andrade. **Técnicas de pesquisa**. 3. ed. São Paulo: Atlas, 1996.

LAKATOS, E. M. ; MARCONI, M. de A. **Fundamentos da metodologia científica**. 6 ed. São Paulo: Atlas, 2009.

MINAYO, Maria Cecília de Souza. Ciência, Técnica e Arte: o desafio da pesquisa social. In: _____. (Org). **Pesquisa Social: teoria, método e criatividade**. 17. Ed. Petrópolis: Vozes, 2000.

NEVES, José Luis. Pesquisa qualitativa: características, usos e possibilidades. **Cadernos de Pesquisas em Administração**, São Paulo, v. 1, n. 3, 1996. Disponível em: <<http://www.ead.fea.usp.br/cad-pesq/arquivos/c03-art06.pdf>>. Acesso em: 22 out. 2011.

NORUZI, Alireza. Folksonomies: why do we need controlled vocabulary?. **Webology**, v.4, n.2, 2007. Disponível em: <<http://www.webology.ir/2007/v4n2/editorial12.html>>. Acesso em: 22 abr. 2012.

NUNES, Cláudio Omar. Algumas considerações acerca da ausência de políticas de indexação em bibliotecas brasileiras. **Biblos**, Rio Grande, v. 16, p. 55-61, 2004. Disponível em: <<http://www.brapci.ufpr.br/documento.php?dd0=0000006479&dd1=8cb17>>. Acesso em: 15 out. 2011.

O'REILLY, Tim. What is Web 2.0? design patterns and business models for the next generation in software. **O'Reilly Media**, Califórnia, 2005. Disponível em: <<http://www.oreillynet.com/pub/a/oreilly/tim/news/2005/09/30/what-is-web-20.html>>. Acesso em: 23 abr. 2012.

PRIMO, Alex. O aspecto relacional das interações na Web 2.0. **E- Compós** (Brasília), v. 9, p. 1-21, 2007. Disponível em: <www6.ufrgs.br/limc/PDFs/web2.pdf>. Acesso em: 22 abr. 2012.

ROBREDO, Jaime; CUNHA, Murilo B. da Cunha. **Documentação de hoje e de amanhã: uma abordagem informatizada da biblioteconomia e sistemas de informação**. São Paulo: Global, 1994.

ROWLEY, Jennifer. **Abstracting and indexing**. 2.ed. London: Clive Bingley, 1988.

SILVA, Fabiano; BLATTMANN, Ursula. A colaboração e a interação na web 2.0. In: CONGRESSO BRASILEIRO DE BIBLIOTECONOMIA, DOCUMENTAÇÃO E CIÊNCIA DA INFORMAÇÃO, 22, 2007, Brasília. **Revista ACB: Biblioteconomia em Santa Catarina**, Florianópolis, v.12, n.2, p. 191-215, jul./dez., 2007.

SILVA, Iara; SANTOS, Cibele dos. Indexação colaborativa na web. **CRB-8 Digital**, São Paulo, v. 4, n. 1, p. 90-98, abr. 2011. Disponível em: <<http://revista.crb8.org.br>>. Acesso em: 05 jun. 2012.

TUMBLR. Disponível em: <<http://tumblr.com>>. Acesso em: 17 abr. 2011.

VAN DER LAAN, Regina Helena. **Tesauro e terminologia**: uma inter-relação lógica. 2002. 196 f. Tese (Doutorado) - Departamento de Programa de Pós-graduação em Letras, Universidade Federal do Rio Grande do Sul, Porto Alegre, 2002.

WAL, Thomas Vander. **Folksonomy definition and wikipedia**. Disponível em: <<http://www.vanderwal.net/random/entrysel.php?blog=1750>>. Acesso em: 24 abr. 2012.