

UNIVERSIDADE FEDERAL DE MINAS GERAIS  
ESCOLA DE CIÊNCIA DA INFORMAÇÃO

PRISCILA OLIVEIRA DA MATA

**O COMPORTAMENTO DA FOLKSONOMIA NO *TWITTER*: A ATINÊNCIA EM  
RECUPERAÇÃO DA INFORMAÇÃO**

Belo Horizonte  
2011

PRISCILA OLIVEIRA DA MATA

**O COMPORTAMENTO DA FOLKSONOMIA NO *TWITTER*: A ATINÊNCIA EM  
RECUPERAÇÃO DA INFORMAÇÃO**

Monografia apresentada ao programa de Especialização do Núcleo de Informação Tecnológica e Gerencial – NITEG, no curso Arquitetura e Organização da Informação da Escola de Ciência da Informação, da Universidade Federal de Minas Gerais, como requisito para a obtenção do certificado de Especialista em Arquitetura e Organização da Informação.

Orientador: Prof. Eduardo Ribeiro Felipe

Belo Horizonte  
2011

M425c

Mata, Priscila Oliveira da.

O comportamento da Folksonomia no *Twitter* [manuscrito] : a atenção em recuperação da informação. / Priscila Oliveira da Mata. – 2011.

64 f., enc.: il., graf., color.

Orientador: Eduardo Ribeiro Felipe.

Área de concentração: Arquitetura e Organização da Informação.

Monografia (Especialização) – Universidade Federal de Minas Gerais, Escola de Ciência da Informação.

Bibliografia: f. : 60-64.

1. Twitter. 2. Organização da informação. 3. Rede de relações sociais. 4. Marketing na internet. 5. Inteligência coletiva. 6. Internet (Redes de computação). 7. Indústria da Internet. 8. Sistemas de recuperação da informação. 9. Colaboração online. 10. Comunicação mediada por computador. 11. Comunicação virtual. 12. Atenção. I. Felipe, Eduardo. II. Universidade Federal de Minas Gerais. Escola de Ciência da Informação. III. Título.

CDU: 025.177



**Universidade Federal de Minas Gerais**  
**Escola de Ciência da Informação**  
**Programa de Pós-Graduação em Ciência da Informação**

Trabalho de Conclusão de Curso de Especialização Arquitetura e Organização da Informação intitulado “O comportamento da *folksonomia* no *Twitter*: a atenção em recuperação da informação”, de autoria de Priscila Oliveira da Mata, aprovada pela banca examinadora constituída pelos seguintes professores:

## **O COMPORTAMENTO DA FOLKSONOMIA NO TWITTER: A ATINÊNCIA EM RECUPERAÇÃO DA INFORMAÇÃO**

---

Prof. Esp. Eduardo Ribeiro Felipe (orientador).  
Escola de Ciência da Informação - UFMG  
Orientador

---

Profa. Dra. Maria Aparecida Moura  
Escola de Ciência da Informação - UFMG

Data de aprovação: Belo Horizonte, 14 de Outubro de 2011

## DEDICATÓRIA

*Dedico este trabalho aos meus pais, fonte inesgotável  
de amor e inspiração em minha vida.*

## AGRADECIMENTOS

A Nossa Senhora Aparecida, luz incessante que ilumina meu caminho.

Ao meu orientador, Professor Eduardo Felipe, por ter aceitado e acreditado no meu projeto, pela orientação precisa e cuidadosa, pela boa vontade, pela motivação, pelos conselhos, pela força e por inspirar a minha vida acadêmica.

Aos meus pais queridos, Norman e Danilo, que sempre se doaram inteiros e renunciaram aos seus sonhos, para que, muitas vezes, pudesse realizar os meus, pelo amor incondicional, pelo incentivo e estímulo e por terem me mostrado desde sempre a importância do estudo em minha vida. A vocês, todo o meu amor.

À bibliotecária Dilma Louback, minha primeira referência na área de C.I., por ter me mostrado a importância do Profissional da Informação.

Aos meus colegas de trabalho da Biblioteca da Faculdade de Letras da UFMG, pela alegre convivência, incentivo e apoio constantes.

À Faculdade de Letras pelo importante apoio.

À minha família: irmão, primos e primas, tios e tias, por sempre me apoiarem e me ajudarem com palavras de incentivo nesta minha caminhada acadêmica.

Aos meus amados sobrinhos, Amanda e Mateus, pela leveza, felicidade e luz que trazem à minha vida.

Aos meus queridos amigos, pela paciência e compreensão nos vários momentos em que estive ausente, em virtude da realização deste trabalho. Vamos comemorar?

Aos meus companheiros de banda Vitrola 101, pela música, pela poesia, pelo carinho, sensibilidade e pela torcida. Vamos nos encontrar para uma *Jam*?

*“A internet é um espaço de comunicação propriamente surrealista, do qual ‘nada é excluído’, nem o bem, nem o mal, nem suas múltiplas definições, nem a discussão que tende a separá-los sem jamais conseguir. A internet encarna a presença da humanidade a ela própria, já que todas as culturas, todas as disciplinas, todas as paixões aí se entrelaçam. Já que tudo é possível, ela manifesta a conexão do homem com a sua própria essência, que é a aspiração à liberdade.”*

*Pierre Lévy*

*“Que seja doce.”  
Caio Fernando Abreu*

## RESUMO

Investigou-se o comportamento da *folksonomia* no *Twitter* no que tange principalmente à credibilidade do conteúdo das informações relacionadas às *hashtags* utilizadas pelos usuários do microblog para classificarem seus *posts*. Analisou-se também o fenômeno lexical da polissemia, um dos fatores interferentes na precisão na recuperação da informação no *Twitter* e os problemas causados pela proliferação de *spams* que tem afetado diretamente a credibilidade das informações que circulam no microblog. Apresentou-se uma revisão de literatura sobre *web*, *web 2.0*, redes sociais, *folksonomia*, semântica e polissemia. Utilizou-se para compor o percurso metodológico a netnografia, que analisa o comportamento dos indivíduos de comunidades virtuais e cibercultura. Este método permitiu analisar as estratégias adotadas pelos usuários do *microblog* para indexarem seus *posts* e entender como estas estratégias interferem na precisão na recuperação de informações e na credibilidade das informações que circulam dentro do *Twitter*.

**Palavras-chave:** *Folksonomia*. *Twitter*. Indexação social. Atinência. Recuperação da informação.



## ABSTRACT

Investigated the behavior of folksonomy on *Twitter* especially regarding the credibility of the information content related to the use *hashtags* Twitter users to rate your posts. Also analyzed the phenomenon of lexical polysemy, one of the factors that determine accuracy in information retrieval on *Twitter* and the problems caused by the proliferation of spam that has directly affected the credibility of the information circulating in the *microblog*. Presented a literature review of web, web 2.0, social networking, folksonomy, semantics and polysemy. Was used to compose the methodological *netnografia* route, which analyzes the behavior of individuals in virtual communities and cyberspace. This method allowed us to analyze the strategies adopted by users of *Twitter* to index your posts and understand how these strategies affect the accuracy in retrieving information and the credibility of the information circulating within *Twitter*.

**Keywords:** Folksonomy. *Twitter*. Social Indexing . Atinência. Information retrieval.



## LISTA DE ILUSTRAÇÕES

Figura 1: Trending Topics .....	13
Figura 2: Variação do significado de palavras No Brasil e em Portugal.....	29
Figura 3: Página inicial do Twitter .....	33
Figura 4: Perfil no Twitter .....	34
Figura 5: Follow e Followers .....	36
Figura 6: Assuntos do dia no Twitter 1.....	37
Figura 7: Assuntos do dia no Twitter 2.....	37
Figura 8: Retweet.....	38
Figura 9: Ferramenta reply no Twitter.....	39
Figura 10: Uso da ferramenta reply no Twitter.....	39
Figura 11: Ferramenta Mensagem Direta.....	40
Figura 12: Tela da Mensagem Direta.....	40
Figura 13: Exemplo de uso das hashtags .....	41
Figura 14: Nuvem de <i>tags</i> no <i>Twitter</i> (14/05/10 a 11/08/10) .....	42
Figura 15: Tópicos que ficaram mais tempo entre os Trending Topics.....	43
Figura 16: Popular trending.....	44
Figura 17: Search <i>Twitter</i> .....	45
Figura 18: Etiquetagem no Twitter .....	46
Figura 19: Hashtag CALABOCAGALVAO .....	47
Figura 20: Hashtag PARANGOLIXO.....	48
Figura 21: Polissemia 1 .....	49
Figura 22: Polissemia 2 .....	50
Figura 23: Polissemia 3 .....	51
Figura 24: <i>Spam</i> no Twitter .....	51
Figura 25: Timeline AvistaHotelsfl .....	52
Figura 26: Report for <i>spam</i> .....	55

## SUMÁRIO

1. INTRODUÇÃO.....	10
1.1 Problema.....	14
2. OBJETIVOS.....	15
2.1 Objetivo geral .....	15
2.2 Objetivos específicos.....	15
3. JUSTIFICATIVA.....	17
4. REFERENCIAL TEÓRICO .....	18
4.1 Web.....	18
4.2 Web 2.0.....	19
4.4 Redes sociais.....	20
4.5 Folksonomia .....	24
4.6 Semântica.....	26
4.6.1 <i>Polissemia e homonímia</i> .....	28
5. METODOLOGIA .....	30
5.1 Netnografia .....	30
5.2 Twitter: ambiente de pesquisa.....	32
5.3 Análise das buscas no Twitter .....	44
5.4 Filtros (Anti-Spam).....	53
6. RESULTADOS .....	57
REFERÊNCIAS.....	60

## 1. INTRODUÇÃO

O advento do serviço *world wide web* (*www*)<sup>1</sup> ou, simplesmente *web*, na rede mundial de computadores, criado pelo físico britânico Tim Berners – Lee em 1989, trouxe um crescimento vertiginoso na quantidade de informações e de usuários conectados à *internet*<sup>2</sup>, o que levou a uma rápida popularização das tecnologias de informação e comunicação<sup>3</sup> numa escala nunca antes vista.

De acordo com a *Internet World Stats*<sup>4</sup>, 1,96 bilhão de pessoas tinham acesso à *internet* em junho de 2010, o que representa 28,7% da população mundial. Segundo estudos da União Internacional de Telecomunicações (UIT)<sup>5</sup>, o número de internautas no mundo dobrou nos últimos cinco anos e vai ultrapassar a marca de 2 bilhões até o final deste ano. Estudos apontam que em 2012 a Internet terá mais de quatro bilhões de usuários com conexão fixa e móvel.

Deste modo, são infinitas as informações circulando na rede a todo instante. E em virtude desse grande volume de informações, os problemas de acesso e recuperação de qualquer tipo de informação neste contexto tornam-se cada vez maiores. Posto isto, cresce a necessidade de viabilizem a organização e recuperação das informações, no âmbito tradicional e semântico em todo o mundo. Vários projetos no sentido de se criarem padrões para estruturação de sites e páginas na *internet* surgiram como iniciativas para a estruturação ou desenvolvimento de metodologias que viabilizem a organização semântica das informações disponíveis na *web* em todo o mundo (FEITOSA, 2006).

A *folksonomia*, objeto de nosso estudo na seção 4.4, surge nesse contexto, caracterizado por um número infinito de informações criadas a todo instante pelos usuários que estão conectados à rede, em um ambiente colaborativo, ambiente esse que é característica intrínseca a *Web 2.0*, que será discutida mais adiante. O termo *Web 2.0* foi criado por Tim O'Reilly em 2004 e reforça o conceito da *internet* de propiciar que seus usuários colaborem efetivamente para a disponibilização de serviços virtuais e organização dos conteúdos (O'REILLY, 2006).

---

<sup>1</sup> Em uma tradução literal, “Teia de Alcance Mundial”.

<sup>2</sup> Infra-estrutura de redes, servidores e canais de orientação que lhe dá sustentação.

<sup>3</sup> Sigla utilizada daqui a diante: TICs

<sup>4</sup> Internet World Stats é um site internacional que disponibiliza estatísticas sobre a população mundial, informações sobre os mercados de telecomunicações e dados recentes sobre a Internet. Disponível em: <http://www.internetworldstats.com/>

<sup>5</sup> A União Internacional de Telecomunicações (UIT) ou International Telecommunication Union (ITU) é uma agência das Nações Unidas que trata dos assuntos relacionados com as tecnologias de informação e comunicação.

A *Web 2.0* tem como uma de suas características o uso cada vez mais comum dos *blogs* e das redes sociais virtuais. O fenômeno das redes sociais teve início em 1997 com a criação do site *Sixdegrees*<sup>6</sup> e hoje conecta cerca de 1,28 bilhões de pessoas no mundo todo. Esse número vem crescendo a cada dia, principalmente em virtude da popularização do acesso à *internet* no mundo e, também, em função do acesso à rede através da telefonia móvel cada vez mais comum nos dias de hoje.

A *internet* atual é caracterizada por um ambiente colaborativo, onde os usuários são ao mesmo tempo produtores e consumidores das informações que circulam na rede. E esse ambiente, como lembra Moura (2009,p.61), só é possível em virtude da existência de tecnologias que

[...] construídas no contexto Web 2.0, como o wiki, os blogs, o Consumer-Generated Média (CGM), ReallySimpleSyndication (RSS), Sistemas de Classificação Distribuída (Distributed Classification Systems – DCSs) e as folksonomias apresentam a oportunidade de compartilhar o conhecimento e ampliar processo interativo entre usuários e informações. (MOURA, 2009, p. 61)

Destarte, esta interação proporcionada pela *Web 2.0* é evidenciada nas redes sociais onde os usuários trocam informações a todo instante, fazendo do ciberespaço um ambiente mais dinâmico, atualizado e potencialmente favorável aos estreitamentos das relações sociais.

Nunca foi tão fácil publicar um conteúdo na *web*. Atualmente, ficamos sabendo dos acontecimentos pela *internet* bem antes mesmo deles serem publicados nas mídias impressas e ou televisivas. Exemplo recente deste fato foi a notícia da morte de Osama Bin Laden, o terrorista mais procurado no mundo. Antes mesmo de ser divulgada pela imprensa, uma pessoa comum percebeu uma movimentação estranha nos arredores da casa onde o terrorista estava escondido. Sem saber ao certo que se tratava de uma operação do exército norte-americano para capturar Bin Laden, na madrugada do dia dois de maio, Sohaib Athar<sup>7</sup> escreveu em seu perfil no *Twitter* (nosso objeto de estudo nos próximos capítulos) que uma forte explosão havia sacudido as janelas de sua casa e que esperava que o estrondo não fosse “o começo de algo degradável”. Sem perceber, Sohaib tinha transmitido ao vivo pelo *Twitter* o ataque ao maior terrorista de todos os tempos. Horas depois, era feito o anúncio da

---

<sup>6</sup> The Web's Hottest New Social Network. Disponível em : <<http://www.sixdegrees.com/>>

<sup>7</sup> Consultor de TI da cidade paquistanesa de Abbottabad. No Twitter, usa o codinome "ReallyVirtual".

morte de Osama pelo então presidente dos Estados Unidos, Barack Obama, em várias emissoras de televisão.

Acontecimentos como este nos mostram o poder das redes sociais como o *Twitter*. Criada em 2006, essa ferramenta, também chamada de *microblog*<sup>8</sup> (embora alguns discordem desta denominação, uma vez que o *Twitter* não apresenta todas as características de um *blog*), vem crescendo consideravelmente a cada dia, tornando-se um fenômeno da *internet*. Em um espaço de apenas 140 caracteres (essa limitação é para viabilizar a integração do *Twitter* com mensagens de texto via celular - SMS), o usuário que tem um perfil na rede pode “*twitter*” acontecimentos importantes ou não e as empresas podem divulgar a sua marca, promoções ou o lançamento de determinado produto.

Desta forma, com os seus benefícios, possibilidades e poder de alcance, o *Twitter* cada vez mais tem conquistado internautas, artistas, jornalistas, intelectuais, chefes de Estado e, mais recentemente, tem chamado a atenção de grandes empresas. O *Twitter* trouxe consigo a popularização de termos como *tags*, *hashtags*, *tweets*, *trending topics*. As *tags* são palavras-chave ou termos associados a um determinado assunto

... (conhecidas também como marcadores ou palavras-chave) são ferramentas usadas por sites, blogs ou usuários para indexar conteúdos da mesma espécie. No *Twitter*, para se indexar assuntos com o mesmo conteúdo, também são utilizadas *tags*, conhecidas como *hashtags*. Os usuários utilizam o sinal # (*hash*) para agregar seu *tweet* a outros que também tenham a mesma palavra-chave. Os tópicos mais comentados na rede social (*Trending Topics*) se constroem a partir desses agregadores. (OLIVEIRA, 2009, p.2 )

Atualmente, no ambiente *web*, é muito comum nos depararmos com as *hashtags*, pois elas possibilitam o agrupamento de uma informação relacionado a alguma palavra-chave (assunto), facilitando assim a sua futura recuperação. No *Twitter*, as *hashtags* mais comentadas ficam agrupadas no *trending topics*.

O *trending topics* ou *TT's* é uma lista dos assuntos mais comentados no *Twitter* em um determinado momento. Atualmente, muitas pessoas estão usando a ferramenta de busca *Twitter Search* para ficar por dentro dos assuntos mais discutidos

---

<sup>8</sup> Microblog ou Micro Blogging é uma forma de publicação de blog que permite aos usuários que façam atualizações breves com aproximadamente 140 caracteres assim como o Twitter e publicá-las para que sejam vistas publicamente ou apenas por um grupo restrito escolhido pelo usuário.

no Brasil e no Mundo. Na figura 1 abaixo, mostramos a lista dos TT's registrados em um determinado momento.



**Figura 1: Trending Topics**

Fonte: [www.twitter.com](http://www.twitter.com)

Os *trending topics* são um exemplo do caráter atemporal do *Twitter*. A cada instante, a lista dos assuntos mais comentados no *blog* é modificada, pois leva em consideração os temas mais postados pelos usuários do *Twitter*. Segundo pesquisa publicada no Mashable<sup>9</sup>, a duração dos *trends* na lista duram em média menos de 40 minutos. Ainda de acordo com a pesquisa, o *Twitter* se preocupa mais com um assunto específico e o alcance de um *tweet* do que quem o postou e com que frequência ele foi postado.

Atualmente, o sucesso dos *trending topics* é tão evidente que as grandes empresas já entenderam que esse pode ser um grande caminho para divulgar a sua marca. Estar entre os assuntos mais comentados possibilita a essas empresas uma forma de divulgação de sua marca ou produto, tendo em vista o alcance que esta ferramenta tem adquirido nos últimos tempos. No entanto, o reconhecimento da importância dos *trending topics*, por parte das empresas, tem gerado uma discussão muito grande acerca da possível manipulação e censura dos *trending topics* pelo *Twitter* para favorecer empresas que pagam para que sua marca e produtos estejam entre os assuntos mais comentados.

Outra questão muito discutida recentemente com relação aos *trending topics* diz respeito ao problema dos *spams*<sup>10</sup>. Alguns usuários e até empresas estão se

<sup>9</sup> Social Media News and Web Tips – Mashable – The Social Media Guide. Referência em notícias sobre mídia digital, social, tecnologia, novos sites e serviços entre outros. Disponível em: <<http://mashable.com/>>

<sup>10</sup> Spam é uma mensagem não solicitada normalmente com propaganda apelativa.



aproveitando dos termos em evidência para disseminar *spams* através do *Twitter*. Com o intuito de conectar audiência dentro da rede, muitos usuários lançam mão das *hashtags* mais comentadas no momento em seus *tweets*, conseguindo assim uma maior visibilidade. Desta forma, tem - se verificado muitos *tweets* sem relação alguma com a *hashtag* usada para representar seu conteúdo.

## 1.1 Problema

Se por um lado, o uso das *hashtags* no *Twitter* vem simbolizar uma tendência atual na organização da informação em ambientes colaborativos, uma vez que é um exemplo claro do uso da *folksonomia*, por outro lado o uso indiscriminado dessas *hashtags* pode prejudicar a precisão na recuperação da informação<sup>11</sup>.

A proposta deste trabalho permeia justamente o universo desta nova forma de organização da informação em universos colaborativos como o *Twitter*. Buscamos analisar o comportamento da *folksonomia* no microblog, no que tange principalmente à credibilidade das informações que estão atreladas às *hashtags* que, em um determinado período, estavam na lista dos *trending topics*. Desta forma, serão verificados problemas como a polissemia e a proliferação de *spams*, que podem ser percebidos no momento em que se realiza uma busca por um termo que esteve entre os assuntos mais comentados no *Twitter*, em virtude do uso da *folksonomia* pelos usuários do microblog. Procuramos entender como e com que finalidade os usuários e empresas têm indexado seus *tweets*, se apenas para conectar audiência dentro da rede ou apenas para divulgar uma marca ou produto. Procuramos ainda, verificar se as informações que estão atreladas aos termos presentes nos *trending topics* realmente dizem respeito àquele termo e se existe de fato um conteúdo informativo nos *posts* dos termos de grande audiência.

E, finalmente, no que se refere aos *spams*, entender como eles estão afetando a credibilidade das informações recuperadas numa busca no *Twitter* e como isso pode afetar o trabalho de monitoramento de redes sociais realizado pelas diversas empresas que estão usando o *Twitter* como ferramenta de marketing. Paralelamente, verificaremos se os administradores do *Twitter* têm censurado as informações que circulam dentro do microblog na tentativa de controlar os problemas causados pelo uso indiscriminado das *hashtags*.

---

<sup>11</sup> A expressão “precisão em recuperação da informação” é o que se entende pelo conceito de atinência.

## 2. OBJETIVOS

A proposta de análise das informações vinculadas à utilização da *folksonomia*, assim como a atinência dos resultados em processos de recuperação da informação será realizada através do atendimento do objetivo geral e aos objetivos específicos, como indicado nas próximas seções.

### 2.1 Objetivo geral

A presente pesquisa tem por objetivo investigar o comportamento da *folksonomia* no *Twitter* no que tange à credibilidade das informações que estão relacionadas às *hashtags* usadas pelos usuários do microblog para indexarem seus *posts* e, que em um determinado momento estavam presentes no *Trending Topics* do *Twitter*.

### 2.2 Objetivos específicos

Neste sentido, para que seja possível obter uma resposta consistente objetiva-se especificamente:

- Investigar o comportamento da *folksonomia* no *Twitter*;
- Entender a *folksonomia* como forma de representação da informação no *Twitter*;
- Identificar quais as estratégias de postagens adotadas pelos usuários do microblog no momento da etiquetagem dos seus *posts*;
- Analisar a credibilidade das informações que circulam no microblog;
- Verificar como a polissemia pode interferir na precisão dos resultados das buscas no *Twitter*;
- Avaliar, a partir de buscas realizadas no *Twitter* com as *hashtags* mais comentadas em determinado período, *#CALABOCAGALVÃO*, *#Gol*, *#Rio*, *#Parangolixo*, *#Idol* e a *hashtag #ufmg*, que no momento da pesquisa, não figurava entre os *Trending Topics*, se o conteúdo dos *posts* recuperados condiz com o assunto das *hashtags*;

- Avaliar se os usuários do microblog estão usando as *hashtags* apenas para aumentar o seu alcance no *Twitter*, divulgar uma marca ou produto e conectar audiência, associando seus *tweets* a assuntos que têm uma maior representatividade na rede;
- Apontar como o problema que o *Twitter* tem enfrentado atualmente com os *spams* pode comprometer o valor informativo da ferramenta;
- Apontar possíveis soluções para o problema dos *spams* no microblog.

### 3. JUSTIFICATIVA

Atualmente, o *Twitter* vem ganhando significativa importância entre os internautas, tornando-se uma importante ferramenta para divulgação de informações, sejam elas relevantes ou não. Além disso, empresas de todos os segmentos do mercado, seja *e-commerce*, cultura, jornalismo entre outros, também estão reconhecendo a importância dessa ferramenta e, cada vez mais, estão utilizando-a para divulgarem suas marcas e produtos. Neste sentido, o *Twitter* vem-se tornando também uma poderosa ferramenta de *marketing*, obrigando as a empresas a investirem em profissionais responsáveis pelo monitoramento da audiência da sua marca dentro do microblog.

O vertiginoso crescimento da rede social *Twitter* e o reconhecimento do seu poder de alcance por parte das empresas e usuários do microblog trouxe alguns problemas como a proliferação de *spams* dentro da rede. Na ânsia de adquirir novos seguidores e divulgar marcas e produtos, usuários e empresas estão lançando mão de atitudes que estão contribuindo para a descaracterização da verdadeira função do *Twitter* que é a de ser uma ferramenta de comunicação rápida, eficaz e uma potencial fonte de informações.

Destarte, a presente pesquisa tem como fator motivador investigar as implicações do uso da *folksonomia* no *Twitter* no que tange principalmente à credibilidade do conteúdo das informações que estão atreladas às *hashtags* usadas pelos usuários do microblog para indexarem seus *posts* e que, em um determinado momento, encontravam-se nos *trending topics* do *Twitter*. Além disso, será avaliado o fenômeno lexical conhecido como polissemia, muito comum em ambientes onde se faz uso da *folksonomia*, sendo um dos fatores que interferem na precisão na recuperação da informação.

Concomitantemente, será analisada também a crescente proliferação de *spams* dentro do *Twitter* em virtude da falta de ferramentas eficazes para combaterem este problema.

## 4. REFERENCIAL TEÓRICO

A seguir será realizado um levantamento bibliográfico sobre a organização da informação na web, *Web 2.0*, redes sociais, *Twitter* e sobre *folksonomia*.

### 4.1 Web

O advento da *internet* trouxe mudanças consideráveis nos processos de tratamento e recuperação da informação, devido à grande quantidade de informações que circulam na rede. Segundo Feitosa (2006), novas formas de organizar e compartilhar os conteúdos disponíveis na *internet* surgem em virtude deste novo cenário trazido pelo uso cada vez maior das TICs.

A *web* muitas vezes está sendo comparada à Alexandria<sup>12</sup> do século 21, pois considera - se que nela se encontra o mais rico repositório de informações na história humana. Desta forma, como lembra Feitosa (2006), fica evidente que a otimização de ferramentas para armazenamento, busca e recuperação de informações na *web* deve ser uma preocupação constante dos pesquisadores.

Levy (2007) avalia a importância do desenvolvimento desse ciberespaço afirmando que:

[...] um mundo virtual pode ser igualmente portador de cultura, de beleza, de espírito e de saber como um templo grego, uma catedral gótica, um palácio florentino, a *Encyclopédie* de Diderot e d'Alembert ou a Constituição dos Estados Unidos. Pode desvendar inéditas galáxias de linguagem, fazer vir à tona temporalidades sociais desconhecidas, reinventar o laço social, aperfeiçoar a democracia, abrir entre os homens trilhas do saber desconhecidas. (LEVY, 2007, p.103)

Segundo Baeza –Yates e Ribeiro Neto (1999), a *web* é altamente interativa, porque não é passiva perante o tempo e o espaço, possibilitando a criação de informações por diversos membros em um ambiente colaborativo. No entanto, apesar de existir enorme interatividade, as pessoas ainda possuem dificuldades para recuperar informações relevantes na *web*. Neste sentido, surgem novos conceitos no âmbito da organização da informação na *web*, como a *folksonomia*, que será comentada mais adiante. Este novo modelo de organização e representação da

---

<sup>12</sup> Entre os anos de 280 a.C. a 416, a biblioteca de Alexandria reuniu o maior acervo de cultura e ciência que existiu na antiguidade.

informação está sendo amplamente utilizada em ambientes colaborativos, característicos da *web 2.0*, que será descrita a seguir.

## 4.2 Web 2.0

O termo *web 2.0*, embora não aceito por todos, surge como uma proposta para definir a segunda geração de serviços de *internet*. Difícil é precisar quando termina a *web 1.0* e quando começa a *web 2.0*, uma vez que são analisados padrões de comportamento na rede e esses padrões estão mudando numa velocidade muito grande.

A principal característica da *web 2.0* é a colaboração, pois nesta nova proposta de *web* as ações são realizadas pelos próprios usuários, de maneira coletiva, de acordo com suas necessidades, ou seja, cada usuário colabora com a criação dos conteúdos disponíveis na *internet*.

No que tange ao caráter colaborativo e interativo da *web 2.0*, Moura (2009) afirma que:

O aprofundamento da discussão do caráter interativo da WEB e da efetiva presença dos usuários de informação no ambiente digital criou os elementos para o surgimento da *Web 2.0*. O termo foi proposto entre 2003 e 2004 e designa a segunda geração de produtos e serviços desenvolvidos no contexto da Internet. Pensada inicialmente como uma estratégia de marketing segmentado, a Web 2.0 privilegiou a efetiva colaboração do usuário no processo de organização e personalização da informação e proporcionou, desse modo, a alteração das lógicas de organização e disseminação da informação, agora não mais centralizadas no papel do mediador da informação. (MOURA, 2009, p. 32)

Segundo Silva e Blattmann (2007, p. 198), a *web 2.0* “pode ser considerada uma nova concepção, que passa agora a ser descentralizada, e na qual o sujeito torna-se um ser ativo e participante sobre a criação, seleção e troca de conteúdo postado em um determinado site por meio de plataformas abertas”.

A *web 2.0* apresenta um ambiente totalmente interativo com inúmeras possibilidades, através das diversas ferramentas e recursos que favorecem o estabelecimento de comunicação e interatividade entre os usuários. Tendo em vista essas diversas possibilidades de comunicação, algumas ferramentas ganharam um

grande destaque em meio aos mais variados públicos. (RUFINO; TABOSA; SOUZA, 2009).

A *web 2.0* pode ser considerada sinônimo de interação. Ela trouxe a possibilidade do usuário interagir com as informações disponíveis na *internet*, por meio de *e-mails*, *blogs*, *wikis*, *podcastings*, *Twitter* e muitas outras ferramentas. Essa interação característica da *web 2.0* está presente nas redes sociais, nosso objeto de estudo no próximo capítulo.

### **4.3 Redes sociais**

Nunca se falou tanto em redes sociais como nestes últimos tempos, em que a *internet* vem-se tornando cada vez mais um modo de vida, possibilitando uma verdadeira revolução na maneira como comunicamos com outras pessoas. As redes promovem interação social e o aumento significativo dos elos informacionais, conectando pessoas de diferentes locais possibilitando, assim, a transferência de informação.

De acordo com Castells (2003), com a evolução social associada à tecnologia da informação, uma nova base material está sendo construída em rede, definindo os processos sociais e moldando a própria estrutura social. Corroborando com essa mesma ideia, Marteleto (2001, p.72) entende rede social como “[...] um conjunto de participantes autônomos, unindo ideias e recursos em torno de valores e interesses compartilhados” Desta forma, em uma rede social, a interação entre os indivíduos proporciona a troca e o compartilhamento de informação e conhecimento. Com o advento das TICs, as redes sociais passaram a existir também no plano virtual, possibilitando uma maior interação e maiores fluxos informacionais entre os internautas.

Ao analisar estruturalmente as redes sociais virtuais, Recuero (2004) afirma que:

[...] a análise estrutural das redes sociais procura focar na interação como primado fundamental do estabelecimento das relações sociais entre os agentes humanos, que originarão as redes sociais, tanto no mundo concreto, quanto no mundo virtual. Isso porque em uma rede social, as pessoas são os nós e as arestas são constituídas pelos laços gerados através da interação social. (RECUERO, p. 4, 2004)

Atualmente, 77 milhões de brasileiros fazem parte de alguma rede social na *internet*<sup>13</sup>. O Brasil é um dos países com maior adesão às redes sociais no mundo. O *Orkut* foi um dos sites que contribuiu para a popularização das redes sociais no país. No entanto, nos últimos tempos ele vem perdendo espaço para outras redes como *Facebook* e *Twitter*. Segundo pesquisa do *Twitalyzer*<sup>14</sup>, o Brasil é o terceiro país onde mais se usa o *Twitter* no mundo. Seguindo essa tendência de crescimento, o *Facebook*, de acordo com pesquisa realizada pelo Ibope Nielsen Online<sup>15</sup>, ultrapassou o *Orkut* e já é a rede social mais acessada no Brasil. Em nenhum outro lugar no mundo existe um entusiasmo tão grande pelas amizades virtuais como no Brasil. Estudos realizados pela empresa E.Life<sup>16</sup> mostram que, para cada quatro minutos na rede, os brasileiros reservam um para atualizar seu perfil e observar o dos amigos.

Entre as 10 maiores redes sociais no mundo, estão o *Facebook*<sup>17</sup>, com 1,2 bilhões de acessos por mês, o *Myspace*<sup>18</sup> mais de 800 milhões de visitas e o *Twitter* aproximadamente 56 milhões de acessos mensais. No Brasil, este ranking é um pouco diferente. De acordo com o infográfico apresentado pela *BBC News*<sup>19</sup>, com pesquisa da *Nielsen Company*<sup>20</sup> em junho de 2010, aponta como redes sociais mais acessadas no Brasil, o *Orkut* com 26 milhões de visitantes únicos no mês, e o *Facebook* e *Twitter*, empatados com 10,7 milhões<sup>21</sup>.

O *Orkut* foi criado em Janeiro de 2004 pelo engenheiro turco da *Google*, Orkut Büyükkökten, com o intuito de ser uma rede social *on-line* onde seus membros criassem novas amizades e mantivessem relacionamentos, bem como fizessem parte de comunidades de assuntos de seu interesse. Quando o *Orkut* foi criado, rapidamente virou sensação no Brasil. Não se falava em outra coisa. A possibilidade de se relacionar virtualmente com pessoas de vários lugares, reencontrar os amigos que tinham se perdido pelo caminho e participar de diversas comunidades com temas variados, fez com que a rede social tornasse febre entre os brasileiros. Durante muito tempo, o *Orkut* foi considerado a principal rede social *online* no Brasil.

---

<sup>13</sup> Pesquisa divulgada pelo Ibope em Agosto de 2011.

<sup>14</sup> Serious Analytics for Social Business. <http://twitalyzer.com/>

<sup>15</sup> <http://br.nielsennetpanel.com/pnl/br/home>

<sup>16</sup> Empresa responsável pela monitoração, análise da mídia gerada pelo consumidor e gestão de relacionamento em mídias sociais na América latina e Portugal.

<sup>17</sup> [www.facebook.com](http://www.facebook.com)

<sup>18</sup> [www.myspace.com](http://www.myspace.com)

<sup>19</sup> Departamento dentro da British Broadcasting Corporation (BBC) responsável pela área de jornalismo e pela produção de seus programas de notícias tanto para televisão como para rádio e internet.

<sup>20</sup> Nielsen Company, uma empresa global de informações e mídia com posições de liderança na indústria de informações de mercado e consumidor, televisão, inteligência on-line, mensuração de telefonia celular, feiras e eventos e publicações comerciais (*Billboard*, *The Hollywood Reporter* e *Adweek*).

<sup>21</sup> Disponível em <<http://www.administradores.com.br>>



Concomitantemente, outras redes foram criadas naquela época, como o *Myspace*, criado em 2003, tornando-se muito popular, principalmente entre músicos e bandas, devido à sua habilidade de hospedar *MP3s*<sup>22</sup>.

Em 2004, foi criado também o *Facebook*, que assim como o *Orkut* é uma rede social *online* onde cada pessoa pode ter o seu perfil, ou seja, os seus dados pessoais, as suas fotos, vídeos, links, notas etc. Os membros desta rede social interagem entre si, visitando os perfis, fazendo amigos, estabelecendo contatos, deixando comentários, enviando mensagens entre si. O site foi criado por Mark Zuckerberg, um programador, empresário e filantropo norte-americano nascido em 1984. No início, tinha como alvo apenas os estudantes da Universidade de Harvard, mas progressivamente foi permitindo a inscrição de estudantes de outras escolas até que, em 2006, estava disponível para todos.

Atualmente, muitos usuários estão trocando o *Orkut* pelo *Facebook*, em virtude da popularização deste último. Muitos usuários falam até em “*orkuticídio*”, referindo-se à iminente morte da rede social *Orkut* ou aos “suicídios” de vários perfis que têm ocorrido na rede social. No entanto, segundo algumas pesquisas, o Brasil ainda é um dos que têm menos perfis cadastrados no *Facebook*. De acordo com a pesquisa realizada pela agência de marketing digital britânica *iCrossing*<sup>23</sup>, a rede social que predomina nos Estados Unidos e Canadá ainda não caiu completamente no gosto dos brasileiros. Porém, em virtude da característica mutante da *internet* e suas ferramentas, este cenário ainda vai mudar e muito.

O *Twitter* é outra rede social que vem se destacando no mundo e principalmente entre os internautas brasileiros. A princípio até pode parecer uma ferramenta, digamos, sem muito propósito, onde o usuário escreve em um espaço que contém 140 caracteres o que ele está pensando ou fazendo naquele exato momento. No entanto, além de ficar sabendo o que determinado usuário está fazendo, o *Twitter* possibilita que seus usuários fiquem atualizados sobre os acontecimentos que ocorrem no mundo. E como isso funciona? Mais adiante será comentado mais sobre o *Twitter*. Antes porém, vamos abordar um recurso muito utilizado nas redes sociais para a organização da informação: as *hashtags*.

---

<sup>22</sup> Padrão para comprimir arquivos de áudio.

<sup>23</sup> Empresa de marketing digital especializada em soluções colaborativas.

#### 4.4 Hashtags

Em ambientes colaborativos, os usuários utilizam as *tags*<sup>24</sup> ou *hashtags* em um processo conhecido como *tagging* (etiquetagem), atribuindo termos a conteúdos da *Web*. É um processo de indexação livre em linguagem natural<sup>25</sup> sem regras, política ou vocabulário controlado.

Atualmente existem vários termos para designar esse processo de etiquetagem. Para alguns autores como Mathes (2004) “etnoclassificação” seria o mais adequado. Há ainda autores, como Peterson (2006) que preferem o termo *social bookmarking*. Hammond et al. (2005) considera mais pertinente os termos “classificação social” ou “classificação distribuída” para designar esse processo de indexação realizado pelos usuários.

De acordo com Catarino e Batista (2007) a vantagem de utilizar as *hashtags* está na

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; BAPTISTA, 2007, p. 48).

Neste sentido, é importante destacar que são distintas as motivações dos usuários ao etiquetar conteúdos na *Web*. De acordo com Cañada (2006), existem quatro perspectivas no processo de etiquetagem realizado pelos usuários. São elas: egocêntrica, amigável, altruísta e populista. A perspectiva egocêntrica diz respeito à descrição pessoal da informação, tendo em vista a sua recuperação futura, sem a preocupação de essa informação ser útil a outros usuários. Assim, como lembra Moura (2009), a indexação realizada funciona como “lembretes mnemônicos que revelam o universo semântico do criador”. No que tange à perspectiva amigável, o autor lembra que ela equivale ao primeiro nível de colaboração, onde os usuários compartilham uma mesma linguagem na descrição dos objetos informacionais. O segundo nível de cooperação refere-se à perspectiva altruísta, onde os usuários tentam estabelecer alguma regularidade e padronização na utilização das *tags*, favorecendo assim a recuperação da informação de forma efetiva para um número maior de usuários. Finalmente, a perspectiva populista utiliza etiquetas de apelo popular para divulgar conteúdos comerciais, incorporando assim a dinâmica colaborativa. De acordo com Moura (2009) essa perspectiva “viabiliza comercialmente o monitoramento das práticas colaborativas na *web* ao vincularem o universo semântico de determinadas etiquetas à oferta de bens e serviços”.

<sup>24</sup> Tag é uma palavra-chave ou termo associado a uma informação.

<sup>25</sup> Linguagem natural, segundo Lancaster (2004, p.250), é sinônimo de discurso comum, isto é, a linguagem utilizada habitualmente na escrita e na fala, e que é o contrário de ‘vocabulário controlado’.

Entender a existência dessas perspectivas nos ajuda a compreender alguns problemas gerados pela utilização da indexação colaborativa em ambientes como as redes sociais. A seguir, faremos um estudo mais detalhado sobre a folksonomia.

#### 4.5 Folksonomia

*Folksonomia* é a tradução do termo *folksonomy* que é um neologismo criado em 2004, pelo então analista de sistemas e consultor norte-americano Thomas Vander Wal, a partir da união de *folk* (povo, pessoas) com *taxonomy*. Desta forma, *folksonomia* é o resultado da atribuição livre e pessoal de *tags* (descritores) a informações ou objetos, visando à sua posterior recuperação. É um processo de indexação colaborativa, ausente da ação de indexadores profissionais ou ferramentas automáticas de indexação.

A *folksonomia* surge como um novo conceito e como uma alternativa para a organização do conhecimento em ambientes colaborativos. Ela é apresentada como uma indexação que tem como objetivo facilitar a recuperação da informação, uma vez que é realizada pelo próprio usuário.

*Folksonomia* é o resultado da atribuição livre e pessoal de etiquetas às informações dos recursos na *Web*, em um ambiente social, compartilhado e aberto a outros, pelos próprios usuários da informação, visando a sua recuperação. Destacam-se, portanto, três fatores essenciais: o primeiro é o resultado de uma indexação livre, feita pelo próprio usuário do recurso; o segundo objetiva a recuperação posterior da informação e o terceiro é desenvolvida num ambiente aberto que possibilita o compartilhamento e, até, em alguns casos, a sua construção conjunta. (CATARINO; BAPTISTA, 2007, p. 49).

Assim como *web 2.0*, a *folksonomia* é realizada de maneira colaborativa, “funciona através da atribuição de *tags* (etiquetas), pelos próprios usuários da *web*, a arquivos disponibilizados *online*. Assim, é o usuário que representa e recupera informações através das *tags* que ele mesmo cria”. (AQUINO, 2008, p. 305). Neste sentido, a principal característica das *tags*, segundo Amaral e Aquino (2008, p. 2), reside no fato de que estas podem ser

“ [...] criadas por qualquer usuário na *web*, ao contrário do que ocorria no início da *web* onde apenas programadores e conhecedores de linguagens de programação editavam *hipertextos*, e porque são criados com base no significado das informações etiquetadas.”.

A finalidade da *folksonomia* seria organizar o caos existente na *web*. Embora a sua característica de liberdade para indexar aponte para a ideia de uma falta de estrutura organizacional, o resultado para quem pesquisa é uma maior facilidade para encontrar termos que as demais linguagens de indexação não conseguem acompanhar em suas tabelas hierárquicas. (SILVA; BLATTMANN 2007, p. 207).

É importante destacar que existe uma grande discussão acerca do conceito de *folksonomia*. Alguns autores a entendem como sendo o resultado de um processo. Outros a enxergam como o próprio processo. Neste sentido

Parece claro que para o criador do termo, *folksonomia* é o resultado de um processo; no entanto, os autores dividem-se em dois grupos: 1) os que entendem a *folksonomia* exatamente como o resultado de um processo, como um produto [...] e 2) os que se referem a *folksonomia* como sendo um sistema, uma metodologia, ou abordagem, ou o próprio processo. (CATARINO; BAPTISTA, 2007, p. 32).

O uso da *folksonomia* como ferramenta de representação da informação é característica intrínseca a *web 2.0* que é caracterizada por um ambiente social e colaborativo. Destarte, existem autores que preferem o termo *social bookmarking*, enfatizando o aspecto colaborativo destas ferramentas. De acordo com Campbell (2006), são ferramentas que possibilitam que usuários marquem suas páginas e atribuam etiqueta para representar seus temas de interesse.

A *web 2.0* trouxe a necessidade de se criarem novas formas para a organização da informação, uma vez que a taxonomia se mostra muito rígida para um ambiente colaborativo.

Dentre as diversas evoluções que estão ocorrendo, destaca-se o que pode ser considerado como um novo paradigma para a organização dos conteúdos dos recursos digitais na Web: a possibilidade de os próprios usuários participarem na organização desses conteúdos. Neste novo paradigma surgem as *folksonomias*. Trata-se de um novo conceito que tem sido utilizado por diversos profissionais e estudiosos da área de informação. No entanto, parece não haver ainda um consenso, quer sobre a utilização deste termo, quer sobre o seu significado. Há os que preferem utilizar outros termos como, por exemplo, classificação social ou *social tagging*. (CATARINO ; BAPTISTA , 2009, p. 47-48).

A *folksonomia* caracteriza-se pela liberdade de indexação dos termos por parte dos usuários. Segundo Catarino e Baptista (2009), trata-se de uma indexação livre em linguagem natural, sem controle de vocabulários e sem política de indexação, ou seja, não há efetivamente a tradução dos termos para uma linguagem artificial.

Porém esta característica de liberdade de indexação e ausência de controle de vocabulários, intrínseca à *folksonomia*, traz muitas discussões acerca das vantagens e desvantagens do uso deste tipo de ferramenta na *web*.

Se por um lado esta parece ser a melhor maneira de organizar as informações em ambientes colaborativos como as redes sociais, onde quem produz e classifica as informações são os usuários, por outro lado, este tipo de organização traz muitos problemas que influenciam na precisão da informação recuperada, denominada *atinência*.

Dentre os diversos problemas que o uso do *folksonomia* pode acarretar, podemos destacar a polissemia, sinonímia, o uso do plural, abreviações e erros ortográficos. Segundo Golder e Hoberman (2006), a liberdade na atribuição tem como resultado um conjunto de palavras-chave que muitas vezes podem ser ambíguas, inexatas e, conseqüentemente, irão afetar a precisão na recuperação da informação.

Na seção seguinte abordaremos mais detalhadamente a Semântica, com o intuito de compreendermos melhor alguns dos problemas semânticos que podem vir ocorrer em virtude do uso *folksonomia*.

#### **4.6 Semântica**

Semântica pode ser entendida como a ciência das significações. É o ramo da linguística que trata do significado das palavras, considerando também as mudanças de sentido, a escolha de novas expressões, o nascimento e morte das locuções. (SIMON, 1990)

O interesse pela semântica vem crescendo nos últimos tempos em virtude da revolução surgida na linguística moderna quando a língua passa a ser vista como “uma estrutura altamente organizada constituída por elementos interdependentes, e da ideia de que as palavras desempenham um papel crucial no acto (sic) de modelar os nossos pensamentos e de os conduzirem por determinados canais.” (ULLMANN, 1999, p.5)

A preocupação com a linguagem remonta ao século V a.C com os filósofos gregos, que desde então já tentavam interpretar os fenômenos naturais e suas relações com as instituições sociais. Buscava-se entender as relações entre as palavras e as coisas que elas nomeavam.

Se aos elementos do mundo correspondiam nomes, um dos aspectos fundamentais dos estudos da linguagem era determinar se as palavras se associavam naturalmente às coisas que se referiam ou se as palavras nomeavam as coisas em decorrência de seu uso na intercomunicação social: a relação entre as palavras e as coisas era arbitrária e resultado de convenção. (MARQUES, 1999, p. 26)

Segundo Marques (1999), as palavras possuem valor simbólico e o seu significado é o valor mental que elas transmitem. As palavras não significam diretamente as coisas a que se referem. Daí nascem os estudos semânticos que irão tratar das diversas formas de interpretação do significado.

De um lado é possível considerar que uma palavra tem tantos sentidos quantas sejam as suas diversas realizações contextuais. De outro lado, pode-se interpretar que a indeterminação inerente ao significado decorre do fato de uma palavra ter um sentido básico, a que se somam fatores contextuais, lógicos, emotivos, combinatórios, evocativos e associativos, que introduzem nuances interpretativas diversas, no mesmo significado básico. (MARQUES, 1999, p.61)

Como bem lembra Marques (1999), uma palavra pode adquirir vários significados em virtude de fatores diversos, levando ao que se convencionou chamar de relações equívocas (homônimos), multívocas (heterônimos) e diversívocas (polinônimos) entre palavras. Daí surgem os fenômenos conhecidos como homonímia, sinonímia e antonímia, que discutiremos mais adiante.

A antonímia é a relação de oposição que se estabelece entre duas palavras que apresentam significados contrários. As palavras velho e novo, terra e mar, falar e ouvir, bom e ruim são palavras ditas antônimas. Quando duas ou mais palavras apresentam significados iguais ou semelhantes, temos uma relação de sinonímia. Como exemplo, podemos citar as palavras novo-recente, distante-afastado.

Outros tipos de relações entre as palavras são a polissemia e a homonímia. Os dois conceitos muitas vezes são confundidos. Alguns autores desconsideram o conceito de homonímia, acreditando-se que exista uma diferenciação apenas no nível da polissemia e monossemia. Outros autores acreditam que a distinção entre polissemia e homonímia está na derivação etimológica das palavras ou da relação/falta de relação entre os significados.

#### 4.6.1 Polissemia e homonímia

De acordo com Lyons (1987), a polissemia pode ser entendida como a propriedade que uma mesma palavra tem de apresentar vários significados além do seu sentido original. O vocábulo “letra”, por exemplo, apresenta vários significados: pode ser um dos sinais gráficos do alfabeto ou mesmo o texto de uma canção. A homonímia, por sua vez, ocorre quando dois vocábulos diferentes, de origens e significados diversos, terminam convergindo para a mesma configuração fonológica. Um exemplo clássico é a palavra “manga”, em que não é possível estabelecer uma relação satisfatória entre os significados. Além disso, derivam de línguas diferentes: a fruta vem do Malaio *manga*, enquanto a parte da vestimenta vem do Latim *manica*.

Homonímia deriva de uma relação equívoca<sup>26</sup> de duas ou mais palavras que possuem a mesma estrutura fonológica<sup>27</sup>, embora tenham significados diferentes. Marques (1999) sugere como exemplo a palavra *nó* que pode ter “conceitos ‘próprios’ distintos”. Pode ser uma unidade de medida, entrelaçamento de uma ou mais cordas ou ainda, partindo para um sentido mais figurado pode corresponder a vínculo, união.

A polissemia vem ganhando grande importância nos estudos linguísticos, principalmente virtude do interesse dos linguistas pelos aspectos voltados à cultura, à cognição e ao uso linguístico. Se antes a Gramática Gerativa<sup>28</sup> fundamentava-se na descontextualização social do funcionamento linguístico e cognitivo, acreditando-se em um único significado para um item lexical, hoje a polissemia passa a ter grande notoriedade nos estudos linguísticos principalmente com o advento da Linguística Cognitiva<sup>29</sup> que, em suas análises, leva em consideração os contextos social, cognitivo e situacional. A partir deste momento, a polissemia, como lembra Amaral (2011), deixa de ser considerada um problema e passa a ser vista como uma característica atrelada à linguagem, seu uso e processamento.

É importante destacar também que uma mesma palavra pode adquirir significados diferentes em diferentes países. Existem algumas palavras usadas em

---

<sup>26</sup> Mesmo nome com sentidos diferentes.

<sup>27</sup> A Fonologia é um ramo da Linguística que estuda o sistema sonoro de um idioma, no que tange a sua organização dentro de uma língua, classificando-os em unidades capazes de distinguir significados, chamadas fonemas.

<sup>28</sup> A gramática Gerativa é um teoria linguística elaborada pelo linguista americano Noam Chomsky, discute aspectos linguísticos como a criatividade do falante e sua capacidade de emitir e de compreender frases inéditas.

<sup>29</sup> A Linguística Cognitiva é uma abordagem da linguagem perspectivada como meio de conhecimento e em conexão com a experiência humana do mundo. De acordo com esta perspectiva, as unidades e as estruturas da linguagem são estudadas, não como se fossem entidades autônomas, mas como manifestações de capacidades cognitivas gerais, da organização conceitual, de princípios de categorização, de mecanismos de processamento e da experiência cultural, social e individual.

Portugal, por exemplo, que nada têm a ver com o significado que elas recebem no Brasil. E há ainda algumas palavras e termos que não se assemelham e outros que, mesmo sendo iguais, significam coisas diferentes.

A seguir, na figura 2, apresentamos alguns desses casos.

<b>Brasil</b>	<b>Portugal</b>
Camisinha	Durex
Durex	Fita cola
Cafezinho	Bica
Fila	Bicha
Homossexual	Paneleiro
Sapatão	Fufa
Pão francês	Cacete
Crianças	Canalhas
Adolescente	Puto
Calcinha	Cueca
Menstruada	Estar com histórias
Absorvente feminino	Penso higiênico
Dentista	Estomatologista
Professor particular	Explicador
Comissária de bordo	Hospedeira
Garis	Almeidas
Salva vidas	Banheiros
Cego	Invisual
Chiclete	Pastilha elástica
Injeção	Pica
Embebedar-se	Enfrascar-se
Impostos	Propinas
Tesão	Ponta
Alô	Está lá?
Mulherengo	Marialva
Peruca	Capachinho

**Figura 2: Variação do significado de palavras No Brasil e em Portugal**

Fonte: <http://linoresende.jor.br/sera-preciso-traduzir>



## 5. METODOLOGIA

### 5.1 Netnografia

A pesquisa foi realizada através do método da “netnografia”, que segundo Kozinets (2002, p. 31) trata-se de uma descrição escrita resultante do trabalho de campo que estuda as culturas e comunidades *online* emergentes, mediadas por computador, ou comunicações baseadas em *internet*.

A netnografia é um ramo da etnografia que analisa o comportamento dos indivíduos de comunidades virtuais e da cibercultura. Originado no campo da antropologia, o método etnográfico “consiste na inserção do pesquisador no ambiente, no dia-a-dia do grupo investigado” (VERGARA, 2005, p. 73, 195). Segundo Kozinets (2002, p. 22),

A netnografia diz respeito à fusão de técnicas etnográficas estabelecidas e inovadoras, adaptadas ao estudo naturalista de comunidades virtuais, e a sua rerepresentação da pesquisa, procurando obter experiências profundas da sociabilidade digital.

Desta forma, foram feitas buscas no microblog *Twitter* com as *hashtags*<sup>30</sup> #CALABOCAGALVAO, #Gol, #Rio, #Parangolixo e #ldol e #ufmg. Com exceção da *hashtag* #ufmg, a escolha das outras *hashtags* deu-se por causa do sucesso alcançado por elas e em virtude do significado e da importância que elas tiveram no momento em que foram criadas. O “CALA BOCA GALVAO” foi um fenômeno do *Twitter* que envolveu o locutor esportivo Galvão Bueno<sup>31</sup> responsável pela transmissão oficial das Copas do Mundo e das Olimpíadas nas transmissões da Rede Globo. O assunto se manteve entre os dez no *trending topics*, uma espécie de ranking dos assuntos que foram mais abordados no *Twitter* em um determinado período.

---

<sup>30</sup> No *Twitter*, as *hashtags* são *tags* utilizadas para indexar assuntos com o mesmo conteúdo. Os usuários utilizam o sinal # (*hash*) para agregar seu *tweet* a outros que também tenham a mesma palavra-chave.

<sup>31</sup> Carlos Eduardo dos Santos Galvão Bueno, mais conhecido como Galvão Bueno (Rio de Janeiro, 21 de julho de 1950), é um narrador, radialista e apresentador esportivo brasileiro. Atualmente, trabalha para a Rede Globo de Televisão.

A *hashtag* “Parangolixo” surgiu quando integrantes da banda de *Heavy Metal*, Angra discutiram no *Twitter* com o vocalista do Parangolé, Léo Santana<sup>32</sup>. A banda paulista de *Heavy Metal* acusa o grupo baiano de axé de ter plagiado o *riff*<sup>33</sup> de guitarra de sua música “Nova Era”. O trecho teria sido usado na música “Tomba Aê”, da banda Parangolé. O assunto chegou ao primeiro lugar no *trending topics* Brasil.

Outra *hashtag* analisada foi “Rio”. Essa *tag* foi comentada em virtude do lançamento do filme *Rio* animação em 3D da 20th Century Fox<sup>34</sup> e *Blue Sky Studios*<sup>35</sup>. No entanto, esta *hashtag*, apresenta uma ambiguidade em virtude dos seus diversos significados. Rio pode ser usada para se referir ao filme, a cidade Rio de Janeiro ou mesmo a rio, curso natural de água.

Analisamos também os resultados das buscas com a *hashtag* “Gol”, referente à empresa de aviação “Gol Linhas Aéreas”. A *hashtag* está sendo muito comentada em virtude da atuação da empresa no *Twitter*, divulgando suas promoções de passagens aéreas pela *internet*.

Fizemos também uma busca com a *hashtag* #Idol relacionada ao programa de calouros American Idol<sup>36</sup>. Ele tornou-se o programa musical com maior publicidade nos Estados Unidos. O programa que visa descobrir o melhor cantor do país é sempre muito comentado no *Twitter*. Os resultados obtidos com as buscas com a *hashtag* nos levaram a necessidade de se fazer uma outra busca, desta vez pelo perfil *AvistaHotelsfl*. Esta necessidade ocorreu em virtude da existência de vários *posts* terem sido indexados com a *hashtag* #Idol, porém, sem relação alguma com o Programa de TV. Verificou-se que a o perfil *Avistahtalsfl* estava se aproveitando da popularidade da *hashtag* apenas para fazer propaganda dos hotéis da rede hoteleira Avista Hotel<sup>37</sup>. Ao analisarmos o perfil *AvistaHotelsfl*, nos deparamos com uma forma de disseminação de *spams*, utilizada pela empresa. Esta forma usa os perfis do *Twitter* para mensagem publicitária. As pessoas ou empresas efetuam vários *replies*<sup>38</sup>,

---

<sup>32</sup> Leandro Silva de Santana (Salvador, 22 de abril de 1988), mais conhecido como Léo Santana, é um cantor e dançarino brasileiro.

<sup>33</sup> Riff é uma progressão de acordes, intervalos ou notas musicais que são repetidas no contexto de uma música, formando a base ou acompanhamento. Geralmente, eles formam a base harmônica de músicas de jazz, blues e rock.

<sup>34</sup> <http://www.foxfilm.com.br>

<sup>35</sup> <http://www.blueskystudios.com>

<sup>36</sup> American Idol (em português Ídolo Americano) é uma competição americana de calouros que teve sua primeira temporada em junho de 2002. Endereço: <http://www.americanidol.com>.

<sup>37</sup> Fundada em 1989, Avista Hotels & Resorts é uma empresa que controla diversos hotéis em Orlando, Flórida.

<sup>38</sup> Reply é uma ferramenta do *Twitter* utilizada para responder a um *tweet* de algum usuário.

geralmente sem contexto algum, para divulgarem seus anúncios. Geralmente, os *spammers*<sup>39</sup>.

A questão da proliferação dos *spams* no *Twitter* tem ascendido uma discussão acerca da necessidade dos desenvolvedores do *blog* criarem mecanismos mais eficazes para conter este problema. No entanto, a criação destes mecanismos suscita um outro debate muito comentado nos últimos tempos, que é a questão da possível existência de alguma forma de censura por parte do microblog. Sabe-se que entre o controle e a censura existe uma linha muito tênue que muitas vezes é ignorada. Neste sentido, é imprescindível se pensar em ferramentas de controle para combater os *spams*, porém não se pode esquecer também da necessidade de não ferir a liberdade de expressão dos usuários do *Twitter*.

Por fim, analisamos também a *hashtag* *#ufmg*, por apresentar características bem diferentes dos outros termos analisados, uma vez que ela caracteriza-se por ser uma *hashtag* estática, ou seja, apresenta uma qualidade mais atemporal do que as outras. Acreditamos que pelo fato dessa *hashtag* não figurar entre os *trending topics no momento da nossa pesquisa*, ela apresenta menos ruídos, como por exemplo, os *spams*.

Através dos resultados das buscas, foram observadas as estratégias usadas pelos usuários para etiquetagem dos *posts* recuperados. Desta forma, foi possível avaliar se os usuários estão usando as *hashtags* apenas para alcançarem um maior alcance no *Twitter* ou se realmente estão indexando de acordo com o conteúdo dos seus *posts*.

A análise dessas estratégias permitiu entender como alguns fatores inerentes ao uso da *folksonomia*, como a polissemia e a intenção dos usuários no momento dos *posts*, tem influenciado a precisão na recuperação das informações dentro do *Twitter*.

## **5.2 *Twitter: ambiente de pesquisa***

*Twitter* é uma rede social de *microblog* onde o usuário escreve pequenos textos, de até 140 caracteres, que se destinam a responder a seguinte questão: *O que você está fazendo?* A mensagem fica disponível via *SMS*, mensageiro instantâneo, *e-mail*, site oficial ou programa especializado. (HIROSH, 2008). É uma forma de comunicação, rápida e ágil, que permite acesso às informações e cobertura de

---

<sup>39</sup> Spammers são pessoas que enviam diversos e-mails ou qualquer outro tipo de mensagem para várias pessoas, sem que elas tenham autorizado.

notícias em tempo real. Além disso, proporciona ao leitor que ele mesmo publique suas informações, ao mesmo tempo em que lê outras.



**Figura 3: Página inicial do Twitter**  
Fonte : [www.twitter.com](http://www.twitter.com)

Apesar de ser um recurso de comunicação na *web* recente, o *Twitter* já atingiu um público muito grande em um curto período. O seu sucesso deve-se principalmente ao fato de ele ser uma plataforma simples, de fácil acesso e que define um novo tipo de comunicação social. Além disso, é dinâmico, interativo, gratuito, público, objetivo, democrático, de fácil acesso e vem-se tornando a cada dia uma importante fonte de informação. Na figura 3, temos a página inicial do *Twitter*.

Desde 2006, o *Twitter* vem crescendo tanto em números como em importância. Se a proposta inicial dos seus desenvolvedores era apenas a de ser uma rede onde as pessoas pudessem trocar experiências como, por exemplo, o que comeu no café da manhã, hoje o *Twitter* atingiu proporções tão significativas a ponto de se tornar um fenômeno virtual onde notícias são divulgadas antes mesmo de chegarem nos meios de comunicação tradicionais.

Na figura 4 a seguir, apresentamos a tela de um perfil usuário do microblog.



**Figura 4: Perfil no Twitter**  
Fonte : Twitter

Uma das funcionalidades mais interessantes do *Twitter* é a atualização, em qualquer lugar que você esteja, por meio de celulares e *smartphones*<sup>40</sup>.

As atualizações de textos são exibidas no perfil do usuário e enviadas àqueles que seguem aquele perfil (*followers*). Os usuários do *Twitter* podem receber essas atualizações através do site oficial ou por *RSS*<sup>41</sup>, *SMS* ou programa especializado.

O fenômeno do *Twitter*, segundo Hiroshi (2008), é conhecido como microblog e consiste de textos curtos que vão resumindo as atividades diárias,

<sup>40</sup> Os *smartphones* são celulares inteligentes, pois possuem diversas funções, além de receber ou fazer ligações. É como se fosse um computador de bolso, já que é possível acessar a internet, mandar e receber e-mails, criar ou receber diversos tipos de arquivos com formatos diferentes, tirar fotos, ouvir e armazenar músicas, além de várias outras ferramentas.

<sup>41</sup> *RSS* (Really Simple Syndication) é um formato baseado na linguagem XML utilizado para a distribuição de conteúdos. Com ele, é possível reunir informações de vários sites em um único ambiente.

disponível em uma página da Internet, em ordem cronológica inversa. O *Twitter* entre blogueiros<sup>42</sup> serve para anunciar *links* da *web* e para fazer pequenos *posts*.

Como afirma Seixas,

O *Twitter* foi o precursor, aquele que definiu o conceito, as novas possibilidades e a nova forma de irrigar o mundo com conteúdo. Permitiu que uma verdadeira legião de programas, sites e mash-ups<sup>43</sup> pudessem proporcionar formas diferentes de publicar e interagir com a quantidade colossal de mini-conteúdos já disponível nesse pequeno, e crescente, universo. Já se percebe que ao redor dessa plataforma se forma uma economia própria, onde ela própria é o início e o fim, o meio e o objetivo. (SEIXAS, 2009, p. 45).

Com o *Twitter*, alguns termos tornaram-se populares entre os internautas e usuários da ferramenta. É muito comum, hoje em dia, no ciberespaço, ver as pessoas comentando sobre os assuntos que estão nos *trending topics*, sobre os *followers* e *following* de determinado usuário, sobre um *retweet* postado ou sobre um *reply* direcionado a determinado *post* no microblog. Todos esses termos são conhecidos por todos aqueles que utilizam o *Twitter*. Mas o que significa cada um destes termos? Qual a sua finalidade? Recorremos ao próprio *Twitter* para tentar responder a essas questões, como será mostrado a seguir.

*Followers* significa as pessoas que seguem determinado usuário. Já *Following* são as pessoas que determinado usuário segue. Na figura 5, temos um exemplo de *following* e *followers* de determinado usuário. Nela podemos constatar que o usuário segue 155 pessoas e 80 pessoas o seguem.

---

<sup>42</sup> Blogueiro ou bloguista ou ainda blóguer ou *blogger* são palavras utilizadas para designar aquele que escreve em blogs.

<sup>43</sup> Um *mashup* é a combinação de dois aplicativos que podem complementar e melhorar a oferta de determinado serviço.



**Figura 5: Follow e Followers**

Fonte: Twitter

Dentre as ferramentas que o *Twitter* possui, os *trending topics* vêm ganhando um lugar de destaque entre os internautas. Os *trending topics* ou *Tópicos da Tendência* é uma lista dos termos mais comentados no microblog em determinado momento. Esta lista é composta pelas *hashtags* e/ou mesmo por nomes de determinados perfis da rede. Estar entre os *trending topics* significa que aquele assunto está sendo comentado por um grande número de pessoas em um dado período. É muito comum ouvirmos, nas mídias tradicionais, que determinado assunto foi um dos mais comentados no *Twitter* naquele dia. Estar entre os tópicos mais comentados ou mesmo fazer comentários sobre algum termo que se encontra na lista pode contribuir para aumentar a audiência de determinado perfil dentro da rede. E, cientes deste fato, empresas e usuários se aproveitando do impacto desta ferramenta para divulgarem seus produtos e marcas ou mesmo para aumentar a audiência do se perfil dentro do *Twitter*.

Atualmente já existem sites que divulgam os assuntos mais comentados no dia no *Twitter*. Um desses sites é o site da G1.Globo.com<sup>44</sup>, sobre Tecnologia e Games. Todos os dias, esse site faz um levantamento dos assuntos que entraram na lista dos *trending topics* em São Paulo, Rio de Janeiro, no Brasil e no mundo (Figura 6 e Figura 7). Mais adiante, falaremos mais sobre o *trending topics*.

<sup>44</sup> <http://www.globo.com/>



Edição especial comemora 15 anos da série 'Resident Evil'

veja também: jogos | downloads

GOOGLE+ **Criminosos usam convites em golpes**

DIZEM ESPECIALISTAS **Twitter precisa melhorar sua segurança**

**segurança digital**

últimas notícias

11 DE JULHO DE 2011

19h34 **Assuntos do dia no Twitter - segunda-feira, 11/7/2011**

19h02 **Multiprocessador que cozinha e 'vitrola combo' são exibidos em feira**

PACOTÃO **Confira como ter segurança na conexão via 3G**

PARA ACESSO **Criminosos sabotam programa para aceitar ':' como usuário**

IPHONE E IPAD **Site que desbloqueia volta a**

Shopping

eFácil Micro System Philco PH6... 12 x R\$32,4

compare preços de

veja todos os produtos >

Figura 6: Assuntos do dia no Twitter 1

Fonte: <http://g1.globo.com/tecnologia/>

Músicas: Ouvir mús... Downloads O QUE SÃO FOLLO... Central de Ajuda de... G1 - Assuntos do di... O que é o Twitter?...

g1.globo.com/tecnologia/noticia/2011/07/assuntos-do-dia-no-twitter-segunda-1172011.html

11/07/2011 19h34 - Atualizado em 11/07/2011 19h34

## Assuntos do dia no Twitter - segunda-feira, 11/7/2011

Entenda os termos que se destacaram na rede social. Veja lista de assuntos no Mundo, Brasil e nas capitais de SP e RJ.

Do G1, no Rio imprimir

Estes são os principais assuntos que entram na lista de **"Trending Topics"** da rede social de microblog Twitter nesta **segunda-feira, 11 de julho de 2011**. O Twitter mede os assuntos do dia em **todo o mundo** e também entre os usuários do **Brasil** e de duas cidades brasileiras: **São Paulo** e **Rio de Janeiro**.

Brasil **Trending Topics**

Trends - Brazil - change  
#designstar Promoted  
Aquamarine  
#ultimatepotter  
Carlitos Tevez  
#forzapalestra  
#aloselecao  
Sherlock Holmes  
Frizzo  
Lucas Piazon

**Aquamarine: atração da Sessão da Tarde**, da TV Globo, desta segunda-feira (11). "Aquamarine" conta a história de duas amigas que encontram uma sereia na piscina de um clube e a ajudam a conquistar seu grande amor.

**#ultimatepotter**: promoção com distribuição de prêmios de uma revista sobre Harry Potter. O desafio é acertar quantas diademas há em uma figura postada no Twitter. Tópico relacionado: **Diademas**.

**Carlitos Tevez: Corinthians estaria negociando a craque argentino**. Atacante do Manchester

Aguardando pagead2.googleyndication.com...

Iniciar SmartRAM G1 - Assunt... TCC 4 Microsof... Imagem - Paint PT 19:42

Figura 7: Assuntos do dia no Twitter 2

Fonte: <http://g1.globo.com/tecnologia/noticia/2011/07/assuntos-do-dia-notwitter-segunda-1172011.html>



Outro termo muito conhecido entre os usuários do microblog é o *tweet* que nada mais é que qualquer mensagem postada no *Twitter*. A explicação para a restrição quanto ao número de caracteres de cada *tweet* é que este número viabiliza a atualização de *status* via mensagem de texto em dispositivos móveis, corroborando assim a ideia de o *Twitter* ser uma ferramenta acessível também através de tecnologias móveis.

O *retweet*, ou "RT" é um recurso que facilita aos usuários compartilhar com todos os seus seguidores, um determinado *tweet* de forma simples e rápida (vide Figura 8).



**Figura 8: Retweet**

Fonte: Twitter

Para responder a um *tweet* de determinado usuário, existe a ferramenta *reply*. As respostas são sempre precedidas de um *@nome de usuário*. Nas próximas figuras (Figura 9 e Figura 10), mostraremos onde se encontra a ferramenta de *reply* e como ela é usada.



Figura 9: Ferramenta reply no Twitter  
Fonte Twitter

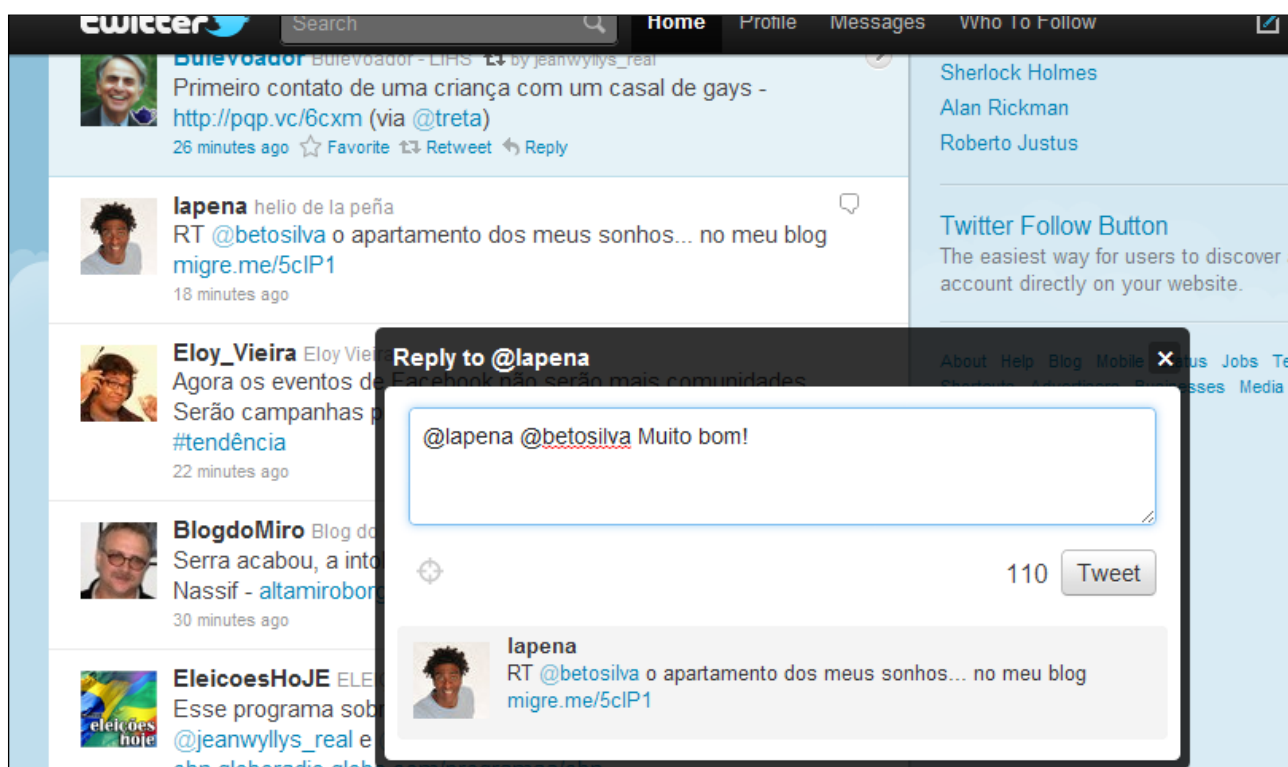


Figura 10: Uso da ferramenta reply no Twitter  
Fonte: Twitter

A *Direct Message* (DM) ou ainda Mensagem Direta (MD) – Figura 11, é uma mensagem privada enviada via *Twitter* por um usuário ao seu seguidor. Diferentemente do *reply*, somente os *followers* (seguidores) de determinado usuário podem enviar uma mensagem direta. Ou seja, um usuário não pode enviar uma mensagem direta a um usuário que não o está seguindo.



Figura 11: Ferramenta Mensagem Direta

Fonte: Twitter

Ao clicar na ferramenta de Mensagem Direta, abre-se uma tela onde será redigido o texto da mensagem, como mostrado na figura 12.



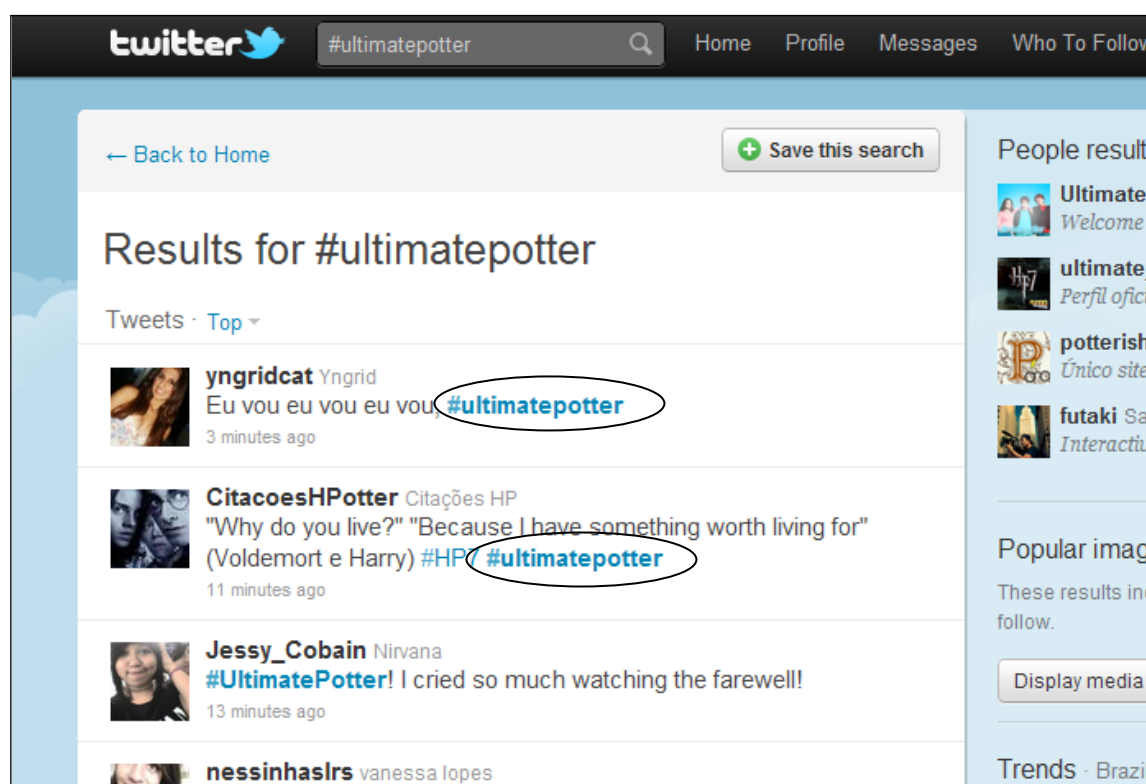
Figura 12: Tela da Mensagem Direta

Fonte: Twitter

Outra característica muito marcante do *Twitter* são as *hashtags*. De acordo com Vitale (2009), as *hashtags* são hoje a maneira mais eficiente para indexar e conectar *tweets* sobre um tema específico. É uma forma fácil de encontrar mensagens relacionadas aos assuntos de seu interesse.

Criar um *hashtag* é simples: basta usar o símbolo # antes de uma palavra. Além de ser uma forma de os usuários indexarem os seus *posts* no *Twitter*, as *hashtags* podem ser uma ótima maneira de aumentar o alcance no *Twitter* e conectar audiência, associando *tweets* a assuntos representativos. (VITALE, 2009).

Na figura abaixo, apresentamos um exemplo de resultado de busca pela *hashtag* “#ultimatepotter”.



**Figura 13: Exemplo de uso das hashtags**

Fonte: Twitter

A audiência de determinado termo no *Twitter* é verificada através do *trending topics* “tópicos da tendência”. Os *trending topics*, como foram apresentados anteriormente, são as palavras mais comentadas em determinado período. Eventos como o Oscar<sup>45</sup> são muito propícios a aparecer nos *trending topics* por se tratarem de

<sup>45</sup> Oscar (*Academy Awards*, no original em inglês) é um prêmio entregue anualmente pela Academia de Artes e Ciências Cinematográficas, fundada em Los Angeles, Califórnia, em 11 de maio de 1927. São entregues anualmente em reconhecimento à excelência de profissionais da indústria cinematográfica, como diretores, atores e roteiristas.

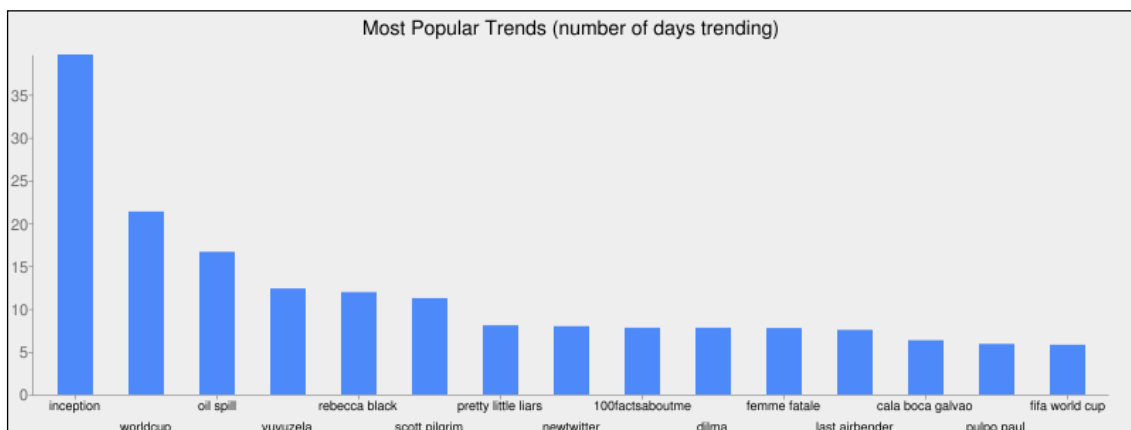




Abaixo, na figura 15, com os termos que ficaram em evidência durante mais tempo e um gráfico mostrando os *trending topics* mais populares em determinado período.

Longest Running Trends	
1.	inception for 27 days, 23 hours and 58 minutes and trended 27 times
2.	#worldcup for 21 days, 8 hours and 26 minutes and trended 105 times
3.	oil spill for 16 days, 15 hours and 47 minutes and trended 195 times
4.	vuvuzela for 12 days, 9 hours and 14 minutes and trended 117 times
5.	pretty little liars for 7 days, 13 hours and 53 minutes and trended 28 times
6.	last airbender for 7 days, 13 hours and 22 minutes and trended 22 times
7.	cala boca galvao for 6 days, 8 hours and 5 minutes and trended 14 times
8.	pulpo paul for 5 days, 21 hours and 45 minutes and trended 9 times
9.	fifa world cup for 5 days, 19 hours and 26 minutes and trended 33 times
10.	mick jagger for 5 days, 3 hours and 15 minutes and trended 20 times
11.	holanda for 4 days, 22 hours and 52 minutes and trended 68 times
12.	cristiano ronaldo for 4 days, 17 hours and 45 minutes and trended 17 times
13.	maradona for 4 days, 8 hours and 48 minutes and trended 38 times
14.	boca galvAo for 4 days, 8 hours and 16 minutes and trended 31 times
15.	lebron james for 4 days, 5 hours and 19 minutes and trended 9 times
16.	comic-con for 4 days, 0 hours and 11 minutes and trended 25 times
17.	felipe melo for 3 days, 21 hours and 23 minutes and trended 27 times
18.	bonamana for 3 days, 19 hours and 37 minutes and trended 19 times
19.	keong racun for 3 days, 9 hours and 47 minutes and trended 13 times
20.	starcraft for 3 days, 7 hours and 3 minutes and trended 20 times
21.	dunga for 3 days, 6 hours and 51 minutes and trended 50 times
22.	cerati for 3 days, 6 hours and 18 minutes and trended 31 times
23.	greyson for 3 days, 5 hours and 33 minutes and trended 18 times
24.	felipe neto for 2 days, 22 hours and 30 minutes and trended 17 times
25.	ios for 2 days, 22 hours and 25 minutes and trended 30 times
26.	#honestyhour for 2 days, 20 hours and 46 minutes and trended 4 times
27.	lee dewyze for 2 days, 13 hours and 26 minutes and trended 6 times
28.	cnectd for 2 days, 13 hours and 7 minutes and trended 7 times
29.	alemania for 2 days, 12 hours and 30 minutes and trended 31 times
30.	clr for 2 days, 12 hours and 1 minutes and trended 145 times
31.	#thingsthatbotherme for 2 days, 11 hours and 30 minutes and trended 1 time
32.	satc2 for 2 days, 11 hours and 18 minutes and trended 7 times
33.	dilma for 2 days, 10 hours and 59 minutes and trended 32 times
34.	kat stacks for 2 days, 9 hours and 27 minutes and trended 18 times
35.	teen choice awards for 2 days, 9 hours and 15 minutes and trended 2 times
36.	#4wordsbeforedeath for 2 days, 7 hours and 59 minutes and trended 5 times
37.	goleiro bruno for 2 days, 6 hours and 46 minutes and trended 10 times
38.	hup holland hup for 2 days, 6 hours and 30 minutes and trended 8 times
39.	iã™you for 2 days, 2 hours and 45 minutes and trended 1 time
40.	#bb11 for 2 days, 2 hours and 42 minutes and trended 48 times
41.	nonton imb for 2 days, 2 hours and 28 minutes and trended 19 times
42.	wimbledon for 2 days, 2 hours and 0 minutes and trended 30 times
43.	ron artest for 2 days, 1 hours and 2 minutes and trended 20 times
44.	avada kedavra for 2 days, 0 hours and 15 minutes and trended 2 times
45.	omaha mall for 1 days, 23 hours and 15 minutes and trended 9 times
46.	paraguay for 1 days, 20 hours and 44 minutes and trended 17 times
47.	jick nonas for 1 days, 20 hours and 1 minutes and trended 1 time
48.	jaden smith for 1 days, 19 hours and 40 minutes and trended 16 times
49.	tca for 1 days, 19 hours and 15 minutes and trended 8 times
50.	uniqlø lucky lineã€¸ for 1 days, 18 hours and 58 minutes and trended 4 times

Figura 15: Tópicos que ficaram mais tempo entre os Trending Topics  
Fonte: <http://youpix.com.br/trending/os-trending-topics-mais-longos-do-twitter/>



**Figura 16: Popular trending**

Fonte : <http://twend.it/top>

Atualmente, as grandes empresas estão reconhecendo o *Twitter* como uma importante ferramenta de divulgação de sua marca. E não é difícil encontrar empresas usando o microblog para divulgar seus produtos. Assim, elas utilizam o recurso das *hashtags* para tentar alcançar os *trending topics* e ganharem mais visibilidade. Um exemplo recente foi o caso da empresa MoonFruit<sup>46</sup> que começou uma campanha que prometia dar 10 *MacBooks Pro* de 13<sup>47</sup> (recém lançados) aleatoriamente para quem colocasse a *tag* *#moonfruit* ao final de suas mensagens no *Twitter*. A promoção foi um verdadeiro sucesso para a empresa. O *#moonfruit* foi parar nos *trending topics* durante alguns dias superando outros eventos de maior importância. O *Twitter* retirou o *#moonfruit* dos *trending topics* acendendo a discussão acerca de qual será o limite para o marketing online utilizando o *Twitter* como ferramenta.

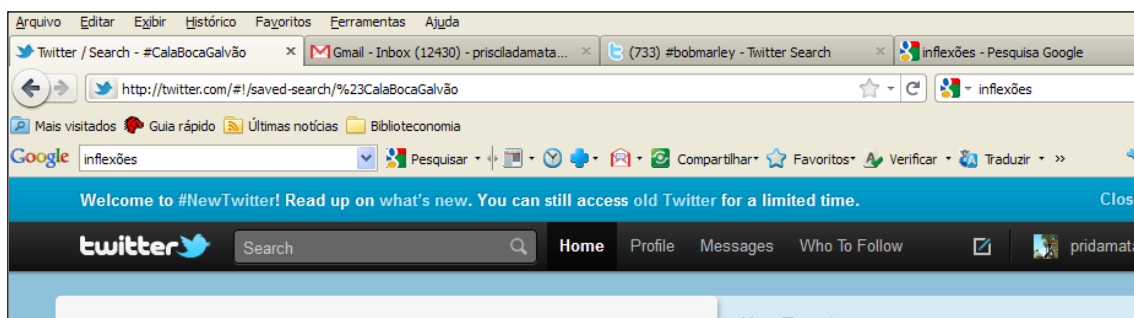
As *hashtags* usadas no *Twitter*, como forma de indexação dos conteúdos dos *posts*, são exemplos da aplicação da *folksonomia* na *web*.

### 5.3 Análise das buscas no Twitter

Na barra de pesquisa “*search*” do *Twitter* (Figura 17), realizamos pesquisas com as *hashtags* *#CALABOCAGALVAO*, *#Gol*, *#Rio*, *#Parangolixo*, *#Idol* e *#ufmg*.

<sup>46</sup> Empresa de hospedagem de sites sediada no Reino Unido.

<sup>47</sup> O *MacBook Pro* de 13” é um computador da Apple mais vendido e mais desejado ao longo do ano de 2010.



**Figura 17: Search Twitter**

Fonte : Twitter

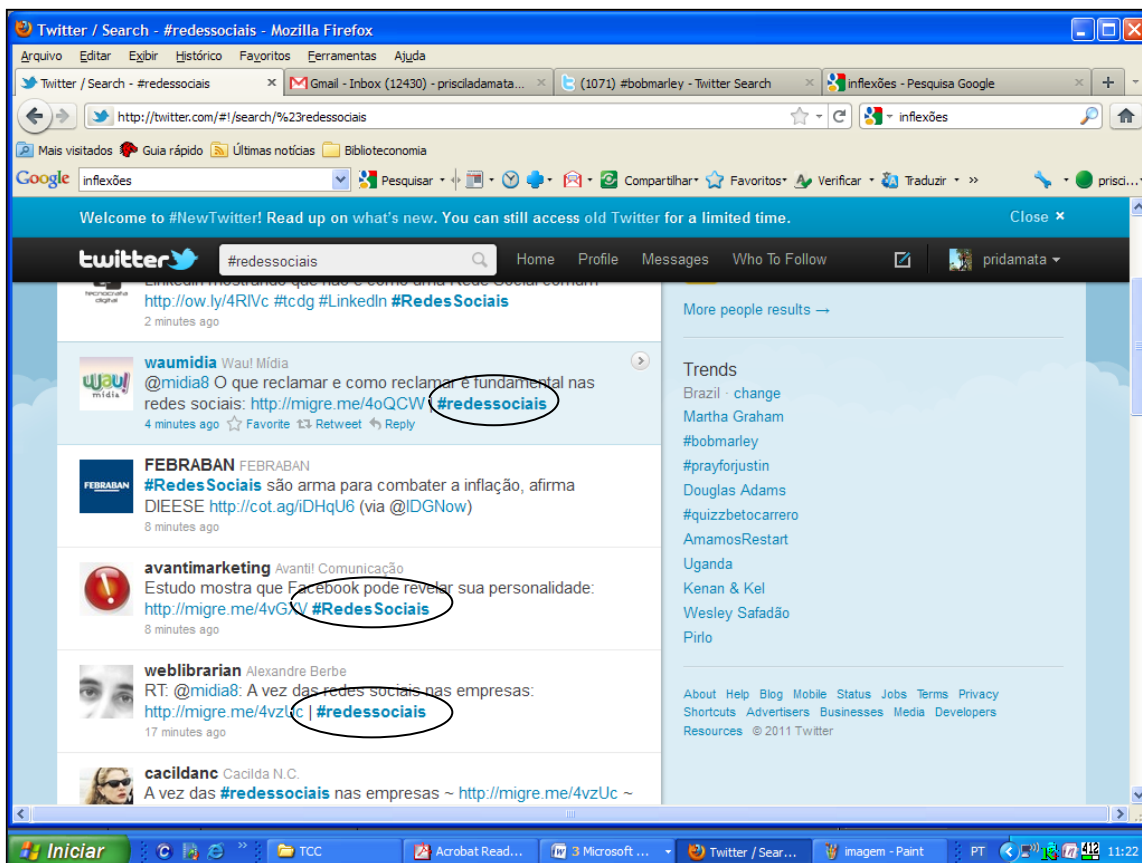
Passaremos então à análise dos resultados obtidos a partir das buscas realizadas utilizando o “*search*” do *Twitter*, levando-se em consideração a forma como os usuários do microblog estão etiquetando os seus *posts*.

É importante lembrar que os resultados das buscas têm um caráter estritamente temporal, uma vez que o *Twitter* caracteriza-se por ser um ambiente totalmente dinâmico. Isso quer dizer que os resultados obtidos dizem respeito à análise das buscas feitas em um determinado período. Desta forma, se realizarmos as mesmas buscas em outro período, podemos encontrar outros resultados.

No *Twitter*, as postagens e as etiquetas dos *posts* são ações realizadas pelos próprios usuários do microblog. Como comentamos anteriormente, as etiquetas utilizadas no *Twitter* são conhecidas como *hashtags* e elas são utilizadas para descrever o conteúdo do que foi postado.

Na figura 18, exemplificamos esse processo de postagem e etiquetagem.





**Figura 18: Etiquetagem no Twitter**  
Fonte Twitter

Como foi dito anteriormente, as *hashtags* facilitam a recuperação da informação no microblog. Atualmente a maioria dos usuários do *Twitter* tem utilizado o processo de etiquetagem, no entanto existem usuários que, por opção ou por desconhecimento do papel das *hashtags*, ainda não fazem uso desse mecanismo.

É importante lembrar também que muitas vezes os usuários usam as *hashtags* apenas com o intuito de conectar audiência dentro do microblog. Este fato corrobora com a perspectiva egocêntrica de etiquetagem proposta por Cañada (2006). Segundo o autor, esta perspectiva está ligada a uma descrição pessoal da informação, sem preocupação alguma com a sua recuperação futura. Neste sentido, no *Twitter*, muitos usuários utilizam *hashtags* que estão sendo muito comentadas naquele momento (*trending topics*) e muitas vezes elas não têm nenhuma relação com o que está sendo postado. Desta forma, quando forem realizadas buscas no blog com aquela *hashtag*, o perfil do usuário aparecerá nos resultados, aumentando assim a sua visibilidade e as suas chances de ganhar audiência e aumentar seus seguidores dentro do microblog.

A *hashtag* #CALABOCAGALVAO, por exemplo, na época do seu auge, foi muito utilizada com esse intuito de conectar audiência. Muitas pessoas, simplesmente postavam a *hashtag* no seu perfil sem mais informação alguma. Tudo bem que podemos dizer que a pessoa estava usando a *hashtag* para desabafar, uma vez que é uma expressão bem sugestiva para isso. No entanto, essa atitude representa um uso indiscriminado dessa ferramenta do *Twitter*, e pode prejudicar a precisão na recuperação da informação no microblog, uma vez que o que será recuperado será apenas uma *hashtag* e não o conteúdo da informação postada que deveria estar atrelada a ela.

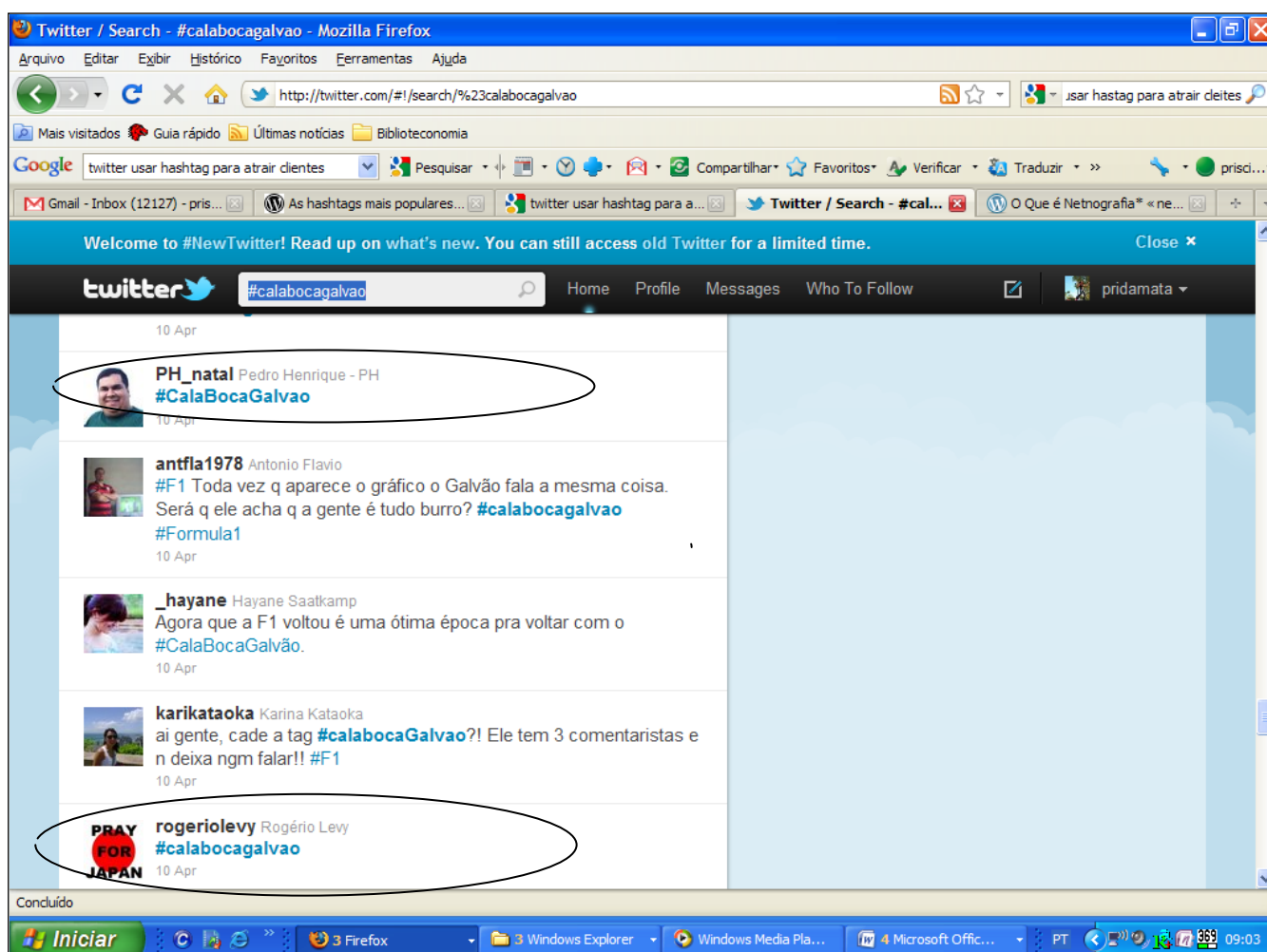
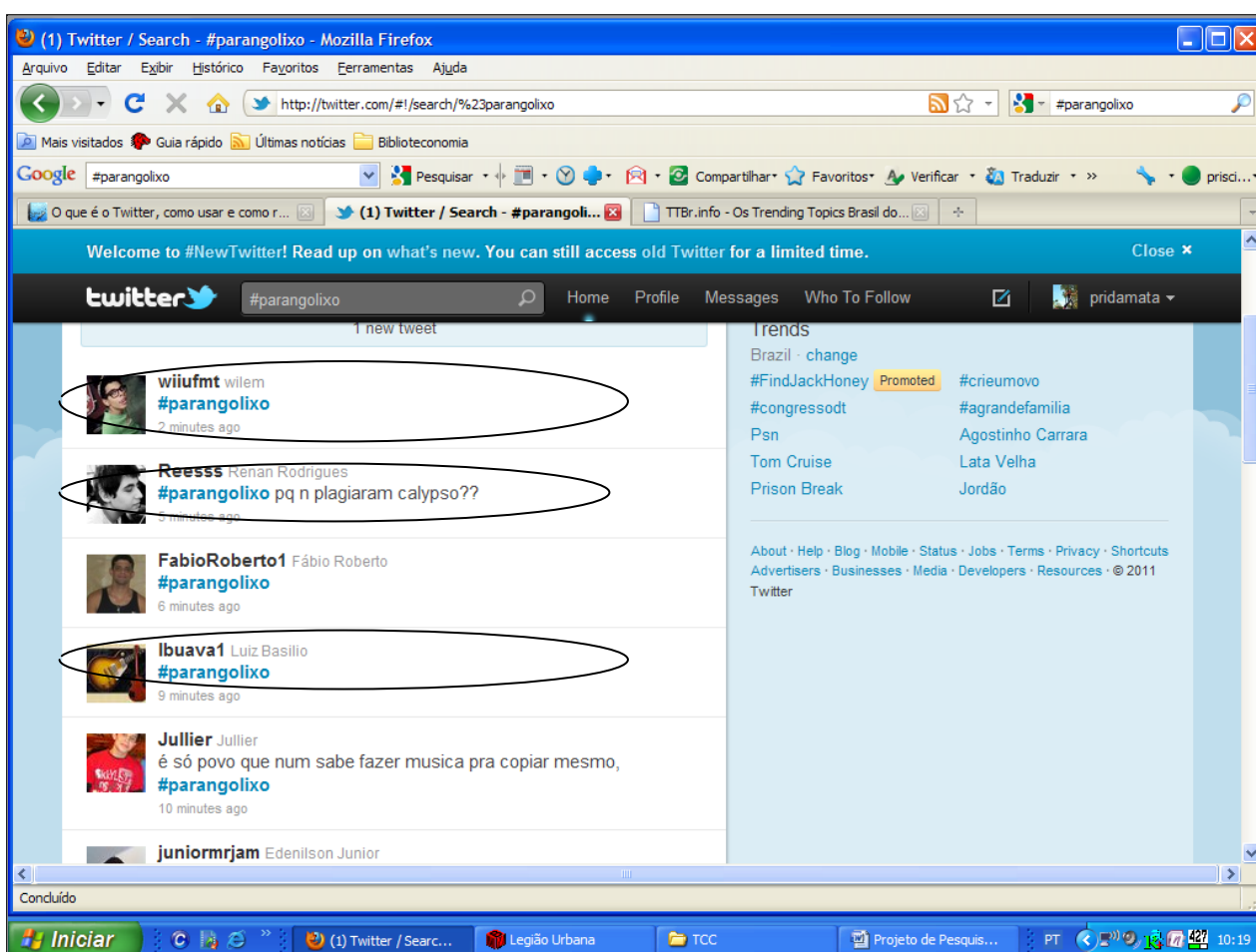


Figura 19: Hashtag CALABOCAGALVAO  
Fonte Twitter

Outra *hashtag* que ficou entre os *trending topics* foi a *#Parangolixo*. Aconteceu quando integrantes da banda Angra acusou o grupo baiano de axé Parangolé de ter plagiado o *riff* de guitarra de sua música “Nova Era”. E em virtude do sucesso da *hashtag*, muitos usuários passaram usá-la em seus *posts*, para protestar contra esse acontecimento, para conectar audiência e ganhar mais visibilidade dentro da rede ou simplesmente para divulgar a *hashtag*.



**Figura 20: Hashtag PARANGOLIXO**  
Fonte Twitter

Outro fato importante que observamos, ao realizarmos as buscas com as *hashtags* no *Twitter*, foi o problema da polissemia. Diversos autores criticam o uso da *folksonomia* em ambientes virtuais justamente em função desses “ruídos” causados pelo uso sem nenhum controle das etiquetas por parte dos usuários, que também

prejudicam a precisão no resultado das buscas. A polissemia é caracterizada pelo “fato de haver uma só forma (significante) com mais de um significado unitário pertencentes a campos semânticos diferentes. [...] cada um desses significados é preciso e determinado.” (BRECHARA, 2004, p. 402).

No *Twitter*, encontramos uma mesma *hashtags* descrevendo assuntos, ideias distintas. Nas duas figuras que se seguem, exemplificamos um caso de polissemia com a *hashtag* #GOL. Como observamos na figura 21, essa etiqueta foi usada para descrever assuntos que tratavam da empresa de aviação Gol Linhas Aéreas, para descrever também o automóvel da marca Gol, da *Volkswagen*, e ainda, para descrever assuntos relacionados a futebol.

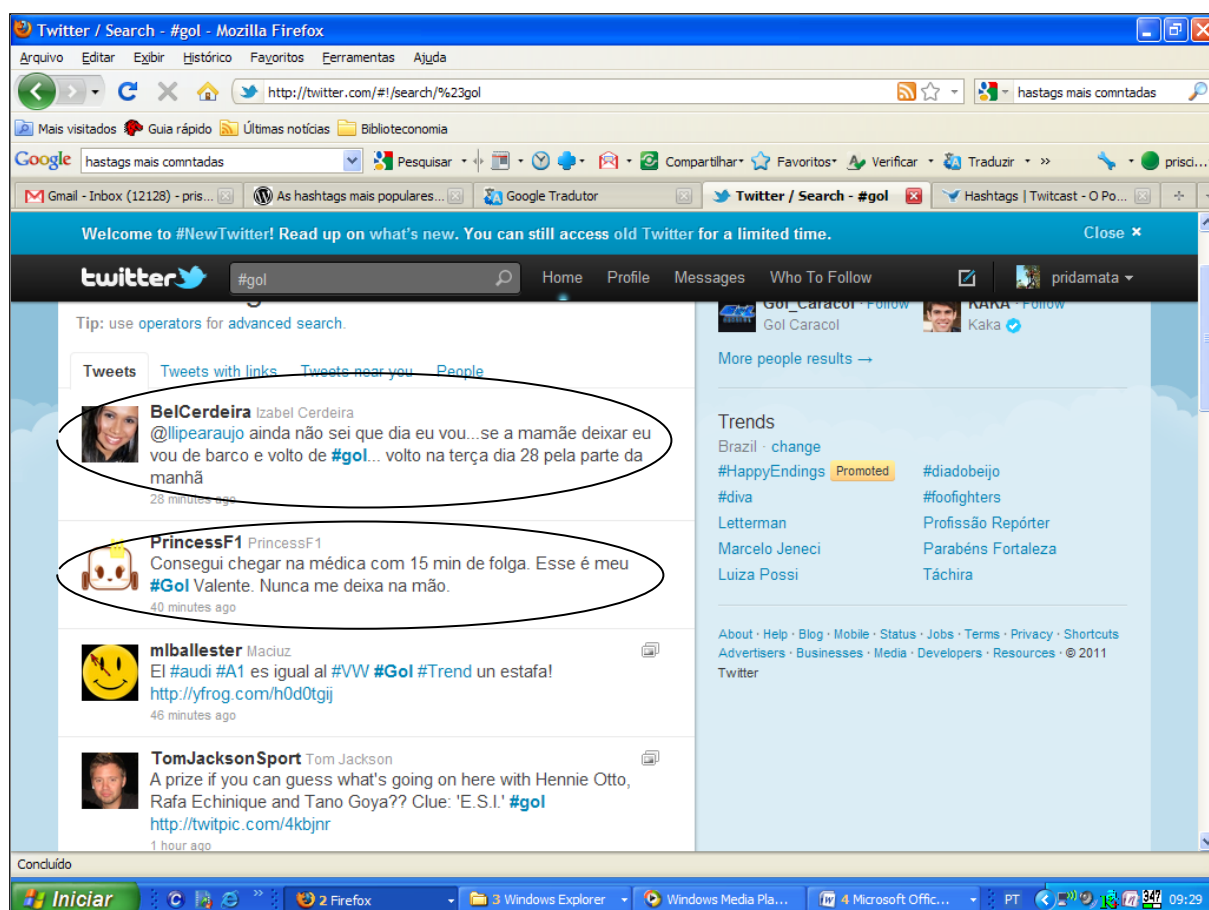
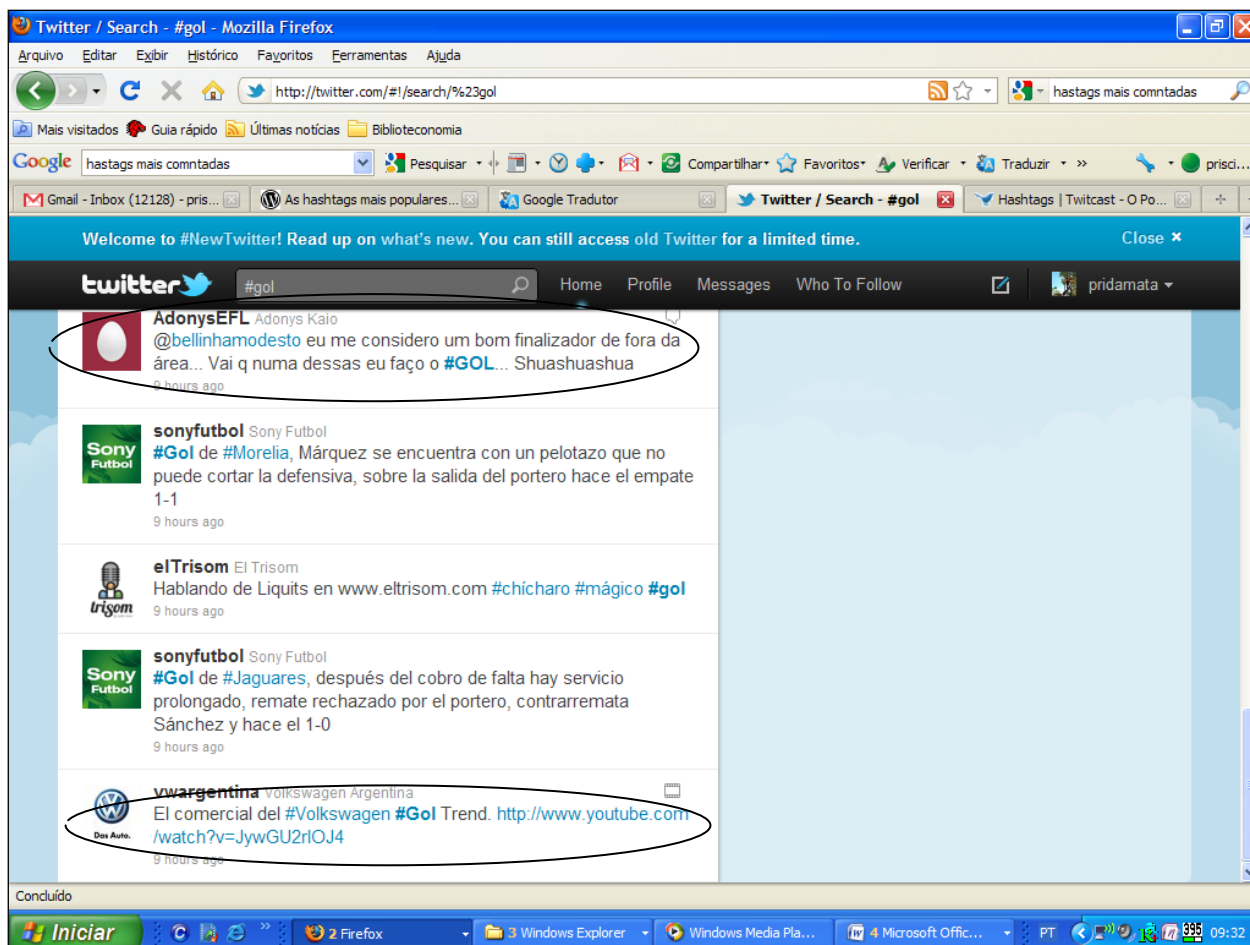


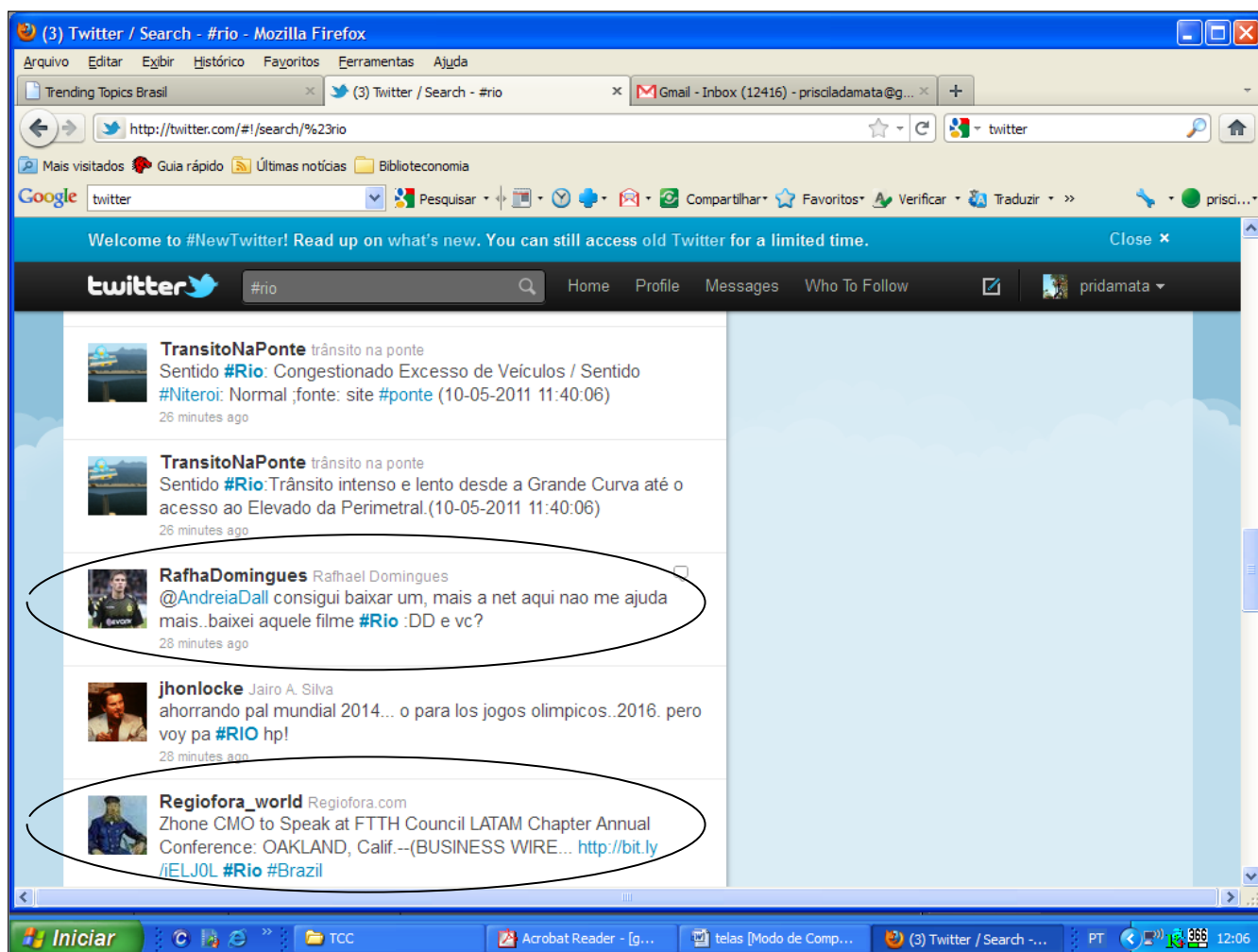
Figura 21: Polissemia 1  
Fonte Twitter



**Figura 22: Polissemia 2**  
Fonte Twitter

Com a *hashtag* #Rio, também observamos alguns “ruídos” em virtude da polissemia. No período em que foram realizadas as buscas, estava sendo lançado no Brasil o filme RIO. Esse foi um sucesso de bilheteria não só no Brasil, mas no mundo inteiro.

O filme foi muito comentado nas redes sociais e principalmente no *Twitter*. No entanto, em virtude do seu nome, a *hashtag* #Rio foi usada para designar não só o filme, mas também, a cidade Rio de Janeiro e algumas vezes até rio, como curso de água natural. Exemplificamos com a figura a seguir este caso de polissemia.



**Figura 23: Polissemia 3**  
Fonte Twitter

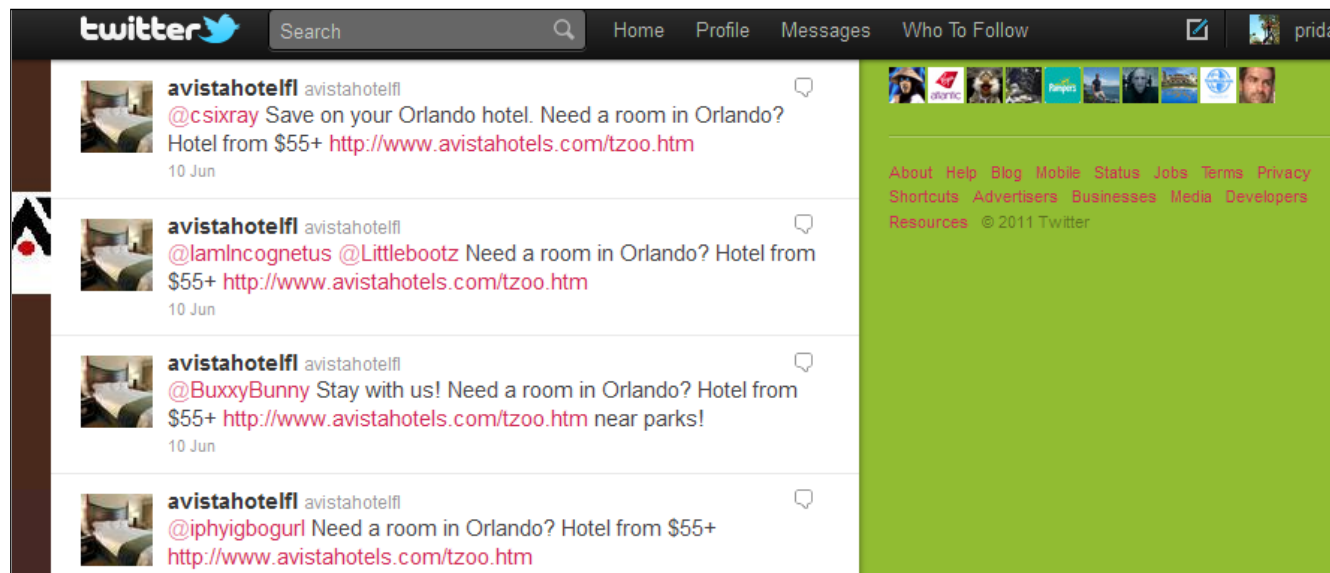
Para entendermos como funciona o mecanismo de proliferação de *spams* no *Twitter*, fizemos uma busca com a *hashtag* #Idol. Ao efetuarmos a busca, deparamo-nos com uma forma muito comum de *spam* que é aquela onde o usuário, seja ele pessoa ou mesmo empresa apodera-se das *hashtags* mais comentadas (*trending topics*) para divulgar seu produto ou marca. A figura abaixo mostra este tipo de *spam*.



**Figura 24: Spam no Twitter**  
Fonte: <http://www.pontomidia.com.br/>



Buscamos também o perfil AvistaHotelfl e percebemos que seus responsáveis utilizam-se de outras formas para propagarem seus *spams* com o intuito de divulgar a sua marca. Na *Timeline*<sup>48</sup> da empresa, observamos vários *replies*, outra forma de disseminação de *spams* no *Twitter*.



**Figura 25: Timeline AvistaHotelfl**

Fonte : Twitter

Os exemplos mostrados sugerem a existência de *spammers* ou robôs<sup>49</sup> no *Twitter*. Os *spammers* utilizam as *hashtags* mais comentadas (*trending topics*) para divulgar anúncios, ou seja, eles utilizam as *tags* mais populares para marcar mensagens publicitárias. É muito comum, ao verificarmos os termos que estão nos *trending topics*, notarmos que entre os *tweets* relacionados aos termos existem vários *spams*.

Os robôs, por sua vez, geralmente são usados por usuários em busca de mais seguidores dentro da rede ou mesmo por algumas empresas. Com o intuito de divulgar sua marca ou produto entre os usuários do *Twitter*, muitas vezes as empresas lançam mão de ferramentas, como os robôs, para alcançar este objetivo. Esses robôs utilizam um script que percorre a rede de seguidores de várias pessoas, deixando o mesmo *spam* como um *reply* para cada membro da rede.

<sup>48</sup> *Timeline* é a linha do tempo nas mídias sociais. É onde se organiza e sequencia as postagens no tempo e conduz o usuário na conversação. O *Timeline* também ajuda aos usuários se orientarem, mostrando quais foram as últimas postagens e mensagens dos usuários das redes sociais.

<sup>49</sup> Robôs ou *Robots*, também conhecidos como "indexadores", "*bots*" ou "*crawlers*" são aplicativos que navegam automaticamente pela internet em busca de documentos, com a finalidade de indexá-los ou monitorar as alterações de conteúdo da página.

O *Twitter* alerta sempre para o uso de certos serviços de robôs para aumentar o número de seguidores dentro da rede. Esses serviços são facilmente encontrados na internet. O problema é que, geralmente, para utilizar estes serviços, o usuário precisa fornecer a sua senha e, uma vez que possuem a senha de acesso de várias contas, os programadores dos robôs conseguem espalhar os *spams* pela rede.

E como solucionar o problema dos *spams* que vem crescendo de forma considerável dentro do *Twitter*? É o que discutiremos no item 5.4.

Para confirmarmos a hipótese de que os termos em menor evidência - que não estão entre os *trending topics* - são menos propícios a apresentarem problemas como a proliferação de *spams*, realizamos buscas com a *hashtag* *#ufmg*, por ela ser considerada uma *hashtag* estática, atemporal e por não estar entre os termos mais comentados no momento da nossa pesquisa. Ao analisarmos os resultados obtidos, percebemos que as informações atreladas ao termo *#ufmg*, em grande parte condiziam com o seu significado descrevendo uma mesma ideia e assunto. Neste sentido, ficou comprovado que as *hahstags* que figuram entre os *trending topics* em determinado momento, tem maiores chances de apresentarem problemas como a proliferação de *spams*.



Figura 26: Hashtag #ufmg

Fonte : Twitter



## 5.4 Filtros (Anti-Spam)

Nos últimos tempos, um dos principais problemas enfrentados pelo *Twitter* tem sido a questão da proliferação de *spams* na rede. Se antes os *spams* eram disseminados apenas em mensagens de emails, hoje com o crescimento da rede social *Twitter*, este problema ganha um novo cenário, com novas formas de ataque, porém, com os mesmos estragos de sempre: mensagens sem nenhum valor, com um único interesse que é vender; uma infinidade de conteúdo não solicitado. Seja através dos *trending topics*, dos *replies* ou mesmo das *direct message* (DM), os *spammers* estão usando o *Twitter* como uma ferramenta, que responde aos usuários e, logo em seguida, faz com que um *tweet spam* seja exibido na sua *timeline*. Eles estão cada vez mais por dentro do funcionamento dessas ferramentas do *Twitter* e, desta forma, estão aprimorando as formas de disseminação de *spams* no blog.

No *Twitter* os *spams* funcionam da seguinte forma: cria-se um perfil falso e esse começa a enviar a mesma mensagem a centenas ou milhares de pessoas diferentes como se fosse um *reply*. Esta tarefa certamente é realizada com a ajuda de um *script*<sup>50</sup>, um *bot*<sup>51</sup> que automatiza esta ação. Desta forma, uma mesma mensagem é enviada a centenas ou milhares de pessoas, em um curto intervalo, geralmente, para pessoas que não são seguidoras daquele perfil, o que é análogo ao *e-mail* não solicitado, o tradicional *spam*.

Para tentar sanar este problema, o *Twitter* conta hoje apenas com um perfil oficial denominado *@Spam* que reúne e divulga informações sobre problemas ou violações de termos de uso dentro da rede. Atualmente, a única forma que o *Twitter*

---

<sup>50</sup> São linguagens de programação interpretadas por um programa, comando a comando, como por exemplo o browser (IE, Firefox, Opera, etc) interpreta uma página de internet.

<sup>51</sup> *Bot*, diminutivo de *robot*, é um utilitário concebido para simular ações humanas, em geral numa taxa muito mais elevada do que seria possível para um editor humano sozinho.

utiliza para combater *spams* e *spammers* é solicitando ao usuário que envie mensagens com denúncias para o perfil *@spam*.

Outra maneira é clicar no link “*report for spam*” em qualquer página suspeita: o recurso aparece sempre abaixo da imagem do perfil, no canto direito da tela. Com isso, a conta suspeita fica impedida de seguir ou responder mensagens para o usuário que a denunciou, e os administradores do *Twitter* recebem um alerta.

Na figura 26, mostramos um exemplo do uso da ferramenta de *report for spam*.



**Figura 26: Report for spam**

Fonte: Twitter

No entanto, este mecanismo que o *Twitter* utiliza para tentar combater os *spams* dentro do site, mostra-se pouco eficiente, uma vez que a ação de denunciar à equipe do *Twitter* um perfil suspeito não impede que este perfil continue enviando *spams* para outros usuários. Isto ocorre porque, quando denunciarmos um perfil suspeito de ser *spam*, ele fica bloqueado e impedido de enviar mensagens apenas para aquele usuário que o denunciou. E enquanto a equipe do *Twitter* analisa a conta suspeita, este perfil poderá continuar proliferando *spams* para outros usuários até que a conta seja suspensa, caso seja realmente considerada, pela equipe do microblog, um perfil *spam*.

Neste sentido, é imprescindível a criação de um filtro *anti-spam* para o *Twitter* que seja realmente eficaz e apresente características específicas inerentes à rede, uma vez que filtros usados em *emails* para combater *spams*, como o

Filtro *Bayesiano*<sup>52</sup> por exemplo, não seriam eficazes para combater este problema no *Twitter*.

Um estudo realizado pelo próprio *Twitter* em março de 2010<sup>53</sup>, apontou que 500.000 mensagens postadas na rede por dia são *tweets* com *links* para sites infectados por vírus, *spams*.

Sabe-se que atualmente existem vários serviços de filtragem de informações para o *Twitter*, como o *twittfilter* e o *philtro beta*. Esses serviços atenuariam, em parte, os problemas causados pelos *spams* no blog. No entanto, além de serem ferramentas criadas por outras empresas e não pelo próprio *microblog*, não são especificamente ferramentas *anti-spams*. O *philtro beta*, que ainda está em sua primeira versão beta, é um serviço que filtra *tweets*. O *twittfilter*, assim como o *philtro beta*, é uma ferramenta para quem quer evitar informações pouco relevantes dos seus contatos. Ela permite que o usuário selecione apenas os assuntos que lhe interessam nos *trending topics*. No entanto, as duas ferramentas estão ainda em fase de testes, o que faz com que elas apresentem um desempenho pouco satisfatório. Solucionar o problema dos *spams* acende outra questão que diz respeito à censura por parte do *Twitter*.

Recentemente o *Twitter* foi acusado de censurar termos relacionados ao *WikiLeaks*<sup>54</sup> - organização que vaza informações confidenciais de governos e empresas. No entanto, o *Twitter* sempre nega que esteja censurando qualquer tipo de termo. Em um comunicado à imprensa enviado ao site americano *Mashable*<sup>55</sup>, o *Twitter* detalha o processo que define o que é *trending topic*. “Às vezes os tópicos que são mais populares não aparecem na lista de ‘trends’, porque a velocidade de conversas (volume de *tweets* em certo momento) não é superior à das horas ou dias anteriores. Alguns tópicos não são disseminados o suficiente para fazer parte da lista dos *trending topics*.”<sup>56</sup>

---

<sup>52</sup> Um filtro bayesiano e vários outros sistemas de classificação automática são baseados no teorema criado por Thomas Bayes. Trata-se de um cálculo utilizado para julgar a probabilidade de um evento ocorrer considerando as circunstâncias e o número de vezes que o evento já ocorreu. É construído a partir de um processo de aprendizagem, onde um usuário o treina apresentando mensagens consideradas spam e mensagens legítimas.

<sup>53</sup> Disponível em: <http://veja.abril.com.br/blog/vida-em-rede/twitter/twitter-intensifica-guerra-anti-spam/>

<sup>54</sup> <http://wikileaks.org/>

<sup>55</sup> <http://mashable.com/>

<sup>56</sup> Fonte: <http://tecnologia.uol.com.br/ultimas-noticias/redacao/2010/12/06/twitter-nega-acusacoes-de-censura-relacionada-aos-termos-do-wikileaks.jhtm>

## 6. RESULTADOS

Tendo em vista as suas infinitas possibilidades, o *Twitter* vem-se tornando objeto de estudo de pesquisadores de diversas áreas, sejam elas Comunicação Social, Computação ou Ciência da Informação. Em ambientes colaborativos, como o *Twitter*, torna-se impossível o uso da taxonomia para organizar as informações tendo em vista a sua rigidez e controle. Neste sentido, a *folksonomia* surge como uma alternativa capaz de organizar o caos gerado em um ambiente caracterizado por uma produção colaborativa com grande volume de publicações em um curto período.

Porém, apesar das vantagens que a *folksonomia* apresenta, também traz diversos problemas que podem diminuir a precisão na recuperação das informações que são indexadas em linguagem natural.

As buscas realizadas no *Twitter* com as *hashtags* #CALABOCAGALVAO, #Gol, #Rio, #Parangolixo e #Idol mostraram alguns desses problemas. São várias as inconsistências e ambiguidades observadas. A polissemia talvez seja um dos principais problemas ocasionados pelo uso da *folksonomia*. São inúmeras palavras com sentidos múltiplos usadas para indexar *tweets*. Como foi visto, a *hashtag* #GOL, por exemplo, pode designar uma empresa de aviação civil, um modelo de uma marca de automóvel ou mesmo lance de futebol.

No entanto, a intenção dos usuários do *Twitter* quando indexam seus *posts* também é um fator que deve ser levado em consideração. Como foi observado na pesquisa, vários usuários indexam seus *posts* com as *hashtags* mais comentadas (*trending topics*) com o intuito de ganhar maior visibilidade dentro da rede, ou no caso das empresas, visando a divulgação de sua marca ou produto. Atitudes como essa fazem com que certas *hashtags* não tenham nenhuma relação com o conteúdo postado, prejudicando assim a precisão nas buscas, uma vez que o conteúdo informacional do *post* não terá relação com a palavra usada para indexá-la.

Ao analisarmos as buscas com a *hashtag* #ufmg verificamos que, por ela não figurar entre os *trending topics* no momento da nossa pesquisa, as informações atreladas a essa *hashtag* são mais consistentes e relevantes. Observou-se também que ela apresenta menos problemas no que se refere à proliferação de *spams* do que em outros termos mais comentados. Como a *hashtag* #ufmg não se encontrava entre os *trending topics* no momento da nossa pesquisa, os usuários não estavam utilizando esse termo para ganharem maior visibilidade dentro do *Twitter*.

A análise das estratégias de indexação adotadas pelos usuários permitiu compreender se eles estão indexando os seus *posts* de acordo com o conteúdo dos

mesmos ou se estão apenas associando *hashtags* mais comentadas, com o intuito de atingir um alcance e uma audiência maior na rede. De acordo com a pesquisa, verificou-se que muitos usuários estão utilizando as *hashtags* mais comentadas, que em determinado momento se encontravam nos *trending topics*, para ganhar maior visibilidade dentro do microblog. Este fato foi comprovado ao analisarmos os resultados das buscas com as *hashtags* #CALABOCAGALVAO e #Idol. Observamos que muitos usuários estavam indexando seus *posts* com as *hashtags* mais comentadas apenas com o intuito de conectarem audiência ou divulgarem uma marca ou produto uma vez que, desta maneira, seu perfil poderia aparecer nos resultados de futuras buscas pela aquela referida *hashtag* no *Twitter*.

Foi possível verificar também a questão da proliferação de *spams* no *site*, uma vez que esse tem sido um grande problema enfrentado pelo *Twitter* recentemente. Muitas empresas estão usando este mecanismo para manter a sua marca e/ou produtos entre os *trending topics* com o intuito de se promoverem na rede. É muito comum, quando buscamos algum termo que esteja em evidência, ou seja nos *trending topics*, ao verificarmos os *tweets* atrelados a ele, é possível perceber várias mensagens fazendo propaganda de determinada marca ou produto. Outra forma muito comum de *spam no Twitter*, como foi mostrado na pesquisa, é o envio de *tweets* de maneira pessoal através de *reply*.

Além disso, a análise da forma de como os usuários estão indexando seus *posts* permitiu ainda compreender como alguns problemas gerados pelo uso da *folksonomia*, como a polissemia, tem afetado a precisão na recuperação da informação no *Twitter*.

Todos esses problemas são fatores que influenciam na atinência em processos de recuperação da informação no *Twitter* e comprometem a finalidade atual da ferramenta que é de ser um canal informativo. Encontrar mecanismos que minimizem os efeitos desses fatores é um grande desafio para os profissionais que trabalham por trás do funcionamento dessas redes.

A presente pesquisa não se esgota em si mesma. Em virtude das várias temáticas aqui abordadas, como a semântica e a *folksonomia*, espera-se que este estudo venha a contribuir e incentivar investigações futuras em diversas áreas. Nesse sentido, a compreensão das estratégias de indexação teve como propósito servir como auxílio para estudos futuros nas áreas de:

- recuperação da informação no *Twitter*, uma vez que, entendendo como é realizado o processo de indexação pelos usuários no momento das suas postagens e os problemas que estão sendo gerados por este processo,

verificaremos também como anda a credibilidade das informações recuperadas numa busca realizada no *Twitter Search*;

- organização da informação no *Twitter*. A pesquisa mostrou os diversos problemas advindos do uso da *folksonomia* no microblog e apontou a necessidade de um estudo mais aprofundado nesta área visando a criação de meios que minimizem os efeitos da utilização da linguagem natural no momento da etiquetagem das informações dentro do microblog;
- comunicação e *marketing* digitais, mais especificamente estudos de mensuração e monitoramento em mídias sociais. A descoberta dos “ruídos” existentes nas informações que circulam dentro do microblog, principalmente em função do uso da *folksonomia*, pode contribuir para resultados mais precisos no que se refere ao monitoramento no *Twitter* por parte de empresas que estão cada vez mais utilizando estes estudos para avaliar o desempenho de determinada marca ou produto dentro do microblog;
- estudos mais aprofundados na área de linguística, mais precisamente na que tange à semântica, em virtude da existência do fenômeno da polissemia, que interfere diretamente na precisão dos resultados na recuperação das informações dentro do microblog;
- estudos voltados para o desenvolvimento de ferramentas *anti-spams* que levam em consideração as especificidades do *Twitter*. Em virtude da crescente proliferação de *spams* dentro do microblog, a criação de ferramentas anti – *spams* eficazes no *Twitter* torna-se fundamental para combater o problema da disseminação de *spams*, uma vez que é um dos fatores que afetam diretamente a credibilidade das informações que circulam no microblog;
- *web* semântica, pois acredita-se que a compreensão da forma como os usuários do *Twitter* estão indexando seus *posts*, como foi mostrado na presente pesquisa, poderá contribuir para a construção de ferramentas semanticamente capazes de diminuir os ruídos ocasionados pelo uso da *folksonomia* no *Twitter* e maximizar a precisão nos resultados das buscas realizadas no site, aumentando assim a credibilidade e a confiabilidade das suas informações que ali circulam.

## REFERÊNCIAS

AMARAL, Bruna Rodrigues do. **Polissemia**: efeitos contextuais no acesso lexical. 2011. 79 f. Dissertação (Mestrado em Estudos Linguísticos) – Faculdade de Letras, Universidade Federal de Minas Gerais, Belo Horizonte, 2011.

AQUINO, Maria Clara. A folksonomia como hipertexto potencializador de memória coletiva: um estudo dos links e das tags no de.licio.ous e no Flickr. **Liinc em Revista**, v. 4, n. 2, p. 303-320, set. 2008.

AQUINO, Maria Clara. Hipertexto 2.0, folksonomia e memória coletiva: um estudo das tags na organização da web. **E-Compôs**, v. 18, n. 18, ago. 2007.

BAEZA-YATES, Ricardo; RIBEIRO NETO, Berthier. **Modeling**. In: \_\_\_\_\_. Modern information retrieval. New York: Addison Wesley, 1999. Cap. 2. Disponível em: <http://www2.dcc.ufmg.br/livros/irbook/chapters/chap2.html> . Acesso em: 12 abr. 2011.

BARRETO, Aldo. Mudança estrutural no fluxo do conhecimento: a comunicação eletrônica. **Ciência da Informação**, v. 27, n. 2, p. 122-127, maio/ago. 1998.

BARTHES, Roland. **O prazer do texto**. São Paulo: Perspectiva, 1973.

BLOG do Hirosh. **Como funciona o Twitter**. Disponível em <<http://www.blogdohiroshi.com/2008/01/como-funciona-twitter/>>. Acesso em 15 mar. 2011.

BECHARA, Evanildo. **Moderna gramática portuguesa**. 37ª ed. rev. e ampl. Rio de Janeiro: Lucerna, 2004.

CAMPBELL, D. G. A phenomenological framework for the relationship between the semantic web and user-centered tagging systems. In: **SIG/CR Classification Research Workshop**, 17, 2006, Austin (USA). *Papers...* Austin, 2006. Disponível em: <<http://www.slais.ubc.ca/users/sigcr/sigcr-06campbell.pdf>>. Acesso em: 13 mar. 2011.

CAÑADA, Javier. Tipologías y estilos en el etiquetado social. **Terremoto.net**, 2006. Disponível em: <<http://www.terremoto.net/tipologias-y-estilos-en-el-etiquetadosocial/>>. Acesso em: 15 dez. 2011.

CASTELLS, Manuel. **A Galáxia da internet**: reflexões sobre a internet, os negócios e a sociedade. Rio de Janeiro: Jorge Zahar, 2003.

CATARINO, Maria Elisabete; BAPTISTA, Ana Alice. Folksonomia: um novo conceito para a organização dos recursos digitais na Web. **DataGramZero**, v.8, n.3, jun. 2007. Disponível em <[http://www.dgz.org.br/jun07/F\\_I\\_art.htm](http://www.dgz.org.br/jun07/F_I_art.htm)> Acesso em: 05 abr. 2011.

\_\_\_\_\_. Folksonomias: características das etiquetas na descrição de recursos da Web. **Inf. Inf.**, v.14, p. 46-67, 2009.

COWAN, James. **O sonho do cartógrafo**: meditações de Fra Mauro na corte de Veneza do século XVI. Tradução de Maria de Lourdes Reis Menegale. Rio de Janeiro: Rocco, 1999. Título original: A mapmaker's dream.

FEITOSA, Ailton. **Organização da informação na web**: das tags à web semântica. Brasília : Thesaurus, 2006. 132 p. (Estudos avançados em Ciência da Informação, v. 2).

FRANÇA, Júnia Lessa; VASCONCELOS, Ana Cristina de. **Manual para normalização de publicações técnico-científicas**. 8. ed. Belo Horizonte. Editora UFMG, 2007.

GOLDER, Scott A.; HUBERMAN, Bernardo A. The structure of collaborative tagging systems. **Journal of Information Science**, 2006. Disponível em: <http://arxiv.org/ftp/cs/papers/0508/0508082.pdf> Acesso em: 10 abr. 2011.

HAMMOND, Tony et al. Social Bookmarking Tools (I): a general review. D-Lib Magazine, v.11, n.4, apr. 2005. Disponível em: <http://www.dlib.org/dlib/april05/hammond/04hammond.html>. Acesso em: 25 nov. 2011.

INTERNET. **World Stats**. Disponível em <http://www.internetworldstats.com/> Acesso em 15 fev. 2011.

KOZINETS, Robert V. The field behind the screen: using netnography for marketing research in online communities. **Journal of Marketing Research**, Chicago, v.39, n.1, fev. 2002. Disponível em <http://www.atyponlink.com/AMA/doi/abs/10.1509/jmkr.39.1.61.18935?journalCode=jmkr>. Acesso em 05 abr. 2011.

LAVILLE, C; DIONNE, J. A pesquisa científica hoje. In: **A construção do saber**: manual de metodologia da pesquisa em ciências humanas. Porto Alegre: Artes Médicas/Belo Horizonte: Editora UFMG, 1999.

LEVY, Pierre. **Cibercultura**. São Paulo: Editora 34, 1999.

\_\_\_\_\_. **A conexão planetária**: o mercado, o ciberespaço, a consciência. São Paulo:[s.n.], 2001.

\_\_\_\_\_. **A inteligência coletiva**: por uma antropologia do ciberespaço. São Paulo: Edições Loyola, 2007.

LYONS, John. **Linguagem e linguística**: uma introdução. Rio de Janeiro: Guanabara koogan, 1987.

MARTELETO, Regina Maria. Análise de redes sociais: aplicação nos estudos de transferência da informação. **Ciência da Informação**, Brasília, v. 30, n. 1, p. 71-81, jan./abr. 2001.

MATHES, Adam. Folksonomies - cooperative Classification and Communication through shared metadata. Computer Mediated Communication – LIS590CMC, Urbana : University of Illinois, 2004. Disponível em: <http://www.adammathes.com/academic/computer-mediated-communication/folksonomies.html>. Acesso em : 25 nov. 2011.

MIRANDA, Antonio. Sociedade da informação: globalização, identidade cultural e conteúdos. **Ciência da Informação**, v.29, n. 2, p. 78-88, maio/ago. 2000.

MONTARDO, Sandra Portela; PASSERINO, Maria Liliana. Estudo dos blogs a partir da netnografia: possibilidades e limitações. Porto Alegre, **RENOTE**: v. 4, n. 2, 2006.



Disponível em: <<http://www.cinted.ufrgs.br/renote/dez2006/index.html>>. Acesso em: 05 fev. 2011.

MOURA, M. A. Folksonomias, redes sociais e as formação para o *tagging literacy*: desafios para a organização das informação em ambientes colaborativos virtuais. *Informação & Informação (UEL.. Online)*.v. 14, p. 25-45, 2009.

\_\_\_\_\_. Informação, ferramentas ontológicas e redes sociais AD HOC: a interoperabilidade na construção de tesouros e ontologias. *Inf. & Soc.*, João Pessoa, v. 19, n. 1, p.59-73, jan./abr. 2009.

\_\_\_\_\_. Informação e conhecimento em redes virtuais de cooperação científica: necessidades, ferramentas e usos. *DataGramZero*, v.10, n.2, abr. 2009.

NASCIMENTO, Geysa Flávia Câmara de Lima. **Folkonomia como estratégia de indexação dos bibliotecários no DEL.ICIO.US**. 2008. Dissertação (Mestrado em Ciência da Informação) – Centro de Ciências Sociais Aplicadas, Universidade Federal da Paraíba, João Pessoa, 2008.

O'REILLY, Tim. Web 2.0: padrões de design e modelos de negócios para a nova geração de software. Tradução: Miriam Medeiros. 2006. Título original: What Is Web 2.0: design patterns and business models for the next generation of software. Disponível em < <http://www.mendeley.com/research/o-que-web-20-padres-design-e-modelos-negcios-para-nova-gerao-software-1/>>. Acesso em: 05 maio 2011.

PETERSON, Elaine. Beneath the Metadata: some philosophical problems with Folksonomy. *D-Lib Magazine*, v.12, n.11, nov. 2006. Disponível em: <<http://www.dlib.org/dlib/november06/peterson/11peterson.html>>. Acesso em: 25 nov. 2011.

PIEIDADE, Maria Antonieta Requião. **Introdução à teoria da classificação**. Rio de Janeiro: Interciência, 1983.

RECUERO, Raquel da Cunha. Redes sociais na internet: considerações iniciais. **E-Compôs**, [s.l.], v.2, [s.n.], 2005. Disponível em <[http://www6.ufrgs.br/limc/PDFs/redes\\_sociais.pdf](http://www6.ufrgs.br/limc/PDFs/redes_sociais.pdf)>. Acesso em 05 maio 2011.

\_\_\_\_\_. Teoria das redes sociais na internet: considerações sobre Orkut, os Weblogs e os Fotologs. In.:IV Encontro dos Núcleos de Pesquisa da XXVII INTERCOM, 27, Porto Alegre. **Anais...** UFRS, 2004. Disponível em: <<http://galaxy.intercom.org.br:8180/dspace/bitstream/1904/17792/1/R0625-1.pdf>> Acesso em 17 jul. 2011.

RUFINO, Airtiane. Classificação: da filosofia à folksonomia. In: Semana de Humanidades UFC/UECE, 6, Fortaleza. **Anais...** Fortaleza: UFC, 2009.

RUFINO, Airtiane; SILVA, Roosewelt Lins. Folksonomia: a classificação colaborativa facilitando a organização do conhecimento. In.: Encontro Nacional de Estudantes de Biblioteconomia, Documentação, Ciência da Informação e Gestão da Informação, 31, Maceió. **Anais Eletrônicos...** Maceió: UFAL, 2008.

RUFINO, Airtiane; TABOSA, Hamilton Rodrigues; SOUZA, Andrezza Abraham Ohana de. Twitter: a transformação na comunicação e no acesso às informações. In: XI

Congresso Regional de Ciências da Comunicação– INTERCOM NORDESTE, 2009, Teresina. **Anais....** São Paulo: Intercom, 2009.

SARACEVIC, Tefko. Ciência da Informação: origem, evolução e relações. **Perspectivas em Ciência da Informação**, v.1, n.1, p.41-62, jan./ jun.1996.

SILVA, Augusto S. da. **O mundo dos sentidos em português**: polissemia, semântica e cognição. Almedina: Coimbra, 2006.

SILVA, Fabiano Couto Corrêa da; 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**, v.12, n.2, p. 191-215, jul./dez., 2007.

SIMON, Maria Lúcia Mexias. **Semântica e pragmática**. Disponível em: <<http://www.filologia.org.br/viisenefil/10.htm>>. Acesso em 15 maio 2011.

SOUZA, Renato Rocha; ALVARENGA, Lídia. A web semântica e suas contribuições para a ciência da informação. **Ciência da Informação**, v.33, n.1,p. 132-141, Jan./abr. 2004.

UOL. **Twitter nega acusações de censura relacionada aos termos do WikiLeaks**. Disponível em: <<http://tecnologia.uol.com.br/ultimas-noticias/redacao/2010/12/06/twitter-nega-acusacoes-de-censurarelacionada-aos-terminos-do-wikileaks.jhtm> > Acesso em 20 jun. 2011.

VERGARA, Sylvia Constant. **Métodos de pesquisa em Administração**. São Paulo: Atlas, 2005.

VILLANI, Leonardo. **Utilizando inteligência artificial para combater spams**. Disponível em <http://www.villani.eti.br/404.html>. Acesso em: 20 jun. 2011

VITALE, Candice. **Como usar as hashtags# no Twitter?** Disponível em <<http://www.heat.com.br/blog/redes-sociais/como-usar-as-hashtags-no-twitter/> > Acesso em 15 abr. 2011.