

**AMANDA MOURA DE SOUSA**

**Organização em sistema caótico:  
uso das *tags* para classificação da informação pelos  
usuários da rede social *Twitter***

**Dissertação de mestrado  
Março de 2012**



UNIVERSIDADE FEDERAL DO RIO DE JANEIRO  
FACULDADE DE ADMINISTRAÇÃO E CIÊNCIAS CONTÁBEIS  
INSTITUTO BRASILEIRO DE INFORMAÇÃO EM CIÊNCIA E TECNOLOGIA  
PROGRAMA DE PÓS-GRADUAÇÃO EM CIÊNCIA DA INFORMAÇÃO – MESTRADO

**AMANDA MOURA DE SOUSA**

**Organização em sistema caótico: uso das *tags* para classificação da informação pelos usuários da rede social *Twitter***

**RIO DE JANEIRO  
2012**

**AMANDA MOURA DE SOUSA**

**Organização em sistema caótico: uso das *tags* para classificação da informação pelos usuários da rede social *Twitter***

Dissertação de Mestrado apresentada ao Programa de Pós-Graduação em Ciência da Informação, convênio entre o Instituto Brasileiro de Informação em Ciência e Tecnologia e a Faculdade de Administração e Ciências Contábeis, Universidade Federal do Rio de Janeiro como requisito parcial à obtenção do título de Mestre em Ciência da Informação.

**Orientadora: Dra. ROSALI FERNANDEZ DE SOUZA**

**Rio de Janeiro  
2012**

S725

Sousa, Amanda Moura de

Organização em sistema caótico: uso das tags para classificação da informação pelos usuários da rede social *Twitter* / Amanda Moura de Sousa. — Rio de Janeiro, 2012.

103 f. : il. color. ; 21 cm.

Dissertação (Mestrado em Ciência da Informação) – Programa de Pós-graduação em Ciência da Informação, Instituto Brasileiro de Informação em Ciência e Tecnologia, Universidade Federal do Rio de Janeiro, Faculdade de Administração e Ciências Contábeis, Rio de Janeiro, 2012.

Orientadora: Rosali Fernandez de Souza.

1. Classificação. 2. *Twitter*. 3. Organização do Conhecimento. 4. Web Semântica. 5. Teoria do Caos. I. Souza, Rosali Fernandez de (orient.). II. Universidade Federal do Rio de Janeiro. Faculdade de Administração e Ciências Contábeis. Instituto Brasileiro de Informação em Ciência e Tecnologia. III. Título.

CDD 025

AMANDA MOURA DE SOUSA

**Organização em sistema caótico: uso das tags para classificação da informação pelos usuários da rede social *Twitter***

Dissertação de Mestrado apresentada ao Programa de Pós-Graduação em Ciência da Informação, convênio entre o Instituto Brasileiro de Informação em Ciência e Tecnologia e a Faculdade de Administração e Ciências Contábeis, Universidade Federal do Rio de Janeiro como requisito parcial à obtenção do título de Mestre em Ciência da Informação.

Aprovada em: \_\_\_\_\_.

BANCA EXAMINADORA

---

Profa. Dra. Rosali Fernandez de Souza  
Dra. em Ciência da Informação - Polytechnic of North London  
Programa de Pós-Graduação em Ciência da Informação Instituto Brasileiro de Informação em Ciência e Tecnologia

---

Prof. Dra. Maria Cecília Magalhães Mollica  
Dra. em Linguística - UFRJ  
Programa de Pós-Graduação em Ciência da Informação Instituto Brasileiro de Informação em Ciência e Tecnologia

---

Profa. Dra. Rose Marie Santini de Oliveira  
Dra. em Ciência da Informação – IBICT/UFRJ  
Escola de Comunicação – Universidade Federal do Rio de Janeiro

Suplentes:

---

Dr. Jorge Calmon Almeida Biolchini  
Dr. em Ciência da Informação – IBICT/UFRJ  
Instituto Brasileiro de Informação em Ciência e Tecnologia

---

Dra. Vera Lucia Doyle Louzada de Mattos Dodebei  
Dra. em Comunicação – UFRJ  
Programa de Pós-Graduação em Memória Social da UNIRIO

Aos meus avós José Leão de Moura e Maria Pureza Cardoso de Moura (*in memoriam*) pelo apoio e carinho que sempre me deram enquanto estiveram ao meu lado.

## AGRADECIMENTOS

À professora Rosali Fernandez pela orientação paciente e delicada.

À banca que avaliou esse trabalho pelas contribuições valiosas.

Aos colegas mestrandos e doutorandos, pelas grandes discussões em sala de aula.

Um agradecimento especial aos colegas Aluf Elias, Camila Belo, Leandro Fonseca e Roberto Lopes pelas excelentes trocas de experiências e conversas animadas.

À equipe do Projeto UTILH<sub>2</sub> e da Divisão de Informação Tecnológica do INT pelo incentivo e apoio.

Aos meus colegas da Biblioteca Plínio Sussekind Rocha do Instituto de Física/UFRJ pela forma como fui acolhida.

Aos meus queridos amigos, por compreenderem minhas ausências em seus aniversários e encontros.

Às mulheres maravilhosas do Estúdio Violanti Dancer pelo encontro com a arte e pelos sábados animados.

À minha família pelo apoio e carinho.

“A ordem da Física Clássica, não é mais o contexto do Universo”.

Edgard Morin.

## RESUMO

SOUSA, Amanda Moura de. **Organização em sistema caótico**: uso das *tags* para classificação da informação pelos usuários da rede social *Twitter*. Rio de Janeiro, 2012. Orientadora: Rosali Fernandez de Souza. 105 f. Dissertação (Mestrado em Ciência da Informação) Faculdade de Administração e Ciências Contábeis, Universidade Federal do Rio de Janeiro, Instituto Brasileiro de Informação em Ciência e Tecnologia, Rio de Janeiro, 2012.

Aborda temas relacionados à classificação e organização do conhecimento, folksonomia, recuperação da informação e às teorias do caos e da complexidade em sistemas, com o objetivo de constituir o campo conceitual necessário para conduzir a investigação. Trata a classificação desde o seu caráter inato, passando pelos estudos filosóficos iniciados na antiguidade, chegando até a abordagem científica da concepção das classificações bibliográficas e das classificações que vem sendo utilizadas para a organização da informação na era da *Web 2.0*. Num segundo momento, a discussão teórica se desenvolve nas teorias do caos e da complexidade em sistemas e suas relações com as redes sociais na Internet, em particular, com a rede social *Twitter* que é uma rede de mensagens postadas pelos usuários do sistema, limitadas a 140 caracteres, sem o agrupamento dos usuários por comunidades ou grupos de interesse, como é comum em outras redes sociais. A pesquisa exploratória e empírica evidencia o uso e as finalidades das *hashtags* no *Twitter*, destacando o papel da colaboração na escolha dos termos. Conclui que a colaboração é essencial para tornar a recuperação da informação mais efetiva, além de atuar na convergência de saberes.

Palavras-chave: Classificação; *Twitter*; Organização do Conhecimento; Web Semântica; Teoria do Caos.

## ABSTRACT

SOUSA, Amanda Moura de. **Organização em sistema caótico**: uso das *tags* para classificação da informação pelos usuários da rede social *Twitter*. Rio de Janeiro, 2012. Orientadora: Rosali Fernandez de Souza. 105 f. Dissertação (Mestrado em Ciência da Informação) Faculdade de Administração e Ciências Contábeis, Universidade Federal do Rio de Janeiro, Instituto Brasileiro de Informação em Ciência e Tecnologia, Rio de Janeiro, 2012.

It discusses issues related to classification and knowledge organization, folksonomy, information retrieval and the theories of chaos and complexity in systems, in order to approach the concepts needed to conduct the investigation. Approaches classification since its innate character, through the philosophical studies started in antiquity, even to the scientific approach to design of bibliographic classifications, and the classifications that have been used for organizing information in the Web 2.0 era. The theoretical discussion is developed on the theories of chaos and complexity in systems and their relationship with social networking sites. In particular with the social network Twitter which is a network of messages posted by users of the system, limited to 140 characters, without forming user groups by communities or groups of interest, as it is common in other social networks. The exploratory and empirical research shows the use and purpose of hashtags in Twitter, highlighting the role of collaboration in the choice of terms. Concludes that collaboration is essential to make information retrieval more effective, as well as acting in the convergence of knowledge.

Keywords: Classification; Twitter, Knowledge Organization, Semantic Web, Chaos Theory.

## LISTA DE FIGURAS

<b>Figura 1:</b> Datas de lançamento dos principais sites de redes sociais.....	40
<b>Figura 2:</b> Tipologia das mídias sociais na Internet .....	42
<b>Figura 3:</b> Mensagens selecionadas para o estudo por perfil de emissores .....	54
<b>Figura 4:</b> Distribuição de mensagens com hashtags por tipo de usuário.....	55
<b>Figura 5:</b> Percentual de mensagens com hashtags segundo total de mensagens de cada tipo de usuário .....	56
<b>Figura 6:</b> Quantitativo de hashtags por mensagem .....	56
<b>Figura 7:</b> Quantitativo de hashtags por mensagem distribuídas por tipos de usuários.....	57
<b>Figura 8:</b> Mensagens contendo hashtags de cada usuário comum .....	58
<b>Figura 9:</b> Total de hashtags usadas por cada usuário comum .....	59
<b>Figura 10:</b> Número de ocorrências de hashtags por usuário comum .....	59
<b>Figura 11:</b> Mensagens contendo hashtags de cada personalidade.....	64
<b>Figura 12:</b> Total de hashtags usadas por cada personalidade .....	65
<b>Figura 13:</b> Ocorrências de hashtags por personalidade .....	65
<b>Figura 14:</b> Mensagens contendo hashtags de cada instituição .....	69
<b>Figura 15:</b> Total de hashtags usadas por cada instituição.....	69
<b>Figura 16:</b> Ocorrências de hashtags por instituição.....	70
<b>Figura 17:</b> Quantidade de sites para o Twitter segundo a finalidade .....	75
<b>Figura 18:</b> Total de citações das hashtags #rio2016 e #london2012 no site <a href="http://hashtagbattle.com/">http://hashtagbattle.com/</a> .....	77
<b>Figura 19:</b> Pesquisa realizada no site HashParty com a hashtag #ThankYouSteve .....	78
<b>Figura 20:</b> Página inicial do site <a href="http://hashtags.org">hashtags.org</a> .....	79
<b>Figura 21:</b> Significado da palavra hashtags.....	80
<b>Figura 22:</b> Rankings do site <a href="http://tagdef.com">tagdef.com</a> .....	81
<b>Figura 23:</b> Definição da hashtag #ff .....	82
<b>Figura 24:</b> Tag cloud das hashtags de @ladygaga .....	83
<b>Figura 25:</b> Impacto positivo, negativo e neutro dos tópicos e hashtags mais recentes .....	84
<b>Figura 26:</b> Exemplo de interação através das hashtags no TweetChat.....	85
<b>Figura 27:</b> <i>Twitoaster</i> .....	86
<b>Figura 28:</b> Busca por #rio2016 no site <a href="http://monitter.com">monitter.com</a> .....	87
<b>Figura 29:</b> Ranking de 2011 no <a href="http://statweestics">statweestics</a> .....	88
<b>Figura 30:</b> <i>Hashtags</i> e tópicos na Holanda no <a href="http://twimmer.com">twimmer.com</a> .....	89
<b>Figura 31:</b> Página inicial do site <a href="http://buzzter">buzzter</a> .....	90

## SUMÁRIO

<b>1 INTRODUÇÃO</b> .....	10
<b>2 CONTEXTUALIZAÇÃO DO TEMA</b> .....	14
2.1 ORGANIZAÇÃO E REPRESENTAÇÃO DO CONHECIMENTO .....	14
2.1.1 Classificação e representação como processos cognitivos .....	15
2.1.2 Classificação na recuperação da informação .....	20
2.2 ORGANIZAÇÃO E RECUPERAÇÃO DA INFORMAÇÃO NA WEB .....	31
2.3 A REDE SOCIAL <i>TWITTER</i> : ORDEM E CAOS .....	39
<b>3 OBJETIVOS</b> .....	51
<b>4 ANÁLISE DO USO DAS HASHTAGS E SITES PARA CLASSIFICAÇÃO DAS MENSAGENS NO <i>TWITTER</i>: METODOLOGIA</b> .....	52
4.1 DELIMITAÇÃO DO UNIVERSO DA PESQUISA .....	53
4.2 DINÂMICA DO USO DAS <i>TAGS</i> NA REDE SOCIAL <i>TWITTER</i> .....	54
4.2.1 As <i>hashtags</i> do grupo usuários comuns .....	58
4.2.2 As <i>hashtags</i> do grupo personalidades .....	64
4.2.3 As <i>hashtags</i> do grupo instituições .....	68
4.3 A COLETIVIDADE NA CLASSIFICAÇÃO DA INFORMAÇÃO: OS SITES DE APOIO À CONSTRUÇÃO DE <i>TAGS</i> .....	74
<b>5 CONSIDERAÇÕES FINAIS</b> .....	92
<b>REFERÊNCIAS</b> .....	96

## 1 INTRODUÇÃO

A Internet exerce papel fundamental no acesso à informação no mundo contemporâneo. A chegada da chamada computação nas nuvens (*cloud computing*) transporta para a virtualidade o que antes dependia de programas instalados em computadores. A chamada *Web 2.0* agrega num único espaço diversos aplicativos para publicação e disseminação da informação, com destaque para as redes sociais on-line, cujo volume de informação que circula diariamente é numericamente representativa, gerando a necessidade de desenvolver uma forma de organizar e classificar essa informação, considerando a complexidade do hipertexto.

O ato de classificar acompanha o cotidiano dos seres humanos, muitas vezes de forma imperceptível, da definição das atividades rotineiras à organização do pensamento para a comunicação. A classificação é, antes de tudo, um processo cognitivo essencial para a escolha dos elementos que melhor representam uma ideia, ou um conjunto de ideias visando expressá-las de forma mais clara para si mesmo ou para outros indivíduos.

Partindo desse ponto, é natural que o conhecimento humano esteja relacionado a um domínio, constituindo, assim, um domínio do conhecimento – que pode ser entendido como categorias ou espaços desenvolvidos para o agrupamento de ideias e conceitos que se relacionam. A concepção de um domínio de conhecimento se torna clara a partir da abordagem filosófica de Aristóteles (384 a. C.) que visava a identificação de categorias de predicados do ser. O estudo desse filósofo foi fundamental para diversas áreas, com destaque para a Biblioteconomia e Ciência da Informação.

Na área da Informação, os estudos em classificação iniciam-se com a Biblioteconomia, a partir do século XIX com o desenvolvimento de diversas teorias e esquemas de classificações bibliográficas. Dentre os principais esquemas podemos destacar a Classificação Decimal de Dewey (1876), a Classificação Decimal Universal (1905) e a Classificação de Dois Pontos (1933). Cabe ressaltar que esses esquemas foram desenvolvidos com a finalidade principal de organizar fisicamente documentos impressos em uma biblioteca.

Na década de 40, a organização da informação torna-se uma preocupação para os profissionais, com o evento chamado de explosão bibliográfica pela Royal Society Conference of Scientific Information (1948). A explosão bibliográfica representa um crescimento, quase exponencial da produção científica e gerou uma grande preocupação no que diz respeito à organização e ao acesso à informação. A partir desse momento fica

evidente que a classificação adquire os objetivos de representar e recuperar informação e não apenas estabelecer a ordem física de documentos impressos.

O traçado do histórico das classificações em Ciência da Informação se faz necessário para esta pesquisa para a compreensão do contexto de desenvolvimento dos principais sistemas de classificação da informação, ou seja, a identificação da motivação de construção desses sistemas e a quem se destinam. Isso é importante para sinalizar algumas transformações da organização da informação impressa e virtual.

Com o desenvolvimento dos sistemas automatizados de informação a partir dos anos 80 e, a partir da década de 90, com o surgimento da *World Wide Web* (www, w3, ou simplesmente *web*) os estudos em classificação e indexação da informação precisaram ser discutidos e transformados, assim como a Internet transformou as relações humanas e o acesso à informação. A partir desse momento, a Internet torna a informação globalizada, proporcionando acesso ao universo de informação on-line através da interligação de documentos através de links (BERNERS-LEE; CAILLIAU, 1992), utilizando uma interface bastante amigável ao usuário.

A nova ordem de organização e disseminação da informação foi instaurada pelo acesso à rede. Até esse momento a produção e transferência eram centralizadas e referenciais, agora não apenas de um indivíduo ao seu receptor, mas de um indivíduo para muitos receptores. Anos mais tarde, em meados dos anos 2000 surge a chamada *Web 2.0*, oferecendo uma nova gama de produtos e serviços de informação baseados em conteúdos elaborados de forma colaborativa e interativa, que favorecem a troca e a organização da informação. Mais precisamente, a *Web 2.0* utiliza a Internet como uma “plataforma” envolvendo *wikis*, aplicações baseadas em folksonomia, redes sociais e tecnologia da informação.

A partir do exposto acima, a *Web 2.0* é uma rede composta por uma pluralidade de domínios, que, por sua vez, se constituem a partir das necessidades de interação e colaboração de seus usuários. Nesse contexto, as redes sociais são, em sua natureza, espaços dinâmicos de interação entre pessoas e instituições de acordo com seus interesses. Diante da variedade de informações que circula nesses canais, instrumentos de classificação, organização e recuperação da informação precisaram ser criados e desenvolvidos como fruto dessa realidade.

A interatividade proposta pela *Web 2.0* se reforçou a partir da reformulação dos conceitos de redes sociais na Internet. Iniciadas em 1997 com o lançamento da rede *Sixdegrees*, as redes sociais conquistam seu espaço na sociedade em 2002, com o lançamento da *Friendster* e, posteriormente com os lançamentos das redes *MySpace* e *Orkut*. Nas redes

sociais na Internet, as relações sociais existentes presencialmente são transportadas para o ambiente virtual, de maneira a ampliar o alcance das relações humanas, visto que é possível conhecer os amigos de um contato, e os amigos em comum que um usuário e outro possam ter.

Em 2006, surge a rede social *Twitter*, que traz consigo um novo modelo de interação. Nessa rede, os usuários são chamados de “seguidores” (*followers*) e se comunicam através de mensagens limitadas em 140 caracteres, sem a presença de comunidades de interesse, fóruns ou grupos de discussão, como ocorre nas redes Facebook e Orkut, por exemplo. Para a classificação e indexação das mensagens é facultado ao usuário atribuir uma *tag* (etiqueta em português) à sua mensagem para classificá-la e representá-la.

Motivadas pela pergunta “O que está acontecendo?”<sup>1</sup>, a divulgação dessas curtas mensagens associam a essa rede diversas outras funções além de relações entre pessoas, como a função de servidor de *microblogging*, *feed* de notícias, divulgação de campanhas publicitárias, coberturas de eventos em tempo real (partidas esportivas e plenárias, por exemplo) entre outras. Essa diversidade de funções gera um volume imenso de mensagens com conteúdos diversos.

A diversidade de funções, o volume e a pluralidade de assuntos e a ausência de comunidades, grupos ou outro modo de filtrar e organizar as mensagens são alguns dos fatores que atribuem à rede uma instabilidade e complexidade no que diz respeito à organização da informação. A partir das considerações acima, a presente pesquisa tem como objetivo estudar as *tags* para a classificação da informação na rede social *Twitter*, analisando o desenvolvimento e o uso das *tags* para a organização da informação, utilizando como pano de fundo a essência imprevisível e caótica da rede social em questão.

A motivação para esta pesquisa teve início a partir do levantamento inicial de informações sobre o tema redes sociais na Internet nas revistas especializadas em Ciência da Informação no Brasil. Foram encontrados poucos artigos sobre o tema e menos ainda sobre o *Twitter*, mesmo com popularização dessa rede no Brasil. Outro incentivo para a realização desse estudo está na emergência de uma extensão da Web baseada na semântica, interligando os significados às palavras, permitindo aos computadores e seres humanos o trabalho em conjunto. Essa extensão, chamada por alguns especialistas de *Web 3.0* tem início na *Web 2.0* com a popularização da folksonomia e de sistemas de classificação e indexação nas redes sociais, como a utilização das *tags* visando a recuperação da informação.

---

<sup>1</sup> Tradução livre do inglês “What’s happening?”

Dentro do contexto apresentado até aqui surgiram algumas questões que estimularam a realização desse trabalho: Existem semelhanças entre os sistemas de classificação já formalizados e o das *hashtags*<sup>2</sup>? Quais as principais diferenças entre eles? Existe ordem em meio ao caos? As *hashtags* possuem importância para uma única mensagem ou há a busca por um padrão de classificação?

Com o levantamento dos dados realizado para esta pesquisa foi possível responder algumas dessas perguntas, assim como informações importantes acerca da adesão ao uso de *hashtags* pelos usuários do *Twitter* e a influência de fatores externos à rede na escolha das palavras que compõem as *hashtags*.

Essa dissertação divide-se metodologicamente em duas partes. A primeira é dedicada à contextualização do tema obtida através da revisão bibliográfica de conteúdos relacionados à organização do conhecimento, classificação, representação da informação, recuperação da informação, Internet, *Web 2.0*, folksonomia, Web Semântica e redes sociais.

A segunda parte compreende o levantamento de *hashtags* no *Twitter* e a busca e recuperação de sites de apoio ao uso do *Twitter*. Os dados coletados foram estruturados e os resultados analisados e interpretados.

A partir dos dados coletados tanto na revisão bibliográfica quanto no levantamento dos dados, são apresentadas considerações sobre o uso das *hashtags* para classificação de mensagens no *Twitter*, observando as características desse sistema sob a ótica das teorias do caos e da complexidade. A abordagem do tema busca estabelecer um paralelo entre a classificação da informação em rede e os sistemas de classificação desenvolvidos para outros tipos de registros de informação, contextualizando a realidade de cada um a fim de identificar semelhanças e diferenças em suas aplicações, e com isto contribuir para a investigação no campo da organização e representação da informação no contexto da Ciência da Informação.

---

<sup>2</sup> As *tags* no *Twitter* são denominadas *hashtag* por ser obrigatória a utilização símbolo # que é chamado de *hash* nos países do Reino Unido.

## 2 CONTEXTUALIZAÇÃO DO TEMA

O referencial teórico que integra o presente trabalho concentra-se em três eixos principais: organização e representação do conhecimento, organização e recuperação da informação na web e na ordem e caos na rede social *Twitter*.

A organização e representação do conhecimento são abordadas em relação à recuperação da informação. São discutidos temas relacionados à classificação enquanto processo cognitivo e aos sistemas formais classificação para a recuperação: sistemas de classificação bibliográficos e não bibliográficos, taxonomias, tesouros e ontologias.

A organização e recuperação da informação na *Web 2.0* destaca a colaboração no desenvolvimento de produtos e serviços on-line, em especial de produtos e serviços para organização da informação. Em seguida são explorados conceitos sobre folksonomia e *tags* para classificação de conteúdo on-line na *Web 3.0*.

A abordagem da ordem e do caos na rede social *Twitter* compreendem um breve histórico das redes sociais na internet as relações da teoria do caos e complexidade com as características da rede social em questão.

### 2.1 ORGANIZAÇÃO E REPRESENTAÇÃO DO CONHECIMENTO

Esta parte da pesquisa destina-se à discussão teórica da classificação enquanto processo cognitivo e das classificações no âmbito da organização do conhecimento para a recuperação da informação.

A classificação como processo cognitivo aborda essa atividade como essencial para a representação do conhecimento a ser transmitido. São discutidos os conceitos de representação, categorização e classificação como coexistentes na organização do conhecimento humano.

As classificações estruturadas para a recuperação da informação são destacadas como sistema formal de organização do conhecimento, incluindo a análise de características de alguns sistemas de classificação.

### 2.1.1 Classificação e representação como processos cognitivos

Classificar é uma ação incorporada ao dia a dia do ser humano, e está presente nas tarefas mais básicas até as mais complexas. Langridge<sup>3</sup> (2006), ao narrar o cotidiano de um casal, observa o papel da classificação em diversas decisões que o homem toma diariamente, a organização dos itens no banheiro e a disposição das notícias na primeira página do jornal, por exemplo, que refletem as preferências e finalidades de classificação para objetivos definidos.

O dicionário Houaiss (2011), aponta cinco definições para o verbo transitivo direto *classificar*:

- Distribuir em classes e nos respectivos grupos, de acordo com um sistema ou método de classificação;
- Determinar a classe, ordem, família, gênero e espécie;
- Pôr em determinada ordem, arrumar (coleções, documentos etc.);
- Aprovar ou ser aprovado em exame seletivo, concurso, competição etc.;
- Emitir opinião ou julgamento a respeito de (outrem ou de si); considerar(-se).

A partir dessas definições, observam-se as noções essenciais do ato de classificar: classe, ordem, arranjo, família, gênero e espécie. A compreensão desses conceitos constitui passo importante para a análise do ato de classificar enquanto atividade humana e científica. O mesmo dicionário define classificação como substantivo feminino que designa ato ou efeito de classificar(-se), além das seguintes acepções em diversas rubricas:

1. Distribuição por classes
2. Rubrica: bibliologia.  
Em bibliotecas, filmotecas, melotecas, gliptotecas e grupamentos documentais afins, sistema de distribuir (em classes, subclasses, temáticas etc.) as unidades e a sua localização.
3. Rubrica: biologia.  
Arranjo sistemático ou método de arranjo dos seres vivos, em uma série hierárquica de categorias ou grupos.
4. Num concurso, exame seletivo, competição etc., posição obtida por um candidato.  
Ex.: sua c. em segundo lugar foi ótima.
5. Rubrica: termo jurídico.  
Definição do delito ou contravenção para determinar a fixação da pena que lhe corresponde
6. Rubrica: sociologia.  
m.q. *peneiramento* ('processo') (HOUAISS, 2011, não paginado).

---

<sup>3</sup> Reimpressão da primeira edição em língua portuguesa do ano de 1977.

Em outro tipo de abordagem, a Enciclopédia Einaudi (c1985) estabelece que “classificar constitui atividade intelectual primitiva” e é construída sobre princípios individualizados dos entes que categorizam. Essa discussão segue as ideias de Jean Piaget (1896-1980) que apresenta três estágios principais na elaboração das classificações:

1. Construção de coleções figurais – capacidade de reconhecer relações de semelhanças, sem considerar a relação parte/todo.
2. Agregados baseados apenas em semelhanças, permanecendo justapostos sem a inclusão em classes mais gerais.
3. Conquista da noção de inclusão, e, portanto, de hierarquização, resultando na construção de formas que “lhes [aos seres humanos] são progressivamente isomorfas.”

Ainda sobre a classificação, Satija (1998, p. 32) identifica sua manifestação nas seguintes atividades: nomeação, definição, análise, generalização, discriminação, distinção, padronização, distribuição, filtragem, demarcação, separação, individualização, identificação, categorização, agrupamento, correspondência, seleção, amostragem, organização, ordenação, posição, correlação, tabulação, mapeamento, desenho, estruturação, coordenação e controle.

Para este estudo, a utilização dos conceitos classe e categoria se dará de formas distintas, pois segundo Langridge (2006) mesmo os termos sendo originalmente sinônimos, as categorias tratam das classes mais gerais de fenômenos. As categorias, portanto, representam parte significativa para a compreensão da classificação.

Na Enciclopédia Einaudi, o verbete “categorizar” discute diversas ideias filosóficas distintas relacionadas ao processo de categorização pelo ser humano e define categorias como:

Representações genéricas da experiência, mais exatamente, critérios que presidem à distribuição e ordenação da pregnância da experiência nos seus diversos aspectos. E as atividades classificatórias prolongam diretamente procedimentos de identificação e reconhecimento que se dão já na representação sensível.

Mas é assim porque as categorias habitam a representação; é por ela proceder a uma categorização mínima de experiência que um R. Bacon considerou semelhança/diversidade e contínuo/discreto como determinações próprias à esfera do sensível. Por si só, a percepção reúne já semelhanças e seleciona contrastes, estabelecendo saliências contra fundos indiferenciados; e, desse modo, se organiza categorialmente, no homem, como em todo animal, quando espontaneamente faz classificações: a “síntese da reconhecimento no conceito” começa a fazer-se dentro da “síntese da apreensão na intuição” [Crítica da Razão Pura, A98-106] (EINAUDI, c1985, v. 41, p.53)

Como é possível observar, a classificação envolve diversos outros processos, com ênfase na categorização e na representação. Esta pesquisa pretende investigar sobre a confluência desses processos como essencial para o desenvolvimento seja das classificações, das categorizações e das representações.

Ainda segundo a Enciclopédia Einaudi (c1985, v. 41, p.12), “em todas as formas de representação uma coisa se encontra no lugar de outra, representar significa ser o outro dum outro que a representação, num mesmo movimento, convoca e revoca”. A partir dessa definição básica, representar, enquanto ação cognitiva se relaciona com a identificação de um objeto em outro, de modo que um possa apresentar o outro.

Os estudos que agregam a classificação, categorização e representação discutidos até aqui têm suas origens nas áreas da educação e da filosofia, o que é justificável, pois são áreas que investigam o ser humano como indivíduo e como ser social, o que envolve a exploração do conhecimento e suas representações. Alguns desses estudos começaram a ser desenvolvidos desde a Antiguidade, sendo retomados ao longo da história.

Aristóteles (384 a. C. – 322 a. C.) assinala um conjunto de escritos denominado *Metafísica*, que se concentra em seus estudos sobre o ser. Entre diversos questionamentos a cerca da existência do ser, Aristóteles identificou algumas categorias que se aplicariam à organização das experiências humanas para que fosse possível compreender como o ser expressa e pensa o conhecimento. Inicialmente, o filósofo concentrou seus estudos em dez categorias, que seriam os princípios básicos que permitiriam o conhecimento das coisas, “partindo de uma perspectiva que concebe o mundo como um todo integrado de causas e efeitos, em que as coisas e as suas propriedades essenciais, seus estados, processos e relações podem e devem ser reveladas pelo trabalho intelectual” (ARANALDE, 2009).

As categorias iniciais de predicados do ser concebidas por Aristóteles foram as seguintes: **substância, qualidade, quantidade, relação, agir, sofrer, lugar e tempo**. Essas categorias seguem uma construção baseada em princípios de lógica. A substância é a categoria imutável, qualidade mais profunda que subsiste no ser, e que sustenta as outras nove. Algum tempo depois, Aristóteles percebeu que seria possível reduzir essas dez categorias em três: **substância, modo e relação**.

No século IV, o filósofo grego Porfírio formula, a partir das categorias de Aristóteles, princípios para a subdivisão de classes, denominados predicáveis ou categoremata: **gênero, espécie, diferença, propriedade e acidente**. Entre os cinco categoremata, as noções de gênero e espécie são encontrados em vários estudos da área. Para Piedade (1983, p. 23), gênero “é

um conjunto de coisas ou ideias que podem ser divididas em dois ou mais grupos ou espécies” e espécies “são os vários grupos resultantes da divisão de um gênero por determinada característica”.

Séculos mais tarde, Francis Bacon (1561-1626) abordava o conhecimento verdadeiro estudando como a razão e o pensamento poderiam eliminar o erro. Sendo assim, a teoria do conhecimento torna-se uma disciplina centrada na filosofia. Mediante o estudo das causas e das formas do erro, acreditava que com o avanço das técnicas e dos conhecimentos, as ciências e a filosofia se desenvolveriam a ponto de propiciarem uma grande reforma no conhecimento humano e na vida humana (CHAUI, 1995). A partir disso pode-se compreender seu objetivo ao escrever *The Advancement of Learning* (1605).

Na obra *The Advancement of Learning* (1605), Bacon produz um sistema de organização que envolve uma análise exaustiva sobre todas as áreas do conhecimento estabelecendo um plano geral sobre elas apontando ainda quais áreas ainda precisavam ser desenvolvidas. O sistema classificatório das ciências proposto por Bacon é dividido em três grupos: 1) Poesia e imaginação; 2) História ou ciência da memória, que se subdivide em natural e civil natural e civil; 3) Filosofia ou ciência da razão, que se subdivide em filosofia da natureza e antropologia. Ao comparar os estudos de Bacon com o dos outros filósofos que integram esta parte do trabalho é possível identificar que as classificações se destinam ao conhecimento já estabelecido, enquanto que as categorias estão mais relacionadas ao pensamento ou às formas de pensar do ser.

Immanuel Kant (1724-1804) concentra seus estudos não mais nas categorias inerentes ao ser, mas categorias relacionadas às formas de pensar do ser. Kant estabeleceu a Tábua das Categorias, que apresenta doze conceitos distribuídos em quatro grupos:

- a) quantidade: unidade, pluralidade e totalidade;
- b) qualidade: realidade, negação e limitação;
- c) relação: inerência-subsistência, causalidade-dependência e comunidade;
- d) modalidade: possibilidade-impossibilidade, existência-não-existência e necessidade-contingência.

Kant propõe a observação do conhecimento a partir de pares dicotômicos de conceitos, seguindo a tendência do pensamento científico da época. Isso significa dizer que o saber transita por diversas perspectivas, muitas vezes antagônicas, sem que isso implique na anulação de um ou outro. Na sua visão, eram esses pontos divergentes que tornavam possível o conhecimento da coisa, tal como ela se apresenta (ARANALDE, 2009).

Na era moderna, os estudos buscam relacionar os modelos cognitivos com a formação de categorias pelas pessoas. Lakoff utiliza a ideia do modelo cognitivo como uma estrutura mental usada para a categorização. A partir dessa concepção, trata dos seguintes temas:

**Semelhanças de família:** a ideia de que os membros de uma categoria podem estar relacionados entre si sem que todos os membros tenham em comum todas as propriedades que definem a categoria. **Centralidade:** a ideia de que alguns membros de uma categoria podem ser "melhores exemplos" dessa categoria do que outros. **Polissemia como categorização:** a ideia de que os significados relacionados de palavras formam categorias e que os significados apresentam semelhanças de família entre si.

**Geratividade como um fenômeno protótipo:** essa ideia diz respeito a categorias que são definidas por um gerador (um membro especial ou subcategoria) mais regras (ou um princípio geral como a semelhança). Nesses casos, o gerador tem o status de um membro da categoria central, ou "prototípico"

**Composição por gradação:** a ideia de que, pelo menos, algumas categorias apresentam graus de adesão e não possuem fronteiras claras.

**Gradação de centralidade:** a ideia de que membros (ou subcategorias) que estão claramente dentro dos limites da categoria podem ainda ser mais ou menos centrais. **Incorporação conceitual:** a ideia de que, as propriedades de determinadas categorias são uma consequência da natureza de capacidades biológicas humanas e da experiência de funcionamento em um ambiente físico e social, é contrastada com a ideia de que conceitos existem independentemente da natureza do corpo de qualquer seres pensantes e independentes de sua experiência.

**Incorporação funcional:** a ideia de que certos conceitos não são apenas entendidos intelectualmente, mas sim, eles são usados automaticamente, inconscientemente, e sem esforço notável como parte do funcionamento normal. Os conceitos utilizados desta forma têm um estado psicológico diferente, e mais importante, do que aqueles que são pensados de forma consciente.

**Nível básico de categorização:** a ideia de que categorias de nível básico não são apenas organizadas em uma hierarquia a partir do mais geral ao mais específico, mas também são organizadas de modo que as categorias que são cognitivamente básicas estão "no meio" de uma hierarquia geral específica. Generalização procede "para cima" a partir do nível básico e a especialização prossegue "para baixo".

**Primazia de nível básico:** a ideia de que categorias de nível básico são funcional e epistemologicamente primárias no que diz respeito aos seguintes fatores: percepção gestáltica, formação da imagem, movimento do motor, organização do conhecimento, caso de processamento cognitivo (aprendizagem, reconhecimento, memória, etc) e o caso de expressão linguística.

**Ponto de referência ou raciocínio "metonímico":** a ideia de que uma parte de uma categoria (ou seja, um membro ou subcategoria) pode permanecer para toda a categoria em determinados processos de raciocínio (LAKOFF, 1987, p. 13-14, tradução nossa).

Classificar muitas vezes implica em posicionar objetos com rótulos de maneira que os mesmos estejam à disposição para uso posterior. De imediato, pode-se pensar que se trata de uma afirmação direcionada para algo físico, mas é algo que acontece com as ideias e pensamentos no momento em que há algum contato com a informação. Sendo assim, há a sempre a busca em relacionar ou aproximar o desconhecido com algo familiar.

Para além das percepções práticas da classificação na cognição, a ciência cognitiva vem se desenvolvendo no sentido de explorar mais a fundo diversas atividades da mente humana, como o raciocínio e a aprendizagem. Segundo Lima (2003, p. 81) o processo cognitivo está associado a diversas atividades mentais como “o pensamento, a imaginação, a recordação, a solução de problemas, a percepção, o julgamento, a aprendizagem da linguagem, entre outras, as quais ocorrem diferentemente em cada indivíduo, dependendo do grau de habilidade de cada um”. Por explorar tantas frentes, fica claro a interdisciplinaridade da ciência cognitiva e a proximidade com a Psicologia, a Linguística, Pedagogia e Ciência da Informação.

Discutir a classificação como processo fundamental para o conhecimento humano revela aspectos determinantes para a compreensão do encadeamento das ideias para diversas atividades. Para esta pesquisa, o estudo das classificações, categorizações e representações no nível da cognição é essencial para conhecer as bases que compõem alguns dos sistemas de classificação da informação produzida e registrada.

### 2.1.2 Classificação na recuperação da informação

Os estudos dos processos de classificação e suas relações com a cognição humana serviram de base para diversos estudos posteriores em classificação. No século XIX, Melvil Dewey retomou a ideia das dez categorias de Aristóteles, porém, a categorização não busca mais a identificação de características do ser, mas o agrupamento de documentos segundo a sua principal característica: o assunto. Para Dewey, o conhecimento humano registrado em livros poderia ser dividido em: 000 - Generalidades e Miscelânea; 100 - Filosofia; 200 - Religião; 300 - Ciências Sociais; 400 - Línguas; 500 - Ciências puras; Ciências aplicadas; 700 - Artes; 800 - Literatura; 900 - História e Geografia.

Entender o processo de classificar se inicia com o conhecimento de seus conceitos básicos. Piedade (1983) dedica um capítulo inteiro à análise do que a autora chama de “conceitos fundamentais”. Segundo Piedade, os conceitos principais da classificação são: característica; classe; categorias; gênero e espécie; facetas, focos e isolados; divisão em cadeia; divisão em fileira; extensão e compreensão; modulação; termo. O presente trabalho apresenta, inicialmente, os conceitos de característica, classe e categorias.

Para Piedade (1983, p. 17) característica é “a qualidade ou atributo escolhido para servir de base à classificação ou à divisão”, ou seja, podem vários tipos de classificação, cada uma obedecendo à uma determinada característica. As classes seriam conjuntos de coisas ou

ideias que possuem uma ou mais características em comum. A autora define categorias como as maiores classes de fenômenos que podem ser geradas, assim como Langridge (2006). Esses conceitos estão presentes já nos estudos mais antigos de classificação.

No século XX, o matemático e bibliotecário indiano Shiyali Ramamrita Ranganathan (1892-1972) dá início ao que Dahlberg (1972) chama de moderna teoria da classificação. Em 1937, Ranganathan escreve o *Prolegomena to Library Classification*, com o objetivo de formular uma teoria para a constituição de sistemas de classificação. Para Ranganathan, a classificação de um livro necessita de mais que a simples determinação prévia de classes de assuntos, mas a criação das classes somente no momento em que o documento é analisado.

Dentre as diversas contribuições de Ranganathan, cabe a análise, em primeiro lugar dos três níveis distintos em que trabalham os classificacionistas: o plano da ideia; o plano verbal; e o plano notacional. No plano das ideias a abordagem parte das ideias e dos conceitos. No plano verbal é analisada a expressão verbal dos conceitos (ex. variações linguísticas). No plano notacional está a escolha das formas abstratas que representarão os conceitos, tais como letras e números. Ainda segundo Dahlberg (1972) outras duas contribuições importantes de Ranganathan para a construção das teorias de classificação foram a abordagem analítico-sintética para a identificação dos assuntos e os seus princípios de sequência útil.

Os princípios de sequência útil são divididos em dois blocos: no primeiro, estão os princípios normativos: princípio da lei da biblioteconomia, princípio da lei de interpretação, princípio da lei da imparcialidade, princípio da lei da simetria, princípio da lei parcimônia, princípio da lei da variação local, princípio da lei da osmose. No segundo bloco, estão os princípios para ordenação em qualquer nível: princípio do posterior-no-tempo, posterior-na-evolução, princípio da contiguidade espacial, princípio da medida quantitativa, princípio da complexidade crescente, princípio da sequência canônica, princípio da garantia literária, princípio da ordem alfabética.

No plano das ideias, Ranganathan identificou cinco categorias fundamentais para a organização do conhecimento, conhecidas como PMEST: P = personalidade, M = matéria, E = energia, S = espaço e T = tempo. Dentre as cinco categorias fundamentais de Ranganathan, Kumar (1981) expõe que as categorias espaço e tempo são utilizadas nas classificações da mesma forma como são compreendidas e são mais facilmente definidas. As outras três categorias - energia, matéria e personalidade - já não são tão simples de serem observadas. Essas categorias devem ser aplicadas ao contexto da classificação enquanto disciplina, ainda de acordo com Kumar (1981). Para compreender melhor essa afirmação é relevante a análise

sobre o que as categorias se propõem a organizar a partir das considerações do autor acerca de cada uma delas:

- **Energia:** essa categoria relaciona-se a uma ação ou realização. Pode manifestar-se entre e por todos os tipos de entidades – inanimadas, animadas, conceituais, intelectuais e intuitivas. Uma das dificuldades ao se trabalhar com essa categoria é que ela pode ocorrer em um mesmo objeto mais de uma vez.
- **Matéria:** essa categoria pode se manifestar de três formas: material (ex. notebook, caneta, etc.); propriedade de uma matéria – possível de ser compreendida pelo autor através do uso de analogias como o caso da doença para a Medicina; e um método relacionado a uma matéria, ainda mais difícil de definição – método volumétrico (apropriação da matéria volume para compor uma outra matéria que seria o método volumétrico). Esse último tipo nem sempre é considerado uma categoria por outros autores.
- **Personalidade:** essa categoria fundamental é a mais complexa para se definir. De maneira geral, essa categoria é uma característica que distingue algo de outras coisas, como por exemplo, a forma de algum órgão do corpo humano. A categoria personalidade denota a coisa em si, formando a base de todas as outras categorias. A definição desta categoria é extremamente difícil e é constantemente confundida com a categoria matéria. Para facilitar a compreensão, um exemplo dado pelo autor é “tratamento de doenças do coração”, onde a personalidade, ou seja, o que diferencia e especifica o tipo de tratamento e de doenças é o coração.

A reflexão de Ranganathan sobre uma teoria da classificação o levou a idealizar um sistema de classificação, a Classificação dos Dois Pontos (*Colon Classification*). Esse sistema previa a materialização de seus princípios descritos no *Prolegomena to Library Classification*, ou seja, priorizava a orientação das chamadas facetas - que seriam as representações de áreas especializadas do conhecimento - sintetizando assuntos em expressões combinadas que formariam uma classe a partir da análise de tópicos de um documento. Barbosa define classificação facetada como:

*Classificação facetada* é o sistema que agrupa termos estruturados, na base da análise de um assunto, para identificação de suas *facet*s, isto é, dos diferentes aspectos nele contidos. A *análise em facet*s coordena conceitos, significando que um assunto, por mais complexo que seja, pode ser representado pela síntese de mais de uma faceta, cada uma indicando conceitos diferentes (BARBOSA, 1972, p. 74).

Ranganathan porém não foi o único a constituir um sistema para classificação bibliográfica. Desde o fim do século XIX, importantes iniciativas surgiram, em meio à

expansão da ciência e do aumento gradativo (até a II Guerra Mundial, quando ocorreu a explosão bibliográfica) das publicações bibliográficas de modo geral. No quadro abaixo, as classificações são sintetizadas para evidenciar as características de cada sistema:

<b>Esquema</b>	<b>Estrutura</b>	<b>Uso</b>	<b>Notação</b>	<b>Regras</b>
Classificação Decimal de Dewey (CDD)	Dez classes principais	Amplamente usado, principalmente em bibliotecas públicas e escolares	Inicialmente enumerativo, introduzindo aos poucos elementos analítico-sintéticos	Não são especificadas
Classificação Decimal Universal (CDU)	Semelhante às classes principais da CDD, com mais assuntos compostos	Bibliotecas especializadas e universitárias	Números decimais e indicadores de facetas especiais.	Flexíveis e adaptáveis a diversas circunstâncias
Classificação da Library of Congress (LC)	4 classes principais.	Biblioteca do Congresso e algumas universidades dos EUA	Mista de letras e números inteiros.	Não são especificadas.
Classificação Bibliográfica de Bliss	20 classes principais.	Algumas bibliotecas do Reino Unido.	Letras maiúsculas com números e letras minúsculas e pontuação para indicação das facetas	Não são especificadas.
Classificação dos Dois Pontos (C C)	40 classes principais arranjadas em: Ciências e Tecnologia, Misticismo, Artes, Humanidades e Ciências Sociais.	Bibliotecas acadêmicas, especializadas e públicas da Índia.	Mista de letras e números e indicadores especiais de faceta.	Série completa de regras explicitadas.

**Quadro 1:** Principais esquemas de classificação bibliográfica

**Fonte:** Adaptado do texto de Langridge (2006).

Como pode ser observado na síntese apresentada no Quadro 1, os estudos para a constituição de sistemas de classificações bibliográficas são orientados pelo seu uso final. O tipo de biblioteca e o seu contexto político e social foram determinantes para a estrutura e para a construção da notação visto que não seria necessário para uma biblioteca escolar muitos indicadores de facetas, ao contrário de uma biblioteca universitária, que é composta por livros de assuntos interdisciplinares. A influência da sociedade fica mais evidente na Classificação dos Dois Pontos, onde Religião cede lugar ao Misticismo, que está muito mais relacionado à cultura oriental do que na ocidental.

As influências externas na elaboração das classificações também trazem alguns problemas. Ainda sobre as classificações bibliográficas, Edson Nery da Fonseca (1972)

identifica algumas discrepâncias lógicas presentes em alguns sistemas de classificação, como na CDD, em que cita a separação da classe História (900) das Ciências Sociais (300). Outro entrave às classificações bibliográficas é a crescente interdisciplinaridade na ciência, o que prejudica uma ou outra área se o sistema for enumerativo e hierárquico. O autor ressalta ainda que, as classificações bibliográficas foram ainda mais necessárias depois da expansão das ciências, porém, a forma como estão elaboradas precisa de reformulações para representar de forma satisfatória o desenvolvimento da ciência.

Sistemas de classificação não são exclusividade das bibliotecas. Ao longo da história foram criados para servir às mais variadas áreas e aos mais diferentes propósitos. Um dos mais antigos e conhecidos desses esquemas vem do campo da Astrologia, o Zodíaco. Esse sistema se propõe a agrupar pessoas pela posição dos astros no momento do seu nascimento.

Entre os sistemas de classificação não bibliográficos que utilizam números e letras para exprimir conceitos, dois são destacados: a Classificação Internacional de Doenças (CID) e a Classificação Internacional de Patentes (CIP).

A CID e a CIP foram escolhidas para análise por duas razões: a primeira é importância dessas classificações para a sociedade, visto que ambos são adotados em escala mundial. A segunda razão é a presença de algumas semelhanças da composição das notações com os sistemas de classificação bibliográficos já discutidos anteriormente.

A CID começou a ser pensada no século XIX, a partir de uma classificação desenvolvida por um comitê dirigido por Jacques Bertillon (1851-1922). Esta classificação foi baseada em estatísticas das causas de mortes de pacientes de hospitais na França, sendo rapidamente adotada em diversos outros países. Em 1900, foi definido um comitê internacional para a revisão da classificação de Bertillon, que está hoje em sua décima revisão (Organização Mundial da Saúde, não paginado). Atualmente, a classificação se expandiu para uma família de classificações: a CID, a Classificação Internacional de Funcionalidade, Incapacidade e Saúde (ICF), Classificação Internacional de Intervenções na Saúde (ICHI). No quadro 2 é apresentado um exemplo de classificação:

**CAPÍTULO V TRANSTORNOS MENTAIS E COMPORTAMENTAIS (F00-F99)**

**F00\* Demência na doença de Alzheimer (G30.-†)**

**F00.0\* Demência na doença de Alzheimer de início precoce (G30.0†)**

Demência na doença de Alzheimer com início antes da idade de 65 anos, com um curso de deterioração relativamente rápido e com transtornos múltiplos e marcantes das funções corticais superiores.

Demência degenerativa primária do tipo Alzheimer, de início pré-senil

Demência pré-senil, tipo Alzheimer

Doença de Alzheimer, tipo 2

**F00.1\* Demência na doença de Alzheimer de início tardio (G30.1†)**

Demência na doença de Alzheimer com início após a idade de 65 anos, e usualmente ao fim do oitavo decênio (70-79 anos) ou após esta idade; evolui lentamente e se caracteriza essencialmente por uma deterioração da memória.

Demência degenerativa primária do tipo Alzheimer, de início na senilidade

Demência senil do tipo Alzheimer

Doença de Alzheimer, tipo 1

**F00.2\* Demência na doença de Alzheimer, forma atípica ou mista (G30.8†)**

Demência atípica do tipo Alzheimer

**F00.9\* Demência não especificada na doença de Alzheimer (G30.9†)**

**Quadro 2:** Hierarquia da CID para Doença de Alzheimer

**Fonte:** Disponível em: <http://www.datasus.gov.br/cid10/v2008/cid10.htm>. Acesso em: 18 ago. 2011.

No Quadro 2, observa-se que a estrutura da classificação é hierárquica e sua notação é composta por números e letras que podem vir acompanhados do asterístico (\*) para a indicação de que trata-se de código optativo para a doença de Alzheimer ou o símbolo do punhal ou cruz († ou +) para a indicação de um código primário da doença, ou seja, o punhal ou cruz são usados quando a doença é a causa básica manifestada em algum órgão específico.

A escolha da doença de Alzheimer como exemplo foi motivada por conter indicadores que orientam o classificador a optar ou não por determinada classificação. A orientação ao classificador utilizada na CID é semelhante à orientação adotada pela CDD, que utiliza, por exemplo, notas indicando outras possibilidades de classificação e os símbolos da cruz (+) para indicar a extinção de uma classe e o uso de colchetes ([]) para indicar o desuso ou remanejamento de uma subdivisão (20ª edição). Essa semelhança na orientação ao classificador presente no dois sistemas indica que algumas características dos sistemas de classificação são compartilhadas independentemente do tipo de registro do conhecimento (livros no caso da CDD e documentação médica no caso da CID).

A Classificação Internacional de Patentes – CIP – foi criada em 1971 pelo Acordo de Estrasburgo e foi adotada no Brasil em 1975, pelo decreto nº 76.472. Segundo Jannuzzi, Amorim e Souza (2007) a classificação foi criada para uniformizar a descrição das invenções nas patentes de modo a representar o conhecimento relativo às invenções. A classificação é

hierárquica e é composta por seções, classes, subclasses, grupos e subgrupos. A compreensão se torna mais clara a partir do exemplo a seguir:

<b>Seção</b>	G -> Física
<b>Classe</b>	02 -> Óptica
<b>Subclasse</b>	F -> Dispositivos ou disposições, nos quais o funcionamento óptico é modificado pela variação das propriedades ópticas do meio que constitui estes dispositivos ou disposições para o controle da intensidade, da cor, da fase, da polarização ou da direção da luz.
<b>Grupo</b>	1/01 -> Dispositivos ou disposições para controle da intensidade, da fase, da polarização ou da cor.
<b>Subgrupo</b>	5 -> Baseado em elementos semicondutores com pelo menos uma barreira de potencial

**Quadro 3:** Hierarquia da CIP para uma patente sobre disposições óticas

**Fonte:** Dados extraídos da CIP em português no site <http://www.inpi.gov.br/>. Acesso em: 16 set. 2011.

A notação completa para a tecnologia apresentada no Quadro 3 é G02F 1/015 e pode ser descrita como *dispositivos ou disposições óticas para controle da intensidade, da fase, da polarização ou da cor baseados em elementos semicondutores com pelo menos uma barreira de potencial*. A classificação não precisa necessariamente chegar até o nível hierárquico do grupo ou subgrupo, porém, a representação não será muito próxima da tecnologia descrita no documento de patente. Como pode ser observado, este sistema de classificação segue alguns princípios comuns a outros sistemas como os bibliográficos, por exemplo.

A exploração da CIP na presente pesquisa é relevante por apresentar padrão hierárquico de construção da notação semelhante aos sistemas bibliográficos e a CID e, ao mesmo tempo, por conservar uma diferença: um documento de patente pode receber quantas classificações forem necessárias para representar a tecnologia descrita no documento. Quanto mais classificações uma patente receber, maior é a complexidade da tecnologia descrita no documento.

Os sistemas de organização do conhecimento descritos até aqui se utilizam de números e letras para compor as notações que representam um conceito. Isso se justifica de duas maneiras diferentes: nos sistemas bibliográficos para estabelecer uma ordem sequencial para estantes de livros; já na CID e na CIP a ideia é universalizar a descrição dos conceitos independentemente da língua em que o documento esteja escrito. Essa forma de representar, no entanto, não é viável para a recuperação de informação por quem não conhece o que esses

símbolos querem dizer. Diante disso, novos sistemas de representação e organização do conhecimento foram desenvolvidos.

Nas Ciências, as classificações são elaboradas a partir da crescente especialização de cada campo, e se fazem essenciais em algumas delas como no caso da Botânica, em especial. Isso se deve ao fato de que o propósito inicial é organizar um conhecimento em classes que podem estar expressas ou não em documentos. Para esse tipo de classificação, a notação pode ou não ser adotada. Em sua tese de doutorado, Cavalcanti (2005) relaciona algumas dessas classificações da botânica e aponta sua proximidade maior com as taxonomias do que com os sistemas bibliográficos:

SOC	ANO	CLASSIFICACIONISTA
Classificação Botânica de Lineu	1753	Carl Lineu
Classificação Botânica de Jussieu	1789	Antoine Jussieu
Classificação de Bentham e Hooker	1862-1883	Bentham e Hooker
Classificação de Engler	1900	Engler
	1845	Carpenter
	1897	Lydekker
	1903	Hertwig
	1729-1832	Cuvier
	1884-1892	Owen

**Quadro 4:** Classificações da Botânica

**Fonte:** Kaula (1984) adaptado por Cavalcanti (2005, p. 123)

O Quadro 4 apresenta a perspectiva histórica das classificações da Botânica. A relevância deste quadro para a pesquisa está relacionada ao conhecimento dos pioneiros no desenvolvimento dos sistemas taxonômicos que são a base para as taxonomias criadas para a recuperação da informação.

Uma outra ciência cuja classificação é essencial é a Zoologia. Segundo Rodrigues (2009), Carl Von Linné (1707-1778) é apresentado como o “pai da taxonomia” por agrupar animais propondo uma nomenclatura binária, além de atribuindo rótulos aos animais, como “fixista e criacionista”. Com o aprofundamento dos estudos nas áreas de Zoologia, Fisiologia, Anatomia e Genética, a descrição dos animais recebe mais detalhes e novas espécies são identificadas, modificando a taxonomia inicial de Linné.

Sobre as taxonomias, Campos e Gomes (2008, p. 3) ressaltam que as taxonomias surgiram como a ciência das leis da classificação de formas vivas e tiveram como função

inicial a classificação de espécies de plantas na Botânica e de animais na Zoologia. Com o desenvolvimento dos sistemas de recuperação da informação (SRI's), as taxonomias “são estruturas classificatórias que têm por finalidade servir de instrumento para a organização e recuperação de informação em empresas e instituições”. As autoras revelam ainda as características essenciais para esse sistema:

- Conter uma lista estruturada de conceitos/termos de um domínio.
- Incluir termos organizados hierarquicamente
- Possibilitar a organização e recuperação de informação através de navegação.
- Permitir agregação de dados, diferentemente das taxonomias seminais, além de evidenciar um modelo conceitual do domínio.
- Ser um instrumento de organização intelectual, atuando como um mapa conceitual dos tópicos explorados em um Sistema de Recuperação de Informação.
- Ser um novo mecanismo de consulta em Portais institucionais, através de navegação (CAMPOS; GOMES, 2008, p. 3).

A partir do exposto acima sobre as taxonomias é importante expor sobre o conceito do **termo**, visto que esse tipo de classificação do conhecimento utiliza as palavras ao invés da combinação de números decimais e letras para a representação de um conceito.

A aproximação entre a Ciência da Informação (CI) e a Linguística ocorre por diversos objetivos. No presente estudo, concorda-se com Mendonça (2000) no que concerne a maior interseção da CI com a Linguística em estudos que se referem à representação do conhecimento, mais precisamente, a relação entre terminologia e representação temática do conhecimento. Para esta conclusão, a autora analisou a presença de artigos que tratam de linguística no periódico Ciência da Informação no período compreendido entre 1972 e 1998.

Em Linguística, o termo é um dos objetos da terminologia. Krieger e Finatto (2004) analisam diversas definições de termo de modo a diferenciá-lo de uma unidade lexical qualquer. Dentre as definições, duas se destacam para a representação do conhecimento:

O termo caracteriza-se no sentido de que para uma noção dada, há, teoricamente, uma única denominação. Esta característica do termo se funda sobre um postulado da terminologia: o da relação de univocidade entre denominação (significante) e noção (significado, relação do tipo reflexiva) (RONDEAU, 1984, p. 19).

Um termo é uma unidade linguística que designa um conceito, um objeto ou um processo. O termo é a unidade de designação de elementos do universo percebido ou concebido. Ele raramente se confunde com a palavra ortográfica (GOUADEC, 1990, p.3).

Das duas definições selecionadas fica claro na terminologia que um termo se caracteriza como tal se for possível defini-lo. Wüster (apud DESMET, 1990, p. 6) lança o

conceito de unidade terminológica como "símbolo convencional que representa uma noção definida num certo domínio fundador". A partir dessa perspectiva clássica de definição em terminologia, Lara (2004) se alinha à esta pesquisa ao argumentar que a definição terminológica "é classificadora, hierarquizante, estruturante; relaciona-se à definição da coisa, ao contrário da definição lexicográfica que se relaciona à palavra e é feita pela identificação de traços semânticos que caracterizam o significado".

O estudo do termo envolve a compreensão de diversas variáveis, entre elas a identificação do usuário de uma terminologia e o contexto sociolinguístico em que estão inseridos. Faulstich (1995, p. 2) entende a socioterminologia "como disciplina descritiva, estuda o termo sob a perspectiva linguística na interação social" Ainda segundo a autora, são obras lexicográficas e terminológicas: os dicionários, os vocabulários, os léxicos, os glossários, as nomenclaturas, os bancos de terminologia e os tesouros.

As noções de termo no escopo da socioterminologia são essenciais na elaboração de vários sistemas de representação do conhecimento, em especial os tesouros e as ontologias. Para Gomes e Campos (2004), os termos são a base para o intercâmbio de informações, especialmente na Web. Para as autoras, o termo é a materialização do conceito, e se constitui pelo par conceito-denominação, ao passo que "se adotarmos a palavra denominação, em vez de termo, teremos mais facilidade em identificar a expressão linguística que designa um conceito, ou seja, para nomearmos um objeto utilizaremos tantas palavras quantas forem necessárias para tal."

Os tesouros tiveram origem em 1852, quando Peter Mark Roget publicou o *Thesaurus of English Words and Phrases*, que era uma coleção de termos organizada segundo as ideias que expressavam, ao contrário dos dicionários, que organizavam termos em ordem alfabética com as suas respectivas definições. Dodebei (2002, p. 64) afirma que "o objetivo de tal estrutura era o de encontrar as palavras pelas quais as ideias pudessem ser mais bem expressas em textos". Ainda segundo Dodebei (2002, p. 66), a partir de 1940, os tesouros começaram a ser usados no processo de recuperação da informação, como sendo "um instrumento capaz de transportar conceitos e suas relações mútuas, tal como expressos na linguagem dos documentos, em uma linguagem regular, com o controle de sinônimos e estruturas sintáticas simplificadas."

O desenvolvimento e a expansão dos sistemas automatizados de informação geraram a necessidade de desenvolver sistemas de classificação mais adequados à recuperação da informação. Os tesouros suprimiram essa necessidade por um tempo, porém, com a crescente

publicação de conteúdos na web, as ontologias começam a ser pensadas como um instrumento para lidar com esse volume de informações.

Almeida e Bax (2003) ao sistematizarem as principais contribuições para as ontologias, expõe que o termo ontologia tem origem na junção de duas palavras do grego “*ontos*”, ser, e “*logos*”, palavra. Em filosofia, é empregado com objetivo de “distinguir o ser como tal”. Para a organização da informação, as ontologias possuem diversas definições. Para este trabalho, será adotada a definição de Borst (1997, p. 12) e a análise feita por Almeida e Bax:

“Uma ontologia é uma especificação formal e explícita de uma conceitualização compartilhada”. Nessa definição, “formal” significa legível para computadores; “especificação explícita” diz respeito a conceitos, propriedades, relações, funções, restrições, axiomas, explicitamente definidos; “compartilhado” quer dizer conhecimento consensual; e “conceitualização” diz respeito a um modelo abstrato de algum fenômeno do mundo real (ALMEIDA; BAX, 2003, p. 9).

As ontologias podem servir a diversos propósitos e podem ser elaboradas seguindo várias metodologias. Biolchini (2001) ao tratar especificamente da construção de uma ontologia para a área médica propõe uma metodologia de organização segundo as interações de conceitos “definidas a partir das unidades terminológicas propriamente ditas, contidas nas bases de conhecimento médico”. Com isso, as ontologias visam a redução das diferenças na representação de conceitos, aproximando os sistemas de informação da realidade dos usuários.

Na web, as ontologias se desenvolvem para minimizar as dificuldades de recuperação de informação pelos sites de busca diante dos vários obstáculos que a busca utilizando linguagem natural apresenta. Pickler (2007) apresenta a polissemia de algumas palavras (por exemplo, a palavra manga que pode ser parte de uma camisa ou uma fruta) e a ausência de um domínio semântico definido para a orientação da busca.

No sentido de avançar na recuperação de informação na web utilizando ontologias, alguns buscadores foram lançados. Um deles é o Swoogle (c2007), que se destina a buscar um termo em uma ontologia a partir de uma ontologia, de um documento, de um termo e entre ontologias. Um outro buscador não se destina à busca em ontologias, mas estrutura a recuperação da informação com base em uma ontologia, o Hakia (c2010). O site agrupa a informação buscada em quadros que indicam a fonte da informação recuperada e a forma como se apresenta a informação: **web em geral, credibilidade** (informação recuperada em sites .org e .gov), **blogs, notícias, Pubmed, Twitter, imagens, vídeos**.

A análise dos diversos sistemas de classificação ao longo da história, mostra que todos partem de dois conceitos fundamentais: os documentos e os usuários, que Souza (2000) diz

em outras palavras “O QUE em termos de informação e PARA QUEM”. A partir disso, se faz necessário a abordagem do contexto de organização da informação na Web atualmente.

## 2.2 ORGANIZAÇÃO E RECUPERAÇÃO DA INFORMAÇÃO NA WEB

A Internet transformou as relações humanas e o acesso à informação. Muito antes da Internet chegar ao grande público, Vannevar Bush (1945) propõe um novo olhar sobre a dinâmica da produção científica considerando o uso de computadores para a recuperação da informação, salientando que a construção da ciência parte de observação e da congregação de diversas experiências trocadas ao longo dos anos e que seria fundamental o aprimoramento constante de tecnologias para a recuperação dessas informações. Esse momento foi marcado pelos microfiches como principal tecnologia de informação.

Na década de 70, a Internet começa a se difundir pelo mundo e, com isso, iniciam-se os estudos para a construção de SRIs on-line. Houve um aumento significativo de SRIs, mas podemos destacar o surgimento dos provedores de bases de dados que oferecem serviços desde o armazenamento de dados até uma equipe de especialistas para atender às necessidades de busca e recuperação da informação dos usuários finais e as instituições mantenedoras da base (HALL, 2010). Identifica-se a partir disso, a valorização dos interesses comerciais acerca da informação.

Na década de 80, passam a ser desenvolvidos os primeiros sistemas de acesso remoto a servidores de bases de dados, eliminando, principalmente, as barreiras geográficas de acesso à informação. Eram utilizados especialmente na área acadêmica e científica e a comunicação on-line era feita utilizando-se de tecnologias de emulação de terminal (Telnet) e do serviço Finger (TEIXEIRA; SCHIEL, 1997).

Na década de 90, a informação torna-se globalizada com o surgimento da World Wide Web (www ou w3), um projeto que permite acesso ao universo de informação on-line através da interligação de documentos através de links (BERNERS-LEE; CAILLIAU, 1992), utilizando uma interface bastante agradável ao usuário. Souza (2000) afirma que “A Internet pode ser definida como a Rede das Redes ou como uma Rede Interativa de Redes, agregando serviços e recursos interativos, sem que possa ser fisicamente visualizada ou localizada em termos de espaço e tempo”. Isso indica que a Internet, enquanto um sistema de recuperação da informação possui características singulares em relação às bases de dados, por exemplo.

A partir do surgimento da web, a informação não estava mais concentrada exclusivamente a um espaço físico. Diversas instituições passaram a oferecer seus serviços na

Internet, como bases de dados textuais e indicadores financeiros. As bases de dados mantidas pela maioria das instituições, no entanto, ainda apresentavam um modelo referenciado, organizado em diretórios e muitas vezes restrito a um público específico para determinada informação.

A real contribuição da Internet para a informação está além do oferecimento de serviços e produtos. Esta afirmação, porém, não faz muito sentido sem apresentação da definição de informação. Pinheiro e Loureiro (1995) apontam que a Ciência da Informação ainda não apresenta uma definição consensual e unívoca para informação. No contexto apresentado até aqui, a definição escolhida é a de Nicholas Belkin (1978, p. 78): “informação é tudo o que for capaz de transformar uma estrutura”.

A chegada da Internet ao grande público através do acesso à rede pelos computadores pessoais deram início ao desenvolvimento de diversos sites especializados em busca e recuperação da informação. Em seu início, os sites de buscas eram organizados na forma de diretórios, onde as páginas eram classificadas visando o público em geral. Um exemplo é o Yahoo!, que classificava as páginas como educação, esporte, entretenimento, viagens, compras (CENDÓN, 2001).

Os sites de busca na Internet construídos a partir dos anos 90 trazem consigo também um método de buscar informação que se adequa a quase todo tipo de usuário do site, por utilizar a linguagem natural para a recuperação da informação. Segundo Hall (2010), dos anos 60 até meados dos anos 80, era necessário o conhecimento de operadores booleanos para executar as buscas e também o conhecimento dos termos utilizados na indexação.

Os sites de busca na web não foram os únicos a transformar o modo como se busca e transmite informação. Outras três ferramentas estabeleceram uma nova dinâmica no acesso e comunicação da informação:

- E-mail – redução para a expressão *eletronic mail* é um método para enviar e receber mensagens seja utilizando sistemas baseados no protocolo SMTP que utiliza a Internet ou para sistemas conhecidos como intranets, em geral gerenciados por empresas e organizações. Este sistema de informação é amplamente usado, tanto para a comunicação formal e informal entre as pessoas.
- Salas de bate-papo ou *chats* – consistem de um espaço destinado ao contato de usuários da Internet por meio do envio de mensagens uns aos outros em tempo real. A ideia básica é uma simulação de conversa que pode ser em grupo ou no modo privado, para apenas uma pessoa. As salas de bate-papo são oferecidas,

em geral, por um site e são organizadas segundo o interesse dos usuários (por idade, por região, cinema, etc.)

- Programas para troca de mensagens instantâneas – permitem o envio de mensagens para uma lista de contatos do usuário. Em seu início, a troca era apenas de mensagens escritas, porém, com o passar do tempo, outros recursos foram acrescentados como conversa por voz e vídeo. Pode-se dizer que o marco inicial se deu em 1996, com o lançamento do ICQ. Nos anos que se seguiram dezenas de outros surgiram com destaque para o AIM e o MSN Messenger.

As tecnologias de comunicação on-line configuram um novo fluxo de conhecimento. Barreto (1998) analisa a questão considerando que na Internet, a geração do conhecimento por parte do indivíduo está integrada ao seu espaço de convivência com outros indivíduos, a partir da identificação dos tipos de comunicação do conhecimento, que seriam oral, escrita tipográfica e eletrônica. Para esses três tipos de comunicação são apresentadas oito características: **fundamental, tempo de transferência, espaço de transferência, armazenamento, relação de audiência, estrutura da informação, interação com o receptor e conectividade (acesso).**

A nova ordem de organização e disseminação da informação foi instaurada pelo acesso à rede. Até esse momento a produção e transferência eram centralizadas e referenciais, ou seja, a concepção e a transferência de informação deveriam ser “intencional, arbitrária e contingente ao atingir o seu destino: criar conhecimento no indivíduo e em sua realidade” (BARRETO, 1999). A web possibilita não apenas a comunicação de um indivíduo ao seu receptor, mas de um indivíduo para muitos receptores.

Embora a Internet tenha revolucionado a forma como os indivíduos se comunicam, passado algum tempo a comunicação de um indivíduo para vários outros indivíduos já não basta para suprir as necessidades de informação na sociedade contemporânea. Chegando aos anos 2000, é preciso que muitos indivíduos se comuniquem com outros muitos indivíduos através da colaboração.

### 2.2.1 *Web 2.0* e a era da colaboração

Em meados dos anos 2000, um novo movimento surge na área de tecnologia da informação. A chamada computação nas nuvens (*cloud computing*) é uma metáfora utilizada para indicar que a computação não está mais confinada a uma estrutura física. Os serviços de

tecnologia da informação são fornecidos como serviços na web, eliminando a necessidade de instalações de softwares e as posteriores atualizações, os quais podem não estar disponíveis para determinadas máquinas, pois as mesmas se tornam obsoletas rapidamente (SOUSA, et. al., 2009).

A chamada *Web 2.0* (O'REILLY, 2004) surge para oferecer uma nova gama de produtos e serviços de informação baseados em conteúdos desenvolvidos através da colaboração e interação, favorecendo a troca e a organização da informação. Mais precisamente, a *Web 2.0* utiliza a Internet como uma “plataforma” envolvendo wikis, aplicações baseadas em folksonomia, redes sociais e Tecnologia da Informação. O'Reilly (2006) estabelece ainda regras para interação com a *Web 2.0*, as quais ajudam a compreender esse universo:

1. Não trate o software como um artefato, mas como um processo de comprometimento com seus usuários - **beta perpétuo**.
2. Abra seus dados e serviços para que sejam reutilizados por outros. Reutilize dados e serviços de outros sempre que possível. – **Pequenas peças frouxamente unidas**.
3. Não pense em aplicativos que estão no cliente ou servidor, mas desenvolva aplicativos que estão no espaço entre eles. – **Software acima de um único dispositivo**.
4. Lembre-se de que em um ambiente de rede, APIs abertas e protocolos padrões vencem, mas isso não significa que a ideia de vantagem competitiva vá embora. – **Lei da Conservação de Lucros, de Clayton Christensen**.
5. A mais importante entre as futuras fontes de fechamento e vantagem competitiva serão os dados, seja através do aumento do retorno sobre dados gerados pelo usuário, sendo dono de um nome ou através de formatos de arquivo proprietários. – **Os dados são um novo “Intel Inside”**.

A aparente simplicidade do conceito *Web 2.0* é decorrente de um processo de modificação longo e complexo das demandas por informação, estimulado, principalmente, pela popularização das redes sociais, que segundo Póvoa (2006), atuam como catalisadores de comunidades.

Gracioso (2007, Não paginado) observa que o cenário atual “permite ao usuário da informação ser ao mesmo tempo o seu autor, editor e classificador”, o que reforça a necessidade do profissional da informação em conhecer as particularidades dessa nova forma de tratar a informação, pois o dinamismo e a flexibilidade dos conteúdos de informação tornam as ideias tradicionais de autoridade, instituição um pouco abstratas.

A nova realidade imposta pela *Web 2.0* modificou a compreensão de autoria e autoridade<sup>4</sup> diante da valorização da comunicação entre os usuários como meio para construção coletiva de inteligência. Miranda, Simeão e Mueller afirmam que

A autoria é uma instituição em crise. Analisando-a em perspectiva, é possível afirmar que está sujeita a interpretações diversas e até divergentes em níveis de conceito e de práxis, em uma abordagem multidisciplinar. Sempre esteve atrelada às noções de cultura e ciência em que se desenvolveu, mas é na civilização ocidental, sobretudo com o advento do liberalismo, do capitalismo e do individualismo triunfantes, que a autoria ganhou foros de direito e as áureas da sociedade. (MIRANDA; SIMEÃO; MUELLER, 2007, p. 35).

A “crise” da autoria ocorre a partir do uso de ferramentas que permitem ao autor realizar a escrita e a pesquisa simultaneamente, que, ainda segundo Miranda, Simeão e Mueller (2007) tornam o processo de produção de informação “mais inteligente” diante do aumento da coautoria pela redução das barreiras geográficas com o uso da rede. Nesse momento destaca-se o conceito de inteligência coletiva proposto por Lévy:

A *coordenação das inteligências em tempo real* provoca a intervenção de agenciamentos de comunicação que, além de certo limiar quantitativo, só podem basear-se nas tecnologias digitais da informação. Os novos sistemas de comunicação deveriam oferecer aos membros de uma comunidade os meios de coordenar suas interações no mesmo universo virtual de conhecimentos. Não seria tanto o caso de modelar o mundo físico comum, mas de permitir aos membros de coletivos mal-situados interagir em uma paisagem móvel de *significações*. Acontecimentos, decisões, ações e pessoas estariam *situados* nos mapas dinâmicos de um contexto comum e transformariam continuamente o universo virtual em que adquirem sentido. Nessa perspectiva, o *ciberspaço* tornar-se-ia o espaço móvel das interações entre conhecimentos e conhecedores de coletivos inteligentes e desterritorializados (LÉVY, 1998, p. 29).

O que se destaca aqui é que a *Web 2.0* atua na convergência de saberes, “eliminando os limites entre os meios, tornando-os solidários em termos operacionais, e erodindo as tradicionais relações que mantinham entre si e com seus usuários” (MIRANDA, 2000, p. 79). A dinamização dos conteúdos oferecidos como produto de uma interação social representada pelos instrumentos de compartilhamento de informação (*tags, social tagging, social bookmarking*, etc.) contidos, especialmente, nas redes sociais por indivíduos que são autores, pesquisadores e gestores da informação. Por fim, Miranda (2000, p. 81) afirma que a “Sociedade da Informação desenvolve-se através da operação de conteúdos sobre a infraestrutura de conectividade”.

---

<sup>4</sup> Em um sentido mais geral, **autoria** é a fonte de uma ideia ou trabalho criativo em qualquer forma, com referência ao seu criador ou originador, por exemplo, o compositor de uma obra musical ou escritor de um romance. Na comunidade acadêmica, a **autoridade** é indicada por credenciais, trabalhos já publicados sobre o assunto, afiliação institucional, prêmios, selo, revisões, padrões de citação, etc. (REITZ, J. M., c2010, Não paginado).

Em Ciência da Informação, a recuperação da informação utiliza, tradicionalmente, a redução dos conceitos para obter o maior número de documentos possíveis e pertinentes ao universo a ser pesquisado. Essa redução é possível pela atividade de indexação, que consiste, basicamente, na atribuição de palavras-chave ao documento, pela extração de palavras do texto (uso da linguagem natural) ou a tradução de termos considerados relevantes em uma linguagem documentária (DODEBEI, 2002).

A característica não linear do hipertexto não possibilita o uso de um produto de indexação – como os tesauros – para sua representação. Na *Web 1.0*, eram usadas as taxonomias para organização do conteúdo e sua recuperação. As taxonomias na web são caracterizadas por diretórios, onde os sites são agrupados por assuntos ou classes e subclasses (AQUINO, 2007) e são utilizados principalmente por sites de buscas (ex. Google, Yahoo!). A escolha das classes ficava a cargo dos desenvolvedores, centralizada na figura do programador.

Em 1989, Vickery e Vickery reformulam a ideia clássica de comunicação informativa acrescentando a influência das fontes, dos canais e outros fatores sociais que tornam esse processo mais complexo do que a noção básica de emissor-mensagem-receptor. Para o autor, o processo de comunicação da informação deve ser entendido como “transferência de significado”, pois um conceito é integrado a uma estrutura de conhecimento pessoal, e expresso em palavras e outros signos linguísticos.

Na *Web 2.0*, a recuperação e a disseminação da informação baseiam-se na comunicação entre os usuários, rompendo a barreira entre os desenvolvedores e usuários. A partir dessa necessidade, emergiu na rede uma nova forma de representar e classificar o conteúdo: a utilização da folksonomia, que seria uma “classificação de informação feita pelos usuários através de meta-dados – as *tags*” (MEIRELLES; MOURA, 2007, p. 16). A classificação da informação através das *tags* torna-se tão relevante para a Internet na atualidade que instaura a era da *Web 3.0*, que será explorada a seguir.

### 2.2.2 Web 3.0: folksonomia e *tags* para a classificação de conteúdo on-line

Etiquetar é uma das ações relacionadas à classificação mais incorporada ao dia a dia. Está nos rótulos dos alimentos, nas roupas e em uma variedade de objetos de uso pessoal ou coletivo e contém informações importantes sobre determinado objeto. Na “sociedade 2.0” etiquetar torna-se ação de disseminação da informação, utilizando-se da folksonomia para a atribuição de termos para representação do conhecimento.

A folksonomia é a tradução do neologismo em inglês criado pela união das palavras *folk* (povo) com *taxonomy* (taxonomia). A folksonomia, no entanto, aponta para direção diferente das taxonomias. Na realidade a ideia é estabelecer relação de antítese às taxonomias, pois, segundo Thomas Van der Wal (2005) - o autor do termo - a folksonomia é o resultado do *tagging* livre e individual de informações e objetos para posterior recuperação pelo próprio. Apesar desse ter sido o propósito inicial da folksonomia, o próprio autor reconhece que outros usos do *tagging* visando a colaboração podem ser desenvolvidos a partir da dinâmica inicial do uso das *tags*.

Algumas pesquisas já identificam o alcance da folksonomia no que se refere à representação da informação. Jakobsen (2007) relata que existem as folksonomias estreitas e as folksonomias amplas<sup>5</sup>. Nas folksonomias estreitas, usuários etiquetam objetos de informação que muitas vezes são de sua autoria, assim, esses objetos recebem poucas *tags*, mas estas *tags* podem ter muitos objetos ligados à elas. Nas folksonomias amplas, ocorre o inverso, usuários classificam objetos de informação que não os pertencem, e estes objetos recebem muitas *tags*.

Em algumas redes sociais da Internet, a ideia inicial do uso da *tag* para recuperação da informação de forma individualizada parece ter se transformado para atender aos princípios básicos da *Web 2.0*: a socialização e a colaboração, ou seja, disponibilizar informações a um grupo de pessoas/instituições de forma a estabelecer comunicação. Desse pressuposto das redes sociais, surgiram as *tags* sociais (*social tagging*) como uma forma de classificar objetos e informações de forma colaborativa.

As *tags* ou etiquetas são definidas no ambiente virtual de acordo com diferentes perspectivas acerca da sua estruturação. Derntl et al. (2010) adotam como consenso geral a definição proposta por Maier e Thalmann (2008): uma *tag* pode ser definida como um composto triplo, que seriam o objeto que é marcado, a palavra-chave usada, e o usuário que atribui a *tag* para o objeto.

Uma outra definição apresenta uma perspectiva diferente para a construção das *tags*. Müller-Prove (2008) afirma que a *tag* é um “par composto de uma etiqueta e um conjunto de objetos (recursos) marcados com esta etiqueta”. Esta definição ajuda na compreensão do porque alguns sistemas que usam *tags* para classificação tentam substituir a iniciativa espontânea do usuário no processo de classificação sugerindo uma etiqueta ou impedindo a repetição das etiquetas para um mesmo objeto.

---

<sup>5</sup> Tradução livre de narrow folksonomies e broad folksonomies, respectivamente.

As *tags* possibilitam ao usuário relacionar qualquer palavra à informação selecionada por ele quando esse serviço é oferecido, de forma descentralizada, sem ser utilizado um vocabulário controlado. Ainda segundo Aquino:

frente ao que poderia parecer uma certa desorganização na web, porém, além de viabilizar a escrita coletiva online algumas das novas ferramentas, que surgem com o momento da web 2.0, agregam uma nova forma de representação, organização e recuperação de informações que vai de encontro aos postulados do autor. Estas ferramentas baseiam-se na folksonomia que é a possibilidade que o indivíduo tem de relacionar qualquer palavra à um dado que armazena nesses serviços. Esta relação é estabelecida através de uma *tag* adicionada ao documento. Assim, a informação fica armazenada e pode ser recuperada através da *tag* que o próprio usuário criou, e não mais através de um vocabulário controlado, o qual muitas vezes é desconhecido de quem faz a busca (AQUINO, 2007, p. 9).

O uso das *tags* para a classificação dos conteúdos nas redes sociais encontra problemas de categorização semelhantes aos esquemas de classificação bibliográfica. Lakoff (1987) resgata a pergunta “porque chamamos coisas diferentes por um mesmo nome?” do filósofo J. L. Austin, para ressaltar a influência da polissemia na escolha de palavras para compor uma categoria. Sendo assim, apesar das *tags* serem a forma mais adequada de categorização para a Web até o momento, não são capazes de transpor esse problema semântico com eficácia.

A forma de amenizar os problemas semânticos na classificação usando as *tags* é a estruturação do sistema de modo a permitir que o sistema ou usuários do sistema recomendem *tags* ou objetos classificados com determinada *tag*. Isso configura as *tags* também como sistemas de recomendação – SRs, que segundo Santini (2010, p. 83) são estruturados para além da solução técnica de organização e filtragem do excesso de informação. Estão inseridos no contexto social de “flexibilidade e “democratização” dos meios de produção, circulação e uso de bens culturais na Internet”.

Pensar em uma classificação que possa organizar conteúdos dinâmicos e com vários sentidos não é uma necessidade exclusiva da folksonomia aplicada às redes sociais na internet. Kumar (1981), ao discutir as ideias de categoria de Ranganathan, reconheceu que um termo pode não ser adequado para classificar algo se não forem determinadas as facetas de uma ideia relacionada a uma categoria fundamental, o que, em outras palavras significa considerar o contexto como elemento a ser observado para a escolha de termos para categorizar uma informação.

A disseminação da folksonomia através do sistema das *tags* caracteriza um dos fenômenos mais importantes da *Web 2.0*, por seu alcance e aceitação dos usuários da grande rede. Tal alcance talvez não fosse conquistado sem a adesão desses usuários às redes sociais.

Nocera e Ursino (2011), consideram a folksonomia e as redes sociais como fenômenos que não podem ser dissociados.

### 2.3 A REDE SOCIAL *TWITTER*: ORDEM E CAOS

As redes sociais existem antes mesmo da Internet. Em sociologia, as teorias de redes sociais iniciam-se no século XIX a partir da compreensão de fato social de Émile Durkheim (2007<sup>6</sup>), em que os valores individuais não são socialmente representativos sem a interação desses valores com os valores de outros indivíduos. Durkheim (2007) considera a totalidade como ideia central de constituição da sociedade, embora essa totalidade seja abordada como um agregado de diversas partes distintas.

Marcel Mauss<sup>7</sup> (2003) aprofunda o pensamento de Durkheim, eliminando a noção de dualidade individual/coletivo. Para Mauss, as diferenças individuais contêm os elementos que irão constituir a totalidade, ou seja, cada indivíduo ou grupo de indivíduos contribuem igualmente para a sociedade. Mauss reconhece ainda a existência de uma obrigação social – a dádiva – que representa as trocas de bens aparentemente voluntárias e gratuitas entre os indivíduos, que, no entanto, são obrigatórias e com interesse no retorno (dar, receber, retribuir).

Norberto Elias (1994) atribui a característica de movimento nas relações humanas que constituem as redes sociais. Segundo Elias (1994) a dinâmica das relações humanas não deve ser reduzida nem às liberdades individuais e nem mesmo na expectativa de retribuição nas trocas de bens, estabelecendo que “efetivamente, cresce o indivíduo, partindo de uma rede de pessoas que existiam antes dele para uma rede que ele ajuda a formar”. Desta maneira, as redes se estruturam a partir de ligações entre indivíduos e também com ruptura dessas ligações.

Em Ciência da Informação, Marteleto (2001, p. 72) entende as redes sociais como um “conjunto de participantes autônomos, unindo ideias e recursos em torno de valores e interesses compartilhados”. As redes, portanto podem envolver uma comunidade local, pesquisadores de uma área científica, etc.

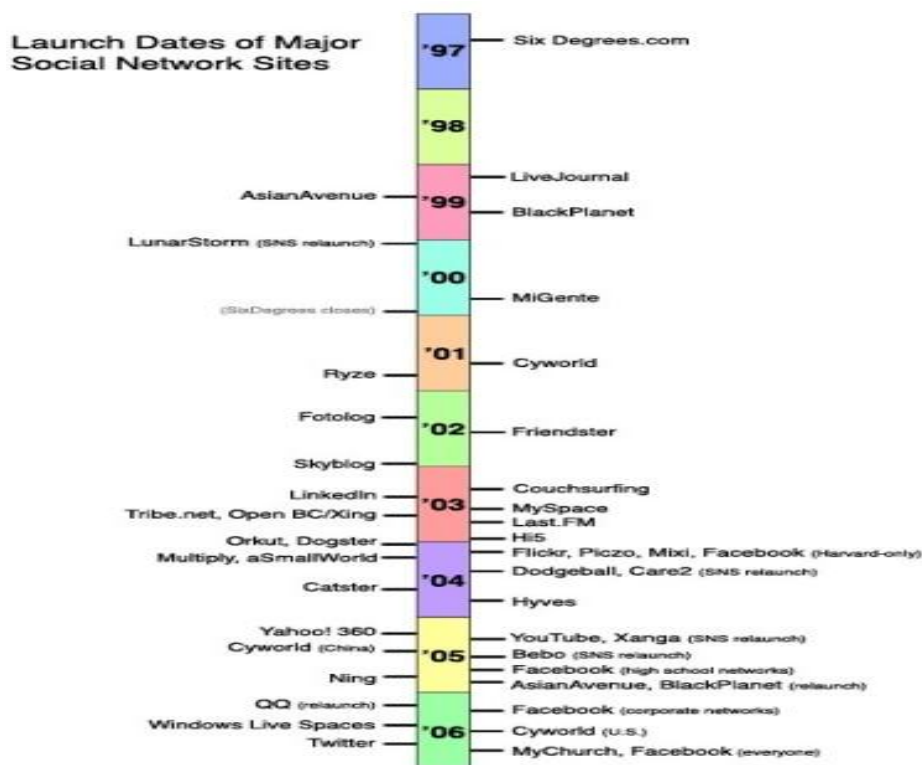
Reproduzindo no mundo virtual a necessidade de união e compartilhamento de interesses que já existia no mundo real, surgem, no final dos anos 90, as redes sociais na Internet. Elas são sistemas para interação dos usuários, numa simulação da forma como os

---

<sup>6</sup> Obra publicada originalmente em 1893.

<sup>7</sup> Obra publicada originalmente em 1924.

relacionamentos acontecem na vida real, ou seja, o contato entre os usuários pode estabelecer redes segundo o grau de amizade, interesses profissionais e acadêmicos, etc. A Figura 1, apresenta a cronologia resumida das principais redes sociais na Internet:



**Figura 1:** Datas de lançamento dos principais sites de redes sociais

**Fonte:** Disponível em: <http://socialmediagraphics.posterous.com/>. Acesso em: 11 nov. 2011.

O marco inicial<sup>8</sup> das redes sociais na Internet na forma como as conhecemos hoje foi no lançamento da rede *SixDegrees* em 1997. Nela, os usuários podem listar amigos, familiares, conhecidos que participavam da rede ou não. Os usuários podiam enviar mensagens e publicar no quadro de avisos de pessoas em primeiro, segundo e terceiro grau e visualizar as conexões deles com qualquer outro usuário do site. As atividades do site foram interrompidas em 2001 e retomadas recentemente (SIXDEGREES, c2010), mas estão disponíveis apenas para quem já era usuário do sistema antes da interrupção.

O modelo utilizado pela rede social é baseado na teoria dos seis graus de separação que apareceu pela primeira vez num conto do autor húngaro Frigyes Karinthy (1887 –1938). A teoria dos seis graus de separação se desenvolve nos Estados Unidos da América na década de 60. A teoria estabelece que são necessários, no máximo, seis laços de amizade para que quaisquer duas pessoas estejam ligadas (MILGRAM, 1967). A teoria foi criada com base no

<sup>8</sup> <http://blog.afridesign.com/2010/09/sixdegrees-com-social-networking-in-its-infancy/>

estudo que identificou os números de laços de conhecimento pessoal entre duas pessoas, com base na análise do envio de cartas entre grupos de pessoas. Essa teoria influenciou o modo como as redes sociais se estruturaram para que um usuário mantenha sua rede de amizades.

No Brasil, a rede social de maior destaque até o momento é o *Orkut*, lançada em 2004 pela empresa Google, sendo o criador da rede o funcionário Orkut Büyükkökten. Segundo dados de outubro de 2011, 50,60% dos usuários<sup>9</sup> desta rede social são brasileiros (ORKUT, c2011). No momento de seu lançamento, o *Orkut* caracterizava-se por um extenso perfil de usuários onde são solicitadas informações sobre diversos aspectos de sua vida. Cada usuário monta sua rede de amigos enviando convites por e-mail e pelo próprio *Orkut* (adicionar aos amigos) e pode interagir com eles deixando recados ou enviando mensagens. Outra opção para interação é a participação em fórum de discussão numa comunidade de interesse. De 2004 a 2011 foram muitas as atualizações, grande parte delas para se aproximar de outra grande rede social, o *Facebook*.

Em 2004, três estudantes da Harvard University criaram uma rede social restrita aos alunos da universidade. Em pouco tempo, a rede foi expandida para o Massachusetts Institute of Technology, Boston University, Boston College e as escolas da Ivy League<sup>10</sup>. Dois anos depois, o *Facebook* foi aberta também para estudantes secundaristas e algumas empresas, e posteriormente foi aberta para o público em geral. A interação entre os usuários se efetiva por meio da publicação de atualizações de status, divulgação de vídeos, fotos e páginas da Internet.

As publicações do *Facebook* podem ser vistas apenas pela rede de amigos ou para todo o público da rede, de acordo com a escolha do usuário. A rede oferece também diversos outros recursos para a interação, com destaque para os jogos, nos quais para quem joga depende da ajuda/colaboração para atingir seus objetivos. Como já considerado, a colaboração é a razão de existir da *Web 2.0*.

A colaboração é exatamente a principal motivação para o uso das redes sociais. Diversas outras redes sociais on-line surgiram para a colaboração entre os usuários de acordo com os mais variados objetivos. Na Figura 2, é possível visualizar o panorama atual das redes sociais segundo a finalidade:

---

<sup>9</sup> <http://www.orkut.com.br/Main#MembersAll>. Para acessar as estatísticas oficiais do Orkut, é necessário *login*.

<sup>10</sup> <http://pt.wikipedia.org/wiki/Facebook>



**Figura 2:** Tipologia das mídias sociais na Internet

**Fonte:** Disponível em: <http://www.fredcavazza.net/2008/06/09/social-media-landscape/>. Acesso em: 12 nov. 2011.

A breve gênese das redes sociais na Internet apresentada indicou o crescimento da adesão desse canal pelos usuários na Internet para a interação dos usuários através da troca de informações. Nos exemplos explorados, as redes sociais oferecem diversas maneiras para os usuários trocarem informações. A rede social de interesse da pesquisa, o *Twitter*, optou por um modelo completamente diferente de interação, o que apresentaremos a seguir.

O *Twitter* foi criado em 2006 por Jack Dorsey. Pode ser definido como “uma rede social e servidor para *microblogging* que permite aos usuários enviar e receber atualizações pessoais de outros contatos (em textos de até 140 caracteres, conhecidos como "*tweets*"), por meio do *website* do serviço, por *SMS* e por softwares específicos de gerenciamento” (WIKIPEDIA, c2011). Consiste, portanto, em um canal estruturado de forma a disseminação de diversos tipos de informação. Huberman, Romero e Wu ao discutirem o papel dessa rede social para diversos setores da sociedade explicam a forma como as mensagens são disseminadas no site:

Os usuários do *Twitter* podem postar atualizações diretas e indiretas. As diretas são usadas quando um usuário escreve a atualização para uma pessoa específica e é sinalizada por um símbolo "@" ao lado de nome da pessoa, enquanto que atualizações indiretas são usadas quando a atualização é para qualquer um que se dispõe a lê-la. Apesar de atualizações diretas serem usadas para se comunicar diretamente com uma pessoa específica, elas são públicas e qualquer pessoa pode vê-las. Muitas vezes, dois ou mais usuários terão conversas postando atualizações dirigido ao outro. Cerca de 25,4% de todas as mensagens são dirigidas, o que mostra que esse recurso é amplamente utilizado entre os usuários de *Twitter* (HUBERMAN; ROMERO; WU, 2008, p.3)

Essa rede social teve como propósito inicial, funcionar como um blog expresso, ou seja, divulgar as principais atividades de uma entidade diariamente, usando como motivação a pergunta "o que está acontecendo?". A experiência de uso, no entanto, revelou potencial para a divulgação de todo tipo de mensagem, incluindo as notícias, campanhas publicitárias, e informes de interesse de uma comunidade.

Ao contrário das redes sociais já abordadas neste trabalho, o *Twitter* não solicita que o usuário preencha um extenso perfil. São solicitadas informações básicas e uma descrição pessoal de no máximo 160 caracteres, o que contrasta com o perfil do Orkut, por exemplo, que solicita informações pessoais, profissionais e até mesmo sentimentais.

No Quadro 5 é apresentado um resumo da terminologia do *Twitter* (MOLLET; MORAN; DUNLEAVY, 2011):

<b>Termo</b>	<b>Definição</b>
Seguir ( <i>Follow</i> )	Seguir um usuário significa que todas as mensagens do usuário seguido aparecerão em seu <i>feed</i> . Para seguir um usuário é necessário clicar no nome de usuário e em seu perfil clicar no botão verde <i>Follow</i> que aparece ao lado direito da tela.
Lista de quem seguir ( <i>Who to follow list</i> )	Lista de sugestões do <i>Twitter</i> contendo pessoas ou organizações que podem interessar a um usuário, com base em pontos de similaridade com o perfil.
Para de seguir ( <i>Unfollow</i> )	Para parar de seguir alguém, basta ao usuário ir à sua lista de seguidos, selecionar o usuário a que deseja parar de seguir e posicionar o mouse sobre o botão <i>Follow</i> até que mude para <i>Unfollow</i> e em seguida clicar.
Bloquear ( <i>Block</i> )	É um recurso para bloquear definitivamente um usuário de visualizar mensagens ou contatar um perfil.
Retweetar ou RT ( <i>Retweet or RT</i> )	Recurso para retransmitir aos seguidores uma mensagem de outro usuário. Quem recebe essa mensagem, visualiza uma seta que indica que essa mensagem não foi escrita por quem a enviou. É como encaminhar uma mensagem de e-mail.
Responder ( <i>Reply</i> )	Para responder a uma mensagem de outra pessoa, é necessário passar o mouse sobre ele e selecionar a opção <i>reply</i> , que aparecerá na coluna <i>@mentions</i> .
@	Usado em mensagens quando se quer citar outro usuário na mensagem. Também é a primeira parte de cada nome de usuário do <i>Twitter</i> - por exemplo <i>@moura_amanda</i> .
Menções	Ao verificar a coluna <i>@mentions</i> , um usuário pode ver quando outros

(Mentions)	usuários o mencionaram em alguma mensagem.
#	<i>Hashtag</i> - usada para categorizar mensagens. Algumas <i>hashtags</i> tornam-se populares figurando entre os tópicos mais cometados do <i>Twitter</i> . Ao clicar em qualquer um dos tópicos listados na página inicial, é mostrada uma lista de <i>tweets</i> contendo o tópico. A inclusão de <i>hashtags</i> populares numa mensagem pode atrair mais atenção para uma mensagem. <i>Hashtags</i> são também utilizadas como parte do “ <i>backchannel</i> ” de comunicação em torno de um evento, seja ele uma conferência, um programa de TV ou um evento global. um evento público pode compartilhar comentários, perguntas, etc durante a realização do evento ao atribuir uma <i>hashtag</i> exclusiva para essas mensagens.
Mensagem Direta ou DM ( <i>Direct Message or DM</i> )	Estas são mensagens privadas enviadas para outros usuários do <i>Twitter</i> . Para enviar, basta clicar no menu <i>Message</i> situado na parte superior da página inicial.
URLs reduzidas ( <i>Shortened URLs</i> )	Um endereço de web é, em geral, extenso. Como uma mensagem pode conter no máximo 140 caracteres, a utilização de sites de encurtamento de links como <i>bitly.com</i> e <i>tinyurl.com</i> são essenciais para fornecer links mais curtos para colar nos <i>tweets</i> .

**Quadro 5:** Terminologia do *Twitter*

Fonte: Adaptado de MOLLET; MORAN; DUNLEAVY (2011, p. 3)

O *Twitter* se diferenciou das redes sociais anteriores por não oferecer a possibilidade de adesão a uma comunidade de interesse. O foco central de uso é confluência mais livre de ideias, pois ao invés de adicionar um usuário à sua conta, o que se faz nessa rede é seguir um usuário com perfil público para receber suas mensagens em tempo real, ou seja, a cada segundo várias mensagens são divulgadas diariamente. A ausência de indicações do universo ao qual uma mensagem pode pertencer, instaura no sistema uma situação de caos.

Diante desse imenso volume de informação circulando pela rede, as questões acerca do caos fazem parte da discussão. Caos, utilizado com frequência como sinônimo de desordem, bagunça. Nas ciências, no entanto, as teorias do caos sugerem imprevisibilidade e complexidade (ARAÚJO, 1995). Em relação ao caos na *Web 2.0* e nas redes sociais da Internet, a definição de Rezende se faz pertinente:

O caos é um estado complexo caracterizado pela (aparente) imprevisibilidade de comportamento e por grande sensibilidade a pequenas mudanças das variáveis do sistema ou das condições iniciais. É observado tanto em sistemas muito simples, quanto em sistemas complexos. A condição essencial para um sistema apresentar estado caótico é ser não linear, isto é, apresentar uma resposta não proporcional ao estímulo. (REZENDE, 1992, p. 30)

Em meio à complexidade e à imprevisibilidade do caos, “A história da evolução humana evidencia uma busca pela regularidade: ordem das estações, sucessão de dias e noites, precisão (ainda que muitas vezes só aparente) do movimento de estrelas e planetas no céu etc.” (ARAÚJO, 1995). Por essa razão, este trabalho explora a ordem em meio ao caos numa rede social.

Pesquisas estabelecendo caos e complexidade às redes sociais não são de interesse exclusivo das ciências sociais. Alguns estudos interdisciplinares em física já relacionam os estudos em complexidade com as redes sociais na Internet. Grabowski, Kruszewska e Kosiński (2008) estudaram as características de cinco redes sociais na Internet:

- Grono – rede social para usuários localizados na Polônia que oferece fóruns de discussão, compartilhamento de fotos, busca de eventos em algumas cidades, além de oportunidades de negócios. Parte das funcionalidades do site estão disponíveis apenas a quem acessa a modalidade paga.
- MMORPGs – site para troca de informação para os praticantes de Jogo de Interpretação de Personagens Online e em Massa para Múltiplos Jogadores (*Massively multiplayer online role-playing game*, sigla MMORPG).
- LastFM – comunidade onde são trocadas informações e recomendações sobre música, de acordo com o gosto musical expresso no perfil de um usuário.
- XFire – é um software para troca de mensagens instantâneas para praticantes de MMORPG. O software permite a troca de informações entre um grupo de usuários interagindo em tempo real.
- Shelfari – é uma rede para “catalogação social” de livros. Os usuários montam suas estantes virtuais de livros que possuem ou que já leram e podem elaborar revisões, avaliações, classificações (através de *tags*) e ainda discutir um livro com outros usuários.

O estudo identificou que as redes de amizade em todos os sistemas possuem características similares, assim como o comportamento humano nesses sistemas, embora as redes estudadas não se destinem a um mesmo objetivo.

Em um estudo mais relacionado à história da ciência, Helbing et al. (2011) relaciona o pensamento complexo à dinâmica “técnico-socio-econômica” da formação de algumas redes sociais no processo de colaboração dentro delas. São relatados casos de sucesso, e são fornecidas diretrizes para a identificação de possibilidades e limitações para a previsão dessa dinâmica. Uma das propostas é o desenvolvimento de uma teoria que possa conectar o indivíduo, o comportamento micro-nível com o agregado, comportamento do sistema em nível macro, e o *feedback* do comportamento macro no micro-nível. Para os autores, o *feedback* pode contribuir para a previsão de modelos para sistemas complexos, em particular em situações de desequilíbrio.

As teorias e leis do caos podem ser aplicadas a diversas áreas do conhecimento, justamente por considerar o evento, ou seja, um fato novo como um elemento capaz de alterar uma estrutura ou uma trajetória. Prigogine (2002) ressalta ainda que inicialmente, o ideal das ciências da natureza era a busca pela certeza através de uma descrição determinista, ao contrário das ciências humanas que sempre se preocuparam com as noções de incerteza, escolha e risco. O autor observa ainda que:

É o modo de descrever o curso do tempo que distingue as duas culturas. Poder-se-ia mesmo pensar em distingui-las pela complexidade de seu objeto: a física ocupar-se-ia então dos chamados fenômenos simples, e as ciências humanas dos fenômenos complexos. Mas hoje em dia a diversidade entre fenômenos simples e complexos tem-se reduzido. Sabemos que as chamadas partículas elementares e os problemas da cosmologia correspondem a fenômenos extremamente complexos [...]. (PRIGOGINE, 2002, p. 13)

Ainda de acordo com Prigogine (2002, p. 12) o caos é consequência de fatores de instabilidade, considera que na Física o determinismo dominou a área por muito tempo, embora “a maior parte dos sistemas de interesse para a física, quer de mecânica clássica quer de mecânica quântica, é de sistemas instáveis. Neles, uma pequena perturbação amplifica-se e trajetórias inicialmente próximas divergem”. A partir disso, por analogia, o *Twitter* pode ser visto como um sistema instável, pois uma única mensagem (evento) pode modificar a direção de uma discussão (trajetória).

As ideias de instabilidade e mudança de trajetória não são as únicas que se relacionam com o caos nas redes. Wersig (1993) insere no contexto de organização do campo da Ciência da Informação as ideias de *atratores* e *atratores estranhos* ao afirmar que um dos passos para estruturar a realidade caótica seria descobrir seus “atratores estranhos e suas contradições ou relações, para então contrastar a estrutura interna dos problemas do campo com as estruturas gerais. Por fim, estratégias têm que ser desenvolvidas para lidar com problemas sob condições caóticas, usando os conceitos disponíveis, ou atratores, para organizá-los”.

Na física matemática, um *atrator* representa um conjunto de comportamentos que levam à evolução em um sistema dinâmico, independente do ponto de partida. Os três principais tipos de atratores são: “**atratores puntiformes**, [que] correspondem a sistemas que atingem equilíbrio estável; **atratores periódicos**, que correspondem a oscilações periódicas; e os assim chamados **atratores estranhos**, que correspondem a sistemas caóticos” (RIPELL; RECH, 2004, p. 6).

Ao considerar o *Twitter* como um sistema caótico, o estudo dos atratores estranhos fornece alguns indícios do comportamento flutuante e contínuo dos usuários do sistema em relação à publicação de mensagens na rede, pois assim como atratores estranhos os usuários

do Twitter convergem em direção a algum ponto embora não se possa afirmar onde se localiza este ponto de convergência com certeza. A descrição do comportamento dos atratores estranhos foi realizada por Edward Lorenz (1963).

O caos traz consigo as noções de imprevisibilidade e mudança sobre determinada ordem. A ordem, que é modificada várias vezes a cada situação de caos ocasionando o surgimento de novas ordens a cada novo evento. Schrödinger (1995) ao analisar as relações da física com a dinâmica da vida considera que o par ordem-desordem caminha junto, pois apesar do modelo estatístico da física indicar a desordem da matéria, em um nível microscópico o que se vê é a ordem das moléculas.

Braga (1995) traz a discussão do caos e da complexidade para a Ciência da Informação. A autora explicita que a informação se revela como uma probabilidade, pois não há como se chegar a uma quantidade exata de toda a informação que circula no mundo. A partir dessa aceção, entende-se o porquê da Ciência da Informação estar sempre acompanhada de expressões como explosão e caos para caracterizar eventos marcantes de sua história. Para a presente discussão o questionamento a seguir se faz pertinente:

Caos é desordem, é o comportamento de pequenas alterações levando a grandes mudanças posteriores, é instabilidade persistente, é imprevisibilidade. Caos parece ser o princípio da informação e um atrator para os sistemas de recuperação da informação. Como determinar qual mensagem será informação para o receptor? Como assegurar a recuperação e a disseminação de algo que pode ocorrer fora dos limites dos sistemas? Pequenas alterações em uma política de seleção ou de indexação levam a grandes alterações na recuperação, pois uma questão negociada de formas distintas pode levar tanto a documentos altamente relevantes, como a irrelevantes. O SRI é sensível às alterações das condições iniciais, uma vez que exhibe asperezas semelhantes às de um fractal. A lei dos 80/20 tem, portanto, uma impressionante invariância na escala: independentemente do tamanho inicial da coleção, 20% dela atendem a 80% da demanda dos usuários. E cortes nessa coleção, por exemplo, reduzindo-a a 20% de seu tamanho inicial, não trarão alterações significativas de padrão, voltando a manifestar-se o fenômeno dos 80/20. Isso sem ir a fundo no processamento da informação no cérebro humano: os milhões de neurônios conectados por bilhões de sinapses, o armazenamento em “memórias”, as ligações e conexões inesperadas de ideias, a imprevisibilidade dos *insights* – características do caos (BRAGA, 1995, p. 3).

As duas questões apontadas por Braga (1995) na passagem acima podem ser adaptadas para o contexto do *Twitter*: *Como determinar qual tweet será informação para o usuário?* e *Como assegurar a recuperação e a disseminação de tweets que remetem a questões fora do sistema?* Essas duas questões apontam para a necessidade dos usuários em buscar um meio de filtrar e ordenar as mensagens do *Twitter* de modo a minimizar os efeitos da natureza caótica do sistema.

Para a classificação das mensagens disponibilizadas no site é facultado aos usuários atribuir uma ou mais *tags* que representem aquilo que se quer dizer. Para atribuir uma *tag* a

uma mensagem, o sistema exige o cumprimento de algumas regras: para ser uma *tag*, a(s) palavra(s) desejadas devem ser precedidas pelo símbolo # e não conter espaços se for necessário o uso de mais de uma palavra para representar um conceito. Alguns nomes próprios, em geral de personalidades bastante conhecidas são reconhecidos automaticamente como *tags*. Esse conjunto de regras explicita algumas necessidades de recuperação da informação de forma sistemática pelos integrantes do sistema.

Segundo Primo (2010), as *tags* começaram a serem usadas de forma espontânea, e a iniciativa inicial partiu do usuário do sistema Nate Ritter, no ano de 2007, quando o mesmo acrescentou o sinal “#” antes de “sandiegofire” a cada mensagem escrita por ele sobre os incêndios em San Diego. No *Twitter*, as *tags* são chamadas de *hashtags*.

Para a reflexão desse universo de classificação e representação da informação, a observação da Teoria do Conceito de Dahlberg (1978) se faz necessária. Para a autora, uma metodologia de classificação deve agregar os métodos indutivo e dedutivo por reconhecer que existem categorias já conhecidas (universais), porém as mesmas não seriam identificadas sem o processo indutivo de análise dos conceitos. A partir disso, compreende-se porque a rede social em questão busca a complementação do uso das *tags* para a classificação e representação da informação com algumas possibilidades de categorização. Santini e Souza identificam dois tipos de *tags*:

As *tags* são etiquetas individuais, ou seja, etiquetas pequenas e de texto livre aplicada a um conteúdo por um usuário por um usuário específico. O vocabulário destas etiquetas não são estruturados e não possuem limites linguísticos. A social *tags*, também chamadas de folksonomia, se referem aos conjuntos de *tags* individuais compartilhados por uma comunidade de usuários. (SANTINI; SOUZA, 2010, p. 4)

A busca por uma classificação colaborativa nessa rede social se tornou mais evidente quando a página começou a divulgar uma lista dos 10 tópicos (os tópicos podem ser *hashtags*, nomes próprios e sequencia de palavras) mais comentados por país. A estatística, ou seja, o quantitativo exato de quantas pessoas usaram uma *tag* não é informado para o público, e algumas *tags*, inclusive, são patrocinadas, o que reflete a tendência de induzir os usuários a seguirem o padrão de representação utilizado pela maioria, ou seja, estimular a interação entre os participantes para a uniformização dos termos para designar categorias.

Os *trending topics* (tópicos mais comentados) revelam mais do que a necessidade de uniformização e padronização terminológica. Naaman, Becker e Gravano (2011) identificaram a capacidade que os tópicos mais utilizados têm de representar eventos em tempo real, ou seja, os *trending topics* revelam acontecimentos de momento que podem ser de interesse local ou global e são alterados rapidamente em função da atualidade dos eventos.

Chang (2010) discute como finalidade das *hashtags* no *Twitter*, a de difusoras de inovação. O autor argumenta que a padronização seria um caminho que agregaria mais inovação ao sistema de *hashtags* do *Twitter* citando como exemplo a presença de duas formas de representar uma mensagem que trata de gripe suína na lista dos tópicos mais comentados do ano de 2009: *#swineflu* versus *Swine Flu*. O autor destaca ainda que a adoção ou não de uma *hashtag* para sua mensagem por parte dos usuários depende do quanto esses usuários foram ou não expostos a determinados assuntos e mensagens que contenham *tags*.

A necessidade de ordem (ou a busca dela) é auto-organização em um sistema como assinala Capra (1996). Para o autor, entender a auto-organização prevê o conhecimento prévio de padrão. O padrão representa o ponto de partida para a compreensão da vida que não é diferente na cibernética, “onde a ideia de um padrão de organização – uma configuração de relações característica de um sistema em particular” é a razão de sua existência. No *Twitter*, a opção para ordenar as mensagens é através do uso das *hashtags*, que surgiram por iniciativa de um usuário para a classificação de uma série de mensagens.

Até meados de 2009 a única maneira de recuperar informação contida nas mensagens de forma sistematizada era buscando as *tags* escolhidas pelos usuários. Com o aumento considerável de usuários do sistema, uma simples busca de *tags* passa a não atingir o objetivo desejado. Assim, surgiram as opções de salvar seus *tweets favoritos* e criar listas, que agrupam usuários que estão sendo seguidos em categorias definidas segundo a lógica pessoal de cada um, limitadas ao máximo de 20 listas. Isso reflete, o fato de que mesmo em sistemas onde há a predominância da descentralização de gerenciamento da informação, busca-se a manutenção de algum padrão inspirado em esquemas de classificação já formalizados.

A opção de criar listas para organizar os usuários conectados constitui recurso relevante para classificação nessa rede. O usuário do sistema cria categorias para incluir usuários que integram ou não uma rede em particular. Os usuários podem ser agrupados segundo suas atividades profissionais, relacionamento com que elabora as listas (ex. amigo, família, etc.), região, etc.

Num sistema em que usuários participam não apenas via computador, mas também via celular, é natural que uma opção para classificação de mensagens sobre uma localidade fosse desenvolvida. Isto significa dizer que os desenvolvedores do sistema já identificaram um padrão bem definido de representação territorial. Numa atualização do sistema, os usuários podem também tornar pública sua localização exata, através de um aplicativo que tem função semelhante à de um Sistema de Posicionamento Global (GPS).

Pereira e Cruz (2010) estudam as relações de afeto na escolha das *tags* por parte dos usuários. No estudo, os autores, em concordância com esta pesquisa investigam a funcionalidade dessas *tags* para o agrupamento e compartilhamento de mensagens que expressem afeto, opinião ou motivação para determinada ação. Uma observação que se pode fazer a respeito destas *tags* é que elas, além de organizar e classificar o conteúdo, agrupam também os usuários por um objetivo comum.

Com base nas questões conceitual-teóricas apresentadas, o trabalho desenvolveu-se no sentido de analisar o uso das *tags* (que no *Twitter* denominam-se *hashtags*) para classificação no *Twitter*.

### 3 OBJETIVOS

#### OBJETIVO GERAL

Investigar a classificação colaborativa da informação nas redes sociais da Internet através do uso das *tags* pelos usuários, considerando a presença de elementos de caos e de complexidade nesses sistemas.

#### OBJETIVOS ESPECÍFICOS

1. Investigar a dinâmica da disseminação da informação e da interação entre os usuários na rede social *Twitter* para identificar elementos das teorias do caos e da complexidade em sistemas.
2. Identificar a lógica do uso das *tags* para a classificação da informação no *Twitter* estabelecendo relação com o desenvolvimento de sites e ferramentas de apoio, destacando a característica colaborativa desse tipo de classificação.

#### 4 ANÁLISE DO USO DAS HASHTAGS E SITES PARA CLASSIFICAÇÃO DAS MENSAGENS NO TWITTER: METODOLOGIA

Neste trabalho, a organização da informação no Twitter é analisada através de pesquisa exploratória, que segundo Gil (2002, p. 41) tem como objetivo primordial “proporcionar maior familiaridade com o problema, com vistas a torná-lo mais explícito ou construir hipóteses”. Outra característica interessante da pesquisa exploratória é destacada por Raupp e Beuren (2006, p.80) “consiste no aprofundamento de conceitos preliminares sobre determinada temática não contemplada de modo satisfatório anteriormente. Assim, contribui para o esclarecimento de questões superficialmente abordadas sobre o assunto”.

A escolha pela pesquisa exploratória está associada a dois fatores: o primeiro deles é a novidade do assunto tratado. O segundo fator que justifica a escolha deste tipo de pesquisa reside na necessidade de realizar novas investigações acerca dos desdobramentos da atual pesquisa, de cunho preliminar.

A presente pesquisa envolve levantamento bibliográfico e a realização de pesquisa empírica de caráter documental, pois a amostra é constituída por mensagens enviadas ao Twitter e por sites desenvolvidos para complementar o uso das *hashtags* no Twitter.

A primeira amostra foi obtida a partir da seleção de mensagens enviadas ao Twitter nos meses de setembro e outubro por usuários de três tipos de perfis: usuários comuns, personalidades e instituições. A análise compreende a exploração dos usos das *hashtags* para classificação considerando o contexto do emissor da mensagem, a identificação de tipos de *hashtags* e de alguns elementos presentes nas teorias de classificação de Ranganathan.

A segunda amostra compreende o levantamento de sites e ferramentas desenvolvidos para complementar o uso do Twitter, com ênfase nos sites e ferramentas que se destinam ao uso das *hashtags* para classificação de mensagens. A análise realizada explorou o papel da colaboração na busca, recuperação e sistematização das *hashtags* para o uso dos usuários do Twitter.

Em paralelo à pesquisa empírica foi adotada a pesquisa qualitativa por proporcionar uma abordagem interpretativa aos dados coletados. Além disso, a pesquisa qualitativa está alinhada aos objetivos deste trabalho por permitir a observação e a descrição do problema sem uma hipótese pré-concebida (FLICK, 2009).

A opção por uma pesquisa qualitativa, não exclui o emprego de métodos quantitativos em alguns momentos da pesquisa. Segundo Morse (1991), o método que agrega características qualitativas e quantitativas é chamado de “triangulação simultânea”. Nesse

método, a interação entre o método quantitativo e qualitativo acontece mais na fase de conclusão da pesquisa complementando um ao outro.

A seguir, o universo da pesquisa é delimitado.

#### 4.1 DELIMITAÇÃO DO UNIVERSO DA PESQUISA

Como já explicitado, a primeira amostra a ser coletada foi a das mensagens no Twitter. Coletadas a partir do próprio site do *Twitter*, de modo a obter dados de três tipos de perfis:

- - Usuários: neste grupo se enquadram usuários que utilizam o *Twitter* para postar mensagens com finalidades diversas como opiniões, notícias, e de cunho profissional, desde que o perfil selecionado não seja de uso exclusivamente profissional.
- - Personalidades: neste grupo se enquadram perfis de celebridades e foram selecionados a partir da lista de perfis mais seguidos segundo o site *Twitaholic.com*.
- - Instituições: neste grupo estão os perfis mantidos por organizações.

A escolha de três tipos de perfis para a obtenção da amostra de mensagens tem o objetivo de explorar diferentes motivações para o uso das *hashtags*. A quantidade de mensagens selecionadas para a amostra foi definida a partir de notícias veiculadas na imprensa, que estimou a população deste estudo em 7,500.000.000 de mensagens enviadas ao site mês de setembro<sup>11</sup>. Um outro estudo elaborado por Daas et al. (2011) afirma que cerca de 14,8% das mensagens postadas em 2010 no idioma holandês contém *hashtags*. Considerando as duas afirmações, foi efetuado o cálculo de uma amostra aleatória simples (SANTOS, c2011), que resultou em uma amostra de mínima de 396 mensagens para a realização do estudo, contudo, o total de mensagens selecionadas para o estudo totalizou 400.

A segunda amostra foi obtida a partir de buscas em blogs e sites especializados em redes sociais, de modo a sistematizar o maior número possível de sites e ferramentas de apoio ao uso do *Twitter* e seleção de *hashtags*. As buscas foram realizadas em agosto de 2011. Foram recuperados mais de 250 sites, sendo que alguns foram recuperados em duplicidade. Após a eliminação destes, o quantitativo final de sites analisados foi de 122.

Do quantitativo final dos sites, alguns estavam direcionados para outras ações do *Twitter* que não estão diretamente relacionadas com o uso das *hashtags*. Estes sites são parte do estudo por terem como objetivo estruturar informação para o coletivo, no entanto, receberam mais atenção os 16 sites que estão relacionados à classificação de mensagens a partir das *hashtags*.

---

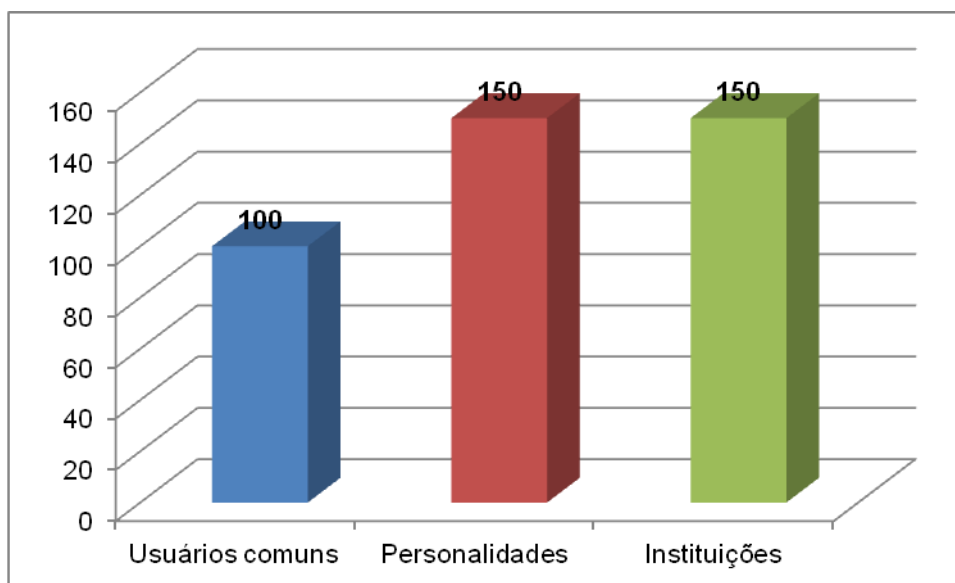
<sup>11</sup> [http://www.oficinadanet.com.br/noticias\\_web/4405/250-milhoes-de-mensagens-dia-e-recorde-do-twitter](http://www.oficinadanet.com.br/noticias_web/4405/250-milhoes-de-mensagens-dia-e-recorde-do-twitter)

A seguir, as amostras são detalhadas e interpretadas.

#### 4.2 DINÂMICA DO USO DAS TAGS NA REDE SOCIAL TWITTER

A base de dados de análise do presente estudo se compõe de 400 mensagens recuperadas no período de setembro a outubro de 2011, a partir de contas com perfis públicos que acessam a rede social *Twitter* de três tipos de perfis:

- Usuários comuns: usuários que não são pessoas famosas e utilizam o *Twitter* com objetivos diversos.
- Personalidades: pessoas muito populares na rede social que priorizam o uso do *Twitter* para fins profissionais.
- Instituições: perfis de pessoas jurídicas que usam o *Twitter* para promover suas atividades.



**Figura 3:** Mensagens selecionadas para o estudo por perfil de emissores

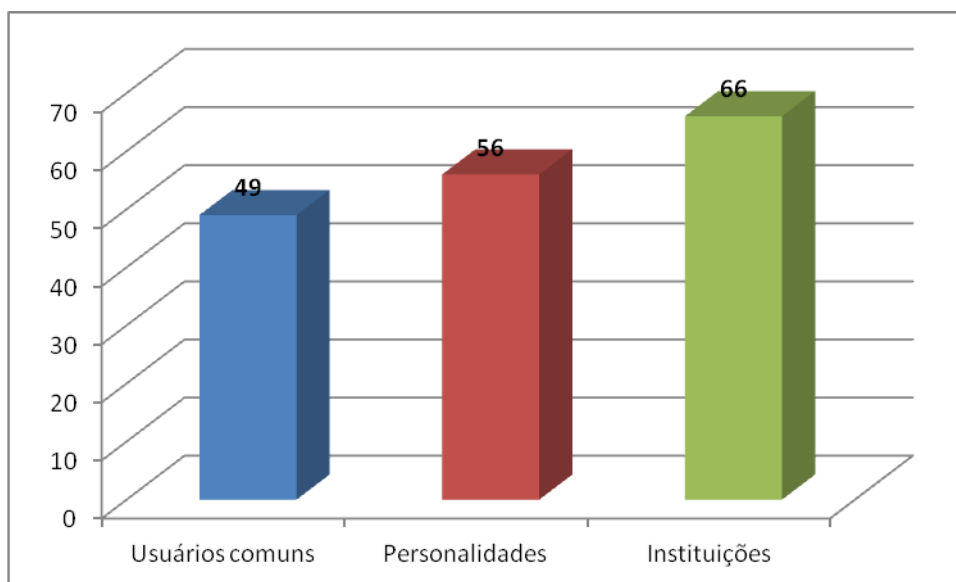
Na Figura 3 é apresentado o quantitativo definido para cada tipo de perfil em função do número de mensagens postadas por tipo de usuários no período. O grupo dos usuários comuns foi o que postou menos mensagens. Isso se justifica, em parte, pela heterogeneidade do grupo e pela ausência de um objetivo claro para postar as mensagens, ao contrário dos outros dois grupos que se direcionam mais diretamente a um público específico.

A seleção dos perfis que teriam as mensagens analisadas seguiu critérios distintos. Para o grupo de usuários comuns, os perfis foram buscados entre os seguidores de contas de instituições jornalísticas como, por exemplo, *New York Times*, por agregarem uma grande variedade de seguidores.

O grupo de personalidades é composto pelas três celebridades mais seguidas segundo o site [www.twitaholic.com](http://www.twitaholic.com), e são elas: Lady Gaga, Justin Bieber e Katy Perry, todos cantores.

O grupo das instituições visou obter informações de três seguimentos que apresentem certa popularidade nas redes sociais: o *YouTube*, por ser a empresa privada mais seguida segundo o site [www.twitaholic.com](http://www.twitaholic.com); a Organização Mundial das Nações Unidas por ser uma organização humanitária que envolve a participação de diversas nações; e Rio 2016 por ser o perfil de um evento esportivo a ser realizado no Brasil.

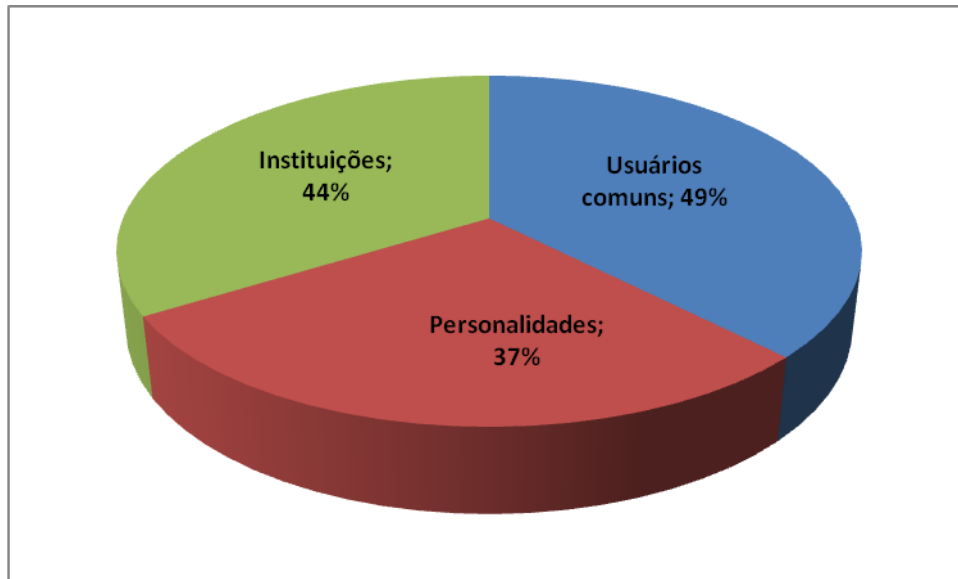
Do total de mensagens analisadas, 170 continham pelo menos uma *hashtag* representando 42% do total de mensagens analisadas. A análise por grupo apresenta na figura 4 a seguinte distribuição:



**Figura 4:** Distribuição de mensagens com hashtags por tipo de usuário

Na Figura 4 é observada a contribuição de cada grupo para o quantitativo geral de mensagens com *hashtags*. O grupo dos usuários comuns utilizou *hashtags* em 49 mensagens, o que representa o uso de *hashtags* em 29% das 171 mensagens que contém *hashtags*. O grupo das personalidades tem participação de 33% no total de mensagens com *hashtags*. O grupo das instituições foi o que apresentou o maior percentual de mensagens com *hashtags*: 39%.

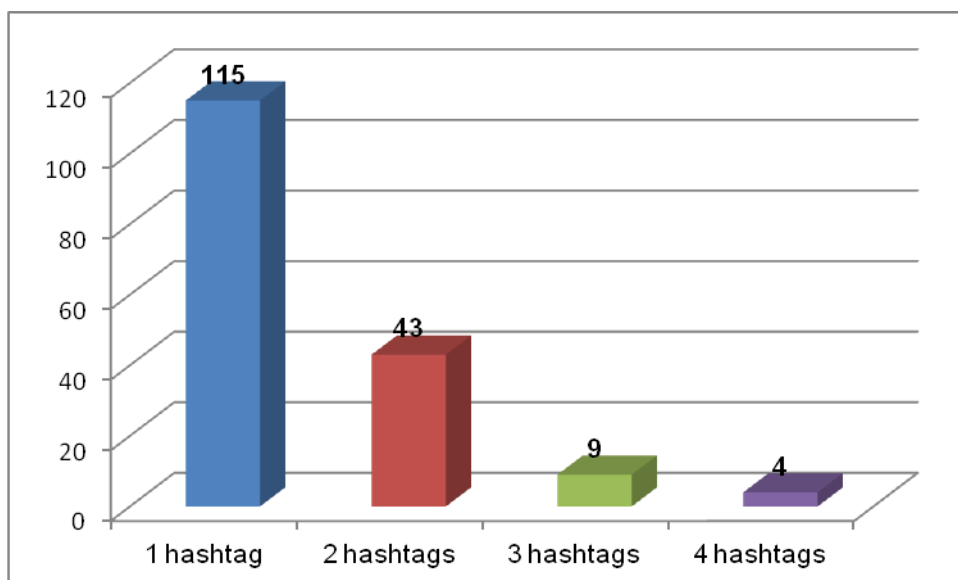
Na Figura 5 é apresentado o percentual de mensagens com *hashtags* segundo total de mensagens de cada grupo:



**Figura 5:** Percentual de mensagens com hashtags segundo total de mensagens de cada tipo de usuário

Na Figura 5, os usuários comuns utilizam *hashtags* em 49% do total de mensagens do grupo (100 mensagens). As personalidades utilizaram *hashtags* em 37% das 150 mensagens selecionadas para o estudo e as instituições em 44% das 150 mensagens.

Entre as mensagens selecionadas, o número de *hashtags* encontrado nas 171 mensagens que as contem, variam do mínimo de uma *hashtag* até o máximo de quatro *hashtags*. A distribuição do quantitativo de *hashtags* por mensagem está representado na figura 6:

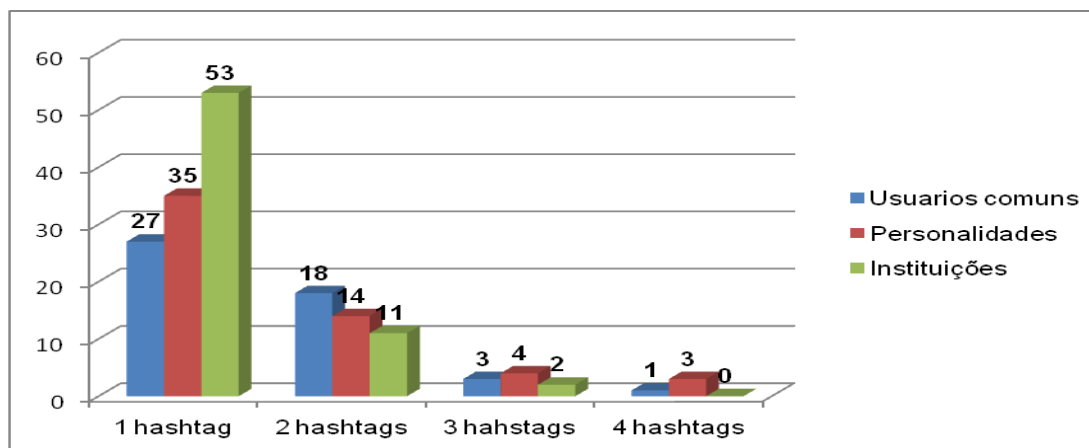


**Figura 6:** Quantitativo de hashtags por mensagem

Na Figura 6, 115 mensagens contem uma *hashtag* para a classificação da mensagem, o que representa 67% do total das mensagens que contem *hashtags*. Mensagens com duas

*hashtags* representam 25% e as mensagens com 3 e 4 *hashtags* representam 5% e 2% do total, respectivamente.

Em relação à mesma análise separada por tipos de usuários, chegou-se ao seguinte resultado:



**Figura 7:** Quantitativo de *hashtags* por mensagem distribuídas por tipos de usuários

A Figura 7 apresenta o número de *hashtags* que cada mensagem contém distribuídas por tipo de usuários. Todos os tipos de usuários utilizaram uma *hashtag* para classificar a maioria das mensagens: 27 mensagens do grupo usuários comuns; 35 do grupo personalidades e 53 do grupo instituições.

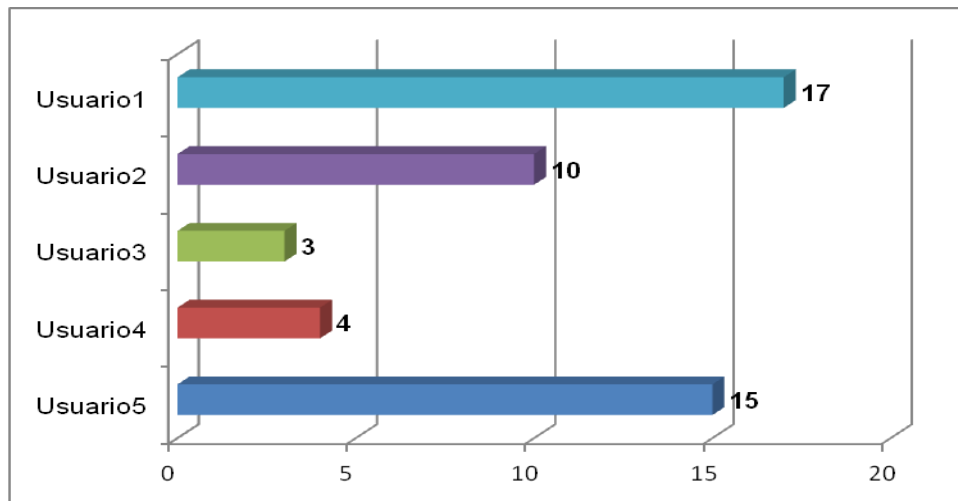
A utilização de duas *hashtags* por mensagem apresenta uma modificação no *ranking*: os usuários comuns utilizaram duas *hashtags* em 18 mensagens superando as personalidades que utilizaram duas *hashtags* em 14 mensagens. O grupo das instituições utilizou duas *hashtags* em 11 mensagens.

O quantitativo de mensagens contendo três e quatro *hashtags* foi menor que o quantitativo de mensagens contendo uma ou duas *hashtags*. O grupo das personalidades utilizou três *hashtags* em quatro mensagens; os usuários comuns em três mensagens e as instituições em duas mensagens. Dois grupos utilizaram quatro *hashtags* em pelo menos uma mensagem: os usuários comuns em uma mensagem e o grupo personalidades em três mensagens.

A análise apresentada até aqui revela algumas características gerais acerca do uso de *hashtags* dentro do contexto de cada grupo. A seguir, essas características são exploradas em detalhe.

#### 4.2.1 As *hashtags* do grupo usuários comuns

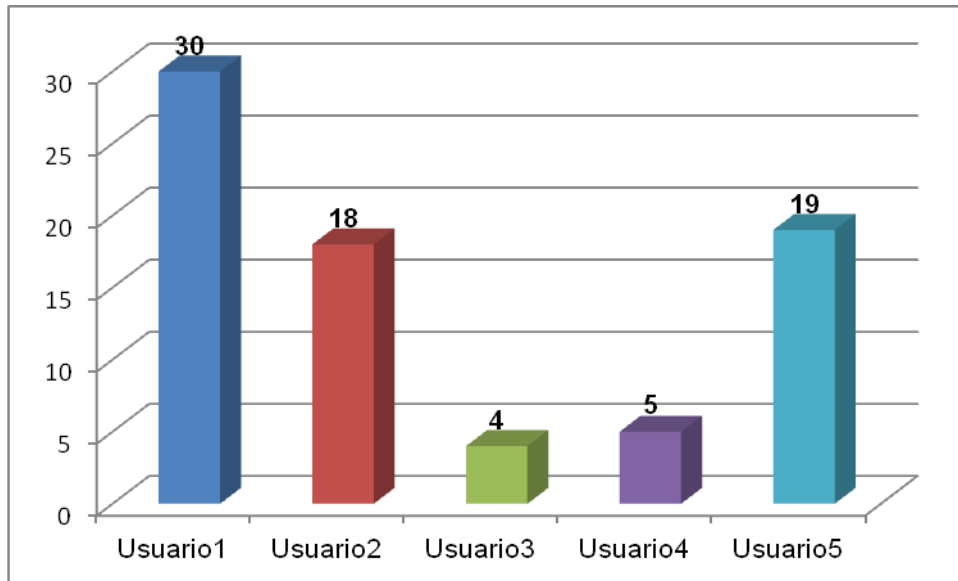
O grupo usuários comuns é composto por cinco perfis, com 20 mensagens de cada usuário. O número igual de mensagens para cada usuário foi estabelecido para que a investigação encontrasse algumas informações de ordem quantitativa de forma equivalente entre os cinco perfis. Por não se tratar de pessoas públicas, os usuários serão identificados como Usuário1, Usuário2, Usuário3, Usuário4 e Usuário5. Na Figura 8, são apresentados os quantitativos de mensagens contendo *hashtags*:



**Figura 8:** Mensagens contendo *hashtags* de cada usuário comum

Como observar-se na Figura 8, o Usuário1 e o Usuario5 utilizam *hashtags* na maioria de suas mensagens, representando 85% e 75%, respectivamente, enquanto que o Usuario2 utiliza *hashtags* em 50% de suas mensagens e os Usuários 3 e 4 utilizam *hashtags* em apenas 15% e 20% de suas mensagens, respectivamente.

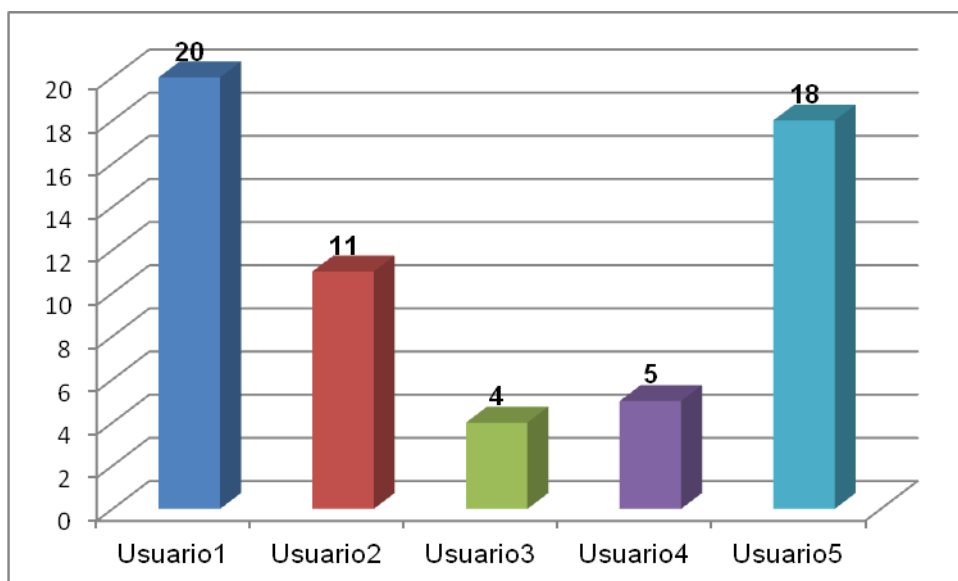
Em complementação a esses dados é relevante analisar o total de *hashtags* utilizadas por cada usuário:



**Figura 9:** Total de hashtags usadas por cada usuário comum

Numa mensagem podem conter quantas *hashtags* for de interesse de quem irá postar a mensagem. Neste grupo, foi possível identificar variação no número de *hashtags* necessárias para classificar uma mensagem. Como é possível observar na Figura 9, o usuario1 utilizou um total de 30 *hashtags*; o Usuário2 utilizou 18 *hashtags*; o Usuário3 utilizou quatro *hashtags*; o Usuário4 utilizou cinco *hashtags* e o usuário5 utilizou 19.

Na Figura 10, o quantitativo apresentado é o da ocorrência de *hashtags*, ou seja, as *hashtags* que se repetiram foram contadas uma vez:



**Figura 10:** Número de ocorrências de hashtags por usuário comum

Na Figura 10, os Usuários 1 e 2 utilizaram algumas *hashtags* por mais de uma vez enquanto que nos Usuários 3 e 4 as *hashtags* não ocorreram por mais de uma vez. O Usuário5

utilizou uma *hashtag* por mais de uma vez. Isso mostra que alguns participantes do *Twitter* escrevem de forma recorrente sobre determinados assuntos, enquanto outros são influenciados mais por determinados eventos, como nos mostra cada caso em particular a seguir.

Usuário1	
<i>Hashtags</i>	Frequência
<a href="#">#acupuncture</a>	1
<a href="#">#art</a>	1
<a href="#">#band</a>	1
<a href="#">#guitar</a>	2
<a href="#">#Guitar</a>	1
<a href="#">#guitarsolo</a>	1
<a href="#">#Jogged</a>	1
<a href="#">#Love</a>	3
<a href="#">#music</a>	5
<a href="#">#Music</a>	1
<a href="#">#qigong</a>	1
<a href="#">#quote</a>	1
<a href="#">#rawfood</a>	1
<a href="#">#Rawfood</a>	1
<a href="#">#song</a>	2
<a href="#">#Taichi</a>	1
<a href="#">#taichi</a>	1
<a href="#">#therapeutic</a>	1
<a href="#">#yang</a>	1
<a href="#">#yoga</a>	2

Quadro 6: *Hashtags* do Usuário1

O Usuário1 constrói suas *hashtags* utilizando uma palavra classificar suas mensagens na maioria dos casos. Ao analisarmos o Quadro 6, é possível identificar o assunto das mensagens, mesmo sem a leitura das mensagens. Percebe-se também que há uma certa regularidade de assuntos abordados em suas mensagens como “música” (*#music* aparece cinco vezes, *#Music* aparece uma vez, *#band* aparece uma vez e *#song* aparece também uma vez) e também a preferência pelo instrumento guitarra (*#guitar* aparece duas vezes, *#Guitar* aparece uma vez e *#guitarsolo* aparece uma vez).

O Usuário1 apresenta um conjunto de *hashtags* que revelam o interesse do Usuário1 em postar mensagens com assuntos relacionados ao estilo de vida saudável baseado em algumas atividades que fazem parte da cultura oriental: *#Taichi*, *#taichi*, *#therapeutic*, *#yang*, *#yoga*, *#qigong* e *#acupuncture*. As *hashtags* *#rawfood*, *#Rawfood* (comida crua) se refere hábitos alimentares.

As *hashtags* #rawfood, #Rawfood e #guitarsolo apresentam duas palavras juntas sem o uso do espaço: *raw food* e *guitar solo*. As palavras se apresentam unidas porque se um espaço for utilizado, somente a palavra que se encontra junto ao símbolo # será reconhecida como uma *hashtag*. Além de unir duas palavras para exprimir um conceito, é facultado ao usuário utilizar o caractere *underline* ( \_ ) para formar uma *hashtag* com duas ou mais palavras.

A preferência por unir as duas palavras sem o uso do caractere *underline* está associada ao limite máximo de 140 caracteres por mensagem, incluindo as *hashtags*. Se o usuário optasse pelo uso do *underline* restariam menos caracteres para redigir a mensagem.

O usuário1 apresenta *hashtags* que representam um mesmo assunto escritas com a inicial em maiúscula e minúscula, como por exemplo, #Taichi e #taichi. A presente pesquisa considera esses casos como *hashtags* distintas seguindo o mesmo critério empregado pelo próprio *Twitter*. Considerando que o sistema não expõe as razões para a adoção desse critério, o controle terminológico de letras maiúsculas e minúsculas não é necessário para a compreensão de uma *hashtag*.

Usuário2	
<i>Hashtags</i>	Frequência
<a href="#"><u>#cantwaitforfutureepisodes</u></a>	1
<a href="#"><u>#CBC</u></a>	1
<a href="#"><u>#egosarentjustforpoliticiansanymore</u></a>	2
<a href="#"><u>#gojetsgo</u></a>	3
<a href="#"><u>#HIMYM</u></a>	1
<a href="#"><u>#hubris</u></a>	2
<a href="#"><u>#mbpoli</u></a>	1
<a href="#"><u>#mbvotes</u></a>	4
<a href="#"><u>#NDP</u></a>	1
<a href="#"><u>#nordman</u></a>	1
<a href="#"><u>#ywg</u></a>	1

Quadro 7: *Hashtags* do Usuário2

O Usuário2 utiliza *hashtags* com construções mais complexas e direcionadas a situações e eventos do cotidiano como em #cantwaitforfutureepisodes (que revela a espera por episódios de uma série televisiva) e #egosarentjustforpoliticiansanymore (opinião sobre política).

As *hashtags* #cantwaitforfutureepisodes e #egosarentjustforpoliticiansanymore apresentam uma construção relevante: se separados os elementos que as compõem, encontramos as frases: *can't wait for future episodes* (não posso esperar por episódios futuros) e *egos aren't just for politicians anymore* (egos não são mais apenas para os políticos),

respectivamente. Para o Usuário2, a extensão das *hashtags* não interfere no limite de 140 caracteres por mensagem, pois as *hashtags* são constituídas para expressar uma opinião que dá continuidade ao sentido da mensagem.

A identificação das *hashtags* do Usuário2 foi possível com a utilização do site *tagdef.com*, que apresenta definições para uma *hashtags* pesquisada no site. As *hashtags* #mbvotes, #NDP, #CBC e #mbpoli estão interligadas e estão relacionadas com a votação para a criação de um novo partido democrático no Canadá, mais precisamente relacionadas à província de Manitoba. A *hashtag* #gojetsgo se refere ao time de hóquei no gelo *Winnipeg Jets*, como forma de incentivo ao time.

Usuário3	
<i>Hashtags</i>	Frequência
<a href="#">#CQC</a>	1
<a href="#">#AyrtonSenna</a>	1
<a href="#">#Senna</a>	1
<a href="#">#SistemOfADown</a>	1

Quadro 8: *Hashtags* do Usuário3

As *hashtags* utilizadas pelo Usuário3 se assemelham às do Usuário1 no que tange o modo de construção, ou seja, escolhe termos que podem ser compreendidos por um grupo significativo de pessoas, por tratar de assuntos relacionados ao esporte e ao entretenimento. #AyrtonSenna e #Senna se referem ao piloto tricampeão da Fórmula 1 Ayrton Senna, falecido em 1994. #CQC é a sigla do programa de televisão Custe o Que Custar. Enquanto que #SistemOfADown se refere à banda de rock System of a Down. Essas *hashtags* representam alguns programas e especiais de televisão transmitidos em setembro e outubro (meses selecionados para a coleta dos dados).

Usuário4	
<i>Hashtags</i>	Frequência
<a href="#">#windowsphoneBR</a>	1
<a href="#">#JR</a>	1
<a href="#">#QueroRomárioNoMinistério</a>	1
<a href="#">#IE10</a>	1
<a href="#">#Windows8</a>	1

Quadro 9: *Hashtags* do Usuário4

Das cinco *hashtags* utilizadas pelo Usuário4, #QueroRomárioNoMinistério se enquadra no que Lemos (2010) identificou como *hashtag* que forma uma comunidade temporária, que são *hashtags* que tem a função de facilitar o acompanhamento de temas

específicos em um espaço de tempo determinado. Essa *hashtag* representa um desejo explícito do usuário diante da veiculação na imprensa de denúncias de corrupção no Ministério do Esporte. As outras quatro *hashtags* são todas produtos: #windowsphoneBR, #IE10 e #Windows8 no caso da empresa Microsoft e #JR é uma abreviatura de Jornal da Record, programa jornalístico do canal de TV aberta Record.

Usuário5	
<i>Hashtags</i>	Frequência
<a href="#">#bigbangtheory</a>	1
<a href="#">#BTJ</a>	1
<a href="#">#champ</a>	1
<a href="#">#derrenbrown</a>	2
<a href="#">#fifa</a>	1
<a href="#">#fryup</a>	1
<a href="#">#guttled</a>	1
<a href="#">#itstakenawhile</a>	1
<a href="#">#manutdvsaldershot</a>	1
<a href="#">#not</a>	1
<a href="#">#nufc</a>	1
<a href="#">#pointless</a>	1
<a href="#">#qpr</a>	1
<a href="#">#raywilkins</a>	1
<a href="#">#simplyunstoppable</a>	1
<a href="#">#spurs</a>	1
<a href="#">#survivorseries</a>	1
<a href="#">#wilfred</a>	1

Quadro 10: *Hashtags* do Usuário5

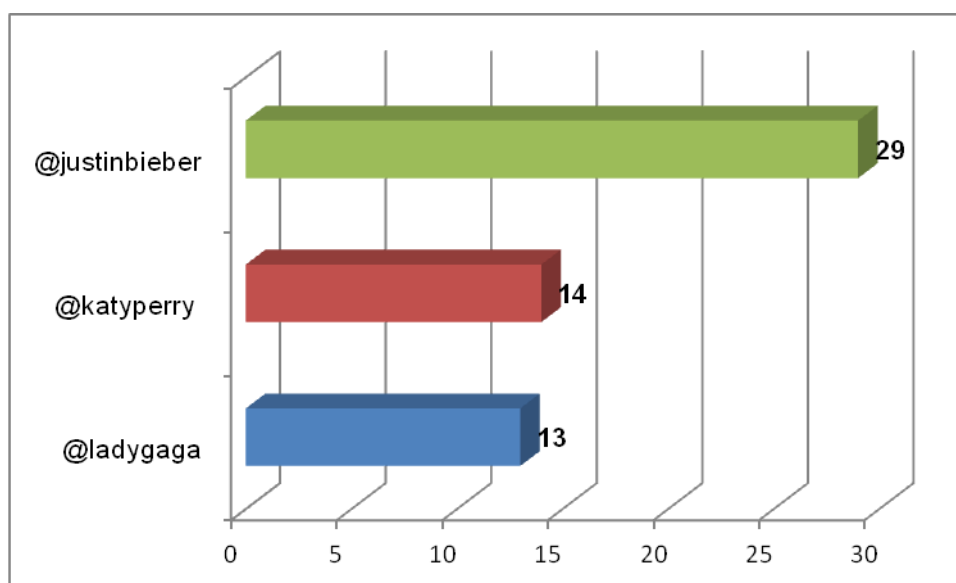
O Usuário5 é o último perfil do grupo de usuários comuns cujas mensagens analisadas apresentam várias *hashtags* sobre esportes, mais precisamente sobre o futebol do Reino Unido: #spurs, #nufc, #qpr. Essas *hashtag* foram conhecidas através das definições fornecidas pelo site *tagdef.com*, e no caso da *hashtag* #spurs foi identificada uma ambiguidade no sentido: #spurs pode remeter ao time de basquete norte-americano *San Antonio Spurs* e ao time de futebol do Reino Unido *Tottenham Hotspur*. Como a mensagem que continha a *hashtag* era sobre futebol, a *hashtag* #spurs representa o time *Tottenham Hotspur*.

Outras *hashtags* representam séries televisivas: #bigbangtheory, #pointless, #survivorseries, #wilfred. Este usuário contribui para este estudo, por conter algumas *hashtags* que são definidas por Primo (2010) como afetivas, que são *hashtags* que expressam

algum sentimento, emoção ou opinião: #not – não em português, #fryup – fritar ou fritando, #itstakenawhile – demorou algum tempo, #simplyunstoppable – “simplesmente incansável<sup>12</sup>”.

#### 4.2.2 As *hashtags* do grupo personalidades

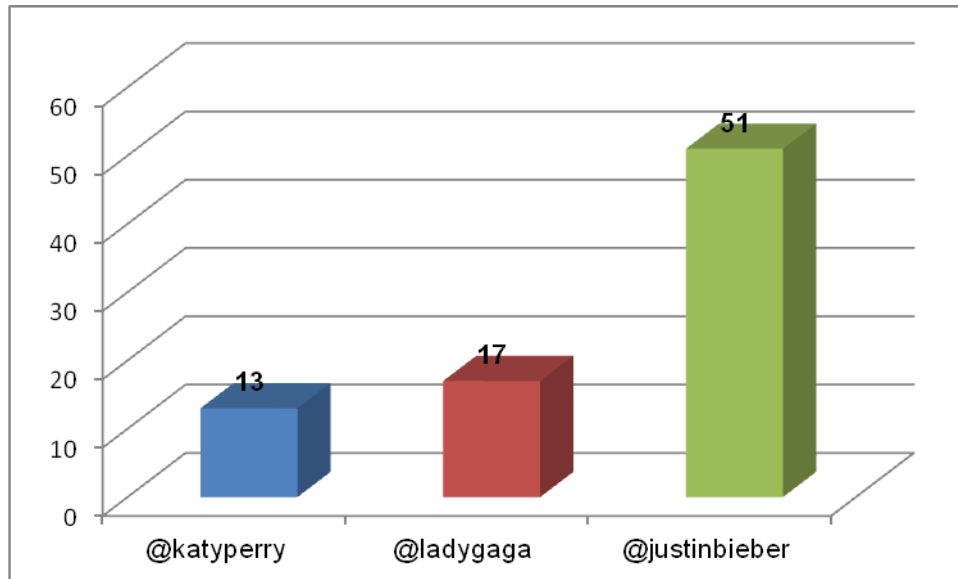
O grupo estudado compreende perfis de pessoas públicas que administram contas no *Twitter*. As personalidades foram selecionadas por serem as três mais seguidas em toda a rede *Twitter* (*Worldwide*) segundo o site *Twitaholic.com* e foram estudadas 50 mensagens de cada perfil. Os três são artistas da música, Lady Gaga e Katy Perry são norte-americanas e Justin Bieber é canadense. Na figura 11 estão discriminados os quantitativos de mensagens com *hashtags* por personalidade.



**Figura 11:** Mensagens contendo *hashtags* de cada personalidade

Na Figura 11 observa-se disparidade entre o cantor Justin Bieber e as cantoras Katy Perry e Lady Gaga. Enquanto o cantor utiliza *hashtags* em 58% das 50 mensagens selecionadas para estudo, Katy Perry e Lady Gaga apresentam *hashtags* em 28% e 26%, respectivamente. Panorama semelhante é obtido em relação ao número total de *hashtags* utilizadas:

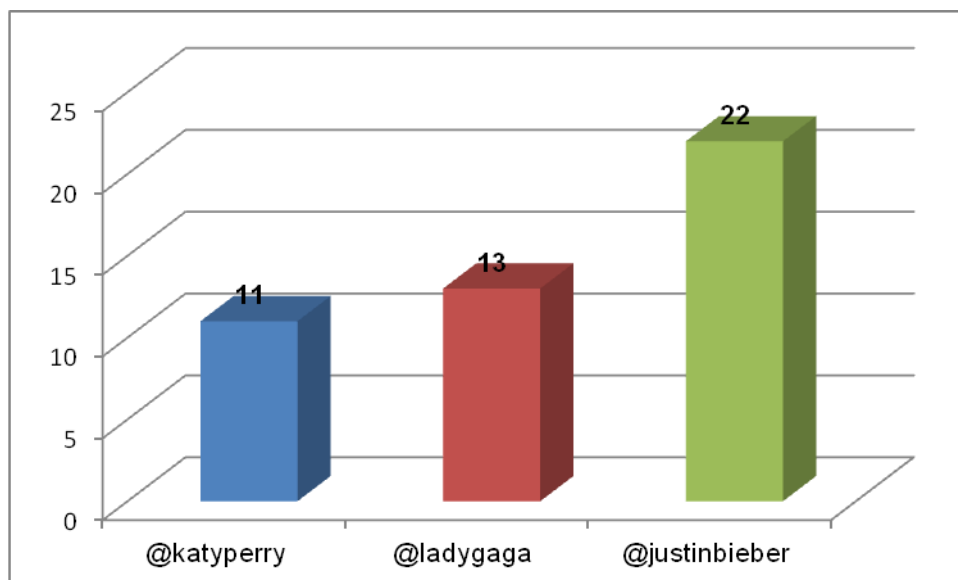
<sup>12</sup> Diante da ausência de tradução literal da palavra *unstoppable* para o português, tradução foi estabelecida através do contexto da que aponta uma longa partida de videogame.



**Figura 12:** Total de hashtags usadas por cada personalidade

Na Figura 12 é observado que a cantora Katy Perry utilizou 13 hashtags, enquanto que Lady Gaga utilizou um total de 17 *hashtags*. O cantor Justin Bieber utilizou um total de 51 *hashtags*, número muito superior ao de Kate Perry e Lady Gaga.

Na Figura 13 o quantitativo apresentado é o da ocorrência de *hashtags*, ou seja, as *hashtags* que se repetiram foram contadas uma vez:



**Figura 13:** Ocorrências de hashtags por personalidade

Na Figura 13 percebe-se que as cantoras utilizam menos *hashtags* que Justin Bieber, porém, variam mais os assuntos tratados que o cantor visto que eliminadas as repetições, o cantor utilizou 22 *hashtags* enquanto que Katy Perry e Lady Gaga utilizaram 11 e 13 *hashtags*, respectivamente. Ao comparar as figuras 12 e 13 identifica-se que o cantor repetiu 29 *hashtags*, enquanto Katy Perry e Lady Gaga repetiram 2 e 4 *hashtags*, respectivamente:

<b>@ladygaga</b>	
<b>Hashtags</b>	<b>Frequência</b>
<a href="#"><u>#where'sgaga</u></a>	1
<a href="#"><u>#NewDeli</u></a>	1
<a href="#"><u>#NewDelhi</u></a>	1
<a href="#"><u>#15millionmonsters.</u></a>	1
<a href="#"><u>#justsayin</u></a>	1
<a href="#"><u>#MARRY</u></a>	1
<a href="#"><u>#GovernmentHooker</u></a>	1
<a href="#"><u>#ThankYouSteve.</u></a>	1
<a href="#"><u>#14MillionMonsters</u></a>	1
<a href="#"><u>#MakeALawForJamey</u></a>	1
<a href="#"><u>#Giantsgame</u></a>	3
<a href="#"><u>#GoBlue</u></a>	3
<a href="#"><u>#Don'tEffWithNY</u></a>	1

**Quadro 11:** *Hashtags* de @ladygaga

A cantora Lady Gaga ocupa o posto número 1 da lista de perfis com mais seguidores (18.372.019 seguidores em números absolutos). O trabalho da cantora é marcado por performances ao vivo com forte apelo visual e suas músicas tratam de questões sobre o amor e a sexualidade. Duas de suas *hashtags* fazem clara referência ao seu trabalho: #14MillionMonsters e #15millionmonsters – que se referem ao crescimento do número de seguidores que a cantora tinha na época, inserindo os fãs no contexto de um de seus discos, o *The Fame Monster*.

Entre as *hashtags* da cantora, #NewDeli, #NewDelhi e #Don'tEffWithNY classificam mensagens relacionadas a um lugar e se enquadram na categoria fundamental espaço de Ranganathan. As *hashtags* #Giantsgame e #GoBlue estão interligadas e tratam da expectativa da cantora em relação a uma partida de futebol americano envolvendo o time *New York Giants*.

Entre as *hashtags* da cantora, #ThankYouSteve forma uma comunidade temporária. Essa *hashtag* foi criada para classificar mensagens de agradecimento ao legado deixado por Steve Jobs, executivo da empresa do ramo da computação Apple, na ocasião da divulgação de seu falecimento.

<b>@justinbieber</b>	
<b>Hashtags</b>	<b>Frequência</b>
<a href="#"><u>#12HOURS</u></a>	1
<a href="#"><u>#24Hours</u></a>	2
<a href="#"><u>#24HOURS</u></a>	1

<a href="#">#2DAYS</a>	6
<a href="#">#3DAYS</a>	2
<a href="#">#3days</a>	2
<a href="#">#48HOURS</a>	3
<a href="#">#4DAYS</a>	2
<a href="#">#6HOURS</a>	1
<a href="#">#9HOURS</a>	1
<a href="#">#brotherfromanothermother</a>	1
<a href="#">#ForReal</a>	1
<a href="#">#halfbeard</a>	1
<a href="#">#HYPED</a>	1
<a href="#">#Listen</a>	1
<a href="#">#Nov1st</a>	3
<a href="#">#nov1st</a>	1
<a href="#">#NOV1st</a>	3
<a href="#">#Tuesday</a>	1
<a href="#">#UnderTheMistletoe</a>	10
<a href="#">#UNDERTHEMISTLETOE</a>	6
<a href="#">#youngboyswag</a>	1

Quadro 12: *Hashtags* de @justinbieber

O quadro de *hashtags* do cantor Justin Bieber é composto por diversas *hashtags* que se repetem em sequencia: as *hashtags* #Nov1st, #nov1st, #NOV1st, #Tuesday, #9HOURS, #6HOURS, #4DAYS, #48HOURS, #3days, #3DAYS, #2DAYS, #24HOURS, #24Hours, #12HOURS estão relacionadas diretamente com as *hashtags* #UnderTheMistletoe e #UNDERTHEMISTLETOE que representam a contagem do tempo que faltava para o lançamento do álbum *Under The Mistletoe*, que é composto por canções de Natal. Estas *hashtags* têm função bem definida quanto à classificação das mensagens, que é a de agrupar mensagens que promovem a data de lançamento de seu trabalho, onde se observa a manifestação da categoria tempo de Ranganathan.

@katyperry	
<i>Hashtags</i>	Frequência
<a href="#">#braziliankatycats</a>	1
<a href="#">#CALIFORNIADREAMSTOUR</a>	1
<a href="#">#everyoneeatscandytoday</a>	1
<a href="#">#grandmagagaftw</a>	1
<a href="#">#KatyInJKT</a>	1
<a href="#">#Legend.</a>	1
<a href="#">#MEOW</a>	3
<a href="#">#queenmomishome</a>	1
<a href="#">#REDRUM</a>	1
<a href="#">#thecaliforniadreamstourgoesginger</a>	1

**#WeFoundLove**

1

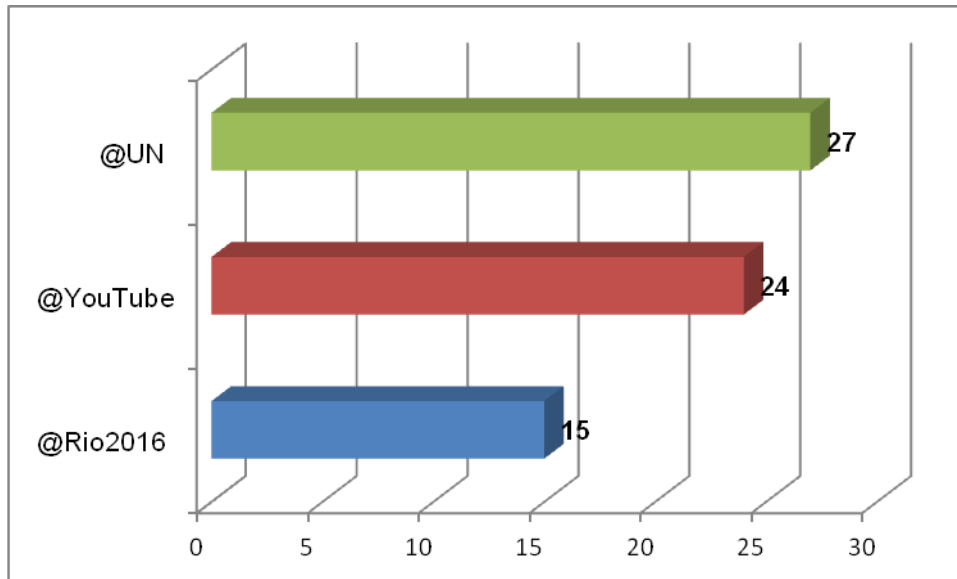
**Quadro 13:** *Hashtags* de @katyperry

A cantora Katy Perry apresenta *hashtags* com objetivo de divulgação semelhantes às do cantor Justin Bieber, porém de forma menos repetitiva. As *hashtags* #thecaliforniadreamstourgoesginger e #CALIFORNIADREAMSTOUR estão relacionadas à mensagens sobre a turnê mundial de shows que estava em curso na época; enquanto que a *hashtag* #MEOW classifica mensagens que promovem o perfume Meow!, assinado pela cantora. A cantora ainda faz alguns comentários sobre a vida pessoal, mais precisamente sobre sua avó e sua mãe: #grandmagagaftw e #queenmomishome.

A *hashtag* #MEOW apresenta ambigüidade no sentido. Usuários que acompanham a carreira da cantora entendem que trata-se do perfume que estava sendo lançado na época da seleção das mensagens. Para quem não tem conhecimento do contexto das mensagens com a *hashtag* #MEOW, é possível interpretá-la como uma onomatopeia para o som dos gatos em língua inglesa.

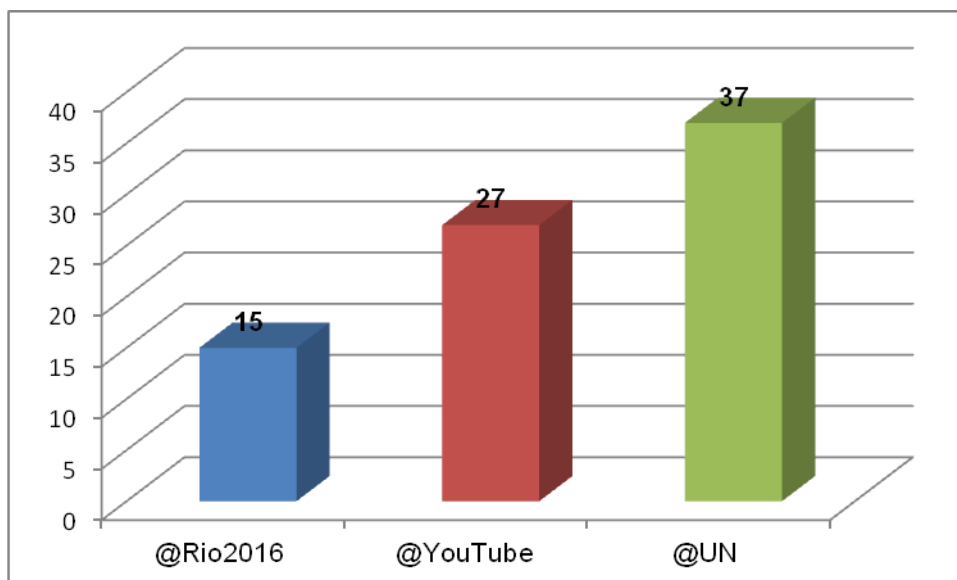
#### 4.2.3 As *hashtags* do grupo instituições

As instituições selecionadas para o estudo foram buscadas segundo suas finalidades: o site de compartilhamentos de vídeos @YouTube representa as grandes empresas do mundo virtual, a Organização das Nações Unidas (United Nations) representa uma organização mundial, e @Rio2016 representa um grande evento, as olimpíadas que se realizarão no Brasil em 2016. Cada uma apresentou o quantitativo de mensagens com *hashtags* como representado na figura 14:



**Figura 14:** Mensagens contendo hashtags de cada instituição

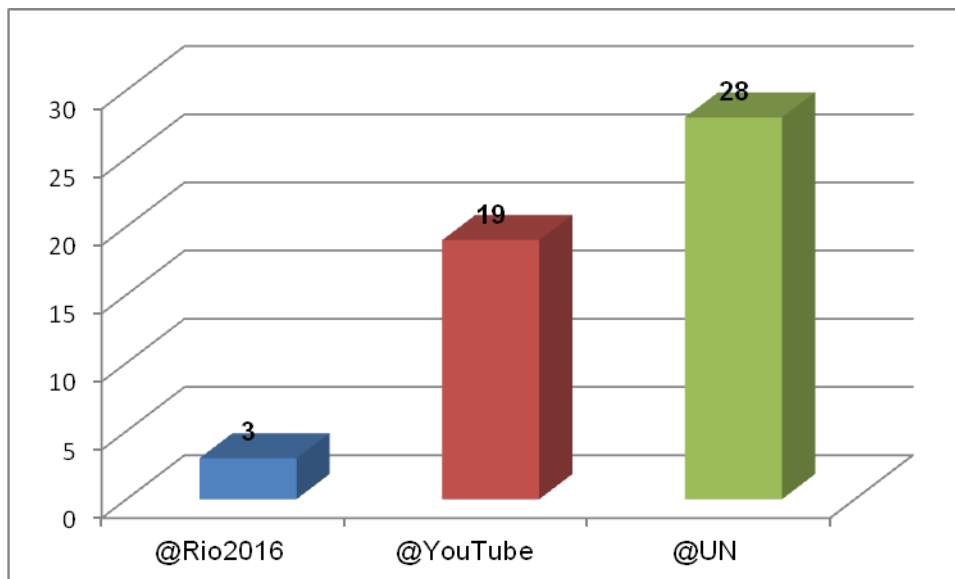
Na Figura 14, observa-se as Nações Unidas com mais mensagens contendo *hashtags*, representando 54% das 50 mensagens analisadas, seguida de YouTube que utilizou *hashtags* em 48% de suas mensagens. Nas mensagens de Rio2016, foram encontradas *hashtags* em 30% de suas mensagens. A Figura 15 apresenta os totais de *hashtags* usadas de cada perfil.



**Figura 15:** Total de hashtags usadas por cada instituição

Na Figura 15 é apresentado o quantitativo de *hashtags* usadas por cada instituição, com as Nações Unidas utilizando 37 *hashtag*, Youtube utilizando 27 *hashtags* e Rio2016 utilizando 15 *hashtags*. Observa que o perfil das Nações Unidas utilizou mais *hashtags* os outros dois perfis.

Na Figura 16, o quantitativo apresentado é o da ocorrência de *hashtags*, ou seja, as *hashtags* que se repetiram foram contadas uma vez:



**Figura 16:** Ocorrências de hashtags por instituição

Na Figura 16, fica mais visível a ocorrência de assuntos tratados nas mensagens de cada um. Enquanto Rio2016 utilizou apenas três *hashtags* distintas para classificar suas mensagens, os outros dois perfis variaram mais seus assuntos, o que de certa forma já era esperado pelas características das instituições: YouTube é um site de compartilhamento de vídeos de qualquer natureza e Nações Unidas trata de questões políticas e humanitárias ao redor do mundo. Nos quadros apresentados a seguir, é possível visualizar a frequência com que as *hashtags* ocorreram.

<b>@ YouTube</b>	
<b>Hashtags</b>	<b>Frequência</b>
<a href="#">#amandaknox</a>	2
<a href="#">#ausopen</a>	2
<a href="#">#CMJ</a>	1
<a href="#">#Dexter</a>	2
<a href="#">#edtech</a>	1
<a href="#">#getmoreinto</a>	10
<a href="#">#GetMoreInto</a>	2
<a href="#">#music</a>	5
<a href="#">#OnTheRise</a>	2
<a href="#">#OracleOpenWorld</a>	2
<a href="#">#SpaceLab</a>	1
<a href="#">#spacelab</a>	3
<a href="#">#StandUpToCancer</a>	1
<a href="#">#ThankATeacher</a>	1

<a href="#">#ThankYouSteve</a>	1
<a href="#">#v8sc</a>	2
<a href="#">#Worldview</a>	2
<a href="#">#YTrockinrioLIVE</a>	2
<a href="#">#ytrockinriolive</a>	1

**Quadro 14:** *Hashtags* de @YouTube

O YouTube é um site de carregamento e compartilhamento de vídeos on-line criado em 2005 e adquirido pelo Google em 2006. No site, podem ser encontrados vídeos com mais variados tipos de conteúdos. Isso está visível na classificação das mensagens no perfil da instituição no *Twitter*. Existem *hashtags* que formam comunidades temporárias como #ThankYouSteve – que representa mensagens de agradecimento ao executivo da Apple Steve Jobs no dia de sua morte (05 de outubro); #ThankATeacher – em comemoração ao dia internacional do professor e #StandUpToCancer – que incentiva pessoas a se engajarem na luta contra o câncer.

São frequentes *hashtags* sobre música, estas representando categorias bem definidas como #music – que trata de musica de uma forma geral, #YTrockinrioLIVE #ytrockinriolive que representam mensagens sobre o festival de música Rock in Rio, realizado no Rio de Janeiro no período pesquisado.

Outras *hashtags* que merecem o destaque são #getmoreinto e #GetMoreInto que apareceram em doze mensagens. Essas *hashtags* podem ser traduzidas como “obtenha mais” e são uma forma de orientar aos internautas a conhecer mais sobre determinado assunto assistindo à vídeos indicados pelo YouTube. As *hashtags* #SpaceLab e #spacelab classificam mensagens sobre um canal existente no site, específico para assistir a vídeos da Estação Espacial Internacional. Estas *hashtags* estão direcionadas para a divulgação e promoção de produtos oferecidos pela instituição.

@UN	
<i>Hashtags</i>	Frequência
<a href="#">#7billion</a>	9
<a href="#">#Argentina</a>	1
<a href="#">#AskUNWomen</a>	1
<a href="#">#Cambodia</a>	1
<a href="#">#children</a>	1
<a href="#">#Cooperatives</a>	1
<a href="#">#Darfur</a>	1
<a href="#">#endpoverty</a>	1
<a href="#">#G20</a>	2
<a href="#">#Halloween</a>	1

<a href="#">#HumanRights</a>	1
<a href="#">#ITUworld11Manifesto</a>	1
<a href="#">#Jobs</a>	1
<a href="#">#Kordofan</a>	1
<a href="#">#Mozambique's</a>	1
<a href="#">#other90cities</a>	1
<a href="#">#Palestine</a>	2
<a href="#">#Senegal</a>	1
<a href="#">#SG</a>	1
<a href="#">#subsidy</a>	1
<a href="#">#Sudan</a>	1
<a href="#">#Syria</a>	1
<a href="#">#thiomersal</a>	2
<a href="#">#tot4unicef</a>	1
<a href="#">#UN</a>	1
<a href="#">#UNAMID</a>	2
<a href="#">#UNESCO</a>	1
<a href="#">#vaccines</a>	1

**Quadro 15:** *Hashtags* de @UN

United Nations (UN) ou Organização das Nações Unidas (ONU) é uma organização internacional cujo objetivo declarado é facilitar a cooperação em matéria de direito internacional, segurança internacional, desenvolvimento econômico, progresso social, direitos humanos e a realização da paz mundial. Na análise das *hashtags* utilizadas observar-se a identificação de categorias relacionadas ao escopo institucional.

Em várias *hashtags* utilizadas é observada a manifestação da categoria fundamental “espaço” de Ranganathan. São elas: #Argentina, #Cambodia, #Mozambique's, #Palestine, #Senegal, #Sudan, #Syria, #Darfur, #Kordofan e #other90cities. Estas mensagens são sobre a situação política desses locais, e em alguns casos sobre ações da organização nesses lugares.

A *hashtag* #7billion se repetiu nove vezes e é uma *hashtag* que forma uma comunidade temporária, ou seja, permite o acompanhamento de uma situação ou evento em um espaço de tempo. Essa *hashtags* é sobre a população mundial, que atingiria os 7 bilhões de pessoas em 31 de outubro de 2011. As mensagens contendo essa *hashtag* discutem basicamente os impactos do tamanho da população mundial às questões da urbanização e o controle da pobreza.

<b>@Rio2016</b>	
<b><i>Hashtags</i></b>	<b>Frequência</b>
<a href="#">#5anos</a>	1
<a href="#">#em5anos</a>	12
<a href="#">#Rio2016</a>	2

**Quadro 16:** *Hashtags* de @Rio2016

O perfil Rio2016 é o *Twitter* oficial dos Jogos Olímpicos de Verão de 2016, oficialmente denominados Jogos da XXXI Olimpíada, será um evento multiesportivo realizado no segundo semestre de 2016, no Rio de Janeiro, Brasil. Este perfil apresentou a menor variedade de *hashtags*. Essa pouca variedade pode ser explicada pelas próprias características do perfil, que apresenta um universo bastante definido, que é o universo de uma competição esportiva.

As *hashtags* #em5anos, #5anos agrupam mensagens que fazem contagem regressiva para a realização do evento, em cinco anos. São *hashtags*, portanto, que formam comunidades temporárias sobre um concurso cultural realizado pela instituição para comemorar os cinco anos para a realização dos Jogos Olímpicos no Rio de Janeiro. A *hashtag* #Rio2016 trata de ações do comitê para a realização do evento.

#### SÍNTESE DOS RESULTADOS:

As *hashtags* analisadas dos três grupos estão atreladas às motivações pelas quais escrevem suas mensagens, que segundo Recuero e Zago são importantes para analisar a difusão da informação nas redes sociais:

Os atores decidem suas interações nessas redes sociais com base nos valores percebidos nessas trocas (Recuero, 2009). Esses valores são compreendidos como o acesso ao capital social construído na rede. O conceito de capital social é frequentemente associado aos valores que circulam nas redes sociais e que são percebidos pelos atores que fazem parte dela (RECUERO; ZAGO, 2010, p.72).

Em relação à troca de valores estabelecida por Recuero e Zago (2010) é relevante a exploração do alcance que se pretende com determinada *hashtag*. Segundo Cañada (2009) são quatro os tipos de etiquetas ou *tags* e alguns exemplos desses tipos foram encontrados na presente pesquisa:

##### 1. Etiquetagem egoísta:

São as que etiquetam algo com termos que visam a recuperação posterior da informação exclusivamente por quem o rotulou, se precisar da informação mais tarde. São rótulos importantes para um contexto individual, mas sem sentido fora ele. São exemplos: #not, #fryup, #itstakenawhile, #simplyunstoppable do usuário5 do grupo usuários comuns); e #grandmgagaftw e #queenmomishome de Katy Perry do grupo personalidades.

##### 2. Etiquetagem amigável

É quando algo é etiquetado para compartilhar com outras pessoas de um mesmo círculo (amigos, colegas, pares) e os termos de uso são consensuais entre os que fazem parte do

círculo e desconhecidos para o resto. Algumas etiquetas podem vir a ser relevantes também para o público em geral. Exemplos: #Taichi, #taichi, #qigong, #rawfood, #Rawfood do usuário1 do grupo usuários comuns.

### 3. Etiquetagem altruísta

É nesse tipo de *tag* que há a necessidade de compartilhar com o resto do mundo, onde a escolha de rótulos acontece segundo a capacidade de serem mais descritivos, conhecidos e geralmente aceitos. Exemplos: #bigbangtheory, #fifa, #survivorseries, #pointless, #wilfred do usuário5 do grupo usuários comuns; #NewDeli, #NewDelhi, #Giantsgame de Lady Gaga do grupo personalidades; #ThankATeacher e #StandUpToCancer de YouTube do grupo instituições; #ThankYouSteve utilizada por Lady Gaga e Youtube; #AskUNWomen, #Cooperatives, #children, #endpoverty, #HumanRights, #Jobs e #vaccines de Nações Unidas do grupo instituições.

### 4. Etiquetagem populista

É quando a escolha das *tags* é direcionada para atrair seguidores. Para atingir a esse fim, são usadas marcas que são populares e atraentes para a maioria das pessoas. Exemplos: #windowsphoneBR, #IE10, #windows8 do usuário4 do grupo usuários comuns; #UnderTheMistletoe e #UNDERTHEMISTLETOE de Justin Bieber do grupo personalidades; #MEOW de Katy Perry do grupo personalidades; #ausopen, #OracleOpenWorld, #YTrockinrioLIVE #ytrockinriolive de YouTube do grupo instituições; #UN, #UNAMID e #UNESCO de Nações Unidas do grupo instituições; e #Rio2016 de Rio2016 do grupo instituições.

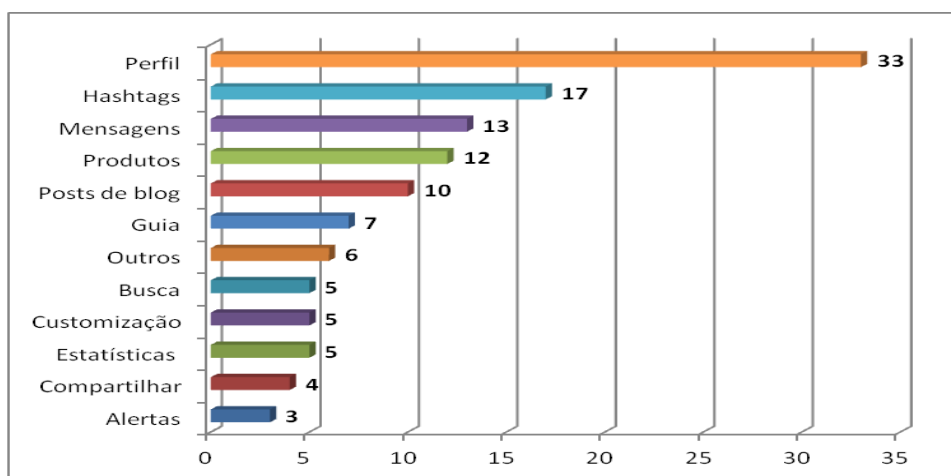
## 4.3 A COLETIVIDADE NA CLASSIFICAÇÃO DA INFORMAÇÃO: OS SITES DE APOIO À CONSTRUÇÃO DE TAGS

As redes sociais na Internet por estarem inseridas no contexto da Web 2.0, se utilizam ferramentas e sites para a construção de uma inteligência coletiva. Tratando especificamente do papel desses sites e ferramentas no contexto do *Twitter*, foram levantados diversos sites que se destinam à otimização do uso do *Twitter* em diversos aspectos, que foram divididos em grupos:

- Produtos – sites, ferramentas e softwares comercializados junto a empresas.
- Posts de blogs – postagens de blogs sobre *Twitter*.
- Alertas – sites e ferramentas que avisam ao usuário sobre novidades em sua conta.

- Busca – sites para realização de buscas em geral no *Twitter*.
- Compartilhamento – sites e ferramentas para o compartilhamento de fotos, vídeos e links nas mensagens.
- Customização – sites para alterar o visual da página do perfil no *Twitter*.
- Estatísticas – sites de estatísticas gerais da rede.
- Guias – Sites de ajuda aos usuários que usam o *Twitter* ou pretendem usar.
- *Hashtags* – Sites para buscar, descobrir, estatísticas, conhecer significados e tudo o que for relacionado às *hashtags*.
- Perfil – Sites para buscar, comparar, gerenciar, estatísticas, descobrir e tudo o que for relacionado ao perfil do usuário.
- Mensagens – Sites e plataformas para envio, visualização, busca, organização e tudo o que for relacionado às mensagens.
- Outros – Sites relacionados ao *Twitter* que não se enquadram em nenhum dos outros tipos.

Os sites foram buscados na Internet em geral (blogs especializados, buscadores – Google e yahoo!) como ponto de partida. Em uma etapa posterior, foi utilizado o site <http://www.similarsitesearch.com/> para encontrar o máximo de sites relacionados ao *Twitter*. Das diversas buscas realizadas no site chegou-se ao quantitativo de 250 sites. Foram eliminadas as repetições, e o quantitativo após os descartes foi de 135 sites. Destes, ainda foram descartados mais 16 sites, que tinham encerrado as atividades, totalizando 120 sites apresentados na figura 17:



**Figura 17:** Quantidade de sites para o Twitter segundo a finalidade

Como observa-se na Figura 17, a maioria dos sites recuperados estão relacionados a perfil de usuário representando 28% do total de sites analisados, seguido de *Hashtags* (14%),

Mensagens (11%), Produtos (10%), Posts de blog (8%), Guia (6%), Outros (5%), Busca (4%), Estatísticas (4%), Customização (4%), Compartilhar (3%), Alertas (3%).

Alinhada com os objetivos desta pesquisa, a análise dos sites se concentra nos 17 sites recuperados para o uso de *hashtags*. Alguns são apenas descritos e outros são mais bem detalhados por estarem mais ligados à proposta desta pesquisa que é identificar a coletividade na construção das *hashtags*. Lemos (2010) cita em seu trabalho que o aparecimento dos serviços de “indexação” de *hashtags* se deve em função da grande quantidade de usos inusitados das *hashtags*.

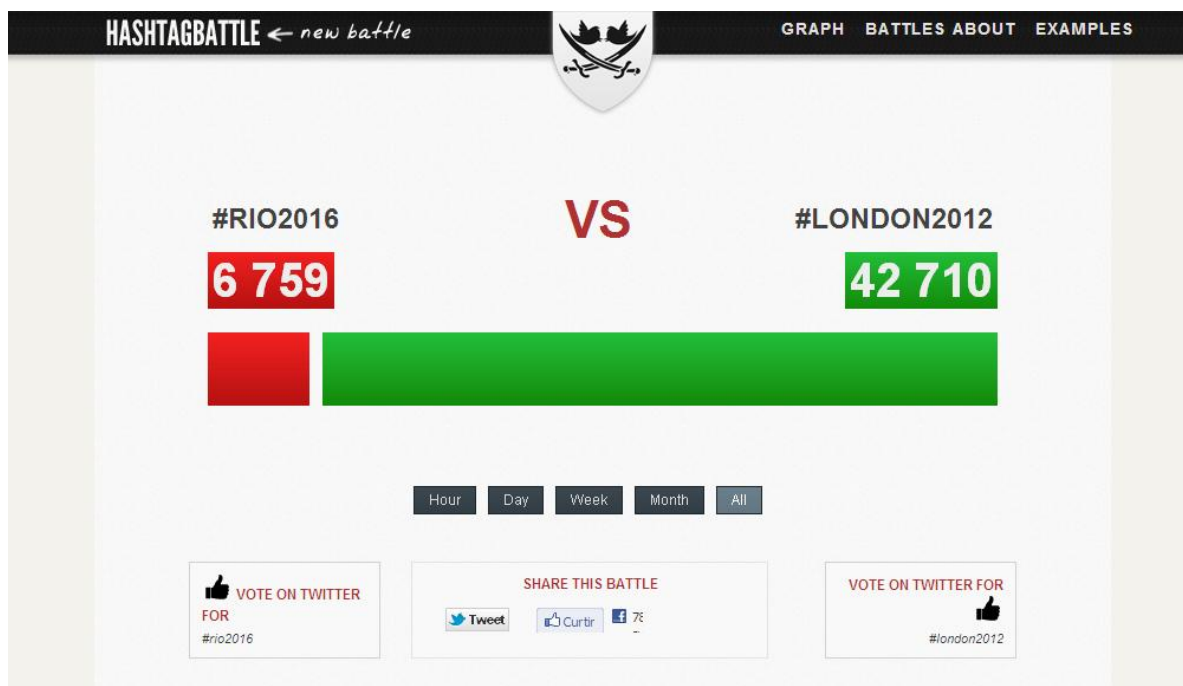
Os 17 sites selecionados para o estudo são:

1. Hashtagbattle - <http://hashtagbattle.com/>
2. HashParty - [www.hashparty.com/](http://www.hashparty.com/)
3. Twitag - <http://twitag.com/>
4. Hashalbum - <http://hashalbum.com/>
5. Hashtags's Blog - <http://hashtags.wordpress.com/>
6. Hashtags.org - <http://hashtags.org/>
7. Twittonary - <http://twittonary.com/>
8. Tagdef - <http://tagdef.com/>
9. My Tweet Cloud - <http://mytweetcloud.com/>
10. Tip Top Insight - <http://feeltiptop.com/>
11. Tweetalarm - <http://www.tweetalarm.com/>
12. TweetChat - <http://tweetchat.com/>
13. Twitoaster - <http://twitoaster.com/>
14. Monitter.com - <http://monitter.com/>
15. Statweestics - <http://statweestics.com/>
16. Twimmer - <http://twimmer.com/>
17. Buzzter - <http://buzztter.com/en>

O primeiro site da lista dos 17 é o *HASHTAGBATTLE* (<http://hashtagbattle.com/>). O site surgiu de uma brincadeira entre um grupo de amigos, em que parte deles estava na França e outra parte estava na Suíça e queriam saber qual dos dois lados venceria uma batalha de *hashtags*.

A “batalha” acontece ao digitar duas *hashtags*, clicar em *fight* e é exibido o número de vezes que cada uma foi usada. As *hashtags* escolhidas como exemplo foram #rio2016 e #london2012 para verificar qual das duas foi a mais citada. Essas duas *hashtags* foram

selecionadas por ambas estarem relacionadas aos Jogos Olímpicos de Verão e também pelo fato da *hashtag* #rio2016 ter aparecido no levantamento de *hashtags* do grupo instituições.



**Figura 18:** Total de citações das *hashtags* #rio2016 e #london2012 no site <http://hashtagbattle.com/>

Na Figura 18, pode-se constatar que #london2012 foi citada 42.710 vezes e #rio2016 foi citada 6.759. Isso representa uma maior popularidade da *hashtag* #london2012 em relação à #rio2016 em função da proximidade de realização deste evento em relação ao outro.

Outro site analisado foi o site *HashParty* (<http://www.hashparty.com/>) que apresentou diversos tipos de análise de *hashtags* como na figura 19:

The screenshot shows the HashParty website interface. At the top, there's a search bar with the hashtag #ThankYouSteve entered. Below the search bar, there's a section for 'Your recent search history' and a link to '#ThankYouSteve: is about... do you know? click here to edit'. The main content is divided into three columns: 'what they said', 'who said it', and 'sweet links'. The 'what they said' column lists tweets from users like noahvalisari, Miss\_Manou, denisiglesias, dmedroso, and others, all mentioning #ThankYouSteve. The 'who said it' column lists users who interacted with the tweets, such as noahvalisari, Miss\_Manou, denisiglesias, dmedroso, and others. The 'sweet links' column contains various links and sections like 'editors' choice hashtags', 'recently searched hashtags', and 'twitter trend hashtags'.

Figura 19: Pesquisa realizada no site HashParty com a hashtag #ThankYouSteve

O site lista mensagens indicando quem escreveu as mensagens, os principais links sobre a *hashtag* pesquisada e as *hashtags* pesquisadas recentemente. O site permite também que os resultados sejam exportados e salvos no formato .csv e o compartilhamento dos resultados encontrados em outras redes sociais. Os resultados encontrados revelam para quem realizou a busca, quem comentou o assunto pesquisado e as *hashtags* recentemente buscadas.

O terceiro site analisado foi *Twitag* (<http://www.twitag.com/>), um buscador de mensagens a partir da digitação de uma *hashtag*. A busca, porém, só recupera mensagens postadas muito recentemente, ou seja, está voltado à recuperação de mensagens sobre acontecimentos de momento e no idioma inglês, embora isso não esteja explícito no site.

O site *Hashalbum* (<http://hashalbum.com/pages/about>) é um site de buscas para descobrir assuntos comentados em tempo real relacionados a mensagens do *Twitter* com fotos. Ao se cadastrar no site, cada vez que um usuário do *Twitter* postar fotos com uma mesma *hashtag*, as mesmas são agrupadas pelo site formando um álbum.

O site *Hashtags's Blog* (sic) (<http://hashtags.wordpress.com/>) é um blog que visa promover outra página para pesquisa e descoberta de *hashtags*, a *hashtags.org* (<http://hashtags.org/>). O blog divulga os mais recentes aprimoramentos na página

*hashtags.org*, além de solicitar colaboração para o desenvolvimento e manutenção da página, seja através de doações ou na criação de forma voluntária de APIs<sup>13</sup>.

O já citado site *hashtags.org* (<http://hashtags.org/>) apresenta uma perspectiva diferenciada de outros sites relacionados ao uso de *hashtags*. Na página principal do site, há uma caixa de busca onde qualquer *hashtag* pode ser digitada para a recuperação e também sugere algumas *hashtags* de forma categorizada, como na figura 20:

The screenshot shows the homepage of *hashtags.org*. At the top, there is a navigation bar with links for 'Hashtags Home', 'About', 'Advertise', and 'Follow Us'. Below this is a search input field with the placeholder text 'Type a hashtag and hit Enter.' Underneath the search field, it says 'Or, select a hashtag from the list below:' followed by several categories of suggested hashtags:

- Business:** #jobs, #business, #networking, #retail, #shoppers, #consumer, #sales, #economy, #technology, #luxury, #news, #internet, #mktg, #marketing, #webmarketing
- Celebrities:** #50cent, #aplusk, #bieber, #britneyspears, #charliesheen, #ladygaga, #obama, #oprah, #osama, #perezhillon, #ryanseacrest, #the\_real\_shag, #theellenshow
- Education:** #edapp, #edtech, #education, #elearning, #lrnchat, #mllearning, #teachers
- Environment / Justice:** #humanrights, #poverty, #hunger, #aid, #sustainability, #health, #green, #eco, #earthtweet, #humantrafficking, #climate, #solar, #fairtrade
- Social Change:** #socialgood, #cause, #volunteer, #4change, #giveback, #dogood, #crisiscommons
- TV Entertainment:** #bones, #dwt, #glee, #gossipgirl, #graysanatomy, #idol, #lrnchat, #mllearning, #sharktank, #survivor, #teachers, #theoffice, #thevoice, #xfactor

On the right side, there is a 'Show Some Love' section with a 'Curtir' button and a notification that 'Amanda Moura e 3.280 outras pessoas curtiram isso.' Below this are social media icons for Twitter, Facebook, LinkedIn, YouTube, and Instagram. Further down is an 'About Hashtags' section explaining that the service is the de facto standard for hashtag information and that the # symbol is used to mark keywords or topics in tweets. At the bottom right is a 'Hashtags Sponsors' section listing 'micfo.com' and 'DomainSherpa'.

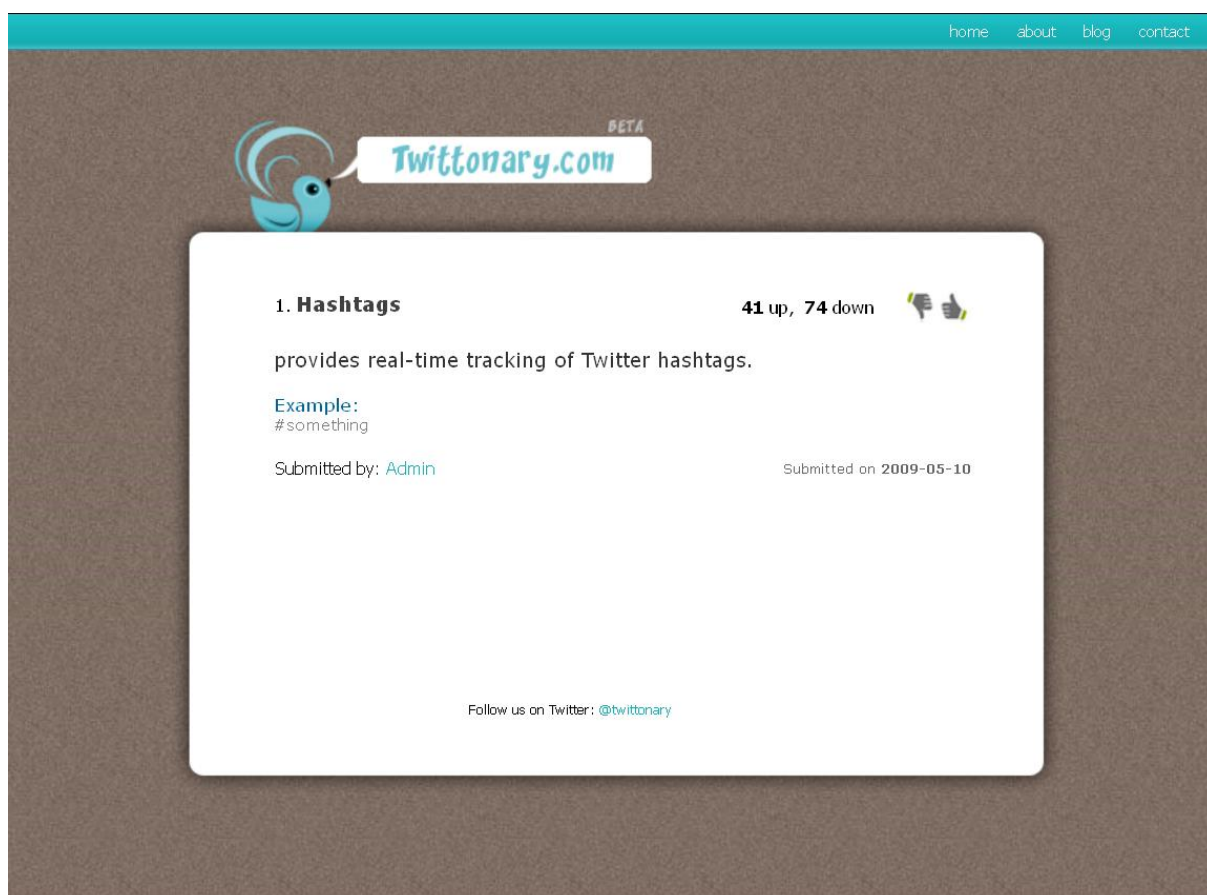
**Figura 20:** Página inicial do site *hashtags.org*

Na Figura 20 são sugeridas algumas *hashtags* para busca, categorizadas segundo os seguintes assuntos: Business (Negócios), Celebrities (Celebidades), Education (Educação), Environment/Justice (Meio Ambiente/Justiça), Social Change (Mudança Social) e TV Entertainment (Entretenimento na TV). Como já mencionado, o modo como as *hashtags* estão organizadas se assemelham aos diretórios criados por sites de buscas como Yahoo! e Google. E esses por sua vez se inspiram na Taxonomia por reconhecerem assim como Campos e Gomes (2008, não paginado) que “em qualquer dos modelos de acesso [à informação], o sucesso da recuperação depende, basicamente, de alguns fatores: É preciso haver um vocabulário-padrão; Os termos precisam ser atribuídos segundo critérios previamente estabelecidos, para assegurar consistência no tratamento dos dados; O software deve ser

<sup>13</sup> *Application Programming Interface* ou Interface de Programação de Aplicativos.

adequado às características de tal serviço de informação”. A diferença está na construção do sistema, que prima pelo coletivo e pela descentralização na definição dos critérios para a seleção dos termos.

A busca pela padronização aparece em dois sites que se destinam a apresentar significados para algumas *hashtags*. O site *The Twitter Dictionary* ou *Twittonary* (<http://twittonary.com/>) fornece explicações sobre termos comumente empregados para denominar serviços e ações realizadas no *Twitter*, como na Figura 21:



**Figura 21:** Significado da palavra hashtags

Os significados são enviados pelos usuários da página. Como no exemplo acima, é identificado quem fez a definição, a data de envio e dois polegares em sinal de positivo e negativo, que servem para avaliar a definição como positiva e negativa. Esse recurso de avaliação está presente em diversas redes sociais e funciona como uma forma de avaliação de conteúdo.

O site tagdef (<http://tagdef.com/>) atua como um dicionário de *hashtags* e foi de importante para a realização dessa pesquisa, por fornecer dados que permitiram a interpretação de algumas *hashtags* em língua inglesa que não eram do conhecimento da

pesquisadora. Além de mostrar rankings de novas *hashtags*, *hashtags* mais usadas do dia, *hashtags* mais usadas na semana e de *hashtags* mais usadas de todos os tempos, se concentra no fornecimento de definições de várias *hashtags*, escritas em vários idiomas, como representadas nas figuras 22 e 23:

The screenshot shows the #tagdef website interface. At the top, there is a search bar with the placeholder text "Enter hashtag or search phrase" and a "Look Up" button. Below the search bar, there are five columns of hashtag rankings:

New Tags	Top today	Top this week	All time high	Trending Tags
#eamsingle	#eksaherada	#eksaherada	#ff	x #pilksq
#hompdyompty	#subtweet	#barkhabarks	#subtweet	x #comeonmurray
#brokenheart	#ff	#oomf	#oomf	x #grillnev2
#plop	#dpda	#subtweet	#justinbieber	x #censuramestatwit
#shikhare	#huaiquipedia	#ff	#fb	x #ovetemdnarubror
#wowgbg	#fb	#yolo	#ctb	x #grillnev2
#cenobiteindustries	#oomf	#ipkatycatallie	#np	x #paidika_traumata
#on	#savetheintern	#sherlocked	#cot	x #paidika_travmata
#workathome	#sherlocked	#bieberconda	#p2	x #pasustres
#homebasedbusiness	#np	#fb	#globomente	x #vuelveramn

At the bottom of the page, there is a footer with the text "Contact info Share search with gog.is. privacy policy" and a "Developers." link. A small note at the bottom right of the trending tags section states: "x Indicates that a a definition i needed."

**Figura 22:** Rankings do site tagdef.com

Discover what hashtags really mean, and add your own definitions in seconds.

Enter hashtag or search phrase

# ff Look Up

**New Tags**

- #teamsingle
- #homptydompty
- #brokenheart
- #plop
- #shikhare
- #wowgbg
- #cenobiteindustries
- #on
- #workathome
- #homebasedbusiness

**Popular definitions**

**Definition of #ff**

Tweet this definition 58 +1

#ff is the same as (short for) [#followfriday](#). 1200 617

Details... Added on May 8, 2009

every friday you can use [#followfriday](#) ([#FF](#)) to suggest people to follow. 708 322

Details... Added on May 26, 2010

**PEGA PREÇO**  
WWW.PEGAPREÇO.COM.BR

Notebook  
pague Barato

\*Imagem ilustrativa

Preços a partir de  
**R\$ 699,90**

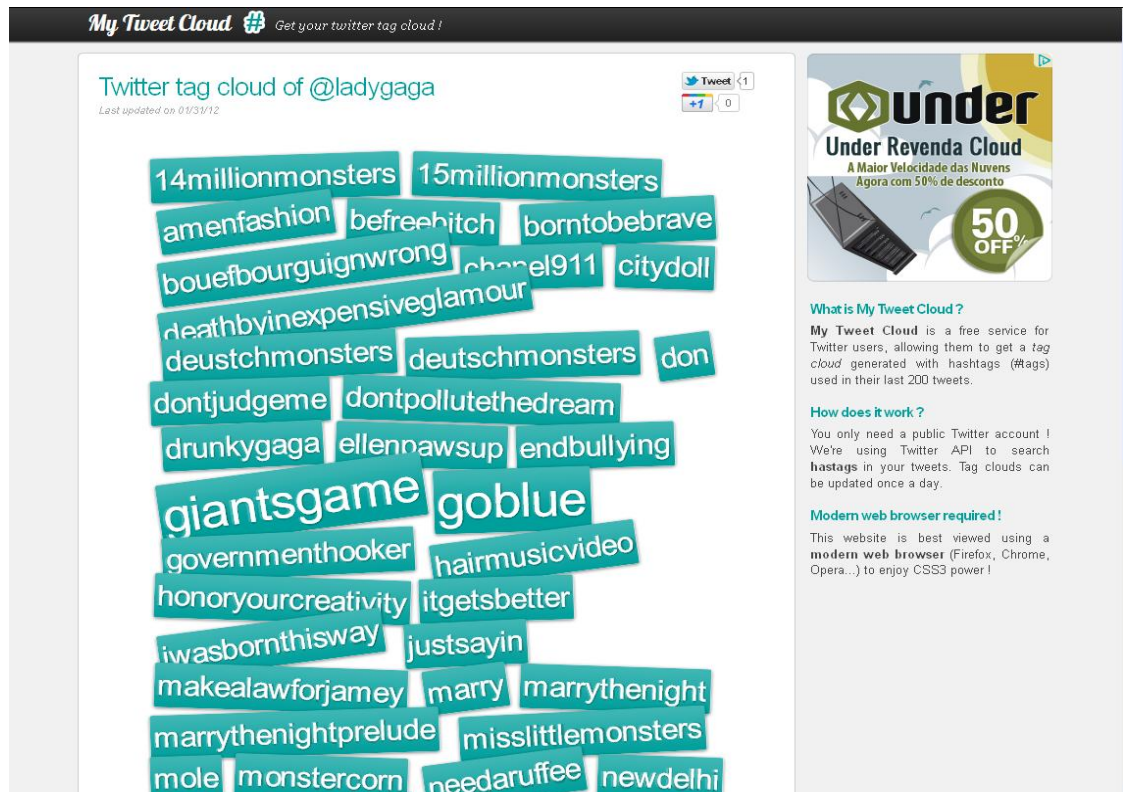
**Figura 23:** Definição da hashtag #ff

As Figuras 22 e 23 representam uma busca por uma regularidade. À medida que são divulgadas as *hashtags* mais utilizadas em determinado período, usuários podem sentir-se estimulados a usar tais *hashtags* de modo a participar de uma discussão que envolve muitos integrantes. Em complementação, oferece também os significados das *hashtags* para que se conheça de que se trata o universo das mensagens classificadas.

Carter, Tsagkias e Weerkamp (2011) propõem um método de tradução semelhante ao empregado pelo site *tagdef.com* por reconhecer que as variantes de *hashtags* podem levar a problemas na análise de conteúdo. Utilizando como exemplo as *hashtags* sobre um desastre em uma mina no Chile, chegam a duas questões:

[primeira questão] etiquetas, ou parte delas, são traduções literais, e, mais interessante, [segunda questão] a especificidade da *hashtag* é diferente por localidade. Quanto à [primeira questão], as pessoas poderiam usar traduções “simples”, traduzindo de #33mineros para #33miners. Quanto a [segunda questão], observamos que as pessoas no Chile referem-se a este evento com *tags* bem específicas, enquanto as *tags* em inglês também incluem termos gerais como #Chile. A partir de uma perspectiva não-chilena, o acidente de mineração é "a única coisa" acontecendo no Chile naquele momento. (CARTER; TSAGKIAS; WEERKAMP, 2011, p. 1).

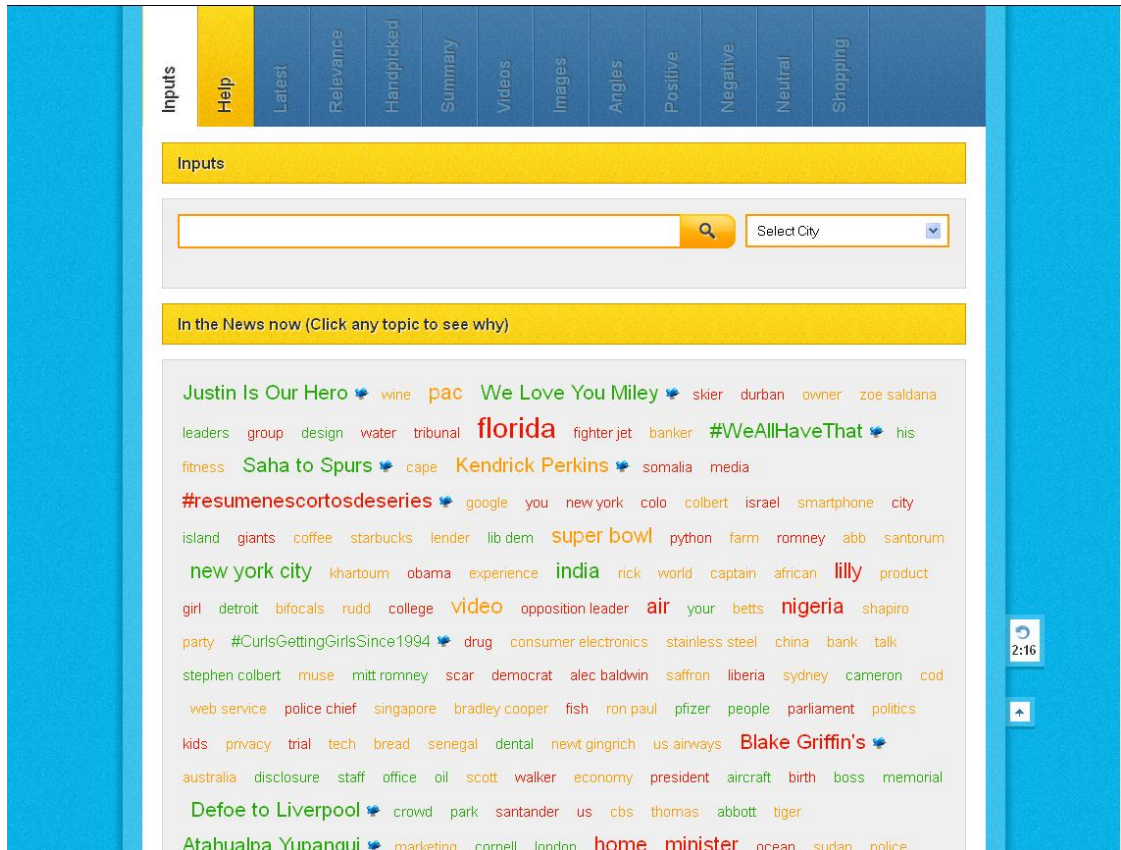
O site *My Tweet Cloud* (<http://mytweetcloud.com/>) se caracteriza pela representação visual de um conjunto de *hashtags* de um usuário da rede. Ao digitar o nome de um perfil como na figura 24, é construída uma *tag cloud* (nuvem de etiquetas) com as *hashtags* utilizadas nas últimas 200 mensagens enviadas à rede por um usuário:



**Figura 24:** Tag cloud das hashtags de @ladygaga

Na figura 24, observa-se que algumas *hashtags* estão escritas com letras maiores que outras *hashtags*. Isto significa que #*giantsgame* e #*goblue* foram mais usadas que as demais, pois este é o padrão encontrado em nuvens de *tags* em diversos outros produtos da *Web 2.0* como os blogs.

O site *Tip Top Insight* se destina à análise de *hashtags* e tópicos mais comentados do *Twitter*. Entre as diversas análises que o site realiza a que mais interessa a este estudo é a análise do impacto das *hashtags* e tópicos, em outras palavras, o site classifica tópicos e *hashtags* como positivas, negativas e neutras, de acordo com o conteúdo da mensagem.



**Figura 25:** Impacto positivo, negativo e neutro dos tópicos e hashtags mais recentes

A figura 25 apresenta as *hashtags* e tópicos em discussão no momento que a página foi recuperada. Os tópicos em vermelho se referem à mensagens de conteúdo negativo sobre o assunto. As que receberam a cor amarela são consideradas tópicos neutros, ou seja, não são feitas afirmações de cunho negativo ou positivo sobre o assunto. Os tópicos na cor verde estão ligados à mensagens de conteúdo positivo sobre o assunto.

Classificar as *hashtags* e os tópicos mais comentados como positivos, negativos ou neutros evidencia a busca de sentido na interpretação dos enunciados. Silveira analisou *hashtags* que foram utilizadas durante um acidente sofrido pelo então candidato à presidência da república José Serra, identificando três eixos semânticos distintos:

Em um primeiro eixo, enquadramos os enunciados que fazem referência aos discursos do campo político eleitoral, ou seja, aqueles que retomam outros discursos do campo político.[...]

Em um segundo eixo, destacamos os enunciados que fazem referência ou colocam em pauta uma reflexão sobre o fazer midiático nos processos eleitorais do Brasil. [...]

E, por fim, em um terceiro eixo agrupamos os enunciados que, a partir das inúmeras repetições, e circulação das duas *hashtags* [#boladepapel facts e #serrarojas], produziu um deslocamento dos sentidos produzidos no campo político e midiático, invertendo os sentidos, ao relacionar os enunciados políticos e midiáticos um aspecto irônico e humorístico, característico dos discursos que circulam no Twitter. Os sentidos produzem-se em uma relação interparafrástica com os enunciados

semelhantes aos destacados acima, mas, agora, mesclam os sentidos de denúncia aos discursos políticos e midiáticos tradicionais, com seus sentidos estabilizados, com sua ordem, com sua recorrência, provocando uma ruptura nos discursos típicos das campanhas eleitorais.[...] (SILVEIRA, 2011, p.3)

O próximo site analisado foi *tweetalarm* (<http://tweetalarm.com/>) que realiza buscas de mensagens contendo determinada *hashtags*. Além da busca, há a opção de configurar um alerta para o usuário ser avisado sempre que a *hashtag* cadastrada for citada em alguma mensagem. É um meio de acompanhar um assunto no universo dos diversos assuntos tratados pelas mensagens todos os dias.

Ouro site que se destina a acompanhar as *hashtags* é o *TweetChat*. A partir da busca por uma *hashtag*, são recuperadas as mensagens mais recentes que a contém. Este diferencia-se do site *Twitag* por promover a interação entre usuários a partir de uma *hashtag*:



**Figura 26:** Exemplo de interação através das hashtags no TweetChat

Na figura 26, encontram-se algumas mensagens contendo a *hashtag* #SouFelizPor e ao lado de cada mensagem há a opção de visualizar o perfil do usuário e respondê-lo, numa espécie de conversa, ou *Chat*.

O site *Twitoaster* (<http://twitoaster.com/>) destina-se a promover a interação entre os usuários a partir dos assuntos discutidos por eles. O site atuava (não é atualizado desde março de 2011) como uma ferramenta de gerenciamento de perfil, auxiliando o usuário a melhorar a forma como se comunica com os outros. Para isso, o site arquivava as conversas em tempo real, fornecendo estatísticas e o contexto das conversações em todo mundo através da divulgação das principais *hashtags* na localidade do usuário, como na figura 27.

The screenshot shows the Twitoaster website interface. At the top, there is a header with the Twitoaster logo and navigation links like 'Home | Twitterers' and 'Log in with twitter'. Below the header, there are several informational sections: 'What is Twitoaster?', 'About Twitoaster', 'Press Page', 'Twitoaster API', and 'WordPress Plugin'. The main content area features a section titled 'Hot Twitter conversations in Brazil' which lists several tweets with their respective engagement metrics (replies, retweets, links). On the right side, there is a sidebar with a search bar, 'Hot Users in Brazil', 'Most active Twitter Users in Brazil', 'Hot Tweets Worldwide', 'Hot Hashtags', and 'About Twitoaster'.

Figura 27: Twitoaster

O desenvolvimento de sites para viabilizar o acompanhamento dos assuntos discutidos no *Twitter* reflete a dificuldade em acompanhar todas as mudanças que surgem quando um novo evento acontece. O que se observa é uma situação de caos. Como discutido no referencial teórico, o caos se sustenta na alteração de uma ordem até estabelecer uma nova ordem, que no caso dos sites *Twitoaster*, *TweetChat* e *Teetalarm* ocorre através do monitoramento das *hashtags* de interesse de um usuário.

O site *monitter.com* destina-se à recuperação e ao monitoramento de novas mensagens contendo uma *hashtag*. À medida que uma nova mensagem contendo a *hashtag* pesquisada é enviada ao *Twitter*, a mesma é recuperada pelo site e organizada em ordem cronológica (as mais recentes no topo). Várias buscas podem ser realizadas ao mesmo tempo, pois o site se divide em colunas para separar os resultados.

monitter tweets containing #rio2016 Add Column Swap Colours Load Trends Get Widgets Contact Tweet Connect with twitter

#rio2016 Paused 9 new

RT @iocmedia: <http://t.co/ePKZghZz> First Rio 2016 sports venue opens to the public #olympics #rio2016  
kcraFitz 12:32 web

RT @iocmedia: <http://t.co/ePKZghZz> First Rio 2016 sports venue opens to the public #olympics #rio2016  
silvia\_munoz 12:30 web

<http://t.co/ePKZghZz> First Rio 2016 sports venue opens to the public #olympics #rio2016  
iocmedia 12:29 web

RT @Rio2016 en: Oscar Niemeyer visits the new Sambódromo, stage of carnival and a Rio 2016™ venue <http://t.co/yZbKr6Bb> #Rio2016  
pgamafilho 11:09 web

IT Infrastructure Specialist - #Rio2016 <http://t.co/44ABdMqL> #olympics  
jumpinsport 6:58 Tweet Button

@MulattoMamba ima kill u and nuke the shit out of #rio2016  
cchang2013 0:47 web

So proud of Tilly being on tv #OlympicBoxer #Rio2016  
Roonzz 21:02 Twitter for iPhone

@Rob Colley @dave\_runcie @cm\_stevenson @kingstonlufbra @alexleishman I aim to make it an Olympic sport #rio2016  
PatrickWareing 20:51 Twitter for iPhone

Had the crowd mad at 300m. #worldchamps2015 #rio2016, you know @lewismooore22  
evansbuck 17:33 Twitter for iPhone

**Figura 28:** Busca por #rio2016 no site monitter.com

Na Figura 28, foi realizada uma busca a partir da *hashtag* #rio2016. As mensagens recuperadas são listadas e ordenadas da mais recente para a menos recente sendo atualizada a cada nova mensagem enviada ao *Twitter* contendo a *hashtag* pesquisada. O site disponibiliza a opção de pausar a busca por novas atualizações.

O site *Statweestics* (<http://statweestics.com/>) fornece estatísticas de popularidade de diversos aspectos do *Twitter*. Estatísticas por localidade, palavras, usuários, língua e pela forma como as mensagens foram enviadas ao site (o site chama esse ranking se *source*) são alguns dos exemplos. As estatísticas de popularidade de *hashtags* são mostradas por períodos: agora, hoje, esta semana, este mês e no ano. Na Figura 29 é mostrado o ranking das 25 *hashtags* mais populares no ano de 2011:



**Figura 29:** Ranking de 2011 no statweestics

Na Figura 29 é mostrado o ranking de *hashtags* acompanhadas do percentual e de setas que indicam se a popularidade aumentou, se manteve estável ou obteve queda de popularidade. Na figura, é possível visualizar as 25 *hashtags* mais citadas nas mensagens, mas o site contabiliza as 500 *hashtags* mais citadas.

O site *twimmer.com* (<http://twimmer.com/>) também é destinado à divulgação de rankings de *hashtags* e tópicos, direcionando-se para as mensagens de usuários localizados na Holanda.

**TWIMMER.com**

De top 100 Nederlandse "trending topics" op Twitter

Het gesprek van de dag in twitterend Nederland  
Ongefilterd en ongecensureerd. Doe er je voordeel mee!

TRENDS TWEEPS

Nieuwsgierig... wat je met je twitter account kunt doen? [Klik hier](#)

Respect! 23 Like 64 +1 8

**PASSAGENS**  
Buenos Aires A PARTIR DE USD 299  
decolar.com  
O MELHOR PREÇO PARA SUA VIAGEM

Laatste uur	Laatste 4 uur	Laatste 24 uur	Laatste 7 dagen	Laatste 14 dagen	Laatste 30 dagen
1. #fctwente	1. #vacature	1. #twexit	1. #twexit	1. #tvoh	1. #tvoh
2. #vacature	2. #fctwente	2. #fctwente	2. #middleschoolmemo	2. #twexit	2. #twexit
3. #boysengirls	3. #ajax	3. #np	3. #np	3. #linkdedownie	3. #fctwente
4. #lachjerot	4. #dtv	4. #oomf	4. #feyaja	4. #dtv	4. #dtv
5. #dtv	5. #np	5. #vacature	5. #oomf	5. #swaggkoning	5. #rt
6. #np	6. #rt	6. #dtv	6. #highschoolmemo	6. #rt	6. #widm
7. #zpp	7. #linkdedownie	7. #linkdedownie	7. #linkdedownie	7. #ajaaz	7. #swaggkoning
8. #ajax	8. #durtevrage	8. #ghettociroclavors	8. #ff	8. #vacature	8. #3fm
9. #veenendaal	9. #nosop3	9. #twitteroff	9. #dtv	9. #np	9. #wetrippymane
10. #twexit	10. #nieuws	10. #ajax	10. #rt	10. #wesleepymane	10. #utrecht
11. #wateenmeisjewil	11. #ede	11. #koud	11. #fctwente	11. #utrecht	11. #nosop3
12. #mijnlevensweg	12. #utrecht	12. #rt	12. #vacature	12. #3fm	12. #linkdedownie
13. #ede	13. #in	13. #nf	13. #twitteroff	13. #durtevrage	13. #win
14. #033	14. #3fm	14. #gtst	14. #nf	14. #gtst	14. #durtevrage
15. #ict	15. #ict	15. #ede	15. #swaggkoning	15. #oomf	15. #fo

7animoveis ENCONTRE JÁ SEU IMÓVEL EXCLUSIVO

Figura 30: Hashtags e tópicos na Holanda no twimmer.com

A Figura 30 apresenta as colunas que informam ao público as 15 *hashtags* e tópicos mais populares registrados no último dia (*Laatste uur*); nas últimas 4 horas (*Laatste 4 uur*); últimas 24 horas (*Laatste 24 uur*); Últimos 7 dias (*Laatste 7 dagen*); Últimos 14 dias (*Laatste 14 dagen*); Últimos 30 dias (*Laatste 30 dagen*).

O site *buzzter* (<http://buzztter.com/>) fornece um quadro com os tópicos mais discutidos do momento e também ferramenta para buscar atualizações de status (como também são chamadas as mensagens no *Twitter*).

**buzztter**  
Snapshot the tweets!

English | Japanese | mobile

iPhone push notification now available. Install [boxcar](#) and subscribe buzztter: [PC](#) or [iPhone](#) (experimental and currently Japanese only).

#WeAllHaveThat Are you tired of ... If you want to ea... thecashjournal .com Tired of working ... #ThingsThatGetMeU... Blake Griffin deadline day #SE7ENcomeback #weallhavethat fr... Kendrick Perkins **awkward moment when** moment when you one friend Super Bowl on loan #Pisces February #Aries oomf January #Aquarius dunk Sometimes the gir... #Capricorn #Sagittarius take a look think they're friend that transfer love being of your life friend who to earn this class Virgo one person Taurus More for tired of days to Tuesday Perkins Scorpio care of to fight Gemini be in the will go silent this morning in class feelings the end of the of being days of You have to because you that can warm trending breakfast classes wake up Blake to be in hard to partners another day March

Try this query: #WeAllHaveThat OR "Are you tired of never having enough money" OR "If you want to earn more money you should start an online business on the side"

Recent updates including buzz-phrases

**lil\_hause** "@WomensHumor: #WeAllHaveThat one night we wish we could remember." (less than a minute ago)

**SJ\_says** RT @WomensHumor: #WeAllHaveThat one friend that we love to get in trouble with! (less than a minute ago)

**AtotB617** #WeAllHaveThat morning sex crave (less than a minute ago)

**veronicawaller** @honeydew87 Are you tired of never having enough money? Do what I did and start an online business. thecashjournal .com (less than a minute ago)

**marianmoran3** @TerraTyler Are you tired of never having enough money? Do what I did and start an online business. thecashjournal .com (less than a minute ago)

**ceciliaatkins01** @Ceee\_TheGreat Are you tired of never having enough money? Do what I did and start an online business. thecashjournal .com (less than a minute ago)

amazon.co.jp

エアカウンター8  
エスケー  
新品 ¥5,475  
ベストプライス  
¥5,980

【Amazon.co.jp限定】ZEPEAL セピーアル  
ZEPEAL (ゼピーアル)  
新品 ¥1,980  
ベストプライス  
¥1,980

【日本特許】家庭用  
掃除機  
タニタ  
新品 ¥2,550  
ベストプライス  
¥2,550

ブラウン オーラ  
ルB 電動歯ブラシ  
Braun (ブラウン)  
新品 ¥1,200  
ベストプライス  
¥880

Figura 31: Página inicial do site buzztter

Na Figura 31 destaca-se a busca com o uso do operador booleano OR e das aspas para realizar a busca. Além disso, há uma breve explicação de como construir uma busca usando os operadores booleanos. Os operadores booleanos já são utilizados em diversos SRI's e bases de dados para aumentar a precisão dos resultados, como, por exemplo, a base de publicações científicas *Web Of Science*.

### SÍNTESE DOS RESULTADOS:

A pesquisa realizada recuperou 120 sites destinados à complementação das mais diversas ações que podem ser realizadas no *Twitter*. Essa quantidade é um recorte pequeno de um universo que aumenta a cada dia, mas a partir da amostra obtida já foi possível a identificação de 12 classes de sites: Produtos; Posts de blogs; Alertas; Busca; Compartilhamento; Customização; Estatísticas; Guias; *Hashtags*; Perfil; Mensagens; e Outros. Isso representa que o desenvolvimento desses sites identifica algumas necessidades dos usuários que não são atendidas pelo *Twitter*.

A descrição dos 17 sites para complementar o uso de *hashtags* identificou alguns exemplos de aspectos das *hashtags* que precisam ser complementados:

1. A necessidade de compreensão do significado das *hashtags* é revelada pela criação de sites para definição de *hashtags*: *tagdef.com*; pelo site de definição de termos do *Twitter*: *Twittonary*; e também pelo site de avaliação de conteúdo como positivo, negativo ou neutro: *Tip Top Insight*.
2. O acompanhamento de mensagens classificadas com determinada *hashtag* através dos sites de monitoramento de mensagens *monitter.com*; *tweetalarm*; *TweetChat*; e *Twitoaster*.
3. A necessidade em conhecer a popularidade de uma *hashtag* através dos sites *Statweestics*; *Twimmer.com*; *My Tweet Cloud*; e *Hashtagbattle*.
4. A busca e a recuperação de mensagens através dos sites *HashParty*; *Twitag*; *Hashalbum*; e *Buzzter*.
5. O agrupamento de *hashtags* por categorias como no site *Hashtags.org* e no blog *Hashtags's Blog*.

## 5 CONSIDERAÇÕES FINAIS

A computação nas nuvens transporta para o ambiente virtual ações que antes eram restritas a certos locais e pessoas, inaugurando a era da *Web 2.0* que se baseia na colaboração para o desenvolvimento de diversos aplicativos e sites para troca de informação na internet. Nesse contexto de colaboração surgem as redes sociais online, numa tentativa de reproduzir as relações sociais no ciberespaço.

As redes sociais na Internet se desenvolveram para além da reprodução das relações sociais existentes fora do ciberespaço. Atualmente, as redes sociais se consolidam como canais para a ampla circulação de informação expressa nas mais diversas formas (vídeo, foto, mensagens, etc.). A diversidade de informação divulgada nessas redes estimulou o desenvolvimento de instrumentos de organização baseados na folksonomia que se utilizam de *tags* para classificação de conteúdos informacionais.

A presente pesquisa explorou a natureza caótica da dinâmica de troca de mensagens na rede social *Twitter* e o uso das *hashtags* para a classificação das mensagens como um meio de estabelecer alguma ordem ao sistema. Como as *hashtags* se popularizaram na rede após o surgimento espontâneo por iniciativa de um usuário.

O desenvolvimento do trabalho contemplou a investigação, com base na literatura, da classificação enquanto processo inato do ser humano e enquanto sistema para organização da informação; da classificação colaborativa baseada na folksonomia, a chamada *Web 3.0*; das relações da teoria do caos com a dinâmica da rede social *Twitter*.

A análise da literatura sobre classificação revelou que a organização da informação nas redes sociais segue uma lógica inerente ao hipertexto, ou seja, os instrumentos para classificação e recuperação da informação nesse espaço estão associados às características básicas da *Web 2.0*: a interação e a colaboração. Esses instrumentos, por sua vez, seguem algumas ideias que integram sistemas de classificação já formalizados que, ao longo da história, foram se adaptando aos diferentes contextos de produção e disseminação de informação e às necessidades de recuperação da informação por parte dos usuários.

Parte relevante da discussão teórica foi a identificação de elementos que conduzem o sistema a uma situação de caos. Cada mensagem enviada à rede caracteriza-se como um evento capaz de alterar a trajetória de uma discussão em curso, que somada à ausência de comunidades ou diretórios para agrupar usuários segundo seus interesses eleva o grau de incerteza na busca e recuperação de informação na rede. Observadas essas questões, fica ainda mais evidente o papel das *hashtags* para estabelecer alguma organização na rede.

O levantamento e análise das *hashtags* permitiu a identificação de três tipos de perfis de usuários: os usuários comuns que acessam o *Twitter* sem um objetivo específico; as personalidades que são celebridades que usam o *Twitter* para promoção de suas atividades profissionais; e as instituições que divulgam mensagens relacionadas à missão e aos produtos e serviços que desenvolvem. Sendo assim, as *hashtags* promovem não só a classificação de uma mensagem, mas também a classificação de perfis de usuários.

O grupo dos usuários comuns, caracterizado por perfis de pessoas anônimas revela que os usos de *hashtags* podem variar muito tanto em relação à seleção dos termos quanto ao alcance que se pretende com determinada classificação. Em resumo, a característica essencial foi a percepção de três tipos básicos de *hashtags*: as que estabelecem categorias como as utilizadas pelo usuário1, as que formam comunidades temporárias sobre determinado assunto como #QueroRomárioNoMinistério e as que representam mensagens de forma afetiva como as do Usuário5.

Embora a intenção inicial desta pesquisa fosse obter dados variados sobre personalidades que acessam o *Twitter*, o grupo das personalidades tornou-se o mais homogêneo, em termos de classes de mensagens muito em função de pertencerem ao mesmo ramo, o da Música Pop americana. Os três artistas utilizam *hashtags* que se dirigem ao seu público e à promoção de seus trabalhos. Como já esperado, utilizam muitas *hashtags* que formam comunidades temporárias.

O grupo de perfis institucionais apresentou *hashtags* bem direcionadas à missão e às principais atividades e projetos em andamento. Como não se tratam de perfis de pessoas, as chamadas *hashtags* afetivas ocorreram menos que nos grupos dos usuários comuns e das personalidades. Isso se deve ao fato de que quanto a precisão dos termos selecionados para classificar uma mensagem, maiores são as chances do produto ou atividade ser visto.

A análise das *hashtags* identificou ainda dois casos de polissemia: a *hashtag* #spurs – que pode representar o time de basquete norte-americano San Antonio Spurs ou o time de futebol do Reino Unido *Tottenham Hotspurs* (usuário5) e a *hashtag* #MEOW de Katy Perry que classifica mensagens que promovem um perfume, mas que também podem remeter à onomatopeia para o som emitido por gatos. A polissemia representa um obstáculo para a precisão na compreensão do sentido e também para a recuperação da informação, como já apontado na contextualização do tema.

O estudo das *hashtags* contemplou ainda a identificação de duas categorias fundamentais de Ranganathan: espaço e tempo. No presente estudo a identificação das categorias personalidade, matéria e energia não foi possível de ser efetuada na amostra por ser

composta de perfis selecionados aleatoriamente. Como observado na contextualização do tema, a dificuldade de identificação dessas categorias é comum na atividade de classificação.

Os problemas relativos à polissemia e às categorizações são o elo entre a amostra de *hashtags* e os 17 sites de apoio ao uso de *hashtags* descritos no estudo. Esses dois problemas representam o desejo de tornar a prática de atribuir *hashtags* a uma mensagem mais eficiente através do desenvolvimento de sites de apoio, isto é, o simples uso de uma *hashtag* não parece ser mais suficiente para classificar mensagens.

A análise dos 17 sites selecionados para a pesquisa revela a busca constante pela organização da informação no *Twitter*. O volume de mensagens e também a popularização do uso das *hashtags* estimulou o desenvolvimento de ferramentas de apoio à busca, para definição de palavras, para conhecer quais *hashtags* foram mais usadas. Criada de forma espontânea como recurso para classificação de mensagens, as *hashtags* na atualidade atuam como catalisadoras de ideias e pessoas.

A construção de sites para o uso de *hashtags* também está relacionada a pouca informação sobre diretrizes ou regras por parte do próprio *Twitter*. O tópico da ajuda sobre o assunto apresenta pouca informação sobre regras para construção de *hashtags*, embora forneça recomendações de boas práticas. Ao invés de abordar com detalhes as características das *hashtags*, a rede social opta por indicar sites onde mais informações podem ser encontradas, entre esses sites, encontra-se *Hashtags.org* que também faz parte do presente estudo.

A colaboração na construção das *hashtags* ficou evidente com a variedade de finalidades de uso. Foram encontrados sites para a busca e recuperação de *hashtags*; para conhecer a popularidade de uma *hashtag* através de *rankings*; avaliar *hashtags*; criar e pesquisar definições para *hashtags*; listar mensagens contendo determinada *hashtag* e etc. O desenvolvimento de sites e ferramentas para otimizar o uso das *hashtags*, em especial os sites que definem e avaliam *hashtags*, explicitam a busca pela padronização na escolha dos termos para representar as mensagens.

Os sites reforçam que o uso de *hashtags* para classificar mensagens no *Twitter* necessita cada vez mais da colaboração para que o termo selecionado para categorizar uma mensagem seja efetivo para a recuperação da informação tanto por parte do emissor da mensagem quanto pelo receptor. O surgimento espontâneo das *hashtags* no *Twitter* transformou-se em estratégia não apenas para organização da informação, mas também em estratégia de divulgação, promoção e convergência de saberes.

Em síntese, o uso das *hashtags* para a organização do caos da rede social *Twitter* é necessário, embora necessite da colaboração que acontece fora do sistema. A capacidade de auto-organização do sistema para a recuperação da informação através das *hashtags* já não é suficiente. Devido a algumas limitações, como o limite de 140 caracteres por mensagem e a ausência de regras que favoreçam a padronização dos termos, os sites e as ferramentas de apoio são cada vez mais necessárias. Essa situação, no entanto, pode ser modificada no futuro ou até mesmo na próxima atualização realizada no *Twitter*.

## REFERÊNCIAS

ALMEIDA, M.B.; BAX, M.P. Uma visão geral sobre ontologias: pesquisa sobre definições, tipos, aplicações, métodos de avaliação e de construção. **Ciência da Informação**, Brasília, DF, v. 26, n. 1. p. 39-45, set./dez. 2003.

ARANALDE, Michel Maya. Reflexões sobre os sistemas categoriais de Aristóteles, Kant e Ranganathan. **Ciência da Informação**, Brasília, DF, v. 38, n. 1, p. 86-108, jan./abr. 2009.

ARAÚJO, V. M. R. H. de. Sistemas de informação: nova abordagem teórico-conceitual. **Ciência da Informação**, Brasília, DF, v. 24, n.1, jan./abr. 1995. Disponível em: <<http://revista.ibict.br/ciinf/index.php/ciinf>>. Acesso em: 13 maio 2011.

AQUINO, Maria Clara. Hipertexto 2.0, folksonomia e memória coletiva: um estudo das *tags* na organização da web. **Revista E-Compós**, v. 9, p. 2-18, ago. 2007. Disponível em: <<http://www.compos.org.br/seer/index.php/e-compos/article/viewArticle/165>>. Acesso em: 15 out. 2009.

REITZ, Joan M. **Online Dictionary for Library and Information Science**, c2010. Disponível em: <[http://www.abc-clio.com/ODLIS/odlis\\_A.aspx](http://www.abc-clio.com/ODLIS/odlis_A.aspx)>. Acesso em: 13 jun. 2011.

BACON, Francis. **The advancement of learning**. London: Cassell & Company, 1893.

BARBOSA, Alice Príncipe. Classificações facetadas. **Ciência da Informação**, Rio de Janeiro, v. 1, n. 2, p. 73-81, 1972.

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

\_\_\_\_\_. A oferta e a demanda da informação: condições técnicas, econômicas e políticas. **Ciência da informação**, Brasília, DF, v. 28, n. 2, p. 168-173, maio/ago. 1999. Disponível em: <<http://www.alternex.com.br/~aldoibict/sensivel.htm>>. Acesso em: 6 out. 2009.

BELKIN, N. J. Information concepts of information for Information Science. **Journal of Documentation**, v. 34, n.1, p. 55-85, 1978.

BERNERS-LEE, Tim; CAILLIAU, Robert. World-Wide Web. In: \_\_\_\_\_. Computing in High Energy Physics, 1992, Annecy, France. Disponível em: <<http://www.freehep.org/chep92www.pdf>>. Acesso em: 28 set. 2009.

BIOLCHINI, Jorge Calmon Almeida. Semântica e cognição em bases de conhecimento: do vocabulário controlado à ontologia. **DataGramZero**, v.2, n.5, 2001. Disponível em: <<http://www.dgz.org.br>>. Acesso em: 25 jun. 2011.

BORST, W. N. **Construction of engineering ontologies**. 1997. Tese (Doutorado). Disponível em: <<http://www.ub.utwente.nl/webdocs/inf/1/t0000004.pdf>>. Acesso em: 21 abr. 2009.

BRAGA, G. Informação, Ciência da Informação: breves reflexões em três tempos. **Ciência da Informação**, Brasília, DF, v. 24, n. 1, p. 84-88, jan./abr. 1995.

BUSH, V. As we may may think. **Atlantic Monthly**, v.176, n.1, p. 101-108, 1945.

CAMPOS, M. L.; GOMES, H. E. Taxonomia e Classificação: princípios de categorização. **DataGramaZero**, v. 9, n. 4, p. 1, ago. 2008. Disponível em: <[http://www.dgz.org.br/ago08/Art\\_01.htm](http://www.dgz.org.br/ago08/Art_01.htm)>. Acesso em: 18 jul. 2010.

CAÑADA, J. **Tipologias y estilos en el etiquetado social**. Disponível em: <<http://www.terremoto.net/tipologias-y-estilos-en-el-etiquetado-social/>>. Acesso em: 29 set. 2011.

CAPRA, Fritjof. **A teia da vida: uma nova compreensão científica dos sistemas vivos**. São Paulo: Cultrix, 1996.

CARTER, Simon; TSAGKIAS, Manos; WEERKAMP, Wouter . Twitter hashtags: Joint Translation and Clustering. In: ACM WEBSCI,11, 2011, Koblenz, Germany. **Proceedings**. p. 1-3. Disponível em: <[http://www.websci11.org/fileadmin/websci/Posters/125\\_paper.pdf](http://www.websci11.org/fileadmin/websci/Posters/125_paper.pdf)>. Acesso em: 30 nov. 2011.

CENDÓN, B. V. Ferramentas de busca na Web. **Ciência da Informação**, DF, v. 30, n. 1, jan./abr. 2001. 1995. Disponível em: <<http://revista.ibict.br/ciinf/index.php/ciinf>>. Acesso em: 06 jun. 2011.

CHANG, Hsia C. A new perspective on twitter hashtag use: diffusion of innovation theory. **Proceedings of the American Society for Information Science and Technology**, v. 47, n. 1, p. 1-4, 2010.

CHAUÍ, Marilena. **Convite à filosofia**. São Paulo: Ática, 1995.

DAAS, P.J.H., VAN DE VEN, M., ROOS, M. **Twitter as a data source for official statistics: first result**. In: GENERAL ONLINE RESEARCH CONFERENCE, 11, Düsseldorf, Germany. 2011. Disponível em: <[http://www.pietdaas.nl/beta/pubs/pubs/Twitter\\_GOR2011\\_pres.pdf](http://www.pietdaas.nl/beta/pubs/pubs/Twitter_GOR2011_pres.pdf)>. Acesso em: 2 dez. 2011.

DAHLBERG, Ingetrault. Teoria do conceito. **Ciência da Informação**, Rio de Janeiro, v. 7, n. 2, p. 101-107, 1978.

\_\_\_\_\_. Teoria da Classificação, ontem e hoje. In: CONFERÊNCIA BRASILEIRA DE CLASSIFICAÇÃO BIBLIOGRÁFICA. **Anais**. Rio de Janeiro : IBICT, 1979. p. 352-370. Disponível em: <[http://www.conexaorio.com/biti/dahlbergteoria/dahlberg\\_teoriam.htm](http://www.conexaorio.com/biti/dahlbergteoria/dahlberg_teoriam.htm)>. Acesso em: 2 jan. 2011.

DERNTL, Michael *et al.* Inclusive social tagging and its support in Web 2.0 services. **Computers in Human Behavior**, v. 27, n. 4, p. 1460–1466, 2011. Disponível em: <<http://www.sciencedirect.com/science/article/pii/S0747563210002906>>. Acesso em: 14 dez. 2011.

DESMET, I. Questões de semântica em terminologia: a problemática da definição terminológica. **Terminologias**, n. 2, p. 4-21, dez. 1990.

DODEBEI, Vera Lúcia Doyle. **Tesauro**: linguagem de representação da memória documentária. Niterói: Intertexto; Rio de Janeiro: Interciência, 2002.

DURKHEIM, Émile. **Fato social e divisão do trabalho**. São Paulo : Ática, 2007.

ELIAS, Norberto. **A sociedade dos indivíduos**. Rio de Janeiro: Jorge Zahar, 1994.

ENCICLOPÉDIA Einaudi. Porto, Portugal: Imprensa Nacional; Casa da Moeda, c1985. v. 41.

FAULSTICH, Enilde. Socioterminologia: mais que um método de pesquisa, uma disciplina. **Ciência da Informação**, Brasília, DF, v. 24, n. 3, p. 281-288, set./dez. 1995.

FLICK, Uwe. **Desenho da pesquisa qualitativa**. Porto Alegre: Artmed, 2009.

FONSECA, Edson Nery da. Apogeu e declínio das classificações bibliográficas. **BITI**, Rio de Janeiro, [200-?]. Disponível em: <[www.conexaorio.com/bit/neri/index.htm](http://www.conexaorio.com/bit/neri/index.htm)>. Acesso em: 6 mar. 2011.

GIL, A. C. **Como elaborar projetos de pesquisa**. São Paulo: Atlas, 2002.

GOMES, Hagar Espanha; CAMPOS, Maria Luiza de Almeida. Tesauro e normalização terminológica: o termo como base para o intercâmbio de informações. **DatagramaZero**, dez. 2004. Disponível em: <http://www.datagramazero.org.br>. Acesso em: 29 jun. 2011.

GOUADEC, D. **Terminologie**: constitution des données. Paris: AFNOR, 1990.

GRABOWSKI, A.; KRUSZEWSKA, N.; KOSIŃSKI, R. A. Properties of on-line social systems. **Eur. Phys. J. B**, v. 66, n. 1, p. 107–113, 2008. Disponível em: <<http://dx.doi.org/10.1140/epjb/e2008-00379-2>>. Acesso em: 15 nov. 2011.

GRACIOSO, L. S. Biblioteca, Web 2.0, Biblioteca 2.0. In: SEMINÁRIO INTERNACIONAL DE BIBLIOTECAS DIGITAIS, 2007, São Paulo. **Anais**. São Paulo: Seminário Internacional de Bibliotecas Digitais, 2007. Disponível em: <<http://libdigi.unicamp.br/document/?code=23460>>. Acesso em: 6 jan. 2011

HALL, J.L. Online retrieval history; how it all began: some personal recollections. **Journal of Documentation**, v. 67, n. 1, p. 182-193, 2011. Disponível em: <<http://dx.doi.org/10.1108/00220411111105506>>. Acesso em: 20 dez. 2011.

HELBING, D. et al. Understanding, creating, and managing complex techno-socio-economic systems: Challenges and perspectives. **Eur. Phys. J. Special Topics**, v. 195, n. 1, p. 165–186, 2011.

HOUAISS, Antônio. **Dicionário Houaiss de Língua Portuguesa**. Disponível em: <<http://houaiss.uol.com.br/busca.jhtm?verbete=classificar&stipe=k>>. Acesso em: 6 maio 2011.

HUBERMAN, B. A.; ROMERO, D. M.; WU, F. Social networks that matter: Twitter under the microscope, 2008. Disponível em: [http://papers.ssrn.com/sol3/papers.cfm?abstract\\_id=1313405](http://papers.ssrn.com/sol3/papers.cfm?abstract_id=1313405). Acesso em: 12 maio 2011.

JANNUZZI, Anna Haydée Lanzillotti; AMORIM, Rita de Cássia Rocha; SOUZA, Cristina Gomes de. Implicações da categorização e indexação na recuperação da informação tecnológica contida em documentos de patentes. **Ciência da Informação**, Brasília, DF, v. 36, n. 2, p. 27-34. ago. 2007. Disponível em: <<http://dx.doi.org/10.1590/S0100-19652007000200003>>. Acesso em: 20 fev. 2011.

JAKOBSEN, Charles Seger. **Collaborative tagging and folksonomies: user driven knowledge sharing**, 2007. Disponível em: <<http://pure.iva.dk/en/publications/collaborative-tagging-and-folksonomies-user-driven-knowledge-sharing%28a96d0230-3ec0-11de-8213-000ea68e967b%29.html>>. Acesso em: 14 out. 2010.

KRIEGER, Maria da Graça; FINATTO, Maria José Bocorny. **Introdução à terminologia: teoria e prática**. São Paulo: Contexto, 2004.

KUMAR, K. Categories. In: \_\_\_\_\_. **Theory of Classification**. 2. ed. Nova Delhi: Vikas Publishing House, 1981. p. 246-258.

LAKOFF, George. From Wittgenstein to Rosch. In: \_\_\_\_\_. **Women, fire, and dangerous things: what categories reveal about the mind**. Chicago: University of Chicago Press, 1987. p. 12-57.

LANGRIDGE, Derek. **Classificação: abordagem para estudantes de Biblioteconomia**. Rio de Janeiro: Interciência, 2006.

LARA, M.L.G. Diferenças conceituais sobre termos e definições e implicações na organização da linguagem documentária. **Ciência da Informação**, Brasília, DF, v. 33, n. 2, p. 91-96, 2004.

LEMOS, Renata. Qotd, por @umairh: a inteligência coletiva no Twitter. **Revista Galáxia**, São Paulo, n. 19, p. 226-239, jul. 2010.

LÉVY, Pierre. **A inteligência coletiva: por uma antropologia do ciberespaço**. São Paulo: Loyola, 1998.

LIMA, Gercina Ângela Borém. Interfaces entre a Ciência da Informação e a Ciência Cognitiva. **Ciência da Informação**, Brasília, DF, v.32, n.1, p. 77-87, 2003.

LORENZ, E. N. Deterministic nonperiodic flow. **Journal of Atmospheric Science**, v.20, p.130-141, 1963.

MAIER, R.; THALMANN, S. Institutionalised collaborative tagging as an instrument for managing the maturing learning and knowledge resources. **International Journal of Technology Enhanced Learning**, v.1, n. 1/2, p. 70-84, 2008.

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

MAUSS, Marcel. Ensaio sobre a dádiva: forma e razão da troca nas sociedades arcaicas. In: \_\_\_\_\_ **Sociologia e antropologia**. São Paulo: Cosac & Naify, 2003.

MEIRELLES, Junia Cristina J. P.; MOURA, Mônica. Web 2.0: novos paradigmas projetuais e informacionais. **InfoDesign**, v. 4, n. 2, p. 12-19, 2007. Disponível em: <[http://www.infodesign.org.br/conteudo/artigos/34/ing/InfoDesign\\_v4\\_n2\\_2007\\_02\\_Meirelles\\_Moura.pdf](http://www.infodesign.org.br/conteudo/artigos/34/ing/InfoDesign_v4_n2_2007_02_Meirelles_Moura.pdf)>. Acesso em: 13 out. 2009.

MENDONÇA, E. S. A lingüística e a ciência da informação: estudos de uma interseção. **Ciência da Informação**, Brasília, DF, v.29, n.3, p.50-70, set./dez. 2000. Disponível em: <<http://www.scielo.br/pdf/ci/v29n3/a06v29n3.pdf>> . Acesso em: 20 maio 2011.

MILGRAM, S. The small world problem. **Psychology Today**, v. 1, n. 1, p. 61-67, May 1967.

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

MIRANDA, Antonio; SIMEÃO, Elmira; MUELLER, Suzana. Autoria coletiva, autoria ontológica e intertextualidade: aspectos conceituais e tecnológicos. **Ciência da Informação**, Brasília, DF, v. 36, n. 2, ago. 2007. Disponível em: <[http://www.scielo.br/scielo.php?script=sci\\_arttext&pid=S0100-19652007000200004&lng=en&nrm=iso](http://www.scielo.br/scielo.php?script=sci_arttext&pid=S0100-19652007000200004&lng=en&nrm=iso)> . Acesso em: 13 nov. 2009.

MIRANDA, Marcos Luiz Cavalcanti de. **Organização e representação do conhecimento: fundamentos teórico-metodológicos na busca e recuperação da informação em ambientes virtuais**. 2005. 353 f. Tese (Doutorado em Ciência da Informação) – Convenio MCT/IBICT-UFRJ/ECO, Rio de Janeiro, 2005.

MOLLETT, Amy; MORAN, Danielle; DUNLEAVY, Patrick. **Using twitter in university research, teaching and impact activities**. London: LSE Public Policy Group, 2011. Disponível em: <<http://eprints.lse.ac.uk/id/eprint/38489>>. Acesso em: 10 dez. 2011.

MORSE, J. M. Approaches to qualitative-quantitative methodological triangulation. **Nursing Research**, v.40, n.1, p. 120-132, 1991.

MÜLLER-PROVE, M. Model und Anwendungsperspektiven des Social Tagging. In: GAISER, B.; HAMPEL, T.; PANKE, S. (eds.). **Good tags - Bad tags**. München: Waxman, 2008.

NAAMAN, M.; BECKER, H.; GRAVANO, L. Hip and trendy: characterizing emerging trends on Twitter. **JASIST**, v. 62, n.2, p. 902–918, 2011.

NOCERA, Antonino; URSINO, Domenico. An approach to providing a user of a “social folksonomy” with recommendations of similar users and potentially interesting resources. **Knowledge-Based Systems**, v. 24, n. 8, p. 1277–1296, 2011.

O'REILLY, Tim. What is Web 2.0. Disponível em: <<http://oreilly.com/web2/archive/what-is-web-20.html>>. Acesso em: 10 out. 2009.

\_\_\_\_\_. Web 2.0 compact definition: trying again. Disponível em: <<http://radar.oreilly.com/archives/2006/12/web-20-compact-definition-tryi.html>>. Acesso em: 4 nov. 2009.

ORKUT. <[www.orkut.com/](http://www.orkut.com/)>. Acesso em: 12 abr. 2011.

PEREIRA, Débora de Carvalho; CRUZ, Ruleanderson do Carmo. Folksonomia e tags afetivas: comunicação e comportamento informacional no Twitter. **DatagramaZero**, v. 11, n. 6, dez. 2010. Disponível em: <[http://www.dgz.org.br/dez10/Art\\_06.htm](http://www.dgz.org.br/dez10/Art_06.htm)>. Acesso em: 9 fev. 2011.

PICKLER, Maria Elisa Valentim. Web semântica: ontologias como ferramentas de representação do conhecimento. **Perspectiva em Ciência da Informação**, Belo Horizonte, v. 12, n. 1, p. 65-83, jan./abr. 2007.

PIEIDADE, Maria Antonieta Requião. **Introdução à teoria da classificação**. 2. ed. rev. aum. Rio de Janeiro: Interciência, 1983.

PINHEIRO, L. V. R.; LOUREIRO, J. M. M. Traçados e limites da Ciência da Informação. **Ciência da Informação**, Brasília, v. 24, n. 1, p. 42-53, 1995.

PÓVOA, Marcello. O que é Web 2.0? Disponível em: <<http://webinsider.uol.com.br/index.php/2006/10/30/o-que-e-web-20/>>. Acesso em: 7 nov. 2009.

PRIGOGINE, Ilya. **As leis do caos**. São Paulo: Editora UNESP, 2002.

PRIMO, Alex. As tags no twitter como informação contextual de afeto. **Dossiê Alex Primo** (blog científico), 9 mar. 2010. Disponível em: <[http://www.interney.net/blogs/alexprimo/2010/03/09/as\\_tags\\_no\\_twitter\\_como\\_informacao\\_conte/](http://www.interney.net/blogs/alexprimo/2010/03/09/as_tags_no_twitter_como_informacao_conte/)>. Acesso em: 10 nov. 2010.

RAUPP, F. M.; BEUREN, I. M. Metodologia da pesquisa aplicada às ciências sociais. In: BEUREN, I. M. (Org.). **Como elaborar trabalhos monográficos em contabilidade**. 3. ed. São Paulo: Atlas, 2006.

RECUERO, Raquel; ZAGO, Gabriela. "RT, por favor": considerações sobre a difusão de informações no Twitter. **Revista Fronteiras: estudos midiáticos**, v. 12, n. 2, maio/ago., 2010. Disponível em: <<http://www.frenteiras.unisinos.br/pdf/88.pdf>>. Acesso em: ago. 2011.

REZENDE, S. M. A dança dos spins. **Ciência Hoje**, v. 14, n. 80, p. 29-32, 1992.

RIPELL, Jair; RECH, Simone T. **Teoria do Caos – Atratores Estranhos**. Disponível em: <[hermes.ucs.br/ccet/deme/emsoares/esac\\_2004\\_4/artigos/teoria\\_caos.doc](http://hermes.ucs.br/ccet/deme/emsoares/esac_2004_4/artigos/teoria_caos.doc)>. Acesso em: 18 maio 2011.

RODRIGUES, Sabrina Páscoli. Uma contribuição para o ensino da Sistemática na sala de aula: relato de experiência sobre a classificação dos animais de Aristóteles e Linné. **História da Ciência e Ensino**: construindo interfaces, v. 2, p. 89-97, 2010. Disponível em: <<http://revistas.pucsp.br/index.php/hcensino/article/view/4156/2866>>. Acesso em: 15 nov. 2011.

RONDEAU, G. **Introduction à la terminologie**. Québec: Gaëtan Morin, 1984.

SANTINI, Rose Marie; SOUZA, Rosali Fernandez de. Classificação colaborativa de conteúdos não-textuais na internet: as novas formas de mediação e organização da informação da Música através da folksonomia. In: ENCONTRO NACIONAL DE PESQUISA EM CIÊNCIA DA INFORMAÇÃO, 9, 2010, Rio de Janeiro. Rio de Janeiro: [s. n.], 2010. Disponível em: <<http://congresso.ibict.br/index.php/enancib/xienancib/paper/download/261/222>>. Acesso em: 2 dez. 2010.

SANTINI, Rose Marie. **Os usuários e a desorganização da cultura**: os sistemas de recomendação e as consequências da classificação para os usos sociais da música na Internet. Rio de Janeiro, 2010. 480 p. Tese (Doutorado em Ciência da Informação) - IBICT/UFF.

SATIJA, M. P. Classification: some fundamentals, some myths, some realities. **Knowledge Organization**, v. 25, n. 1/2, p. 32-35, 1998.

SANTOS, Glauber Eduardo de Oliveira. **Cálculo amostral**: calculadora on-line. Disponível em: <<http://www.calculoamostral.vai.la>> . Acesso em: 22 nov. 2011.

SCHRÖDINGER, Erwin. **What is Life?** The physical aspects of the living cell with mind and matter and autobiographical sketches. Cambridge: Cambridge University Press, 1995.

SILVEIRA, Juliana da. O visível e o invisível no funcionamento discursivo do Twitter. In: SEMINÁRIO DE ESTUDOS EM ANÁLISE DO DISCURSO, 5. O acontecimento do discurso: filiações e rupturas, Porto Alegre, 20-23 set 2011. Disponível em: <<http://www.discurso.ufrgs.br/anaisdosead/5SEAD/SIMPOSIOS/JulianaDaSilveira.pdf>>. Acesso em: 30 nov. 2011.

SIXDEGREES. <<http://sixdegrees.com>>. Acesso em: 12 dez. 2011.

SOUSA, Flávio R. C.; MOREIRA, Leandro O.; MACHADO, Javam C. **Computação em nuvem: conceitos, tecnologias, aplicações e desafios**. Disponível em: <[www.ufpi.br/ercemapi/arquivos/file/minicurso/mc7.pdf](http://www.ufpi.br/ercemapi/arquivos/file/minicurso/mc7.pdf)>. Acesso em: 25 dez. 2011.

SOUZA, Rosali Fernandez de. A classificação como interface da Internet. **DataGramZero**, v. 1, n. 2, abr. 2000. Disponível em: <[http://www.dgz.org.br/abr00/Art\\_01.htm](http://www.dgz.org.br/abr00/Art_01.htm)>. Acesso em: 5 maio 2011.

TEIXEIRA, Cenidalva Miranda de Sousa; SCHIEL, Ulrich. A internet e seu impacto nos processos de recuperação da informação. **Ciência Da Informação**, Brasília, DF, v. 26, n. 1, p. 65-71, jan. 1997.

TWITTER. <<https://twitter.com/>>. Acesso em: 10 jan. 2012.

WIKIPEDIA, c2011. Disponível em: <<http://pt.wikipedia.org/wiki/Twitter>> . Acesso em: 11 jun. 2011.

VICKERY, Brian; VICKERY, Alina. Semantics and retrieval. In: \_\_\_\_\_. **Information Science in theory and practice**. London : Bowker-Saur, 1989. p. 133-179.

WAL, Thomas Van der. Folksonomy Entries. Disponível em: <<http://www.vanderwal.net/random/category.php?cat=153>>. Acesso em: 21 nov. 2010.

WERSIG, G. Information science: the study of postmodern knowledge usage. **Information Processing & Management**, v.29, n.2, p. 229-239,1993.