

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

Jacy Pereira Gonçalves

TAGS:

PODEM TER VALOR DE TERMO?

Porto Alegre

2012

Jacy Pereira Gonçalves

TAGS:

PODEM TER VALOR DE TERMO?

Trabalho de conclusão de curso de graduação apresentado ao Departamento de Ciências da Informação da Universidade Federal do Rio Grande do Sul como requisito parcial para a obtenção do grau de Bacharel em Biblioteconomia.

Orientadora: Prof^a. Dr. Regina Helena van der Laan

Porto Alegre

2012

UNIVERSIDADE FEDERAL DO RIO GRANDE DO SUL

Reitor: Carlos Alexandre Netto

Vice-Reitor: Rui Vicente Oppermann

FACULDADE DE BIBLIOTECONOMIA E COMUNICAÇÃO

Diretor: Ricardo Schneiders da Silva

Vice-Diretora: Regina Helena van der Laan

DEPARTAMENTO DE CIÊNCIAS DA INFORMAÇÃO

Chefe: Ana Maria Mielniczuk de Moura

Chefe Substituta: Sônia Elisa Caregnato

COMISSÃO DE GRADUAÇÃO DO CURSO DE BIBLIOTECONOMIA

Coordenadora: Samile Andréa de Souza Vanz

Coordenadora Substituta: Glória Isabel Sattaminni Ferreira

Dados Internacionais de Catalogação na Publicação (CIP)

G635r Gonçalves, Jacy Pereira

Tags: podem ter valor de termo? / Jacy Pereira Gonçalves. –
2012.

53 f. : il.

Trabalho de conclusão de curso (graduação) – Universidade
Federal do Rio Grande do Sul, Curso de Biblioteconomia, 2012.

“Orientadora: Regina Helena van der Laan”

1. Biblioteconomia. 2. *Tags*. 3. Terminologia. I. Título. II. van
der Laan, Regina Helena. III. Universidade Federal do Rio
Grande do Sul. Curso de Biblioteconomia

CDU 025.43

Bibliotecária Jaqueline Insaurriaga Silveira – CRB 10/1392

Departamento de Ciências da Informação

Rua Ramiro Barcelos, 2705

Campus Saúde

Bairro Santana

Porto Alegre – RS

CEP 90035-007

AGRADECIMENTOS

A Deus, razão da minha existência.

À Ufrgs, pela excelência de ensino gratuito proporcionado.

A minha Orientadora, Doutora Regina Helena van der Laan, por compartilhar seu conhecimento e experiência, tanto profissional como pessoal. Obrigado pelas conversas brilhantes, pela paciência, dedicação e por acreditar que eu seria capaz de realizar este trabalho.

A Doutoranda, Rita do Carmo Laipelt, por me auxiliar nas dúvidas e problemas em informática;

A minha Mãe, Vera, pelo amor incondicional e pela força espiritual que me manteve segura.

Ao meu Pai, Rui, por todo apoio, principalmente material. Espero poder retribuir um dia.

A minha Irmã, Juliana, e meu Cunhado, Adriano, pela presença marcante nesta etapa.

Ao meu Sobrinho lindo, Miguel (Bilé), que chegou para encher minha vida de alegria.

Ao meu Namorado, Lennon, pela paciência, carinho e amor demonstrado neste período. Obrigada por estar ao meu lado, pra me consolar, me animar, me apoiar e por sempre acreditar em mim. Te amo muito.

As Gurias da Biblioteca da Escola de Administração da Ufrgs, Ana, Evelin, Jaqueline e Tânia, pela simpatia e disposição em ajudar sempre.

A todos os professores, familiares, amigos e colegas que de alguma forma acompanharam esta trajetória.

RESUMO

Apresenta pesquisa realizada para analisar se as *tags* atribuídas por usuários no *CiteUlike* sobre Oncologia possuem estatuto terminológico. Organiza o *corpus* textual composto por 48 (quarenta e oito) artigos. Elabora o *corpus* de estudo que foi constituído por um conjunto de 104 (cento e quatro) *tags*. Valida o *corpus* de estudo no discurso especializado e nos Descritores em Ciências da Saúde, o DeCS. Adota como metodologia um estudo exploratório descritivo de amostragem não probabilística com abordagem qualitativa. Relata que do *corpus* de estudo, composto por 104 (cento e quatro) rótulos, 75 (setenta e cinco) *tags* são termos específicos da Oncologia por serem validadas na situação discursiva e *in vitro* e 29 (vinte e nove) *tags* não foram validadas devendo passar por novos estudos. Conclui, a partir dos resultados obtidos, que o pressuposto de pesquisa pôde ser comprovado e que algumas *tags* atribuídas no *CiteUlike* da área de Oncologia são termos.

Palavras-chave: *Folksonomia*. *Tags*. Terminologia. Termo.

ABSTRACT

Presents research conducted to examine whether the tags assigned by users in Oncology CiteULike about terminology have status. Organizes the corpus consists of 48 (forty eight) articles. Prepares the body of study that consisted of a set of 104 (one hundred and four) tags. Validates the corpus study on communication and Descriptors in Health Sciences, DeCS. Methodology adopts as an exploratory study of non-probability sampling with a qualitative approach. Reports that the corpus of study consisting of 104 (one hundred and four) labels, 75 (seventy five) tags are specific terms of Oncology to be validated in the discursive situation and in vitro and 29 (twenty nine) tags were not validated and should undergoing further studies. Concludes from the results obtained, the assumption that research can be proven and that some CiteULike tags assigned in Oncology are terms.

Keywords: Folksonomy. Tags. Terminology. Term.

LISTA DE FIGURAS

Figura 1: Nuvem de <i>tags</i>	17
--------------------------------------	----

LISTA DE SIGLAS

BIREME	Biblioteca Regional de Medicina
DeCS	Descritores em Ciências da Saúde
EMBRAPA	Empresa Brasileira de Pesquisas Agropecuárias
INCA	Instituto Nacional de Câncer
MeSH	<i>Medical Subject Headings</i>
OMS	Organização Mundial da Saúde
RI	Recuperação da Informação
TCT	Teoria Comunicativa da Terminologia
TGT	Teoria Geral da Terminologia
UFSCar	Universidade Federal de São Carlos
USP	Universidade de São Paulo

SUMÁRIO

1 INTRODUÇÃO	10
2 OBJETIVOS	12
2.1 OBJETIVO GERAL.....	12
2.2 OBJETIVOS ESPECÍFICOS.....	12
3 REFERENCIAL TEÓRICO	13
3.1 <i>FOLKSONOMIA</i>	13
3.2 <i>TAGS</i>	16
3.3 REDES SOCIAIS E O COMPARTILHAMENTO DA INFORMAÇÃO.....	18
3.4 TERMINOLOGIA.....	21
3.5 TERMO E VARIAÇÃO TERMINOLÓGICA.....	23
3.6 REPRESENTAÇÃO E RECUPERAÇÃO DA INFORMAÇÃO.....	25
4 METODOLOGIA	28
4.1 TIPO DE ESTUDO E ABORDAGEM.....	28
4.2 OBJETO DE TRABALHO.....	28
4.3 ÁREA TEMÁTICA.....	28
4.4 AMBIENTE DE COMPARTILHAMENTO.....	30
4.5 AMBIENTE DE GESTÃO TERMINOLÓGICA.....	31
4.6 <i>CORPUS</i> TEXTUAL.....	31
4.7 <i>CORPUS</i> DE ESTUDO.....	32
4.8 VALIDAÇÃO.....	32
4.9 REGISTRO DOS DADOS.	33
4.10 PROCEDIMENTO PARA O TRATAMENTO DOS DADOS.....	34
4.11 PLANO DE ANÁLISE E INTERPRETAÇÃO DOS DADOS.....	34
4.12 LIMITADOR DO ESTUDO.....	34
5 CONSTITUIÇÃO E ANÁLISE DOS <i>CORPUS</i>	35
5.1 LEVANTAMENTO DO <i>CORPUS</i> TEXTUAL.....	35
5.2 COLETA DAS <i>TAGS</i>	35
5.3 VALIDAÇÃO DAS <i>TAGS</i> EM SITUAÇÃO DISCURSIVA.....	36
5.4 VALIDAÇÃO DAS <i>TAGS</i> NO DECS.....	37

5.5 TAGS QUE NÃO ESTÃO NEM NO DISCURSO NEM NO DECS.....	38
5.6 TAGS QUE ESTÃO NO DISCURSO E NO DECS.....	39
6 CONSIDERAÇÕES FINAIS.....	40
REFERÊNCIAS.....	41
APÊNDICE A- MODELO DE QUADRO SÍNTESE PARA COLETA DAS TAGS.....	46
APÊNDICE B - MODELO DE QUADRO SÍNTESE DE VALIDAÇÃO NO DISCURSO.....	47
APÊNDICE C - MODELO DE QUADRO SÍNTESE DE VALIDAÇÃO NO DECS.....	48
APÊNDICE D - TAGS QUE ESTÃO NO DISCURSO E NO DECS.....	49
APÊNDICE E - TAGS EXPRESSAS APENAS NO DISCURSO.....	51
APÊNDICE F - TAGS EXPRESSAS APENAS NO DECS.....	52
APÊNDICE G - TAGS QUE NÃO FORAM VALIDADAS.....	53

1 INTRODUÇÃO

Desde os tempos antigos, o Homem sempre necessitou atribuir um nome para identificar o que o cercava. Com o passar dos séculos esta necessidade aumentou, pois a comunicação ficou cada vez mais complexa entre os diferentes povos, quer nas relações sociais como nas econômicas ou políticas. Um dos fatores que mais impulsionaram essa necessidade foi o avanço científico e tecnológico. Ele promoveu o aumento significativo de novos conhecimentos, produtos e conceitos que precisavam ser representados ou identificados por um nome.

Os avanços tecnológicos, principalmente, na segunda metade do século passado e início deste século, que possibilitaram a criação de redes virtuais e a disponibilização de conteúdos de acesso livre aos diferentes públicos, trazem em seu bojo uma nova maneira de ver a informação e a sua representação.

As redes de compartilhamento possibilitam a disponibilização de informações de forma livre sem a necessidade de um profissional especializado. E também, que diferentes usuários atribuam etiquetas, denominadas *tags*, para representar estas informações.

A representação da informação passou a ser realizada por usuários que empregam palavras de seu acervo linguístico, ou seja, de seu conhecimento. O resultado de etiquetar a informação na *web* é denominada *folksonomia*. Considerando que as *tags* podem ser unidades lexicais do acervo linguístico de cada usuário, se entende que as mesmas podem ser termos.

A partir do pressuposto que *tags* podem ser termos realizar-se-á a presente pesquisa tendo como foco a área da saúde, especificamente Oncologia. A escolha desta área médica se justifica uma vez que o câncer é uma das principais causas de mortalidade no mundo. Estima-se que até o ano de 2030, cerca de 11,5 milhões de mortes no mundo ocorra devido a esta enfermidade. (ORGANIZAÇÃO MUNDIAL DA SAÚDE, 2008).

O motivo que justifica a escolha do tema ocorre pelo fato das *tags* descreverem conteúdos por pessoas não especializadas. Esta prática, até então, realizada principalmente pelo bibliotecário, hoje, está ocorrendo de forma livre na *web*. Neste

contexto, ele pode ser considerado importante para a área da Ciência da Informação, uma vez que se abre mais uma fonte de identificação de prováveis termos para a construção de vocabulários controlados.

Diante do exposto, a questão de pesquisa é: **as tags expressas no CiteUlike, sobre Oncologia, podem ser termos?**

Para a composição do estudo foram realizadas as seguintes etapas: organização do *corpus* textual, constituído por um grupo de artigos coletados no *CiteUlike*; elaboração do *corpus* de pesquisa composto por um conjunto de *tags*; validação das *tags* em situação discursiva e através do vocabulário controlado de Descritores em Ciências da Saúde (DeCS).

O trabalho está estruturado da seguinte forma: capítulo 1, Introdução, apresentação das informações gerais do trabalho, capítulo 2, Objetivos, objetivo geral e os objetivos específicos; capítulo 3, Referencial Teórico, que aborda os temas discutidos no trabalho: *Folksonomia*, *Tags*, Redes sociais, Compartilhamento da informação, Terminologia, Termo, Variação Terminológica, Representação e Recuperação da informação; capítulo 4, Metodologia, descrição do tipo de estudo, objeto de estudo, área temática, ambiente de compartilhamento, ambiente de gestão terminológica, *corpus* textual, *corpus* de estudo, validação, registro dos dados, procedimentos para o tratamento da coleta, plano de análise e interpretação dos dados e o limitador do estudo; capítulo 5, Constituição e Análise dos *Corpus*, levantamento do *corpus* textual, coletas das *tags*, validação das *tags* em situação discursiva, validação das *tags* no DeCS, *tags* que não estão nem no discurso nem no DeCS e *tags* que estão no discurso do especialista e no DeCS; Capítulo 6, Considerações Finais, seguida pelas referências e pelos apêndices.

2 OBJETIVOS

Para a execução do trabalho foram estabelecidos os seguintes objetivos:

2.1 OBJETIVO GERAL

Verificar se as *tags* sobre Oncologia atribuídas pelos usuários no *CiteUlike* possuem estatuto terminológico.

2.2 OBJETIVOS ESPECÍFICOS

- a) organizar o *corpus* textual;
- b) elaborar o *corpus* de pesquisa;
- c) validar as *tags* na situação discursiva;
- d) validar as *tags* no DeCS.

3 REFERENCIAL TEÓRICO

A fundamentação teórica deste trabalho aborda os seguintes assuntos: *Folksonomia*, *Tags*, Redes Sociais, Compartilhamento da Informação, Terminologia, Termo, Variação Terminológica, Representação e Recuperação da Informação.

3.1 FOLKSONOMIA

A *folksonomia* é o resultado da atribuição de etiquetas realizada por usuários para descrever conteúdos disponíveis na *web*. Este sistema foi cunhado por Thomas Vander Wal, sendo que o prefixo *folk*, derivado da língua inglesa, significa pessoas ou povo agregado à *taxonomy*. Segundo Vander Wal (2007, tradução nossa):

Folksonomia é o resultado da atribuição livre e pessoal às informações ou objetos (qualquer coisa com um *URL*) visando a recuperação. A atribuição de etiquetas é feita num ambiente social (normalmente compartilhada e aberta aos outros). *Folksonomia* é criada a partir do ato de marcação pela pessoa que consome a informação.

Desta forma, se entende *folksonomia* como sendo o resultado de uma criação espontânea feita pelo sujeito que constrói elementos que representem características relacionadas a um objeto. Essa representação se dá através da *tag*.

De acordo com Gouvêa e Loh (2007, p. 1), a *folksonomia* é: “[...] uma espécie de classificação (taxonomia) feita por usuários comuns e não por especialistas.” Nela, os sujeitos não precisam ser profissionais especializados na área, ou seja, são pessoas comuns que atribuem palavras-chave que consideram ter relação com um determinado conteúdo.

Para os autores uma característica marcante da *folksonomia* é a possibilidade de qualquer indivíduo associar rótulos que possam ter relação ou não com o objeto ao qual está sendo marcado. Isto por que o sujeito não é obrigado a conhecer o que está

¹ Texto original. Folksonomy is the result of personal free tagging of information and objects (anything with a URL) for one's own retrieval. The tagging is done in a social environment (usually shared and open to others). Folksonomy is created from the act of tagging by the person consuming the information.

rotulando, mas sim faz esta marcação conforme sua percepção. Além de permitir a rotulação livre de conteúdos, a *folksonomia* também possibilita que diferentes usuários compartilhem uma mesma *tag* e que intervenham na criação de novas etiquetas.

Conforme Santana e Santana (2010), a *folksonomia* “[...] permite que outras pessoas adicionem novas *tags* e que esse conteúdo descritivo gerado seja compartilhado com todos os usuários do sistema, ou seja, é um sistema de indexação colaborativa”. Através da discussão dos autores se pode entender que a indexação colaborativa ocorre a partir da ação do usuário em atribuir rótulos a um conteúdo possibilitando que outros sujeitos possam interferir nesta marcação sendo que, um mesmo conteúdo pode ser rotulado por diferentes indivíduos.

Na forma tradicional de representar um determinado conteúdo o bibliotecário normalmente utiliza ferramentas que orientam este trabalho, tais como: tesouros e cabeçalhos de assuntos. No ambiente *web* este processo pode ser realizado sem o auxílio de instrumentos, pois conforme já foi salientado, esta atribuição é realizada livremente pelo usuário.

Para Kato e Silva ([2000], p. 2), a *folksonomia*: “[...] permite que cada usuário classifique a informação com uma ou mais palavras-chave sem a obrigatoriedade de elencar apenas termos de um determinado vocabulário controlado”. Isto ocorre por que o conteúdo é relacionado naturalmente pelos usuários o que pode ser um canal de recuperação de informação.

Embora não seja função primordial da *folksonomia* a recuperação da informação, da mesma forma em que uma biblioteca, isto pode ocorrer uma vez que os indivíduos atribuem rótulos empregando unidades lingüísticas de seu conhecimento particular.

Em situações que estão referindo-se a um falante de uma área especializada, é muito provável que este sujeito rotule as informações empregando unidades lexicais especializadas. Dessa forma, outros usuários conhecedores da terminologia da área terão a possibilidade de recuperar estas informações.

Conforme Catarino e Baptista (2007) a *folksonomia* é:

[...] o resultado da etiquetagem dos recursos da *Web* num ambiente social (compartilhado e aberto a outros) pelos próprios usuários da informação visando a sua recuperação. Destacam-se portanto três fatores essenciais: 1) é resultado de uma indexação livre do próprio usuário do recurso; 2) objetiva a

recuperação a posteriori da informação e 3) é desenvolvida num ambiente aberto que possibilita o compartilhamento.

Através da colocação das autoras se compreende que a *folksonomia* é uma consequência gerada a partir da ação do usuário em representar um objeto através da etiqueta o que posteriormente gera a recuperação destes conteúdos. Pelo fato destes rótulos estarem disponíveis livremente na web, provavelmente esses conteúdos serão recuperados por outros sujeitos.

Aquino (2007, p.10) afirma que a *folksonomia*:

[...] trata-se de um mecanismo de representação, organização e recuperação de informações que não é feito por especialistas anônimos, o que muitas vezes pode limitar a busca por não trazer determinadas palavras-chave, mas sim um modo onde os próprios indivíduos que buscam informação na rede ficam livres para representá-la, organizá-la e recuperá-la, realizando estas ações com base no senso comum e tendo assim um novo leque de opções ao efetuar uma pesquisa para encontrar algum dado.

Neste sentido se a *folksonomia* pode ser considerada como um instrumento de representação e organização da informação, então ela também será um facilitador no processo de recuperação de conteúdos a diferentes usuários, uma vez que a linguagem está mais próxima que ele conhece.

Alguns autores enfatizam que a *folksonomia* possibilita a interação entre recursos na web. Conforme Santarém Segundo e Vidotti (2011, p. 88), a *folksonomia* pode ser caracterizada como: “[...] uma forma de inserir e relacionar recursos por meio da descrição dos mesmos pelas palavras-chave, de forma aberta, que tem como principal objetivo facilitar o processo de gerenciamento e recuperação das informações em ambientes digitais”.

Esse processo de recuperação poderá ser possibilitado através das palavras-chave atribuídas pelos próprios usuários interessados em determinados conteúdos. Porém, essa recuperação só ocorrerá de modo totalmente satisfatório se a representatividade da palavra-chave, em relação ao recurso na qual está descrevendo, tenha sentido para outros usuários.

A partir dessa discussão é possível compreender que um objeto poderá tornar-se recuperável ou não, pois a *folksonomia* é dirigida pelos diversos usuários que, muitas

vezes, não atribuem *tags* as quais sejam realmente significativas em relação ao objeto que está rotulando.

3.2 TAGS

A crescente demanda de informações na *web* impulsionou a criação de mecanismos que pudessem servir como canais de interação entre os diferentes conteúdos. Essa interação de informações na *Internet* é possibilitada através de métodos que facilitam a representação de informações, como é o caso da *tags*.

Tags, de modo geral são palavras-chave empregadas por usuários não especializados para nomear um recurso. Apesar da literatura e das diferentes áreas do conhecimento apresentar outras definições para *tag*, adotar-se-á, como forma de reportar-se a ela, nesta pesquisa como: *tags*, marcador, palavras-chave, rótulo e etiqueta.

Segundo Vander Wal (2007, tradução nossa)², existem três princípios básicos que regem às *folksonomias*: “1) *tag*; 2) objeto que está sendo marcado; e 3) identidade.” Desta forma, a *tag* é considerada como base fundamental na construção da *folksonomia* por ser o resultado da atribuição feita pelos usuários. O autor ressalta o princípio da identidade, o qual se pode compreender como sendo a relação estabelecida entre o sujeito e o conteúdo a ser rotulado, além da importância da marcação do objeto através da *tag*.

Uma característica relevante das *tags* é o fato delas serem atribuídas pelo próprio usuário. Para Rappetti (2007, p. 18), “as *tags* (etiquetas) são metadados, as palavras-chave para se classificar o conteúdo da informação, porém com o diferencial de serem incluídas pelos usuários desta informação”. A autora destaca o fato destes marcadores serem atribuídos pelos próprios sujeitos que consomem a informação. Este diferencial pode gerar mudanças no processo da representação da informação, que outrora era realizada, geralmente, por bibliotecários.

As *tags* podem destacar-se umas das outras, porque quanto mais uma *tag* for empregada pelos usuários, maior destaque ela obterá. É possível visualizar esse

² Texto original. 1) *tag*; 2) object being tagged; and 3) identity

destaque através da chamada nuvem de *tags* que é definida por Galdo, Vieira e Rodrigues (2009), como:

[...] uma forma de visualização para facilitar a recuperação de informações. Nas nuvens de “*tags*”, o tamanho das letras e a cor das palavras dispostas em imagens geradas dinamicamente, modificam-se à medida que informação é acessada, gerada ou classificada, podendo indicar a popularidade de uma “*tag*”, ou a frequência da palavra nos documentos da coleção.

Pode-se dizer que a nuvem é a representação visual das *tags* na qual se destacam graficamente conforme a frequência em que os usuários atribuem ou acessem uma palavra-chave. Esta nuvem fica disponível para que todos os usuários possam visualizá-la.

Figura 1- Nuvem de *tags*



Fonte: *CiteUlike* (2012).

Na figura 1, é possível perceber que o rótulo *review* se sobressai em relação ao rótulo *cancer*. Este destaque pode influenciar os usuários a consumir uma informação pela ênfase que ela recebeu em relação a outros assuntos. Isto por que, ao visualizar as *tags*, os sujeitos percebem quais conteúdos estão sendo considerados mais relevantes.

Com a folksonomia, a informação ganha relevância por meio de um “filtro social colaborativo”. Ou seja, se centenas de pessoas consideram uma mesma informação relevante e a classificam com uma mesma tag, existe uma chance muito maior de ela ser encontrada facilmente. (KATO; SILVA, [2000], p.2).

É evidente que alguns assuntos se destacam em relação a outros devido à popularidade que possuem. Porém, conforme os autores, as *tags* podem tornar um conteúdo mais representativo se comparado a outros devido ao número de acessos que este marcador recebeu.

Neste sentido, as *tags*, além de popularizar um conteúdo também podem aproximar usuários, devido a este “filtro social colaborativo” que possibilita aos diferentes sujeitos compartilhar os mesmos interesses informacionais de forma que podem relacionar, gerar e compartilhar novos conteúdos.

3.3 REDES SOCIAIS E O COMPARTILHAMENTO DA INFORMAÇÃO

A *Internet* promoveu grandes transformações no âmbito das relações humanas. Antes do surgimento deste advento, os sujeitos se relacionavam num cenário onde a informação era disseminada de modo gradual. Hoje, este contexto foi modificado e através das ferramentas oferecidas pela *web* e do aprimoramento das redes sociais a comunicação entre os sujeitos propaga-se de forma muito dinâmica.

As redes sociais são ambientes nos quais os usuários interagem, consomem, produzem e compartilham informações. Para Marteleto (2001, p.72), as redes sociais representam “[...] um conjunto de participantes autônomos, unindo idéias e recursos em torno de valores e interesses compartilhados”. Estes espaços se popularizaram largamente através da Web 2.0.

A *Web 2.0* é a segunda geração de serviços online e caracteriza-se por potencializar as formas de publicação, compartilhamento e organização de informações, além de ampliar os espaços para a interação entre os participantes do processo. (PRIMO, 2007, p. 1).

Ao mesmo tempo em que os usuários consomem também produzem informações e compartilham entre si uma vez que estes ambientes permitem a interação das pessoas e a construção coletiva de novos conhecimentos.

A interação é, sem dúvida, uma das características marcantes das redes sociais. Conforme Recuero (2005, p. 9), “esses sistemas funcionam com o primado fundamental da interação social, ou seja, buscando conectar pessoas e proporcionar sua comunicação e, portanto, podem ser utilizados para forjar laços sociais.”

Estes laços se devem ao fato dos usuários compartilharem os mesmos interesses em relação a uma determinada informação. Assim, os sujeitos são atraídos a ambientes que possuam afinidade em relação as suas opiniões, comportamento, crenças, entre outros.

Nas redes sociais, há valorização dos elos informais e das relações, em detrimento das estruturas hierárquicas. Hoje o trabalho em rede é uma forma de organização humana presente em nossa vida cotidiana e nos mais diferentes níveis de estrutura das instituições modernas. (MARTELETO, 2001, p.72).

Dentre as particularidades das redes sociais já citadas, cabe ressaltar que elas são mais que espaços de compartilhamento, são ambientes que geram transformações no proceder dos sujeitos. Conforme Alves e Barbosa (2010, p.117), “os novos fenômenos relacionados ao conhecimento humano, como a criação de redes sociais e de ambientes digitais interativos, têm contribuído para a transformação de costumes e comportamentos com relação à informação”.

Desta forma, se percebe claramente que as redes assumiram um papel de espaços que promovem a comunicação e compartilhamento da informação, como também de ambientes que podem influenciar na conduta dos sujeitos no processo da apropriação da informação.

A partir das redes sociais, os usuários podem obter um conhecimento mais amplo em relação ao que está sendo produzido e que tipo de informação as pessoas

estão se interessando. Esses ambientes também se destacam pela grande quantidade de informações que compartilham em um curto período.

O compartilhamento é ato de transmitir a informação de forma que ela possa ser absorvida por outras pessoas. Segundo Alcará et al. (2009, p. 172), “entre indivíduos o compartilhamento é um processo que permite a uma pessoa transformar o conhecimento para que possa ser entendido, absorvido e utilizado por outros atores”. Desta forma, o compartilhamento ocorre quando um indivíduo converte seu conhecimento de forma que possa ser compreendido por outras pessoas.

Sendo assim, a rede social pode ser considerada um importante meio para transmitir e compartilhar informações. Para Tomáel, Alcará e Di Chiaro (2005, p. 93), “as redes sociais constituem uma das estratégias subjacentes utilizadas pela sociedade para o compartilhamento da informação e do conhecimento, mediante as relações entre atores que as integram”.

Para Alves e Barbosa (2010, p.119) o compartilhamento de informação se caracteriza: “[...] como um processo capaz de promover a integração entre as pessoas com vistas ao aprendizado contínuo, à transmissão mútua de conceitos e habilidades, bem como à geração de novos conhecimentos que promovam a inovação”. Neste sentido os autores introduzem outro aspecto característico do compartilhamento: a inovação através dos novos conhecimentos. A partir da integração e do acúmulo de informações propostas por diferentes sujeitos é possível que seja favorável ao desenvolvimento e a produção de conhecimentos que ainda não foram discutidos ou que ainda não são tão frequentes.

Quando um sujeito transmite seu conhecimento a outras pessoas não implica dizer que ele terá perdas em relação às informações que possui. Para Alcará et al. (2009, p. 172), “compartilhar não significa ceder à posse do conhecimento e sim usufruir dele juntamente com os pares.” O sujeito ao interagir com outros indivíduos não estará acarretando dano a si, mas sim estará agregando novos aprendizados e aprimorando seus conhecimentos.

3.4 TERMINOLOGIA

A história da humanidade mostra que o Homem sempre precisou identificar os objetos que o rodeavam para facilitar suas relações sociais e econômicas. Nos séculos XVIII e XIX esta necessidade se tornou cada vez mais fundamental. Isto por que, neste período, a sociedade presenciou uma série de mudanças tecnológicas com grande impacto no âmbito social, econômico e científico. Houve, também, o surgimento de novos conceitos e nomenclaturas nas áreas especializadas do conhecimento que assim provocaram dificuldades na comunicação.

Diante do exposto, os cientistas passaram a considerar fundamental o estudo da Terminologia para eliminar os problemas na comunicação. Segundo Pavel et al. (2003, texto eletrônico), a terminologia é uma: “disciplina lingüística que estuda os conceitos e os termos usados em linguagem de especialidade”. Nesta perspectiva se salienta os estudos da Terminologia moderna através do engenheiro Eugen Wüster (1898-1977) quando este em 1931 publica a obra *Die Internationale Sprachnormung In Der Technik, Besonders In Der Elektronik*.

O pensamento wusteriano era regido por aspectos normativos e metodológicos com o objetivo de eliminar as ambigüidades ocorridas nos discursos das áreas especializadas do conhecimento. A partir do seu trabalho surge a Teoria Geral da Terminologia (TGT).

Para Wüster (1998), o termo é regido pelo princípio da univocidade (um conceito é representado por um único termo) e caracterizado pela monorreferencialidade (um termo só pode representar um único conceito). Desta forma para a TGT não pode ocorrer à polissemia.

A padronização terminológica que em princípio possibilitaria mais precisão na comunicação entre os especialistas nem sempre foi condição de êxito. Conforme van der Laan (2002, p. 50), isto pode ser explicado: “[...] por que em situações reais há uma multiplicidade de registros que mantém seu caráter especializado, mas apresentam características de unidades lexicais manifestadas em outras situações discursivas”.

A observação que pode ocorrer variações denominativas no discurso especializado ou que unidades lexicais do léxico geral poderiam adquirir estatuto

terminológico gerou os estudos de Maria Tereza Cabré e seus colaboradores, que resultaram na Teoria Comunicativa da Terminologia (TCT). Segundo Cabré (1999, p. 120, tradução nossa)³ :

A TCT não concebe a terminologia como uma matéria autônoma, mas mantém seu caráter intrinsecamente interdisciplinar e defendendo a generalização, tenta explicar dentro de uma teoria da linguagem que, por sua vez, se insere em uma teoria da comunicação e do conhecimento.

Para autora, a Terminologia não é uma área que pode ser estudada isoladamente, mas sim que é necessário lançar mão de áreas afins do conhecimento, como: ciências da língua, ciências da cognição e as ciências sociais. Conforme Cano (2002), a TCT:

[...] o objeto de estudo da Terminologia são as unidades terminológicas utilizadas nos diversos domínios da especialidade. Os termos não são unidades isoladas, constituindo um sistema próprio, mas unidades que se incorporam ao léxico do falante, à medida que adquire conhecimentos especializados.

Para Cabré (1993), o termo é uma unidade lexical que apresenta caráter poliédrico, em outras palavras, possui uma dimensão lingüística, uma dimensão cognitiva e uma dimensão comunicativa. Assim, o termo é uma unidade lexical que pertence a um domínio de especialidade e que transmite e divulga o conhecimento científico.

A Terminologia é um termo polissêmico entendido como disciplina, prática e produto. Para Pavel (2003), a terminologia é entendida como uma disciplina linguística dedicada ao estudo científico dos conceitos e dos termos usados nas linguagens de especialidade. Como prática é a atividade profissional que abrange o estudo científico dos conceitos e dos termos em uso nas linguagens de especialidade. E por fim, como produto, é o conjunto de palavras técnicas que pertencem a uma ciência, a uma atividade profissional, ou a um grupo social.

³ Texto original. A TCT não concebe a terminología como una matéria autônoma, sino que, manteniendo su caráter instrínsecamente interdisciplinar, y abogando por la generalización, intenta explicarla dentro de una teoría del lenguaje que, a su vez, se inserte em uma teoría de la comunicación y del conocimiento.

Conforme Cabré (1995, tradução nossa)⁴, a terminologia é compreendida por três noções:

a) a disciplina; b) a prática e c) o produto gerado por esta prática. Como disciplina, é a matéria que lida com termos especializados; como prática, é o conjunto de princípios para a recolha de termos, como um produto, é o conjunto de termos de uma especialidade.

Desta forma, o estudo da Terminologia vem ao encontro da proposta deste trabalho, uma vez que, objetiva-se estudar as *tags* entendendo-as como prováveis termos.

3.5 TERMO E VARIAÇÃO TERMINOLÓGICA

O termo é uma unidade lexical expressa no discurso dos especialistas. Conforme Faulstich (2006, p. 28), o termo pode ser definido como: “signos que encontram sua funcionalidade nas linguagens de especialidade de acordo com a dinâmica das línguas”.

Segundo van der Laan (2002, p. 63), o termo é uma: “unidade lexical básica da terminologia que expressa um conceito, um objeto ou um processo.” Em outras palavras, o termo é uma designação verbal que representa um conceito de uma área especializada.

O termo pode manifestar-se em diferentes níveis discursivos, ou seja, no discurso especializado, no discurso técnico ou de vulgarização científica. Desta forma, os termos transmitem e representam um conhecimento em diferentes níveis de discursos gerando o que se denomina variação terminológica. A variação ocorre quando para um mesmo conceito existe mais de uma forma de representá-lo, ou seja, mais de um registro.

Para Faulstich (2006, p. 28), “quando os termos tiverem as mesmas condições

⁴ Texto original. a) a la disciplina; b) a la práctica; c) al producto generado por esa práctica. Como disciplina, es la materia que se ocupa de los términos especializados; como práctica, es el conjunto de principios encaminados a la recopilación de términos; como producto, es el conjunto de términos de una determinada especialidad.

de uso, serão considerados variantes um do outro. Nesse caso, eles têm formas parcial ou totalmente diferentes para um mesmo significado referencial e estão disponíveis para o uso corrente”.

Conforme a autora, “a polifuncionalidade da unidade lexical, no discurso científico, no discurso técnico ou no discurso de vulgarização científica pode produzir mais de um registro ou mais de um conceito para o mesmo termo”. (FAULSTICH, 1998a, apud VAN DER LAAN, 2002, p. 51)⁵. Isto, porque a polissemia possibilita que uma unidade lingüística adquira um novo significado.

Este fato pode ser explicado porque a língua é um fenômeno em constante movimento. Neste sentido, Arruda (2011, p. 8) afirma que: “a variação da língua faz parte de sua natureza interpessoal e da própria natureza heterogênia da sociedade humana”.

Diante do exposto, pode-se afirmar que a variação também ocorre no discurso especializado. Para van der Laan, Ferreira e Bonotto (2007, p. 818): “o falante, mesmo o especializado, constantemente cria termos e re-significa termos já existentes, o que resulta no surgimento de variantes”.

Conforme Faulstich (1998b), existem dois grupos de variantes terminológicas: variantes lingüísticas e as variantes de registro. O primeiro grupo é entendido como, “[...] o fenômeno propriamente lingüístico determina o processo de variação.” (FAULSTICH, 1996, texto eletrônico). Este grupo leva em consideração o significado atribuído pelo sujeito na sua forma de se comunicar e que conseqüentemente acarretará modificações na forma em que se dará o registro dos termos.

Para a autora as variantes lingüísticas estão classificadas como: a) variante terminológica fonológica, quando o registro pode ser feito conforme a fala; b) variante terminológica morfológica, quando ocorre alternância na estrutura morfológica; c) variação terminológica sintática, quando ocorre alternância entre duas construções sintagmáticas; d) variante lexical, quando ocorre à supressão de um item na construção verbal sem alterar a idéia; e) variante terminológica gráfica, quando ocorre alterações na grafia, conforme regras da língua.

⁵ FAULSTICH, 1998a apud VAN DER LAAN, 2002, p. 51.

O segundo grupo se refere às variantes de registro. Conforme Faulstich (1996) são aquelas cuja: “[...] variação decorre do ambiente de ocorrência, no plano horizontal, no plano vertical e no plano temporal em que se realizam os usos lingüísticos dos termos”.

Para a autora, este grupo está subdividido em: a) variante terminológica geográfica, ocorre quando para um mesmo conceito existem várias formas de expressá-lo por sujeitos que compartilham a mesma língua, porém que se situam em regiões diferentes; b) variante terminológica de discurso, ocorre nas expressões para um mesmo conceito nos diferentes níveis de discurso (técnico, científico ou de vulgarização científica); c) variante terminológica temporal, ocorre quando um termo, com o passar de um período de tempo, cai em desuso sendo substituído por outro.

Pode-se dizer que as variantes terminológicas de registro consideram a forma em que os termos ocorrem no discurso dos especialistas e no modo em que são expressos, ou seja, são observados tanto na oralidade como no seu registro.

3.6 REPRESENTAÇÃO E RECUPERAÇÃO DA INFORMAÇÃO

A grande quantidade de informações produzidas, principalmente, após o advento da imprensa, refletiu fortemente no modo como estes dados eram tratados. Foi necessário estabelecer formas de representar a informação a fim de que pudesse ser encontrada.

A representação é o processo na qual se descreve as características de um objeto de forma que este possa ser posteriormente encontrado. Segundo Alvarenga (2003, p. 21), a representação:

Destaca-se como uma instância do processo cognitivo humano aquela que culmina com a representação primária do conhecimento, situando-se no âmbito do registro do pensamento em um suporte documental, incluindo as etapas de percepção, identificação, interpretação, reflexão e codificação, etapas que são envolvidas no ato de se conhecer um novo ser ou coisa, ou aprofundar-se no conhecimento de um ser ou uma coisa já conhecida, utilizando-se dos sentidos, da emoção, da razão e da linguagem.

A autora destaca os sentidos, a emoção, a razão e a linguagem como pontos influenciadores na representação. O sentido, como sendo aquela atribuição do ser humano que percebe e que reconhece as coisas. A emoção, como a faculdade capaz de raciocinar, de compreender e de chegar a conclusões acerca de um objeto. E por fim, a linguagem, como sendo a forma de comunicar ou de expressar uma determinada idéia.

Neste âmbito, se percebe que a representação é um procedimento que compreende etapas complexas. Conforme Novellino (1996, p. 38), “o processo de representação da informação envolve dois passos principais: 1) análise de assunto de um documento e a colocação do resultado desta análise numa expressão lingüística. 2) atribuição de conceitos ao documento analisado”. Para a autora, a representação de um conteúdo compreende a interpretação da informação descrita e representada por uma unidade lexical.

Na esfera informacional, a representação de um objeto é fundamental para que, futuramente, este possa ser recuperado. Neste sentido Novellino (1996, p. 38) afirma que a representação: “[...] funciona então como um artifício para enfatizar o que é essencial no documento considerando sua recuperação, sendo a solução ideal para organização e uso da informação”.

Assim, se pode dizer que um dos principais objetivos da representação do conteúdo é a sua recuperação. A recuperação da informação (RI) para Gonzalez e Lima ([2000], p. 2), “consiste na busca de documentos relevantes a uma dada consulta que expressa a necessidade de informação do usuário”. Logo, o sujeito realiza uma pesquisa com êxito ao recuperar informações que vão ao encontro de seus interesses.

Entretanto, a RI é um processo que não envolve apenas a busca por documentos, mas que considera as características subjetivas do sujeito que o influenciam quando este realiza a busca de uma determinada informação.

A RI, por levar em consideração também os aspectos subjetivos dos usuários de informação, presentes não só na formulação de termos para busca, mas também no valor (relevância) dado às respostas recebidas de um sistema de busca, não pode abranger apenas instrumentos quantitativos em suas atividades. Assim como outros sistemas de informação, um sistema de busca, ou de RI, é composto por registros de informação, usuários e o contexto de sua utilização. (TEIXEIRA; DUQUE, 2010 p. 3).

O processo de recuperação da informação foi modificado através do advento da *web*, pois a informação passou a estar disponibilizada livremente nestes espaços. Assim, os próprios usuários que utilizam estes ambientes representam a informação através das *tags* e isto tem se tornado uma prática cada vez mais comum na *Internet*.

Conforme Pereira e Cruz (2012), “a Web [*sic*] hoje permite o uso de linguagens mais flexíveis e de padrões cada vez mais aceitos de representação da informação. Isso a transforma em uma rede de conhecimento, e não apenas em um espaço onde co-habitam dados sem conexão”.

Esta representação de conteúdos realizada na *web* através das *tags* permite dizer que se um rótulo representa um conteúdo é possível que ele recupere uma informação mesmo que o usuário não tenha tal intenção. Isto pode ocorrer uma vez que os usuários empregam rótulos de acordo com a sua linguagem.

De tal modo cabe ressaltar as transformações ocorridas no processo da recuperação da informação na *web* que foram geradas através da *folksonomia* e das *tags*. Santarém Segundo e Vidotti (2011, p. 89), afirmam que:

A *folksonomia* mudou o paradigma em relação à recuperação da informação na Web. Portanto, trata-se de um recurso muito rico, que contribui de forma acentuada para o fortalecimento e solidificação da Internet como plataforma para a construção de informação coletiva.

As *folksonomias* e as *tags*, de certa forma, colaboraram para que a recuperação da informação na *web* se tornasse um processo mais dinâmico, uma vez que, existe a possibilidade da interação e do compartilhamento entre diferentes recursos.

4 METODOLOGIA

A metodologia desta pesquisa foi subdividida nas seguintes seções: definição do tipo e objeto de estudo; área temática; ambiente de compartilhamento; ambiente de gestão terminológica; *corpus* textual; *corpus* de estudo; validação; registro; procedimento para o tratamento e análise dos dados; e por fim os limitadores deste estudo.

4.1 TIPO DE ESTUDO E ABORDAGEM

O trabalho consiste em um estudo exploratório descritivo, de amostragem não probabilística com abordagem qualitativa. A escolha da pesquisa qualitativa deu-se pelo fato dela compreender:

A interpretação dos fenômenos e a atribuição de significados são básicas no processo de pesquisa qualitativa. Não requer o uso de métodos e técnicas estatísticas. O ambiente natural é a fonte direta para coleta de dados e o pesquisador é o instrumento-chave. É descritiva. Os pesquisadores tendem a analisar seus dados indutivamente. O processo e seu significado são os focos principais de abordagem. (SILVA e MENEZES 2001, p. 20)

Com este estudo, buscou-se realizar a pesquisa a fim de verificar se *tags* possuem estatuto terminológico.

4.2 OBJETO DE TRABALHO

O objeto de estudo deste trabalho foi constituído pelo conjunto de *tags* expressas em *site* de compartilhamento social na área da Saúde, especificamente da Oncologia.

4.3 ÁREA TEMÁTICA

A Oncologia é a especialidade que estuda o desenvolvimento dos tumores malignos (cancros) e as formas de tratar esta doença. A palavra Oncologia é de origem grega que vem de *oykos*, que significa volume ou tumor.

Esta especialidade no Brasil é conhecida como Cancerologia e considerada uma área muito complexa da Saúde por se tratar de um ramo multidisciplinar que lança mão de outros especialistas como: pediatras, psiquiatras, profissionais da patologia, entre outros. (ONCOGUIA, 2003).

Conforme estudos da Organização Mundial da Saúde (OMS), o tipo de câncer que mais atinge a população mundial é o de pulmão. Este fato se deve ao elevado consumo de tabaco consumido diariamente pelas pessoas, o que diminui drasticamente a qualidade de vida.

A ocorrência do câncer está ligada a alguns fatores, tais como: estilo de vida saudável, radiação, exposição a agentes cancerígenos em locais de trabalho e contato com algumas infecções. A ocorrência também varia conforme a região, como no caso de câncer de mama, próstata e cólon que são mais freqüentes em países desenvolvidos. Já em países em desenvolvimento surgem outros tipos de câncer, como estômago, fígado e colo do útero. Espera-se que do ano de 2007 ao ano de 2030 o número de mortes por esta doença aumente cerca de 45%. Esta expectativa ocorre não só pelo aumento de novos casos como também pelo aumento e envelhecimento da população mundial. (ORGANIZAÇÃO MUNDIAL DA SAÚDE, 2008).

No Brasil, as estatísticas realizadas no ano de 2012 apresentadas pelo Instituto Nacional de Câncer (INCA), organização responsável por ações voltadas a prevenção e controle da doença, apontam que: o tipo de câncer mais comum que ocorre no sexo feminino é o de mama com 52.680 novos casos sendo mais frequente nas mulheres da região sudeste e é o que mais acomete as mulheres de todo o mundo. A pesquisa ainda revelou que o tipo de câncer mais frequente no sexo masculino é o de próstata com cerca de 60.180 novos casos no Brasil, sendo que, é o segundo tipo de câncer que mais afeta os homens no mundo. (INSTITUTO NACIONAL DE CÂNCER, 2012).

Salientando o crescimento desta área é possível compreender que é um importante campo de estudo a ser explorado e pesquisado pelos Bibliotecários uma vez que se confirmado a hipótese desta pesquisa de que *tags* podem ser termos, se abre mais um canal para coleta de prováveis termos para a construção de linguagens documentárias na área da Saúde. Isso significa dizer que existem mais possibilidades

de fontes, além das tradicionais, para o profissional da informação criar linguagens de padronização que auxiliem na comunicação dos especialistas.

4.4 AMBIENTE DE COMPARTILHAMENTO

O ambiente de compartilhamento utilizado nesta pesquisa para a composição do *corpus* textual foi o *CiteUlike*, que é uma rede de compartilhamento de informação disponível *on-line*. Ele é um *site* que utiliza o princípio de *bookmarks*, método que permite aos usuários da *Internet* organizar seus recursos *on-line*, como uma biblioteca só que na *web*, e possibilita desenvolver troca de referências entre os pesquisadores. (CITEULIKE, 2010).

O *CiteUlike* foi criado no ano de 2004 por Richard Cameron e é um ambiente que permite ao usuário armazenar, organizar e compartilhar trabalhos acadêmicos. É possível também extrair automaticamente detalhes da referência de um conteúdo quando o usuário encontra um documento na *web* que lhe pareça interessante. (CITEULIKE, 2010).

Para realizar uma pesquisa no *CiteUlike*, é preciso inserir um termo de busca, e em poucos segundos o sistema recupera documentos relacionados ao assunto solicitado. O *site* apresenta as referências desses documentos e o usuário ao acessá-las visualiza informações, como: o local de acesso do documento, comentários, quantidade de usuários que atribuíram rótulos, as *tags* que foram atribuídas aos documentos, entre outros.

O ambiente admite que o usuário, que possua cadastro, insira novos conteúdos no *site*. Para realizar este procedimento é necessário: acrescentar os dados referentes ao conteúdo que está inserindo, tais como: autor, título e data. (CITEULIKE, 2010).

Deve-se ressaltar que o cadastro é importante para os sujeitos que apenas realizam pesquisas e que não possuem interesse de inserir artigos, pois nem todos os documentos estão disponíveis livremente. Alguns recursos são de acesso restrito e estão disponíveis apenas para pessoas que estiverem inscritas no *site*.

4.5 AMBIENTE DE GESTÃO TERMINOLÓGICA

Para a organização do *corpus* textual desta pesquisa foi necessário o auxílio de um ambiente que facilitasse o processo da pesquisa e da prática terminológica. Nesta pesquisa foi utilizado o e-Termos, acrônimo de Termos Eletrônicos, ambiente colaborativo disponível *online* que possibilita a gestão terminológica automática permitindo a criação de produtos terminológicos, seja para a pesquisa acadêmica ou interesses afins. (E-TERMOS, 2009).

O site oferece seis etapas para o gerenciamento terminológico, sendo que para a presente pesquisa foram adotadas apenas as duas primeiras, descritas a seguir:

- a) Compilação automática de *corpus*: realiza a compilação automática do *corpus* utilizando um conjunto de palavras-chave, para a consulta e coleta de textos;
- b) Análise da qualidade do *corpus*: realiza a compilação e análise quantitativa e qualitativa do *corpus*.

O projeto é o produto final da Tese de doutorado de Leandro Henrique Mendonça de Oliveira e foi estruturado a partir das parcerias entre a Empresa Brasileira de Pesquisas Agropecuárias (Embrapa) Informática Agropecuária, a Universidade de São Paulo (USP) e Universidade Federal de São Carlos (UFSCar). O e-Termos tem como objetivo ser uma ferramenta que auxilie estudantes e pesquisadores, a fim de facilitar, através da automação, as tarefas e etapas na produção de trabalhos terminológicas. (E-TERMOS, 2009).

4.6 CORPUS TEXTUAL

O *corpus* textual foi composto pelo conjunto de documentos referentes à Oncologia coletados no *CiteUlike*. Para a composição desse *corpus* foi realizada as seguintes etapas: busca com o termo *Oncology* e a verificação dos documentos recuperados conforme os critérios estabelecidos nesta pesquisa, expostos a seguir:

- a) Possuir ênfase no tema Oncologia;
- b) ser artigo;
- c) disponível para acesso completo, em formato pdf;
- d) escrito no idioma inglês;
- e) compartilhado por, no mínimo, dois usuários.

4.7 CORPUS DE ESTUDO

O *corpus* do estudo foi constituído pelo conjunto de *tags* coletadas no *corpus* textual deste trabalho. Esta coleta gerou um total de 529 (quinhentas e vinte e nove) *tags*. Nesta listagem foram encontrados rótulos que não se enquadravam com o foco central desta pesquisa.

Diante do exposto, para a organização do *corpus* de estudo foram adotados os seguintes critérios:

- a) selecionar apenas as *tags* da área de Oncologia representadas por, no mínimo, dois usuários;
- b) excluir *tags* que forem repetidas, siglas, acrônimos, abreviaturas e muito genéricas em relação à área da Oncologia;
- c) validar as *tags* através no DeCS, buscando seu equivalente em português (Brasil).

4.8 VALIDAÇÃO

O conjunto de *tags* foi validado em duas situações: discursiva e *in vitro*. Na situação discursiva as *tags* foram verificadas no *corpus* textual a fim de analisar se estas foram empregadas pelos especialistas. Na validação *in vitro* as *tags* foram verificadas tendo como respaldo o vocabulário utilizado na área da Saúde: o DeCS. Este vocabulário estruturado foi construído com o intuito de padronizar a indexação dos diferentes documentos referentes às Ciências da Saúde, tais como: artigos de revistas científicas, livros, periódicos, anais de congressos, entre outros.

A escolha do DeCS se deu pelo fato dele ser considerado uma importante ferramenta na padronização da linguagem de indexação utilizada na área da Saúde. Além disso, ele possui relevância no âmbito internacional e o reconhecimento como um importante instrumento para servir como uma linguagem única na indexação de documentos referentes à área das Ciências da Saúde.

O DeCS foi criado pela Biblioteca Regional de Medicina (BIREME) e desenvolvido a partir do *Medical Subject Headings* (MeSH), vocabulário controlado para a indexação de artigos, com o intuito de possibilitar o uso da terminologia comum na área da Saúde a fim de proporcionar uma consistência maior na recuperação da informação. (BIBLIOTECA VIRTUAL EM SAÚDE, 2012).

O vocabulário é composto por 30.895 descritores de várias especialidades da Saúde e disponível em três idiomas: português, inglês e espanhol. Os conceitos são organizados através da estrutura hierárquica, o que permite ao usuário pesquisar tanto termos mais gerais quanto mais específicos em uma mesma hierarquia (BIBLIOTECA VIRTUAL EM SAÚDE, 2012).

4.9 REGISTRO DOS DADOS

Para o registro dos dados foram utilizados formulários, criados a partir da planilha no Excel. A escolha se deu pelo fato do formulário ser um instrumento simples e que possibilita uma maior facilidade na coleta.

Devido à quantidade elevada de informações registradas, geradas a partir do levantamento de artigos e das *tags*, foi necessário sintetizar os dados obtidos nos formulários através de quadros síntese para a melhor compreensão dos dados

Para a realização da coleta foram utilizados três formulários. O primeiro formulário teve como objetivo registrar as *tags* coletadas no *corpus* textual onde se registrou os seguintes dados: a *tag* atribuída, o número de usuários que a compartilharam e a data da coleta.

O segundo formulário foi utilizado para a validação das *tags* em situação discursiva. Neste formulário, foram registradas as seguintes informações: a *tag*

atribuída, se a *tag* foi expressa no *corpus* textual, o número de ocorrências e a data da coleta.

O terceiro formulário serviu para a validação das *tags* no DeCS e foi organizado com os seguintes campos: a *tag* atribuída, se estava no DeCS, à equivalência em português, e a data de registro.

4.10 PROCEDIMENTO PARA O TRATAMENTO DOS DADOS

A análise dos dados desta pesquisa foi estruturada em duas etapas descritas a seguir:

- a) validação das *tags* no *corpus* textual;
- b) validação das *tags* no DeCS;

4.11 PLANO DE ANÁLISE E INTERPRETAÇÃO DOS DADOS

O plano e interpretação dos dados ocorreram na forma descritiva, através da apresentação do tema conforme a bibliografia consultada. Em seguida, foi realizada a interpretação dos dados por meio das informações registradas nos formulários de coleta.

A partir dos dados coletados, foi possível, então, afirmar se as *tags* expressas no *corpus* de pesquisa possuem ou não estatuto terminológico, ou seja, valor de termo. Estes dados foram verificados, analisados e apresentados na forma textual para melhor expor os resultados obtidos.

4.12 LIMITADOR DO ESTUDO

O idioma das *tags* e dos documentos foi um limitador, porque o ambiente de compartilhamento se encontra disponível apenas na língua inglesa. Para evitar que nenhum significado fosse perdido, se adotou o critério de tradução por equivalência, ou seja, se buscou equivalentes na língua portuguesa (Brasil).

5 CONSTITUIÇÃO E ANÁLISE DOS *CORPUS*

Este capítulo descreve como foram constituídos e analisados o *corpus* textual composto pelo conjunto de artigos coletados no *CiteUlike* e o *corpus* de estudo constituído pelas *tags*. Esta seção foi subdividida em tópicos para uma melhor compreensão das etapas realizadas nesta pesquisa.

5.1 LEVANTAMENTO DO *CORPUS* TEXTUAL

O *corpus* textual foi constituído após o levantamento de documentos no ambiente de compartilhamento do *CiteUlike*, empregando o termo *Oncology*, tema central da pesquisa. Obteve-se um total de 1.000 (mil) documentos.

Esta listagem foi analisada e verificou-se que 925 (novecentos e vinte e cinco) não correspondiam a um dos critérios da pesquisa quanto ao número de usuários, pois haviam sido compartilhados por apenas um sujeito. Desta forma chegou-se a uma nova listagem composta por 75 (setenta e cinco) documentos.

Destes 75 (setenta e cinco) documentos, 27 (vinte e sete) tinham restrição de acesso e foram excluídos conforme o estabelecido nos critérios de seleção do *corpus* textual. Assim, o *corpus* textual ficou constituído por um total de 48 (quarenta e oito) artigos.

5.2 COLETA DAS TAGS

As *tags* foram coletadas no *corpus* textual desta pesquisa e geraram uma listagem de 529 (quinhentos e vinte e nove) rótulos registrados em planilhas no Excel elaboradas especificamente para este fim.

A listagem foi analisada observando-se o significado da *tag* a fim de se obter a tradução e o seu equivalente para o português do Brasil. Este critério foi estabelecido, pois muitas palavras quando traduzidas para o português (Brasil), perdem seu significado.

Ao verificar a listagem de 529 (quinhentas e vinte e nove) *tags*, de acordo com os critérios estabelecidos na pesquisa, constatou-se que: 281 (duzentas e oitenta e uma)

eram repetições, 44 (quarenta e quatro) eram abreviaturas, siglas ou acrônimos e 100 (cem) se referiam a expressões muito gerais e não significativas para a área específica deste trabalho.

Desta forma foi gerada uma nova listagem composta por 104 (cento e quatro) *tags*.

5.3 VALIDAÇÃO DAS TAGS EM SITUAÇÃO DISCURSIVA

Considerando o pressuposto de pesquisa de que *tags* poderiam ser termos e, portanto os ambientes de compartilhamento seriam fontes de coleta e de identificação de candidatos a termos, iniciou-se a análise das *tags*.

Foi verificada cada uma das *tags* em relação a sua ocorrência nos textos que constituem o *corpus* textual deste trabalho. Esta validação se justifica, uma vez que, um autor especializado ao redigir seus trabalhos deveria empregar a terminologia da sua área de pesquisa. Desta forma, partiu-se do princípio que toda a *tag* encontrada nos textos são termos.

Das 104 (cento e quatro) *tags* que constituem o *corpus* de pesquisa, 64 (sessenta e quatro) *tags* constam no discurso dos especialistas representando um total de 62 por cento. De acordo com van der Laan (2002, p. 63), “[...] uma unidade lexical adquire valor de termo dentro de um determinado contexto, ou seja, uma situação discursiva real”. Diante do exposto pode-se afirmar que essas 64 (sessenta e quatro) *tags* são termos, uma vez que foram expressas no discurso especializado.

Verificando a incidência das *tags* no *corpus* textual percebeu-se que alguns rótulos obtiveram maior destaque em relação a outros. Para Krieger e Finato (2004, p. 207): “a maior ou menor presença de determinadas unidades lexicais, tidas como conceitualmente mais relevantes, pode oferecer um tópico de pesquisa com textos especializados vinculado à observação de completude maior ou menor dos textos sob estudo”.

As *tags* que se destacaram em relação à incidência no *corpus* textual foram: *Cancer* com 3.274 (três mil duzentas e setenta e quatro) incidências; *Tumor* com 562

(quinhentas e sessenta e duas) incidências; *Oncology*, termo que foi o argumento de busca para o levantamento de artigos, com 372 (trezentas e setenta e duas) incidências, entre outras.

5.4 VALIDAÇÃO DAS TAGS NO DECS

Além da validação das *tags* no discurso especializado, também se realizou a validação das mesmas *in vitro*, empregando como referência o vocabulário da área da Saúde, o DeCS.

O DeCS é um vocabulário estruturado que organiza, descreve e promove o acesso a informação. Os vocabulários estruturados são: “[...] coleções de termos, organizados segundo uma metodologia na qual é possível especificar as relações entre conceitos com o propósito de facilitar o acesso à informação”. (BIBLIOTECA VIRTUAL EM SAÚDE, 2012).

Desta forma entende-se a importância que este vocabulário possui no âmbito da organização e tratamento da informação na área da Saúde uma vez que estes são:

[...] usados como uma espécie de filtro entre a linguagem utilizada pelo autor e a terminologia da área e também podem ser considerados como assistentes de pesquisa ajudando os usuários a refinar, expandir ou enriquecer suas pesquisas proporcionando resultados mais objetivos. (BIBLIOTECA VIRTUAL EM SAÚDE, 2012).

Todavia, é importante salientar também que o DeCS representa descritores utilizados nas Ciências da Saúde. Para Krieger e Finato (2004, p. 60): “os descritores, unidades que sustentam a linguagem documentária, possuem certas qualidades que permitem compará-los aos termos”.

Da mesma forma que foi realizada a validação em situação discursiva, nesta etapa cada *tag* também foi analisada uma a uma para verificar se estavam presentes no DeCS. Constatou-se que do *corpus* de pesquisa constituído por 104 (cento e quatro) *tags*, 66 (sessenta e seis) rótulos foram validados *in vitro* o que significa dizer que representam 63 por cento do conjunto de *tags*.

5.5 TAGS QUE NÃO ESTÃO NEM NO DISCURSO NEM NO DECS

Nesta categoria foram analisados os resultados obtidos em relação às *tags* que foram atribuídas por usuários no *CiteUlike*, porém que não estão no discurso dos especialistas e no DeCS.

De acordo com Strehl (2001, p. 111) este fato pode ter ocorrido por que: “[...] nas *folksonomias* a atribuição de marcadores está relacionada com a identificação da significação particular dos documentos para cada um dos usuários do sistema, não sendo, necessariamente, representativas dos conceitos ali tratado.” O objeto é alvo de diversos olhares que interpretam uma mesma ideia com unidades lexicais que nem sempre representam os conteúdos ao qual estão se referindo.

Desta forma, pode-se dizer que algumas *tags* não foram validadas nos textos e no DeCS por não terem relação com os assuntos ali registrados. Porém, conforme Krieger e Finato (2004, p. 207, grifo nosso): “[...] no caso de se verificar que seja baixa ou **ausente** a reiteração de uma determinada palavra conceitualmente importante, poder-se-ia explorar se o grau de coesão lexical, num texto sob consideração, seria ou não prejudicado”.

Neste sentido, foi observado que 29 (vinte e nove) *tags* não foram encontradas em nenhuma das situações representando um total de 28 por cento do conjunto. Esta situação pode ser justificada uma vez que o *corpus* textual desta pesquisa foi constituído por um número limitado de artigos.

Diante do exposto, entende-se que este conjunto de *tags* podem ser unidades lexicais típicas da área médica, mas que neste trabalho não ocorreram nem na situação discursiva nem *in vitro*. Como exemplo, é possível destacar a *tag Chemotherapy* (quimioterapia em português) que segundo o Instituto Nacional de Câncer (2012) significa: “método que utiliza compostos químicos, chamados quimioterápicos, no tratamento de doenças causadas por agentes biológicos”.

Através desse exemplo é provável que a palavra *Chemotherapy* pertença ao léxico da especialidade de Oncologia. Para verificar se realmente estas 29 (vinte e nove) *tags* são termos sugere-se que este conjunto de rótulos passe por um novo

estudo a fim de validá-las. Uma vez que as mesmas podem estar denominando novos conceitos da área ou constituírem variantes de termos já consagrados.

5.6 TAGS QUE ESTÃO NO DISCURSO E NO DECS

Nesta categoria foram verificados os rótulos que estavam presentes na situação discursiva e no DeCS. Para van der Laan (2002, p.63), “um termo é a designação de um conceito por uma expressão lingüística em uma linguagem de especialidade”. Desta forma, as *tags* que ocorrem no discurso dos especialistas são termos.

Cruzando os dados obtidos entre as *tags* validadas na situação discursiva e no DeCS verificou-se que 55 (cinquenta e cinco) *tags* estão expressas nas duas situações, ou seja, são termos por estarem duplamente validadas, sinalizando que o pressuposto de pesquisa é verdadeiro.

Ao verificar o *corpus* de estudo percebeu-se que nove *tags* foram expressas apenas no discurso do especialista. Entre as *tags* que ocorreram nesta categoria destacam-se: *Psyco-Oncology* com 38 (trinta e oito) ocorrências; *Cancer Therapy* com 14 (catorze) ocorrências; e *Cancer Imaging* com nove ocorrências. De acordo com van der Laan (2002, p. 63), “[...] uma unidade lexical adquire valor de termo dentro de um determinado contexto, ou seja, uma situação discursiva real”.

Também se pôde observar que 11 (onze) *tags* estão presentes apenas no DeCS, sendo elas: *Advace Care Planning*, *Biomarkers*, *Clinical-Trial*, *Continuum Of Care*, *Drug Discovery*, *Gene Expression*, *Molecular-Biology*, *Motivation*, *Mutation*, *Oncogenese* e *Prognosis*. Conforme Krieger e Finato (2004, p. 61): “o descritor, tal como o termo, revela sua funcionalidade no processo comunicacional”.

Neste sentido, por entender que o DeCS é uma importante ferramenta e que possui relevância em relação à organização da informação na área da Saúde é possível dizer que as 11 (onze) *tags* validadas neste vocabulário também são termos específicos da área de Oncologia. Isto porque na perspectiva deste trabalho, termos são unidades lexicais manifestas do discurso dos especialistas e que se entende por descritor como sendo uma unidade de representação de informações especializadas. Logo, o descritor é um termo.

6 CONSIDERAÇÕES FINAIS

Este trabalho teve a intenção de proporcionar uma discussão sobre a possibilidade das *tags* possuírem estatuto terminológico, considerando o fato de que esta prática de atribuir rótulos em ambientes virtuais, por pessoas não especializadas, tem se tornado cada vez mais comum no ambiente da *web*.

A partir da coleta realizada e de todas as etapas propostas neste trabalho foi possível obter resultados relevantes que prestaram sustentação para comprovar o pressuposto de pesquisa que questionava se *tags* podem ser termos. A análise realizada no *corpus* de estudo mostrou que ocorreram quatro situações distintas: *tags* que foram validadas no discurso, *tags* que foram validadas no DeCS, *tags* que não foram expressas nem no discurso nem no DeCS e *tags* que estão no discurso do especialista e no DeCS.

Deste modo, ao analisar o *corpus* de estudo constituído por 104 (cento e quatro) *tags* foi constatado que: 75 (setenta e cinco) *tags*, ou seja, 72 por cento dos rótulos analisados nesta pesquisa podem ser considerados termos por terem sido validados na situação discursiva e no vocabulário da área da saúde através do DeCS. Então, confirma-se o pressuposto de que algumas *tags* atribuídas por usuários no *CiteUlike* sobre Oncologia são termos.

Contatou-se também que 29 (vinte e nove) *tags*, ou seja, 28 por cento dos marcadores não puderam ser validados. Isto não significa que as mesmas não tenham estatuto terminológico. Para confirmar a sua condição de termos se faz necessário outros estudos para validação das mesmas.

Diante do exposto foi possível afirmar ainda que os ambientes de compartilhamento podem ser mais uma fonte de coleta e identificação de prováveis termos. Esta constatação pode ser relevante para a comunidade científica da área da Ciência da Informação uma vez que surge mais uma fonte de consulta para respaldar a construção de vocabulários controlados.

REFERÊNCIAS

ALCARÁ, Adriana Rosecler, et al. Fatores que influenciam o compartilhamento da informação e do conhecimento. **Perspectivas em Ciência da Informação**, Belo Horizonte, v. 14, n. 1, p. 170-191, jan./abr. 2009. Disponível em: <<http://www.scielo.br/pdf/pci/v14n1/v14n1a12.pdf>> Acesso em: 14 out. 2011.

ALVARENGA, Lídia. Representação do conhecimento na perspectiva da ciência da informação em tempo e espaço digitais. **Encontros Bibli**, Florianópolis, n. 15, p. 18-40, jan./jun. 2003. Disponível: <<http://www.periodicos.ufsc.br/index.php/eb/article/view/1518-2924.2003v8n15p18/5233>> Acesso em: 18 out. 2011.

ALVES, Alessandra. BARBOSA, Ricardo Rodrigues. Influências e barreiras ao compartilhamento da informação: uma perspectiva teórica. **Ciência da Informação**, Brasília, DF, v. 39, n. 2, p.115-128, maio/ago., 2010. Disponível em: <<http://revista.ibict.br/ciinf/index.php/ciinf/article/viewFile/1822/1372>> Acesso em: 05 abr. 2012.

AQUINO, Maria Clara. Hipertexto 2.0, folksonomia e memória coletiva: um estudo das tags na organização da web. **E-Compós**, Brasília, v. 9, ago. 2007. Disponível em:< <http://revistas.ffclrp.usp.br/incid/article/view/50/pdf>>. Acesso em: 28 abr. 2012.

ARRUDA, Francisco Edmar Cialdine. Socioterminologia e ensino. **Sociodialeto**, Mato Grosso do Sul, v.1, n.5, p. 1-15, nov. 2011. Disponível em: <<http://www.sociodialeto.com.br/edicoes/10/13122011122432.pdf>> Acesso em: 28 abr. 2012.

BIBLIOTECA VIRTUAL EM SAÚDE. **DeCS**. Disponível em: <<http://decs.bvs.br/P/decsweb2012.htm>> Acesso em: 05 abr. 2012.

_____. **O que são vocabulários estruturados**. Disponível em: <<http://decs.bvs.br/P/aboutvocabp.htm>> Acesso em: 09 jun. 2012.

CABRÉ, Maria Teresa. **La Terminología**: teoría, metodología, aplicaciones. Barcelona: Antártida/Empúries, 1993.

_____. La terminología hoy: concepciones, tendencias y aplicaciones. **Ciência da Informação**, Brasília, v. 24, n. 3, 1995. Disponível em: <<http://revista.ibict.br/cienciadainformacao/index.php/ciinf/article/view/487/442>> Acesso em: 19 out. 2011. [documento não paginado].

CANO, Waldenice Moreria. Uma aplicação pedagógica da terminologia: um dicionário escolar em ciências. **GELNE**, [Nordeste], v.4, n.1, 2002. Disponível em: <http://www.gelne.ufc.br/revista_ano4_no1_43.pdf> Acesso em: 29 abr. 2012

_____ **La Terminología: representación y comunicación.** Elementos para uma teoria de base comunicativa y outros artículos. Barcelona: IULA; Universitat Pompeu Fabra, 1999.

CATARINO, Maria Elisabete; BAPTISTA, Ana Alice; Folksonomia: um novo conceito para a organização dos recursos digitais na Web. **DataGramZero**, [Brasil], v. 8, n. 3, jun. 2007. Disponível em: <http://www.dgz.org.br/jun07/Art_04.htm> Acesso em: 04 nov. 2011. [documento não paginado].

CITEULIKE. **Nuvem de tag.** Disponível em: <<http://www.citeulike.org/search/all?q=oncology>> Acesso em: 02 jul. 2012.

CITEULIKE. **What is CiteULike?** Disponível em: <<http://www.citeulike.org/>> Acesso em: 04 dez. 2011.

E-TERMOS. **O projeto e-terms.** Disponível em: <<http://www.etermos.cnptia.embrapa.br/>> Acesso em: 03 abr. 2012.

FAULSTIC, Enilde Leite de Jesus. Variantes terminológicas: princípios lingüísticos de análise e método de recolha. In: **Réflexions méthodologiques sur le travail en terminologie et en terminotique dans les langues latines.** Nice: Realiter/Université de Nice Sophia-Antipolis, 1996. Disponível em: <<http://www.realiter.net/spip.php?article631>>. Acesso em: 28 abr. 2012.

_____ Termo e variação: tendências no português do Brasil. In: **Socioterminologia.** Brasília: UnB, 1998a. (Excerto, parte II).

_____ Entre a sincronia e a diacronia: variação terminológica no código e na língua. In: **VI Simpósio Ibero-Americano de Terminologia,** 1998b, Havana. Terminologia, Desenvolvimento e Identidade Nacional. Lisboa: Colibri, 1998b. p. 61-74. Disponível em: <<http://e-groups.unb.br/il/liv/enilde/documentos/HAVANA98.pdf>>. Acesso em: 14 out. 2011.

_____ A socioterminologia na comunicação científica e técnica. **Ciência e Cultura,** São Paulo, v. 58, n. 2, p. 27-31, jun. 2006. Disponível em: <<http://cienciaecultura.bvs.br/pdf/cic/v58n2/a12v58n2.pdf>> Acesso em: 04 dez. 2011.

GALDO, Alessandra; VIEIRA, Angel Freddy Godoy; RODRIGUES, Rosângela Schwarz. Classificação social da informação na web: tecnologia, informação e gente. **DataGramZero**, [Brasil], v.10, n.8, dez. 2009. Disponível em: <http://dgz.org.br/dez09/Art_03.htm>. Acesso em: 20 abr. 2012. [documento não paginado].

GONZALEZ, Marco; LIMA, Vera Lúcia Strube de. **Recuperação de Informação e Expansão Automática de Consulta com Thesaurus:** uma avaliação. Disponível em:

<<http://www.inf.pucrs.br/~gonzalez/docs/expanconsul.pdf>> Acesso em: 18 out. 2011. p. 1-8

GOUVÊA, Cleber; LOH, Stanley. Folksonomias: identificação de padrões na seleção *tags* para descrever conteúdos. **RESI**, Curitiba, Paraná, v. 6, n. 2, p. 1-8, 2007. Disponível em: <<http://revistas.facecla.com.br/index.php/reinfo/article/view/214/118>> Acesso em: 19 out. 2011.

INSTITUTO NACIONAL DE CÂNCER. **Estimativa de incidência de câncer no Brasil para o ano de 2012**. Disponível em: <<http://www.inca.gov.br/estimativa/2012/index.asp?ID=5>> Acesso em: 09 jun. 2012.

_____ **Quimioterapia**. Disponível em: <http://www.inca.gov.br/conteudo_view.asp?id=101> Acesso em: 09 jun. 2012.

KATO, David; SILVA, Gledson. **Folksonomia**: características, funcionamento e aplicações. Disponível em: <<http://www.terraforum.com.br/biblioteca/Documents/artigo-david-gledson.pdf>> Acesso em: 25 abr. 2012. p. 1-8.

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

MARTELETO, Regina Maria. Análise de redes sociais: aplicação nos estudos de transferência da informação. **Ciência da Informação**, Brasília, v. 30, n. 1, p. 71-81, jan./abr. 2001. Disponível em: <<http://dici.ibict.br/archive/00000204/01/Ci%5B1%5D.Inf-2004-261.pdf>> Acesso em: 20 out. 2011.

NOVELLINO, Maria Salet Ferreira. Instrumentos e metodologias de representação da informação. **Informação & Informação**, Londrina, v.1, n.2, p. 37-45, jul./dez. 1996. Disponível em: <<http://www.uel.br/revistas/uel/index.php/informacao/article/view/1603/1358>> Acesso em: 21 nov. 2011.

ONCOGUIA. **O que é Oncologia?** Disponível em: <<http://www.oncoguia.com.br/site/interna.php?cat=2&id=474&menu=2>> Acesso em 15 nov. 2011.

ORGANIZAÇÃO MUNDIAL DA SAÚDE. **Será que aumenta ou diminui o número de casos de câncer no mundo?** Disponível em <<http://www.who.int/features/qa/15/es/index.html>> Acesso em: 10 abr. 2012.

PAVEL, Sílvia, et al. **O que é terminologia**. Disponível em: <<http://www.btb.termiumplus.gc.ca/didacticiel-tutorial/lecon-lesson-1/lecon-lesson-1-2-2-por.html>> Acesso em: 17 abr. 2012. [documento não paginado].

PEREIRA, Débora de Carvalho; CRUZ, Ruleandson. Folksonomia e tags afetivas: comunicação e comportamento informacional no Twitter. **DataGramZero**, [Brasil], v.11, n.6 dez. 2010. Disponível em: < http://www.dgz.org.br/dez10/Art_06.htm> Acesso em: 15 abr. 2012.

PRIMO, Alex Fernando Teixeira. O aspecto relacional das interações na Web 2.0. **E-Compós**, [Brasília], v. 9, p. 1-15, 2007. Disponível em: < <http://www6.ufrgs.br/limc/PDFs/web2.pdf>> Acesso em: 05 dez. 2011.

RAPETTI, Luciano. **Folksonomia**: organização e uso da informação na web. 2007. 55 f. Trabalho de conclusão (graduação) - Universidade Federal do Rio Grande do Sul. Faculdade de Biblioteconomia e Comunicação, Porto Alegre, 2007. Disponível em: <<http://www.bibliotecadigital.ufrgs.br/da.php?nrb=000667018&loc=2010&l=c950363a71234fcd>> Acesso em: 17 nov. 2011.

RECUERO, Raquel da Cunha. Redes Sociais na Internet: considerações iniciais. **E Compós**, [Porto Alegre], v. 2, p. 1-20, 2005. Disponível em: <http://www6.ufrgs.br/limc/PDFs/redes_sociais.pdf> Acesso em: 17 nov. 2011.

SANTANA, Anderson de; SANTANA, Ana Lucia de Viveiros de. *Folksonomia*: uma análise de sua operacionalidade e sua possível aplicabilidade na ciência da informação. In: **XVI Seminário Nacional de Bibliotecas Universitárias**, 2010, Rio de Janeiro, Anais..., out. 2010. Disponível em: < http://www.gapcongressos.com.br/eventos/z0070/trabalhos/final_222.pdf>. Acesso em: 19 out. 2011. [documento não paginado].

SANTAREM SEGUNDO, José Eduardo; VIDOTTI, Silvana Aparecida Borsetti Gregório. Rede de tags para recuperação da informação no contexto da representação iterativa. **InCID**, Ribeirão Preto, v. 2, n. 1, p. 86-109, 2011. Disponível em: < <http://revistas.ffclrp.usp.br/incid/article/view/50/pdf>>. Acesso em: 08 nov. 2011.

SILVA, Edna Lúcia da; MENEZES, Estera Muszkat. **Metodologia da pesquisa e elaboração de dissertação**. 2001. Disponível em: < http://projetos.inf.ufsc.br/arquivos/Metodologia_de_pesquisa_e_elaboracao_de_teses_e_dissertacoes_4ed.pdf> Acesso em: 21 nov. 2011. 121 p.

STREHL, Leticia. As folksonomias entre os conceitos e os pontos de acesso: as funções de descritores, citações e marcadores nos sistemas de recuperação da informação. **Perspectivas em Ciência da Informação**, Belo Horizonte, v.16, n.2, p.101-114, abr/jun. 2011. Disponível em: <<http://www.scielo.br/pdf/pci/v16n2/07.pdf>> Acesso em: 15 abr. 2012.

TEIXEIRA, Fábio Augusto Guimarães; DUQUE, Cláudio Gottschalg. A Recuperação da Informação e a colaboração de usuários na Web: novas oportunidades para a Comunicação. In: **II Congresso Internacional de Comunicación 3.0**, 2010, Salamanca, Anais..., p. 1-13, out. 2010. Disponível em: <<http://campus.usal.es/~comunicacion3punto0/comunicaciones/047.pdf>>. Acesso em: 21 nov. 2011.

TOMÁEL, Maria Inês; ALCARÁ, Adriana Rosecler; DI CHIARA, Ivone Guereiro. Das redes sociais à inovação. **Ciência da Informação**, Brasília, v. 34, n. 2, p. 93-104, maio/ago, 2005. Disponível em: <<http://www.scielo.br/pdf/ci/v34n2/28559.pdf>> Acesso em: 01 nov. 2011.

VAN DER LAAN, Regina Helena. **Tesouro e terminologia**: uma inter-relação lógica. 2002. 262 f. Tese (doutorado) - Universidade Federal do Rio Grande do Sul. Instituto de Letras. Programa de Pós-Graduação em Letras, Porto Alegre. Disponível em: <<http://www.lume.ufrgs.br/handle/10183/3487>>. Acesso em: 19 out. 2011.

VAN DER LAAN, Regina Helena; FERREIRA, Glória Isabel Sattamini; O Binômio Representação/Recuperação da Informação em Interface com a Teoria Comunicativa da Terminologia - TCT. In: **Congresso Internacional da Associação Brasileira de Lingüística**, 2007, Belo Horizonte. Anais..., 2007. p. 818-819.

VANDER WAL, Tomas. **Folksonomy**. Disponível em: <<http://vanderwal.net/folksonomy.html>> Acesso em: 17 nov. 2011. [documento não paginado]

WÜSTER, Eugen. **Introducción a la Teoría General de la Terminología y a la Lexicografía Terminológica**. Barcelona, Institut Univertari de Lingüística Aplicada/Universitat Pompeu Fabra, 1998.

APÊNDICE D - TAGS QUE ESTÃO NO DISCURSO E NO DECS

Número	<i>Tags</i>
1	Angiogenesis
2	bayes
3	bias
4	biochemistry
5	bioinformatics
6	biology
7	body
8	brain
9	breast
10	Breast-cancer
11	cancer
12	clinical trials
13	colon cancer
14	colorectal
15	colorectal cancer
16	drug
17	esophageal
18	esophagus
19	genetics
20	genomics
21	infection
22	kinase
23	mammography
24	medical
25	medicine
26	metabolism
27	methodology
28	microrna
29	mirna
30	molecular
31	neuroblastoma
32	nursing
33	oncogene
34	oncogenes

35	oncology
36	ovarian
37	pathway
38	patient
39	pharmacogenomics
40	prediction
41	rates
42	regression
43	rna interference
44	screening
45	seeking
46	specimen
47	spectroscopy
48	surveillance
49	survival
50	systems biology
51	therapy
52	treatment
53	trials
54	tumor
55	tumors

APÊNDICE E - TAGS EXPRESSAS APENAS NO DISCURSO

Número	<i>Tags</i>
1	anti-cancer
2	biomarker
3	biomathematics
4	cancer imaging
5	cancer therapy
6	improved patient care
7	oncogenesis
8	psycho-oncology
9	quality cancer care

APÊNDICE F - TAGS EXPRESSAS APENAS NO DECS

Número	<i>Tags</i>
1	advance care planning
2	biomarkes
3	clinical trial
4	continuum of care
5	drug discovery
6	gene expression
7	molecular biology
8	motivation
9	mutation
10	oncogenese
11	prognosis

APÊNDICE G - TAGS QUE NÃO FORAM VALIDADAS

Número	<i>Tags</i>
1	acute myeloid leukemia
2	advace care
3	bioimage
4	cancer biology
5	cancer clinical
6	cancer communication
7	cancer drug development
8	cancer drugs
9	cancer crosstalk
10	câncer drug disc
11	cancer system biology
12	chemotherapy
13	drug combination
14	drug disc
15	drug targets
16	epigenetics
17	gene set analysis
18	gene-signatures
19	genetic interactions
20	hepatotoxicity
21	mathematical biology
22	mirna and cancer
23	molecular photogenesis
24	pathway analysis
25	personalized medicine
26	plasma cell myeloma
27	psysician patient
28	reporter gene
29	research method