

CENTRO UNIVERSITÁRIO SENAC

Jéssica Cintra Pinto Silva

Robson Amaral Bacarin

**Uso de Folksonomia para Expansão de
Consultas na Web**

São Paulo

2009

JÉSSICA CINTRA PINTO SILVA
ROBSON AMARAL BACARIN

Uso de Folksonomia para Expansão de Consultas na Web

Trabalho de conclusão de curso apresentado ao Centro Universitário SENAC – Campus Santo Amaro, como exigência parcial para a obtenção do título de Bacharel em Ciência da Computação.

Orientador Prof. Fabrício J. Barth

São Paulo

2009

Silva, Jéssica Cintra Pinto

Bacarin, Robson Amaral

Uso de Folksonomia para Expansão de Consultas na Web /

Jéssica Cintra Pinto Silva e Robson Amaral Bacarin – São Paulo,

2009

64 fls.

Trabalho de Conclusão de Curso – Faculdade Senac de Ciências

Exatas e Tecnologia

Alunos: Jéssica Cintra Pinto Silva e Robson Amaral Bacarin

Título: Uso de Folksonomia para Expansão de Consultas na Web

Trabalho de conclusão de curso apresentado ao Centro Universitário SENAC – Campus Santo Amaro, como exigência parcial para a obtenção do título de Bacharel em Ciência da Computação.

Orientador Prof. Fabrício J. Barth

A banca examinadora dos Trabalhos de Conclusão em sessão pública realizada em 18/12/2009 considerou o(s) candidato(s):

1) Examinador (a): _____

2) Examinador (a): _____

3) Presidente: _____

AGRADECIMENTOS

Agradecemos ao nosso orientador: Prof. Fabrício J. Barth pela ajuda, dedicação e paciência ao longo deste ano.

Aos nossos familiares, a fé e esperança que depositaram em nós até o fim dessa etapa. E também o auxílio para todas as nossas necessidades.

Aos nossos amigos, a amizade, o companheirismo, o apoio e a convivência maravilhosa que tivemos durante esses anos.

Aos nossos professores, a paciência e a colaboração constante para o nosso crescimento.

"A imaginação é mais importante que a inteligência."

Albert Einstein

RESUMO

Este trabalho apresenta um método de expansão de consultas para sistemas de recuperação de informação multimídia na Web que leva em consideração tags de sistemas de folksonomia. Para demonstrar o objetivo proposto de maneira prática e funcional foi desenvolvido um protótipo de um sistema de recuperação de informação multimídia na Web. A proposta apresentada é avaliada utilizando um método qualitativo de pesquisa. A partir dos resultados obtidos nesta avaliação concluí-se que a utilização do conceito de folksonomia pode de fato trazer grandes melhoras na recuperação da informação. Os resultados encontrados indicam que 79% dos usuários consideraram os resultados do sistema como bons ou ótimos, e apenas 21% considerou os resultados entre razoáveis e ruins.

ABSTRACT

This paper presents a method of query expansion for multimedia information retrieval systems on the Web that takes into account tags of folksonomy systems. To demonstrate the proposed objectives in a practical and functional way, a prototype system was designed for multimedia information retrieval on the Web. The proposal presented here is evaluated using a qualitative method. From the results of this assessment, it is concluded that the use of the concept of folksonomy can in fact bring great improvements in information retrieval. The results indicate that 79% of all users considered the results of the system as good or excellent and only 21% considered the results as poor.

LISTA DE ILUSTRAÇÕES

Figura 1: <i>Tagging</i>	17
Figura 2: Nuvem de <i>tags</i>	18
Figura 3: Consulta em sites de busca x Consulta no protótipo	21
Figura 4: Seqüência de passos para realizar uma consulta no protótipo.....	22
Figura 5: Exemplo de contagem de <i>tags</i> para a freqüência simples.....	25
Figura 6: Exemplo de freqüência composta.....	27
Figura 7: Padrão HTML para recuperação de <i>tags</i>	31
Figura 8: Tela do <i>Folkaliza</i> com os resultados de Bookmarks	32
Figura 9: Preferências e criação do objeto “ <i>Photos</i> ” do <i>Flickr</i>	33
Figura 10: Recuperando <i>dados através da API</i> do <i>Flickr</i>	34
Figura 11: Tela do <i>Folkaliza</i> com os resultados de Imagens	34
Figura 12: Tela do <i>Folkaliza</i> com os resultados de Vídeos.....	36
Figura 13: Código XML retornado de uma consulta do <i>Yahoo</i>	38
Figura 14: Tela do <i>Folkaliza</i> com os resultados de Documentos	38
Figura 15: Dados do acesso ao site.....	41
Figura 16: Novos acessos x retorno.....	42
Figura 17: Dados da divulgação do site	42
Figura 18: Repostas dos usuários no questionário de avaliação	42
Figura 19: Gráfico das ferramentas de busca utilizadas	43
Figura 20: Gráfico das ferramentas conhecidas.....	44
Figura 21: Gráfico das ferramentas utilizadas.....	44
Figura 22: Gráfico de mídias mais procuradas na internet.....	45
Figura 23: Gráfico da satisfação dos resultados obtidos no <i>Folkaliza</i>	46
Figura 24: Gráfico dos resultados diferenciados	46
Figura 25: Gráfico do tempo de resposta dos resultados/mensagens	47
Figura 26: Gráfico de clareza na navegabilidade do sistema.....	47
Figura 27: Gráfico da estabilidade do site	49
Figura 28: Gráfico de situações críticas na utilização do <i>Folkaliza</i>	50
Figura 29: Gráfico da funcionalidade do site	50
Figura 30: Gráfico da utilidade das informações contidas no <i>Folkaliza</i>	51

LISTA DE TABELAS

Tabela 1: Exemplos para redução dos termos.....	23
Tabela 2: Lista de <i>stopwords</i> utilizadas em inglês e português	24

SUMÁRIO

1	Introdução	13
1.1	Objetivos.....	14
1.2	Protótipo	14
1.3	Estrutura do Trabalho	15
2	Folksonomia.....	16
2.1	A criação, recuperação e compartilhamento de <i>tags</i>	17
2.2	Vocabulário.....	18
2.2.1	Flexões gramaticais e de gêneros	18
2.2.2	Idiomas	19
2.2.3	Polissemia.....	19
2.2.4	Sinonímia	19
2.2.5	Intenção do usuário.....	20
2.3	Sugestão de <i>Tags</i>	20
3	Proposta.....	21
3.1	Métodos auxiliares.....	22
3.1.1	Algoritmo de <i>Stemming</i>	23
3.1.2	Stopwords	24
3.1	Frequência Simples.....	25
3.2	Frequência Composta	26
3.4	Recuperação Multimídia.....	28
4	Protótipo.....	29
4.1	Armazenamento dos Dados.....	29
4.2	<i>Delicious</i>	30
4.2.1	Recuperando <i>tags</i>	30
4.2.2	Recuperando Bookmarks.....	31
4.3	Flickr.....	33
4.4	<i>Youtube</i>	35
4.5	Yahoo	37
5	Método de avaliação	39
6	Resultados	41
6.1	Perfil da amostra	43
6.2	Desempenho do sistema.....	45

6.3 Usabilidade do sistema.....	47
7 Conclusão	52
8 Trabalhos Futuros	523
REFERÊNCIAS.....	55
GLOSSÁRIO.....	58
APÊNDICE A - Expansão de Consultas na Web	59
APÊNDICE B - Página inicial do <i>Folkaliza</i>	63
APÊNDICE B - Imagem de uma consulta realizada no <i>Folkaliza</i>	64

1 Introdução

Nem sempre os resultados obtidos pelos sistemas de recuperação de informação disponíveis na Web são os esperados pelo usuário que realizou uma consulta. Isto acontece por que: (i) em parte dos casos a consulta é ambígua, isto é, não especifica a real intenção de busca e, portanto, não reflete a necessidade de informação do usuário, e; (ii) o sistema de recuperação de informação faz uso apenas do valor sintático das palavras inseridas na consulta bem como das palavras contidas nos documentos, sem levar em consideração nenhum outro fator.

Atualmente, na Web é muito comum encontrar blogs, fóruns, documentos, imagens e vídeos que se utilizam de *tags*, conhecidas como etiqueta ou palavra-chave, para categorizar seu conteúdo. Também são muito comuns sites que centralizam estas *tags* e as disponibilizam a outros usuários com o objetivo de facilitar a organização e recuperação de itens. Com isso o usuário pode visualizar as *tags* de outros usuários, assim como identificar o grau de popularidade de cada tag no sistema, e acessar diversas informações relacionadas. O uso de tags para marcar determinados recursos na Web de maneira colaborativa é chamado de folksonomia.

Entende-se por Folksonomia a classificação de conteúdo na web feita de forma colaborativa pelos usuários que utilizam um sistema baseado em tal conceito. É utilizado em qualquer site que de alguma forma faz o compartilhamento de informações entre os usuários, já que cada um pode categorizar o conteúdo e posteriormente recuperá-lo. Pode-se visualizar a aplicação de folksonomia em grandes sucessos da web, como: *Delicious*¹, *Flickr*² e o *Youtube*³.

Alguns autores (MARTINS, 2009) destacam que o conceito de folksonomia pode ser útil para a organização da informação da seguinte maneira: ela reúne informações que se relacionam, e torna mais fácil o encontro da informação para

¹ <http://delicious.com/>

² <http://www.flickr.com/>

³ <http://www.youtube.com/>

os usuários, de maneira ágil, simples e eficiente visto que está classificada de acordo com o assunto a que diz respeito.

1.1 Objetivos

O objetivo deste trabalho é desenvolver um método de expansão de consultas para sistemas de recuperação de informação multimídia na Web que leva em consideração *tags* de sistemas de folksonomia. A idéia por trás do objetivo é eliminar parte da ambigüidade das consultas fornecidas pelo usuário e obter resultados relevantes. Além disso, espera-se que este método possa fornecer resultados diferenciados para o usuário, ou seja, resultados que outras ferramentas de recuperação de informação na Web não fornecem, principalmente levando-se em conta que quando um usuário tem a possibilidade de filtrar a sua consulta, a chance de obter um resultado mais satisfatório aumenta consideravelmente, isto é, ele será direcionado ao assunto ao qual tinha a intenção de buscar.

Para um melhor entendimento, pode-se considerar um exemplo em que um usuário digita a palavra câmbio. Nesse caso um resultado convencional traria conteúdos relacionados a “câmbio de carro” e “troca de moedas” misturados. Já com a expansão da consulta, o usuário pode optar se deseja os resultados relacionados a “câmbio de carro” ou “troca de moedas”, eliminando boa parte do conteúdo que ele não deseja consultar.

A recuperação de informação multimídia é utilizada no trabalho por ser um recurso muito eficiente se comparado com a capacidade de armazenamento de informação pelo usuário, pois a sua percepção para identificar e memorizar uma informação através de um gráfico ou imagem é muito mais eficaz que um simples texto.

1.2 Protótipo

A criação de um protótipo mostrou-se necessária para demonstrar o objetivo proposto de maneira prática e funcional. Foi criado um website para que os usuários possam realizar consultas, através da metodologia proposta neste trabalho e o resultado é composto por diferentes tipos de mídias, como: imagens, vídeos e textos em geral (documentos ou páginas HTML).

Este protótipo reflete uma necessidade que tem sido observada no comportamento dos usuários ao utilizar grandes sites colaborativos. Tal necessidade está associada a exibição de resultados multimídia de maneira mais conveniente.

1.3 Estrutura do Trabalho

Este trabalho está estruturado da seguinte maneira: são apresentados conceitos e exemplos relacionados à folksonomia no capítulo 2; no capítulo 3 é apresentada a proposta de trabalho; no capítulo 4 são apresentados detalhes sobre a implementação do protótipo; no capítulo 5 é descrito o método de avaliação e os resultados do questionário; no capítulo 6 são apresentados os resultados do questionário; no capítulo 7 é apresentada a conclusão do trabalho e por fim no capítulo 8 são apresentados opções para trabalhos futuros.

2 Folksonomia

A *Folksonomia* é caracterizada pela construção colaborativa da classificação do conteúdo a partir do vocabulário natural dos usuários que a utilizam. É uma analogia ao termo *Taxonomia*, de origem grega e que significa “a ciência de classificar coisas”. Para a classificação dos itens são utilizadas *tags*, que podem ser denominadas como palavras-chave ou etiquetas, e tem a finalidade de descrever ou mesmo categorizar um determinado conteúdo. É uma meta-informação que está visível a todos os usuários de um sistema de folksonomia. Através de um conjunto de tags, o usuário consegue descrever o conteúdo de um item que está relacionado a tag (PEREIRA, 2009).

O prefixo “Folks” (do inglês, **folks = pessoas**), indica que a classificação destas coisas, em especial neste caso, da informação, é definida por pessoas e para pessoas.

Para exemplificar, abaixo são listados alguns sites baseados em folksonomia na chamada Web 2.0:

- *Youtube*: site onde os usuários podem compartilhar vídeos pessoais. Este site permite que várias pessoas, autores dos vídeos ou não, possam marcar os vídeos com quantas *tags* desejarem;

- *Delicious*: site onde os usuários podem compartilhar sua lista de endereços de páginas online; A cada endereço podem estar associadas inúmeras tags;

- *Flickr*: site onde os usuários compartilham fotos. Da mesma maneira que acontece no *Youtube*, os usuários deste site podem atribuir às fotos inúmeras tags. No caso do Flickr, apenas os proprietários das fotos podem atribuir tags às imagens, e;

- *Twitter*: site de microblog com posts curtos onde a informação é classificada com as chamadas *hashtags*, como por exemplo: “#folksonomia”. Neste caso, as *hashtags* são inseridas dentro do conteúdo do item, o que é ligeiramente diferente dos demais exemplos.

Todos estes sites descritos acima permitem a classificação do conteúdo através das *tags*, o compartilhamento das mesmas entre todos os usuários e,

além disso, pode-se também encontrar o conteúdo publicado por outro usuário no sistema de busca pela sua *tag*.

2.1 A criação, recuperação e compartilhamento de *tags*

A prática de *tagging* pelos usuários pode ser vista na figura 1. Os usuários buscam informação na web, interagem com outros usuários através de sistemas baseados em folksonomia e cada um gera uma *tag* a partir do seu vocabulário pessoal.

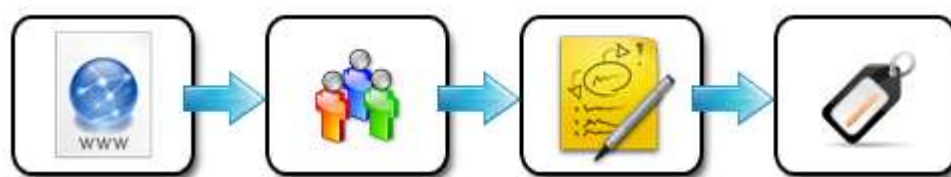


Figura 1: *Tagging*

Ao recuperar conteúdo através de uma *tag*, não é necessário que apenas uma palavra seja utilizada para recuperá-lo, visto que é possível adicionar mais de uma palavra para categorizar um conteúdo. Por exemplo: ao procurar pela palavra “desenvolvimento” serão encontrados resultados sobre, desenvolvimento de software, desenvolvimento sustentável, desenvolvimento da educação, desenvolvimento da economia, entre outros assuntos. Com isso a recuperação da informação por meio das *tags* é mais eficiente para encontrar resultados diferenciados, pois é a maneira que os usuários têm de organizar as informações na web, para que cada vez mais os resultados sejam encontrados de uma forma mais ágil e fácil (MARTINS, 2009).

Desta forma, é possível visualizar e compartilhar as *tags* com os outros usuários, permitindo a visualização de grau de popularidade de cada uma e também o acesso ao conteúdo relacionado. Conforme mostra a figura 2, as palavras que possuem maior destaque entre os usuários aparecem em uma nuvem de *tags* que estão disponíveis nos sites. Pode-se perceber que as *tags* que possuem maior número de ocorrências ficam em destaque nessa nuvem, seguindo níveis de popularidade até o de menor tamanho.

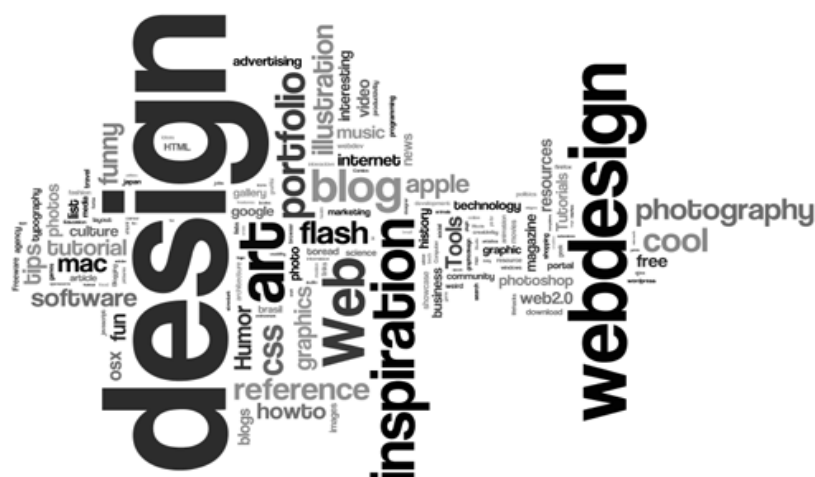


Figura 2: Nuvem de *tags*

2.2 Vocabulário

Com a facilidade que a folksonomia traz para que os usuários participem desse tipo de sistema, podem ser citadas como *vantagens* a facilidade na recuperação da informação e o surgimento de muitas informações relevantes em um curto espaço de tempo, tornando as aplicações cada vez mais ricas e interativas (AQUINO, 2007).

No entanto, sistemas baseados em folksonomia possuem algumas *desvantagens* para a recuperação e classificação da informação, dado que o aumento de marcações nos sistemas atinge grandes proporções, o que ocorre pela característica do sistema em não haver um controle para adicionar *tags*.

Como *desvantagens* pode-se citar: flexões gramaticais e de gêneros, idiomas, polissemia, sinonímia, intenção do usuário. Para melhorar essa prática descontrolada de categorizar conteúdo nos sistemas baseados em folksonomia, alguns deles vem adotando a prática de sugestão de *tags*, para evitar um crescimento além do inesperado de *tags*. Estes tópicos serão explicados a seguir de forma mais detalhada.

2.2.1 Flexões gramaticais e de gêneros

É muito comum encontrar *tags* referenciadas a um mesmo conteúdo no plural e singular, tais como: livro e livros, e animal e animais.

Também é muito freqüente *tags* que representam a mesma coisa, mas são escritas nos gêneros (masculino e feminino), exemplos: cidadão e cidadã, aluna e aluno, homem e mulher (também palavras que possuem radicais diferentes).

Em uma busca baseada em folksonomia os usuários tendem a utilizar em abundância as flexões gramaticais para categorizar conteúdo. Tais maneiras de escrita interferem diretamente no resultado da consulta, podendo inclusive fornecer um resultado diferente do esperado por conta da diferenças expostas acima.

2.2.2 Idiomas

Como a tendência de conhecer mais de um idioma é cada vez maior, os sistemas também tendem a disponibilizar informações em diversos idiomas. As redes sociais, mais famosas, estão disponíveis para todos os países, com isso é muito natural ter uma mesma *tag* em diferentes idiomas. Pode-se também categorizar um mesmo conteúdo com *tags* em idiomas diferentes, para o caso que o usuário deseja divulgar o conteúdo em outros idiomas que não são nativos. Por exemplo, o site do Ministério da Saúde pode ter as seguintes *tags*: saúde, health, salud e ministério.

Portanto, o idioma utilizado para rotular a informação influencia fundamentalmente na efetividade da busca, dado que a tradução de uma consulta muitas vezes pode não representar fielmente a intenção da busca.

2.2.3 Polissemia

Define-se como polissemia o fenômeno de uma mesma palavra possuir vários sentidos. Para exemplificar, a palavra '*jaguar*' possui três significados que podem se referenciar a: um animal, um carro e um sistema operacional.

Por causa desse fenômeno, uma busca pela palavra '*jaguar*' pode ter até três direcionamentos, que o usuário pode procurar por todos ou por um sentido em específico.

2.2.4 Sinonímia

Sinonímia é o ato de relacionar o mesmo significado para palavras diferentes. Os sinônimos são bastante utilizados pelos usuários para classificar a

informação, pois é comum classificar um mesmo conteúdo com várias palavras diferentes, e se as palavras forem analisadas poderá ser notado que também é comum existir mais de uma palavra que possui o mesmo significado.

Para esse caso pode-se citar a palavra 'casa', que tem seus vários sinônimos: lar, moradia, habitação.

2.2.5 Intenção do usuário

A criação das *tags* pode estar vinculada a caráter individual ou coletivo. Quando o usuário deseja marcar um conteúdo para que a recuperação da informação seja fácil para si, ele se utiliza de *tags* pessoais como: nome de pessoas, títulos de trabalhos, entre outros (AQUINO, 2007). Quando o usuário cria *tags* com a intenção de compartilhá-las com os demais usuários do sistema, as mesmas possuem um caráter coletivo, e para isso são utilizadas *tags* mais genéricas como: substantivos, adjetivos.

Logo, *tags* utilizadas para caráter individual prejudica a recuperação de informação em sistemas colaborativos, já que não é levado em consideração a classificação de conteúdo para todos os usuários, mas sim para si próprio.

2.3 Sugestão de Tags

Muitos dos sistemas colaborativos da Web têm adotado a prática de sugestão de *tags*. Mesmo que o campo de classificação de *tags* seja aberto, o sistema fornece as *tags* mais frequentes para determinado conteúdo.

Com isso a tendência dos usuários criarem *tags* diferentes dos demais é menor. A adoção dessa prática tende a trazer benefícios ao sistema como um todo, pois a recuperação de informação que utiliza termos mais genéricos proporciona ao usuário assuntos mais abrangentes. No entanto, utilizar termos mais específicos, que geralmente são criados pelos os usuários sem interferência dos sites, poderia classificar melhor determinado conteúdo (MARTINS, 2009).

3 Proposta

Neste capítulo é apresentado o desenvolvimento de um método para expansão de consultas que consiste em recuperar informação multimídia na Web levando em consideração *tags* de sistemas de folksonomia. Foram utilizadas duas técnicas para compor o método de expansão de consultas que são explicados detalhadamente a seguir, são eles: frequência simples e composta.

A proposta deste trabalho engloba a criação de um protótipo que o usuário consegue recuperar informação na web diferente dos sites casuais, onde o usuário submete uma consulta para o protótipo, escolhe uma opção de expansão para compor a sua consulta, denominada “consulta expandida”, e visualiza os resultados multimídia (figura 3).

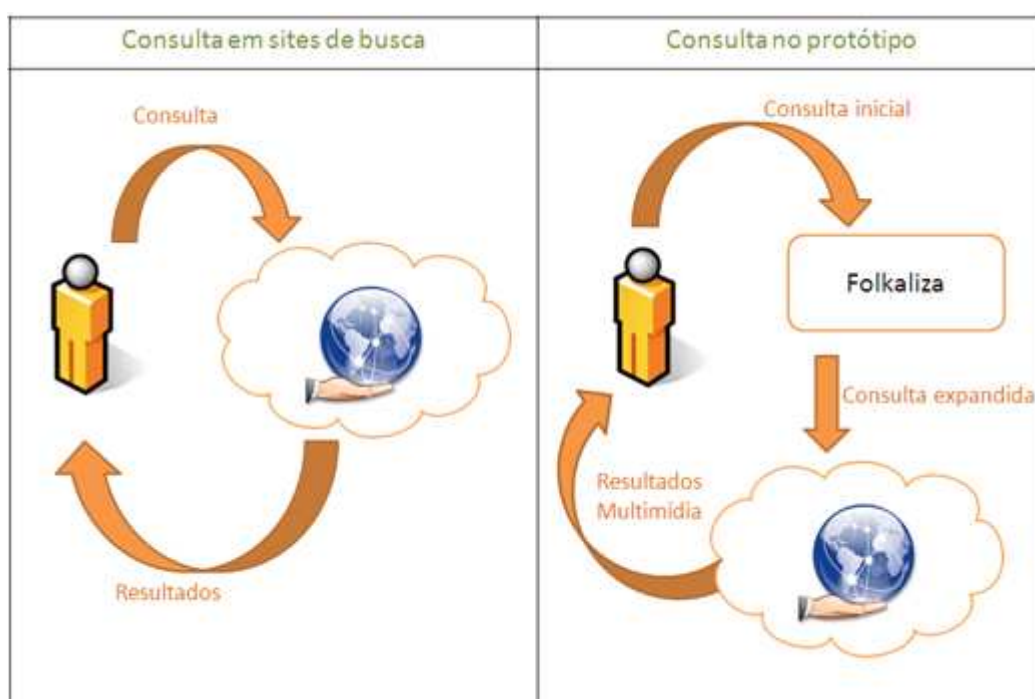


Figura 3: Consulta em sites de busca x Consulta no protótipo

O usuário digita uma ou mais palavras-chave e após escolher a opção para compor a expansão da consulta, são retornadas cinco *tags* mais relacionadas a ela. Com isso o usuário escolhe uma opção fornecida pelo site para compor a sua consulta final. No fim dessa interação é obtida uma consulta, sendo esta

composta pela consulta do usuário e pela opção da expansão da consulta. Essa seqüência de passos pode ser vista na figura 4.

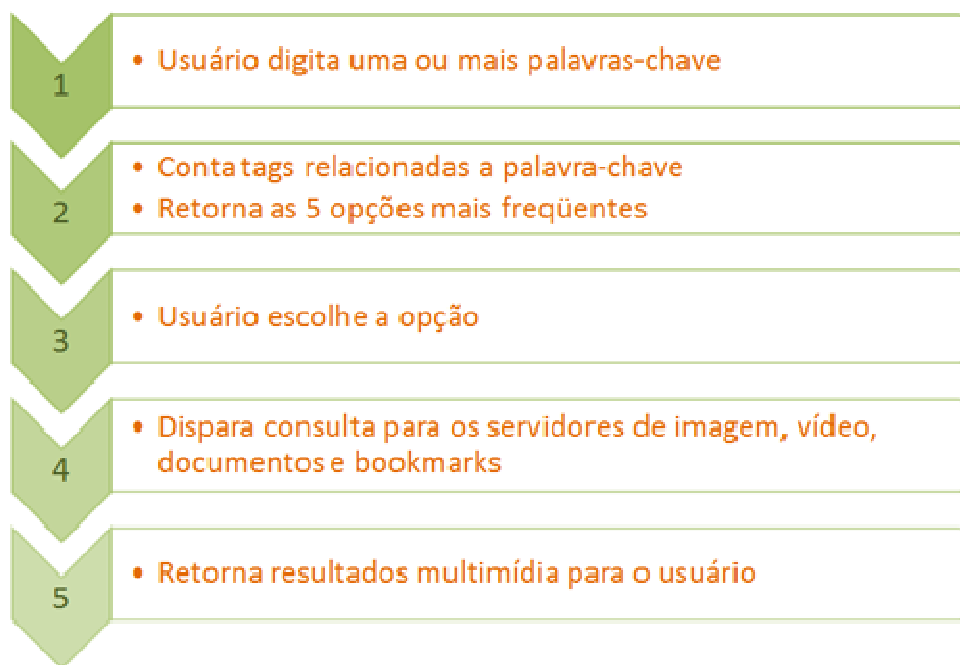


Figura 4: Seqüência de passos para realizar uma consulta no protótipo

Com isso o usuário pode visualizar um resultado final nos modos: texto, imagem e vídeo.

A ação de contar tags relacionadas (item 2 da seqüência de passos do protótipo) é implementada por dois métodos: freqüência simples e freqüência composta. A freqüência simples é utilizada quando o usuário digita apenas um termo para a consulta. Já a freqüência composta é utilizada quando o usuário digita duas ou mais palavras-chave para a consulta.

3.1 Métodos auxiliares

Para conseguir um resultado melhor, foram pesquisados alguns métodos para analisar as *palavras* que o usuário digita no campo de busca. Esses métodos são compostos por algoritmos que realizam um tratamento sobre a consulta inicial do usuário e seu uso foi fundamental para remoção de palavras comuns e que não adicionam significado a intenção de consulta do usuário. Foram analisados dois métodos auxiliares, que são: algoritmo de *Stemming* (ORENGO; HUYCK, 2001) e as *Stopwords* (YATES;NETO, 1999).

3.1.1 Algoritmo de *Stemming*

Foi necessário identificar um método que faz um tratamento na consulta do usuário para eliminar palavras com variações diferentes, mas que possuem o mesmo significado. Por exemplo, a palavra *balão* que pode ser encontrada tanto no singular e no plural, mas o seu significado é o mesmo alterando apenas a maneira de classificação do usuário.

Portanto, o primeiro método analisado foi o algoritmo de *Stemming*, que tem por objetivo reduzir termos derivados em apenas um termo no seu radical, retirando as diversas variações. Essa regra somente seria aplicada a termos escritos em português, a tabela 1 mostra algumas reduções aplicáveis nos termos para o plural, gênero, advérbios, entre outros que não foram citados na tabela 1.

Tabela 1: Exemplos para redução dos termos

Regra de redução	Sufixo	Tamanho mínimo da palavra sem sufixo	Substituir por	Exemplo
Plural	“ões”	3	“ão”	Balões - Balão
	“ais”	1	“al”	Normais – Normal
	“ns”	1	“m”	Bons – Bom
	“óis”	2	“ol”	Lençóis - Lençol
Feminino	“ã”	2	“ão”	Irmã – Irmão
	“ima”	3	“imo”	Prima - Primo
	“ora”	3	“or”	Professora-Professor
	“ica”	3	“ico”	Prática - Prático
Advérbio	“mente”	4	-	Felizmente - Feliz

Na implementação do código de redução, foi observado que não seria uma técnica aplicável, pois o *Delicious* possui *tags* em português e inglês. Com isso poderiam ocorrer reduções de palavras incorretas devido a mistura de idiomas. Como exemplo tem-se a palavra Ballons em inglês, se fosse aplicada a regra de redução do plural em português se transformaria em Ballom, que não seria a correta. Por isso foi abortada a implementação dessa lógica, já que poderiam ocorrer diversas reduções, ao termo para o radical, que são incorretas, tanto do

português para o inglês e como do inglês para o português. No entanto, seria uma ótima opção, caso o protótipo lidasse apenas com apenas um idioma.

3.1.2 Stopwords

Foram analisadas também as “*stopwords*”, palavras digitadas freqüentemente pelos usuários, mas que não influenciam no resultado de busca. Nessa lista, podem existir palavras desconsideradas dependendo da área de aplicação do sistema, como o site não se restringe a um tema são considerados apenas artigos, preposições e algumas conjunções. Com isso foram inclusas no desenvolvimento do sistema, as listas de *stopwords* em português e inglês. A seguir, na tabela 2, é possível observar alguns exemplos de palavras eliminadas das consultas elaboradas pelo usuário.

Tabela 2: Lista de *stopwords* utilizadas em inglês e português

Stopwords			
Português	agora	aquele	aqui
	como	este	foi
	início	maioria	muito
	nosso	outro	parte
	por	qual	seu
	tenho	um	usa
Inglês	about	after	all
	as	been	can
	but	could	for
	from	had	her
	if	into	last
	of	only	so

A aplicação desse método na consulta do usuário é interessante, pois apenas as palavras que influenciam a busca do usuário permanecem na consulta, reduzindo assim a quantidade de termos para serem analisados no site quando for feita a intersecção das *tags* utilizadas na expansão de consultas.

3.2 Freqüência Simples

A técnica de filtragem de *tags* por freqüência consiste em identificar quantas repetições ocorrem para as *tags* relacionadas à consulta e tem por objetivo encontrar as cinco mais citadas, por usuários diversos, e que estão relacionadas com o termo pesquisado.

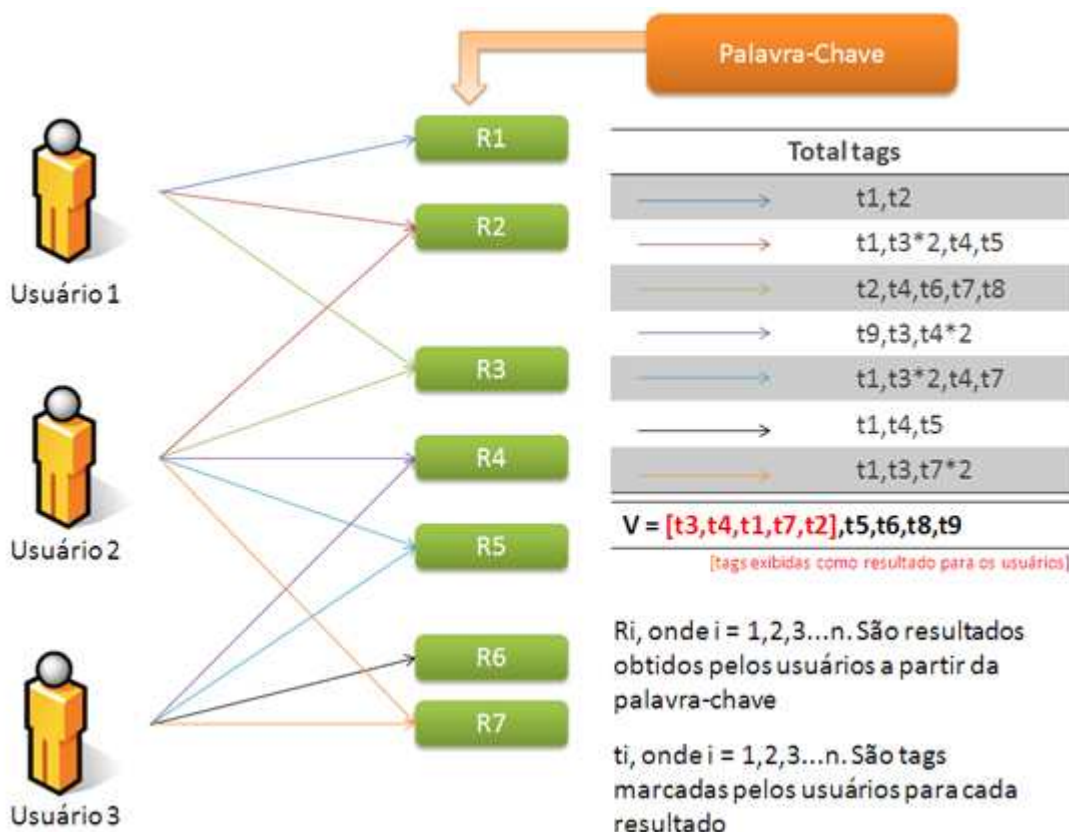


Figura 5: Exemplo de contagem de *tags* para a freqüência simples

A figura 5 representa como as cinco *tags* mais freqüentes são selecionadas. Quando o usuário elabora uma consulta são recuperados resultados diversos, representados por R_i, classificados por usuários do sistema, que no caso da figura 5 são exemplificados por três usuários. Cada um pode adicionar quantas *tags* desejar para qualquer conteúdo, então para cada R_i é obtida uma lista com a totalização de *tags*, que está representado por t_i na figura e está de acordo com a cor da seta correspondente ao R_i. São armazenados, em tempo de execução, no vetor V a soma da freqüência de cada *tag* da ordem da maior freqüência para a menor. Por exemplo, a t₃ que ocupa o primeiro índice do

vetor, foi citada uma vez em R_4 e R_7 e duas vezes em R_2 e R_5 , e por outro lado a t_9 que foi citada apenas uma vez em o R_4 está no último índice. Assim as *tags* que serão fornecidas para o usuário como opção de consulta são t_3 , t_4 , t_1 , t_7 e t_2 .

O motivo pelo qual optou-se por limitar em cinco *tags* utilizadas é o fato de que se fosse utilizada apenas uma, poderia não refletir na expansão desejada pelo usuário, enquanto utilizar um número muito grande poderia gerar confusão e menor desempenho de resposta na listagem das *tags* para expansão, como foi observado no protótipo. Pode ser citada a consulta por “michael jackson”, onde as opções de consulta poderiam variar entre: rei, pop, michael, jackson, michaeljackson, mjackson.

As *tags* citadas muitas vezes pelos usuários possuem maior relevância se comparadas com as *tags* pouco referenciadas, pois muitos usuários estão utilizando a mesma *tag* para determinado conteúdo. Porém essa técnica apresenta uma falha quando a *tag* pesquisada estiver relacionada a um assunto pouco abordado pelos usuários, pois geralmente as opções de *tags* relacionadas a um assunto pouco abordado são poucas e podem comprometer o resultado da consulta, de forma que não apresente nenhuma opção de expansão.

Outro detalhe muito importante a ser analisado é a questão da relação de frequência geral \times frequência diária. A frequência geral pode ser entendida como a quantidade de vezes que uma *tag* é citada levando em consideração um período de tempo maior, já a frequência diária leva em consideração *tags* citadas em apenas um dia. Portanto a frequência geral é mais precisa, já que leva em consideração o “tempo de vida” da utilização de uma *tag*.

Existem muitas *tags* que são citadas com muita frequência em determinadas épocas do ano, pois o assunto pelo qual ela se relaciona está em destaque. Para exemplificar esse caso, pode-se comentar sobre épocas de eleição, pois a tendência de criar *tags* que estejam relacionadas ao assunto é mais provável se comparado quando não está em destaque.

3.3 Frequência Composta

A técnica de frequência composta, ou refiltragem, é baseada na técnica de frequência simples, e poderia ser aplicada para uma consulta que contém mais de um termo. Por exemplo, na figura 5 supõe-se que o usuário elaborou a consulta

“Eu gosto de filmes de ação”, sem aplicar nenhum filtro para essa palavra seriam encontradas as *tags* mais freqüentes para cada termo desta consulta, ou seja, encontrar as cinco *tags* mais freqüentes para “eu”, “gosto”, “de”, “filmes” e “ação”. É possível perceber que existem termos desnecessários para a consulta, que não irão retornar um conteúdo relevante. Portanto, nesse caso é utilizado o método auxiliar das *stopwords* que excluem as palavras que não influenciam na consulta do usuário. Então aplicando tal método na consulta são obtidos os termos reduzidos “filmes ação”.

Assim é aplicado o método de freqüência simples para “filmes” e “ação”, e são selecionadas as *tags* mais freqüentes que estão no vetor V de cada uma. Com isso são armazenadas em um novo vetor V, as *tags* que se encontram tanto para o termo “filmes” quanto para “ação” e se encontram na região de intersecção entre o círculo verde e laranja representada na figura 6.



Figura 6: Exemplo de freqüência composta

Entretanto, há uma desvantagem em relação a este método, pois devido ao fato de buscar tags mais freqüentes para cada termo da consulta, o desempenho do sistema fica comprometido. E analisando o protótipo em termos de performance e tempo de resposta, uma consulta composta por mais de cinco termos, levou muito tempo para retornar um resultado. Por esse motivo, optou-se por não utilizar este método no protótipo.

3.4 Recuperação Multimídia

Neste trabalho foram recuperados artefatos multimídia, pois com essa funcionalidade o trabalho como um todo é enriquecido de informações devido as diferentes características e formas de representação dos artefatos, que podem ser vídeos, imagens e áudios (YATES;NETO, 1999). Ao recuperar informação multimídia a capacidade de percepção do usuário, com relação ao entendimento do conteúdo relacionado, é maior do que informações disponibilizadas em formato de texto.

É importante ressaltar que a única fonte de informação utilizada para recuperar esse tipo de conteúdo são as *tags* (meta-informações) cadastradas por usuários nos seus respectivos sites. Cabe lembrar que, é possível encontrar resultados não consistentes devido a categorização do usuário, ou até mesmo conteúdo impróprio para a consulta do usuário no protótipo.

4 Protótipo

O protótipo desenvolvido, intitulado Folkaliza⁴ tem como objetivo coletar *tags* do sistema de folksonomia *delicious* e expandir a busca do usuário fornecendo algumas possibilidades de termos para sua escolha. Quando o usuário selecionar o termo, automaticamente são recuperadas fotos do *flickr*, vídeos do *Youtube*, bookmarks do *delicious* e conteúdo texto que podem ser documentos (limitados em arquivos com as extensões “.pdf” e “.doc”) e páginas HTML, recuperados através do sistema de busca do *Yahoo*.

O resultado para a consulta do usuário por documentos é limitado nos formatos citados anteriormente, pois são formatos mais comuns e que possivelmente contenham conteúdo relevante para a sua consulta, descartando outros formatos como: **.rar**, **.xls**, entre outros.

Para o desenvolvimento do protótipo foi utilizada a linguagem de programação C# (ou C Sharp)⁵ e o banco de dados MySQL⁶. Foram utilizadas as APIs do *Yahoo* e do *Flickr*, já para o *Delicious* e *Youtube* foi utilizado um *crawler* (TORRES, 2006), que tem a finalidade de extrair informações das páginas de maneira que os sites sejam indexados em sistemas de busca.

4.1 Armazenamento dos Dados

Foram armazenados algumas informações no banco de dados. Para toda consulta realizada por usuários do *Folkaliza*, foram armazenadas os termos digitados, bem como a lista de *tags* relacionadas a ele a fim de manter um histórico de consultas.

No início deste trabalho, foi feita uma modelagem do banco de dados, para que todos os dados relacionados a consulta fossem armazenados, desta forma antes de procurar na base de dados do *Delicious*, primeiramente, uma consulta ao banco de dados do Folkaliza seria realizada. No entanto, foi possível perceber,

⁴ <http://www.jessicacindra.com.br/folkaliza/>

⁵ <http://msdn.microsoft.com/pt-br/vcsharp/default.aspx>

⁶ <http://www.mysql.com/>

que devido ao fato dos dados mudarem freqüentemente, as informações contidas no banco de dados não estariam atualizadas.

Assim, os dados são armazenados no banco, a fim de manter um histórico de consultas e não tornar esse banco de dados como fonte de informação para as consultas do protótipo.

4.2 *Delicious*

O *Delicious*⁷ é um site onde os usuários podem adicionar e pesquisar sites favoritos, chamados também de bookmarks. É uma ferramenta muito útil dado que os usuários podem catalogar bookmarks e acessá-los a qualquer momento. Além de compartilhar seus bookmarks com os usuários que participam da rede social, é possível ver quantos usuários adicionaram o mesmo link e suas respectivas *tags*.

Possui um número muito grande de usuários, de ordem de grandeza superior a 500.000, e o conteúdo em geral é muito variado, com mais de dois milhões de bookmarks compartilhados (NOLL;YEUNG, 2009).

Esse sistema de folksonomia foi escolhido para coletar *tags* dos usuários, por apresentar uma amostragem de usuários muito grande, de diversos perfis de utilização, que categorizaram inúmeros assuntos, sendo o sistema ideal para coletar *tags* para a expansão de consulta.

4.2.1 Recuperando *tags*

A recuperação das *tags* é fundamental na estruturação da consulta expandida e é realizada após a inserção dos termos que compõem inicialmente a consulta elaborada pelo usuário,

O termo consultado é submetido como um parâmetro a uma página da internet cujo endereço é [http://delicious.com/search?p="termo"](http://delicious.com/search?p=), adicionalmente alguns parâmetros são passados ao endereço para a definição de um período extenso ao qual desejam recuperar as *tags*. Com o código HTML gerado, foi possível analisá-lo para que fosse identificado um padrão (figura 6) que pudesse

⁷ <http://delicious.com/>

ser utilizado para desenvolver um crawler a fim de automatizar a recuperação de tags. Então foi possível coletar as tags relacionadas a este termo contidas no código HTML desta página utilizando este crawler.

```

<div class="tagdisplay">
  <ul class="tag-chain">
    <li class="tag-chain-item off first">
      <a class="tag-chain-item-link noplay" rel="tag"
        href="/tag/cursos;_ylt=A0wNBqD7exJLFTwB_K5WRh54;_ylv=3">
        <span class="tag-chain-item-span">cursos</span></a></li>
    <li class="tag-chain-item off">
      <a class="tag-chain-item-link noplay" rel="tag"
        href="/tag/design;_ylt=A0wNBqD7exJLFTwB_K5WRh54;_ylv=3">
        <span class="tag-chain-item-span">design</span></a></li>
    <li class="tag-chain-item off">
      <a class="tag-chain-item-link noplay" rel="tag"
        href="/tag/educa%C3%A7%C3%A3o;_ylt=A0wNBqD7exJLFTwB_K5WRh54;_ylv=3">
        <span class="tag-chain-item-span">educação</span></a></li>
    <li class="tag-chain-item off"><a class="tag-chain-item-link noplay" rel="tag"
      href="/tag/courses;_ylt=A0wNBqD7exJLFTwB_K5WRh54;_ylv=3">
        <span class="tag-chain-item-span">courses</span></a></li>
    <li class="tag-chain-item on last"><a class="tag-chain-item-link noplay" rel="tag"
      href="/tag/senac;_ylt=A0wNBqD7exJLFTwB_K5WRh54;_ylv=3">
        <span class="tag-chain-item-span">
          senac</span></a></li>
  </ul>
</div>

```

Figura 7: Padrão HTML para recuperação de tags

Realizando a consulta no site do *Delicious* pelo termo “senac”, um dos resultados retornou as tags: cursos, design, educação, courses e senac. Assim, são recuperadas todas as tags relacionadas ao termo através de uma página HTML, utilizando o padrão citado na figura 7.

Inicialmente a recuperação das tags relacionadas seria feita utilizando uma API do *Delicious* desenvolvida em C#. Ao realizar alguns testes, foi possível perceber que efetuando muitas consultas no *Folkaliza* em um intervalo de tempo curto, retornava um erro de limite de requisição.

Portanto, devido a essa limitação técnica optou-se por utilizar um *crawler*, para que não ocorresse nenhum problema impactante ao realizar consultas, já que esse módulo está relacionado com o objetivo principal deste trabalho, podendo comprometer a experiência do usuário.

4.2.2 Recuperando Bookmarks

Para a recuperação de bookmarks é feita uma interação com o site do *Delicious*, onde a consulta do usuário é submetida a um parâmetro embutido em

uma URL [http://delicious.com/search?p=""&chk=&sd6M=&ed=&fr=delicio_us&lc=1&atags=&rtags=&context=all](http://delicious.com/search?p=) de **consulta+opção de expansão**.

Tal interação é como se o usuário realizasse uma consulta pelo próprio site do *Delicious*. Após realizar uma consulta através da URL, um *crawler* recupera informação da consulta do usuário, de maneira semelhante a recuperação de *tags*, onde é encontrado um padrão no código HTML e por fim são selecionados os dados utilizados para os resultados do *Folkaliza*.

A figura 8 apresenta como os resultados de bookmarks são apresentados no *Folkaliza*.

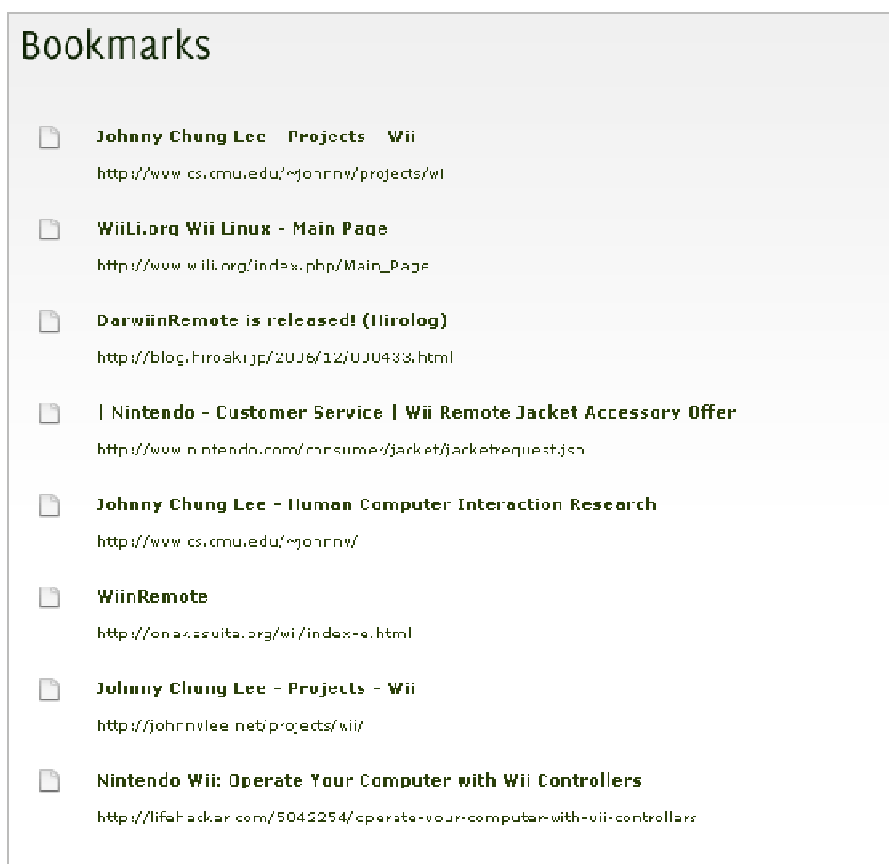


Figura 8: Tela do *Folkaliza* com os resultados de Bookmarks

4.3 Flickr

O *Flickr*⁸ é um site onde os usuários podem disponibilizar fotos, bem como tipos diferentes de documentos gráficos como: desenhos e ilustrações. Os usuários categorizam através de tags as imagens, ou mesmo álbuns, adicionados para tornar a recuperação de imagens um processo mais fácil e ágil para os usuários do site.

A API⁹ do *Flickr* está disponível para uso em site com fins não comerciais, e para utilizá-la foi necessário obter uma chave¹⁰ para a sua utilização no *Folkaliza*.

Para recuperar imagens do site, foi necessário definir as opções de busca, sendo elas a consulta expandida e a quantidade de fotos por página e criar um objeto do tipo *Photos*, conforme é mostrado na figura 9.

```
PhotoSearchOptions searchOptions = new PhotoSearchOptions();  
  
searchOptions.Tags = Session["expanded"].ToString();  
searchOptions.PerPage = 10;  
  
Photos photos = flickr.PhotosSearch(searchOptions);  
long pages = photos.TotalPages;
```

Figura 9: Preferências e criação do objeto “Photos” do Flickr

Com as opções definidas, a consulta é submetida a base de imagens do *Flickr* e os resultados retornados são apresentados no *Folkaliza*. A figura 10 demonstra o algoritmo utilizado para criar uma tabela de dados com os resultados que serão exibidos.

⁸ <http://www.flickr.com/>

⁹ <http://www.flickr.com/services/api/>

¹⁰ <http://www.flickr.com/services/api/keys/>

```

DataTable dt = new DataTable();
dt.Columns.Add("WebUri");
dt.Columns.Add("Title");
dt.Columns.Add("SquareThumbnailUri");
dt.Columns.Add("RawTags");
for (int i = 0; i < photos.PhotoCollection.Count; i++)
{
    if (i == results) break;
    PhotoInfo tag = flickr.TagsGetListPhoto(photos.PhotoCollection[i].PhotoId);
    string tags = "";
    for (int j = 0; j < tag.Tags.TagCollection.Length; j++)
    {
        if (j == 7) break;
        if (j != tag.Tags.TagCollection.Length - 1)
            tags = tags + tag.Tags.TagCollection[j].TagText + ", ";
        else
            tags = tags + tag.Tags.TagCollection[j].TagText;
    }
    dt.Rows.Add(photos.PhotoCollection[i].WebUri, photos.PhotoCollection[i].Title,
        photos.PhotoCollection[i].SquareThumbnailUri, tags);
}
PhotoRepeater.DataSource = dt;
PhotoRepeater.DataBind();
return dt;
}

```

Figura 10: Recuperando dados através da API do Flickr

No resultado das imagens, são utilizados os dados: url, miniatura e título da imagem, bem como as tags adicionadas para a imagem. Na figura 11 pode ser visualizado a tela do *Folkaliza*, quando o usuário procura resultados por imagem.

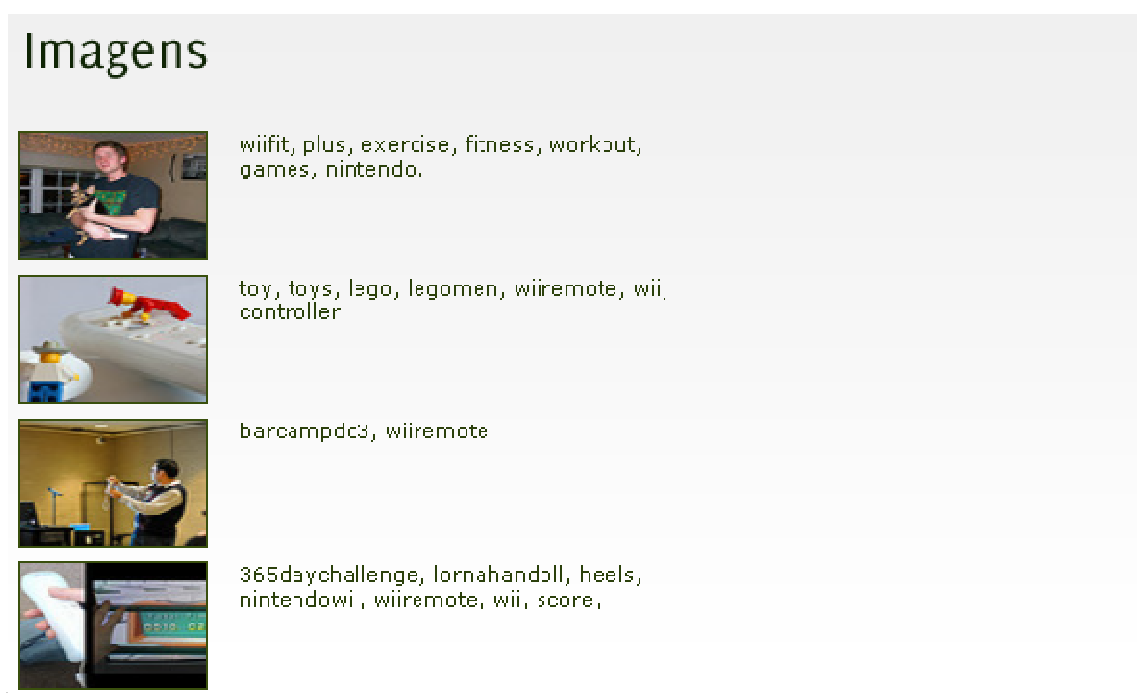


Figura 11: Tela do *Folkaliza* com os resultados de Imagens

Durante os testes, realizados em paralelo com o desenvolvimento, e na avaliação do sistema pelos usuários foi notado que em muitas consultas através

do *Folkaliza* não retornava nenhum resultado, no entanto a partir de consultas realizadas pelo site do *Flickr* foi possível observar o retorno de muitas imagens. O retorno de imagens diferentes deve-se ao fato de que a API do *Flickr* utiliza uma base de dados diferente da utilizada no site para que o seu desempenho não seja comprometido. A atualização dessa base de dados utilizada para a API não é feita automaticamente, ou seja, não é igual a base original, causando então essa diferença nos resultados.

4.4 Youtube

O *Youtube*¹¹ é um portal de armazenamento de vídeos adicionados por usuários, e que são compartilhados publicamente para todo usuário que desejar assisti-los. Os vídeos podem ser vistos de forma online, sem necessidade de realizar download.


Com relação à recuperação da informação no formato de vídeo, o sistema *Folkaliza* interage com o portal de vídeos *Youtube*, semelhante a interação do *Delicious*, onde a consulta é submetida a este site através de um parâmetro passado na sua URL ([http://www.youtube.com/results?search_query=consulta + opção de expansão&search_type=&aq=f](http://www.youtube.com/results?search_query=consulta+opção+de+expansão&search_type=&aq=f)). Assim um *crawler* recupera o código identificador dos vídeos encontrados na página, sendo possível adicionar a funcionalidade de visualizar os vídeos na própria página do *Folkaliza*.

Apesar de haver uma API para desenvolvimento utilizando o *Youtube* esta não foi utilizada devido a não oferecer um método que retornasse somente o código do vídeo para ser utilizado no reprodutor de vídeos do protótipo, sendo assim optou-se pela implementação de um *crawler*. No resultado de vídeos, são utilizados os dados: url, miniatura e título do vídeo, a tela do *Folkaliza* pode ser vista na figura 12.

¹¹ <http://www.youtube.com/>

Vídeos

Head Tracking for Desktop VR Di...
★★★★★




YouTube

0:00 / 4:45

Head Tracking for Desktop VR Displays using the WiiRemote

Johnny Lee: Wii Remote hacks
★★★★★



YouTube

0:00 / 6:14

Johnny Lee: Wii Remote hacks

Figura 12: Tela do *Folkaliza* com os resultados de Vídeos

Nem todos os vídeos podem ser exibidos pelo protótipo, e isso ocorre devido a restrição que o usuário pode impor ao vídeo, ao fazer o upload, de somente permitir sua visualização no próprio *Youtube*, ou seja, qualquer aplicação externa fica impedida de exibir este vídeo.

4.5 Yahoo

O *Yahoo*¹² é um portal que disponibiliza um buscador para a web, e foi escolhido para integrar o protótipo por oferecer a possibilidade de executar filtros avançados através de sua API.

Apesar de já ocorrer à consulta de bookmarks com o *Delicious*, optou-se pela recuperação de páginas HTML que de fato ofereçam conteúdo relacionado à consulta e que sejam documentos, visto que ao recuperar bookmarks são recuperados sites que contém links para outros sites.

Foi definida uma restrição dos formatos de conteúdo a serem recuperados, levando em consideração os mais utilizados para texto e esta restrição seleciona os seguintes formatos para a consulta: Microsoft Word (DOC), Portable Document Format (PDF) e Hyper Text Markup Language (HTML).

Para que seja possível obter resultados de páginas e documentos é realizada a submissão da consulta expandida ao buscador do Yahoo através de sua API, chamada *Yahoo Search Boss*¹³, destinada principalmente a desenvolvedores que tenham interesse em desenvolver soluções otimizadas de busca.

O funcionamento desta API é bastante simples. Basicamente foi necessário submeter a consulta já expandida a esta API através da URL http://boss.yahooapis.com/ysearch/web/v1/consulta+opção-de-expansão?appid=AICRBtrV34G_ICpjSKwHWuMB2VNEuPFI_BhhDN3AeA3cTYRdTts4PnLN4speHjanJ2g--&format=xml&type=tipo do arquivo.

Realizando uma consulta, por exemplo, com o termo computação e a opção de expansão senac, é obtido XML representado pela figura 13.

¹² <http://www.youtube.com/>

¹³ <http://developer.yahoo.com/search/boss/>

```

<result>
  <abstract>
    <b>...</b> CorelDraw (36h) - <b>SENAC</b> (2000) Introdução a Visual Basic (4h)
    - <b>SENAC</b> (2001) <b>...</b> Encontro Nacional dos Estudantes de
    <b>Computação</b>, ENEC, Campina Grande/PB, 2008 <b>...</b>
  </abstract>
  <clickurl>
    http://lrd.yahooapis.com/_ylc=X3oDMTV1aXFxNXJzBF9TAzIwM1MxNTI3MDIEYXBwaU
    QDQWxDUKJ0clYzNEdfbENvalNLdOhXdU1CM1ZORXVQRmxfQmhoRE4zQWVBM2NUUVVJkVHM0UG5MTjRzc
    GViamFuSjJnLSOEY2xpZW50A2Jvc3MEc2VydmljZQNC1NTBHNsawNOaXRzZQRzcmNvdmlkA21oXDVyV
    Wd1QXUzZ2ZoSEtORnlLMzYuZ31SdG9zRXNTbnljQUNwbEQ-/SIG=12dju9hbo/**http%3A//users.cjb.net
    /gustavovanderlinden/curriculum_gustavo2008_word.doc
  </clickurl>
  <date>2008/08/17</date>
  <dispurl>
    <b>users.cjb.net</b>/<b>...</b>
    <br>
    /curriculum_gustavo2008_word.doc
  </dispurl>
  <size>344576</size>
  <title>DOC (Word)</title>
  <url>
    http://users.cjb.net/gustavovanderlinden/curriculum_gustavo2008_word.doc
  </url>
</result>

```

Figura 13: Código XML retornado de uma consulta do Yahoo.

Para os resultados de documentos do *Folkaliza* são utilizados os dados: qual é o formato do documento, o título do documento e a URL. A figura 14 mostra um exemplo do resultado.



Figura 14: Tela do *Folkaliza* com os resultados de Documentos

5 Método de avaliação

A proposta apresentada neste trabalho é avaliada utilizando um método qualitativo de pesquisa. Uma amostra aleatória de pessoas foi convidada a utilizar o sistema implementado, e a responder um questionário após a utilização. O questionário respondido foi estruturado em: (i) questões para identificar o público que acessou e respondeu o questionário; (ii) questões para avaliar o desempenho do método de recuperação de informação do sistema, e; (iii) questões para avaliar a usabilidade do sistema. O questionário completo pode ser visto no Apêndice A.

Por ser um sistema disponibilizado na internet, o acesso ao Folkaliza não está restrito a nenhum tipo de público. Em termos de avaliação isso deve ser considerado, pois se o usuário já faz utilização de ferramentas semelhantes, então a idéia do sistema e a forma de utilizá-lo não será totalmente nova para o usuário. Por isso, é necessário identificar o perfil de quem respondeu o questionário.

Tradicionalmente, o desempenho de um sistema computacional é medido levando em consideração o tempo e a quantidade de memória utilizada para retornar uma resposta. No entanto, a avaliação de sistemas de recuperação de informação também deve considerar a qualidade das informações recuperadas para cada uma das consultas realizadas no sistema. Por isso, o questionário utilizado para avaliar o sistema Folkaliza possui questões para avaliar a qualidade das respostas fornecidas pelo sistema e o tempo de resposta.

Para avaliar a usabilidade do sistema, foram utilizadas algumas das dez heurísticas definidas por Jakob Nielsen (NIELSEN, 2009). Os objetivos principais eram identificar se os usuários estavam satisfeitos com a navegação e encontrar possíveis sugestões. Foram analisadas as seguintes heurísticas:

- *Visibilidade do usuário:* o usuário deve ser informado das situações que ocorrem no sistema e receber avisos em tempo real;
- *Consistência e padronização:* mede se o usuário não tem surpresas ao longo da navegação no sistema, sendo essas surpresas referentes a mensagens, situações ou ações;
- *Prevenção de erros:* o sistema deve possuir mensagens objetivas para o usuário e um cuidado, em termos de design, quando um

erro acontece no sistema. Para que o usuário não se perca durante a navegação no sistema ou mesmo fazer algo por engano é necessário eliminar situações de erro, que podem ser a apresentação de mensagens de erros claras ou a confirmação de ações.

- *Perceber ao invés de memorizar:* a navegação no sistema deve ser intuitiva, o usuário não pode “aprender” a utilizar o sistema na primeira vez e seguir sempre os mesmos passos das próximas vezes.

- *Eficiência de uso:* o sistema deve permitir acessibilidade tanto aos usuários experientes como os inexperientes;

- *Estética e Design:* avaliar se o sistema possui informações ou funcionalidades irrelevantes ao sistema, que quando desabilitadas no sistema não causam muito impacto; e

- *Ajuda e informações gerais:* avaliar se o sistema possui informações que ajude o usuário a entender e navegar no sistema naturalmente. Bem como informações relacionadas a contato, quando necessário.

6 Resultados

Os resultados deste trabalho foram obtidos utilizando um questionário que ficou disponível no período de 13/11/2009 à 27/11/2009 e foi divulgado através de listas de e-mail e twitter¹⁴. Para acompanhar e monitorar o acesso ao site foram utilizadas as ferramentas *Google Analytics*¹⁵ e o sistema de monitoramento do *Google Docs*¹⁶.

O período em que o site teve o maior número de acessos foi perto do início da divulgação, nos três primeiros dias. O site teve um total de 158 acessos únicos e totalizou 208 visitas no site, como é possível visualizar na figura 15.



Figura 15: Dados do acesso ao site

¹⁴ <http://www.twitter.com>

¹⁵ <http://www.google.com.br/analytics/>

¹⁶ <http://docs.google.com/>

O número de retorno dos usuários ao site foi baixo, 76,44% visitou apenas uma vez e 23,55% retornaram de 2 a 8 vezes o site (figura 16).

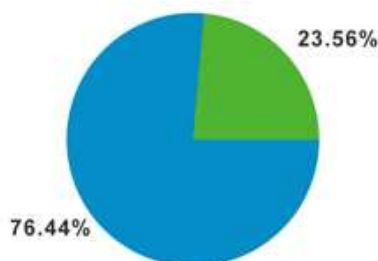


Figura 16: Novos acessos x retorno

Na figura 17 é possível visualizar alguns dados sobre os meios utilizados para acessar o Folkaliza. Os usuários acessaram o site através de: link direto, questionário de avaliação, servidores de e-mails, twitter, Orkut, entre outros.



Figura 17: Dados da divulgação do site

O questionário teve um total de 47 respostas. Durante o período em que ficou disponível, o maior número de respostas foi nos primeiros dias, porém houve um aumento na quantidade de respostas quando foi enviado em 20/11/2009 o segundo e-mail de divulgação. Na figura 18 é possível visualizar este comportamento.



Figura 18: Repostas dos usuários no questionário de avaliação

6.1 Perfil da amostra

As primeiras perguntas, de 1 a 7, tinham por objetivo identificar qual era o público que estava respondendo ao questionário. A primeira pergunta estava relacionada ao acesso do usuário a ferramentas de busca disponíveis na internet. Por unanimidade, 100% dos usuários utilizam algum sistema de busca.

A segunda pergunta questiona qual a ferramenta de busca que os usuários mais utilizam. Foram citadas as principais utilizadas atualmente, que são: Google, Yahoo, Bing¹⁷ e os usuários poderiam citar alguma outra ferramenta. Era possível escolher apenas uma resposta. Com 96% o *Google* é ferramenta mais utilizada, correspondente a um total de 45 respostas. Os 4%, que correspondem as 2 respostas restantes, foram citados pelos usuários a ferramenta de busca do *Altavista*¹⁸ e também consultas realizadas através dos navegadores, nesse caso pelo *Firefox*¹⁹ (figura 19).

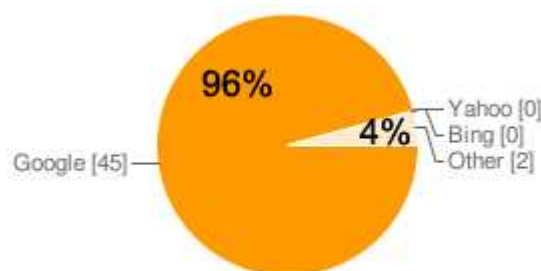


Figura 19: Gráfico das ferramentas de busca utilizadas

A terceira pergunta tinha por objetivo identificar se o usuário tinha conhecimento sobre as ferramentas da web 2.0, que utilizam *tags* para categorização de conteúdo. Foram citados os sites *Delicious*, *Flickr*, *Youtube* e *Wikipedia* e os usuários podiam escolher mais de uma opção de resposta. Os sites mais conhecidos, totalizando 100%, são *Youtube* e *Wikipedia*, a seguir o *Flickr* com 70% das respostas e por último com 53% o site do *Delicious*. (figura 20)

¹⁷ <http://www.bing.com>

¹⁸ <http://br.altavista.com/>

¹⁹ <http://br.mozdev.org/>

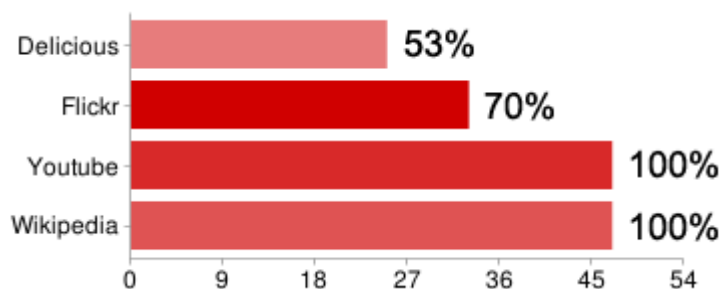


Figura 20: Gráfico das ferramentas conhecidas

A quarta pergunta era semelhante a terceira, porém foi pesquisado se utilizavam os sites *Delicious*, *Flickr*, *Youtube* e *Wikipedia*. Nesta questão, os usuários também podiam escolher mais de uma resposta e nesse caso observou-se que os sites mais utilizados, foram o *Youtube* e *Wikipedia* com 96%, a seguir o site do *Flickr* com 23% e por último o site do *Delicious* com 17% (figura 21).

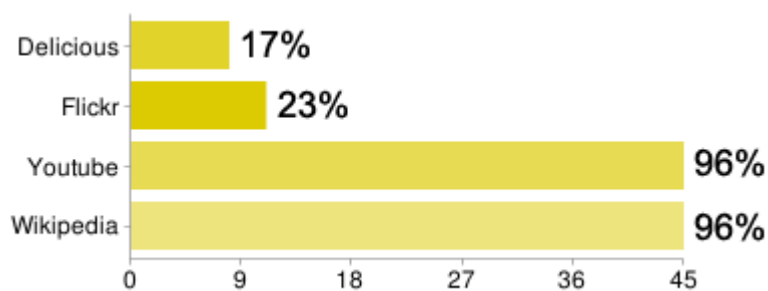


Figura 21: Gráfico das ferramentas utilizadas

A quinta pergunta, era aberta e, tinha a intenção de saber para qual finalidade esses sites eram utilizados pelos usuários. Ao analisar as respostas, foi possível perceber que o assunto entre os usuários era muito diferente. Muitos citaram que utilizam esses sites com a intenção de recuperar informações para fins acadêmicos ou escolares, e buscam por novas informações ou conhecimentos técnicos relacionadas a sua, respectiva, área de atuação no mercado de trabalho. Também citaram que o uso desses sites facilita a recuperação da informação, de forma que os usuários têm a capacidade de armazenar e compartilhar informações, referências para outros sites ou até mesmo para arquivos. A minoria dos usuários citou que utilizam para fins de entretenimento ou, para a tradução de textos, por exemplo.

A sexta pergunta tinha por objetivo analisar quais tipos de mídias os usuários costumam procurar na internet. Para identificar se as mídias utilizadas no *Folkaliza* eram relevantes. Foi possível escolher mais de uma opção.

De acordo com a pesquisa, 91% dos usuários citaram que documentos em geral é a mídia que mais costumam pesquisar. Posteriormente ficaram as mídias em imagens, vídeos e áudio, respectivamente com 77%, 74% e 64% da opinião dos usuários. Uma minoria, correspondente a 9%, citou arquivos no formato *torrent* e perfis de usuários na internet como outras mídias (figura 22).

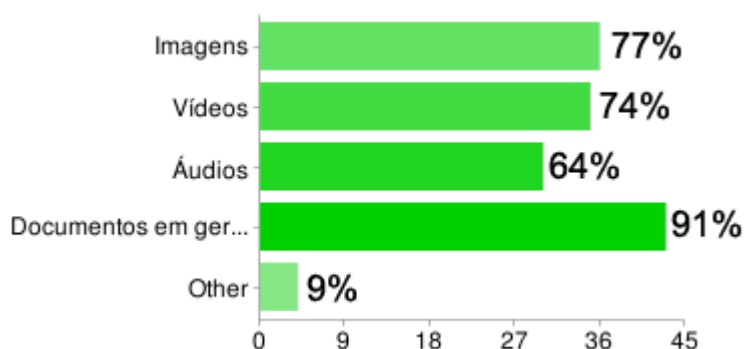


Figura 22: Gráfico de mídias mais procuradas na internet

Na sétima pergunta, o objetivo era identificar se os usuários consideravam útil a recuperação de informação multimídia. Todos os usuários consideram útil resultados que são compostos por diferentes mídias.

6.2 Desempenho do sistema

O objetivo das perguntas 8, 9 e 19 era de coletar informações sobre o desempenho do *Folkaliza*. Na oitava pergunta os usuários tinham que estipular uma nota, considerando a escala de 1 à 4, onde 1 era um resultado ruim e 4 um resultado ótimo. Esta pergunta tinha por objetivo quantificar o nível de satisfação dos resultados obtidos no *Folkaliza*, que está completamente relacionada ao objetivo principal deste trabalho. A opinião em relação ao sistema foi satisfatória pois 79% consideraram a consulta realizada como um resultado positivo, ou seja, os resultados oscilavam entre ótimo e bom, já 21% tenderam para um resultado contrário que se encontra em um nível regular a ruim. Na figura 23 é possível visualizar detalhes sobre os resultados desta questão.

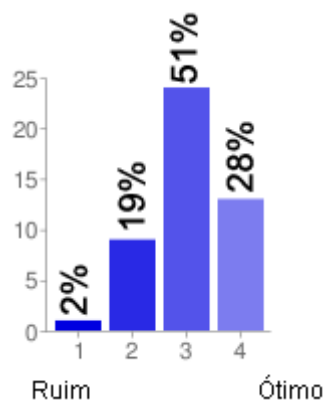


Figura 23: Gráfico da satisfação dos resultados obtidos no *Folkaliza*

Juntamente com o nível de satisfação dos resultados, são analisados também o quanto esses resultados são diferentes dos encontrados em outros sites de recuperação de informação. O resultado da avaliação ficou muito dividido, 51% consideraram que o *Folkaliza* retornou resultados diferenciados, por outro lado 49% responderam que os resultados que foram encontrados não eram diferenciados se comparados com outras ferramentas de busca (figura 24).

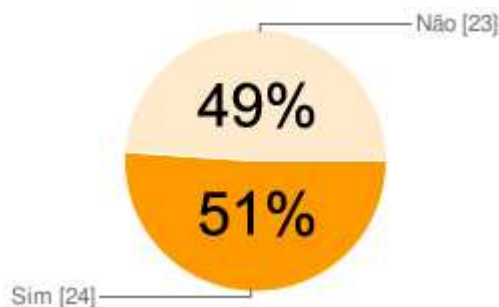


Figura 24: Gráfico dos resultados diferenciados

O tempo de resposta de mensagens e resultados foi abordado na questão dezenove, os usuários deveriam escolher apenas uma resposta. Um total de 72% considerou que o sistema não demorou em retornar a sua consulta, porém 28% foram contrários a essa opinião (figura 25).

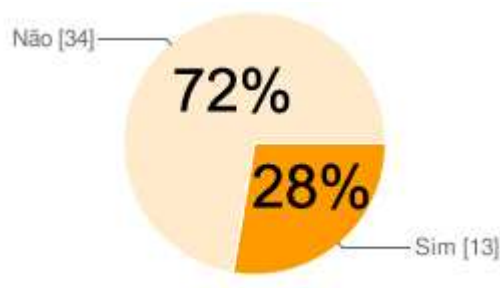


Figura 25: Gráfico do tempo de resposta dos resultados/mensagens

6.3 Usabilidade do sistema

A partir da décima questão até a décima oitava, são analisados aspectos de usabilidade do sistema e são baseados nas heurísticas descritas na sessão 5. Para essas perguntas os usuários deveriam responder “Sim” ou “Não”, para as questões 10, 12, 14 e 16, e poderiam comentar nas perguntas abertas sobre considerações relacionadas a pergunta anterior.

A décima questão identifica se o usuário se sentiu confuso em qualquer ação realizada no *Folkaliza*. Na figura 26 é possível visualizar o resultado.

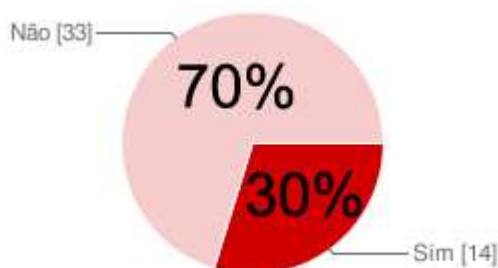


Figura 26: Gráfico de clareza na navegabilidade do sistema

Boa parte dos usuários, 70% no total, não se sentiram confusos, no entanto 30% responderam que se sentiram confusos em alguma situação, principalmente relacionado a expansão de consultas.

Muitos citaram que estavam confusos por conta dos termos encontrados como opção de expansão que eram iguais a um termo da consulta, bem como termos que tinham o mesmo significado, mas estavam em diferentes idiomas. Também relacionado a opção de expansão, ficou confuso o fluxo de navegação para obter um resultado, por exemplo o usuário gostaria de ter um resultado ao

clicar apenas no expandir ao invés de ter que escolher a opção “NÃO DESEJO UTILIZAR OS ITENS ACIMA”. Um problema também citado por muitos usuários foi ao realizar uma nova consulta, pois os resultados da consulta anterior permaneciam na tela, mas tinham consultado coisas diferentes do que ele estava visualizando. Também foram realizadas consultas muito extensas ou sem ligação nenhuma, com isso não era retornada nenhuma opção de expansão, apenas os campos vazios e a escolha de não utilizar nenhum item de expansão.

Tiveram problemas com algumas consultas que não retornavam nenhuma opção de expansão, porém eram retornados termos vazios juntamente com a opção de não utilizar as opções fornecidas pelo *Folkaliza*.

Um dos usuários citou um fato importante, a não sinalização de qual opção de expansão foi escolhida, caso ele fosse alternando as opções para visualizar todos os resultados não conseguia identificar a opção escolhida anteriormente. Com os comentários dos usuários, foram notados que para alguns casos a aplicação do método auxiliar das *stopwords* não era aplicada da maneira correta. Pode-se citar como exemplo uma consulta por “São Paulo”, o termo “são” era retirado da consulta, fazendo com que as opções da expansão de consultas fossem relacionadas a escritores, time de futebol ou blogs pessoais, e na verdade o usuário esperava por uma opção relacionada a cidade de São Paulo.

Alguns outros fatores foram citados por uma minoria de usuários, como: frustração em não identificar qual fonte de informação que estava sendo utilizada no *Folkaliza*, não ter a opção de selecionar um idioma para realizar as consultas, a navegação nos menus superior e de navegação não ser muito intuitiva e, por fim, mensagens dispostas nas páginas internas.

A décima segunda pergunta identifica se encontraram alguma situação que o sistema não agiu da maneira certa ou esperada. Como pode ser visto na figura 27, 85% opinaram que o sistema não se comportou de maneira inesperada, no entanto 15% responderam que sim.

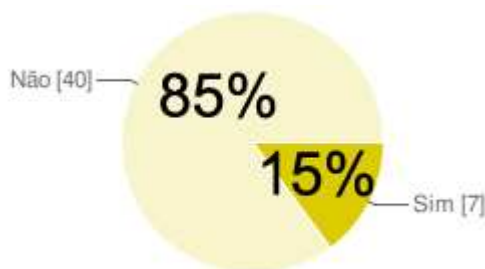


Figura 27: Gráfico da estabilidade do site

Desses 15% de usuários muitos persistiram nos comentários anteriores, relacionados a opção de expansão de consultas, relacionados a não atualização dos resultados quando é realizada uma nova consulta; um dos usuários citou a pesquisa por termos que não são totalmente relacionados, como “futebol”, “dunga” e “fermat”, que não retornaram nenhuma opção para a expansão; e relacionado ao tempo de resposta, foi citado por poucos a demora para retornar os resultados, este comportamento foi notado no navegador *Internet Explorer*²⁰.

Outro item não muito comentado pelos usuários foram os resultados da expansão da consulta. Foi citado o resultado por imagens, onde o usuário realizou diversas consultas sem encontrar nenhuma imagem, no entanto se fosse pesquisado diretamente no site do *Flickr* eram encontradas várias imagens, esta falha foi detalhada a sessão 4.3. Para o resultado de bookmarks, um dos usuários comentou que os endereços ultrapassavam o limite definido para a disposição dos resultados na tela, esse problema foi observado nos demais resultados. Também foram encontrados resultados não condizentes com a consulta realizada pelo usuário. Pode-se mencionar que a consulta por termos muito específicos, podem não retornar nenhum resultado relacionado, pois a prática de sugestão de *tags* que é utilizada pelo *Delicious* elimina a criação de tais termos pelos usuários.

A décima quarta questão, estava relacionada a navegabilidade do sistema, para avaliar se o usuário conseguiu navegar naturalmente desde o primeiro contato com o sistema. Em um total de 89% de opiniões, os usuários não encontraram nenhuma situação que comprometeu a navegação no *Folkaliza*, por outro lado 11% encontraram dificuldades para realizar consultas no sistema (figura 28).

²⁰ <http://www.microsoft.com/brasil/windows/Internet-explorer/default.aspx>

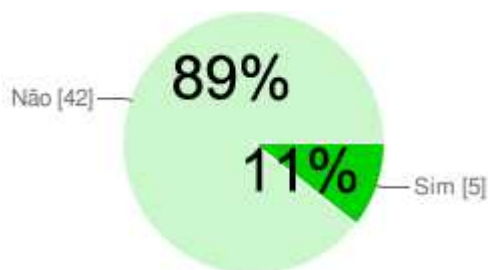


Figura 28: Gráfico de situações críticas na utilização do Folkaliza

Os comentários citados com relação a décima quarta questão, foram os mesmos citados anteriormente, relacionados a expansão de consultas, nas questões que se referem a aspectos de usabilidade.

A décima sexta questão procurou identificar se o usuário achou que o *Folkaliza* disponibiliza alguma informação ou funcionalidade que não é necessária. Um total de 87% opinou que o sistema não possui nenhuma informação ou funcionalidade desnecessária, no entanto 13% dos usuários responderam que sim (figura 29).

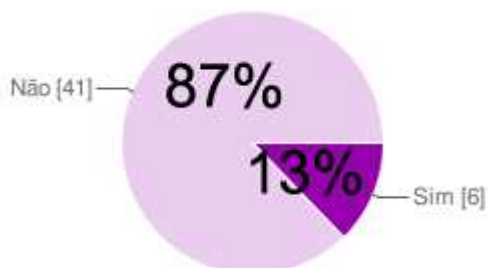


Figura 29: Gráfico da funcionalidade do site

Para essa questão foram feitas considerações sobre os resultados de imagens, onde acharam desnecessário a aparição de todas as *tags* relacionadas a imagem. Também citaram que o menu de navegação era desnecessário, já que na página principal é possível obter os primeiros resultados de cada mídia para a consulta, possibilitando que os usuários pudessem visualizar mais resultados diretamente na primeira página. Persistiu o comentário relacionado a expansão de consultas, citaram que a escolha de uma opção poderia ser opcional deixando o resultado com a consulta inicial do usuário. Também foi levantada a possibilidade de não aparecer nenhum resultado, caso a consulta do usuário não fosse relevante.

A décima oitava questão tinha por objetivo identificar se as mensagens, tanto de ajuda quanto informação do status da consulta, foram úteis e claras aos usuários que utilizaram o *Folkaliza*. O usuário deveria escolher apenas uma resposta. Um total de 91% considerou que as mensagens serviram de ajuda e situaram os usuários durante a consulta, enquanto 9% discordaram da questão (figura 30).

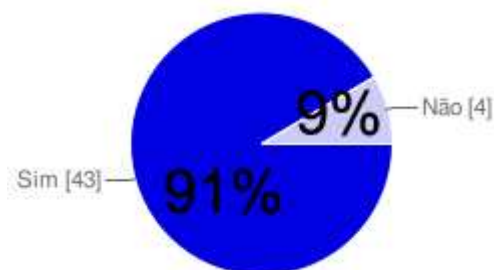


Figura 30: Gráfico da utilidade das informações contidas no *Folkaliza*

A última pergunta era do tipo aberta e os usuários poderiam comentar sobre qualquer aspecto que não foi abordado no questionário de avaliação. Foram citados alguns itens de melhoria ao site que são abordados em trabalhos futuros no capítulo 7.

Por fim foram feitos, por uma minoria de usuários, alguns comentários positivos sobre o *Folkaliza* no campo aberto do questionário. Citaram que o aspecto visual do site estava muito simples, porém objetivo e fácil de ser utilizado. Elogiaram a iniciativa em disponibilizar um sistema que faz recuperação multimídia e, ainda, possui resultados diferenciados, se comparados a outras ferramentas de busca, já que o usuário interage com o sistema de forma a refinar a sua consulta. Um dos usuários chegou a questionar se estava sendo utilizando threads para a expansão da consulta, pois na sua experiência com o *Folkaliza* os resultados eram retornados entre 2 a 5 segundos.

7 Conclusão

Este trabalho apresentou o desenvolvimento de um método de expansão de consultas para sistemas de recuperação de informação multimídia na Web que leva em consideração *tags* de sistemas de folksonomia. A criação de um protótipo mostrou-se necessária para demonstrar o objetivo proposto de maneira prática e funcional. Foi criado um website para que os usuários pudessem realizar consultas, usando o método de expansão de consultas proposto neste trabalho.

A proposta apresentada neste trabalho foi avaliada utilizando um método qualitativo de pesquisa. A partir dos resultados obtidos nesta avaliação, concluí-se que a utilização do conceito de folksonomia pode de fato trazer grandes melhoras na recuperação da informação. Os resultados encontrados indicam que 79% dos usuários consideraram que os resultados do sistema foram bons ou ótimos, e apenas 21% considerou os resultados entre razoável e ruim.

8 Trabalhos Futuros

Com base nas opiniões obtidas do questionário de avaliação, foi possível perceber que há uma série de melhorias que poderão ser considerados para trabalhos futuros, tanto relacionadas a implementação quanto usabilidade que estão detalhadas a seguir.

Uma funcionalidade muito importante que poderia ser implementada, é um método para complementar o método de expansão de consultas contida no protótipo. Isso foi percebido ao realizar consultas diretamente nos sites utilizados neste trabalho, como: *Delicious* e *Flickr*. Uma consulta realizada no protótipo com um termo popular retornava bons resultados, caso a consulta fosse realizada com um termo muito específico não eram encontrados resultados muito relevantes, diferentemente dos sites citados.

Os métodos auxiliares de *Stemming* e *StopWords*, foram pesquisados afim de melhorar a consulta do usuário. O primeiro método não foi aplicado, devido ao fato de que as opções da expansão estão em diferentes idiomas e ao testá-lo foram feitas alterações incorretas. O segundo foi implementado neste trabalho, no entanto foram percebidas algumas falhas, sendo ideal fazer melhorias nesse método

A técnica de filtragem composta não foi implementada, porém não utilizada neste trabalho, devido ao tempo de resposta muito alto. Outra melhoria poderia ser adição desta técnica, juntamente com um algoritmo que otimize a consulta para que os resultados sejam encontrados mais rapidamente. Talvez utilizar a consideração feita por um usuário, de utilizar *Threads* no método de expansão de consultas.

Um aperfeiçoamento nas opções de expansão pode ser feito, de forma que o usuário possa ter a opção de escolher um idioma para os seus resultados, fazendo com que os métodos auxiliares consigam ser inseridos, ou que ele possa descartar opções que não atendem o seu objetivo da consulta, e em tempo real seja sugerido ao usuário, novas opções de expansão para a sua consulta.

Podem ser melhorados aspectos de usabilidade do protótipo, afim de que não tenham problemas para identificar o fluxo de navegação do site. Por exemplo, a opção da consulta poderia ser sinalizada, para que o usuário não tenha que

memorizar qual item que foi escolhido. Criar um nível de agrupamento para consultas já feitas no protótipo, a fim de sugerir algumas sugestões de busca para os usuários. Limitar caracteres para os resultados da consulta, para que não haja a quebra de layout. Ocultar “módulos” que o usuário não deseja recuperar informação, por exemplo esconder os resultados de vídeos. Deixar os links do site mais intuitivos, de forma que o usuário entenda que pode navegar através deles.

REFERÊNCIAS

AQUINO, Maria Claro. Hipertexto 2.0, folksonomia e memória coletiva: Um estudo de *tags* na organização da web. **Revista da Associação Nacional dos Programas de Pós-Graduação em Comunicação**, vol. 9, p. 2 – 18, 2007.

BT, Sam Judson. **Execução de Flickr com .NET**. Disponível em: <<http://msdn.microsoft.com/pt-br/library/cc518056.aspx>>. Acesso em: 08 set. 2009.

CHAVES, Eduardo O. C. **A Definição de Multimídia**. Disponível em: <<http://www.chaves.com.br/TEXTSELF/MULTIMED/mm11.htm>>. Acesso em: 8 dez. 2009.

DAMME, Céline Van; HEPP, Martin; COENEN, Tanguy. **Quality Metrics for Tags of Broad Folksonomies**. Disponível em: <<http://www.heppnetz.de/files/VanDammeHeppCoenen.pdf>>. Acesso em: 14 dez. 2009.

GOUVÊA, Cleber; LOH, Stanley; GARCIA, Luís Fernando F. **Metódos para Seleção Automática de Tags para Descrição de Notícias na Web**. Vila Velha: WebMedia, 2008.

GOUVÊA, Cleber; LOH, Stanley. Folksonomias: Identificação de Padrões na Seleção de *Tags* para Descrever Conteúdos. **Revista Eletrônica de Sistema de Informação**, n. 2, edição 11, 2007.

LEMOS, Manoel. **Entendendo as Tags**. Disponível em: <<http://blog.blogblogs.com.br/2007/03/20/entendendo-as-tags/>>. Acesso em: 20 jun. 2009.

MARTINS, Carla. **Folksonomia: classificação cooperada**. Disponível em: <<http://www.carlamartins.com/blog/2009/02/folksonomia-classificacao-cooperada/>>. Acesso em: 21 ago. 2009.

MATHES, Adam. **Folksonomies – Cooperative Classification and Communication Through Shared Metadata**. Disponível em: <<http://www.adammathes.com/academic/computer-mediated-communication/folksonomies.html>>. Acesso em: 29 jun. 2009.

NIELSEN, Jakob. **Tem Usability Heuristics**. Disponível em: <http://www.useit.com/papers/heuristic/heuristic_list.html>. Acesso em: 13 nov. 2009.

NOLL, Michael G.; YEUNG, Ching-man Au. **How SPEAR Identifies Domain Experts within Delicious**. Disponível em: <<http://blog.delicious.com/blog/2009/08>>. Acesso em: 6 dez. 2009.

ORENGO, V.M; HUYCK, C. **A Stemming Algorithm for the Portuguese Language**. Disponível em: <<http://www.computer.org/portal/web/csdl/doi/10.1109/spire.2001.10024>>. Acesso em: 14 dez. 2009.

P., Jason Morrison. **Tagging and Searching: Search Retrieval Effectiveness os Folksonomies on the Web**. Tese (Master of Science) - B.A., Ohio Wesleyan University, 2001.

PEREIRA, Henrique Costa. **Folksonomia e a maneira com que nós colocamos ordem nas coisas**. Disponível em: <<http://revolucao.etc.br/archives/folksonomia-e-a-maneira-com-que-nos-colocamos-ordem-nas-coisas/>>. Acesso em: 13 jun. 2009.

Portuguese Stopwords. Disponível em: <<http://www.ranks.nl/stopwords/portugese.html>>. Acesso em: 15 set. 2009.

SEGARAN, Toby. **Programming Collective Intelligence – Building Smart Web 2.0 Applications**. First Edition. United States of America: O'Reilly, 2007.

SILVA, José Vanderlei. **Gerenciamento do Vocabulário do Usuário em Sistemas Baseados em Tagging**. Dissertação (Pós-Graduação em Ciência da Computação) – Universidade Estadual de Maringá, 2009.

Sinonímia, Antonímia, Homonímia e Paronímia. Disponível em: <<http://falabonito.wordpress.com/2007/05/29/sinonimia-antonimia-homonimia-e-paronimia/>>. Acesso em: 05 ago. 2009.

Stopwords for Request Manager. Disponível em: <<http://www.oclc.org/americalatina/pt/support/documentation/resourcesharing/using/searchrequests/stopwords.htm>>. Acesso em: 15 set. 2009.

TORRES, Bruno. **O básico da web – termos e definições**. Disponível em <<http://www.obasicodaweb.com/termos-e-definicoes/>>. Acesso em: 6 dez. 2009.

VAHL, José; MACHADO, Eduardo. **Tagging e Folksonomia**. Disponível em: <<http://www.aqueleblogdesoa.com.br/2008/06/tagging-e-folksonomia/>>. Acesso em: 20 jun. 2009.

WAL, Thomas Vander. **Folksonomy**. Disponível em: <<http://vanderwal.net/folksonomy.html>>. Acesso em: 16 out. 2009.

XU, Shengliang; BAO, Shenghua; FEI, Ben; SU, Zhong; YU, Yong. **Exploring Folksonomy for Personalized Search**, Disponível em: <<http://portal.acm.org/citation.cfm?id=1390363>>. Acesso em: 14 dez. 2009.

YATES, Ricardo Baeza; NETO, Berthier Ribeiro. **Modern Information Retrieval**. First Edition. United States of America: Addison Wesley, 1999.

GLOSSÁRIO

API: Uma biblioteca que fornece códigos prontos para serem utilizados.

Stopwords: Palavras que não são relevantes para se realizar busca por uma informação.

Tags: Palavras utilizadas para rotular determinada informação de modo a classificá-la de acordo com um assunto.

XML: Formato de arquivo utilizado para armazenar informação de forma organizada e classificada em níveis.

Post: Termo empregado para designar a adição de informação a um blog.

Renderizar: Ação do navegador para apresentar ao usuário o conteúdo recebido pelo servidor.

Threads: são sub-programas que realizam tarefas simultâneas para finalizar um programa principal. É utilizada para agilizar o tempo de resposta, já que não precisam esperar cada sub-programa terminar pois trabalham em paralelo.

APÊNDICE A

Questionário de Avaliação do Sistema

O questionário a seguir esteve disponível em:

<http://spreadsheets.google.com/viewform?formkey=dE03UExDS1NacEhwZHpUVXcxREZjX3c6MA>

Expansão de Consultas na Web

Bem vindo ao questionário de avaliação do sistema Web "Folkaliza".

Folkaliza é um sistema que tem por objetivo expandir a consulta do usuário, para que possa encontrar resultados mais satisfatórios do que uma consulta normal.

Você pode acessar o Folkaliza em <http://www.jessicacindra.com.br/folkaliza>

Obrigado por responder o nosso questionário de avaliação!

***Obrigatório**

1. Você utiliza ferramentas de busca disponíveis na Internet? *

Google, Yahoo, Bing, Internet Explorer, Firefox

- Sim
 Não

2. Que ferramenta de busca você mais utiliza? *

- Google
 Yahoo
 Bing
 Outro:

3. Você conhece alguns dos sites listados abaixo? Marque cada opção que você conhece. *

- Delicious
 Flickr
 Youtube
 Wikipedia

4. Você usa alguns dos sites listados abaixo? Marque cada opção que você utiliza. *

- Delicious
 Flickr
 Youtube
 Wikipedia

5. Para qual finalidade você utiliza estes sites? *

Busca em geral, entretenimento e etc.

6. Você costuma procurar quais tipos de mídia na internet? *

- Imagens
- Vídeos
- Áudios
- Documentos em geral (páginas da internet, “.doc”, “.pdf”)
- Outro:

7. Você considera útil uma ferramenta de busca que retorna mais de um tipo de mídia para a sua consulta? *

- Sim
- Não

8. O Folkaliza conseguiu retornar resultados relevantes para a sua consulta? *

Uma nota utilizando a escala (1= ruim, 2 = regular, 3 = bom, 4 = ótimo)

1	2	3	4	
Ruim	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	Ótimo

9. O Folkaliza retornou resultados inesperados para a sua consulta? *

Resultados que não seriam esperados de se encontrar em buscadores normais.

- Sim
- Não

10. Utilizando o Folkaliza, você se sentiu confuso em alguma situação? *

- Sim
- Não

11. Se a resposta anterior for afirmativa, você gostaria de detalhar em qual situação ficou confuso? Utilize o campo texto abaixo.

12. O sistema se comportou de forma inesperada em algum momento? *

- Sim
- Não

13. Se a resposta anterior for afirmativa, você gostaria de detalhar a situação que o Folkaliza se comportou de maneira inesperada? Utilize o campo texto abaixo.

14. Você encontrou alguma situação que não sabia como continuar a interagir com o sistema? *

- Sim
- Não

15. Se a resposta anterior for afirmativa, gostaria de especificar em qual situação você não conseguiu interagir no Folkaliza? Utilize o campo abaixo.

16. Há alguma informação ou funcionalidade que você achou desnecessária utilizando o Folkaliza? *

- Sim
- Não

17. Se a resposta anterior for afirmativa, gostaria de citar os itens desnecessários no Folkaliza? Utilize o campo abaixo.

18. As mensagens distribuídas no sistema são úteis e claras? *

- Sim
- Não

19. Em algum momento o sistema demorou, além do esperado, para retornar uma

mensagem e/ou resultados? *

- Sim
- Não

20. Você gostaria de comentar algum aspecto que não foi avaliado no questionário?

Caso afirmativo, utilize o campo abaixo.

APÊNDICE B

Página inicial do *Folkaliza*

Home | Ajuda | Fale Conosco

Folkaliza beta

DOCUMENTOS | IMAGENS | Vídeos | BOOKMARKS

Home

Busca

Digite palavras separadas por espaço.

EXPANDIR →

Expansão Consulta

Para expandir a sua busca, digite uma ou mais tags (termos) no campo texto ao lado e clique em expandir.

Serão disponibilizadas 5 itens de expansão, clique em um item e depois em pesquisar.

Navegação

- Imagens
- Vídeos
- Bookmarks
- Documentos
- Ajuda
- Fale Conosco

Contato

Jéssica Cintra
jessica.cintra@gmail.com

Robson Bacarin
robson.bacarin@gmail.com

Sobre o Protótipo

O **Folkaliza** tem por objetivo melhorar a busca do usuário através da expansão de consultas por tags.

O resultado esperado da expansão da consulta é trazer resultados mais precisos e inesperados para o usuário, além de contar com vários tipos de mídias em um mesmo ambiente.

Questionário

Sua opinião é muito importante, para melhorar o funcionamento do Folkaliza.

Você pode responder o nosso questionário de avaliação, que tem por objetivo avaliar funcionalidades em geral e usabilidade do sistema.

Para respondê-lo **Clique aqui**.

Contamos com a sua participação!

Imagem de uma consulta realizada no *Folkaliza*

Home | Ajuda | Fale Conosco

Folkaliza

beta

DOCUMENTOS | IMAGENS | VÍDEOS | BOOKMARKS

Home

Busca

Digite palavras separadas por espaços.

EXPANDIR →

Expansão Consulta

Você pode expandir a sua consulta através das páginas abaixo:

- ▶ multiback
- ▶ interface
- ▶ hack
- ▶ remote
- ▶ wii

▶ NÃO DESEJO UTILIZAR OS ITENS ACIMA

Documentos

- Initial Treatment Design Document
- New Albany Branch CPL Videogame List
- Design Text Bed

- Wii Operations Manual
- Wii Operations Manual
- Wii Console By Robert Wesley Thibault

- Wii Remote - Wikipedia, the free encyclopedia
- Wii at Nintendo
- Wii.com

[ver mais documentos](#)

Imagens

- barcamp03, wiiremote
- 3d3d3dchallenge, nintendo, haxli, nintendo, wiiremote, wii, score, nintendo, wiiremote, nintendocontroller
- nintendo, wiiremote, nintendocontroller
- nintendo, wiiremote, wii, nintendo, wiiremote, wii, nintendo, wiiremote, wii, nintendo, wiiremote
- wiiremote, nintendo, nintendo, wiiremote

[ver mais imagens](#)

Vídeos

- Head Tracking for Desktop VR Displays using the WiiRemote
- Black Nintendo and nintendo Unboxing
- Wii Remote Reverse
- Johnny Lee: Wii Remote Hacks
- Will It Blend? - Wii Remote

[ver mais vídeos](#)

Bookmarks

- Johnny Chung Lee - Projects - Wii
<http://www.cs.cmu.edu/~johnny/projects/>
- Wii.org Wii Linux - Main Page
http://www.wii.org/index.php/Wii_Page
- GamecubeRemote is released! (Hirelog)
<http://blog.hirelog.jp/2004/12/20041215.html>
- Nintendo - Customer Service | Wii Remote Jacket Accessory Offer
<http://www.nintendo.com/customer-service/jacket-accessory-offer>
- Johnny Chung Lee - Human Computer Interaction Research
<http://www.cs.cmu.edu/~johnny/>
- WiiRemote
<http://enakuruta.org/wiiremote.html>
- Johnny Chung Lee - Projects - Wii
<http://johnnylee.net/projects/wii/>

[ver mais bookmarks](#)

Navegação

- Imagens
- Vídeos
- Bookmarks
- Documentos
- Ajuda
- Fale Conosco

Contato

Jéssica Cidre
jessica.cidre@gmail.com

Rafael Barreira
rafael.barreira@gmail.com

Sobre o Protótipo

O *Folkaliza* tem por objetivo melhorar a busca de usuários através da expansão de consultas por tags.

O resultado esperado da expansão da consulta é trazer resultados mais precisos e relevantes para o usuário, além de contar com vários tipos de mídia em um mesmo ambiente.

Questionário

Sua opinião é muito importante, para melhorar o funcionamento do *Folkaliza*.

Você pode responder o nosso questionário de avaliação, que tem por objetivo avaliar funcionalidades em geral e usabilidade do sistema.

Para respondê-lo: [Clique aqui!](#)

Obrigado com a sua participação!