



UNIVERSIDADE ESTADUAL DO NORTE DO PARANÁ
FACULDADES LUIZ MENEGHEL

PRISCILA DA SILVA VELASCO DE OLIVEIRA

**NAVEGAÇÃO DE FACETAS POR METADADOS
COMO APOIO À BUSCA NA WEB**

Bandeirantes
2009

PRISCILA DA SILVA VELASCO DE OLIVEIRA

**NAVEGAÇÃO DE FACETAS POR METADADOS
COMO APOIO À BUSCA NA WEB**

Trabalho de Conclusão de Curso submetido às Faculdades Luiz Menghel da Universidade Estadual do Norte do Paraná, como requisito parcial para a obtenção do grau de Bacharel em Sistemas de Informação.

Orientador: Prof. Jose Reinaldo Merlin

Bandeirantes

2009

PRISCILA DA SILVA VELASCO DE OLIVEIRA

**NAVEGAÇÃO DE FACETAS POR METADADOS
COMO APOIO À BUSCA NA WEB**

Trabalho de Conclusão de Curso submetido às Faculdades Luiz Meneghel da Universidade Estadual do Norte do Paraná, como requisito parcial para a obtenção do grau de Bacharel em Sistemas de Informação.

COMISSÃO EXAMINADORA

Prof. Jose Reinaldo Merlin Orientador
UENP – *Campus* Luiz Meneghel

Prof.^a Me. Daniela de Freitas G. Trindade
UENP – *Campus* Luiz Meneghel

Prof.^a Me. Cristiane Yanase H. de Castro
UENP – *Campus* Luiz Meneghel

Bandeirantes, ___ de _____ de 2009.

AGRADECIMENTOS

Aos professores desta Instituição que com competência e dedicação ampliaram meus horizontes de conhecimento.

Deixo aqui também um agradecimento todo especial a professora Christiane Enéas da Silva, que com muita competência, dedicação e amizade participou deste momento em minha vida como minha co-orientadora neste projeto, me dando apoio nos momentos difíceis, não me deixando desistir do meu objetivo, Chris meu muito obrigado a você.

RESUMO

A pesquisa por informação é uma atividade fundamental na vida das pessoas. As tecnologias da *Web* fornecem a infra-estrutura necessária para o desenvolvimento de novas abordagens para auxiliar as pessoas a encontrar a informação desejada. A investigação de modelos de navegação para a *Web* pode fornecer informações para melhorar as interfaces e os mecanismos de sistemas de navegação. As pesquisas envolvidas neste trabalho foram analisar as implementações de buscas e navegação para a visualização de informações e imagens, e a verificação da importância do design como mediador entre usuário e informação, no sentido de viabilizar a aquisição de conhecimento e o seu compartilhamento.

Palavras Chaves: Visualização de Informações. Design da Informação. Design de Interface. Sistemas de Busca. Recuperação de Informações.

ABSTRACT

The search for information is a fundamental activity in people's lives. Web technologies provide the infrastructure necessary for the development of new approaches to help people find the desired information. The investigation of models of navigation for the Web can provide information to improve the interfaces and mechanisms for navigation systems. The research involved in this work were to analyze the implementations of search and navigation for viewing information and pictures, checking the importance of design as a mediator between User and information in order to facilitate the acquisition of knowledge and its sharing.

Keywords: Information View. Information Design. Interface Design. Search Engines. Information Retrieval.

FIGURAS

Figura 1	Crescimento do número total de <i>sites</i> disponíveis na Internet	pág.12
Figura 2	Exemplo de Buscas em sites na <i>Web</i>	pág.18
Figura 3	Site <i>Clusty.com</i>	pág.19
Figura 4	Navegação por Facetas hierárquicas no Flamenco	pág.25
Figure 5	Navegação por Facetas hierárquicas no Flamenco e subcategorias	pág.27
Figure 6	Visualização do site eBay Express, com as informações apresentadas em facetas	pág.27
Figure 7	Seleção da faceta Accessories no site eBay Express	pág.28
Figure 8	- Refinamento da consulta no site eBay Express	pág.28
Figure 9	- Opções disponíveis para refinamento da consulta no eBay Express	pág.29
Figura 10	Caminho de breadcrumbs feito por facetas no site www.tokstok.com.br	pág.32
Figura 11	Visualização dos resultados de busca mostrados pelo Grokker.	pág.33

LISTA DE SIGLAS, ACRÔNIMOS E ABREVIATURAS

RI	Recuperação de Informação
Web	World Wide Web

SUMÁRIO

1. INTRODUÇÃO	9
1.1 Delimitação do Tema	10
1.2 Objetivo	10
1.3 Motivação.....	10
1.4 Organização do Trabalho.....	11
2. SOBRE O SISTEMA DE BUSCA.....	12
2.1 Recuperação da Informação	13
3. COMPORTAMENTO DE INTERAÇÃO DOS USUÁRIOS COM SISTEMAS DE BUSCA NA <i>WEB</i>	15
4. A IMPORTÂNCIA DO CONTEXTO NOS RESULTADOS DE BUSCA NA <i>WEB</i> ...	17
5. <i>CLUSTERING</i> X CATEGORIAS DE FACETAS PARA EXPLORAÇÃO DA INFORMAÇÃO	18
6. NAVEGAÇÃO POR FACETAS DE METADADOS.....	21
6.1 Terminologia.....	22
6.2. Navegação por Hierarquia de Facetas	24
6.3 Layout de Rótulos por Facetas.....	29
6.4 Exposição de Facetas	30
6.5 Incorporação da Pesquisa pela Palavra Chave.....	31
6.6 <i>Design Breadcrumbs</i>	31
6.7 Discussão.....	32
7. GROKKER	33
8. BUSCA E NAVEGAÇÃO DE FACETAS DE METADADOS DE IMAGENS.....	34
9. CONSIDERAÇÕES FINAIS	35
REFERÊNCIAS.....	36

1. INTRODUÇÃO

Com a internet presente, cada vez mais, na vida das pessoas, atividades cotidianas são planejadas e decisões são tomadas, muitas vezes baseadas em consultas à web. Para realizar essas consultas as pessoas utilizam os sistemas de busca. São eles que permitem que os usuários encontrem as informações que procuram em meio a uma infinidade de informações num *site*. No entanto, com o crescimento exponencial de informações disponibilizadas na rede, as consultas apresentam cada vez mais resultados. Os resultados aparecem listados de forma textual e sequencialmente em inúmeras páginas, exigindo dos usuários um esforço extra no sentido de refinar e localizar o que procuram.

A área de buscas na *web* tem sido objeto de pesquisa sob vários enfoques, com contribuições de diversos campos do saber. As pesquisas têm-se dado principalmente no âmbito da Interação Humano-Computador, Ciência da Computação, Recuperação de Informações, Mineração de Dados e Design de Interfaces.

O domínio da aplicação explorado neste trabalho é um dos modelos de buscas exploratórias em *sites* comerciais, que visam apresentar uma nova forma que como a pesquisa de dados dentro de um site *e-commerce* pode ser entendida e manipulada (SILVA, 2009).

O usuário pode selecionar um conjunto de facetas, baseado nos metadados das páginas, e como resultado ele recebe um conjunto de *links* que apontam para páginas que satisfazem à seleção das facetas escolhidas. Esta é uma maneira de atribuir semântica aos *links*. No contexto da *Web*, a classificação em facetas surge como um novo paradigma de navegação. Uma das vantagens da navegação baseada em facetas é a possibilidade de se combinar um ou mais valores (SILVA, 2009). Na tentativa, de uma navegação mais eficiente na *Web* este trabalho propõe mecanismos para auxiliar o usuário na navegação em repositórios de informação baseados na *Web*. Nesse contexto, observa-se a possibilidade da utilização de metadados para facilitar a descrição e descoberta da informação.

1.1 DELIMITAÇÃO DO TEMA

O presente trabalho é uma pesquisa bibliográfica sobre as técnicas de buscas e navegação por informações na *Web*, mostrando seus conceitos e definições, aprofundando-se na técnica por buscas por facetas em *sites* comerciais.

1.2 OBJETIVO

O objetivo é investigar mecanismos para antecipar os passos que poderiam ser tomados para acelerar a busca por facetas em *sites* comerciais, para que estas buscas fiquem mais rápidas e eficientes para auxiliar a navegação em interfaces *web*.

O conteúdo deste trabalho, através de uma revisão bibliográfica, detalha técnicas de buscas de informações que visam contribuir na melhoria do suporte para acesso das pessoas à informação considerada relevante.

1.3 MOTIVAÇÃO

De forma resumida, a motivação para a realização da pesquisa aqui relatada surgiu a partir da soma das seguintes constatações:

- Os sistemas de busca têm importância crescente na localização de informações na *web*;
- Os usuários de sistemas de busca utilizam apenas duas ou três palavras-chave por consulta, e na maioria das vezes não olham além da primeira página de resultados;
- Os sistemas de busca não atendem bem nos casos de busca exploratória e não exploratória sendo que estas correspondem a até um terço de todas as consultas realizadas na *web*.

Existe uma nova geração de sistemas de busca que utilizam técnicas de visualização para mostrar seus resultados, porém de forma ainda bastante embrionária.

Ainda quanto à questão da abundância de informações na web, Meirelles e Moura (2007) ressaltam a posição de Bonsiepe¹ quanto à importância do design como mediador entre usuário e informação, no sentido de viabilizar a aquisição de conhecimento e o seu compartilhamento:

[...]conhecimento como experiência acumulada precisa ser comunicado e compartilhado entre indivíduos. O processo de comunicar e compartilhar está ligado a apresentação do conhecimento e essa é, ou pode vir a ser, um assunto do Design.”
(Bonsiepe, 2000 apud Meirelles e Moura, 2007)

Encurtar o caminho entre o buscar e o achar é o grande desafio dos sistemas de busca e diversas áreas de pesquisa vêm convergindo esforços no sentido de pensar novas formas para que a tarefa de localizar uma informação na web seja mais bem sucedida.

1.4 ORGANIZAÇÃO DO TRABALHO

Este trabalho está relacionado em 9 seções, a seção 1 apresenta a síntese do quadro teórico para a elaboração deste trabalho, a seção 2 apresenta os aspectos sobre os sistemas de busca, a seção 3 apresenta a síntese dos estudos sobre comportamento de interação dos usuários com sistemas de busca na Web, a seção 4 trata da importância do contexto nos resultados de busca na Web, a seção 5 apresenta uma comparação entre clustering e categorias de facetas para exploração da informação, a seção 6 apresenta as recomendações, terminologia, navegação por hierarquia, layout de rótulos, exposição, incorporação da pesquisa, design e discussões para navegação por facetas de metadados, a seção 7 apresenta o Grokker uma outra técnica de buscas na Web, a seção 8 apresenta outra técnica de busca e navegação por facetas de metadados de imagens e finalmente a seção 9 apresenta as considerações finais deste trabalho.

¹ **Gui Bonsiepe** é designer graduado pela Hochschule für Gestaltung de Ulm, Alemanha. Trabalha no campo do design de informação para o ensino e gestão de conhecimento em empresas públicas e privadas.

2. SOBRE O SISTEMA DE BUSCA

“Os gregos faziam suas perguntas ao oráculo de Delfos no famoso templo de Apolo. Esse era um local de troca de idéias e informação. Os oráculos da era da informação são os sistemas de busca na *web* onde milhões de usuários submetem consultas diariamente as suas bases de dados” (LUNARDI, 2008).

As tecnologias de informação e de comunicação abrem novas perspectivas à sociedade do futuro. A Internet possibilita hoje uma difusão rápida da informação que, uma vez produzida, circula instantaneamente, e pode ser acessada por toda a sociedade. Segundo a pesquisa de julho de 2007 do Internet Systems Consortium, Inc. já existem 489 milhões de *sites* disponíveis na Internet.

Internet Systems Consortium, Inc. 2007

Crescimento da Internet

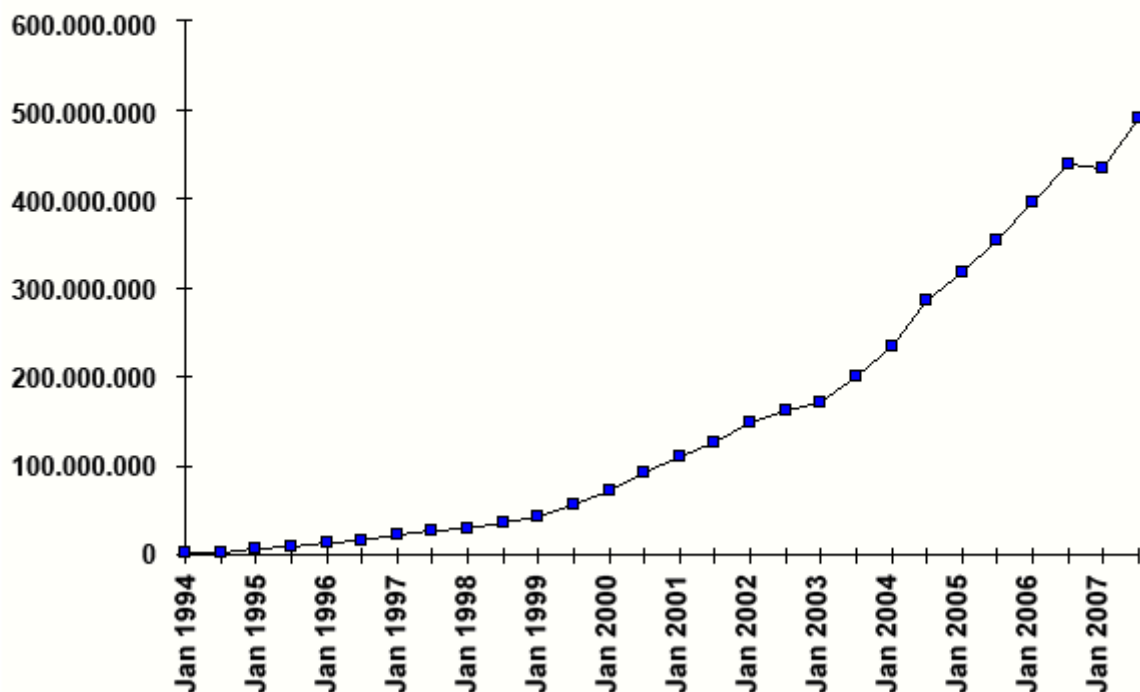


Figura 1 - Crescimento do número total de *sites* disponíveis na Internet.

Para se localizar algo em meio a toda essa informação disponível, é imprescindível um mecanismo que permita a pesquisa nesse universo. Esse é o objetivo dos sistemas de busca que surgiram logo após a criação da Internet. Os sistemas de busca se propõem a permitir a pesquisa e a localização de dados

específicos em meio a toda essa informação crescente, apresentando os resultados de forma organizada, rápida e eficiente.

2.1 RECUPERAÇÃO DA INFORMAÇÃO

Os sistemas de busca nada mais são do que sistemas de recuperação de informação.

Recuperação da Informação (RI) é uma subárea da ciência da computação que trata da representação, armazenamento, organização e do acesso aos itens de informação (BAEZA-YATES e RIBEIRO NETO, 1999). Um item de informação é geralmente constituído de texto como documentos, páginas na web, livros, entre outros, podendo conter outros tipos de dados, como fotografias, gráficos e figuras. Segundo Macedo (2001), o principal objetivo de um sistema de RI, é a seleção, num universo de documentos disponíveis, do conjunto de documentos relevantes para uma necessidade de informação do usuário.

As interfaces homem-máquina existentes atualmente, não permitem que um sistema de RI obtenha as informações diretamente da mente do usuário. Portanto, o usuário precisa traduzir a sua necessidade de informação utilizando uma linguagem formal específica de um determinado sistema, o que representa uma das grandes dificuldades para o usuário.

A necessidade de informação do usuário deve ser traduzida para uma consulta que possa ser processada por um sistema de RI. A consulta deverá ser formulada através de uma palavra-chave ou um conjunto de palavras-chave. O tipo de formalização exigida para a consulta, depende também do tipo de sistema de RI que estiver sendo utilizado pelo usuário.

A dificuldade na formulação da necessidade de informação pelo usuário, ocorre também em grande parte por se tratar de uma “necessidade visceral” (MACEDO, 2001), ou seja, o usuário está consciente que precisa da informação, mas não consegue nem sequer a sua definição em linguagem natural. Portanto, transpô-la, para a linguagem suportada pelo sistema de RI é muito mais difícil. Conseqüentemente, o usuário pode vir a formular uma consulta inadequada e a probabilidade do sistema retornar documentos não relevantes para a necessidade do usuário aumenta.

Através da consulta formulada pelo usuário, o sistema de recuperação de informação é capaz de selecionar as informações (documentos) relevantes para a necessidade do usuário. A forma utilizada pelo sistema para selecionar a informação relevante é identificar a similaridade entre as informações armazenadas no sistema com a necessidade de informação descrita na expressão da consulta. Esta comparação pode ser problemática, porque um documento pode ser relevante à consulta do usuário, mas não ser relevante para o usuário (que pode ter formulado incorretamente a sua necessidade de informação). Após determinar quais os documentos de uma coleção são relevantes à consulta do usuário, os sistemas de RI retornam o resultado da consulta em uma lista, onde os documentos estão ordenados de acordo com uma hierarquia.

Muitos sistemas de RI são fechados, ou seja, funcionam em ambientes que permitem o controle do que é introduzido no sistema.

Na revisão das bibliografias, têm-se como foco deste trabalho as buscas pré-determinadas em um *site*, ou seja, sistemas de RI que são fechados, em que o usuário poderá escolher opções para consultar, podendo refinar as informações.

Existe uma série de termos utilizados para denominar os sistemas de recuperação de informação na *web*. Em português, o termo mais comum é sistema de busca que provém do inglês *search engine*. Também são utilizados os termos mecanismo de busca, ferramenta de busca, ou ainda, motor de busca como sinônimos (MOURA, 2000). Muitas vezes o termo sistema de busca é usado de forma genérica e engloba também facetadas de metadados, *clustering*, busca de imagens, dentre outras. Estes sistemas, embora tenham modos de funcionamento diferentes, tratam do mesmo problema, que é a recuperação, em um universo de documentos, do conjunto de documentos relevantes para uma necessidade da informação do usuário (MOURA, 2001). A seguir será feita uma breve descrição de cada um deles.

Nesse trabalho, o termo sistema de busca é utilizado para definir os sistemas cuja indexação, técnicas de buscas, refinamento de informações é feito exclusivamente através de programas de computadores.

Esses sistemas também são conhecidos como mecanismos de busca. Existe uma base de dados onde armazenam referências da informação indexada e uma interface que permite ao usuário efetuar sua consulta e apresentar os resultados obtidos.

Os mecanismos de buscas das informações armazenados têm o objetivo de permitir que essas informações sejam localizadas de maneira rápida e eficiente, em que o usuário visualiza as possibilidades disponíveis para uma consulta e submete ao sistema. Essa consulta é feita em uma interface que é responsável pela interação do usuário com o sistema. Vale destacar que ao realizar uma pesquisa, o usuário não está pesquisando diretamente a *web* e sim em uma base de dados a partir de um *site* da *web* e que essas tecnologias com os mecanismos de buscas levantadas no trabalho deverão ser implementadas pelos desenvolvedores da interface, e sendo assim, o programa de computador desenvolvido com estes mecanismos podem ser disponibilizados através da construção de um componente reutilizável.

3. COMPORTAMENTO DE INTERAÇÃO DOS USUÁRIOS COM SISTEMAS DE BUSCA NA *WEB*

O funcionamento básico de sistemas de buscas é a recuperação das informações que preenchem os requisitos, que são disponibilizados para usuário por ocasião da consulta, através de facetas hierárquicas. Existem outras etapas para a recuperação de informações que vai da análise até a indexação das páginas, e que não serão abordadas profundamente neste trabalho.

As principais constatações desses estudos sobre o comportamento de busca dos usuários foram que as consultas tendem a ser curtas com em média duas ou três palavras. Os usuários raramente utilizam as ferramentas de busca avançada ou operadores lógicos para aumentar a precisão das consultas, e ainda, os usuários de sistemas de busca geralmente não olham além da primeira página de resultados. Cerca de 20 a 30% de todas as consultas realizadas na *web* são exploratórias por natureza.

Quando os usuários iniciam uma busca a falta de uma visão geral dos resultados é particularmente problemática e outro problema evidenciado é a ausência de estudos cujo enfoque principal está no design gráfico da informação (LUNARDI, 2008).

Em resumo, as principais constatações do estudo sobre o comportamento de busca na *web* foram:

1. As consultas tendem a ser curtas e logo são potencialmente vagas e ambíguas; a maioria dos estudos reporta consultas com em média duas ou três palavras;

2. Os usuários raramente utilizam as ferramentas de busca avançadas, como por exemplo, os operadores booleanos para aumentar a precisão das consultas. Essas facilidades são utilizadas em menos de 20% das consultas, e em estudos mais recentes, em menos de 10%;

3. Os usuários de sistemas de busca geralmente não olham além da primeira página de resultados de uma consulta; e apesar de esses estudos não serem muito recentes, eles avaliaram um longo período, e o comportamento dos usuários se manteve constante ao longo do tempo.

Estudos identificaram que os usuários geralmente desenvolvem estratégias para lidar e compensar essas situações onde as necessidades de informação são vagas. Por exemplo, os usuários submetem uma consulta tentativa e passam a navegar a partir dos resultados retornados contando apenas com sua habilidade em interpretar pistas contextuais e navegar entre documentos. Resumindo, exploram a informação disponível procurando seletivamente e obtendo passivamente pistas sobre qual o próximo passo a ser seguido. O objetivo da exploração de informações é o refinamento de uma necessidade vaga de informação que através da interação com diversas fontes de informação leva a um entendimento maior do problema.

Um estudo realizado por Teevan et al. (2004) também mostrou que ao invés de ir direto ao assunto através das consultas por palavras-chave, os usuários de sistemas de busca na *web* navegam até o assunto de interesse através de pequenos passos. Esses passos são localizados, e são dados apoiados no conhecimento contextualizado desses usuários, mesmo quando eles sabem exatamente o que estão procurando.

Esse comportamento foi bastante comum principalmente entre os participantes da pesquisa que tinham pouco conhecimento a respeito da estrutura da informação consultada.

Algumas vantagens foram observadas nesse comportamento de busca realizado através de pequenos passos. Dessa forma os usuários não precisam especificar muito, o que estão procurando, e também conseguem observar os

resultados em um contexto que permite que ele seja compreendido de uma maneira mais clara.

Segundo Marchioni (2006), essa estratégia envolve várias consultas e uma pesquisa e interativa nos resultados obtidos. Torna necessária uma escolha seletiva de um caminho a ser percorrido de forma a se obter pistas sobre os passos seguintes. Por esse motivo, esse comportamento foi denominado busca exploratória e não exploratória e compreende uma mistura de sorte, aprendizagem e investigação.

Interfaces de sistemas que visam auxiliar esse processo de buscas já vêm sendo desenvolvidas há algum tempo (HEARST, 2000) e continuam sendo alvo de discussões e estudos recentes. O que mostra que ainda existem oportunidades nas áreas de pesquisa e de desenvolvimento para melhorar as interfaces de busca atuais.

4. A IMPORTÂNCIA DO CONTEXTO NOS RESULTADOS DE BUSCA NA WEB

Os sistemas de busca são bastante eficientes na geração de listas de resultados relevantes. E quando os usuários realizam consultas por temas conhecidos, geralmente encontram o que procuram na primeira página de resultados.

A fim de proporcionar uma visão geral dos resultados em contexto, alguns estudos vêm sendo realizados investigando o uso de categorias semânticas geradas de forma automática para organizar as várias páginas de resultados retornadas pelos sistemas de busca.

Nesse estudo foram desenvolvidas e avaliadas interfaces que organizaram os resultados apresentados pelos sistemas de busca tradicionais em categorias. Os resultados continuaram sendo listados segundo a ordem de relevância do sistema. No entanto, eles foram agrupados visualmente mostrando os itens de uma mesma categoria juntos.

Essas interfaces foram comparadas com as interfaces tradicionais onde os resultados são apresentados em listas. E em todas as comparações, as interfaces organizadas por categorias, foram mais eficientes. Nelas os usuários encontraram os resultados procurados de forma mais rápida.

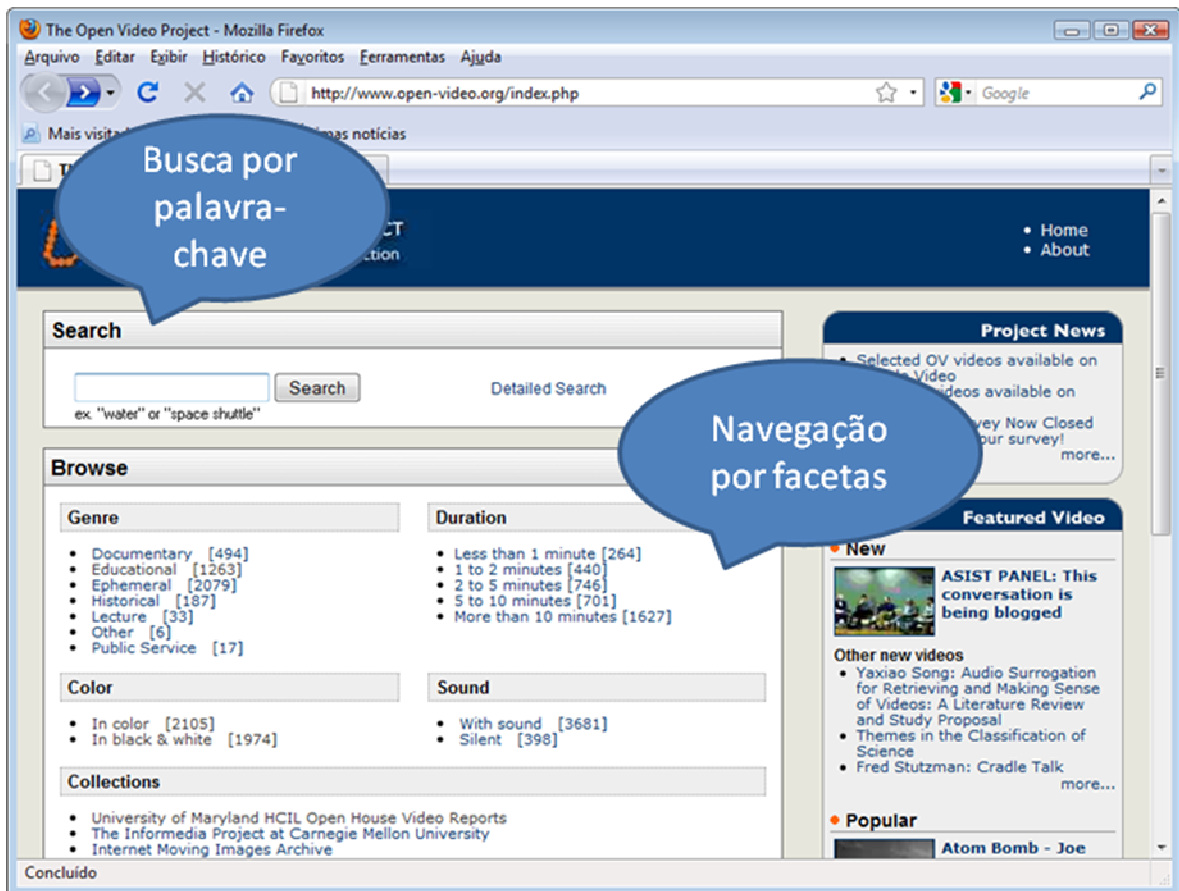


Figura 2 – Exemplo de Buscas em sites na Web

Fonte: www.open-video.org

5. CLUSTERING X CATEGORIAS DE FACETAS PARA EXPLORAÇÃO DA INFORMAÇÃO

CLUSTERING refere-se ao agrupamento de itens de acordo com algum ponto de similaridade. No agrupamento de documentos, a semelhança é tipicamente itens entre os recursos, em que as características são tipicamente palavras e frases. Uma das melhores implementações de agregação de resultados da *Web* pode ser encontrada em *Clusty.com*.

O *Clusty* é um exemplo de sistema de busca que apresenta os resultados organizados em categorias.

O sucesso dos resultados apresentados por esses estudos abriu caminho para o desenvolvimento de uma nova área de pesquisa. Recentemente surgiram sistemas de busca que apresentam formas de visualização gráfica de seus resultados, que por sua vez são organizados em categorias semânticas.

Essas formas de visualização têm sempre o objetivo de mostrar uma visão geral dos resultados das consultas e o contexto desses resultados, conforme é mostrado na **Figura 3**.

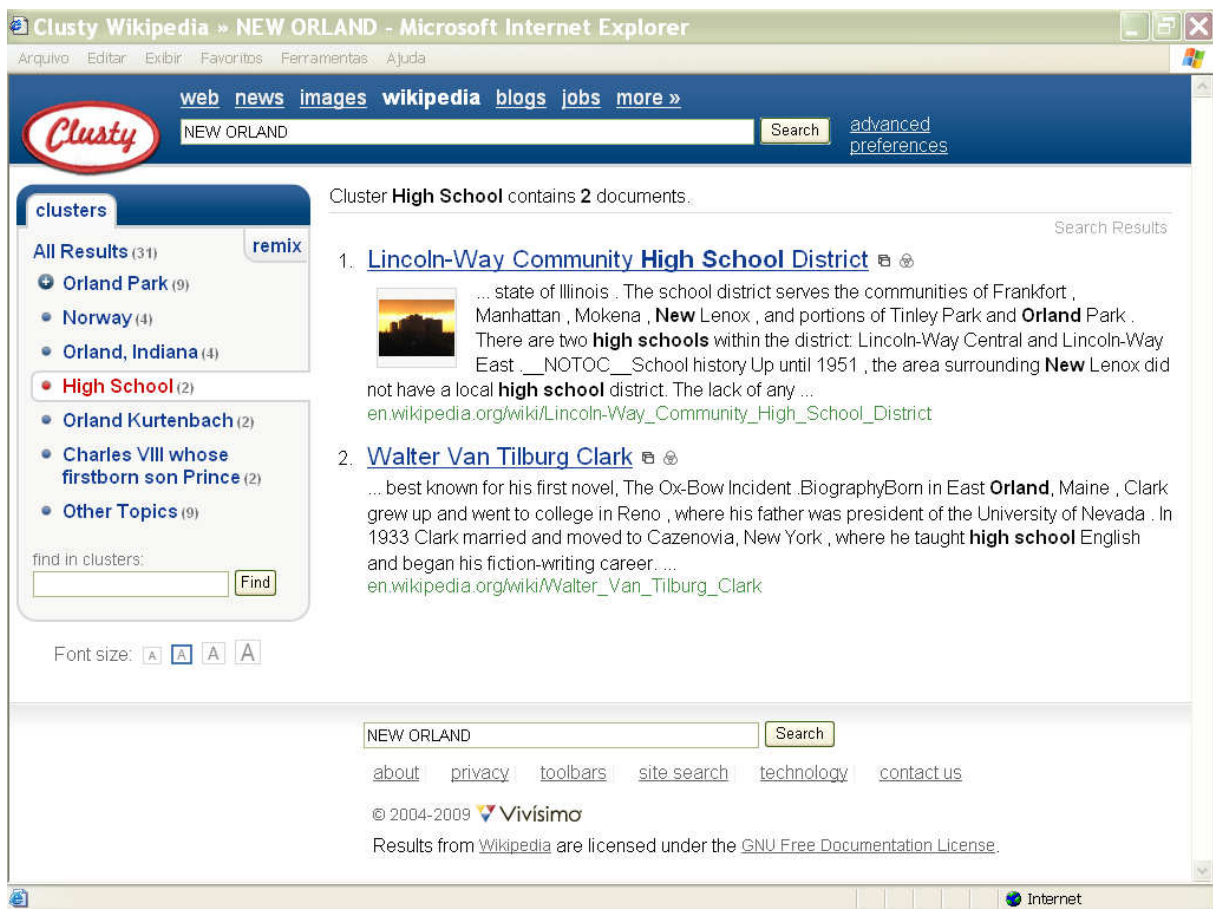


Figura 3 – Site Clusty.com

Clustering pode ser útil para clarificar e refinar uma consulta de vagas, mostrando aos usuários os temas dominantes dos resultados retornados. Clustering também funciona bem para ambíguas consultas de siglas. Por exemplo, ACL pode representar Anterior Cruciate Ligamento, *Association for Computational Linguistics*, *Atlantic Coast Line Railroad*, entre outros. Infelizmente, pelos algoritmos

de *clustering* são resultados de pesquisas imperfeitas, não resulta a informação de forma nítida, sendo assim, para agrupar todas as ocorrências de cada sigla em um *cluster*, e permitir aos usuários a emissão do acompanhamento de consultas que retornem apenas documentos do sentido pretendido (por exemplo, "reunião ACL" vai voltar reuniões para múltiplos sentidos do termo).

As desvantagens do *clustering* incluem a falta de previsibilidade, e a fusão de várias dimensões simultaneamente, incluindo a dificuldade de rotular os grupos. Algum algoritmo de criação de *clusters* gerados em torno das frases dominante apresenta uma incompatibilidade, o que torna compreensíveis para os rótulos, mas o conteúdo não corresponde necessariamente aos rótulos.

Para ilustrar estas deficiências, considere um exemplo de receita, onde a exploração de coleção de receitas são tarefas comuns de buscas. Por exemplo, a receita de "frango" na *Clusty.com* transforma-se a variedade dos seguintes grupos:

- Salada
- Peito de frango
- Churrasco / grelhado
- Receitas sopa
- Saudável
- Comida *Diet*
- Receitas fáceis de frango
- Italiano

Esta lista está incompleta e inconsistente. Na busca por Frango foi retornado churrasco / grelhado, mas não retornou cozido e frito? Peito de Frango e por que não, de pernas e asas? Por receitas italianas, mas não indiana, tailandesa, ou o francês? Além disso, aprofundando a hierarquias raramente revela resultados intuitivos. Esses tipos de erros são bastante típicos para o resultado das buscas, e o resultado da usabilidade que os usuários não gostam de agrupamentos desordenados como estes, preferindo hierarquias compreensíveis em que as categorias são apresentadas em níveis uniformes de granularidade. As facetas hierárquicas é um conjunto de rótulos significativos organizados de modo a refletir os conceitos relevantes para um domínio específico. Eles normalmente são criados manualmente, embora a atribuição de categorias de documentos possa ser

automatizada para certo grau de precisão. Sistemas que geram categorias eficientes têm as características de serem coerentes e (relativamente) completos e, portanto, representam uma vantagem sobre os resultados imprevisíveis de agregação; e segundo (HEARST, 2006) em estudos de usabilidade realizados que comparam os dois, foi constatado que os participantes preferem categorias.

A questão, a saber, é que tipo de estrutura para a criação de categorias é mais eficaz para a exploração e consulta das coleções das informações.

Há um crescente reconhecimento de que a organização estritamente hierárquica de categorias é ineficaz, uma representação alternativa, intermediária em complexidade e muito rica em termos de flexibilidade, tornou-se influente nos últimos anos. Esta representação é conhecida como categorias hierárquicas facetada. A idéia principal é bastante simples. Ao invés de criar uma hierarquia de grande categoria, construir um conjunto de hierarquias de cada categoria que corresponde a uma faceta diferente relevante para a coleção para ser navegado. No caso do frango (e outras) receitas, estas categorias hierárquicas pode incluir o tipo prato (principal, sopa, salada, sobremesa), o tipo de Ingrediente (carne, legumes, cereais, especiarias), o método de preparo (cozido, frito, grelhado, fácil), o tipo de gastronomia (italiano, indiano, francês, e assim por diante). Cada faceta tem uma hierarquia de termos associados com ele.

Após a concepção de facetas hierárquicas, a cada item da coleção pode ser atribuído muitos rótulos para a hierarquia.

6. NAVEGAÇÃO POR FACETAS DE METADADOS

Nesta seção são apresentadas as recomendações de *design* para interfaces para navegação por facetas de metadados, ou consultas por facetas hierárquicas. Segundo (SILVA, 2009) a hierarquia de facetas de metadados é considerada o melhor modelo de dados para interfaces de busca, considerada uma intermediária na complexidade entre hierarquia e representação do conhecimento. Todos os *websites*, especialmente sites para comércio eletrônico, frequentemente tem utilizado a navegação por categoria de informação; em muitos casos, suas aplicações tem sido inconsistente, incompleta, ou com outros problemas.

O domínio da aplicação explorado neste trabalho é um dos modelos de buscas em sites comerciais, esta forma de pesquisas por facetas envolvida, visa

apresentar uma nova forma que a pesquisa de dados dentro de um site *e-commerce* pode ser entendida e manipulada (SILVA, 2009).

O usuário pode selecionar um conjunto de facetas, baseado nos metadados das páginas, e como resultado ele recebe um conjunto de *links* que apontam para páginas que satisfazem à seleção das facetas escolhidas. Esta é uma maneira de atribuir semântica aos *links*. No contexto da Web, a classificação em facetas surge como um novo paradigma de navegação. Uma das vantagens da navegação baseada em facetas é a possibilidade de se combinar um ou mais valores (SILVA, 2009).

6.1 TERMINOLOGIA

A tecnologia envolvida para o desenvolvimento deste trabalho é a classificação em facetas, que busca ser um meio flexível para catalogar itens de informação. Vem sendo utilizada em diversos cenários como uma alternativa às classificações hierárquicas e enumerada, inerentemente mais rígida e de difícil manutenção. Porém, a idéia da classificação em facetas é um meio natural de organização da informação.

Facetas referem-se às categorias utilizadas para catalogar itens de informações em uma coleção. Vem sendo utilizadas em diversos cenários como uma alternativa às classificações hierárquicas e enumeradas, inerentemente mais rígidas e de difícil manutenção. Para se compreender a classificação por facetas, considere como exemplo um catálogo de filmes. A maneira mais comum de se classificar filmes é pelo gênero: filmes de ação, de suspense, de terror, comédias, etc. Porém, essa é apenas uma das possíveis opções para se classificar um filme segundo a sua faixa etária: livre, maiores de 15, maiores de 18, maiores de 21, etc. Também é possível classificá-lo pela duração: curta-metragem ou longa-metragem. Estes diferentes pontos de vista para classificação são as chamadas facetas. Cada faceta (gênero, faixa etária, duração) possui diferentes valores. Por exemplo, a faceta gênero pode possuir os valores: ação, suspense, terror, comédia, etc. Por sua vez, a faceta faixa etária pode possuir os valores: livre, maiores de 15, maiores de 18, maiores de 21, e assim por diante (PRAZERES, 2005).

A classificação em facetas surge como um novo paradigma de navegação. Normalmente, navega-se de uma página para outra seguindo-se *links*

unidirecionais. Na navegação em facetas, pode-se navegar de acordo com os valores das facetas.

Considerando outro exemplo, cada faceta tem um nome, como ingredientes para uma receita ou coleção de Gastronomia. Uma faceta pode ser plana ou hierárquica. Em ambos os casos, associada com uma faceta existe um conjunto de rótulos (PRAZERES, 2005).

Exemplos de hierarquias de facetas: Ingredientes inclui Legumes, Pães e Cereais, e assim por diante. Pode haver rótulos que são englobados pelas Hortaliças, como Cenoura, Alface. Assim, refere-se a partes da hierarquia dentro de uma faceta, como a subcategoria ou sub-hierarquia da faceta. Portanto, a receita para Massa Primavera receberia rótulos do formulário de Gastronomia dentro da subcategoria Italiano, dentro da subcategoria Ingrediente dentro da subcategoria Massa e ainda, dentro da subcategoria Ingrediente está a subcategoria Vegetais, e assim por diante.

Quando um rótulo dentro de uma faceta é selecionado dentro da interface, todos os itens que foram atribuídos a esse rótulo são recuperados, para selecionar um rótulo dentro de uma hierarquia é equivalente a um filtro, no qual todas as subcategorias abaixo deste rótulo são selecionadas.

Quando rótulos de diferentes partes da interface são selecionados, o sistema constrói dinamicamente a consulta, que é um conjunto separado dos rótulos selecionados e suas subcategorias abaixo.

Prazeres et al. (2005) propõem uma alternativa de navegação com semântica associada e utilizando o conceito de facetas. O usuário pode selecionar um conjunto de facetas, baseado nos metadados das páginas, e como resultado ele recebe um conjunto de *links* que apontam para páginas que satisfazem à seleção das facetas escolhidas. Esta é uma maneira de atribuir semântica aos *links*.

O termo “metadados” é utilizado para denotar “dados sobre dados” e seu propósito original é descrever documentos. (Aqui não distinguimos entre *dados* e *documentos*.) Há tipos diferentes de metadados descritivos: técnico (para descrever formato, padrões de codificação, etc.); administrativo (para descrever direitos de propriedade intelectual, condições de acesso, etc.); e conteúdo (temática, escopo, autoria, etc.). Essas descrições caracterizam e explicam os dados. Metadados ajudam a compreender o que são os dados e como utilizá-los. Os metadados

possuem dois componentes: Um formato e um conjunto de valores (BUCKLAND, 2006).

6.2. NAVEGAÇÃO POR HIERARQUIA DE FACETAS

Uma grande questão sobre a forma de apoio hierárquico de facetas de metadados é a forma de expor as categorias hierárquicas sem confundir o usuário. Uma abordagem é a de expor uma faceta sub-hierárquica quando o usuário passa o mouse sobre ela, usando um menu *fly-away*, ou seja, são refinados os itens da subcategoria dessa faceta, como é feito em alguns arquivos de navegação do sistema operacional. Uma vantagem desta abordagem é que permite ao usuário ver todas as opções abaixo de um rótulo, detalhando com precisão o rótulo desejado. Isso pode se transformar em uma desvantagem, entretanto, se o número de opções é muito extenso. O menu obscurece tudo no visor, requer deslocamento (o que é ruim para uma seleção no menu), e torna a multi-seleção difícil (SILVA, 2009).

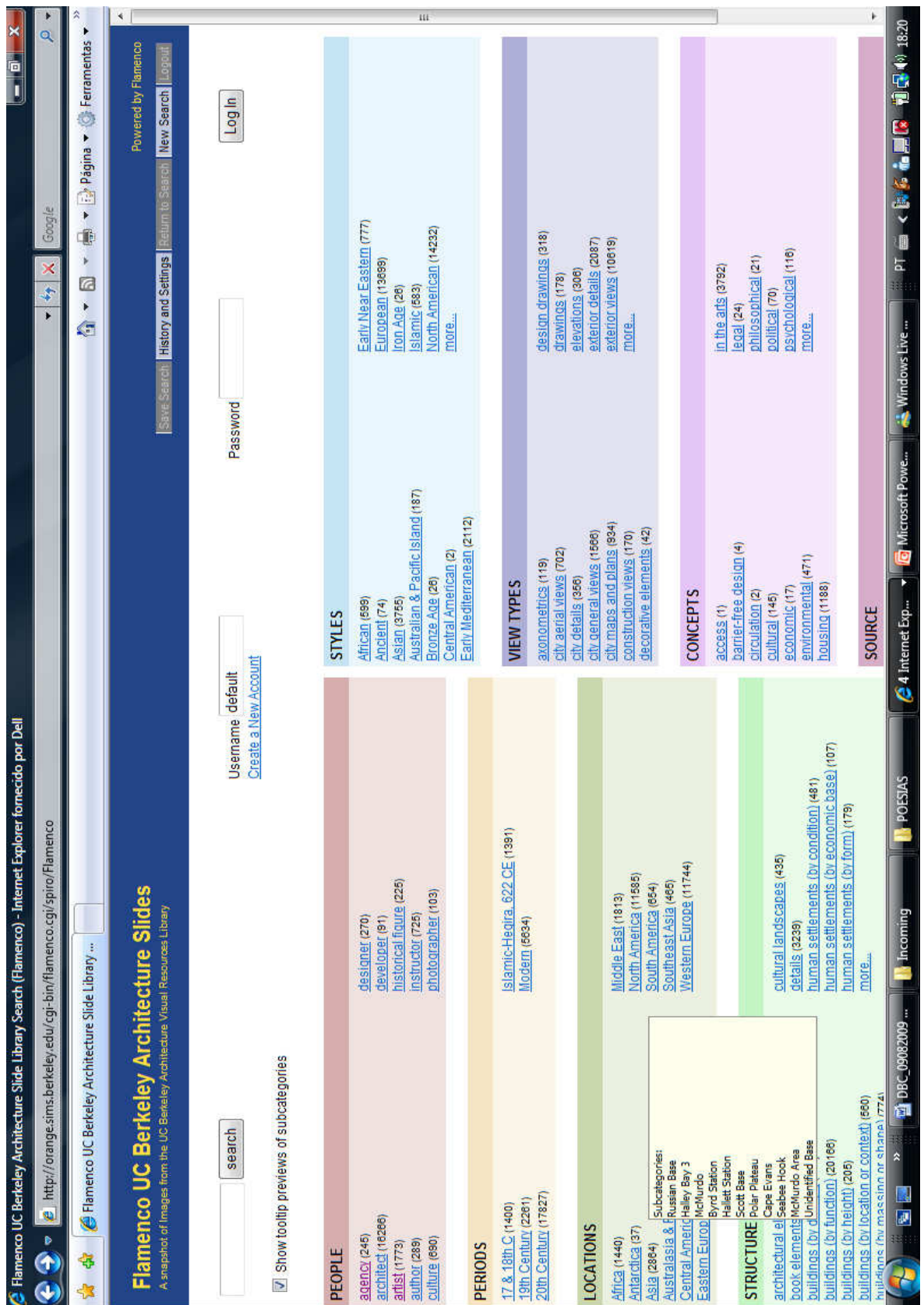


Figure 4 - Navegação por Facetas hierárquicas no Flamenco

Outra desvantagem desta abordagem é que as consultas prévias devem ser construídas para todos os níveis da hierarquia, para evitar que o usuário escolha involuntariamente um item para o qual os resultados estão vazios. Por último, essa abordagem é oposta a um dos benefícios da apresentação progressiva da hierarquia, que permite ao usuário ver os resultados agrupados por conceitos de nível superior para obter uma compreensão do conteúdo da coleção.

Outra abordagem para apresentar hierarquia é construir uma interface semelhante à do Windows Explorer, no qual as pastas expandem e fecham, e que permite que vários níveis sub-hierárquicos sejam simultaneamente visualizados da raiz ao nível selecionado. Isso tem, pelo menos, dois grandes lados. Primeiro, se muitas das hierarquias são expandidas, o componente de navegação pode ficar muito grande e exigem uma extensa rolagem. Segundo, que é ainda mais importante, os usuários serão familiarizados com a idéia de um item simultaneamente residente em várias pastas; o Explorer não tem suporte para essa funcionalidade.

No Flamenco (HEARST, 2006), adotaram um passoapasso para a abordagem dinâmica de acessar abaixo os itens das subcategorias do rótulo selecionado, sendo visível o nível superior à subcategoria. Além disso, quando o mouse é clicado sobre um rótulo, imediatamente os subitens são exibidos, de modo que o usuário pode, de fato, ver três níveis simultaneamente. Por exemplo, se o usuário seleciona o rótulo Localização e dentro desse rótulo tem a subcategoria América do Norte em seguida a subcategoria Estados Unidos, os itens que são visíveis abaixo dela seria Alabama, Alaska, etc. E dentro desses subitens se for selecionado um item, será apresentada uma lista de cidades desse estado. Assim, três níveis de informação são simultaneamente visíveis.

LOCATIONS	
Africa (1440)	Middle East (1813)
Antarctica (37)	North America (11585)
Asia (2884)	South America (654)
Australasia & F	Southeast Asia (465)
Central America	Western Europe (11744)
Eastern Europe	
Subcategories:	
Russian Base	
Halley Bay 3	
McMurdo	
Byrd Station	
Hallett Station	
Scott Base	
Polar Plateau	
Cape Evans	
Seabee Hook	
McMurdo Area	
Unidentified Base	

Figure 5 - Navegação por Facetas hierárquicas no Flamengo e subcategorias

No painel do lado direito da Figura 5, os resultados são agrupados por subcategorias, que foi selecionada a partir do rótulo LOCATIONS, na faceta Antarctica. Neste exemplo, os resultados foram inicialmente agrupados, de acordo com o estado ANTARCTICA, onde são apresentadas 37 subcategorias. Estes resultados podem, por sua vez, serem agrupados por subcategorias de uma faceta diferente, como PEOPLE, conforme mostrado na **Figura 4**.

Uma abordagem diferente à exposição da hierarquia é demonstrado na interface do *eBay Express*. Nesse sistema, se um rótulo dentro de uma faceta hierárquica é escolhido, o passo seguinte é mostrar a faceta separada.

The screenshot shows the eBay Express homepage with a search bar and navigation links. Below the search bar, there are several category facets arranged in a grid. The facets include: Antiques, Art, Baby, Books, Business & Industrial, Cameras & Photo, Cell Phones & PDAs, Clothing, Shoes & Accessories, Coins & Paper Money, Collectibles, Computers & Networking, and Crafts. An arrow points to the 'Accessories' link under the 'Books' category.

Figure 6 – Visualização do site eBay Express, com as informações apresentadas em facetas
Fonte: www.ebay.com

The screenshot shows the eBay website interface. At the top, there's a search bar and navigation links like 'Buy', 'Sell', 'My eBay', 'Community', and 'Help'. Below the search bar, there's a 'Categories' dropdown menu with 'Accessories' selected. The main content area shows '9,011 results found' for the search 'Accessories'. A 'Refine search' sidebar on the left lists various subcategories like 'Address Books', 'Blank Diaries & Journals', 'Book Covers', etc. The main list of items includes:

Item Name	Price	Time Left
Mighty Bright XtraFlex2 Clip-On Light (Black, Kindle)	\$14.98	2d 22h 57m
5 Mixed Styles Tibetan Silver Bookmarks W3758	\$7.50	4d 13h 58m

Figure 7 – Seleção da faceta Accessories no site eBay Express

Fonte: www.ebay.com

Por exemplo, na **Figura 6**, o usuário escolheu no rótulo Books, em seguida a faceta *Accessories*. As subcategorias relacionadas ao rótulo *Accessories* são apresentadas na **Figura 7**.

The screenshot shows the eBay website interface with the search results refined to 'Address Books'. The main content area shows '254 results found'. A 'Refine search' sidebar on the left includes filters for 'Price', 'Condition', and 'Seller'. The main list of items includes:

Item Name	Price	Time Left
COACH Metallic Gold 2009 Address Book and Daily Planner	\$27.00	11h 48m
5 Address Books for Clear checkbook Covers BOOK NEW	\$4.50	2d 23h 45m
Italian Leather Pocket Journal Book 3x4 Brown Soft	\$12.99	15h 34m

Figure 8 - Refinamento da consulta no site eBay Express

Fonte: www.ebay.com

Em seguida, na seleção pelo usuário da subcategoria *Address Books*, são apresentados os itens dessa subcategoria, e como opção o refinamento da pesquisa (**Figura 8**). Uma restrição atual é que as hierarquias não devem ir mais fundo do que níveis.

The image shows two sections of a search refinement interface. The top section, titled "Refine search", has a yellow header and contains three expandable categories: "Price" with two input fields for a price range, "Condition" with radio buttons for "New", "Used", and "Not specified", and a "Choose more..." link; and "Seller" with a "Specify sellers..." link. The bottom section, titled "Preferences", has a grey header and contains one expandable category: "Buying formats" with radio buttons for "Auction", "Buy It Now", and "Include Store inventory", and a "Choose more..." link.

Figure 9 - Opções disponíveis para refinamento da consulta no site eBay
Fonte: www.ebay.com

6.3 LAYOUT DE RÓTULOS POR FACETAS

No início dos estudos do Flamenco perguntavam para os usuários sobre as suas preferências quanto ao *layout* das etiquetas dentro das facetas. Embora houvesse diferenças nas opiniões, a principal tendência parecia ser a preferência por colunas do que por lista de rótulos.

Um problema com o desenho orientado a coluna é o tamanho variável de rótulos das facetas. Se algumas facetas têm rótulos compridos e outros são muito curtos, forçando uma escolha para o número de colunas uniforme leva ao

desperdício de espaço. Na interface do Flamenco resolveram este problema por computar automaticamente o número de colunas para mostrar dentro dos itens, baseado no comprimento do maior rótulo visível (**Figura 4**).

O Flamenco também fornece um mecanismo para ajustar a ordem em que os rótulos são apresentados, quer por ordem alfabética, numérica ou por número de acessos para o rótulo visível na coleção. No entanto, nos casos em que há linhas quando não há muitos rótulos de uma faceta a serem exibidos, pode ser útil mostrar o mais saliente ou o que ocorre mais freqüentemente nas primeiras consultas e com a opção de mostrar numa ordem previsível. O sistema eBay Express utiliza essa estratégia. Na **Figura 7** dentro do rótulo *Categories*, são visualizadas as facetas e suas respectivas subcategorias ordenadas em ordem alfabética.

6.4 EXPOSIÇÃO DE FACETAS

Invariavelmente surge a questão de como muitas facetas são apresentadas e em que ordem devem ser mostradas. A atual implementação do Flamenco expõe todas as facetas, uma vez que tem sido aplicado principalmente em domínios, como bibliotecas digitais e de coleções de imagem, nos quais o conjunto de dados de exploração e aprendizagem sobre a estrutura por categoria é de alta prioridade. O Flamenco também inclui uma personalização da funcionalidade que permite aos usuários esconder as facetas que não são de interesse (SILVA, 2009).

Para algumas tarefas, tais como compras on-line, a plena exploração dos rótulos não é desejável, e a tarefa pode ser mais dirigida. Assim, na interface do *eBay Express*, os engenheiros tinham determinado antecipadamente qual subconjunto de facetas são os mais interessantes para a maioria dos usuários, e inicialmente expuseram opções adicionais sobre uma lista compacta mais abaixo. Eles também posicionaram a expansão das facetas ao lado da parte superior da tela, no *sweet spot*, ou sinal que chama a atenção (a região de uma página da Web que os usuários tendem a olhar primeiro, uma vez que normalmente contém o mais importante conteúdo da página) (SILVA, 2009).

6.5 INCORPORAÇÃO DA PESQUISA PELA PALAVRA CHAVE

Um componente chave para o sucesso para a navegação por facetas (que infelizmente nem sempre é implementado devidamente) é a integração da busca por palavra-chave, no Flamenco, a busca por palavra-chave pode ser executada a qualquer momento e pode ser incorporada à pesquisa dentro da consulta *breadcrumb*. A consulta pode mostrar categorias descrevendo cada item; o usuário pode selecionar a categoria que corresponda ou simplesmente manter a palavra-chave na consulta, combinando todos os itens que contêm o termo. Outro resultado sugere que nos demais casos eles podem organizar os resultados de uma consulta ampla pela palavra-chave e depois utilizou as facetas. Então, para obter sucesso, na composição das consultas por facetas, é necessário gerar essas consultas dinamicamente pela palavra chave que está sendo selecionada, e prever as disponibilidades dos itens referentes ao rótulo da faceta que está sendo consultado; esse processo vai acontecendo em subcategorias até a visualização dos metadados referentes às facetas selecionadas (SILVA, 2009).

6.6 DESIGN BREADCRUMBS

Breadcrumbs são recursos de navegação que mostram a trilha de acesso a uma página em relação à estrutura do *site* em que está situada. O nome em inglês é uma referência à história de João e Maria, que, para não se perder na floresta, jogam pedaços de pão para formar uma trilha que sinalize o caminho de volta. “É difícil traduzi-la, mas podemos dizer que são “trilha de orientação” ou “ barra de localização”.

O nome se justifica pois as “breadcrumbs” são muito úteis para orientar o deslocamento dos usuários entre páginas *Web* baseados nos caminhos percorridos, sem o uso dos botão “Voltar” ou da ferramenta de busca.

A navegação por facetas já mantém o caminho com cada faceta na separação do componente visual. Ambos igualam a noção da consulta consistindo na conjunção de diferentes categorias de diferentes níveis de hierarquia, e permite a expansão flexível da consulta, como mostrado na **Figura 10**.

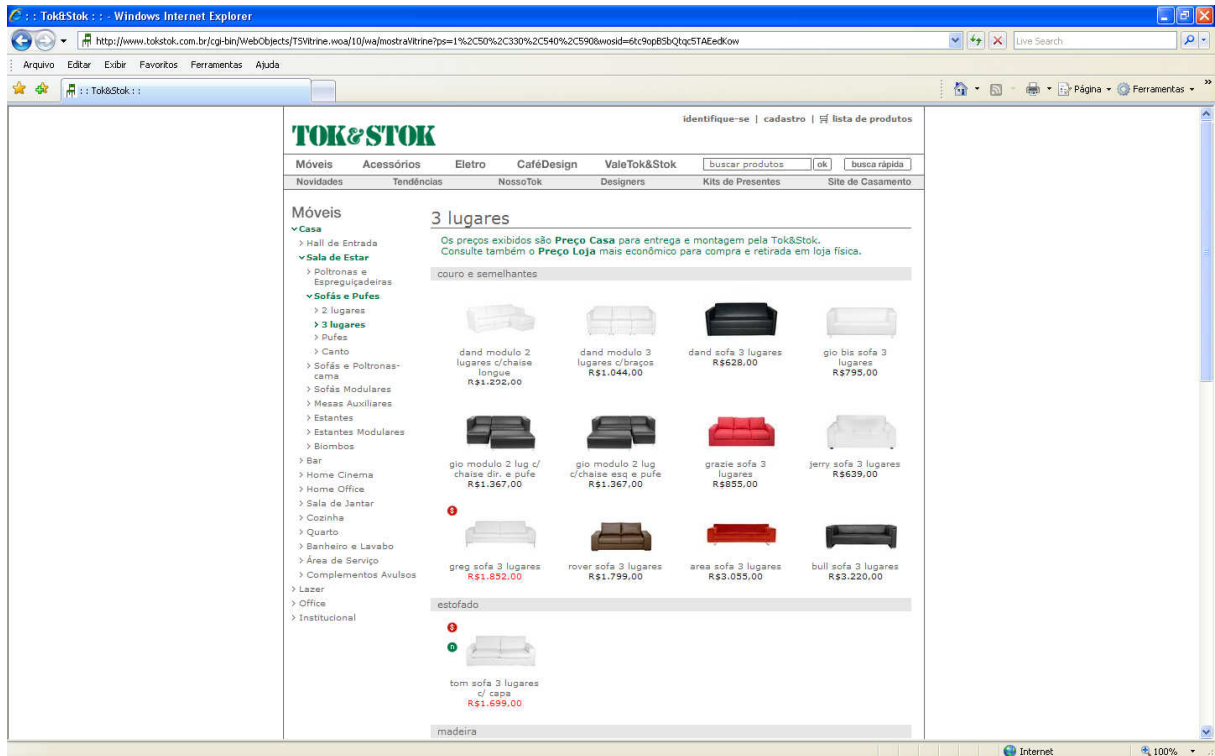


Figura 10 – Caminho de breadcrumbs feitos por facetas no site www.tokstok.com.br.

6.7 DISCUSSÃO

Outra abordagem para revelar hierarquia é construir uma interface semelhante à do Windows Explorer como as pastas que expandem e fecham, e que permitem que vários níveis sub-hierárquicos sejam simultaneamente visualizados da raiz ao nível selecionado. Isto tem, pelo menos, dois grandes lados. Primeiro, se muitas das hierarquias são expandidas, o componente de navegação pode ficar muito grande e exigem uma extensa rolagem.

Em segundo lugar, que é ainda mais importante, os usuários serão familiarizados com a idéia de um item simultaneamente residentes em várias pastas, o Explorer não suporta essa funcionalidade.

Outra desvantagem desta abordagem é que as consultas prévias devem ser construídas para todos os níveis da hierarquia, com medo de que o usuário escolha involuntariamente um item para o qual os resultados estão vazios. Por último, se opõe a um dos benefícios da progressiva amostra da hierarquia, o que

permite ao usuário ver os resultados agrupados por conceitos de nível superior, a fim de obter uma compreensão do conteúdo da coleção (HEARST, 2006).

7. GROKKER

Um desses sistemas de busca que apresenta a visualização gráfica dos seus resultados é o Grokker¹. O Grokker agrupa os resultados em uma hierarquia e apresenta as categorias em um diagrama de *Euler*². Os círculos maiores representam as categorias principais e os menores as subcategorias desses grupos. Os usuários podem explorar os resultados através de um zoom nas categorias

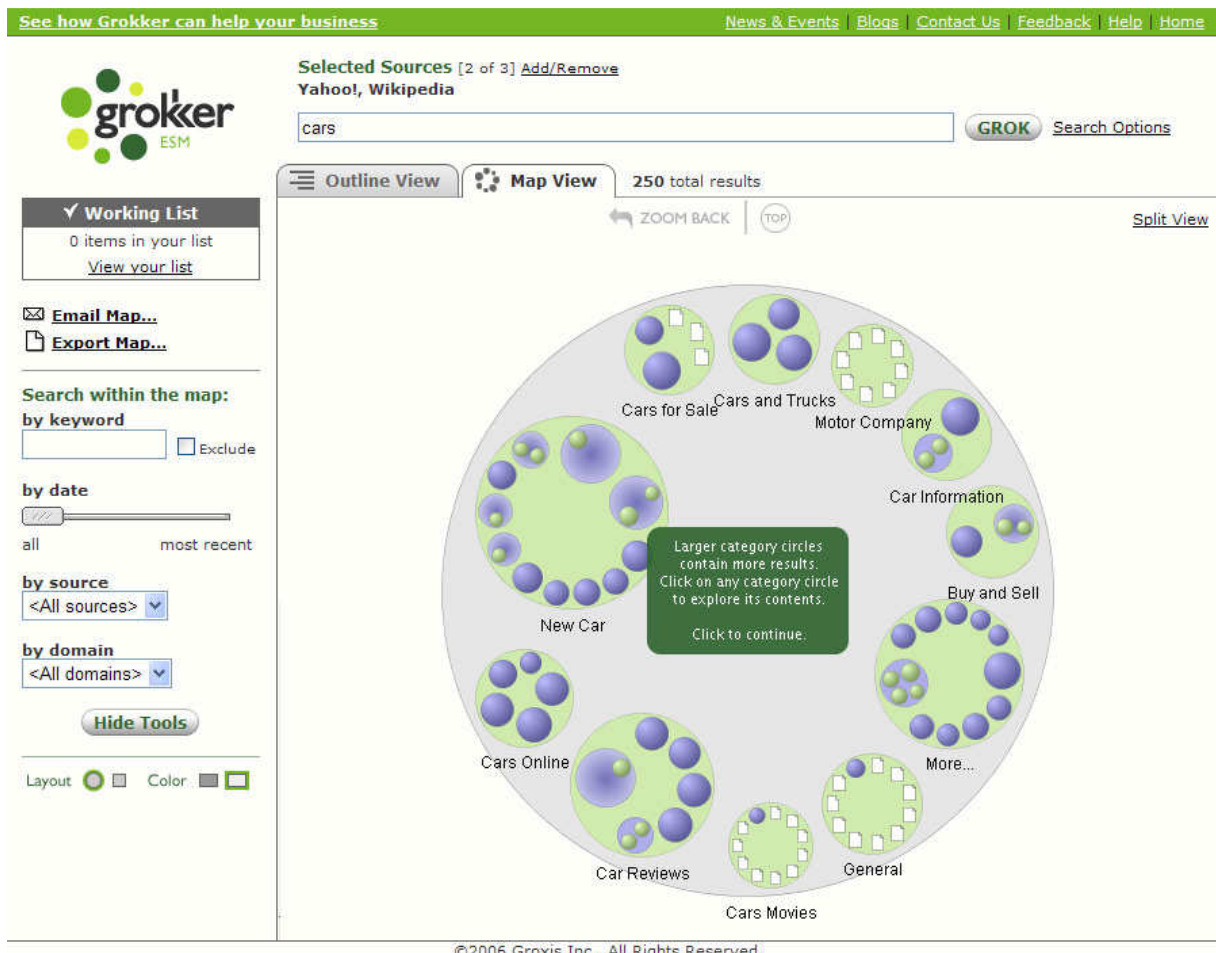


Figura 11 - Visualização dos resultados de busca mostrados pelo Grokker.

Fonte: www.grokker.com.

¹ www.grokker.com

² Um diagrama de *Euler* é similar a um *diagrama de Venn*, mas não precisa conter todas as zonas (onde uma zona é definida como a área de intersecção entre dois ou mais contornos). Assim, um diagrama de Euler pode definir um *universo de discurso*, isto é, ele pode definir um sistema no qual certas intersecções não são possíveis ou consideradas.

8. BUSCA E NAVEGAÇÃO DE FACETAS DE METADADOS DE IMAGENS

Coleções de imagens estão surgindo para buscas em páginas web, e muitos pesquisadores desenvolveram interfaces de usuário para navegar e pesquisar essas coleções. Provavelmente, a interface de busca mais conhecida de imagem atualmente é utilizada pelos sistemas de buscas por imagens na Internet, em que os usuários digitam os termos-chave, e as imagens são mostradas em uma tabela ordenada por alguma medida de relevância. Estes sistemas podem ser eficazes para a busca de itens muito específicos (HEARST, 2000).

No site do Flamenco foi desenvolvida uma interface para grandes coleções de imagens que permite aos usuários navegar com nitidez ao longo das dimensões conceituais que descrevem as imagens. A interface usa metadados de facetas hierárquicas e gera dinamicamente as visualizações das consultas, sem problemas com a categoria que integra a navegação, pesquisando palavras-chave... Em estudos realizados com 32 estudantes de história da arte que precisavam navegar por uma coleção de imagens de belas artes, encontraram-se resultados com uma maior preferência por categorias diferenciadas para a interface sobre as informações, baseado neste estudo, esta tendência por categorias diferenciadas mostra uma direção promissora as pesquisas de imagens em interfaces.

9. CONSIDERAÇÕES FINAIS

Após uma busca exploratória sobre o tema foi constatado que as previsões de buscas por facetas apresentam sugestões sobre possíveis ações, facilitando a seleção de busca de um metadado dentro das hierarquias de facetas.

O usuário engajado em atividades de buscas pode precisar de auxílio para descobrir novas associações e tipos de conhecimento, resolver problemas complexos, ou desenvolver a compreensão da terminologia e da estrutura do espaço de informação. No contexto da Web, muitos desses requisitos podem ser alcançados por meio do uso sistemático da infra-estrutura da Web Semântica. No entanto, as pessoas esperam (ou têm a expectativa) de que a Web também seja útil em outras situações para as quais os sistemas e ferramentas disponíveis deveriam fornecer suporte mais adequado.

A pesquisa por informação é uma atividade fundamental na vida das pessoas. As tecnologias da Web Semântica fornecem a infra-estrutura necessária para o desenvolvimento de novas abordagens para auxiliar as pessoas a encontrar a informação que elas estão procurando. A investigação de modelos de navegação para a Web Semântica pode fornecer informações para melhorar as interfaces e os mecanismos da próxima geração de sistemas de navegação (PANSANATO, 2007).

Para trabalhos futuros existem muitos testes para promover a utilização em maior escala de buscas por facetas de metadados em geral. As hierárquicas facetadas de metadados são utilizadas em muitos sites comerciais, como por exemplo, *e-Bay* e *Shopping.com* e estes sites estão experimentando a utilização com diferentes variações de implementações. No entanto, é um conceito para ser utilizado amplamente e regularmente pelos sites comerciais, sendo assim as técnicas implementadas poderão ser ampliadas e utilizadas regularmente em vários outros domínios que necessitam implementar buscas e consultas mais eficientes em sites *web*.

REFERÊNCIAS

BAEZA-YATES, R. & RIBEIRO NETO, B. **Modern Information Retrieval**. ACM Press, Addison Wesley, 1999.

BUCKLAND, M.K. **Descrição e Pesquisa: Metadados como Infra-estrutura, Brazilian Journal of Information Science**. Universidade da Califórnia. Artigo recebido em 2006.

HEARST, M.A. **Clustering versus Faceted Categories for Information Exploration**. Artigo **COMMUNICATIONS OF THE ACM** April 2006/Vol. 49, No. 4 Disponível em: <<http://flamenco.berkeley.edu/papers/cacm06.pdf>>, "Acesso em: 15 ago 2009".

HEARST, M.A. **Design Recommendations for Hierarchical Faceted Search Interfaces**, <<http://flamenco.berkeley.edu/papers/faceted-workshop06.pdf>>, "Acesso em: 13 ago 2009".

HEARST, M.A. **Next generation web search: Setting our sites**. IEEE Data Engineering Bulletin: Special Issue on Next Generation Web Search, 2000.

LUNARDI, Márcia S.; CASTRO José Muanis F. de; MONAT, André S. **A utilização de nuvens de texto para a visualização de resultados de sistemas de busca na web**. Anais do 8º Congresso Brasileiro de Pesquisa e Desenvolvimento em Design - P&D Design 2008, São Paulo, SP. * Artigo Premiado

MACEDO, J. **Recuperação de Informação Textual Distribuída por Fontes Autônomas com Sobreposição**. (Tese de Doutorado) Universidade do Ninho, Portugal, Julho, 2001.

MARCHIONI, G. **Exploratory search: From finding to understanding**. Communications of the ACM, v. 49, n. 4, p. 41-46, 2006.

MEIRELLES, Junia Cristina J. P., MOURA, Mônica. **Web 2.0: novos paradigmas projetuais e informacionais**. Infodesign - Revista Brasileira de Design da Informação v. 4, n. 2, p. 12-19, 2007.

PANSANATO, L. T. E. **Um modelo de navegação exploratória para a infra-estrutura da Web Semântica**. 164 p. Tese (Doutorado) – ICMC/USP, São Carlos, 2007.

PANSANATO, L.T.E. **Navegação em repositórios de informação baseados na Web: metadados como suporte a passos de navegação em uma estratégia de orientação**. Usp – São Carlos. 2004.

PRAZERES, C.; Lucrédio, D.; Fortes, R.; Teixeira, C. **Uma Proposta de Navegação na Web Utilizando Facetas**. Submetido ao WebMedia 2005.

SILVA, C.E. **Um componente Java para busca de facetas de metadados.** Monografia (Latu-Sensu) – UTFPR, Cornélio Procópio, 2009.

TEEVAN, J., ALVARADO, C., ACKERMAN, M. S. & KARGER, D. R.

The Perfect Search Engine Is Not Enough: A Study of Orienteering Behavior in Directed Search. Proceedings of the SIGCHI conference on Human factors in computing systems, p. 415-422, Vienna, Austria, 2004.