



UNIPAC

UNIVERSIDADE PRESIDENTE ANTÔNIO CARLOS
FACULDADE DE CIÊNCIAS DA SAÚDE DE BARBACENA

CURSO DE CIÊNCIA DA COMPUTAÇÃO

Veruska Jaqueline Lopes

TÉCNICAS DE ANÁLISE PARA A USABILIDADE DE DIFERENTES MÁQUINAS DE BUSCA

BARBACENA
NOVEMBRO DE 2004

VERUSKA JAQUELINNE LOPES

**TÉCNICAS DE ANÁLISE PARA A USABILIDADE DE DIFERENTES
MÁQUINAS DE BUSCA**

Trabalho de Conclusão de Curso
apresentado à Universidade Presidente
Antônio Carlos, como requisito parcial
para obtenção do grau de bacharel em
Ciência da Computação.

ORIENTADORA: Prof^a. Lorena Sophia Campos de Oliveira

BARBACENA
NOVEMBRO DE 2004

Veruska Jaquelinne Lopes

TÉCNICAS DE ANÁLISE PARA A USABILIDADE DE DIFERENTES MÁQUINAS DE
BUSCA

Trabalho de Conclusão de Curso
apresentado à Universidade Presidente
Antônio Carlos, como requisito parcial
para obtenção do grau de bacharel em
Ciência da Computação.

Aprovada em ____/____/____

BANCA EXAMINADORA

Prof. Lorena Sophia Campos de Oliveira (Orientadora)

Universidade Presidente Antônio Carlos

Prof. Gustavo Campos Menezes

Universidade Presidente Antônio Carlos

Prof. Luís Augusto Mattos Mendes

Universidade Presidente Antônio Carlos

Dedico este a todos os meus colegas do 8º período por lutarmos juntos, nessa grande conquista.

Dedico este projeto àqueles que querem ter um conhecimento maior sobre o funcionamento e diferenças entre máquinas de buscas, àqueles que desejam se aprimorar nas pesquisas realizadas na *Web* para obterem maiores sucessos e também aqueles que apenas de curiosidade podem estar lendo e aprendendo.

Agradeço em primeiro lugar a Nossa Senhora Aparecida e ao meu bom Deus Pai, por estarem sempre ao meu lado me proporcionando força e inteligência, para a realização deste.

Aos meus Pais Betinho e Leleca, por sempre me apoiarem, graças a eles pude fazer uma faculdade. Aos meus irmãos Jê e Chris. A Jê me ajudou nessa jornada, fornecendo materiais e explicações.

Ao meu adorável sobrinho Caio, pelas horas de descontração, quando o *stress* aumentava. A galera do fundão, pelas brincadeiras e alegrias. A Rafa, a Dani, a Rose, o Itamar e o Fausto por serem meus amigos e me apoiarem nas horas tristes. Ao Marcos e claro a minha orientadora Lorena, pelos materiais e idéias.

Obrigada a todos que de certa forma contribuíram para a realização desta grande conquista.

Lista de Figuras

Figura 1 - Tarefas que podem ser executadas por um usuário num SRI....	25
Figura 2 – cachorro	31
Figura 3 – pulga	31
Figura 4 - cachorro E pulga	31
Figura 5 - cachorro NÃO pulga	31
Figura 6 - cachorro OU pulga	32
Figura 7 - gato E pulga E cachorro	32
Figura 8 - (pulga OU gato) NÃO cachorro	33
Figura 9 – Gráfico de idade das pessoas entrevistadas	42
Figura 10 – Gráfico de sexo das pessoas entrevistadas	42
Figura 11 – Gráfico de escolaridade das pessoas entrevistada	43
Figura 12 – Gráfico de familiaridade com a <i>Internet</i> das pessoas entrevistada	44
Figura 13 – Gráfico de frequência de utilização de site de busca	44
Figura 14 – Gráfico de frequência de utilização de máquina de busca	45
Figura 15 – Gráfico de “encontra o que procura”	46
Figura 16 – Gráfico de “Você sabe o que são operadores booleanos”	46
Figura 17 – Gráfico de “Você sabe como utilizar os operadores booleanos”	47
Figura 18 – Gráfico da maior dificuldade na utilização de máquinas de busca	48
Figura 19 – Gráfico de “O que você acha das interfaces das máquinas de	

busca”	49
Figura 20 – Gráfico de “Você sabe como as máquinas de busca procuram informação na <i>Web</i> ”.	50
Figura 21 – Gráfico “O que você acha que poderia mudar para melhorar os resultados encontrados pelas máquinas de busca”	51
Figura 22 – Interface da máquina de busca Google	56
Figura 23 – Interface da máquina de busca Yahoo	57
Figura 24 – Interface da máquina de busca Radar Uol	57
Figura 25 – Interface da máquina de busca Todo Br	58
Figura 26 – Interface da máquina de busca Alta Vista	58
Figura 27 – <i>Internet explorer</i> - entrando no site de busca inteligente.....	97
Figura 28 – <i>interface</i> da máquina de busca inteligente	97
Figura 29: Proposta de uma máquina de busca inteligente para <i>Web</i>	101

Lista de Tabelas

Tabela 1 - Operadores booleanos	30
Tabela 2 – Exemplos de argumentos de pesquisas mais refinados	39

Lista de Siglas

FI – Filtragem de Informação

RI – Recuperação de Informação

SRI – Sistema de Recuperação de Informação

WEB – Wolrd Wide Web

SUMÁRIO

INTRODUÇÃO	16
CAPÍTULO 2 – HISTÓRIA	20
2.1 – A história da recuperação	20
2.2 – Recuperação de informação	20
2.3 – Relevância	21

2.4 – Modelo para recuperação de informação	22
2.4.1 – Modelo booleano	22
2.4.2 – Modelo vetorial	23
2.4.3 – Modelo probabilístico	23
2.5 – A tarefa do usuário	24
2.6 – A Web e as bibliotecas digitais	25
CAPÍTULO 3 – MÁQUINAS DE BUSCA	27
3.1 – O que são máquinas de busca	27
3.2 – Para que serve as máquinas de busca	28
3.3 – Como utilizar as máquinas de busca	28
3.4 - Operadores booleanos	29
3.4.1 – Funcionamento da pesquisa utilizando operadores booleanos	30
3.4.2 - Operadores implícitos	33
3.5 – Tipos de pesquisa	34
3.5.1 - Pesquisa com uma única palavra	34
3.5.2 - Pesquisa com várias palavras	35
3.5.3 - Pesquisa com uma frase exata	36
3.5.4 - Pesquisa por prefixo	36
3.5.5 - Pesquisa por exclusão	37
3.5.6 - Pesquisa por proximidade	38
3.5.7 – Argumentos de pesquisa mais refinados	38
3.6 - Termos especiais de pesquisa	39
3.7 – Pesquisa I – “Avaliação da usabilidade de máquinas de busca”	41
3.7.1 – Resultado da pesquisa I	42

3.8 – Pesquisa II – “Utilização de diferentes máquinas de busca”	53
3.8.1 – Etapas seguidas pelo usuário	54
3.8.2 – Conclusões	56
4 – TIPOS SISTEMAS DE BUSCA	59
4.1. Diretórios	59
4.1.1 – Como é formada a base de dados	61
4.2 – Exemplo de sistema de busca – diretório.....	61
4.2.1 – Yahoo	62
4.2.2 – História	62
4.2.3 – Características	63
4.2.4 – Como é feita a busca	64
4.3 – Mecanismo de busca	65
4.3.1 – Robô	65
4.3.2 – Base de dados	66
4.3.3 – <i>Software</i>	67
4.3.4 – Como é formada a base de dados	68
4.4 – Exemplos de mecanismo de busca	69
4.4.1 – Google	69
4.4.2 – História	69
4.4.3 – Características	70
4.4.4 – Como é feita a busca	71
4.5 – Metabuscadores	72
4.6 – Exemplo de metabuscador	73
4.6.1 – Radar Uol	73
4.6.2 – História	73

4.6.3 – Característica	74
4.6.4 - Como é formada a base de dados	74
4.7 – Comparação entre as categorias de busca	75
4.7.1 – Diretórios	75
4.7.2 – Mecanismos de busca	76
4.7.3 – Metabuscadores	76
4.7.4 – Resultados	77
5 – INTRODUÇÃO A AGENTES INTELIGENTE	78
5.1 – O que são agentes inteligente	78
5.2 – Como são utilizados	79
5.2.1 – Possíveis aplicações	79
5.3 – Características	81
5.3.1 – Autonomia	81
5.3.2 – Mobilidade	82
5.3.3 – Cooperação	82
5.3.4 – Comunicabilidade	83
5.3.5 – Aprendizagem	83
5.3.6 – Reatividade	84
5.3.7 – Habilidade social	84
5.3.8 – Pró-atividade	84
5.4 – Comunicações entre agentes	85
5.5 – Tipos de agentes	86
5.5.1 – Agentes colaborativos	86
5.5.2 – Agentes de interface	87
5.5.3 Agentes assistentes	87

5.5.4 Agentes de recuperação de informação	88
5.6 – Questões básicas de um agente	89
5.6.1 – Resolução de problemas	89
5.6.2 – Representação do conhecimento	90
5.6.3 – Aprendizado	90
5.7 – Proposta de um agente inteligente para a filtragem de informações em máquinas de busca – FILVE2004	91
5.7.1 – Compreendendo as necessidades dos usuários	91
5.7.2 – Motivação	92
5.7.3 – Características	93
5.7.4 – Objetivos do FILVE2004	94
5.7.5 – Funcionamento	96
5.7.5.1 – Exemplo de funcionamento do filve2004	96
CAPÍTULO 6 – UMA NOVA MÁQUINA DE BUSCA	99
6.1 – Esquema geral da máquina de busca inteligente	101
CONCLUSÃO	102
7.1 – Trabalhos futuros	103
BIBLIOGRAFIA	104
ANEXO I – Formulário Pesquisa I	108
ANEXO II - Resultado pesquisa II	112

INTRODUÇÃO

O grande volume de informação tomada disponível consolidou a *Web* (*World Wide Web*) como a principal fonte de referência para busca de informação. Entretanto, isto introduziu, no contexto da *Web*, um problema antigo, que é a dificuldade de se encontrar informação de interesse ou relevante [2]. Para enfrentar este problema, foram criados os mecanismos de busca e os diretórios. Enquanto os diretórios apresentam uma classificação precisa, elaborada manualmente, a partir de um conjunto pequeno de domínios, os mecanismos de busca possibilitam busca ampla em um conjunto muito maior.

Uma máquina de busca percorre a *Web*, coletando páginas diversas e fazendo uma cópia local das mesmas. Sobre esta base são gerados índices, que são utilizados

para responder as consultas dos usuários. Este processo permite ao usuário localizar rapidamente as páginas de seu interesse.

O grande problema é que na maioria das vezes o resultado obtido através de pesquisas realizadas nas máquinas de busca é muito grande, e com isso a dificuldade de encontrar realmente o que se quer se torna uma tarefa muito difícil.

Serão apresentadas técnicas que ajudam os usuários a efetivarem suas buscas, com utilização de operadores booleanos, obtendo assim sucessos nas pesquisas realizadas.

Foi realizada uma pesquisa de campo justamente para mostrar que a maior dificuldade dos usuários é realmente procurar diante de tantas URLs.

Esse projeto tem como objetivo resolver este problema, utilizando um agente inteligente para atuar justamente nessa deficiência que é a filtragem de informação.

Este trabalho se constitui num embrião de algo que pode ser, sem dúvida, muito interessante. A cada instante uma gigantesca quantidade de informações das mais diversas naturezas são colocadas à disposição de quem as desejar na *Web*. Quanto maior a quantidade de informação maior é a dificuldade de selecionar e tratar estes dados. Ferramentas que facilitem esta tarefa serão de fato inestimáveis para quem necessita ou deseja utilizar as informações disponibilizadas na rede.

Este trabalho irá tratar sistemas de busca com o mesmo significado de máquinas de busca.

A Justificativa deste trabalho é devido ao grande número de informações encontradas hoje na *Web*, se torna necessário não apenas a recuperação de informação, mas sim a filtragem de informações. Efetivar uma pesquisa em máquinas de busca é uma tarefa simples, porém encontrar o que se realmente procura é uma tarefa árdua.

Este trabalho foi desenvolvido utilizando-se como metodologia a pesquisa bibliográfica, a pesquisa da *Internet*, pesquisas realizadas com usuários e análise

comparativa de sistemas de busca. A partir das observações e características levantadas na literatura foi proposta uma nova máquina de busca inteligente, capaz de filtrar informações de acordo com o perfil de cada usuário.

Com o advento da *Internet* e principalmente com a grande influência atual que ela representa para o mundo, cresce a variedade de sistemas de busca voltados para vários tipos de RI. O aperfeiçoamento das máquinas de busca tem sido foco em vários estudos.

No entanto, o crescimento do número de sites de busca é inversamente proporcional à qualidade de informações relevantes que eles fornecem.

O desenvolvimento de uma comparação entre sistemas de buscas (mecanismos, diretórios e metabuscadores) contribuirá para uma melhor qualidade de utilização e servirá como base para uma proposta de uma máquina de busca mais eficiente.

Os principais objetivos almejados com esse trabalho são:

- Explicar o surgimento e para que servem as máquinas de busca;
- Apresentar técnicas e operadores booleanos para auxiliar usuários em pesquisas;
- Desenvolver uma comparação entre sites de busca, fornecendo assim, vantagens e desvantagens;
- Realizar pesquisas para demonstrar quais os sistemas de busca mais utilizados;
- Verificar as deficiências dos sistemas de busca;
- Propor um agente para tratar tais deficiências;
- Analisar todo o conteúdo e propor uma máquina de busca inteligente.

Os próximos capítulos descrevem todas as etapas percorridas para o desenvolvimento desse trabalho

No capítulo 2, Será abordada a história da recuperação, a recuperação de informação, o significado de relevância, os três modelos para a recuperação de informação, que são o modelo booleano, o modelos vetorial e o modelo probabilístico. Será apresentada a tarefa do usuário, a *Web* e as bibliotecas digitais.

No capítulo 3, serão abordadas as máquinas de busca, para que servem, como utilizá-las, com o intuito de auxiliar os usuários serão apresentados também os operadores booleanos, o funcionamento da pesquisa utilizando operadores booleanos, os operadores implícitos, os tipos de pesquisas de campo com exemplos e termos especiais para realização de consultas na *Web*. Conterá também as pesquisas realizadas com usuários e seus respectivos resultados.

No capítulo 4, Serão abordados os sistemas de busca, que serão divididos em três categorias: - Diretórios; - Mecanismo de busca; - Metabuscadores. Em cada categoria será apresentado um exemplo de máquina de busca mais utilizado, conforme o resultado da pesquisa de campo realizada no capítulo anterior, contendo sua história, características e como é realizado a busca. Haverá uma comparação entre os sistemas de busca e serão mostrados os resultados.

No capítulo 5, serão abordados os conceitos sobre agente inteligente, o que são, como utilizá-los, algumas aplicações, as suas características, a comunicação entre vários agentes. Serão apresentados quatro tipos de agente, as questões básicas de um agente e também será proposto um agente inteligente, para a filtragem de informação em máquina de busca.

No capítulo 6, Será proposto uma nova máquina de busca utilizando agente inteligente.

No capítulo 7, será apresentada a conclusão desse trabalho, sugestão para futuros trabalhos e a bibliografia.

CAPÍTULO 2 – HISTÓRIA

2.1 – A história da recuperação

“Recuperação de Informação ou *Information Retrieval* (RI ou IR) é a atividade de recuperar itens de informação armazenados em um meio que possa ser acessado por computador. Um item de informação é geralmente constituído de texto (tais como documentos diversos, páginas *Web*, livros, etc.), embora possa conter outros tipos de dados tais como fotografias, gráficos e figuras. RI lida com a representação,

armazenamento, organização e acesso a itens de informação. A representação e organização dos itens de informação devem prover ao usuário acesso facilitado à informação de seu interesse” [7].

2.2 – Recuperação de Informação

Por aproximadamente 4000 anos o homem tem organizado informações para serem recuperadas e usadas posteriormente, como por exemplo, uma tabela de conteúdo de um livro. Como o acervo de livros cresceu, uma estrutura teve de ser criada para acessar de forma mais rápida as informações armazenadas nos livros.

Uma estrutura de dados antiga e muito utilizada para rápida RI, é uma coleção de palavras selecionadas ou conceitos, com as quais estão ligados apontadores às informações relacionadas – o índice. Por séculos, índices foram criados manualmente como uma categorização de hierarquias, isto nos dá uma idéia de ordenação. Mais recentemente, o advento dos computadores tornou possível a construção automática de índices.

Índices criados automaticamente produzem uma visão do problema de RI que é mais relacionado ao sistema do que à necessidade do usuário. Existem duas diferentes visões do problema de RI: uma centrada no computador e outra centrada no homem. Na visão centrada no computador, o problema consiste principalmente em construir índices eficientes, processar consultas de usuários com o melhor desempenho possível e desenvolver algoritmos de ordenação que possam trazer qualidade à resposta do usuário. “Na visão centrada no homem, o problema principalmente consiste no estudo da conduta do usuário, no entendimento de suas principais necessidades e em como

este entendimento afeta a organização e as operações do Sistema de Recuperação de Informação (SRI)” [3].

2.3 – Relevância

O SRI deve de alguma forma “interpretar” o conteúdo das informações encontradas nos documentos de uma coleção e ordená-los de acordo com um grau de relevância para o usuário. Relevância é a palavra central de um SRI. É objetivo do SRI recuperar todos os documentos que são relevantes a uma consulta de um usuário, e o menor número possível de documentos não relevantes.

2.4 – Modelo para recuperação de informação

Os modelos clássicos de recuperação de informação apresentam estratégias de busca de documentos relevantes para uma consulta. Tanto a consulta feita pelo usuário, quanto os documentos que compõem a coleção a ser pesquisada, são representados pelos seus termos. Os Modelos Clássicos são booleano, vetorial e probabilístico.

2.4.1 – Modelo Booleano

O modelo booleano é um dos modelos clássicos que considera uma consulta como uma expressão booleana convencional, que liga seus termos através de conectivos lógicos AND, OR e NOT (Ver cap. 3.4). No modelo booleano um documento é considerado relevante ou não relevante a uma consulta. Não existe resultado parcial e não há informação que permita a ordenação do resultado da consulta.

Este modelo é muito mais utilizado para recuperação de dados do que para recuperação de informação. É bom para quem entende bem de álgebra booleana, mas o usuário na maioria dos casos não entende. Tem as vantagens de ter expressividade completa se o usuário souber exatamente o que quer e ser facilmente programável e exato. Porém, a saída não é ordenada.

2.4.2 – Modelo Vetorial

O modelo espaço-vetorial (ou simplesmente vetorial) foi desenvolvido por Gerard Salton, para ser utilizado num SRI chamado SMART. No modelo vetorial cada documento é representado como um vetor de termos e cada termo possui um valor associado que indica o grau de importância (peso) deste no documento. A consulta do usuário também é representada por um vetor. Desta forma, os vetores dos documentos podem ser comparados com o vetor da consulta e o grau de similaridade entre cada um deles pode ser identificado. Os documentos mais similares (mais próximos no espaço) à consulta são considerados relevantes para o usuário e retornados como resposta para ela.

Deste modelo pode-se dizer que atribuir pesos aos termos melhora o desempenho e que os documentos são ordenados de acordo com seu grau de similaridade com a consulta. Mas possui a desvantagem de ser um modelo generalizado e de que um documento relevante pode não conter termos da consulta.

2.4.3 – Modelo Probabilístico

Possui esta denominação justamente por trabalhar com conceitos provenientes da área de probabilidade e estatística. Neste modelo os termos indexados dos documentos e das consultas não possuem pesos pré-definidos. A ordenação dos documentos é calculada pesando dinamicamente os termos da consulta relativamente aos documentos. É baseado no princípio da ordenação probabilística (*Probability Ranking Principle*). Neste modelo, busca-se saber a probabilidade de um documento ser ou não relevante para uma consulta. Tal informação pode ser obtida assumindo-se que a distribuição de termos na coleção seja capaz de informar a relevância provável para um documento qualquer da coleção.

As vantagens deste modelo são: os documentos são ordenados de forma decrescente por suas probabilidades de serem relevantes, por obrigar o Princípio da Ordenação Probabilística. As desvantagens são: Assume a independência entre os termos. Não há como calcular a ordem (dado um documento este é relevante para a consulta se o termo está presente) ao iniciar a execução do sistema (uma vez que os documentos ainda não são conhecidos). O modelo não faz uso da frequência dos termos no documento.

2.5 – A tarefa do usuário

O usuário de um SRI tem que traduzir sua necessidade de informações em uma consulta, escrita na linguagem fornecida pelo sistema. Geralmente isto implica em especificar um conjunto de palavras que conduzam à semântica de sua necessidade. Neste caso, o usuário está buscando por informações úteis executando uma tarefa de recuperação.

Em outro caso, um usuário que não possua uma consulta com interesses muito específicos. Por exemplo, interesse em “carros velozes”, tanto faz se forem carros de Fórmula 1, ou outro tipo de corrida, ou se forem carros velozes para uso doméstico, etc. Uma ferramenta que auxilie este usuário a navegar por diversos documentos de uma coleção de documentos é mais interessante, pois será mais abrangente.

Existe uma distinção entre duas tarefas que podem ser executadas pelo usuário de um SRI: a recuperação de informação ou a navegação entre documentos.

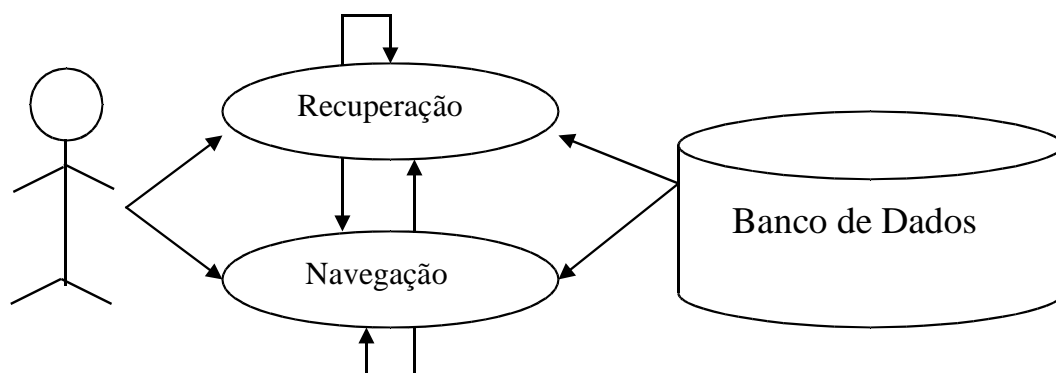


Figura 1 - Tarefas que podem ser executadas por um usuário num SRI

“SRI clássicos normalmente permitem recuperação de informação rápida. Sistemas de hipertexto são geralmente criados para permitir navegação rápida. Bibliotecas digitais modernas e interfaces para a *Web* devem tentar combinar estas duas tarefas” [3].

2.6 – A *Web* e as bibliotecas digitais

As bibliotecas estão entre os primeiros tipos de instituição a adotar sistemas de RI. A primeira geração de sistemas apenas automatizava tecnologias já utilizadas, como uso de catálogos. Na segunda geração foram acrescentadas funções de busca, pelo uso de palavras-chaves e alguns tipos de consultas mais complexas. Na terceira geração, que é a que está sendo atualmente desenvolvida, o principal foco está em criar interfaces gráficas, formulários eletrônicos, características de hipertexto e arquiteturas de sistemas abertos.

Se considerarmos as máquinas de busca da *Web* atualmente, concluímos que elas continuam usando índices similares aos utilizados por bibliotecários há séculos atrás. Porém, três mudanças drásticas ocorreram durante o avanço da tecnologia computacional e no crescimento da *Web*:

Primeira, se tornou muito mais barato ter acesso a várias fontes de informação. Isto permite que seja realizado um número de pesquisas tão grande, como nunca foi possível anteriormente.

Segunda, o avanço em todos os tipos de comunicação digital produziu um acesso ainda maior às redes. Isto implica que as fontes de informações estão disponíveis, mesmo que localmente distantes e que o acesso pode ser realizado rapidamente.

Terceira, a liberdade de divulgar qualquer tipo de informação que uma pessoa julgue útil, isto aumenta cada vez mais a popularidade da *Web*.

Pela primeira vez na história, pessoas têm acesso livre a uma enorme quantidade de publicações de médio e pequeno porte.

“Basicamente baixo custo, maior acesso e liberdade para publicações têm permitido às pessoas usarem a *Web* e as bibliotecas digitais modernas como a forma mais interativa de comunicação” [3].

CAPÍTULO 3 – MÁQUINAS DE BUSCA

3.1 – O que são máquinas de busca

A grande dificuldade da *Internet* é encontrar realmente o que se procura, sem travar uma verdadeira guerra com parâmetros ou ter que digitar a mesma pesquisa em diversos sites. A melhor dica é recorrer aos sistemas de busca, que são de grande

utilidade para se encontrar as informações necessárias, desde que sejam usados de forma correta.

Uma máquina de busca percorre a *Web*, coletando páginas diversas e fazendo uma cópia local das mesmas. Sobre esta base são gerados índices, que são utilizados para responder as consultas dos usuários. Este processo permite ao usuário localizar rapidamente as páginas de seu interesse [2].

As máquinas de busca indexam as informações da *Web*, o que nos permite fazer consultas ao banco de dados da *Web*.

Máquinas de busca são sistemas que têm por objetivo encontrar informação de interesse dos usuários na *Web*. Em termos gerais, elas coletam continuamente os dados disponíveis na *Web* e montam uma grande base de dados que é processada para aumentar a rapidez na RI. “Sem as máquinas de busca seria praticamente impossível encontrar informação na *Internet*, uma vez que há mais de um bilhão de páginas espalhadas em todo o mundo” [14].

3.2 – Para que serve as máquinas de busca

A informação na *Web* está organizada e guardada como hipertexto, páginas HTML ligadas por links, que são exibidas na tela de nosso computador quando selecionamos uma palavra, frase, ícone, desenho marcado como botões nas páginas. Clicando em cada um destes botões, somos levados à página relacionada ao link.

Navegar na *Internet* é muito fácil. Mas, para tornar a navegação mais produtiva e orientada, a *Web* fornece alguns recursos.

“O grande volume de informação tomada disponível consolidou a *Web* como a principal fonte de referência para busca de informação. Entretanto, isto introduziu, no contexto da *Web*, um problema antigo, que é a dificuldade de se encontrar informação de interesse ou relevante” [2]. Para encontrar algo na *Internet* deve-se conhecer o endereço da página, ou através de *sites* conhecidos como máquina de busca.

Para enfrentar este problema, foram criados as máquinas de busca e os diretórios.

Portanto se não existissem as máquinas de busca seriam impossível conseguir buscar informações na *Web*.

3.3 – Como utilizar as máquinas de busca

1. Escolhemos uma máquina de busca. Há várias.
2. Escolhida uma máquina de busca, deve-se digitar em um campo próprio, fornecido pela máquina de busca, a palavra ou palavras relacionada(s) à informação procurada.
3. Digitada a palavra, e pressionado o botão de busca (*search*, nas máquinas de busca estrangeiras), a máquina de busca começa a procurar todas as páginas da *Web* que contém a palavra indicada. Esta procura, às vezes, é demorada.
4. Encontradas as páginas que contém a palavra indicada, a máquina de busca, fornece as URLs - ou seja os endereços que pressionados levam à informação desejada. Mas, na medida em que cresce a quantidade de informação na *Web*, as consultas por palavra-chave estão também se tornando improdutivas. É freqüente que

em resposta a uma consulta, a máquina de busca forneça 1.000 ou mais URLs. Então o que fazer? Alguns conselhos:

a). Deve-se escolher palavras chave precisas. Evitando buscas de uma só palavra! Realize sempre buscas com mais de uma palavra.

b). Faça buscas compostas. Com isto, as buscas ficam mais precisas. Mas, nem sempre é fácil definir o que buscar de forma tão fácil como a apresentada. Portanto, analise muito bem no assunto da consulta para escolher palavras que representem o que realmente quer procurar na *Web*.

3.4 – Operadores Booleanos

George BOOLE foi um matemático inglês que viveu no século XIX. Ele criou um sistema de álgebra e foi um dos precursores da lógica moderna.

Mas o que interessa mesmo é o fato de ele haver criado um modo de conduzir o raciocínio utilizando algumas expressões simples e monossilábicas que ficaram conhecidas como **operadores booleanos**.

A tabela 1 mostra os operadores booleanos aceitos pela maioria dos sistemas de busca, apresenta os operadores em português, os equivalentes em inglês e mais os sinais gráficos correspondentes aceitos por alguns sistemas de busca.

Português	Inglês	Sinais gráficos
E	AND	+

		&
Não	Not	-
OU	OR	
		!

Tabela 1 - Operadores booleanos

3.4.1 – Funcionamento da pesquisa utilizando operadores booleanos

Considere dois animais muito chegados entre si: cachorro e pulga. O usuário necessita fazer uma pesquisa na *Web* que contenham informações sobre esses dois bichos. Antes de iniciar a pesquisa, o usuário deverá ver o que realmente lhe interessa:



Figura 2 - cachorro



Figura 3 - pulga

- Cachorro **E** pulga. Indica que você quer a presença dos dois animais, cachorro e pulga, numa mesma página da *Web*.

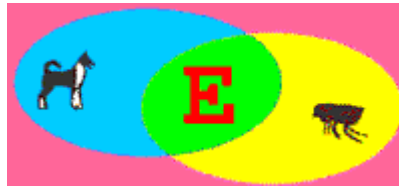


Figura 4 - cachorro **E** pulga

Resultado da sua pesquisa está na região verde.

- Cachorro sem pulga, melhor dizendo: cachorro **NÃO** pulga. Indica que você busca as páginas da *Web* em que há referência ao cachorro, mas **não** à pulga.

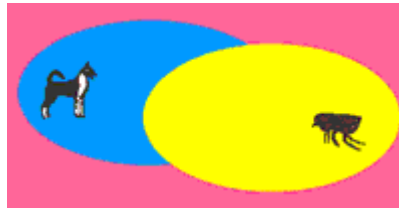


Figura 5 - cachorro **NÃO** pulga

Resultado da sua pesquisa esta na região azul.

- Qualquer dos dois animais: cachorro **OU** pulga. Neste caso, você procura as páginas da *Web* que contiverem referências a um **ou** ao outro animal.



Figura 6 - cachorro **OU** pulga

O resultado da sua pesquisa está na região verde.

Com esses operadores booleanos podem ser criadas pesquisas mais complexas. Veja só como fica ao acrescentar, por exemplo, um gato a essa pesquisa.

Agora as atenções se voltam para cachorros, gatos e pulgas.

- Primeira pesquisa: gatos, pulgas e cachorros. Devem ser mostradas as páginas que contiverem as três palavras.

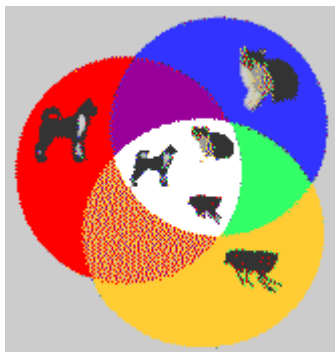


Figura 7 - gato **E** pulga **E** cachorro

O resultado da sua pesquisa está na região branca.

- Segunda pesquisa: pulgas ou gatos, mas sem cachorros.



Figura 8 - (pulga **OU** gato) **NÃO** cachorro

O resultado da sua pesquisa está na região verde

Os parênteses servem para informar ao sistema de busca que primeiro ele deve selecionar todas as páginas que têm pulga **ou** gato e depois ir procurar, nas páginas selecionadas, as que também têm cachorro e excluí-las.

Há sistemas de busca que somente reconhecem os operadores **AND**, **NOT** e **OR**, assim em inglês e com letras maiúsculas. Outros sistemas somente reconhecem os operadores em sua forma gráfica, os sinais gráficos. E há, também os sistemas que reconhecem os operadores em português. Para ter certeza da forma correta de realizar a pesquisa, consulte a página de ajuda do sistema.

3.4.2 – Operadores implícitos

São operadores que não necessitam estar na sua pesquisa, mas o sistema de busca interpreta como se eles estivessem. Quase todos os sistemas de busca os reconhece. [11]

Cada sistema tem um jeito de reconhecer os operadores explícitos. Exemplo, pesquisar as palavras, ***aluguel, casa, praia***, cada sistema vai interpretar a pesquisa de uma forma diferente:

- **O Google** interpreta como se existisse o operador **OR** entre elas: aluguel **OR** casa **OR** praia.
- **O AltaVista** interpreta como se existisse o operador **OR** entre elas: aluguel **OR** casa **OR** praia.
- **O Yahho** interpreta como se existisse o operador **OR** entre elas: aluguel **OR** casa **OR** praia.
- **O Todo BR** interpreta como se existisse o operador **OR** entre elas: aluguel **OR** casa **OR** praia.

- **O Radar UOL**, você tem uma caixa de onde pode ser selecionada uma das seguintes alternativas disponíveis:
- todas as palavras. Equivalente ao operador **E**;
- qualquer uma das palavras. Equivalente ao operador **OU**;

3.5 – Tipos de Pesquisa

3.5.1 – Pesquisa com uma única palavra

Essa é a forma mais simples de recuperação de informação na *Internet*, onde o argumento de pesquisa é composto de uma única palavra.

Exemplo, o usuário está solicitando a recuperação de todos os documentos que possuem pelo menos uma ocorrência da palavra computador, com essa pesquisa ele irá encontrar inúmera páginas, que talvez não o ajudará em nada.

3.5.2 – Pesquisa com várias palavras

Uma outra forma de se efetuar uma pesquisa é digitar uma expressão formada por duas ou mais palavras separadas por conectores, denominados operadores booleanos. A finalidade de um conector é especificar a relação entre as palavras ou termos da pesquisa efetuada pelo usuário.

Por exemplo, ao digitar: *computador e programa*.

O usuário está informando, à máquina de busca, que deseja pesquisar todos os documentos na *Internet* que possuem obrigatoriamente a palavra computador e a palavra programa. Em outras palavras, só serão incluídos no resultado da pesquisa, aqueles documentos que possuem pelo menos uma ocorrência de cada uma dessas duas palavras. Conseqüentemente, não serão considerados os documentos que possuem apenas uma das palavras ou nenhuma delas.

3.5.3 – Pesquisa com uma frase exata

Neste tipo de pesquisa, o usuário solicita à busca por documentos que possuem a frase exata (completa) que foi digitada na caixa de diálogo. Ou seja, os documentos selecionados deverão possuir todas as palavras e na mesma ordem em que elas se encontram na frase.

Para efetuar uma pesquisa por frase exata, o usuário deverá digitar a frase em questão entre aspas duplas ("), como por exemplo: **"Criação de software"**

Cabe ressaltar que nesse caso, os operadores booleanos não são interpretados como tal mas sim como parte da frase sendo pesquisada. Sendo assim, na pesquisa:

"Criação e manutenção de redes"

A palavra **e** não possui obviamente a função de operador booleano entre os termos: Criação e manutenção de redes.

3.5.4 – Pesquisa por prefixo

A pesquisa por prefixo é útil quando o usuário não sabe precisamente a ortografia de uma ou mais palavras que fazem parte do seu argumento de pesquisa. Nesse caso, ele pode digitar a parte inicial da palavra seguida de um asterisco (*). Por exemplo, no argumento de pesquisa: comput*.

A máquina de busca poderá recuperar os documentos que possuem as palavras: **computador, computadores, computação, computacional, etc.**

A pesquisa por prefixo pode ser utilizada na pesquisa por várias palavras como no exemplo a seguir que solicita a recuperação de documentos que possuem a palavra software e as palavras que iniciam com o prefixo comput. Exemplo: comput* e software

Da mesma forma que pode ocorrer com a pesquisa com uma única palavra, o problema a pesquisa por prefixo é que, dependendo do argumento de pesquisa solicitado pelo usuário, pode-se selecionar um número muito elevado de documentos que satisfazem a consulta. Se isso ocorrer, não só o usuário poderá esperar um tempo muito grande para a pesquisa de todos os documentos como também terá um maior

esforço para encontrar aquele(s) documento(s) que de fato o interessa(m) dentro da lista de documentos selecionados.

3.5.5 – Pesquisa por exclusão

Na pesquisa por exclusão, o usuário solicita os documentos que possuem pelo menos uma ocorrência do primeiro termo do argumento, mas não possuem obrigatoriamente o segundo termo. Para isso, ele deverá usar o conectivo NÃO após o operador booleano, **E** que separa os dois termos, como no exemplo a seguir que solicita os documentos que possuem a palavra *computador* mas não possuem a palavra *software*:: computador e não *software*

3.5.6 – Pesquisa por proximidade

A pesquisa por proximidade é semelhante à pesquisa utilizando o operador booleano **E** no sentido de que os dois termos envolvidos devem obrigatoriamente aparecer nos documentos recuperados.

A diferença é que a pesquisa por proximidade, o usuário solicita os documentos que possuem os dois termos digitados próximos entre si.

Para submeter uma pesquisa por proximidade, o usuário deverá usar o conectivo **PERTO** entre os dois termos, como no exemplo: computadores **perto** desempenho

Cabe ressaltar que o significado desse operador pode ser diferente dependendo da forma que determinado produto (máquina de busca) o implementou. Dependendo da máquina de busca, a ordem relativas dos dois termos não é importante. Sendo assim, no caso do exemplo anterior, ele iria recuperar os documentos nas duas ordens de aparição das palavras como em: computadores de alto desempenho ou análise de desempenho de computadores

3.5.7 – Argumentos de pesquisa mais refinados

Os operadores booleanos, bem como os operadores de proximidade (PERTO) e de exclusão (NÃO) podem ser combinados formando argumentos de pesquisa mais refinados. O operador **E** tem precedência em relação ao operador **OU**. No entanto, o usuário pode fazer uso de parênteses tanto para alterar essa precedência quanto para garantir a precisão da consulta a ser efetuada. A tabela 2 apresenta exemplos de argumentos de pesquisa:

Computador e desempenho e fabricante	Recupere os documentos que possuem obrigatoriamente todas as três palavras: Computador, Desempenho e Fabricante.
Computador e (Desempenho ou Fabricante)	Recupere os documentos que possuem a palavra Computador e pelo menos uma das outras duas

	palavras, ou seja, ou Desempenho ou Fabricante ou ambas.
Computador e desempenho e não Fabricante	Recupere os documentos que possuem a palavra Computador e desempenho, mas não possuem a palavra <i>Fabricante</i> .

Tabela 2 – Exemplos de argumentos de pesquisas mais refinados

3.6 - Termos especiais de pesquisa

object: *class*

Localiza páginas que contêm um objeto específico criado por outro programa. A pesquisa *object:dinheiro* localizaria páginas utilizando objetos chamados dinheiro.

domain: *domainname*

Localiza páginas dentro do domínio específico. A pesquisa *domain:com*, localizaria páginas de sites comerciais.

host: *hostname*

Localiza páginas em um computador específico. A pesquisa *host:www.shopping.com*, localizaria páginas no computador do Shopping.com.

image: *filename*

Localiza páginas com imagens tendo um nome de arquivo específico. A pesquisa *image:praias*, localizaria páginas com imagens chamadas praias.

like:URLtext

Localiza páginas similares ou relacionadas ao URL especificado. A pesquisa `like:sfpl.lib.ca.us/`, localizaria sites de bibliotecas universitárias e públicas.

link:URLtext

Localiza páginas com um link para uma página com um texto do URL especificado. A pesquisa `link:www.bacaninha.com.br`, localizaria todas as páginas que fazem link com `bacaninha.com.br`

title:text

Localiza páginas que contêm a palavra ou frase especificada no título da página. A pesquisa `title:Deus` localizaria páginas com a palavra Deus no título.

filetype:pdf

Localiza páginas que estejam em formato .pdf A pesquisa `inteligência filetype:pdf` localizaria páginas apenas sobre inteligência e que estiverem no formato .pdf. pode ser usado qualquer formato (.doc, .ppt, etc.)

3.7 - Pesquisa para avaliação da usabilidade de máquinas de busca

Foi realizada uma pesquisa de campo no período de 10 à 20 de agosto de 2004, na Universidade Presidente Antônio Carlos no Campus Magno em Campolide, com alunos dos cursos de Ciência da Computação, Psicologia e Meio Ambiente. Foram entrevistadas 100 pessoas, porém apenas 64 formulários foram devidamente preenchidos.

As questões foram elaboradas para descobrir se usuários sabem utilizar as máquinas de busca, saber se eles obtêm resultados relevantes, se conhecem os termos especiais de pesquisa quais as verdadeiras dificuldades.

O objetivo da pesquisa é descobrir quais as máquinas de busca mais utilizadas, e através desse resultado fazer comparações entre as 3 mais usadas, verificando suas vantagens e desvantagens.

Esta também teve como principal objetivo descobrir a maior dificuldade de usuários na utilização de máquinas de busca, e através dessa dificuldade propor um melhoramento.

O modelo do formulário utilizado na pesquisa de campo está descrito no Anexo I.

3.7.1 – Resultado da Pesquisa

1) Gráfico de idade das pessoas entrevistada

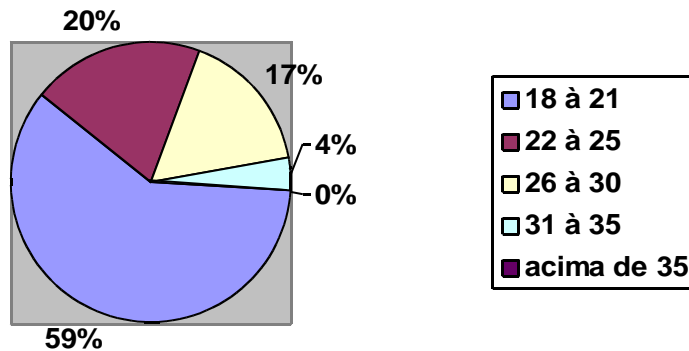


Figura 9 – Gráfico de idade das pessoas entrevistada

Conclusão

Como se pode observar o gráfico acima mostra que a maioria da utilização de máquinas de busca são realizada por pessoas entre a faixa etária de 18 à 21 anos.

2) Gráfico de sexo das pessoas entrevistadas

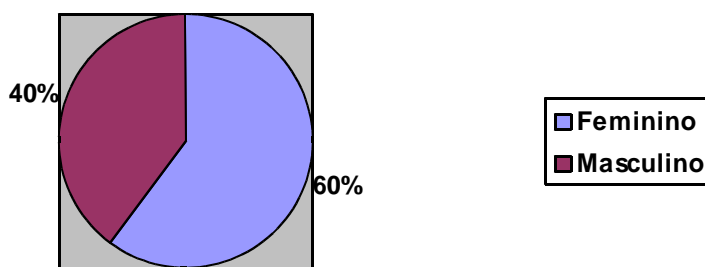


Figura 10 – Gráfico de sexo das pessoas entrevistadas

Conclusão

Como se pode observar o gráfico acima mostra que a maioria da utilização de máquinas de busca são realizadas por pessoas do sexo feminino.

3) Gráfico de escolaridade das pessoas entrevistada

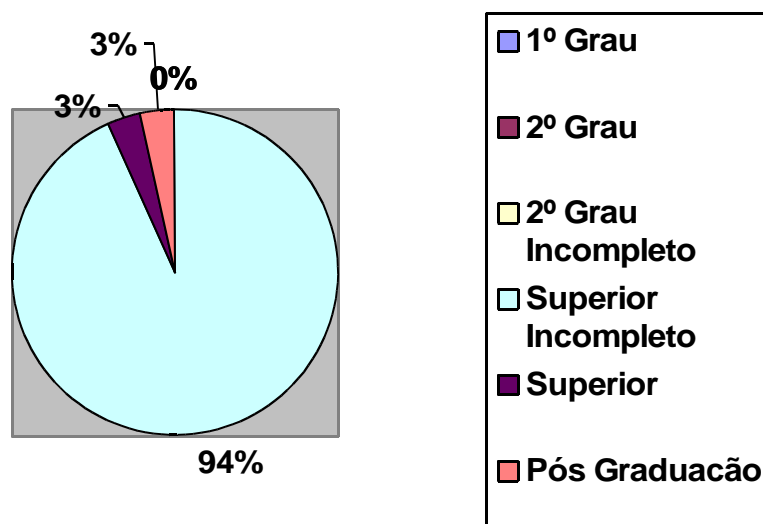


Figura 11 – Gráfico de escolaridade das pessoas entrevistada

Conclusão

Como se pode observar no gráfico acima mostra que a maioria das pessoas entrevistadas, possui a escolaridade superior incompleto, devido ao cenário onde foi realizada a pesquisa, na Universidade.

4) Gráfico de familiaridade com a *Internet* das pessoas entrevistada

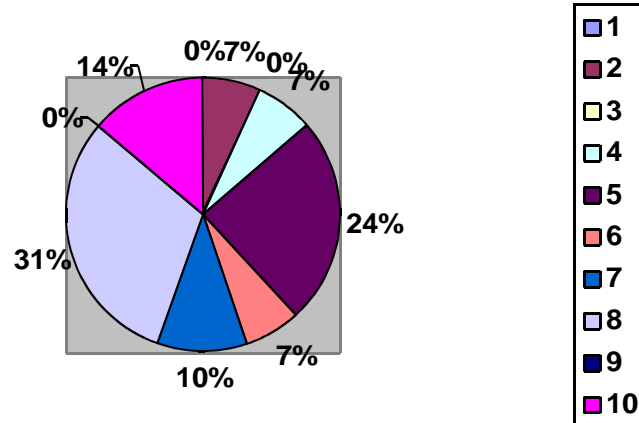


Figura 12 – Gráfico de familiaridade com a *Internet* das pessoas entrevistada

Conclusão

Como se pode observar no gráfico acima mostra que mesmo com uma grande avanço da tecnologia, muitas pessoas ainda não possui uma familiaridade adequada com a *Internet*.

5) Gráfico de freqüência de utilização de site de busca



Figura 13 – Gráfico de freqüência de utilização de site de busca

Conclusão

Como se pode observar no gráfico acima, as pessoas estão cada vez mais utilizando as máquinas de busca, pois se não as existissem, seria praticamente impossível encontrar informações na *Web*.

6) Gráfico de frequência de utilização de máquina de busca

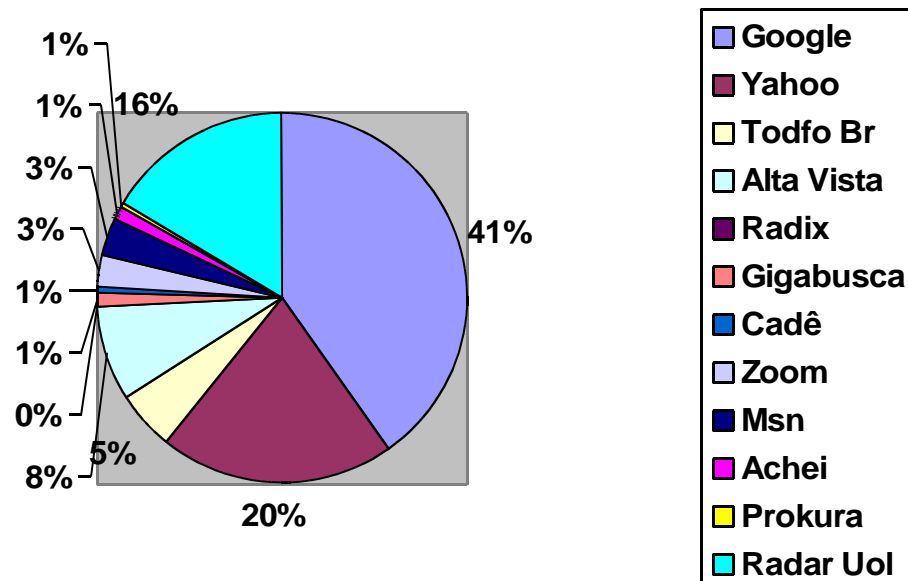


Figura 14 – Gráfico de frequência de utilização de máquina de busca

Conclusão

Como se pode observar no gráfico acima, o mecanismo de busca Google é o primeiro mais utilizado, logo depois vem o diretório Yahoo, e o metabuscador Radar Uol (apesar de não haver a opção na pesquisa do metabuscador Radar Uol, ele apareceu como unânime na opção aberta **Outras**). A partir desse gráfico, será feita uma comparação entre os mecanismos de buscas, diretórios, e metabuscadores.

7) Gráfico de “encontra o que procura”

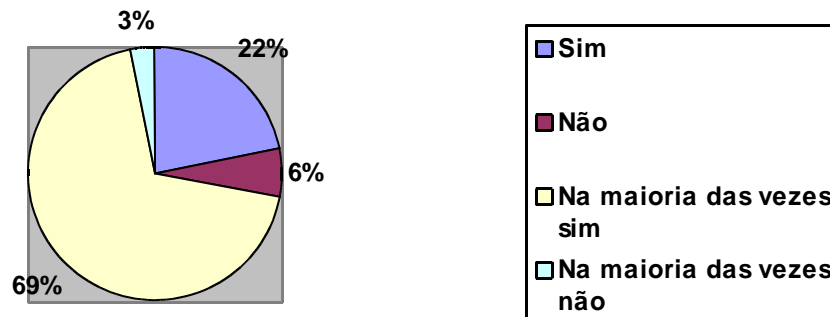


Figura 15 – Gráfico de “encontra o que procura”

Conclusão

Como se pode observar no gráfico acima, a maioria das pessoas sempre encontram o que procuram, mas para que isso ocorra, os usuários tem que possuir um conhecimento sobre como efetivar suas buscas com sucesso, e para isso devem utilizar-se sempre das ajudas encontradas nas máquinas de busca.

8) Gráfico de “Você sabe o que são operadores booleanos”

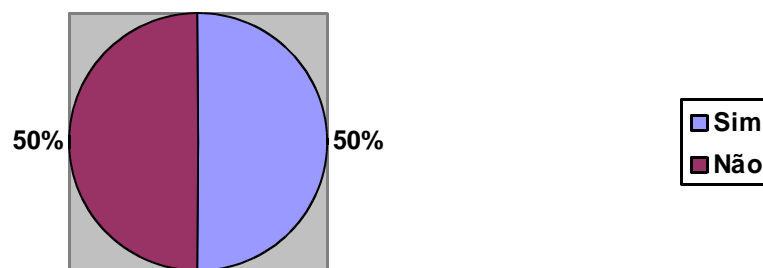


Figura 16 – Gráfico de “Você sabe o que são operadores booleanos”

Conclusão

Como se pode observar no gráfico acima, a metade das pessoas entrevistadas não sabe o que são operadores booleanos, por isso eles foram descritos no capítulo 3.4. para auxiliar aos usuários a efetivarem sua pesquisa com sucesso.

9) Gráfico de “Você sabe como utilizar os operadores booleanos”

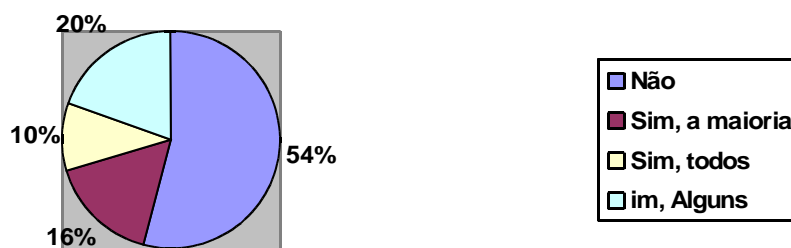


Figura 17 – Gráfico de “Você sabe como utilizar os operadores booleanos”

Conclusão

Como se pode observar no gráfico acima, a maioria das pessoas entrevistadas não sabe como utilizar os operadores booleanos, por isso no capítulo 3.4.1 foi apresentado detalhadamente o funcionamento da pesquisa utilizando operadores booleanos. Através desse funcionamento os usuários irão aprender e poderão obter sucesso em suas buscas.

10) Gráfico da maior dificuldade na utilização de máquinas de busca

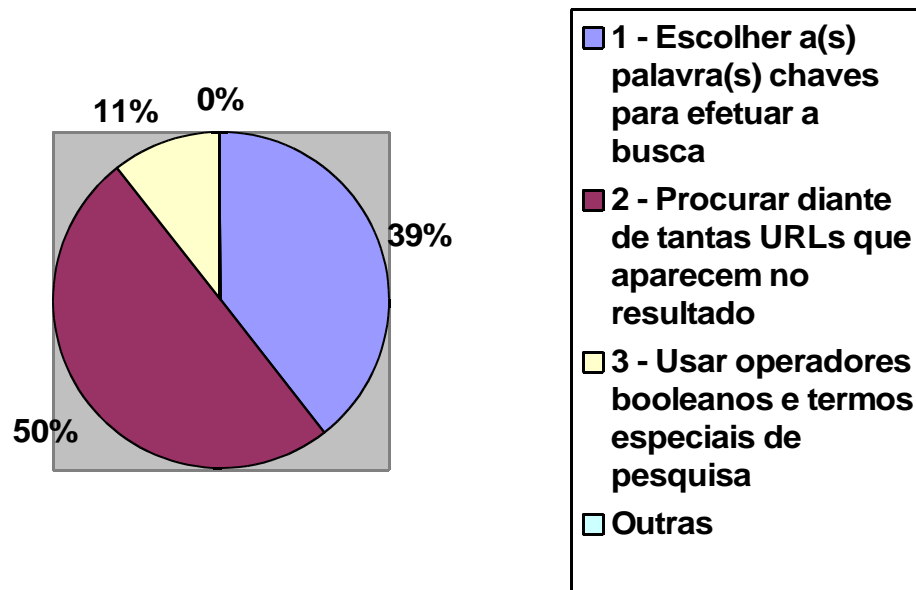


Figura 18 – Gráfico da maior dificuldade na utilização de máquinas de busca

Conclusões

1 – Não há como a máquina de busca resolver, pois na maioria das vezes o usuário não sabe quais as palavras chaves devem fazer parte de sua busca. Porém se os usuários tivessem o hábito de ler as ajudas das máquinas de busca isso o ajudaria na escolha das palavras chaves, e talvez essa dificuldade diminuiria.

2 – Com a máquina de busca proposta no capítulo 5, essa dificuldade praticamente desaparecia, pois o agente proposto agirá exatamente nessa dificuldade, filtrar as informações. Com isso as URLs diminuiriam e irá ficar mais fácil de procurar diante os resultados.

3 – Essa dificuldade foi tratada no capítulo 3, com exemplos e técnicas especiais, sendo que nas ajudas das máquinas de busca o usuário também encontrariam esse tipo de informações.

4 – Apesar da questão ser aberta, não houve nenhum formulário preenchido, com uma resposta.

11) Gráfico de “O que você acha das interfaces das máquinas de busca”

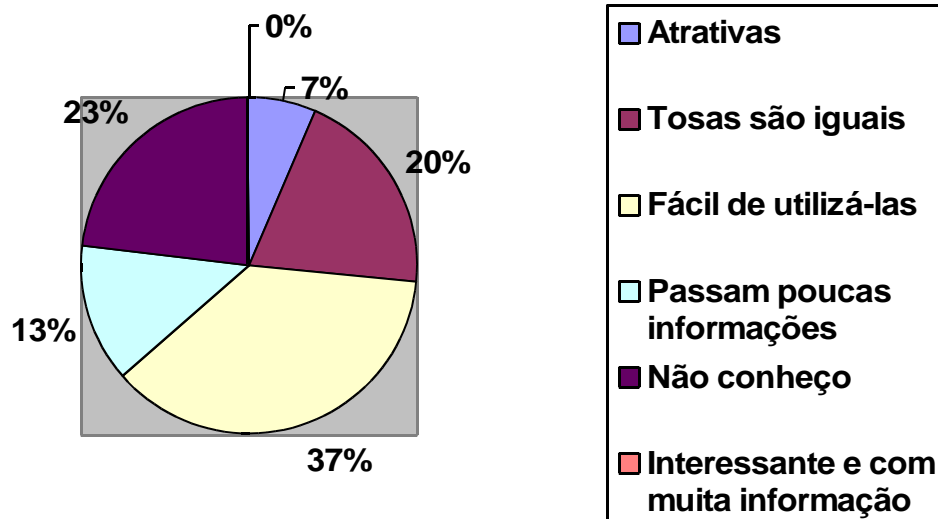


Figura 19 – Gráfico de “O que você acha das interfaces das máquinas de busca”

Conclusão

Como se pode observar no gráfico acima, a maioria das pessoas, acham a utilização da interface em máquinas de buscas muito fáceis, porém pode-se concluir que todas são praticamente iguais. Pois ainda existe aquelas pessoas que não a conhece, porém concluí-se que não é que os usuários não a conhecem, mas não assimilaram a palavra interface com o real significado.

12) Gráfico de “Você sabe como as máquinas de busca procuram informação na Web”.

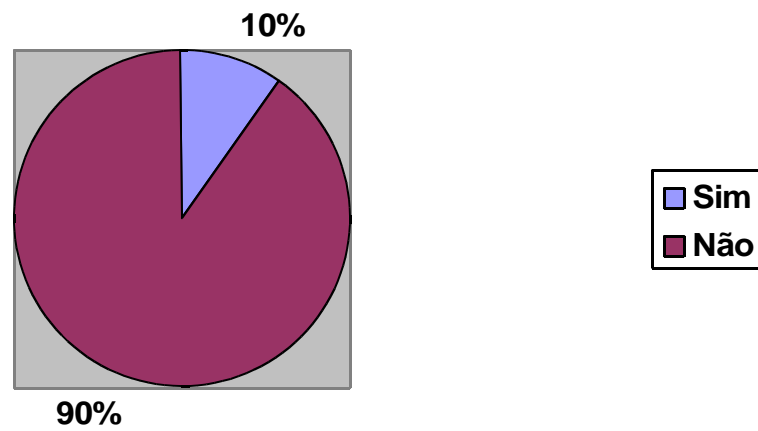


Figura 20 – Gráfico de “Você sabe como as máquinas de busca procuram informação na *Web*”.

Conclusão

Como se pode observar no gráfico acima, a maioria das pessoas não tem o conhecimento de como as máquinas de busca efetivam suas busca na *Web*, apesar da alternativa sim da pesquisa ser aberta, não houve nenhuma qualificação de resposta.

13) O que você acha que poderia mudar para melhorar os resultados encontrados pelas máquinas de busca

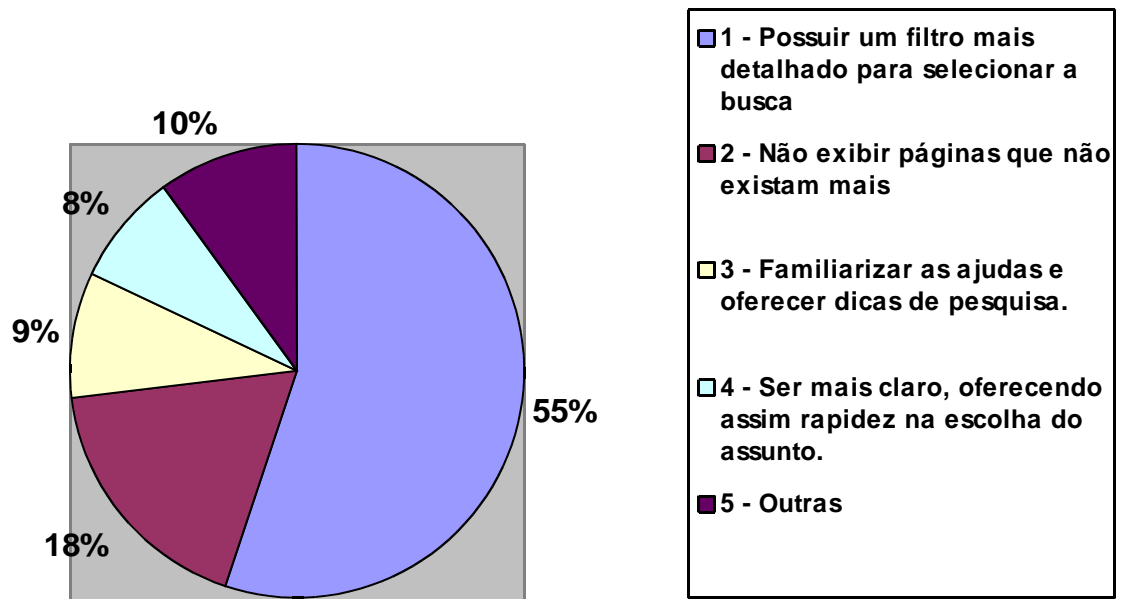


Figura 21 – Gráfico “O que você acha que poderia mudar para melhorar os resultados encontrados pelas máquinas de busca”

Conclusões

1 – Isso o agente FILVE2004 proposto no capítulo 4 irá fazer, filtrar as buscas de acordo com o perfil de cada usuário, com isso as buscas serão mais detalhadas.

2 – O mecanismo de busca google possui um robô chamado de GoogleBot - que, periodicamente, varre toda a *Internet*, navegando por todos os *links* para eliminar as repetições, excluir os *links* quebrados e atualizar a classificação do PageRank, com esse robô o google é o melhor mecanismo de busca para acabar com essas páginas inexistentes, porém ele não consegue eliminá-las completamente.

3 – Isso já existe, porém os usuários não têm o hábito de quando forem utilizar as pesquisas, primeiro dar uma olhada nas ajudas. Elas oferecem dicas e são bem claras, para ajudar os usuários a efetivarem suas buscas.

4 – Essa dificuldade está diretamente ligada a nº 1, pois para ser mais claro o sistema deve retornar resultados relevantes e só através do agente FILVE2004 esses resultados serão filtrados.

5 – Essas outras respostas não foram tão significantes para a pesquisa em questão, e através da leitura desse Trabalho de Conclusão de Curso, essas dificuldades irão desaparecer. Segue abaixo alguns exemplos de respostas de Outras.

- Possuir uma tecla para achar algoritmos;
- Ter máquinas só para esportes;
- Entre outras.

Obs.: As respostas foram analisadas com o seguinte critério, qual a real dificuldade, do usuário, foi analisadas as respostas e classificando as que tem o mesmo significado, porém com palavras diferentes.

14) Existem os Termos especiais de pesquisa: Qual o termo você acha que deveria existir.

As respostas obtidas não trouxeram nada de novo, todas as técnicas (termos) apresentadas já existiam nas máquinas de busca, porém os usuários não possuem esses conhecimentos, porque não fazem uso de ajudas encontradas em máquinas de busca e nem de busca avançadas, que possuem todas essas técnicas.

3.8 - Pesquisa II “Utilização de diferentes máquinas de busca”

Serão testadas as máquinas de busca mais populares de acordo com a pesquisa I realizada no capítulo anterior, que são:

- Google
- Yahoo
- Radar Uol
- Altavista
- TodoBr

Será usado operadores booleanos (And,+, -) e termo especial (“ ”) para efetuar as buscas na *Web*.

Será escolhido um tema solto e um tema composto para começar a efetuar as busca.

Serão apresentados resultados de cada máquina de busca.

Serão mostrados os passos que o usuário seguiu, para a realização da pesquisa e conclusões sobre os resultados encontrados.

3.8.1 – Etapas seguidas pelo usuário

1:

O usuário quer saber quanto custa passagens para fazer um curso no exterior, mas não sabe como realizar essa pesquisa, portanto começa sua pesquisa com palavra única.

PASSAGENS

2:

Através dos resultados acima o usuário decidiu que o que quer encontrar não são apenas passagens, mas sim *passagens aéreas*, então resolve fazer uma busca com palavra composta. PASSAGENS AÉREAS

3:

O usuário decide refinar mais sua pesquisa, pois foram encontradas milhares de URLs, usa-se a técnica especial de aspas dupla “ ”. Com isso obtém resultados mais precisos.

4:

Através dos resultados da pesquisa realizada com “ ”, o usuário decide procurar apenas as passagens aéreas que estão em promoção. Com isso obtém-se resultados mais precisos que os anteriores.

5:

Através dos resultados da pesquisa realizada com o operador booleano + , o usuário encontra muitas palavras que fazem parte da sua pesquisa, daí então resolve retirá-las

com o operador booleano. Com isso obtém-se resultados, que realmente está procurando.

6:

Através dos resultados da pesquisa realizada com operador booleano -, o usuário encontra o que está pesquisando, curso no exterior. Decide então a usar operadores booleano and, e com isso refinará mais as suas buscas

7:

O objetivo final do usuário foi encontrado, porém ele teve que passar por várias etapas e se não conhecesse bem os operadores booleanos e técnicas especiais não teria sucesso em sua busca.

3.8.2 – Conclusões

Essa pesquisa foi realizada, para fins de comparações de eficiência em máquinas de busca.

GOOGLE

Reconhece as aspas duplas " ", fornecendo assim a frase exata. O operador booleano + é interpretado como se fosse: a palavra tem que aparecer obrigatoriamente em meus resultados. O operador booleano – é interpretado como se fosse: retire do meu resultado todas as páginas que contenham essa palavra. O operador **AND** é interpretado como a junção das palavras. O parêntese () tem o papel das aspas dupla.



Figura 22 – Interface da máquina de busca Google

YAHOO

Idem ao Google, porém quando se efetiva as buscas, 1, 2 e 3, o Yahoo lhe apresenta um pequeno índice para adiantar a sua busca. A figura 23 mostra esse índice está.

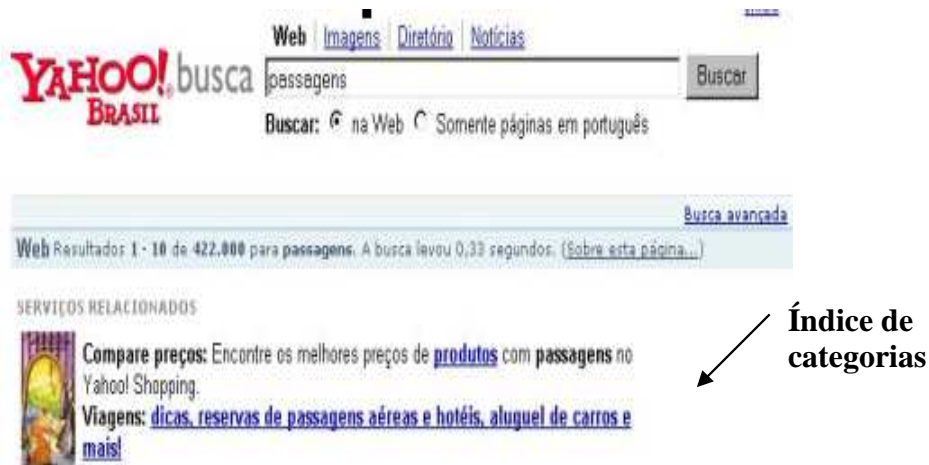


Figura 23 – Interface da máquina de busca Yahoo

RADAR UOL

No Radar Uol é um pouco diferente, mas é bem mais fácil de utilizá-lo, como mostra a figura 24. O usuário não precisa ter conhecimento de operadores nem de técnicas especiais.



Figura 24 – Interface da máquina de busca Radar Uol

TODO BR

No Todo Br a busca não obteve sucesso, pois esta máquina é voltada somente para a *Web* brasileira. O Todo BR também não reconhece o operador booleano – , interpreta-o como se fosse o operador +.



Figura 25 – Interface da máquina de busca Todo BR

ALTA VISTA

Idem Google



Figura 26 – Interface da máquina de busca Alta Vista

OBS.:

A máquinas de busca Google e Radar Uol foram às únicas a encontrarem resultados na busca de nº 7 que é ***“Promoção de passagens aéreas para curso no exterior”***.

CAPÍTULO 4 – TIPOS DE SISTEMAS DE BUSCA

Um sistema de busca é um conjunto organizado constituído de computadores, índices, bases de dados e algoritmos tudo isso reunido com a missão de:

- Analisar e indexar as páginas da *Web*;
- Armazenar os resultados dessa análise e indexação numa base de dados;
- Através de uma consulta de usuário, o sistema de busca vai pesquisar a sua base de dados e fornecer os resultados da pesquisa ao usuário.

Os Sistemas de buscas podem ser divididas em três categorias:

- Diretórios;
- Mecanismo de busca;
- Metabuscadores.

4.1 – Diretórios

São os sistemas de busca nos quais a indexação das páginas da *Web* é realizada por humanos.

Um diretório tem dois componentes principais:

- Uma base de dados, também chamada de índice ou catálogo;
- Um programa de computador que faz a pesquisa na base de dados.

“Os diretórios geralmente têm uma base de dados de menor tamanho que os mecanismos de busca. Mas isso não significa que eles produzam, necessariamente, resultados inferiores. Muito pelo contrário. Devido ao modo de criação de sua base de dados, geralmente eles têm um índice de relevância bem maior.” [11]. Os diretórios de modo geral, são menos eficazes, mas podem ser úteis quando se procura algo bem específico. Eles são capazes de cadastrar todo site e não páginas separadas.

“Os diretórios foram os primeiros sistemas propostos para organizar e localizar as informações na *Web*, vindo a preceder os sistemas de busca por palavras-chave” [5].

O funcionamento dos diretórios pode ser explicado como:

- O *Webmaster* ou administrador do site solicita sua inclusão, podendo sugerir um título e um texto descritivo e ainda uma categoria na qual o site deve fazer parte;
- Um editor ou membro da equipe editorial analisará o conteúdo do site verificando se existem erros nos códigos das páginas, *links* quebrados ou qualquer coisa que impossibilite um usuário de visualizar a página corretamente;
- O editor acata as sugestões de texto descritivo ou sugere um apropriado para o site;
- O site é cadastrado em uma categoria que mais se enquadra ao conteúdo existente;

“O descobrimento e a seleção das informações são realizados em sua maioria por profissionais especializados, os editores que aplicam critérios de qualidade para avaliar se um site pode ser indexado ou não no diretório. Os editores descobrem novos

sites a partir de sugestões do usuário (cadastro do site pelo usuário), através de pesquisas na *Internet* como listas de anúncios de novas páginas, ou ainda, pelo uso de robôs que coletam novas URLs na *Web*". [5].

A análise dos editores tem por finalidade evitar que se faça a indexação de informações sem valor, lixos que jamais serão procurados por usuários.

4.1.1 – Como é formada a base de dados

O usuário é quem deve tomar a iniciativa. Sempre informar ao diretório o título, o URL, a descrição do conteúdo, a categoria a que o site pertence e mais algumas informações complementares.

Caso o diretório disponha de uma equipe de editores, eles irão até o site e farão uma "vistoria" nele. Uma vez aprovada a inclusão, o site é inscrito no índice do diretório.

Mas também existem os diretórios sem editores. Eles aceitam as informações do jeito que o usuário mandar e a arquivam lá no índice do diretório.

“A busca se baseará na descrição do texto definido. A grande vantagem de se usar os diretórios é que quando o site aparecer na listagem, o usuário visualizará o texto descritivo e terá uma clara noção do conteúdo existente.” [16]

4.2 – Exemplo de Sistema de busca - Diretório

Existem vários Diretórios, mas, será apresentado de acordo com a pesquisa I ¹, o mais utilizado que é o yahoo (www.yahoo.com.br).

4.2.1 – Yahoo

O nome Yahoo veio da hierárquica do Oracle, mas, Filo e Yang, insistiram em selecionar o nome porque eles gostavam da definição geral de yahoo: "rude, natural".

4.2.2 – História

Yahoo começou com um passatempo de dois estudantes e evoluiu em uma marca global. Os dois fundadores de Yahoo!, David Filo e Jerry Yang, começaram o guia no campus em fevereiro de 1994, isso para manter rastro dos interesses pessoais deles na *Internet*. Quando as categorias se tornaram muito grandes, eles desenvolveram subcategorias e nasceu o conceito do Yahoo! [16]

1 – Pesquisa I realizada, com o propósito de saber quais as máquinas de busca mais utilizadas, suas deficiências e dificuldades de usuários

Os números que aparecem entre parênteses ao lado das categorias do Yahoo representam o número de categorias sob aquele título. O número não inclui entradas em nenhuma categoria de referência cruzada. As categorias de referência cruzada são aquelas seguidas pelo símbolo "@". O símbolo "@" que aparece no final de uma categoria significa que este título está listado em diversos lugares na hierarquia do Yahoo!. [16]

4.2.3 – Características

Algumas sintaxes do Yahoo! que o usuário pode escolher e usar.

Palavras obrigatórias e proibidas

- Colocando + antes de uma palavra faz com que esta palavra apareça obrigatoriamente em todos os resultados da busca. Por exemplo: *Roberto +Carlos*, o resultado da busca apresentará todos os sites que contenham as palavras *Roberto Carlos*, não necessariamente nessa ordem;

- Colocando – antes de uma palavra faz com que a palavra não apareça em nenhum dos resultados da busca. Por exemplo: *casseta –planeta*; o resultado da busca apresentará todos os sites que contenham a palavra *casseta* e em hipótese nenhuma a palavra *planeta*;

Restrições a partes do documento

t: - restringe a busca somente aos títulos dos documentos. Por exemplo: *t:Deus*, o resultado da busca apresentará todos os sites que em seu título apareça à palavra *Deus*;

u: - restringe a busca as URLs dos documentos somente. Por exemplo: *u:telemig*, o resultado da busca apresentará os sites da *telemig*;

Busca pela frase exata (" "). Por exemplo: "*redes de computadores*", o resultado da busca apresentará todos os sites que contenham obrigatoriamente a frase *redes de computadores*;

4.2.4 – Como é feita a busca

Se o Yahoo! não encontra resultados para sua busca em sua base de dados, ele irá oferecer resultados do Google, uma máquina de busca que trabalha em parceria do Yahoo!. O Google é especializado em indexar toda e qualquer página *Web* que puder encontrar. Isso o confere um monte de dados brutos, como outras máquinas de busca. Dessa forma, ele pode prover bons resultados para buscas muito específicas e, muitas vezes, resultados fracos em buscas mais genéricas. O mais provável, porém, é que se os seus termos de busca não encontraram nada no Yahoo!, você está fazendo uma busca muito específica. Nesse caso, os resultados apresentados pelo Google devem ser úteis. [16]

A busca do Yahoo! examina três áreas principais: as categorias, os sites listados no Yahoo! e as páginas *Web* indexadas pelo Google. Os resultados aparecem em ordem do mais relevante para o menos relevante.

O Yahoo!Brasil primeiro encontra as ocorrências com a(s) palavra(s)-chave e depois ordena os resultados de acordo com seu grau de relevância dentro de cada área específica. Os resultados da seguinte forma:

- Múltiplas ocorrências de palavras-chave: documentos que tenham mais ocorrências da(s) palavra(s)-chave obterão uma posição melhor na classificação do que aqueles que tiverem menos ocorrências;
- Posicionamento no documento: documentos com ocorrência de palavras no título são listados antes daqueles com ocorrência de palavras no corpo do documento ou na URL;
- Generalidade da Categoria: categorias mais elevadas na hierarquia do Yahoo! (exemplo: categorias mais genéricas) são classificadas antes do que aquelas mais profundas na hierarquia (exemplo: categorias com foco mais restrito).

4.3 – Mecanismo de busca

São os sistemas de busca baseados no uso exclusivo de programas de computador para a indexação das páginas da *Web*. Os mecanismos de busca são aqueles que utilizam aplicativos conhecidos como aranhas (*softwares* que navegam automaticamente na *Web*). Tais aranhas recolhem toda a informação e vão indexando tudo num banco de dados. São bons quando se deseja uma pesquisa abrangente.

Eles possuem três componentes principais: Robô, base de dados, software.

4.3.1 – Robô

É um programa que "visita" os sites ou páginas armazenadas na *Web*. Ao chegar em cada site, o programa robô "pára" em cada página dele e cria uma cópia ou réplica do texto contido na página visitada e guarda essa cópia para si. Essa cópia ou réplica vai compor a sua base de dados.

Os robôs também conhecidos como spiders (aranhas) ou *Web crawlers* (rastejadores), são programas que percorrem a estrutura da *Web*, recolhendo informações por eles consideradas relevante sobre as páginas que encontram. Essas informações são indexadas em uma base de dados que será explorada posteriormente utilizando o mecanismo de busca.

Eles Utilizam o protocolo HTTP para recuperar documentos dos servidores. Eles são responsáveis por revisar cada site encontrado e guardar certos dados: endereço, o título da página, o texto que a compõe e as etiquetas com palavras chaves (meta *tags*), incluída por um *webmaster* para ajudar as aranhas a classificar o site.

“Os robôs ao visitarem uma página, primeiramente verificam se a mesma já foi visitada anteriormente ou se é uma página nova para ele. Caso, a página tenha sido recolhida (indexada), o robô verifica se ocorreu alguma modificação desde a ultima visita, e se ocorreu, atualiza a informação armazenada sobre a página na base de dados”. [8].

“Os robôs estão bastante mal documentados na literatura principalmente pelo fator comercial, onde os mecanismos tendem esconder a forma de trabalho do robô”. [8].

4.3.2 – Base de dados

Tudo o que um robô encontra vai para a segunda parte do mecanismo: a base de dados. Se uma página é atualizada, a base de dados também é atualizada com as novas informações.

A base de dados é constituída das cópias efetuadas pelo robô. “Os itens coletados pelos robôs durante o processo de rastreamento são encaminhados aos indexadores que extraem a informação das páginas e as armazenam na base de dados, às vezes também denominada índice ou catálogo” [5].

O tamanho das base de dados variam entre os mecanismos de busca, cada mecanismo de busca possui sua base de dados. Já o conteúdo das base de dados são diferentes devido ao modo de indexação que cada mecanismo de busca utiliza. Na base de dados podem ser encontrados endereços das páginas, títulos, cabeçalhos, resumos, tamanho, e as palavras contidas nos documentos.

“Os mecanismos de busca com a maior base de dados tendem a ser os mais populares. Porque o tamanho da base de dados é responsável pela amplitude da pesquisa, quanto mais documentos ou páginas estiverem armazenados na base de dados do sistema, mais itens o sistema recuperará”. [8].

“Algumas vezes pode ser que demore um tempo até que as novas páginas ou atualizações que um robô encontra sejam adicionadas à base de dados. Assim, uma página pode ter sido rastreada pelo robô, mas não ainda indexada. E até que ela seja indexada, ela não estará disponível para os usuários quando estes realizam uma busca”. [16].

4.3.3 – Software

Software de interface faz a interação do usuário com o mecanismo de busca. É através dele que o usuário fornece sua(s) necessidades de informação para ser(em) pesquisada(s) na base de dados pelo *software* de busca.

Software de busca é um programa que verifica nas milhões de páginas gravadas na base de dados se existem compatibilidades entre o conteúdo e o que está sendo buscado. Com isso, ele monta a listagem de sites na ordem que ele acredita ser mais relevante.

“Portanto, é um programa de busca propriamente dito. Esse programa de busca é acionado cada vez que alguém realiza uma pesquisa. Nesse instante, o programa sai percorrendo a base de dados do mecanismo em busca dos endereços - os URL - das páginas que contém as palavras, expressões ou frases informadas na consulta. Em seguida, os endereços encontrados são apresentados ao usuário”.[11]

4.3.4 – Como é formada a base de dados

Para criar a base de dados de um mecanismo de busca, o programa robô sai visitando os sites da *Web*. Ao passar pelas páginas de cada site, o robô anota os URL existentes nelas para depois ir visitar cada uma dessas URL. Visitar as páginas, fazer

as cópias e repetir a mesma operação: cópia e armazenamento, na base de dados, do que o robô encontrar nesses sites.

A outra maneira de o mecanismo de busca encontrar os sites na *Web* é o "dono" do site informar, ao mecanismo de busca, qual o endereço, o URL, do site. Todos os mecanismos de buscas têm um quadro reservado para o cadastramento.

4.4 – Exemplos de Mecanismo de busca

Existem vários mecanismos de busca, mas, será apresentado de acordo com a pesquisa ¹ o mais utilizado que é o google (www.google.com.br).

4.4.1 – Google

Google é um trocadilho com a palavra 'googol', que foi inventada por Milton Sirota, sobrinho do matemático americano Edward Kasner, para designar o número representado por 1 seguido de 100 zeros. O uso do termo Google reflete a missão da empresa de organizar o enorme montante de informações disponíveis na *Web* e no mundo. [10]

4.4.2 – História

O Google foi fundado por Larry Page e Sergey Brin, dois estudantes Ph.D de Stanford em 1998. A companhia privada anunciou, em junho de 1999, ter assegurado \$25 milhões em consolidação de dívida flutuante de patrimônio líquido. Seus sócio incluem Kleiner Perkins Caufield & Byers e Sequoia Capital. Google presta serviços através de seu próprio site público, www.google.com. A companhia também oferece soluções para busca na rede, em associação com provedores de conteúdo. [10].

4.4.3 – Características

Para inserir uma consulta no Google, basta digitar algumas palavras descritivas e pressionar a tecla 'enter' (ou clicar no botão Pesquisa Google) para obter uma lista com os resultados relevantes.

O Google pesquisa apenas páginas que correspondam exatamente aos termos da pesquisa, assim você pode tentar usar versões diferentes de termos de pesquisa. Por exemplo, se uma pesquisa para "hotel em Florianópolis" não apresentou o que você estava procurando, tente usar "hotéis em Florianópolis ". Ou pode-se tentar reformular a consulta. Por exemplo, pesquisas sobre "passagens de avião econômicas" e "passagens aéreas econômicas" retornam conjuntos de resultados diferentes.

Outra característica do google é que ele não diferencia letras maiúsculas de minúsculas, nem palavras com acento e sem acento, ele sempre acha a palavra exata.

Determinados caracteres servem como conectores de frases. O Google reconhece hífens, barras, pontos finais, sinais de igual e apóstrofes como conectores de frases que funcionam como aspas. Por exemplo: o substantivo guarda-chuva será tratado como uma frase, mesmo se as duas palavras não estiverem entre aspas.

Às vezes é útil escolher determinadas palavras ou frases para excluir de uma pesquisa. Você deseja todos os resultados relevantes, exceto os que contenham uma determinada palavra ou frase. O Google suporta essa funcionalidade "não" com o sinal de menos ("-"). Use esse sinal para excluir propositadamente um termo de sua pesquisa. Certifique-se de incluir um espaço antes do sinal de menos e o Google irá ignorar todas as páginas contendo essa palavra.

Algumas palavras, quando seguidas por dois pontos, têm significados especiais para o Google. Atualmente, o Google suporta um operador especial desse tipo.

4.4.4 – Como é feita a busca

Toda o sistema do Google é baseada na avançada tecnologia PageRank que classifica, num verdadeiro *ranking*, todos os sites da *Web* de acordo com a sua importância. “O PageRank utiliza a vasta rede de *links* da *Internet* como ferramenta organizacional. Basicamente, o sistema interpreta que um *link* de uma página X para uma página Y é um voto da página X para a página Y. E, além da quantidade de votos, também considera o "peso" de cada voto, de modo que quanto maior for a "importância" (ou classificação) da página X no PageRank, mais forte será o voto concebido para a página Y. Em suma, a importância de uma página é dada pela soma da importância daquelas páginas que a referenciam”. [10]

“Toda essa apuração de votos é calculada por uma equação de 500 milhões de variáveis e mais de 2 bilhões de termos. Desta forma, quando é feita uma consulta, o Google vai buscar as páginas que satisfazem a sua pesquisa e depois organiza e exhibe o resultado baseado na classificação do PageRank, o que eleva bastante a probabilidade de que o resultado que você espera apareça entre os primeiros” [10]. Por se basear na natureza excepcionalmente democrática da *Web*, utilizando a inteligência coletiva de milhões de internautas, e por utilizar sofisticados métodos de automação, o Google impede a interferência humana e a manipulação dos resultados, tornando-se um meio honesto, objetivo e imparcial de encontrar informações na *Web*. Para garantir que o seu banco de dados mantenha-se sempre atualizados, o Google dispõe de uma espécie de robô - o GoogleBot - que, periodicamente, varre toda a *Internet*, navegando por todos os *links* para eliminar as repetições, excluir os *links* quebrados e atualizar a classificação do PageRank. [10]

O princípio básico que o Google segue é encontrar as páginas que contenham todas as palavras que você colocou na sua busca. A ordem das palavras interfere no resultado, por isso aconselha-se sempre colocar do mais genérico ao mais específico.

“Os métodos complexos e automáticos de procura do Google impedem interferência humana. Diferente de outros serviços de busca, Google é estruturado, assim ninguém pode obter uma listagem maior ou resultados comercialmente alterados. Uma busca do Google é um modo honesto e objetivo para encontrar *websites* de alta qualidade, facilmente”. [10]

4.5 – Metabuscadores

“Os sistemas de metabusca também chamados de multibuscadores são sistemas que localizam a informação em outros sistemas de busca (mecanismos e diretórios) simultaneamente e combinam os resultados encontrados em uma só lista de resultados” [4].

São sites que possuem um sistema que procura um termo em diversos mecanismo de busca ao mesmo tempo, aumentando a chance de o usuário encontrar a palavra procurada, mas fornecendo muitas vezes um número excessivo de páginas.

“Estes sistemas não utilizam robôs para encontrar as páginas na *Web* e também nem o usuário necessita cadastrar o seu site, pois, os sistemas de metabuscadores não possuem uma base de dados, próprio, uma vez que, utilizam os dados de outros sistemas de busca”. [8].

4.6 – Exemplo de Metabuscador

Existem vários metabuscadores, mas, será apresentado de acordo com a pesquisa¹ o mais utilizado que é o Radar Uol (www.radaruol.com.br).

4.6.1 – Radar Uol

4.6.2 – História

O radar Uol foi criado para facilitar as busca na *Web*. Depois de vários estudos sobre o funcionamento de máquinas de busca o Radar Uol inovou e invadiu o mercado on-line para entrar na disputa para recuperar informações na *Web*. Com essa nova técnica (multibusca) as pesquisa se tornaram mais atrativas.

4.6.3. – Característica

O radar Uol utiliza tanto a base de dados dos diretórios quanto à dos mecanismos de busca.

A barra Uol possibilita você criar sua própria Multibusca, usando os sites que quiser. Cada tipo de busca (por livros, CDs ou na *Web*, por exemplo) utiliza sites específicos, que podem ser definidos pelo usuário. O usuário pode definir quanto tempo quer esperar pelos resultados das buscas: 10, 30, 60 ou 90 segundos. [15]

Durante a busca, o botão "Interromper busca" fica disponível para que você cancele sua busca e veja os resultados recebidos até aquele momento.

Com a barra Uol o usuário pode escolher quais as máquinas de busca deseja procurar sua pesquisa, selecionando-as assim. Isso ajuda aos usuários a encontrarem o que desejam.

4.6.4 - Como é formada a base de dados

O Radar Uol não possui base de dados, mais encontra tudo o que se procura, pois ele utiliza a base de dados de vários mecanismos de busca e diretórios.

4.7 - Comparação entre as categorias de busca

4.7.1 - Diretórios

As vantagens dos diretórios é que eles são mais fáceis de serem utilizados, principalmente para usuários leigos, por possuírem categorias específicas. Eles também permitem ter uma visão geral do volume e conteúdo do índice, muitos diretórios indicam em cada um dos seus grupos quantas referências e subcategorias há nela. Outra vantagem é que as informações disponíveis passaram por um processo de seleção de qualidade e com isso os resultados de uma pesquisa são mais precisos.

As desvantagens dos diretórios são que eles possuem uma pequena cobertura da *Web*, ou seja, poucas páginas indexadas na sua base de dados, a única exceção é o Yahoo e outra desvantagem é que muitas informações indexadas nos Diretórios estão tornando-se rapidamente desatualizados, pois, não há nenhum mecanismo automático que faça as suas atualizações. Atualizar manualmente as informações indexadas torna-se uma tarefa impossível.

4.7.2 - Mecanismos de Busca

As vantagens dos mecanismos de busca são que eles permitem pesquisas amplas, possuem informações atualizadas. Outra vantagem é que existem mecanismos especializados praticamente para todas as áreas de conhecimento.

As desvantagens dos mecanismos de busca são que a cada mecanismo tem a sua própria sintaxe para “expressar” a consulta, o que representa uma das grandes dificuldades para o usuário. Outra desvantagem é que eles retornam resultados pouco precisos, sendo que, as informações indexadas não passaram por um processo de qualidade.

4.7.3 - Metabuscadores

As vantagens dos metabuscadores são que eles realizam buscas em vários sistemas ao mesmo tempo e acabam tendo uma cobertura bem maior da *Web* e eles também possibilitam ao usuário escolher em quais sistemas de busca o sistema deverá efetuar a consulta.

A desvantagem dos metabuscadores é a sua limitação em relação à interface que não permite utilizar os recursos específicos de cada sistemas de busca, ou seja, o

usuário não pode refinar a consulta. Outra desvantagem encontrada é relativa aos resultados obtidos, obtém-se uma maior cobertura sem um aumento de qualidade.

4.7.4 – Resultados

Através dessa pequena comparação, pode-se perceber que os sistemas de busca têm muito que evoluir, pois existem vantagens e desvantagens de cada sistema.

Para concluir qual sistema é o melhor, é uma tarefa que necessitaria um estudo mais profundo, porém os que mais se aproximam de uma realidade mais promissora seria os mecanismos de busca, por utilizarem robôs e possuírem uma base de dados relativamente maior do que os demais.

O grande problema dos mecanismos de busca é a filtragem de informação, que será resolvida utilizando uma nova técnica a partir de agente inteligente que será proposta no capítulo seguinte.

CAPÍTULO 5 – INTRODUÇÃO A AGENTES INTELIGENTES

“A idéia de agentes é executar tarefas baseadas em computador. As pesquisas feitas na área tiveram como objetivo obter uma comunicação de alto nível entre o homem e a máquina” [9].

Neste capítulo será apresentado o conceito de agentes inteligentes, suas características, comunicação entre agentes, possíveis aplicações, tipos de agentes, questões básicas de um agente e será proposto um agente inteligente para a filtragem de informações em máquinas de busca.

5.1 – O que são agentes inteligentes

“Um agente inteligente é um *software* que ajuda as pessoas, agindo em seu lugar” [9]. Um agente inteligente é um programa informático que apresenta um comportamento racional geralmente associado ao ser humano, que envolve a análise da informação que possui e obtém com o objetivo de atingir novas conclusões com base nas quais pode fundamentar as suas ações.

O agente inteligente é uma entidade criada por um programa de software. Está embutido num ambiente com o qual interage recolhendo informação e tomando decisões.

“Agentes são qualquer entidade computacional que pode ser percebida em um ambiente através de sensores e atuar sobre este ambiente produzindo efeitos. A estes Agentes pode-se adicionar comportamento racional, que junto com o comportamento autônomo passam a ser considerados Agentes inteligentes. Tecnologias como lógica Fuzzy, Computação Evolutiva e Redes Neurais são algumas das técnicas utilizadas para este fim” [9].

5.2 – Como são utilizados

“Os Agentes podem ser utilizados de diferentes maneiras reduzindo assim o trabalho e a sobrecarga de informações. Eles deixam a complexidade e a dificuldade da tarefa transparente ao usuário, executam tarefas no lugar do usuário, podem treinar ou ensinar o usuário assim como monitorar procedimentos e eventos” [9].

O Agente pode ser projetado para desempenhar qualquer tarefa que lhe seja delegada. Isso faz com que eles ajudem no trabalho de encontrar e filtrar informações, seja numa base de dado local, numa grande corporação ou em incontáveis pontos distribuídos pelo mundo.

“De uma forma geral, a função de um agente inteligente é resolver problemas do mundo real, realizar tarefas que seres humanos são capazes de resolver ou realizar melhor que os sistemas computacionais existentes atualmente” [1].

5.2.1 – Possíveis aplicações

Assistentes pessoais: Sistemas completos e independentes que planejam, organizam e otimizam tarefas de produtividade pessoais. Aplicações possíveis incluem filtragem de correio eletrônico e sistemas de classificação, programadores automáticos de reuniões, além de diversos filtros de informação.

Sistema para atender visitantes: O atendimento e programação de atividades, reuniões e eventos com a participação dos visitantes no local são organizados e coordenados através da operação de agentes locais com os assistentes pessoais dos visitantes.

Exploração de informação (*Data Mining*): Agentes específicos em informação fornecem um contexto para busca de dados em grandes bancos de dados ou outras fontes de informação, que em cooperação com agentes pessoais, selecionarão a informação útil. Este campo é um dos que estão em mais rápido desenvolvimento no momento, pela explosão da necessidade de buscar informações através de redes de computadores.

Gerenciamento de redes: Agentes colaborativos coletam e trocam informação local sobre as estatísticas da rede para alcançar automação e otimização das decisões nas tarefas de administração de rede tais como roteamento, acesso, provisão de serviços, monitoramento e avaliação estatística, dentro de uma visão global.

Filtragem de Notícias: Ajuda ao usuário a selecionar artigos de uma fonte contínua de informações. À medida que mais e mais informações se tornam disponíveis na rede, o usuário se torna desesperado por ferramentas para ajudá-lo na filtragem das informações e para encontrar artigos do seu interesse. NewT é um sistema que ajuda o usuário a filtrar notícias. Através dele o usuário pode criar um ou mais agentes de notícias e treiná-los através de exemplos de artigos que pode ou não ser selecionados. Um usuário pode criar por exemplo, agente para informações de negócios, um para notícias policiais, um outro para informações sobre computadores e um outro para notícias esportivas. Um agente é inicializado fornecendo para ele exemplos positivos e

negativos de artigos para serem recuperados. O usuário pode programar o agente explicitamente fornecendo templates de artigos que devem ser recuperados.

5.3 – Características

Um agente pode ser definido a partir de suas características básicas. Um agente não precisa possuir todas estas características ou atributos, embora suas capacidades estejam diretamente associadas a presença delas. Segue algumas das características existentes, as mais relevantes.

5.3.1 – Autonomia

“Autonomia refere-se ao princípio de que os agentes podem agir baseados em seus próprios princípios, sem a necessidade de serem guiados por humanos” [6]. Os agentes possuem estados e metas internos, agindo de maneira a atingir estas metas em favor de seus usuários. O elemento chave da autonomia é a pró-atividade (ver cap 5.3.8), que é a sua habilidade de tomar iniciativas, sem a necessidade de agir em virtude de uma mudança de seu ambiente [6]

5.3.2 – Mobilidade

A capacidade de poder se mover através de uma rede de computadores parece ser interessante para agentes que auxiliam seus usuários na busca de informações, principalmente dentro da *Internet*. Entretanto, este atributo pode causar sérios problemas de sobrecarga na rede, uma vez que eles trafegam entre as máquinas conectadas. Outro problema que está relacionado com segurança é que um agente móvel pode conter problemas de código ou até mesmo estar transportando um vírus de computador. Sendo assim, a implantação de agentes móveis deve ser acompanhada de processos de autorização, além da garantia de que a memória e os recursos da máquina estarão protegidos [13].

5.3.3 – Cooperação

“Cooperação pode ser entendida como a capacidade que os agentes tem de trabalharem em conjunto de forma a concluírem tarefas de interesse comum. Acredita-se que a cooperação entre agentes é fundamental, sendo a razão principal para a existência de um ambiente multi-agente” [19]. Para permitir esta cooperação, o agente deve ser dotado de uma certa habilidade social, capacitando-o a interagir com outros agentes e possivelmente humanos através de alguma linguagem de comunicação [19].

5.3.4 – Comunicabilidade

Quando existe mais de um agente envolvido, há uma necessidade óbvia por um modelo de comunicação. Entretanto, o conceito de comunicabilidade não estabelece apenas a troca de informações entre agentes.

“Agentes podem se comunicar com outras entidades além de agentes, incluindo-se humanos e o seu ambiente” [6].

Esta característica será apresentada mais detlhada no capítulo 5.4.

5.3.5 – Aprendizagem

Um dos atributos que mais caracterizam agentes inteligentes é a capacidade de aprender. Uma real autonomia só pode estar presente quando um agente possui a habilidade de avaliar as variações de seu ambiente externo e escolher qual a ação mais correta. Entretanto, mesmo quando um agente não reconhece nenhuma ação a ser executada, é esperado que ele procure encontrar uma saída. A questão não é acertar sempre, mas aprender continuamente por experiência, seja através de sucessos ou de fracassos.

“O aprendizado pode ser também um processo iterativo. Nestes casos, o treinador pode fornecer ao agente o conhecimento através de uma seqüência de

instruções ou informá-lo apenas quando o agente não possuir o conhecimento necessário” [1].

5.3.6 – Reatividade

Reatividade é a habilidade que um agente tem de reagir a mudanças no seu ambiente. Para tal, o agente deve ser capaz de perceber seu ambiente e atuar sobre ele. Este atributo está presente em praticamente todas as definições de agente.

5.3.7 – Habilidade Social

“Esta característica está diretamente associada com a característica de comunicabilidade, uma vez que representa a habilidade de interagir com outros agentes. Isso indica a necessidade de uma linguagem comum para a comunicação de agentes” [1].

5.3.8 – Pró-atividade

Esta característica pode ser também denominada iniciativa, uma vez que representa um comportamento independente. As ações são dirigidas pelo objetivo e não simplesmente por mudanças no seu ambiente. O agente que implementa esta característica possui maior flexibilidade, pois é capaz de resolver problemas causados por situações inesperadas.

5.4 – Comunicações entre agentes

Um agente teoricamente pode ser programado em qualquer linguagem convencional e utilizando-se qualquer linguagem particular, entretanto, a adoção de uma linguagem orientada à objetos ou uma linguagem apropriada para a programação de agentes e uma linguagem de comunicação padrão, certamente garantirão uma maior produtividade e interoperabilidade ao produto final.

“Muitas das informações que os agentes devem compartilhar devem ser usáveis em ambas as direções, por exemplo, ser necessário computar uma quantidade a vezes uma quantidade b em um momento e precisar computar a quantidade b vezes a em outro momento”. [1]

“Os sistemas modernos de computação freqüentemente envolvem múltiplos computadores interagindo de forma distribuída” [6]. Neste sentido, a habilidade de comunicação é um importante atributo que os agentes inteligentes devem possuir. ambientes que possuem mais de um agente, praticamente exigem o intercâmbio de

informações. Neste sentido, torna-se claro a necessidade de uma linguagem de comunicação comum.

“A linguagem de comunicação entre agentes utiliza normalmente uma das seguintes abordagens:

- **Procedural:** a comunicação acontece através de diretivas. Tanto comandos individuais quanto programas completos podem ser transmitidos e executados no lado receptor. Exemplos, Linguagens baseadas em *scripts*, tais como TCL e Telescript.
- **Declarativa:** a comunicação ocorre através da troca de estruturas declarativas, tais como definições, asserções e outras. Exemplos: linguagem KIF e KQML” [1].

5.5 – Tipos de agentes

Serão apresentados 4 (quatro) tipos de agentes, sendo que o de recuperação de informação é o propósito maior para esse trabalho.

5.5.1 – Agentes Colaborativos

Em sistemas colaborativos, cada agente contribui com sua própria técnica inteligente para a solução de um problema complexo. Agentes colaborativos enfatizam autonomia e cooperação com outros agentes de forma a executar tarefas para seus donos. Neste ambiente, torna-se clara a necessidade de negociação para estabelecer acordos e comprometimentos mútuos. Apesar de aprendizado não ser a principal ênfase da operação de agentes colaborativos, eles podem demonstrar um aprendizado limitado. Para que um processo de colaboração possa acontecer, é clara a necessidade de se definir uma linguagem comum para a comunicação entre agentes.

5.5.2 – Agentes de Interface

Os agentes de interface podem ser conhecidos também como sendo agentes que podem aprender ou assistentes pessoais.

Agentes de interface suportam e providenciam assistência, tipicamente para o usuário aprender a usar uma aplicação em particular, como um sistema operacional, por exemplo. O agente observa o usuário e monitora suas atividades na interface, aprendendo maneiras novas de executar tarefas, e sugerindo maneiras melhores de executá-las. Os agentes de interface aprendem para oferecerem um melhor auxílio aos seus usuários.

5.5.3 – Agentes Assistentes

Agentes assistentes são agentes pessoais inteligentes. Apesar de possuírem semelhanças com os agentes de interface, os assistentes adicionam facilidades de ajuda, diagnóstico e orientação, além de poderem executar tarefas autonomamente, sem a intervenção humana.

Um assistente deve ser capaz de identificar situações onde o usuário pode precisar de ajuda e, então, fornecer alguma informação adicional para auxiliá-lo. Outra facilidade é a habilidade de representar seu usuário de forma autônoma, utilizando uma base de crenças criada a partir dos interesses de seu usuário. Desta forma, humanos poderão participar de um processo cooperativo envolvendo outros humanos e agentes.

5.5.4 – Agentes de Recuperação de Informação

Estes agentes são capazes de buscar informação de uma forma inteligente. É importante notar que estes agentes não devem ser confundidos com simples mecanismos de busca utilizados na *Internet*. O objetivo não é simplesmente encontrar informações que satisfaçam um conjunto de palavras-chave, mas espera-se que este tipo de agente possa reconhecer padrões de informação e encontrar aquelas mais relevantes. Além disso, este agente deve poder operar em modo autônomo, realizando filtragens e em alguns casos aplicando inferências. Neste caso, o agente consegue transformar pedaços de informação em conhecimento altamente produtivo para seu usuário. Estes agentes tem ampla aplicação em organizações que possuem grande volume de informação espalhada geograficamente ou em vários bancos de dados.

5.6 – Questões básicas de um agente

Os agentes são projetados para resolver uma infinidade de problemas em uma infinidade de ambientes, resultando em diversas implementações. Não importando qual combinação de ambientes e problemas, toda implementação de agentes deverá lidar com três questões básicas que são resolução de problemas, representação do conhecimento e aprendizado.

5.6.1 – Resolução de problemas

Os agentes inteligentes devem resolver problemas do mundo real. esses agentes aplicam alguma técnica de raciocínio sobre o conhecimento e as percepções sensoriais para produzir respostas úteis. Essas técnicas são as de técnicas de resolução de problemas.

A técnica de resolução de problemas é fator decisivo no projeto de agentes inteligentes por dois motivos principalmente. Primeiro, ela restringe os tipos de problemas que é possível resolver, ou seja, está relacionada aos limites teóricos e uma determinada implementação. Segundo, ela define a eficiência com que os problemas serão resolvidos, ou seja, está relacionada às limitações práticas dos recursos computacionais (neste sentido prático, a resolução de problemas está intimamente ligada à teoria da complexidade de algoritmos).

5.6.2 – Representação do conhecimento

Para que o conhecimento possa ser armazenado é essencial que se possa representá-lo. Representação do conhecimento pode ser definida como um conjunto de convenções sintáticas e semânticas que torna possível descrever coisas.

Para operar satisfatoriamente, um agente inteligente precisa ter algum conhecimento sobre o domínio do problema que ele irá tratar e sobre o ambiente o qual ele está inserido. Esse conhecimento, qualquer que seja seu conteúdo, precisa estar disponível ao agente de uma forma que possa ser usado pelo mecanismo de resolução de problemas, assim como nossos cérebros armazenam informações que usamos em nossa atividade cognitiva.

5.6.3 – Aprendizado

A aprendizagem é a aquisição de conhecimento através da experiência. A descoberta de representações poderosas, eficientes e robustas para problemas não triviais é um grande desafio. É justamente aí que a aprendizagem desempenha seu maior papel. Ela é necessária porque nem sempre é possível embutir previamente no agente todo o conhecimento necessário para seu pleno funcionamento por tempo indeterminado.

A capacidade de aprendizagem de um agente está intimamente ligada à sua autonomia. Sem ela, o agente é incapaz de se adaptar automaticamente às mudanças

do ambiente, não exibindo, portanto, qualquer autonomia. Portanto, se queremos que nossos agentes operem adequadamente, e de maneira cada vez melhor, através de ambientes dinâmicos, sem intervenção humana, precisamos dotá-los de mecanismos de aprendizagem.

5.7 – Proposta de um agente inteligente para a filtragem de informações em máquinas de busca – FILVE2004

O FILVE2004 será proposto de uma forma dissertativa com o objetivo de melhorar as pesquisas realizadas em máquinas de busca. Hoje em dia, devido ao número de informações presentes na *Web*, se torna uma tarefa difícil encontrar informações realmente interessantes.

Ele será responsável pela filtragem de informações encontradas nas máquinas de busca, o que realmente for interessante, somente isso será mostrado ao usuário.

5.7.1 – Compreendendo as necessidades dos usuários

Devido o resultado obtido na pesquisa I realizada no capítulo 3.7, pode-se perceber que a maior dificuldade de usuários é encontrar o que realmente desejam

diante de tantas URLs. Foi para minimizar este descontentamento que o FILVE2004 foi proposto.

5.7.2 – Motivação

“Desde os últimos três anos, em cada seis meses, a informação disponível na *Internet* duplica e, estima-se que este ritmo de crescimento aumente nos próximos anos. O NECRI (*NEC Research Institute*) efetuou um estudo na *Internet* que relata que o número de páginas existentes na *Web* aumenta de uma forma tal que se estima que até que uma nova página esteja pesquisável num motor de procura são necessários seis meses. Este mesmo instituto afirma ainda que na *Web* existem mais de 800 milhões de páginas pesquisáveis e que mesmo os melhores motores de procura só percorrem um sexto das mesmas” [12]. Assim, registra-se uma forte necessidade de qualificar a informação na *Web*, devido aos aspectos acima referidos, que existe urgência em criar meios de pesquisa de informação que permitam obter a informação desejada.

O FILVE2004 surgiu como um novo paradigma que permite a filtragem de tantas informações.

A idéia principal do FILVE2004 é filtrar informações relevantes para o usuários, cada usuário possui um perfil e o FILVE2004 será capaz de distingui-los e atuar de acordo com cada perfil.

5.7.3 – Características

O FILVE2004 deverá ter:

- **Autonomia** (Ver capítulo 5.3.1). O FILVE2004 deve ser capaz de tomar iniciativas em prol do usuário, sem depender de ser guiado. Exemplo, quando o usuário digitar em sua pesquisa uma palavra, o FILVE2004, deve ser capaz de comparar as palavras digitadas pelo usuário, ver se possui alguma palavra semelhante em sua base de dados, se não possuir, deve ter a iniciativa de sair procurando em outros lugares (outras bases de dados, outros agentes).
- **Cooperação** (Ver capítulo 5.3.3). O FILVE2004 deve ser capaz de se interagir com outros agentes, através de uma linguagem específica, tornando um ambiente mais social. Exemplo, quando o FILVE2004 não encontra a pesquisa digitada pelo o usuário, ele procura se interagir com outros agentes, para trocarem informações e conseguir o que realmente esta precisando.
- **Comunicabilidade** (Ver capítulo 5.3.4). O FILVE2004 deve ser capaz de comunicar-se com outros agentes, trocando informações e com isso adquirindo conhecimento. Exemplo, quando o FILVE2004 não encontra em sua base de dados o que o usuário esta procurando além de decidir qual a melhor opção deve executar e interagir com outros agentes, ele também se comunica com outros agentes para ver se algum deles possui as palavras digitadas pelo usuário, com isso ele se comunica com vários agentes e adquire mais conhecimento.
- **Aprendizagem** (Ver capítulo 5.3.5). O FILVE2004 deve ser capaz de estar sempre aprendendo. Através de seu ambiente, terá habilidade de perceber mudanças, de encontrar resultados e se interagir com os usuários. Exemplo, diante de uma nova palavra digitada pelo usuário o FILVE2004, entrará em

processo de aprendizagem, pois ele comparará em sua base de dados, caso não ache as palavras, irá se comunicar, interagir e trocar informações com outros agentes e ambiente diferentes, com isso estará em processo de aprendizagem.

5.7.4 – Objetivos do FILVE2204

A informação produzida nos dias de hoje é cada vez maior. Agora, mais do que nunca, é necessário filtrar informações para que elas se tornem realmente relevantes.

Este agente tem como os principais objetivos:

- Verificar o perfil do usuário;
- Verificar a(s) área(s) de interesse;
- Atribuir valores de relevância a páginas pesquisadas;
- Filtrar informações

Verificando o perfil do usuário

Quando um usuário efetivar uma pesquisa, o FILVE2004 estará observando-o e adquirindo assim conhecimento, com isso ele classificará o perfil do usuário. Caso ele já tenha outro perfil em sua base de dados ele comparará e verá qual o perfil que melhor se adapta com o usuário, além de pesquisar em sua base de dados, o FILVE2004 deverá trocar informações com outros agentes, tornando a escolha do perfil

mais adequada. Portanto o FILVE2004 aprende observando o perfil do usuário, através das pesquisas realizadas pelos usuários, ele determinará um perfil.

Verificando áreas de interesse

Os perfis de usuários serão implementados anteriormente, por isso o FILVE2004 irá observar o comportamento dos usuários e através das pesquisas realizadas, classificá-los de acordo com cada perfil e também irá selecionar de acordo com cada perfil as áreas de interesse do usuário. Portanto com as pesquisas realizadas, o FILVE2004 atribuirá ao perfil do usuário as áreas de maior interesse.

Atribuindo valores de relevância

Esta é a maior tarefa do FILVE2004, porque terá que classificar o tempo que o usuário permaneceu na página, observar se foi copiado algo ou se imprimiu. Depois de armazenar esses dados atribuir um grau de relevância para a página de acordo com a sua classificação². Portanto, O FILVE2004 dará um grau mais alto de relevância, naquela página que o usuário permaneceu mais tempo e que se observou interesse pelo conteúdo da página. O FILVE2004 irá observar cuidadosamente todos os aspectos dos usuários.

Filtrando informações

De acordo com o perfil e da área de interesse dos usuários, o FILVE2004 deverá filtrar as pesquisas utilizando o grau de relevância que foi atribuído por ele através de observações e aprendizado. Portanto para que haja filtragem nas pesquisas o FILVE2004 tem que ter o perfil do usuário, isto é, sua área de interesse e com isso poderá fornecer informações relevantes.

2 – Classificação – Classificar um objeto é determinar com que grupo de entidades, já classificadas

anteriormente, esse objeto apresenta mais semelhança. [16]

5.7.5 – Funcionamento

O FILVE2004 funcionará da seguinte maneira:

Observando e aprendendo com seu ambiente; Ele observará diversos usuários utilizando máquinas de busca e através desse processo, classificará perfis para usuários. Porém isso só ocorrerá depois que o usuário realizar cerca de 20 pesquisas, pois o FILVE2004 estará observando o perfil do usuário para classificá-lo.

Com os perfis pré estabelecido, quando um usuário pesquisar algo FILVE2004 irá construir sua base de dados de acordo com as necessidades de cada usuário, perfil, área de interesse e grau de relevância.

Portanto, através do conjunto desses dados, o FILVE2004 será capaz de filtrar informações e associá-las de acordo com a necessidade de cada usuário.

5.7.5.1 – Exemplo de funcionamento do FILVE2004

O usuário digitará: www.buscainteligente.com.br

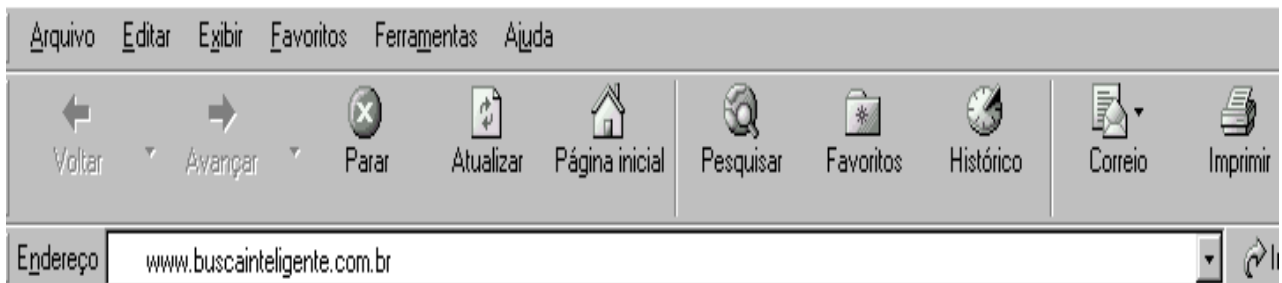


Figura 27 – *Internet explorer* - entrando no site de busca inteligente

Aparecerá a interface da máquina de busca



Figura 28 – *interface* da máquina de busca inteligente

O FILVE2004 irá comparar se as palavras chaves digitadas pelos usuários, já existem em outros perfis de usuários, se já existe, ele irá selecionar o perfil e verificar as áreas de interesse. Achando a(s) área(s) de interesse irá mostrar o usuário as páginas cujo o grau de relevância é maior, justamente nesse processo é que ocorre a filtragem de informação, pois só serão mostrados para os usuários as páginas filtradas pelo FILVE2004 e com o nível de relevância alto.

Caso as palavras não pertencerem a nenhum perfil de usuário encontrado, o FILVE2004 irá procurar se interagir com outros agentes, através de uma linguagem

específica. Verificando se outros agentes possuem perfis de usuário que se adapte ao usuário em questão, caso o FILVE2004 consiga esse perfil, colocará em evidência os itens já pesquisados por outros usuários com o mesmo perfil, caso contrário, armazenará os itens de interesse de acordo com o usuário em questão.

Quando um usuário, efetuar sua pesquisa, por exemplo, as palavras *Inteligência Artificial*, e pressionar a tecla **Buscar**, o FILVE2004 filtrará todas as páginas de sua base de dados na área de informática. Pois ele comparará as palavras digitadas (*inteligência Artificial*) com as áreas de interesse de cada perfil e irá selecionar o perfil do usuário como *informática*. Caso ele não encontre nenhuma página ele procurará se interagir com outros agentes na mesma área para obter resultados relevantes.

Quando apresentado os resultados obtidos para o usuário, o trabalho do FILVE2004 não termina, virá a parte mais importante, porque ele será responsável por classificar o grau de relevância de cada página. Ele irá gravar todas as ações do usuário, se o usuário copiou ou imprimiu algo na página, quanto tempo ele ficou navegando na página, depois de obter esses dados o FILVE2004 apresentará um grau de relevância. Isso ajudará em futuras pesquisas, pois através dessa gravação e do grau de relevância que o FILVE2004 construirá sua base de dados.

Se um outro usuário, depois de um certo tempo acionar a busca inteligente, na mesma área que o usuário anterior, o mesmo assunto, a busca inteligente retornará como os primeiros resultados, os que possuírem grau de relevância maior, o grau que o FILVE2004 estabeleceu, com isso ficaria muito mais fácil encontrar informações realmente interessante.

CAPÍTULO 6 – UMA NOVA MÁQUINA DE BUSCA

Todo o processo de Recuperação de Informação (RI) foi descrito no capítulo 2. Neste capítulo está descrito tudo o que é necessário para se entender o funcionamento de uma máquina de busca. A seguir será proposta a máquina de busca inteligente, para a recuperação de páginas na *Web*, sendo que, além dessa forma descritiva, a figura 5, também ajuda a entendê-la:

- A primeira coisa a ser feito é buscar páginas na *Web* para que elas possam ser recuperadas mais tarde. Na máquina de busca proposta, usa-se uma arquitetura centralizada, isto é, um *software* chamado Crawler (robô, como nos mecanismo de busca) ele fica responsável por esta tarefa. Os Crawlers atravessam a *Web* enviando páginas novas ou atualizadas para o principal servidor onde elas estão indexadas.
- O segundo passo é justamente fazer a indexação dessas páginas capturadas pelos crawlers. Ainda no indexador, é feita a eliminação das *stopwords* (palavras com pouca representação semântica), dos espaços e da pontuação.
- Com isso cria-se o arquivo invertido (uma lista de palavras, sendo que cada uma delas aponta para uma página).
- Depois de criado o arquivo invertido, o administrador do banco de dados já pode especificar os documentos que serão usados, as operações que serão realizadas no texto, e o texto modelo.
- A interface de consulta é como nas tradicionais. Há uma caixa onde uma ou mais palavras podem ser digitadas. Essas palavras são palavras-chave e devem ser de forma que o sistema entenda o que deve ser recuperado.

- No processamento, foi escolhida a aplicação do FILVE2004, descrito no capítulo 5, o sistema de recuperação de informação usando agente inteligente foi a solução encontrada diante o resultado da pesquisa I realizada no capítulo 3, o agente agirá na parte onde é encontrada a dificuldade maior dos usuários, que é a filtragem de informações. Com isso o agente recuperará apenas páginas consideradas por ele interessantes para cada usuário, isso dinamiza o uso de máquinas de busca, fazendo que as páginas recuperadas sejam atualizadas e precisas.
- A interface de resposta usualmente consistirá de uma lista contendo em cada tela as dez páginas com maior relevância, sendo que cada página será especificada por um título, uma pequena descrição e sua URL.
- A ordenação é feita sem acesso ao texto, somente ao índice. As páginas estarão ordenadas de forma decrescente em relação às suas relevâncias. Essa ordenação será de acordo com a pontuação feita pelo FILVE2004.

6.1 – Esquema geral da máquina de busca inteligente

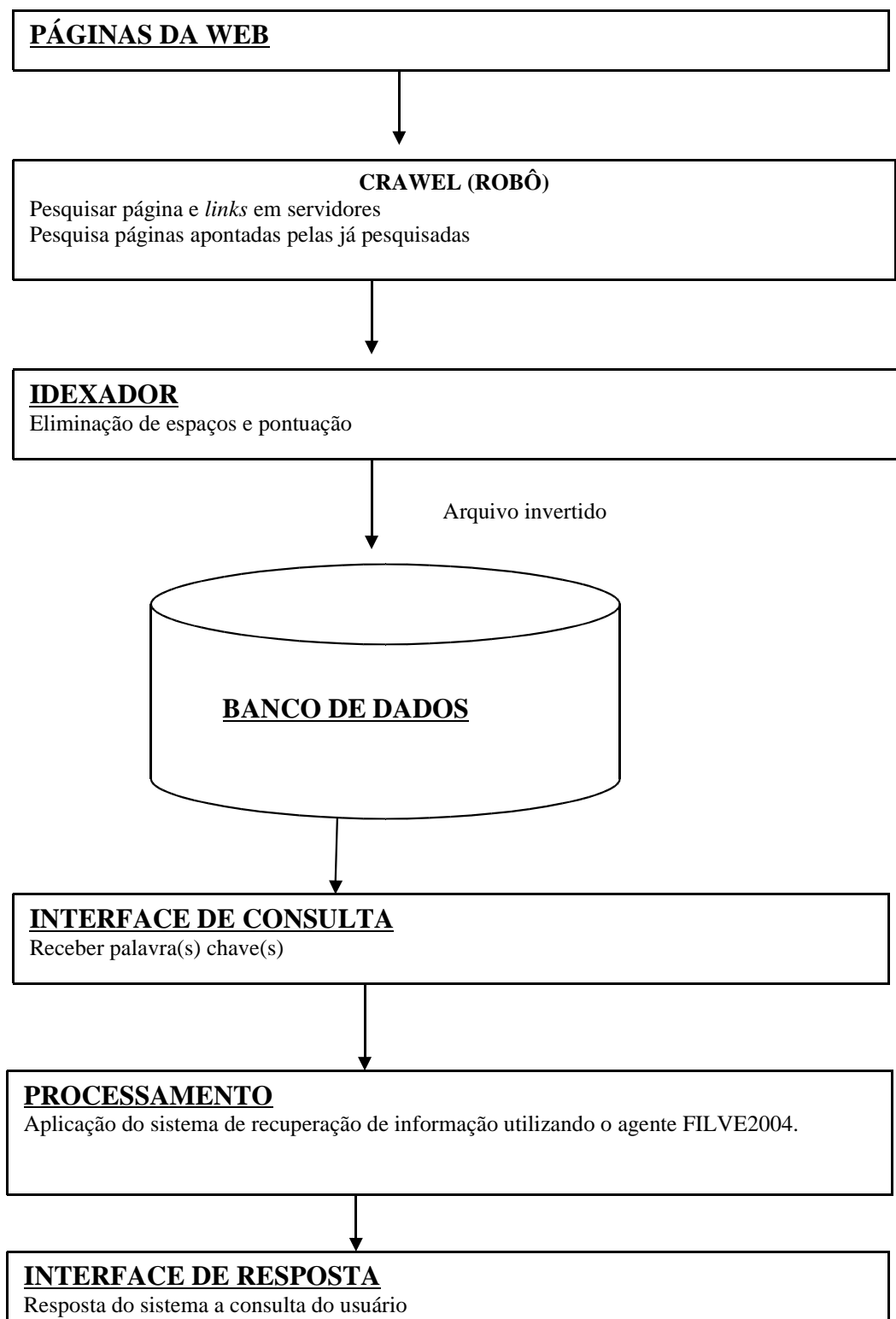


Figura 29: Proposta de uma máquina de busca inteligente para Web.

CONCLUSÃO

Os sistemas de busca foram criados para ajudar usuários a encontrar informações presentes na *Web*.

Obter sucesso em uma busca, não é apenas uma tarefa do sistema, mas sim do usuário, este tem que saber o que realmente está buscando e escolher palavras chaves relevantes. Então conclui-se que devido ao número de URLs encontradas hoje em dia em máquinas de busca, os usuários devem ser mais preparados para utilizá-las, e este trabalho apresentou técnicas especiais para usuários obterem sucessos em suas pesquisas.

Este trabalho trouxe como principal objetivo mostrar as diferenças entre sistemas de busca e através de dificuldades comuns entre usuários obteve-se a proposta de melhoramento do funcionamento das máquinas de busca existentes.

Portanto, este trabalho não apenas propôs um melhoramento em sistemas de busca, mas sim é um grande começo para melhorar a relação homem e sistema de busca.

Com todas as informações presentes nesse trabalho, as pesquisas realizadas na *Web* já não serão mais uma tarefa árdua e sim agradável, pois todos terão o conhecimento necessário para a utilização desses sistemas.

Portanto a proposta apresentada no capítulo 6, seria uma máquina de busca inteligente, e através dela os usuários encontrariam o que realmente são informações relevantes de acordo com cada perfil de usuário.

TRABALHOS FUTUROS

A implementação de uma base de dados para armazenar áreas de interesse do usuário de acordo com seus perfis.

A criação e programação do agente inteligente FILVE2004, apresentada no capítulo 5. Seria um excelente trabalho.

A implementação da máquina de busca inteligente proposta no capítulo 6, seria de uma utilização maravilhosa.

Para a realização de uma das sugestões acima será necessário partir de dados obtidos nesse trabalho

BIBLIOGRAFIA

- [1] ABREU, CHARLES LOPES DE. Trabalho Monográfico - Universidade Federal de Mato Grosso, UFMT. Disponível em <<http://www.ufmt.br/cacomp/Downloads/monografias/IASimbolicaConexionista.pdf>>. Acesso 12 de outubro de 2004.
- [2] AKWAN INFORMATION TECHNOLOGIES. Disponível em <<http://www.Akwan.com.br>>. Acesso em 22 de março de 2004.
- [3] BAEZA-YATES, R.; RIBEIRO-NETO, B. **Modern Information Retrieval**. New Iork: Ed. A. Wesley, 1999. disponível em <<http://www.sims.berkeley.edu/~hearst/irbook.>> Acesso em 16 de julho de 2004
- [4] BLATTMAN, Ursula. FACHIN, Gleisy R. B. RADOS, Gregório J. Varvakis **Recuperar a Informação Eletrônica pela Internet**. Universidade Federal de Santa Catarina, 2000. Disponível em: <<http://www.ced.ufsc.br/~ursula/papers/buscanet.html>>. Acesso em 08 de maio de 2004.
- [5] CENDÓN, B.V. **Ferramentas de busca na Web**. *Ciência da Informação Online*, Brasília, v.30, n.1, p.39-49, jan. abr. 2001. Disponível em: <<http://www.ibict.br/cionline/300101/30010106.pdf>>. Acesso em: 05 maio. 2004

[6] COSTA. Marcelo Thiry Comicholi da. **Uma Arquitetura baseada em Agentes para suporte as Ensino a Distância**. 199. 151 f. Dissertação (Doutorado em Engenharia de Produção). Universidade Federal de Santa Catarina. Florianópolis, 1999.

[7] DCC UFMG. Disponível em <<http://www.dcc.ufmg.br/~nivio/cursos/ri04/>>. Acesso em 10 de maio de 2004.

[8] DETERS. Janice Inês, ADAIME. Silsomar Flores. **Um estudo comparativo dos sistemas de busca na Web**. Anais do V Encontro de Estudantes de Informática do Tocantins. Palmas, TO. outubro, 2003. pp. 189-200. Disponível em <<http://www.ulbra-to.br/ensino/43020/artigos/anais2003/anais/sistemasbuscaWeb-encoinfo2003.pdf>>. Acesso em 13 de abril de 2004.

[9] GILBERT, D. (1997). **Intelligent Agents: The Right Information at the Right Time**. IBM Corporation Research Triangle Park, NC USA. Disponível em <<http://w3.informatik.gu.se/~dixi/agent/read.htm>>. Acesso em 15 de outubro de 2004.

[10] GOOGLE BRASIL. Disponível em <<http://www.google.com.br>>. Acesso 15 de março de 2004.

[11] MOURA. Gevilacio Aguiar Coêlho de. **Sistemas de busca da Web: diretórios e mecanismos de busca**. © Copyright 1999 – 2004. Disponível em <<http://www.quatrocantos.com>> . Acesso 06 de maio de 2004.

[12] OLIVEIRA, Alexandre. MARCOS Adérito, VAASAN, Thanigai. e HORNUNG, Johanna. **TOPIC MAPS NA VISUALIZAÇÃO DE INFORMAÇÃO NO ENSINO E TREINO**. Universidade do Minho, Dep. Sistemas de Informação, Campus de Azurém, Guimarães. Outubro 2000. Disponível em <http://www.dsi.uminho.pt/publicacoes/ficheiros/artigo_topic_maps_final.pdf>. Acessado em 15 de outubro de 2004.

[13] SILVA, Helena Pereira da. THIRY, Marcello. ABREU, Aline França de. **Monitoramento automatizado na *Internet*; uma resposta ao desafio de melhores serviços a custos baixos para as Bibliotecas universitárias**. Universidade Federal de Santa Catarina. Programa de Pós-Graduação em Engenharia de Produção. Santa Catarina – Brasil. Disponível em <<http://snbu.bvs.br/snbu2000/docs/pt/doc/t001.doc>>. Acesso em 20 de outubro de 2004.

[14] TODO BR O BRASIL NA *INTERNET* disponível em <<http://www.todobr.com.br>>. Acesso em 20 de março 2004.

[15] UOLBUSCA. Disponível em <<http://www.radaruol.com.br>>. Acesso em 10 de julho de 2004.

[16] VIDAL. Luís Alfredo. Datamining – mineração de dados.

[17] YAHOO BRASIL COMERCIAL *WEB*. Disponível Em <<http://br.Yahoo.com/ycomWeb/how.html>>. Acesso em : 05 de maio de 2004.

[18] YAHOO BRASIL. Disponível em <<http://www.yahoo.com.br>>. Acesso em 10 de março de 2004.

[19] WOOLDRIDGE, M. JENNINGS, N. R. **Intelligent Agents: Theory and Practice**. Submitted to the Knowledge Engineering Review, 1994. disponível em <<http://www.cs.iastate.edu>>. Acesso em 15 de outubro de 2004.

ANEXO I

NOME:

1) IDADE: 18 a 21 22 a 25 26 a 30 31 a 35 acima de 35

2) SEXO: () Masculino () Feminino

3) Grau de escolaridade:

1º Grua Superior incompleto

2º Grau Superior

2º Grau incompleto Pós Graduação

4) Sua familiaridade com a *Internet* é

Baixa Alta

1 2 3 4 5 6 7 8 9 10

5) Com que frequência você utiliza sites de busca

Nunca usei Raramente Frequentemente

6) Qual a Máquina de busca que você utiliza?

Google Yahoo TodoBr

Altavista Radix GigaBusca

Cadê Zoom Msn

Achei	Prokura	Outra
1	1	1
2	2	2
3	3	3
4	4	4
5	5	5
6	6	6
7	7	7
8	8	8
9	9	9
10	10	10
11	11	11
12	12	12
13	13	13
14	14	14
15	15	15
16	16	16
17	17	17
18	18	18
19	19	19
20	20	20
21	21	21
22	22	22
23	23	23
24	24	24
25	25	25
26	26	26
27	27	27
28	28	28
29	29	29
30	30	30
31	31	31
32	32	32
33	33	33
34	34	34
35	35	35
36	36	36
37	37	37
38	38	38
39	39	39
40	40	40
41	41	41
42	42	42
43	43	43
44	44	44
45	45	45
46	46	46
47	47	47
48	48	48
49	49	49
50	50	50
51	51	51
52	52	52
53	53	53
54	54	54
55	55	55
56	56	56
57	57	57
58	58	58
59	59	59
60	60	60
61	61	61
62	62	62
63	63	63
64	64	64
65	65	65
66	66	66
67	67	67
68	68	68
69	69	69
70	70	70
71	71	71
72	72	72
73	73	73
74	74	74
75	75	75
76	76	76
77	77	77
78	78	78
79	79	79
80	80	80
81	81	81
82	82	82
83	83	83
84	84	84
85	85	85
86	86	86
87	87	87
88	88	88
89	89	89
90	90	90
91	91	91
92	92	92
93	93	93
94	94	94
95	95	95
96	96	96
97	97	97
98	98	98
99	99	99
100	100	100

7) Você sempre encontra o que procura?

Sim Na maioria das vezes sim.

Não Na maioria das vezes não.

8) Você sabe o que são operadores booleanos?

Sim

Não

9) Você sabe como utilizar os operadores booleanos?

Sim, todos Sim alguns.

Não Sim, a maioria.

10) Qual a maior dificuldade que você encontra na utilização das máquinas de busca?

Escolher a(s) palavra(s) chaves para efetuar a busca.

Procurar diante de tantas URLs que aparecem no resultado.

Usar operadores booleanos e termos especiais de pesquisa.

Outras. _____

11) O que você acha das interfaces das máquinas de busca?

Atrativas Passam poucas informações

Todas são iguais. Não conheço

Fácil de utilizá-las Interessante e com muita informação.

12) Você sabe como as máquinas de busca procuram informação na *Web*?

Sim. _____

Não

13) O que você acha que poderia mudar para melhorar os resultados encontrados pelas máquinas de busca?

14) Existem os Termos especiais de pesquisa:

AND	Pão AND Francês encontra documentos tanto com a palavra Pão quanto com a palavra Francês.
OR	Pão OR Francês encontra os documentos contendo ou Pão ou Francês. Os documentos encontrados podem conter ambos os termos, mas não necessariamente.
AND NOT	Pão AND NOT Francês encontra documentos com Pão, mas não contendo Francês.
NEAR	Pão NEAR Francês encontraria documentos com Pão Francês, mas provavelmente nenhum outro tipo de Francês.
*	O asterisco é um curinga; quaisquer letras podem tomar o lugar do asterisco. Compu* encontraria os documentos com Computador, computação, computacional.
()	(amendoim AND manteiga) AND (gelatina OR geléia) localiza documentos com as palavras 'amendoim manteiga e gelatina' ou 'amendoim manteiga e geléia' ou ambos.
“ ”	Pesquisa com uma frase exata. Os documentos selecionados deverão possuir todas as palavras e na mesma ordem em que elas se encontram na frase.

Qual o outro termo você acha que deveria existir.

ANEXO II

RESULTADO PESQUISA II

1) MÁQUINA DE BUSCA GOOGLE

PALAVRA PESQUISADA: Passagens

TIPO DA PESQUISA: Palavra única

Nº DOS RESULTADOS OBTIDOS: 239.000

ASSUNTO:

Passagens rodoviárias

Passagens aéreas

Passagens para a china

Companias aéreas

2) MÁQUINA DE BUSCA GOOGLE

PALAVRA PESQUISADA: Passagens Aéreas

TIPO DA PESQUISA: Palavra composta

Nº DOS RESULTADOS OBTIDOS: 60.200

ASSUNTO:

Passagens aéreas

Reservas de passagens aéreas

Venda de passagens aéreas

Companias de passagens aéreas internacionais e nacionais

3) MÁQUINA DE BUSCA GOOGLE

PALAVRA PESQUISADA: “Passagens Aéreas”

TIPO DA PESQUISA: Uso de técnica especial “ ”

Nº DOS RESULTADOS OBTIDOS: 55.800

ASSUNTO:

Passagens aéreas

Venda de passagens aéreas

Companias de passagens aéreas internacionais e nacionais

Promoção de passagens aéreas

4) MÁQUINA DE BUSCA GOOGLE

PALAVRA PESQUISADA: “Passagens Aéreas” + Promoção

TIPO DA PESQUISA: Uso do operador booleano +

Nº DOS RESULTADOS OBTIDOS: 7.910

ASSUNTO:

Passagens aéreas promoção

Apex Travel promoção de passagens aéreas e hotéis

Concorra a 50 passagens por mês

Reservas online de passagens aéreas

5) MÁQUINA DE BUSCA GOOGLE

PALAVRA PESQUISADA: “Passagens Aéreas” +promoção -hotéis

TIPO DA PESQUISA: Uso do operador booleanos -

Nº DOS RESULTADOS OBTIDOS: 4.430

ASSUNTO:

Concorra a 50 passagens por mês

Premiação de promoção

Onde comprar passagens aéreas em promoção

Bilhetes de passagens aéreas emitidos para curso no exterior

6) MÁQUINA DE BUSCA GOOGLE

PALAVRA PESQUISADA: “Passagens Aéreas” +promoção –hotéis and(curso and internacional)

TIPO DA PESQUISA: Uso do operador booleanos and

Nº DOS RESULTADOS OBTIDOS: 747

ASSUNTO:

Promoção curso no exterior mais passagens aéreas

Curso internacional de treinamento

Preço de passagens aéreas que esta em curso

Curso de francês no exterior com promoção de passagens aéreas

7) MÁQUINA DE BUSCA GOOGLE

PALAVRA PESQUISADA: “Promoção de passagens aéreas para curso no exterior”

TIPO DA PESQUISA: Uso de técnica especial “ ”

Nº DOS RESULTADOS OBTIDOS: 1

ASSUNTO:

Agencia de passagens aéreas para curso no exterior

MÁQUINA DE BUSCA YAHOO

1) MÁQUINA DE BUSCA YAHOO

PALAVRA PESQUISADA: Passagens

TIPO DA PESQUISA: Palavra única

Nº DOS RESULTADOS OBTIDOS: 522.000

ASSUNTO:

Varing

Passagens aéreas nacional e internacional, pacotes turísticos

Seja seu próprio agente de viagem

Agencias BRA

2) MÁQUINA DE BUSCA YAHOO

PALAVRA PESQUISADA: Passagens Aéreas

TIPO DA PESQUISA: Palavra composta

Nº DOS RESULTADOS OBTIDOS: 114.000

ASSUNTO:

Passagens aéreas nacional e internacional, pacotes turísticos

Passagens aéreas preços

Passagens aéreas informa sobre companhias e rotas.

American Airlines

3) MÁQUINA DE BUSCA YAHOO

PALAVRA PESQUISADA: “Passagens Aéreas”

TIPO DA PESQUISA: Uso de técnica especial “ ”

Nº DOS RESULTADOS OBTIDOS: 100.000

ASSUNTO:

Passagens aéreas nacional e internacional, pacotes turísticos

Passagens aéreas consulta de preços

Passagens aéreas informa sobre companhias e rotas.

TAM

4) MÁQUINA DE BUSCA YAHOO

PALAVRA PESQUISADA: “Passagens Aéreas” + Promoção

TIPO DA PESQUISA: Uso do operador booleano +

Nº DOS RESULTADOS OBTIDOS: 14.100

ASSUNTO:

Passagens aéreas BRA

Promoção de passagens aéreas

Apex Travel promoção de passagens aéreas e hotéis

TAM promoção

5) MÁQUINA DE BUSCA YAHOO

PALAVRA PESQUISADA: “Passagens Aéreas” +promoção -hotéis

TIPO DA PESQUISA: Uso do operador booleanos -

Nº DOS RESULTADOS OBTIDOS: 2.120

ASSUNTO:

Passagens aéreas BRA

Promoção de passagens aéreas

TAM promoção

Varing

6) MÁQUINA DE BUSCA YAHOO

PALAVRA PESQUISADA: “Passagens Aéreas” +promoção –hotéis and(curso and internacional)

TIPO DA PESQUISA: Uso do operador booleanos and

Nº DOS RESULTADOS OBTIDOS: 375

ASSUNTO:

Promoção curso no exterior mais passagens aéreas

Curso de turismo emissão de passagens aéreas

Programa da ONU oferece bolsa para jovens

STB, curso no exterior, passagens aéreas

7) MÁQUINA DE BUSCA YAHOO

PALAVRA PESQUISADA: “Promoção de passagens aéreas para curso no exterior”

TIPO DA PESQUISA: Uso de técnica especial “ ”

Nº DOS RESULTADOS OBTIDOS: 0

ASSUNTO:

MÁQUINA DE BUSCA RADAR UOL

1) MÁQUINA DE BUSCA RADAR UOL

PALAVRA PESQUISADA: Passagens

TIPO DA PESQUISA: Palavra única

Nº DOS RESULTADOS OBTIDOS: 92.800

ASSUNTO:

Passagens de ônibus

Passagens aéreas

Sistema de Controle de **Passagens** Mato Grosso

INFO Online - Plantão Info - GOL

2) MÁQUINA DE BUSCA RADAR UOL

PALAVRA PESQUISADA: Passagens Aéreas

TIPO DA PESQUISA: Palavra composta

Nº DOS RESULTADOS OBTIDOS: 139.927

ASSUNTO:

Passagens Aereas VASP, TAM, VARIG, GOL

Salvador Passagens Aéreas

Pacotes turísticos

Se você quer viajar, Malacomalça é a resposta:

3) MÁQUINA DE BUSCA RADAR UOL

PALAVRA PESQUISADA: “Passagens Aéreas”

TIPO DA PESQUISA: Uso de técnica especial “ ”

Nº DOS RESULTADOS OBTIDOS: 10.335

ASSUNTO:

Passagens Aereas VASP, TAM, VARIG, GOL

Passagens Aéreas Nacionais - Passagens Aéreas Internacionais

Pacotes turísticos

Reserva de Passagens Aéreas

4) MÁQUINA DE BUSCA RADAR UOL

PALAVRA PESQUISADA: “Passagens Aéreas” + Promoção

TIPO DA PESQUISA: Uso do operador booleano +

Nº DOS RESULTADOS OBTIDOS: 3

ASSUNTO:

Decolar.com, agência de viagens

Passagens aéreas é com malasemalça

Passagens aéreas, pacotes nacionais e internacionais

5) MÁQUINA DE BUSCA RADAR UOL

PALAVRA PESQUISADA: “Passagens Aéreas” +promoção -hotéis

TIPO DA PESQUISA: Uso do operador booleanos -

Nº DOS RESULTADOS OBTIDOS: 816

ASSUNTO:

Promoção de Passagens Aereas

Emissor Autorizado Passagens Aéreas

Transportes Aéreos em promoção

Passagens Aéreas Reservas

6) MÁQUINA DE BUSCA RADAR UOL

PALAVRA PESQUISADA: “Passagens Aéreas” +promoção –hotéis and(curso and internacional)

TIPO DA PESQUISA: Uso do operador booleanos and

Nº DOS RESULTADOS OBTIDOS: 79

ASSUNTO:

STB, curso no extrrrior, passagens aéreas

Student Travel Bureau Informações sobre cursos no exterior

Combine um ano de experiência internacional

Doação de passagens aéreas para que atletas brasileiros possam competir no exterior

7) MÁQUINA DE BUSCA RADAR UOL

PALAVRA PESQUISADA: “Promoção de passagens aéreas para curso no exterior”

TIPO DA PESQUISA: Uso de técnica especial “ ”

Nº DOS RESULTADOS OBTIDOS: 3

ASSUNTO:

Aprenda inglês online

SNT International College escola de inglês em Bournemouth na Inglaterra

StudyGlobal , aprenda inglês no exterior.

MÁQUINA DE BUSCA TODOBR

1) MÁQUINA DE BUSCA TODOBR**PALAVRA PESQUISADA:** Passagens**TIPO DA PESQUISA:** Palavra única**Nº DOS RESULTADOS OBTIDOS:** 1.500**ASSUNTO:**Seja seu próprio agente de viagemAgencia de turismo de aventuraPassagens aéreas nacional e internacional.Viagens vôos pacotes turísticos.**2) MÁQUINA DE BUSCA TODOBR****PALAVRA PESQUISADA:** Passagens Aéreas**TIPO DA PESQUISA:** Palavra composta**Nº DOS RESULTADOS OBTIDOS:** 1.500**ASSUNTO:**América Airlines em 10 X no cartãoGol Passagens aéreasVenda de passagens aéreasPassagens nacional e internacional

3) MÁQUINA DE BUSCA TODOBR

PALAVRA PESQUISADA: “Passagens Aéreas”

TIPO DA PESQUISA: Uso de técnica especial “ ”

Nº DOS RESULTADOS OBTIDOS: 1.500

ASSUNTO:

Venda de passagens aéreas

Agencia de turismo de aventura

Passagens aéreas nacional e internacional.

Viagens vôos pacotes turísticos.

4) MÁQUINA DE BUSCA TODOBR

PALAVRA PESQUISADA: “Passagens Aéreas” + Promoção

TIPO DA PESQUISA: Uso do operador booleano +

Nº DOS RESULTADOS OBTIDOS: 151

ASSUNTO:

Apex Travel promoção de passagens aéreas e hotéis

LIG Tur

TAM lança promoção

O ministro suspendeu a promoção da Gol

5) MÁQUINA DE BUSCA TODOBR

PALAVRA PESQUISADA: “Passagens Aéreas” +promoção -hotéis

TIPO DA PESQUISA: Uso do operador booleanos -

Nº DOS RESULTADOS OBTIDOS: 178

OBS.: O Todo Br não considera o operador booleano -, considera como se fosse o operador booleano +)

ASSUNTO:

Comparações de preços e hotéis no brasil

Roteiros aéreos , reservas de hotéis

Apex Travel promoção de passagens aéreas e hotéis

Carnaval no Rio 2005

6) MÁQUINA DE BUSCA TODOBR

PALAVRA PESQUISADA: “Passagens Aéreas” +promoção –hotéis and(curso and internacional)

TIPO DA PESQUISA: Uso do operador booleanos and

Nº DOS RESULTADOS OBTIDOS: 41

ASSUNTO:

Bex intercâmbio cultural

Passagens aéreas reserva de hotéis

Crescimento de 15,8% de passagens aéreas

Prazo de emissão de passagens aéreas

7) MÁQUINA DE BUSCA TODOBR

PALAVRA PESQUISADA: “Promoção de passagens aéreas para curso no exterior”

TIPO DA PESQUISA: Uso de técnica especial “ ”

Nº DOS RESULTADOS OBTIDOS: 0

ASSUNTO:

MÁQUINA DE BUSCA ALTAVISTA

1) MÁQUINA DE BUSCA ALTAVISTA

PALAVRA PESQUISADA: Passagens

TIPO DA PESQUISA: Palavra única

Nº DOS RESULTADOS OBTIDOS: 535.000

ASSUNTO:

Central de vendas em São Paulo

Passagens aéreas nacional e internacional.

Venda de passagens aéreas.

Informe sobre companhias e rotas.

2) MÁQUINA DE BUSCA ALTAVISTA

PALAVRA PESQUISADA: Passagens Aéreas

TIPO DA PESQUISA: Palavra composta

Nº DOS RESULTADOS OBTIDOS: 111.500

ASSUNTO:

Passagens aéreas nacional e internacional.

Informe sobre companhias e rotas.

Consulta de preços de passagens aéreas

América Airlines consulta de preços

3) MÁQUINA DE BUSCA ALTAVISTA

PALAVRA PESQUISADA: “Passagens Aéreas”

TIPO DA PESQUISA: Uso de técnica especial “ ”

Nº DOS RESULTADOS OBTIDOS: 99.800

ASSUNTO:

Passagens aéreas nacional e internacional.

Informe sobre companhias e rotas.

Consulta de preços de passagens aéreas

Companhias nacional e internacional.

4) MÁQUINA DE BUSCA ALTAVISTA

PALAVRA PESQUISADA: “Passagens Aéreas” + Promoção

TIPO DA PESQUISA: Uso do operador booleano +

Nº DOS RESULTADOS OBTIDOS: 11.400

ASSUNTO:

Passagens aéreas em promoção.

Viagie para o nordeste com promoção.

Falsa promoção no nome da Gol

Passagem aéreas no brasil

5) MÁQUINA DE BUSCA ALTAVISTA

PALAVRA PESQUISADA: “Passagens Aéreas” +promoção -hotéis

TIPO DA PESQUISA: Uso do operador booleanos -

Nº DOS RESULTADOS OBTIDOS: 1.570

ASSUNTO:

Passagens aéreas em promoção.

Itaú promote passagens aéreas

Promoção Gol

Promoção TAM e Varing

6) MÁQUINA DE BUSCA ALTAVISTA

PALAVRA PESQUISADA: “Passagens Aéreas” +promoção –hotéis and(curso and internacional)

TIPO DA PESQUISA: Uso do operador booleanos and

Nº DOS RESULTADOS OBTIDOS: 285

ASSUNTO:

Curso intensivo

Curso turismo emissão de passagem.

Tudo sobre curso no exterior

Passagens aéreas e promoção de cursos.

7) MÁQUINA DE BUSCA ALTAVISTA

PALAVRA PESQUISADA: “Promoção de passagens aéreas para curso no exterior”

TIPO DA PESQUISA: Uso de técnica especial “ ”

Nº DOS RESULTADOS OBTIDOS: 0

ASSUNTO:

