

3 Mecanismos de busca na Internet

Neste capítulo serão detalhados os conceitos relativos aos mecanismos de busca, com o objetivo de apontar a importância desta ferramenta de recuperação de informações para os usuários. Estes conceitos serão explicados de forma breve, com o objetivo de inserir o usuário dentro deste assunto específico.

3.1.O que é um mecanismo de busca?

Para Silveira (2002), o mecanismo de busca é “um banco de dados que ajuda as pessoas a encontrar informações na Internet de acordo com palavras ou termos digitados pelos usuários”.

Os mecanismos de busca vasculham a Internet diariamente, e armazenam todas as informações encontradas num banco de dados de forma organizada. Quando o usuário faz uma busca na Internet, esse banco de dados é acessado e retorna com todas as informações relacionadas à palavra ou termo pesquisado.

3.1.1.O cenário de buscas na Internet

Os mecanismos de busca vêm crescendo a cada ano e grandes empresas do mercado de informática e tecnologia aplicada a Internet estão fazendo altos investimentos no desenvolvimento dessas ferramentas.



Figura 1 – Grandes marcas que estão investindo pesado no mercado de mecanismos de busca na Internet: Google, Yahoo! e Microsoft.

Segundo a revista Exame (2005), o Google, o mecanismo de busca mais utilizado no mundo, vale hoje 109 bilhões de dólares, superando o investimento de mídia de grandes empresas como a Oracle, Time Warner e Coca-cola. Além disso, teve um faturamento de 5,2 bilhões de dólares nos últimos doze meses, sendo que o Yahoo, seu concorrente, faturou 4,8 bilhões de dólares.

Apontamos com estes números a grande importância dos mecanismos de busca no cenário mundial, demonstrando um novo modelo de negócios na Internet com faturamento real e promissor.

3.1.2.As estatísticas sobre os mecanismos de busca

Inúmeras pesquisas têm apontado os mecanismos de busca como primeira opção dos usuários ao se conectarem na Internet.

Os usuários de Internet classificam a pesquisa como a sua atividade mais importante, recebendo uma nota média de 9,1 em 10, na pesquisa realizada pelo Jupiter Research (1999).

O Instituto de Pesquisas Jupiter Media Metrix (2001) revelou que 28% dos e-consumidores acessam os mecanismos de busca e digitam o nome do produto que querem comprar na Internet. Muito mais do que em canais de shopping (5%) ou clique de propaganda (4%).

Um estudo conduzido pela Forrester Research (2001), afirma que mecanismos de busca na Internet alcançam usuários motivados a achar o que estão procurando. Estes usuários são muito valiosos para qualquer sítio que queira capitalizar sobre esse comportamento.

Os usuários que visitam um sítio como resultado de uma busca têm interesses específicos nos produtos e serviços que este sítio oferece (INTERNET WORLD, 2003).

E-consumidores que encontram um sítio através de uma busca têm cinco vezes mais chance de comprar os produtos e serviços oferecidos do que se tivessem encontrado o sítio através de um *banner* (OVERTURE, 2003).

Segundo Thurow (2003), o mecanismo de busca é a porta de entrada mais usada pelos usuários para encontrar um sítio, sendo que geralmente nas pesquisas relativas a este aspecto variam de 42% a 86% o número de usuários que utilizam os mecanismos de busca para encontrar um sítio.

Para Kenski (2004) os mecanismos de busca tornaram-se os principais intermediários entre os usuários e o enorme manancial de informações disponíveis na Internet.

3.1.3. Classificação de páginas

Cada mecanismo de busca possui suas próprias regras relativas à classificação de páginas. Porém, nenhum desses mecanismos divulga os critérios adotados para os usuários, já que a tecnologia de cada mecanismo é o diferencial dessas empresas com relação à concorrência.

Um mesmo sítio pode aparecer nas primeiras posições dos resultados de uma busca num determinado mecanismo de busca e nos outros nem aparecer, ou aparecer na última página de resultados.

3.2. Tipos de mecanismos de busca

Para entender um pouco mais sobre os tipos de mecanismo de busca é preciso conhecer as classificações e o tipo de tecnologia utilizada por cada um deles.

A seguir estão as três classificações relacionadas a mecanismos de busca na Internet, explicando o funcionamento de cada mecanismo de busca e as relações entre eles. São estes: mecanismos de busca baseado em robôs (*spider* ou *crawlers*), diretórios ou catálogos de busca, mecanismos de metabusca e mecanismos híbridos.

3.2.1. Mecanismos de busca baseados em robôs (*spider* ou *crawlers*)

São os mecanismos de busca baseados em robôs (*spider* ou *crawlers*) que vasculham a Internet atrás de informações e classificam os sítios de acordo com os *links* e os conteúdos encontrados nas páginas dos sítios, como é o caso do Google.

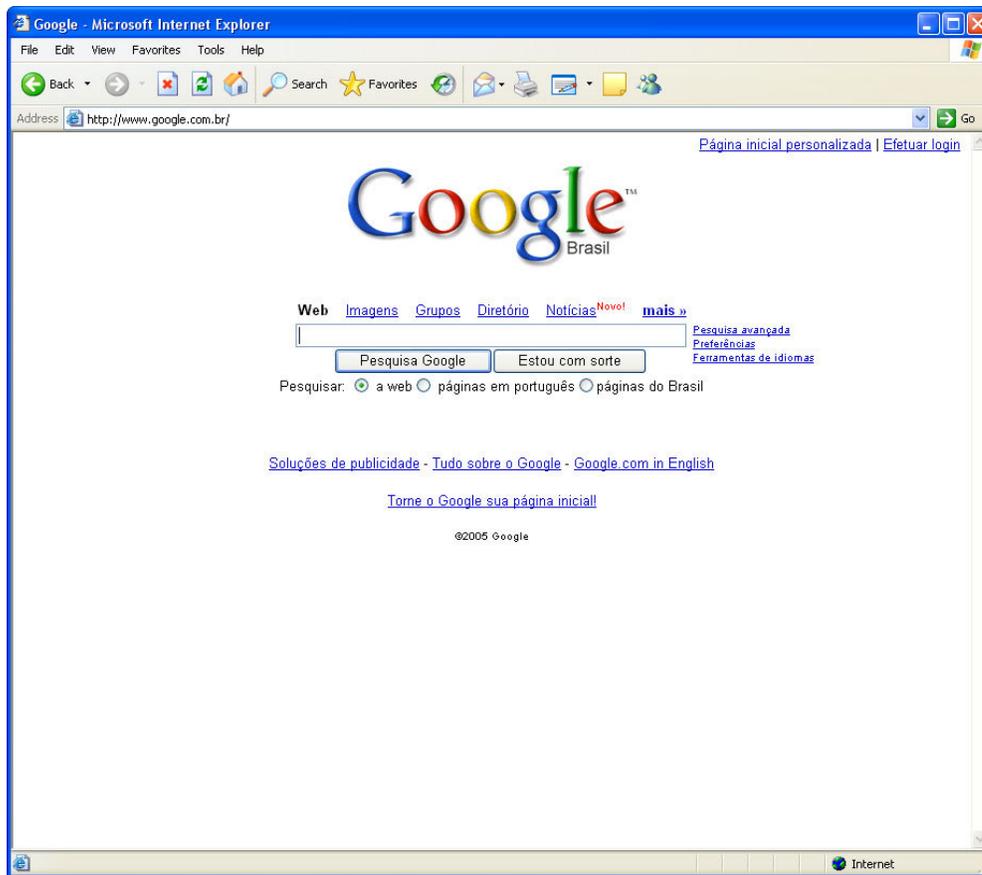


Figura 2 – Google (01/02/2007) <http://www.google.com.br>

Qualquer sítio precisa ser rastreado pelo robô para aparecer nos resultados de busca desses mecanismos.

Segundo Thurow (2003), os robôs executam três ações básicas:

1- Primeiro eles acham as páginas do sítio (processo chamado *spidering* ou *crawling*) e constroem uma lista de palavras e frases encontradas em cada página;

2- Com esta lista eles fazem um banco de dados e encontram as páginas exatas que eles devem procurar, inserindo o sítio vasculhado no banco de dados geral organizado pelas características encontradas em suas páginas. A máquina que insere o sítio no banco de dados geral se chama *indexer*;

3- Depois disso o robô já está apto a encontrar este sítio quando o usuário final digitar na busca uma palavra ou expressão relativa ao conteúdo encontrado no sítio. Essa etapa é chamada *query processor*.

3.2.2. Diretórios ou catálogos de busca

São os mecanismos de busca que utilizam pessoas vasculhando e analisando os sítios que se cadastram nesses mecanismos.

Após a aprovação do sítio nesta análise, este será inserido pela equipe editorial responsável pelo catálogo de acordo com os dados fornecidos durante o cadastramento. Uma boa classificação neste tipo de ferramenta depende, basicamente, da informação que foi fornecida durante o cadastramento do sítio.

É chamado de diretório porque seu banco de dados é dividido e organizado em diretórios, como é o caso do Yahoo! e do Cadê.



Figura 3 – Yahoo! (30/12/2006) <http://www.yahoo.com.br>

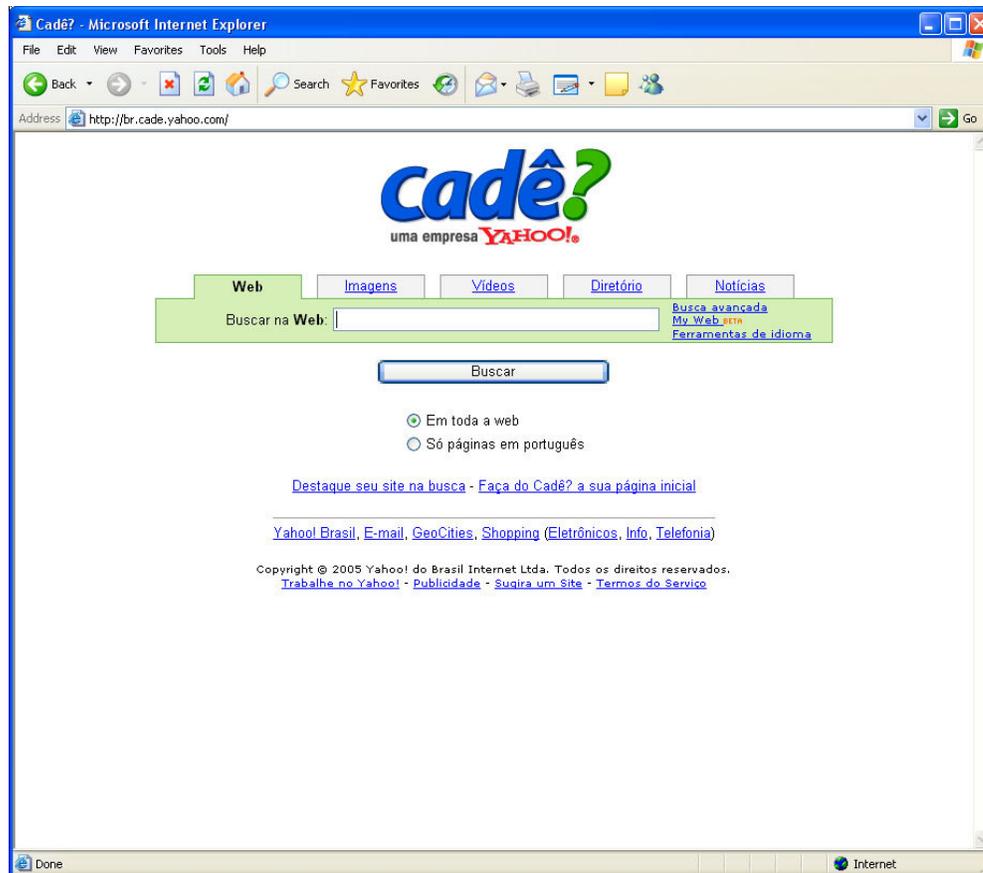


Figura 4 – Cadê (20/09/2006) <http://www.cade.com.br>

Pode ser ainda denominado *human-based* (base humana), já que todo material publicado nestes mecanismos tem que passar por aprovação humana.

Geralmente os resultados das buscas realizadas nos diretórios de busca são um conjunto de sítios que focam no tema buscado pelo usuário.

3.2.3. Mecanismos de metabusca

Os mecanismos de busca que utilizam resultados de outros mecanismos de busca para expor em suas listas de resultados são classificados como **mecanismos de metabusca**. Isso é muito utilizado por empresas que não querem desenvolver sua própria solução de busca e contratam estes serviços de empresas que já têm esta tecnologia desenvolvida. Um dos exemplos mais conhecidos é o mecanismo de busca do portal UOL (Universo On-Line), que recentemente unificou sua ferramenta de busca ao Meta Miner e à ferramenta de busca do BOL, tornando-se o maior mecanismo de metabusca do Brasil, denominado UOL Buscas.

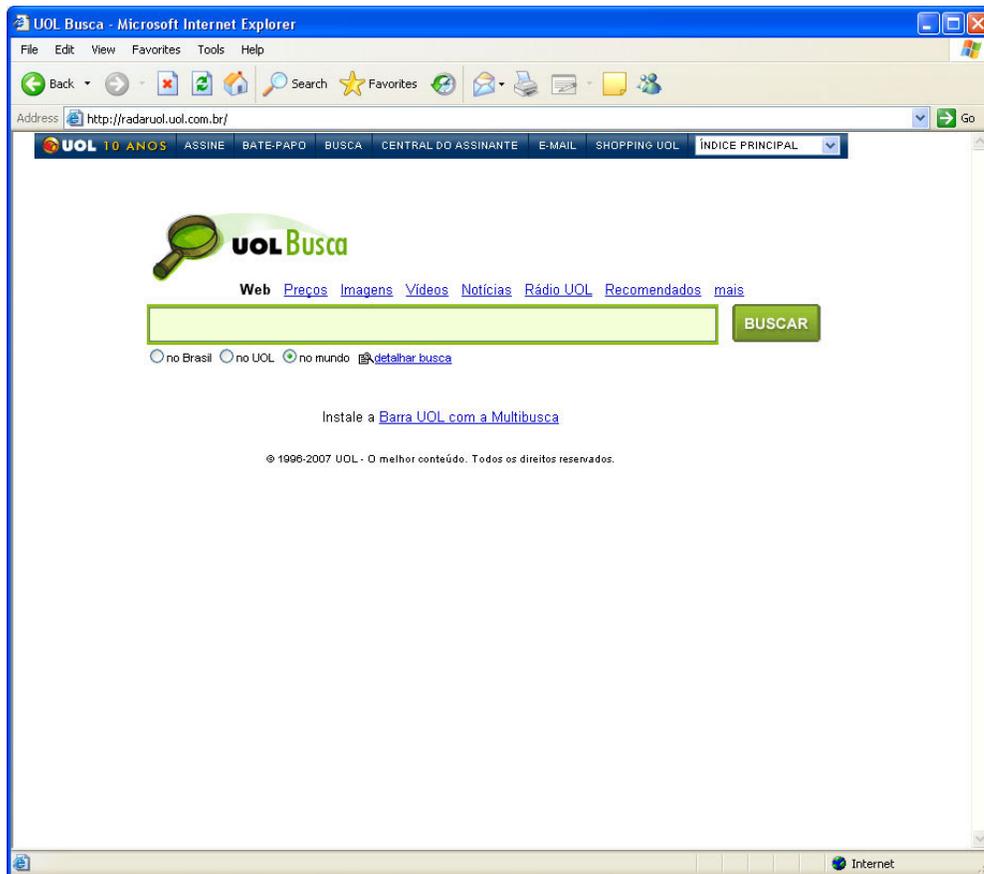


Figura 5 – UOL Buscas (01/02/2007) <http://radaruol.uol.com.br>

Estes mecanismos utilizam os resultados de outros mecanismos de busca e os reorganizam em uma lista só, juntando e agrupando todos os resultados.

3.2.4. Mecanismos híbridos

Os mecanismos de busca que utilizam robôs e humanos analisando e cadastrando sítios são classificados como **híbridos**.

Apesar de o Google ser o mais conhecido mecanismo de busca baseado em robôs, atualmente ele já vem utilizando humanos em seu processo de aprovação de sítios de conteúdo duvidoso. O Yahoo!, que é o mais conhecido mecanismo de busca classificado como diretório, desenvolveu um robô próprio e desde 2004 vem se utilizando dele para aumentar o número de resultados e a exatidão de suas buscas.

Esta tendência adotada pelos dois maiores mecanismos de busca mundiais coloca em evidência a necessidade de os mecanismos de busca se preocuparem, cada vez mais, em conseguir responder exatamente ao que o usuário quer quando digita uma palavra ou frase nos campos de busca.

Porém, é importante ressaltar que deve continuar havendo predominância de uma classificação perante a outra, pois isso define a forma do usuário realizar uma busca. Isso porque procurar com foco em diretórios e categorias é diferente de procurar com foco em um banco de dados amplo criado pelos robôs.

3.3.Principais mecanismos de busca na Internet

Existem milhares de mecanismos de busca na Internet, mas 90% das buscas são efetuadas em menos de 10 mecanismos de busca, segundo Kavinski (2003).

De acordo com esse dado cabe agora apontar os principais mecanismos de busca no mundo e depois, focar nos principais mecanismos. Vale lembrar que essa pesquisa deve ser feita regularmente, porque com a velocidade das mudanças ocorridas na Internet podem acontecer mudanças significativas em poucos meses.

No mundo

Os principais mecanismos de busca na Internet mundial, no ano de 2006, são: (SILVEIRA, 2002; SEARCH ENGINE WATCH, 2006).

- . Google (www.google.com)
- . Yahoo! (www.yahoo.com)
- . MSN Search (www.msn.com)
- . AltaVista (www.altavista.com)
- . AOL (www.aol.com)
- . Overture (www.overture.com)
- . Inktomi (www.inktomi.com)
- . AllTheWeb (www.alltheweb.com)
- . Open Directory (www.dmoz.com)
- . Lycos (www.lycos.com)

Podemos destacar o Google como o principal mecanismo de busca da Internet mundial, com busca baseada em robôs (*spiders*) e o Yahoo!, como o principal diretório de busca e o que mais direciona tráfego gerado por sítios de busca considerando a Internet mundial.

Das 4,5 bilhões de buscas realizadas em julho de 2005 nos EUA, o Google foi responsável por 46,2% do total, seguido pelo Yahoo! com 22,5%, MSN Search com 12,6%, AOL com 5,4% e Outros com 13,3% (NIELSEN NETRATINGS/SEARCH ENGINE WATCH, 2005).

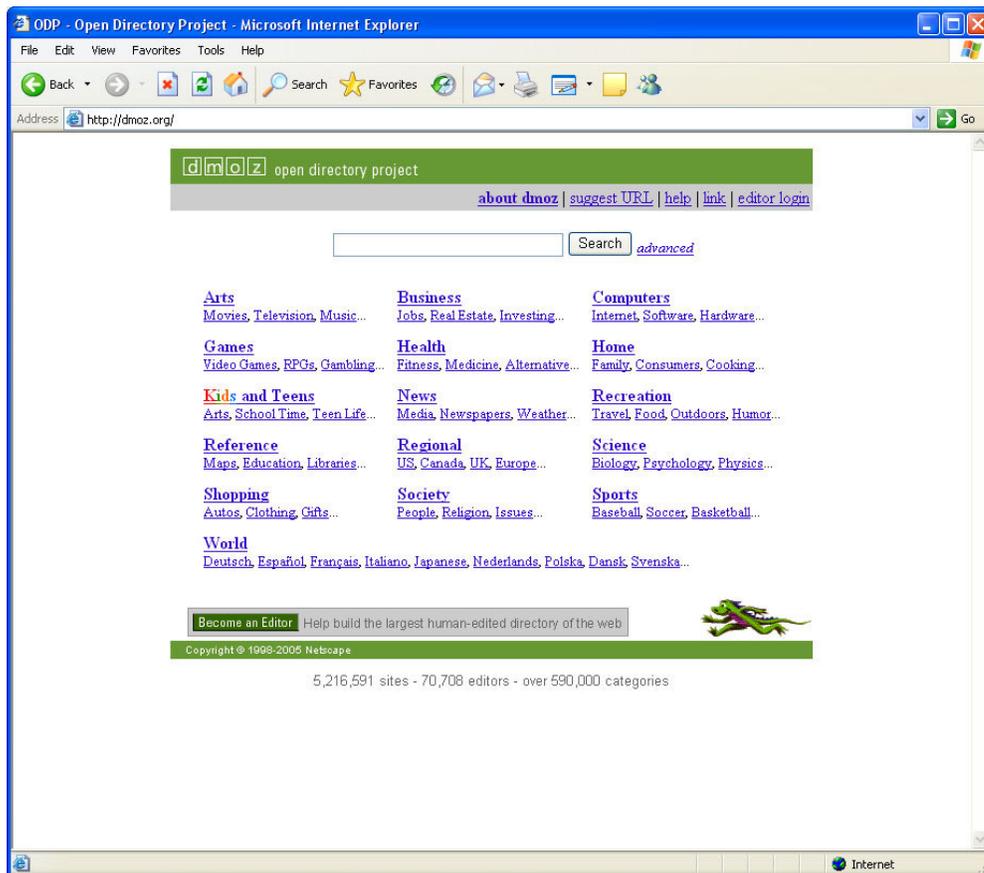


Figura 6 – Open Directory (30/11/2005) <http://www.dmoz.com>

Embora o Open Directory não seja muito conhecido no Brasil, ele tem suma importância, porque fornece informações para outros mecanismos de busca bastante utilizados (SILVEIRA, 2002).

No Brasil

Os principais mecanismos de busca na Internet brasileira, no ano de 2006, são: (SILVEIRA, 2002; SEARCH ENGINE WATCH, 2006).

- . Google BR (www.google.com.br)
- . Yahoo! Brasil (www.yahoo.com.br)
- . Cadê (www.cade.com.br)
- . MSN Busca (www.msn.com.br)
- . AltaVista Brasil (www.altavista.com.br)
- . Zoom (www.zoom.com)
- . Achei (www.achei.com.br)
- . Terra (www.terra.com.br)
- . UOL Buscas (radaruol.uol.com.br)
- . Radix (www.radix.com.br)

No Brasil, seguimos a tendência do sucesso mundial do Google, como o principal mecanismo de busca utilizado. O UOL Buscas é um importante mecanismo de metabusca, pois utiliza resultados de outros mecanismos e também utiliza buscas especializadas em diversas áreas. O Radix é o único mecanismo de busca baseado nos robôs (*spiders*) que foi desenvolvido no Brasil.



Figura 7 – Radix (30/06/2006) <http://www.radix.com.br>

O portal Terra e o sítio de busca Zoom da Globo.com utilizam soluções de busca desenvolvidas pelo Google.

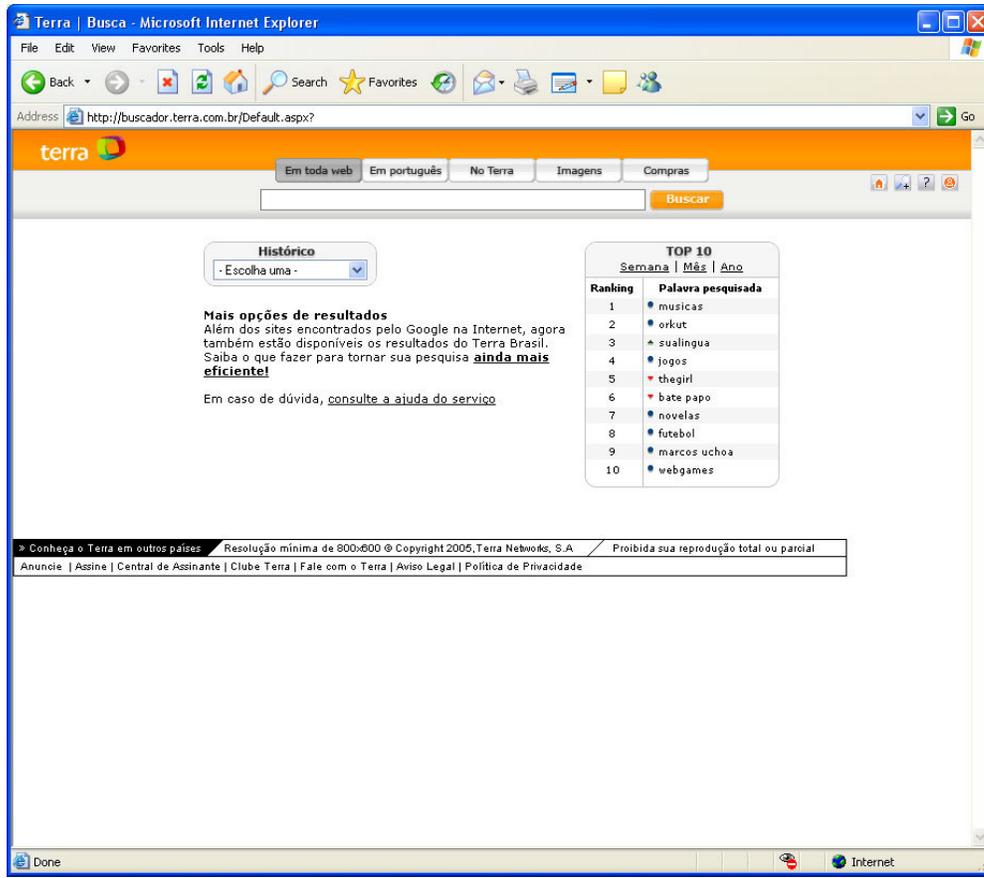


Figura 8 – Terra Busca (01/02/2007) <http://buscador.terra.com.br>

3.4. Entendendo os mecanismos de busca

Para entender o funcionamento dos mecanismos de busca precisamos conhecer seus endereços na Internet, saber a tecnologia usada por cada um e, o principal, saber o que estamos procurando, por exemplo, como usar palavras-chave relacionadas ao tema procurado.

Para começar uma busca na Internet você precisa somente acessar o mecanismo de busca escolhido por você e clicar buscar, após ter escrito um termo ou frase para ser achado.

Segundo Silveira (2002), basta o usuário digitar uma expressão ou palavras individuais relacionadas ao assunto que está pesquisando. O tipo de resultado depende do tipo de busca escolhida pelo usuário e também do algoritmo (conjunto de regras) de pesquisa empregado pelo mecanismo de busca.

Quando o usuário começa a entender o funcionamento dessa ferramenta, ele consegue encontrar resultados mais precisos e em menos tempo de busca.

3.4.1. Tipos de pesquisa

Existem diversos tipos de pesquisa que podem ser realizadas na Internet por cada usuário. Os mecanismos de busca oferecem sempre uma pesquisa para usuários iniciantes (sempre o mais simples e rápido possível), e uma pesquisa avançada, que pode incluir diferentes combinações de palavras-chave ou frases do assunto buscado.

A diferença entre as duas está nos resultados; na pesquisa avançada o usuário consegue ser mais específico e diminuir o número de resultados buscados.

a) Avançada

A busca avançada permite que o usuário utilize diferentes combinações lógicas de palavras para uma busca. Entender as operações de busca avançada pode fazer o usuário ganhar tempo e precisão em suas buscas.

Alguns exemplos de operações de busca avançada:

. ***E*** – ao colocar o ***E*** entre as palavras, o usuário indica para o mecanismo de busca que ele pode responder sua pergunta com todas as palavras digitadas na busca. Ex. perfume e cosmético (só aparecerão resultados que tiverem as duas palavras).

. ***OU*** – ao colocar a palavra ***OU*** entre as palavras, o usuário indica para o mecanismo de busca que ele pode responder sua pergunta com uma palavra ou outra das que foram digitadas na busca. Ex. perfume ou cosmético (podem aparecer resultados só de perfume, só de cosmético ou dos dois juntos).

. **“”** – ao colocar aspas antes e depois das palavras, o usuário indica para o mecanismo de busca que ele pode responder sua pergunta com esta palavra exatamente do mesmo modo que foi digitada na busca. Ex. “Perfume Ralph Lauren” (só podem aparecer resultados que tenham este nome no texto exatamente com a grafia que foi escrita pelo usuário).

. **+ –** ao colocar o **símbolo de adição** antes de uma palavra, o usuário indica para o mecanismo de busca que ele precisa responder sua pergunta com esta palavra. No caso do símbolo ser **de subtração** antes de alguma palavra

significa que esta palavra não pode aparecer nos resultados da busca. Ex. +perfume –Dior (só podem aparecer resultados que tenham perfume no texto, sendo que não pode ter nada relacionado a Dior).

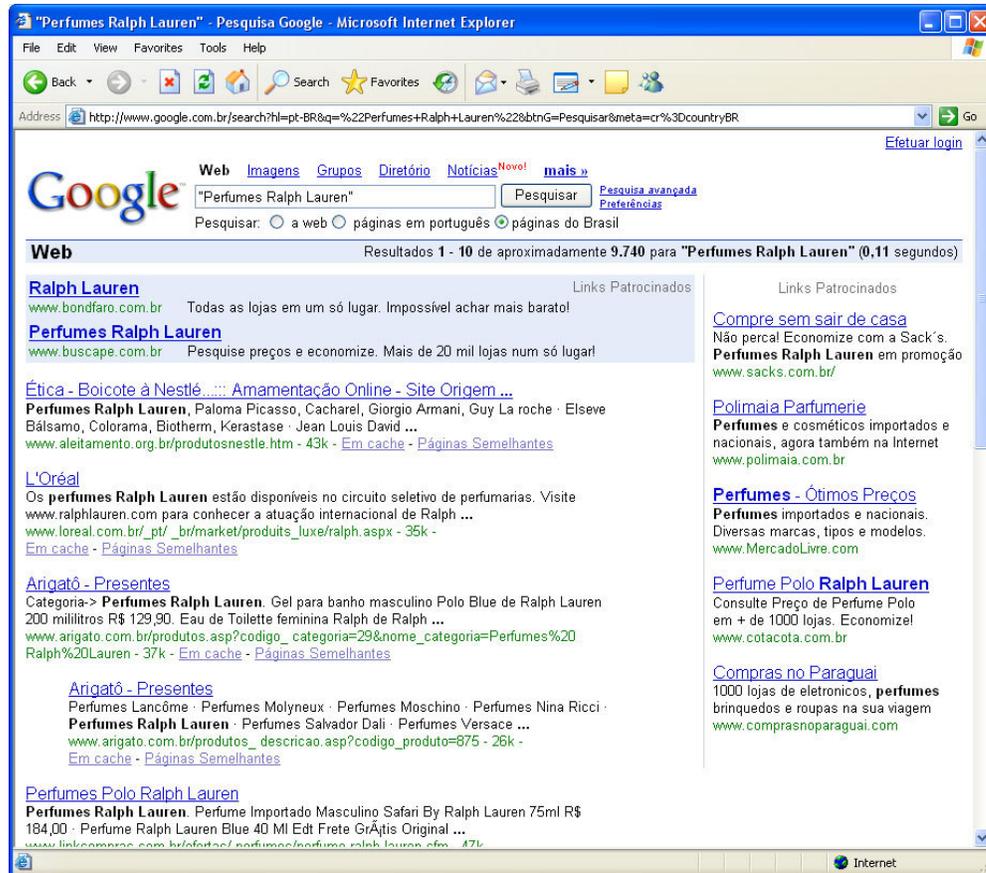


Figura 9 – Exemplo de busca avançada, utilizando aspas, realizada no Google (30/11/2006).

b) Especializada

Os principais mecanismos de busca disponibilizam diferentes opções de pesquisa ao usuário. Estas pesquisas podem ser em todas as páginas na Internet ou serem especializadas em determinadas áreas e seções da Internet.

O Google, por exemplo, apresenta opções de busca por imagens, por livros, por grupos, entre outras formas de busca. A cada semestre este buscador aparece com novidades relacionadas à maneira de se encontrar informações na Internet.

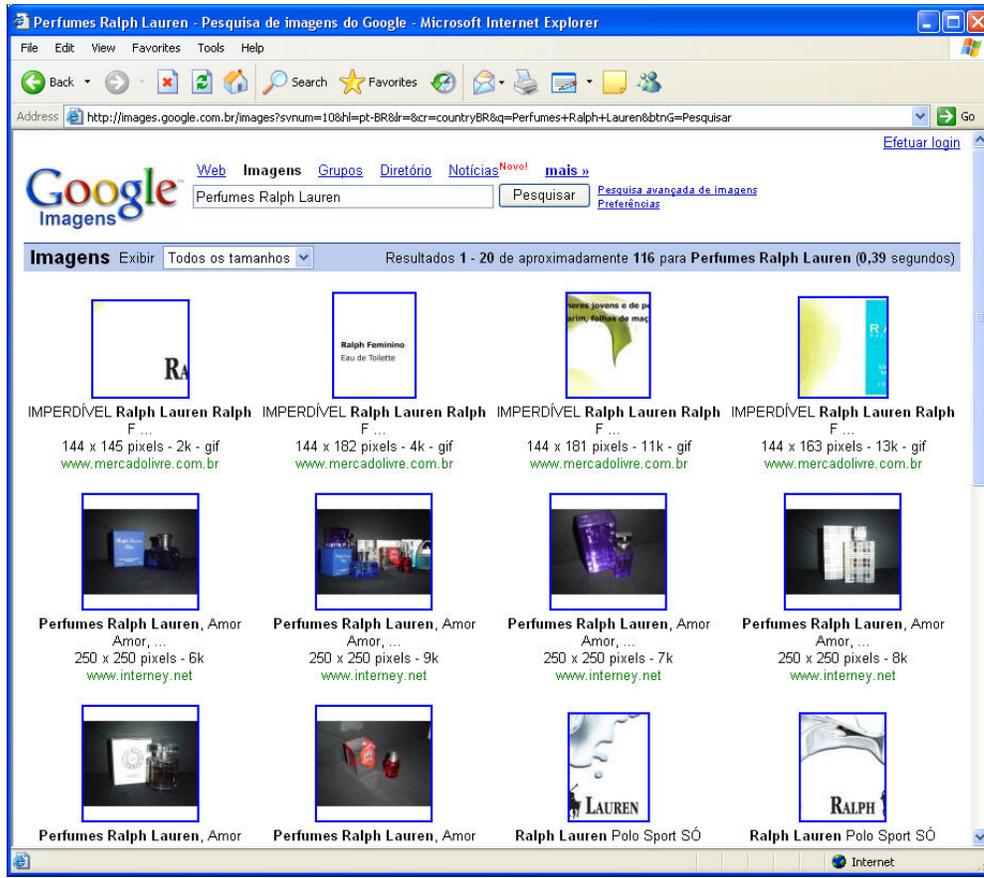


Figura 10 – Exemplo de busca especializada, por imagens, realizada no Google (30/11/2006).

Esse tipo de pesquisa torna os resultados de busca mais específicos e fáceis de serem realizados pelos usuários, tanto iniciantes como experientes.

c) Singular e plural

Normalmente a pesquisa em uma palavra no plural pode levar a resultados totalmente diferentes dessa mesma palavra no singular. Somente alguns mecanismos de busca fazem correlação entre singular e plural.

Quando realizar uma busca deve-se estar atento a essas diferenças e se for preciso buscar tanto no singular, quanto no plural do termo ou frase desejada.

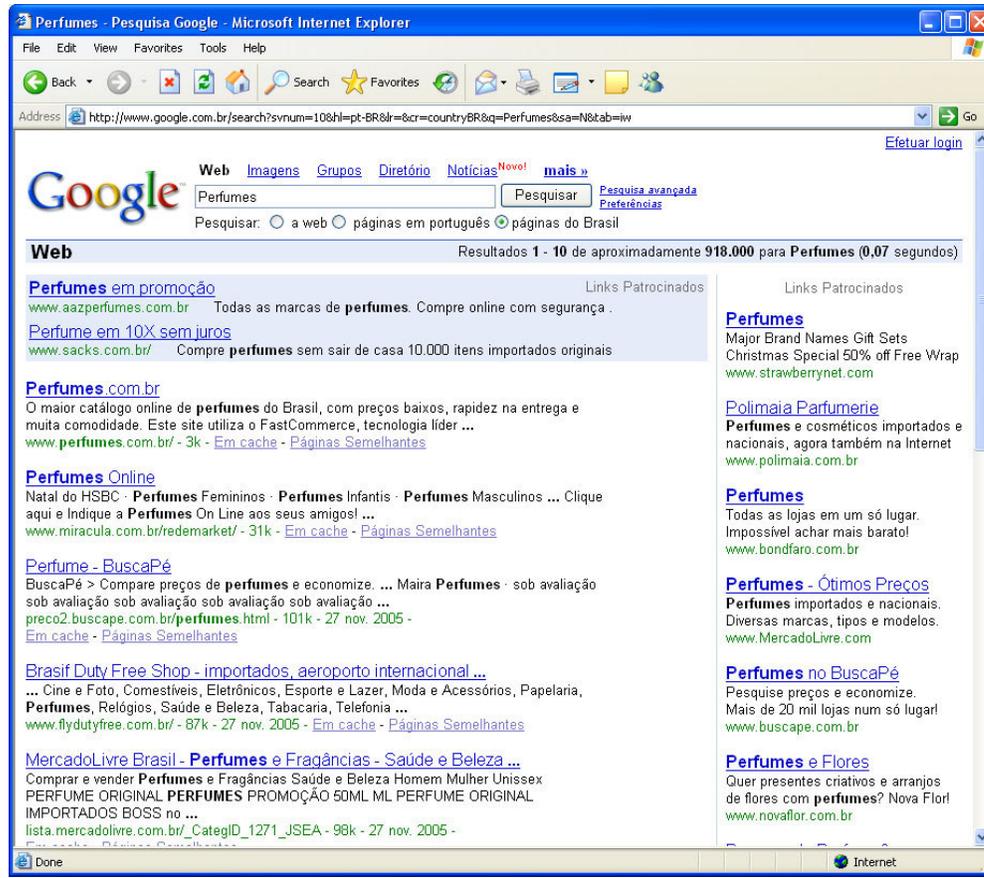


Figura 11 – Pesquisa com palavra no plural realizada no Google (30/11/2006).

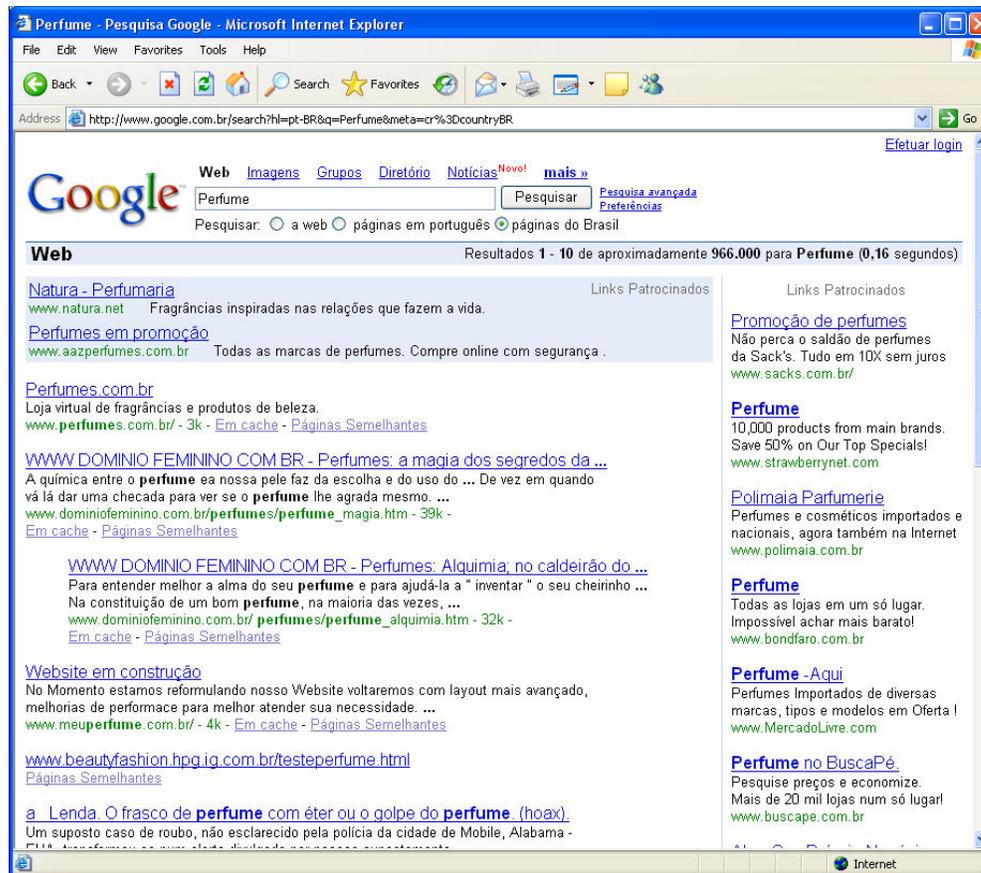


Figura 12 – Pesquisa com palavra no singular realizada no Google (30/11/2006).

d) Palavras desprezadas

Existem algumas palavras que são desprezadas pelos mecanismos de busca. Conjunções, preposições e artigos encaixam-se neste grupo de palavras que podem ser desprezadas na hora em que o usuário for digitar um termo ou frase a ser buscado.

Por exemplo, ao pesquisar por: **perfume de mulher**, o usuário encontrará o mesmo resultado caso procurasse: **perfume mulher**, porque a preposição **de** será desprezada pelo mecanismo de busca.

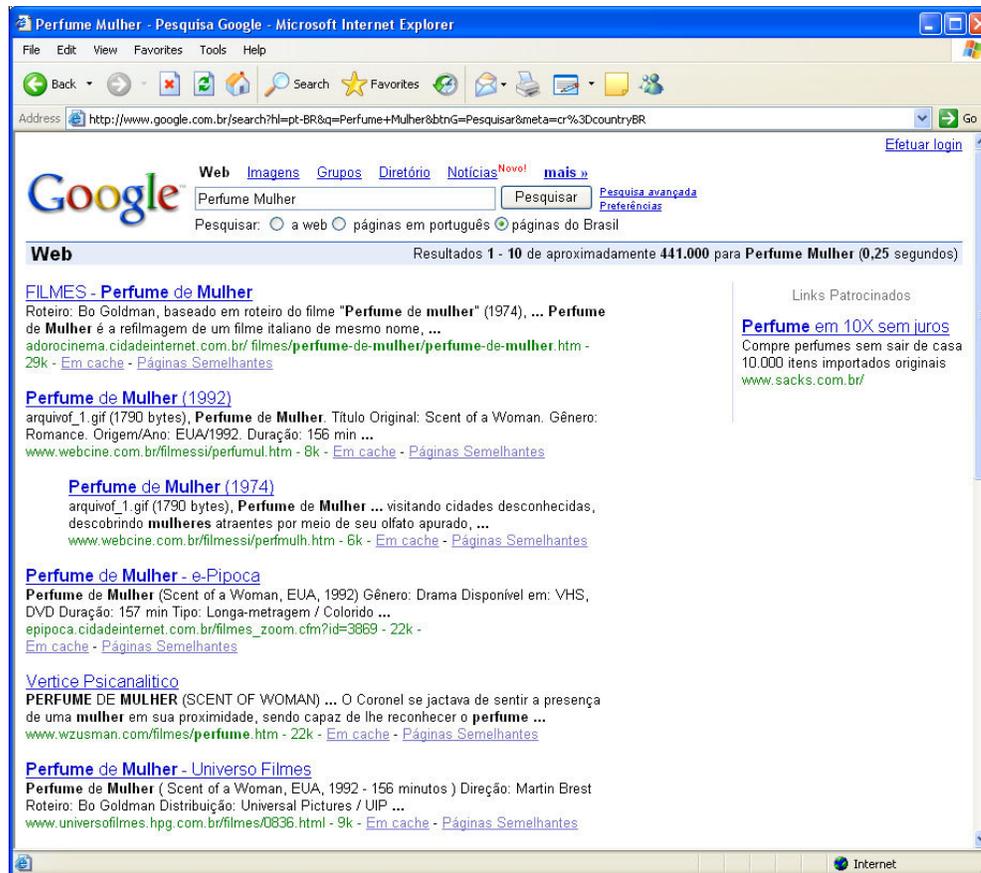


Figura 13 – Pesquisa sem preposição realizada no Google (30/11/2006).

3.4.2. Algoritmos de busca

Cada mecanismo de busca utiliza diferentes critérios e regras para classificar os sítios e assim expor os resultados numa lista gerada a partir de uma palavra ou frase digitada pelo usuário.

Silveira (2002) diz que vários critérios de classificação são usados simultaneamente, mas não é possível obter uma regra clara porque os mecanismos de busca mantêm em segredo os seus algoritmos de classificação. Alguns deles divulgam regras gerais e recomendações sobre seu método de classificação, porém grande parte das regras só pode ser deduzida com exaustivas experiências.

3.4.3. Resultados de busca

Os resultados de buscas na Internet deixaram de ser somente um conjunto de informações agrupadas em ordem de relevância adotada pelos mecanismos de busca, para se transformarem em uma mina de oportunidades a ser explorada por empresas e usuários.

A maioria dos usuários busca nos primeiros 10 resultados dos mecanismos de busca e somente 1% acessa além da terceira página de resultados (AMBERGREEN, 2002).

Segundo Carvalho (2005), um estudo realizado por Brin e Page (os criadores do Google), em 1998, relatou que usuários de Internet não tinham paciência para observar todos os resultados de uma pesquisa efetuada. Naturalmente consultam até os 10 primeiros resultados. Desta pesquisa nasceu o Google, que revolucionou a maneira dos usuários pesquisarem na Internet.

3.5. Conclusões parciais do capítulo

Neste capítulo verificou-se a importância dos mecanismos de busca na Internet. Com o constante crescimento de páginas e informações colocadas na Internet, os mecanismos de busca têm papel fundamental para que estas informações não se percam, nem sejam esquecidas.

Foram listados os tipos de mecanismos de busca e os tipos de busca que o usuário pode fazer nesses mecanismos. Ao oferecer essas ferramentas para o usuário, os mecanismos de busca tentam oferecer resultados cada vez mais precisos, em um menor espaço de tempo de resposta para as buscas.

Apontou-se também que o bom posicionamento nos resultados de busca é fundamental, já que muitos usuários não passam nem da primeira página de resultados nos mecanismos de busca.

Entender como funcionam os mecanismos de busca e a forma do usuário buscar pelas informações nesses mecanismos nos leva a uma infinidade de oportunidades, perguntas e indagações que serão respondidas no capítulo a seguir, que apontará como otimizar um sítio para que este apareça nos primeiros resultados dos mecanismos de busca no seu determinado segmento na Internet.

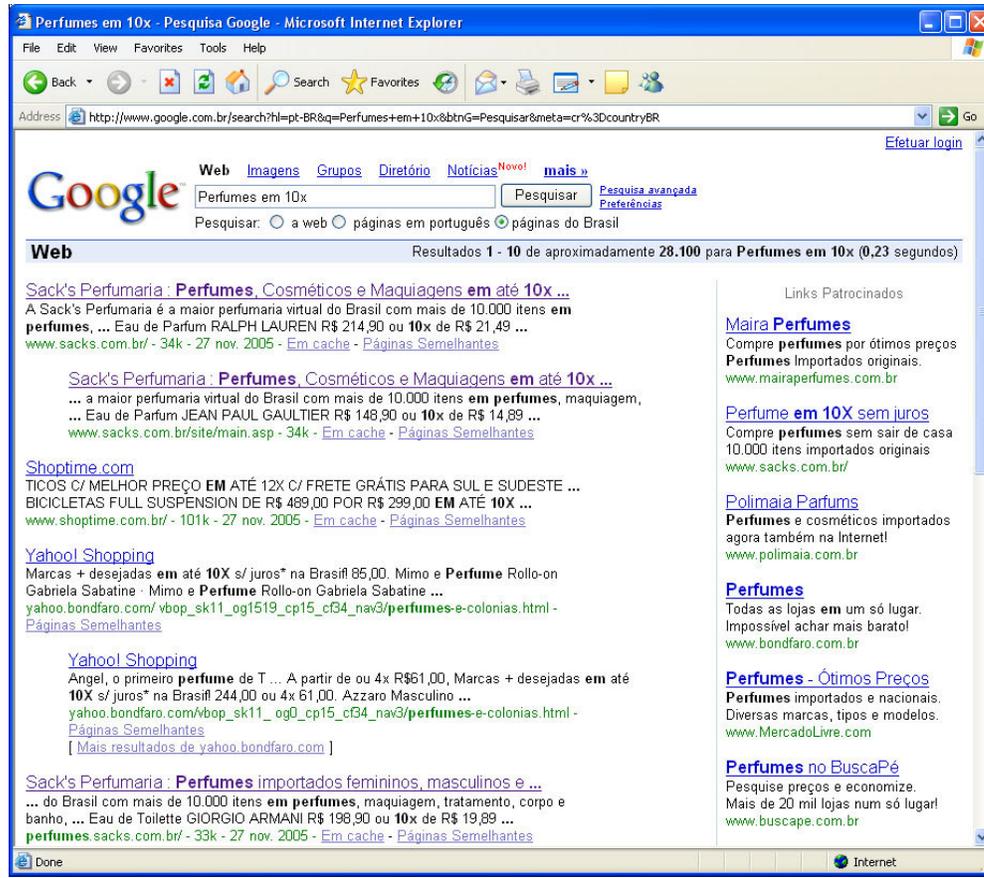


Figura 14 – Resultados de busca do Google, importância de estar nas primeiras posições na lista de resultados (30/11/2006).

3.6.Referências Bibliográficas do capítulo

CARVALHO, Lúcio. *Como se destacar na Internet dominando o Google*. Rio de Janeiro: Publicanet, 2005.

FOLHA DE SÃO PAULO. *Profundezas da rede*. Caderno Informática, 31 de Novembro, 2005.

GUROVITZ, Helio. O Google quer dominar o mundo. *Revista Exame*, ed. 855, nov. 2005.

HALTLEY, R. *Making sure your site is top*. UK: Newsquest Regional Press, 2002.

KAVINSKI, Alexandre. *Como anunciar sem gastar: os buscadores*. Webinsider. Webinsider. Disponível em: <http://www.webinsider.com.br>. Acesso em: 20 set. 2003. 11hs.

_____. *Promovendo seus sítios nos buscadores*. Webinsider. Disponível em: <http://www.webinsider.com.br>. Acesso em: 20 set. 2003. 11:20hs.

KENSKI, Rafael. O mundo Google. *Revista SuperInteressante*, ed. 201, jun. 2004.

SILVEIRA, Marcelo. *Web marketing: usando ferramentas de busca*. São Paulo: Novatec, 2002.

THUROW, Shari. *Search engine visibility*. USA: New Riders, 2003.

WEBWORLD. *Publicidade online em buscadores pagos funciona*. Informativo Semanal, 3 ago. 2001.

ZHANG, Jin; DIMITROFF, Alexandra. *The impact of webpage content characteristics on webpage visibility in search engine results*. USA: University of Wisconsin Milwaukee, 2003.