

Segredos do Google

Desvende os recursos não revelados do poderoso sistema de busca!

Desvendando o Google

- Coloque o seu site em primeiro lugar em uma pesquisa
- Faça pesquisas avançadas de textos e imagens
- Use as ferramentas de tradução
- Visualize o Google no celular
- E muito mais

Invasão de Privacidade

Saiba como a Internet possibilita monitorar informações pessoais de qualquer um, rastrear cartas e entregas, consultar dados da lista telefônica e pesquisar em bancos de dados secretos

Hackeando pelo Google

Como saber tudo sobre um site, incluindo as informações pessoais de seus donos, sem precisar invadi-lo

Segredos do

Google

Desvende os recursos
não revelados do
poderoso sistema
de busca!



© 2004 by Digerati Books

Todos os direitos reservados e protegidos pela Lei 5.988 de 14/12/1973. Nenhuma parte deste livro, sem autorização prévia por escrito da editora, poderá ser reproduzida ou transmitida sejam quais forem os meios empregados: eletrônicos, mecânicos, fotográficos, gravação ou quaisquer outros.

Diretor Editorial

Luis Matos

Assistente Editorial

Monalisa Neves

Redator

Tadeu Carmona

Revisão

Rosangela Cesar

Capa

Daniele Oliveira

Arte e Projeto Gráfico

Patricia Fernandez Ferrari

SEGREDOS DO GOOGLE



Digerati Comunicação e Tecnologia Ltda.

Rua Haddock Lobo, 347 - 12º andar

CEP 01414-001 São Paulo/SP

(11)3217-2600 Fax (11) 3217-2617

www.digerati.com

Diretores

Alessandro Gerardi - (gerardi@digerati.com.br)

Luis Afonso G. Neira - (afonso@digerati.com.br)

Alessio Fon Melozo - (alessio@digerati.com.br)

ISBN: 85-89535-43-6

Impressão e Acabamento

Oceano Indústria Gráfica Ltda.

PREFÁCIO

Parece até lista telefônica da China. São nada menos do que 1,3 bilhão de páginas reunidas em um único local. O número de consultas que recebe é quase igual ao número de habitantes do Brasil. Difícil matar a charada? Então aí vai mais uma dica: seu nome é formado por 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. Agora ficou fácil.

Seu uso é tão difundido que ele originou um verbo amplamente difundido nos Estados Unidos: “googlar”. Sinônimo de pesquisa, o Google é um poderoso sistema de busca na Internet que propicia acesso fácil e rápido a qualquer tipo de informação disponibilizada na Grande Rede em qualquer canto do mundo.

É certo que ele teve antecessores que sobrevivem até hoje, como Altavista, Yahoo, Cadê e cia.. No entanto, é inegável sua superioridade perante a concorrência. Tal é o fenômeno Google, que motivou a criação desse livro focado em ensinar como explorar todo o potencial e as peculiaridades desse poderoso sistema de busca.

Sistema esse que, muitas vezes, nos deixa preocupados, uma vez que é quase impossível ter privacidade na Internet. Basta acessar algum serviço como o Auxílio à Lista (<http://www.auxilio-a-lista.com.br/>), e escolher o Estado. Você será direcionado para o portal da companhia telefônica, no qual pode digitar o nome de uma pessoa para ter acesso a seu endereço e telefone. É a tecnologia contribuindo para a falta de privacidade...

Mas essa mesma tecnologia permite coisas boas como, por exemplo, navegar por um verdadeiro museu reunindo a maioria das páginas já postadas na Internet desde a sua criação. Quer fazer o teste? Acesse o portal www.archive.org, digite o endereço do site que pretende pesquisar e veja toda a coleção de páginas e layouts que já foram usados em seus anos de existência. Sem dúvida, algo impressionante...

Este livro foi escrito para quem deseja conhecer os recursos do Google, e também para aqueles que pretendem fazer o papel de investigador pela infovia e desvendar a vida de uma pessoa ou empresa na Web.

Em síntese, o público-alvo é aquele que, como você, tem sede de conhecimento, buscando acesso rápido ao grande bem da humanidade nos próximos anos, a característica que vai distinguir aqueles que alcançam o sucesso ou não, e que pode ser resumida em uma palavra: informação. Como se vê, o lema “conhecimento é poder” nunca foi tão verdadeiro...

Luis Matos
luismatos@digerati.com.br
Diretor Editorial

Índice

SEGREDOS DO GOOGLE

PRIMEIRA PARTE

O básico do Google	06
Buscas avançadas no Google	14
O Google fora do computador	23
Calendários	27
Datas e horários na Internet	31
Documentos dinâmicos com o Google	35
Colocando seu site em primeiro lugar no Google	38
A volta ao mundo em 80 dias com o Google	45

SEGUNDA PARTE

Você está no Google – O Google e o acesso a informações pessoais e profissionais	50
O uso de caracteres coringa	53
Descobrimo donos de sites	56
Analisando logs via Google	62
Rastreando cartas e entregas	65
Google ajudando os crackers	68
Password generator com Google	73
“Fuçando” bancos de dados no Google	76

TERCEIRA PARTE

Tudo o que eu preciso saber	80
Pesquisa de domínio: como saber tudo sobre um site sem precisar invadi-lo	83
As várias faces do Google	87
Cuidado com os links falsos	91
Dados pessoais via ICQ e MSN	93

Primeira Parte:

Bem-vindo ao Universo Google



O BÁSICO SOBRE O GOOGLE

Ferramentas de busca são recursos relativamente antigos no mundo da informática. No velho MS-DOS, por exemplo, bastava digitar:

```
C:\ find /N "débito" c:\textos\banco.rtf
```

para visualizar na tela todas as linhas de texto em que ocorria a palavra “débito”. A busca ficava restrita ao arquivo indicado – por exemplo, **banco.rtf**. Nada fora do documento setado era trazido para o resultado da busca. Já era alguma coisa, em um tempo em que o computador era usado para um pouco mais do que substituir as calculadoras e máquinas de escrever.

Os sistemas de busca de arquivos mais modernos, como os encontrados no Windows, MacOS e Linux, são mais expansivos: pode-se buscar dados no computador inteiro, na rede em que ele está interligado e até em um diretório armazenado na Web com apenas alguns cliques.

Essa maneira de procurar informações, apesar de ter sua utilidade, poderia ser comparada à onisciência (a capacidade de saber e conhecer tudo), só que limitada ao interior da sua casa e ao outro lado da rua – lugares que você deveria conhecer bem. Da mesma maneira, saber o que está armazenado em seu próprio disco rígido ou em seus CDs de backup deveria ser natural, se a maioria das pessoas fosse minimamente organizada – o que, infelizmente, não corresponde à realidade. Assim, o que os sistemas de busca tradicionais fazem é ruminar informação que você deveria ter, ou já deveria saber onde procurar...

As informações estão lá fora

Os serviços online, em meados da década de 80, já eram mais ou menos conhecidos. Além da Internet – um projeto militar que começava a tornar-se uma rede de serviços ligando as universidades americanas e algumas universidades da Europa entre si –, havia diversos outros serviços de diretório que possibilitavam o acesso a contas telefônicas, informações de bibliotecas e departamentos públicos e downloads de arquivos. No caso da Internet, era imprescindível ser universitário, militar ou trabalhar em um dos dois segmentos. Os serviços de diretório “alternativos”, como o Teletexto, fornecido no Estado de São Paulo pela antiga Telesp, podiam ser assinados por qualquer pessoa, desde que ela, é claro, se comprometesse a pagar as taxas mensais do serviço, que eram quase milionárias.

Foi no ambiente universitário que, em 1995, dois estudantes da Universidade de Stanford, **Sergey Brin**, de 23 anos, especialista em desenho de aplicativos Web e graduado em Engenharia Eletrônica, e **Larry Page**, de 24 anos, *expert* em tratamento de dados e licenciado em Informática e Ciências Matemáticas, se conheceram. Interessados em conseguir acesso ao curso de Doutorado em Ciências Informáticas

da Universidade de Stanford – um dos mais concorridos dos Estados Unidos, e talvez do mundo – os dois estudantes acabaram por descobrir outros pontos de interesse em comum, entre eles um projeto ambicioso: criar um algoritmo de extração de dados que possibilitasse a extração de grandes volumes de informação.

Esse projeto era restrito aos sites e bancos de dados da Biblioteca Digital da Universidade de Stanford. Para que fosse possível fazer buscas fora da rede de Stanford, era necessário construir um novo motor de buscas, com a capacidade de ler páginas de toda a Web em busca de informação, e montando logo em seguida uma lista de links, acompanhada do contexto de cada uma das páginas encontradas.

O que é um motor de busca?

Um motor de busca é uma espécie de catálogo mágico. Mas, diferente dos livros de referência comuns, nos quais está acessível a informação que alguém organizou e registrou, o catálogo do motor de busca está em branco, como um livro vazio. Ao se realizar uma consulta, a lista de ocorrências de assunto é criada em poucos segundos por meio do trabalho de um conjunto de softwares de computador, conhecidos como *spiders* (aranhas), que vasculham toda a Web em busca das ocorrências de um determinado assunto em uma página. Ao encontrar uma página com muitos links, os *spiders* embrenham-se por eles, conseguindo, inclusive, vasculhar os diretórios internos – desde que eles sejam públicos, ou seja, tenham permissão de leitura para usuários – dos sites nos quais estão trabalhando. Motores de busca muito refinados são capazes de saber exatamente que atualizações houve em um site usando esse método de scanner.

Foi assim que, em 1996, Larry e Sergey lançaram o BackRub. Baseado em Java e Python (você pode ver um link com uma dúvida enviada por Larry a uma lista de discussão especializada no endereço: <http://groups.google.com/groups?hl=en&lr=&ie=UTF-8&safe=off&thread=page-0701962007020001@qwerty.stanford.edu&rnum=1&prev=/groups?selm=page-0701962007020001@qwerty.stanford.edu>), o BackRub rodava em algumas máquinas Sun e Intel, localizadas no interior da Universidade. Tanta tecnologia (para a época) era utilizada para amparar a principal diferença do programa em relação aos mecanismos de busca que haviam surgido no ano e meio em que o “pré-Google” estava sendo desenvolvido: o BackRub era capaz de pesquisar os links listados dentro das páginas pesquisadas, aumentando consideravelmente o número de resultados.

O nome Google só seria adotado em 1997, ano em que o projeto deixou de utilizar as máquinas de Stanford. O Google funcionou “como era possível” até a primeira metade de 1998, quando, impulsionados pela compra “milagrosa” de vários terabytes de disco a um preço muito baixo, os sócios decidem montar o CPD (Centro de Processamento de Dados) da empresa na casa (mais especificamente, no quarto) de Larry.

Em finais de 1998, a Google Inc. foi fundada oficialmente, utilizando como capital (e recursos para saudar dívidas com a compra de hardware) 100 mil dólares oferecidos por Andy Bechtolsheim, um dos fundadores da Sun, e mais 1 milhão de dólares doados por amigos e parentes. Nesses primórdios, o Google era assim:



O que significa Google?

Ao procurar pelo termo “Google” utilizando o próprio, você encontrará, exatamente, 54.200.000 ocorrências do termo. Mas, curiosamente, nenhum desses links explica o que quer dizer a palavra Google. É que Google é um termo forjado, retirado do termo *googol*, inventado pelo Dr. Edward Kasner, da Universidade de Columbia. O Dr. Kasner pretendia batizar, com um nome sonoro e fácil de recordar, a centésima potência do número 10, ou um número 1 seguido de 100 zeros. Não satisfeito com esse número absurdo, o cientista criou o *googolplex*, que equivale a um *googol* seguido de um *googol* de zeros. Seja como for, a única utilidade do *googol*, desde o momento de sua invenção, foi a de servir de inspiração para o Google, aproximando a idéia de um número extenso com a da elasticidade inesgotável dos limites da Web. Afinal, não existe nada no Universo (nem estrelas, nem grãos de poeira, nem de átomos) que sequer chegue perto de um *googol*. Já o *googolplex* corresponde a um valor tão absurdo que seria necessário preencher todo o Universo conhecido somente para escrever o algarismo por extenso.

Como funciona o Google

O Google é, essencialmente, um mecanismo de busca de palavras e links por toda a Internet, utilizando diversos recursos de filtragem e catalogação de resultados.

Mas, o que garantiu o sucesso do Google, já que existia uma infinidade de buscadores no momento em que ele ganhava fôlego? Além de seus algoritmos de

extração de dados, que tornam qualquer busca significativamente mais rápida do que qualquer procura realizada com outros sistemas de busca, a interface ajuda muito. O *front-end* do Google é composto apenas por texto e links em HTML, o que faz com que a página retorne buscas quase imediatamente, mesmo em combinações de hardware-software muito antigas, como um 486 com o Internet Explorer 4, ou navegadores muito simples, como o browser em modo texto Links, utilizado por várias distribuições Linux. Essa leveza, além de facilitar a vida do usuário, possibilita que o Google jamais tenha sido retirado do ar por problemas de utilização, e isso contando que recebe cerca de 200 milhões de consultas por dia.

Fazer uma pesquisa no Google é tão simples quanto carregar a página digitando www.google.com.br. No Windows, basta digitar o termo desejado para a pesquisa: suponhamos que você queira pesquisar alguma coisa sobre *gravadores*. Digite esse termo na caixa de verificação e clique em **Pesquisa Google** ou aperte o botão Enter.



Como você pode ver na figura acima, foram retornados alguns milhares de resultados. Isso é comum não só no Google, mas também em boa parte dos mecanismos de busca. Esse comportamento reproduz uma lei da matemática, aparentada à teoria dos conjuntos: quanto menos delimitadores há em uma consulta a um conjunto de elementos, maior será o número de elementos presentes no resultado final. É como se, sendo menos seletivos ao montar uma lista de convidados para a festa (“gordos”, “magros”, “pode trazer um amigo”, “não precisa trazer presente”) acabássemos convidando a cidade inteira.

Esse é um outro grande diferencial do Google. Seu motor de busca é capaz de adicionar, sem muito esforço, diversos filtros tanto a novas buscas quanto depurando buscas já existentes. Vamos conhecer alguns deles:

A lógica booleana

Outro diferencial fica por conta da engrenagem lógica do Google, que utiliza a **lógica booleana**. Batizada assim em homenagem ao matemático britânico George Boole, esse sistema estabelece a possibilidade de busca de palavras em um texto, condicionando a exibição de resultados a valores lógicos:

- um valor deve ser sempre verdadeiro ou falso;
- um valor não pode ser verdadeiro e falso ao mesmo tempo;
- matematicamente verdadeiro pode ser definido como “1”, e falso, como “0”.

Visando não delimitar a busca em demasia, são utilizados elementos para incrementar as funções booleanas ou algoritmos de busca utilizados pelo Google. Desses elementos, os mais conhecidos são:

AND – Esse delimitador é utilizado para incluir, em uma expressão, todos os elementos que serão propostos em uma consulta (as operações booleanas são chamadas dessa maneira, apesar de não serem usadas somente em motores de busca). Esse delimitador é utilizado em sua forma booleana na maioria dos buscadores, como, por exemplo, no Altavista.



O Google não utiliza o delimitador AND de forma explícita: basta digitar os termos separados por um espaço para sua inclusão na busca.



OR – Delimitador de variável. O OR (sempre em maiúscula) é utilizado para encontrar páginas em que se encontre um ou outro termo de uma busca. Ao procurar, por exemplo, pelos termos *livros* e *Digerati* separando-os com OR, teremos acesso a todas as páginas nas quais todos os termos ocorram, mesmo que não de forma conjunta. Esse delimitador é ideal para termos não-corriqueiros, como expressões científicas ou literárias.

NOT (-) – Esse operador é utilizado para suprimir um determinado termo de uma busca, servindo, portanto, como uma espécie de filtro de conteúdo. No Google, é utilizada a forma de sinal (-) aplicada antes de um termo. Ao procurar, por exemplo, pelo termo Linux, mas querendo excluir os termos Conectiva e Unix, utiliza-se o seguinte comando:



e obtém-se o seguinte resultado:



ASPAS (“ ”) – As aspas são utilizadas na lógica booleana para garantir que uma expressão completa (ou conjunto de termos) seja incluída na busca. É muito útil para o caso de expressões em português. Ao procurar, por exemplo, pelos termos *placa de captura ATI*, teremos como resultado páginas em que pode aparecer tanto a expressão completa quanto apenas os termos *placa* ou *captura*, sem nenhuma menção a *ATI*.

Utilizando-se, porém, o conjunto de termos entre aspas – “*placa de captura ATI*” – formaremos uma expressão exata, que será utilizada para delimitar o resultado da busca, como pode ser visto abaixo:



Outros delimitadores de busca

Existem outros delimitadores de busca simples, mas que são próprios do Google. Eles podem ser localizados logo abaixo da barra de pesquisas, em uma linha que começa com a palavra **Pesquisar**.

O padrão dessa linha é fazer a pesquisa em toda a *Web*, buscando informações em todos os servidores disponíveis e que contenham páginas ativas. Utilizando-a ao buscar termos técnicos-padrão ou internacionalizados, como por exemplo *ADSL*, teremos um retorno absurdo de páginas, incluindo algumas em holandês, russo e japonês:



Essa enxurrada de resultados, sobretudo em línguas herméticas, pode ser contida. Na página inicial do próprio Google, basta clicar na opção *páginas em português* para receber resultados somente em nossa língua-pátria.

Mesmo assim, esses resultados incluirão diversas páginas de Portugal (e algumas de Angola), o que é uma perda de tempo ao se pesquisar, por exemplo, sobre serviços públicos, ou procurar o menor preço para determinado produto.

Para resolver esse problema, o Google disponibiliza a função *páginas do Brasil*, que faz com que a pesquisa seja filtrada somente em páginas localizadas em servidores brasileiros ou terminadas com o sufixo .br.



BUSCAS AVANÇADAS NO GOOGLE

Tudo o que foi mostrado no capítulo anterior não corresponde a 2% do que o Google consegue fazer. Existem diversas opções de busca avançada – e a maioria sem sequer sair da página inicial do Google – que podem ser utilizadas para incrementar pesquisas e procurar muito mais do que links e páginas Web.

Google: o pescador de imagens

Nem só da busca de textos e referências escritas é feito o Google. O maior buscador de informações do mundo também conseguiu, em menos de 2 anos, tornar-se a principal ferramenta de busca de imagens utilizada na Internet.

A possibilidade de se pesquisar imagens na Web já existia antes: o pioneiro nesse tipo de serviço era o Altavista, que possui um serviço de busca de imagens desde fevereiro de 2000. O Altavista (<http://www.altavista.com/image/default>), aliás, possui muito mais ferramentas de filtragem de imagens do que o Google (<http://images.google.com.br/imghp?hl=pt-BR>): é capaz de diferenciar fotos de elementos gráficos, botões e banners, além de possuir um filtro de cores e um de resolução de imagem com mais de 22 opções.



Por que, então, ao fazer uma pesquisa de imagens relacionadas ao termo *Linux*, o Altavista encontrou “somente” 91000 imagens, e o Google, mais de 409.000? Não pelo sistema de busca de imagens, que é muito parecido: na verdade, o que são procuradas são referências textuais a figuras ou elementos gráficos contidos em páginas HTML. Ao procurar por imagens JPEG, por exemplo, o Google ou o Altavista procuram pela ocorrência da expressão *.JPG* no código da página. Ao procurar, então, por Linux, o Google identificará, em uma página qualquer, uma imagem que corresponda à busca requerida assim:

```
opentit40112001 - WordPad
Formato: [Barra] Editar Formatar Alinhar
Imagem 100%

<HTML>
<HEAD>
  <META NAME="GENERATOR" CONTENT="Adobe PageMill 3.0 Win">
  <TITLE>Untitled Document</TITLE>
</HEAD>
<BODY BGCOLOR="#ffffff" LINK="#000000" ALINK="#000000" VLINK="#000000">

<PRE><A HREF="bioinfo_pics/linux-bioinformatics.b.gif"><IMG SRC="bioinfo_pics/linux-bioinformatics.gif"
WIDTH="170" HEIGHT="107" ALIGN="RIGHT" BORDER="0" NATURALSIZEFLAG="3"></A><B><FONT SIZE="+1" FACE="Arial">
The Bioinformatics group focuses on high-throughput informatics solutions<br>
for the organization and exploitation of genomic data.
</FONT></B><FONT SIZE="+1" FACE="Arial"> </FONT><FONT FACE="Arial"><A
HREF=""></A></FONT>

<FONT FACE="Arial">Tel: (304) 975 4321</FONT>
<FONT FACE="Arial">Fax: (304) 975 4839</FONT></PRE>

<H6><br></H6>

<H6><br></H6>

<H6><br></H6>

<H6><br></H6>

<PRE><CENTER><HR></CENTER></PRE>

<H6><CENTER><I><FONT FACE="Arial">This Website was last updated
on May 23, 2000</FONT></I></CENTER></H6>

<H6><CENTER><I><FONT FACE="Arial">Questions or Comments about
this Website should be addressed to <A HREF="mailto:jnowak@gene.pitt.edu">Jacek Nowak</A></FONT></I></CENTER></H6>

</BODY>
</HTML>

Seu editor ainda, pressione F1
```

Como isso não ajuda muito o usuário, o Google trata de enviar as informações necessárias para que o browser da pessoa que faz a consulta consiga visualizar a página de resultados e as imagens encontradas, como a página do exemplo abaixo:



Pesquisa avançada de imagens

Ao lado da caixa de pesquisa Google existe um link intitulado **Pesquisa avançada de imagens**, que é utilizado para a filtragem da busca. Como dissemos anteriormente, essa busca não é tão específica (para não dizer preciosa) quanto a oferecida pelo Altavista, mas pode economizar muito tempo, principalmente de quem necessita de um grande volume de imagens sobre um mesmo assunto em pouco tempo.

Ao clicar nesse link, o usuário é direcionado para a página *Pesquisa avançada de imagens*, que é dividida em duas partes: *Procurar resultados* (em azul), onde é possível refinar a busca com indicadores booleanos, e a seção de filtros relacionais, na qual é possível utilizar outros tipos de indicadores.



Procurar resultados relacionados com todas as palavras equivale a escrever palavras soltas no campo de pesquisas do Google. Ao pesquisar, por exemplo, os comandos *disco* e *linux*, surgirão diversas figuras relacionadas a ambas as palavras.

Procurar resultados relacionados à frase exata equivale a escrever palavras entre aspas (“ ”) na página inicial do Google. Ao procurar, por exemplo, por *Star Wars*, surgirão todas as imagens relacionadas à série que o Google consiga encontrar.

Procurar resultados relacionados a qualquer uma das palavras equivale a fazer uma pesquisa de texto utilizando o delimitador OR.

Procurar resultados não-relacionados às palavras é um recurso que pode ser utilizado para excluir um determinado termo de uma busca que pode se tornar muita extensa. Ao procurar por *linux*, por exemplo, podemos excluir todas as figuras do simpático pinguim mascote do sistema preenchendo esse espaço com a palavra *tux*.

Para criar papéis de parede são necessárias imagens com alta resolução. Ao fazer uma busca no Google procurando, por exemplo, por *linux*, pode ser que encontremos desde imagens gigantescas até ícones de 2 KB.

Para resolver esse problema, a busca avançada de imagens disponibiliza diversos filtros. Em **Tamanho**, por exemplo, é possível escolher o comprimento das imagens que se deseja utilizar. Optando por *grande*, teremos figuras em alta resolução, ideais para papel de parede, aproveitamento em slides etc..

O ideal para papéis de parede é utilizar figuras JPG: uma figura destas, em alta resolução, ocupa algumas dezenas de KB, ao passo que seu equivalente em BMP pode ocupar alguns MB. Para retornar figuras somente em JPG ou PNG, a pesquisa avançada utiliza a caixa **Tipos de arquivo**. Se a procura é por imagens em tons de cinza ou preto e branco, é possível utilizar o filtro *Coloração*.

Para finalizar, podemos conseguir todas as imagens de um determinado domínio. Ao digitar, por exemplo, o domínio *digerati.com.br*, conseguiremos todas as imagens contidas na página inicial desse domínio.

Localizando figuras ocultas

Apesar de pesados, às vezes pode ser necessário conseguir alguns arquivos BMP. Embora não esteja presente no filtro *Tipos de arquivo*, podemos fazer buscas por essa extensão. Basta digitar o tema da busca – por exemplo, *brazil* – acrescentando, na caixa de verificação, o termo *filetype:bmp*. Todos os arquivos com essa extensão e relacionados ao Brasil serão listados pelo diretório de imagens do Google.

Buscas avançadas de textos

Existem vários tipos de buscas avançadas também para textos. Além das buscas tradicionais por termos e páginas Web relacionadas a eles, é possível procurar por documentos, planilhas do Excel, arquivos PDF, títulos de páginas, entre muitas outras coisas.

Essas opções, assim como as encontradas na busca avançada de imagens, são acessíveis através de um link um pouco escondido, o **Pesquisa avançada**, que está ao lado do logotipo do Google ou da caixa de busca, em outras páginas.



A área **Procurar resultados** é semelhante à área de mesmo nome localizada na busca de imagens e, como ela, substitui a utilização de parâmetros booleanos por uma interface gráfica. Você pode acompanhar a explicação de seu funcionamento no tópico anterior.

É claro que, além dessas, existem diversas outras ferramentas que podem ser utilizadas para a realização de busca de informações em bancos de dados. Esses recursos não utilizam os delimitadores booleanos “tradicionais”, mas algumas implementações da própria equipe de desenvolvedores do Google.

Idioma

Em **Idioma**, podemos juntar aos elementos de busca booleana tradicionais a possibilidade de se realizar a busca somente em uma determinada língua. Se, por exemplo, quisermos executar uma busca sobre memórias flash de 32 MB, muito usadas em câmeras digitais baratas, computadores de bolso e celulares, podemos simplesmente fazer uma busca em **Procurar resultados em todas as palavras** utilizando o termo em inglês, *flash card 32 MB*, já que o idioma da Rainha pode ser considerado a língua “oficiosa” (não oficial, mas reconhecida por todos como tal) da Web.

Suponhamos agora que queremos saber quanto custam essas memórias em Taiwan ou Hong Kong (principalmente para comparar com os preços que o contrabandista local passa para você). Para fazer a busca com esse parâmetro, vá ao campo **Idioma** e, em **Exibir páginas escritas em**, escolha **chinês (simplificado)**.

Após clicar em **Pesquisa Google**, a pesquisa mostra um imenso número de páginas da China (você pode ver os caracteres *mandarim* no título em cada página). O diferencial é que, ao escrever os termos em inglês, você foi direcionado para páginas com informações nos dois idiomas e se livrou de ter de aprender chinês em uma hora. Esse procedimento é válido para todos os idiomas e pode ser útil também para conseguir e-books em outros idiomas, digitando *books* na busca e escolhendo o idioma desejado, como grego ou espanhol, por exemplo.

Formato de arquivo

Quem nunca deixou para procurar dados muito específicos que precisavam ser inseridos em uma planilha ou gráfico na última hora, ou precisou de uma apresentação em PowerPoint para ser “customizada” (um eufemismo para a cópia de estrutura ou conteúdo que outros fizeram) que acuse o vizinho ou atire a primeira pedra. Principalmente com a possibilidade de se fazer isso pelo Google, de uma maneira muito especializada.

Usando a linha de pesquisa **Formato do arquivo**, o Google é capaz de realizar uma filtragem buscando somente determinados formatos de arquivo. Esses formatos podem ser escolhidos na segunda caixa da linha e compreendem os tipos de documentos que mais são disponibilizados na Web, afora as páginas em HTML. Essa busca inclui, além de arquivos do Power Point e Excel, documentos PDF (que são o padrão atual de documentação digital na Internet) e arquivos em texto que podem ser editados pela Web, como arquivos DOC ou RTF.

Suponhamos, por exemplo, que você deseje fazer uma busca por planilhas de Excel que detalhem gastos com combustível em uma pequena empresa. Além de inserir os termos *gastos* e *combustível* na busca, vá à linha **Formato de arquivo**, selecione a opção **Apenas** e, em seguida, o tipo de documento que deseja buscar – planilhas XLS, no nosso caso. Ao clicar em **Pesquisar**, você encontrará todas as planilhas compartilhadas na Internet de que precisa.

Um aviso: por descuido de administradores de sistema ao implementar políticas de compartilhamento de arquivos, muitas planilhas ou documentos publicados na Web contêm informações que não são de domínio público. Portanto, cui-

dado ao utilizar planilhas, textos ou apresentações colhidas no Google como base para seus próprios documentos: você pode entrar em uma “fria”. Se, por exemplo, ao buscar a melhor maneira de criar uma página de telefones para a Intranet (rede de informações interna) do seu departamento, você busca por telefones e documentos RTF e acaba esbarrando em um documento do mês passado, com listas de telefones de todos os juízes federais do Brasil, deixe o documento onde está. Ou melhor: apague todos os cookies e o Histórico de seu browser, como se você nem tivesse utilizado a Internet naquele dia.

Documentos expostos

Você não sabe, mas seus documentos podem estar expostos no Google. Passar por uma situação destas, aliás, é muito fácil: é só manter algumas configurações de rede e Internet do Windows inalteradas após sua instalação.

Antes de tudo, compartilhar arquivos no Windows é muito perigoso, já que ele não faz nenhuma distinção entre compartilhar um diretório para sua rede interna e para toda a Web – principalmente se o compartilhamento está em uma máquina com acesso direto, tanto discado quanto de banda larga, à Internet. Mas, se não há outro jeito, o ideal é deixar o compartilhamento o mais seguro possível.

No Windows 98/ME, recomenda-se atribuir proteção com senha para os diretórios. Para isso, clique com o botão direito do mouse no diretório escolhido, clicando, em seguida, em **Compartilhamento**. Logo após, clique na opção **Depende da Senha** para as opções de compartilhamento. O Windows XP, infelizmente, não funciona dessa maneira. Para habilitar senhas nos computadores com este sistema operacional, é necessário, antes de tudo, mudar algumas configurações. Para isso, vá ao menu **Iniciar > Windows Explorer**, menu **Ferramentas > Opções de pasta** e desabilite a opção **Usar compartilhamento simples de arquivo (recomendável)**. Agora basta compartilhar o arquivo e selecionar a quantidade limite de usuários e também quais usuários terão permissão para utilizá-los.

Se quiser criar um usuário específico, acesse o **Painel de controle**, clique no ícone **Usuários** e, em seguida, clique em **Criar Usuário**. Crie um usuário convidado – por exemplo, *Rede*. Os serviços de compartilhamento do Windows 2000 (bem melhores do que os do XP) funcionam de maneira semelhante.

Data

O tempo, depois do dinheiro, é um dos fatores que mais pesam no mundo moderno. Além de fazer tudo muito depressa, às vezes temos de nos ater ao que foi feito por último, ou ao que há de mais recente sobre determinado assunto, deixando tudo o que é “velho” de lado.

Nesse ponto, o conteúdo da Internet possui algumas falhas: tendo quase uma década de existência pública, a Web conseguiu acumular, ao lado de conteúdo atualizado diariamente, velharias de seus inícios. Isso acontece porque muitos servidores (principalmente servidores públicos) jamais são limpos ou têm revisão de seu conteúdo. Isso faz com que, ao fazermos buscas simples – por exemplo, por *Homem-Aranha* –, encontremos muita coisa do segundo filme, lançado recentemente, ao lado de notícias de lançamento do primeiro filme, de dois anos atrás.

O Google possui uma ferramenta que permite a filtragem. Apesar de não possuir opções para fazer busca por artigos ou páginas de datas definidas (o que acreditamos, no entanto, que não demorará a ocorrer), esse filtro pode facilitar muito a vida do usuário, principalmente se ele quiser restringir a linha do tempo porque ele próprio não tem tempo!

Assim, procurando por *cotação do dólar paralelo* na página de **Pesquisa avançada**, podemos ir à linha **Data** e clicar em **nos últimos 3 meses** para obter as informações mais recentes sobre esse tema. Também é possível conseguir informações dos **últimos 6 meses** e do **último ano**.

Existem combinações interessantes: juntando a busca *cotação do dólar paralelo* por data à busca por tipo de arquivo – por exemplo, somente planilhas do Excel –, é possível adquirir um comparativo de flutuação do dólar nos últimos três meses, prontinho para usar!

INDICADORES ECONÔMICOS GERAIS	ÍNDICES OU VALORES	VARIAÇÃO(%)			BASE
		MES	ANO	12 MESES	
Índice	307,815	1,40	5,54	7,87	Ago94=100
Índice	310,1520	1,31	5,33	7,04	Ago94=100
Índice	280,9020	0,71	2,86	5,00	Ago94=100
Índice	178,9900	0,43	2,25	4,56	Jun96=100
Índice	242,7722	0,57	1,83	4,45	Jul94=100
Índice	2290,8000	0,51	2,75	5,15	Dez93=100
Índice	2248,3700	0,54	2,76	5,01	Dez93=100
SALÁRIO MÍNIMO	260,0000	0,00	8,33	8,33	R\$ (REAL)
Índice	EXTINTA*	-	-	-	R\$ (REAL)
DÓLAR OFICIAL (1)	3,1291	6,26	6,30	5,51	R\$ (REAL)
DÓLAR PARALELO (1)	3,1700	5,87	4,82	2,59	R\$ (REAL)
Taxa	1,0201	0,146	0,59	3,02	TAXA (R\$)
Índice	2378,9000	0,40	2,63	4,09	Dez93=100

Fontes: FGV, IBGE, DIEESE, FPE, PEACAPIM e BANCO CENTRAL DO BRASIL.
 Elaboração: BANCO DE DADOS CENC.
 (1) Valor de venda, cotação do último dia útil do mês.
 (*) Evolução MP. IRTSMT de 27/10/2000.
 (...) Dado não disponível.

Ocorrências

O Google assume que existem várias maneiras de se enunciar um tema em uma página Web. Você já deve ter notado isso ao perceber que o Google retorna uma busca simples com links tanto para conteúdos de páginas quanto para títulos de sites ou nome pelo qual a página é hospedada no servidor. Basta pesquisar por termos cotidianos, como por exemplo, *tortas de limão*.

Existem maneiras, contudo, de dividir essa pesquisa por localização. Na página *Página avançada*, o Google possui a ferramenta **Ocorrências**, que permite a busca tanto por elementos pertencentes à própria página, como seu título ou conteú-

do, quanto por maneiras como o servidor em que a página está hospedada, ou outros servidores da Internet, a enxergam. Assim, procurar por um local que venda tortas de limão fica muito mais fácil (procure por *Ocorrências, no título da página*), a não ser que você prefira fazer a receita em casa (*Ocorrências, no corpo da página*).

Domínio

Alguns domínios (endereço universal dado a uma página da Internet, e exclusivamente a ela) fazem jus a esse nome, conseguindo ser tão vastos que, percorrê-los link a link em busca de uma informação pode, tranqüilamente, levar alguns anos. Que o digam os softwares de download, bem sortidos, mas nem sempre muito organizados.

Pense, por exemplo, que você precisa de uma versão recente do Linux, como o Gentoo. Buscar por *Gentoo* ou *Linux* no Google trará tantos resultados que, ao encontrar uma página que sirva, provavelmente a distribuição não servirá mais (nem você, se passou o horário de trabalho procurando-a).

Para resolver esse problema, o Google disponibiliza a ferramenta **Pesquisa avançada, Domínio**. No caso acima, basta clicar na opção **Apenas** da linha **Domínio**, digitar o endereço www.isolinux.org (o maior repositório de sistemas baseados em Unix que pode existir). Assim, conseguiríamos não uma só, mas muitas versões do sistema.

Fica claro que poderíamos fazer o caminho contrário, substituindo a opção **Apenas** por **Não exibir resultados do site ou domínio**. Esse tipo de busca pode ser utilizado para fugir de lugares-comuns ou perigosos, evitando que uma pesquisa por *cartas* caia no site dos correios ou em uma página dedicada a cassinos.

Similares no Google

No capítulo anterior, mostramos várias opções de busca avançada do Google utilizando operadores booleanos diversos. Mas esse não é o único padrão de buscas reconhecido por essa página: o Google também é capaz de reconhecer similares e até encontrar um “atalho” para uma página dentro de uma outra página.

Vejam, por exemplo, o recurso **Similares**, localizado na página de **Pesquisa avançada**, na área **Pesquisa por página específica**. Pesquisar por meio desta ferramenta permite o encontro de páginas que podem ter conteúdo parecido ao de um site ou domínio vasculhado à exaustão, no qual não encontramos o que estávamos procurando. Se, por exemplo, procuramos livros publicados pela Digerati e não conseguimos encontrá-los em uma determinada livraria, basta ir à linha **Similares, Pesquisar páginas semelhante** a digitando, em seguida, o nome da livraria. Surgirá uma lista de endereços com a mesma função ou assunto do termo procurado.

Esse recurso também pode ser utilizado por pessoas que estiverem em viagem pelo exterior ou que necessitem de informações sobre um determinado serviço fora do Brasil. Se, por exemplo, digitamos nessa linha de pesquisa o endereço www.correios.com.br, o Google retornará os endereços de diversos serviços de correio do mundo, além de serviços de entrega de encomendas e comércio alfandegário.

Rastreando links

A maioria dos métodos de busca do Google procura por informações diretas. Vimos isso ao procurar informações sobre data, nomes no cabeçalho da página, tipos de arquivos etc..

Mas, além disso, o Google também é capaz de buscar por informações indiretas, vasculhando a Web, página por página, até encontrar uma que faça referência a uma determinada página, em vez de acessar a página diretamente. Ao fazer isso o buscador faz duas coisas interessantes: primeiro, utiliza o Back Rub – que, como vimos, foi o primeiro mecanismo de busca do Google. Em segundo lugar, em vez de vasculhar links com seus *spiders* até encontrar a página desejada, o Google faz o trabalho “pela metade”, limitando-se a criar uma listagem dos locais onde o site foi encontrado.

Esse recurso pode ser utilizado preenchendo a linha **Links**, no campo **Encontrar páginas com link para a página**. Ao procurar, por exemplo, por links para o site da Receita Federal (www.receita.fazenda.gov.br), teremos acesso a diversos endereços que remetem a esta fonte. Outra dica interessante: se você é Web designer, pode utilizar esse recurso para monitorar a inserção de links para sua página em sites de terceiros. Afinal de contas, se você cuida de uma página sobre jóias, não será nada agradável ver seu link colocado em uma página sobre segurança patrimonial ou estatísticas de assaltos.

O GOOGLE FORA DO COMPUTADOR

Olhando de fora, o Google é um universo a parte, como seu próprio nome denuncia (veja o quadro no primeiro capítulo para saber o que significa o termo *google*). Apesar de fechado no mundo da rede e dos computadores, o Google poderia, com uma dose razoável de ficção científica, ser encarado como uma entidade, se não real, ao menos em formação, no melhor estilo de *Matrix*. Afinal de contas, as viagens virtuais desse buscador por todo o mundo acabam desembocando em lugares reais (servidores administrados por pessoas que trabalham em prédios, depois vão para casa, onde moram outras pessoas que também devem usar o Google...), são utilizadas por pessoas reais (que interagem com coisas mais tangíveis que o Google) e que por fim, acabam procurando por informações que podem ser a respeito de coisas reais, como outras pessoas, casas, mercadorias etc..

Pensar sobre isso pode parecer complicado e abstrato – afinal de contas, fora da tela do computador, não temos nenhuma notícia da existência do Google. Precisamos ter um PC ou um notebook, conectado à Internet, quietinho em algum lugar, para que o Google se faça presente, para não dizer real. Isso até começar a surgir a informática móvel...

O Google no celular

O celular foi um grande avanço nas comunicações. Imagine só: após quase um século presas às linhas de telefonia fixas, as pessoas finalmente podiam, a um custo relativamente baixo, ser encontradas em qualquer lugar, não só do seu bairro ou da cidade onde vivem, mas do mundo inteiro (tudo bem, isso tem alguns inconvenientes, principalmente se você deve dinheiro a alguém). Um colega de trabalho viajou para o exterior e você precisa falar com ele com urgência? Se ele levou consigo o celular (da empresa, é claro!), basta ligar para tirar suas dúvidas (da empresa, é claro...).

Em 1999, um consórcio de grandes empresas de telefonia, sediado na Europa, teve uma idéia inusitada: e se, além de ser usado como telefone, o celular também pudesse ser utilizado para buscar ou receber informações, como uma espécie de computador de bolso ligado à Internet? Foi assim que nasceu o WAP.

WAP

O WAP ou *Wireless Access Protocol* (Protocolo de Acesso sem Fio) é uma tecnologia que permite que dispositivos que não utilizem fios, como celulares ou computadores de bolso antigos, como os velhos *Palm Pilot*, acessem informações localizadas em servidores de e-mail, listas de contatos em servidores corporativos e sites na Internet. Para que qualquer uma dessas modalidades de serviço torne-se compatível com o WAP, é necessário criar um servidor WAP e páginas utilizando a linguagem WML (*Wireless Markup Language*), que é o seu equivalente do HTML utilizado em sites comuns.

Devido à lentidão dos modelos de celulares do final dos anos 90 e à escassez de serviços, o WAP acabou desaparecendo, mas abriu caminho para o GSM e outras formas de acesso sem fio mais sofisticadas.

Acompanhando uma enxurrada de serviços de e-mail para WAP (a maioria, na verdade, existe até hoje, e, por incrível que pareça, muita gente ainda usa WAP para trabalhar com correio eletrônico), o Google também resolveu lançar seu buscador nos celulares, possibilitando que mais de 4 milhões de resultados pudessem ser acessados usando os teclados numéricos. Com o “Google de celular”, pode-se fazer tudo o que o Google faz em um computador, e talvez um pouco mais.



Acessar o Google via celular é muito fácil:

1 – Acesse o browser WAP de seu telefone e conecte-se à Internet.

2 – No campo **Go To URL/Website** ou **Ir** (se seu navegador estiver em português), digite www.466453.com (o endereço do Google em conexões wireless).

3 – A página do Google que irá surgir chama-se *Google Number Search*. Afinal, você deve utilizar o teclado numérico do celular para “digitar” o termo que deve ser buscado. Utilize 0 para espaços e 1 para parênteses.

Além do WAP, existe outro serviço que permite o acesso ao Google via celular e que pode ser utilizado tanto em celulares de 2,5G (os famosos GSM, que tiram fotos, fazem downloads de música e que as operadoras conseguem vender mais do que as montadoras vendem PCs) quanto em outros modelos mais antigos, mas compatíveis com o acesso à Web. A maneira de realizar o acesso é ligeiramente diferente:



1 – Faça a conexão com o seu provedor (para acessar esse tipo de serviço, você deve ter um). Logo em seguida, acesse o menu do telefone móvel.

2 – Selecione a opção **Acesso à web** ou o equivalente nas opções do menu. Pressione o botão **Aceitar**.

3 – Insira a palavra *Google* quando for pedida a inserção de uma URL. Se o browser não aceitar *Google* como endereço, digite www.google.com/wml. Clique em **OK**.



E os palmtops?

Os palmtops (computadores de mão) são substitutos modernos tanto dos animais de estimação quanto dos gênios da lâmpada. Até quem achava ser “frescura” ter um Palm ou PocketPC e escrever com a inefável canetinha, hoje se rendeu aos encantos desse bicho que serve para jogar, lembrar nomes de pessoas e recados e, é claro, passear pela Internet. Para realizar essa última façanha, você deve ter um modem (ou um *smartphone*, aparelho que é celular e palmtop ao mesmo tempo), uma conta de provedor (pode ser até um gratuito, como o Yahoo ou o IG) e, por fim, um browser.

O passo seguinte é difícilíssimo: conectado à Web, digite o endereço www.google.com/palm. A montagem do Google em um Palm III com 8 MB de RAM é quase instantânea.

Pequenos brilhantes

Com exceção dos pockets A-20 e E-100 da Casio e de todos os modelos de Palm até o Palm IIIe, a maioria dos palmtops já vem acompanhada de um browser para acesso à Internet. No quesito aplicativos agregados (softwares que já acompanham o sistema operacional) temos boas ferramentas, como o tradicional *Pocket Internet Explorer*, da Microsoft, muito parecido com a versão 4.0 desse navegador, que acompanhava os discos de instalação do Windows 95. Para os usuários do Palm (principalmente Palm Zire) a opção de acesso fica por conta do WebPro, que possibilita, inclusive, a configuração de um proxy.

Se seu palmtop ou *smartphone* não vem com um browser, existem diversas opções, gratuitas e leves (menos de 300 KB) para que você possa se conectar com o Google em qualquer lugar. Para Windows CE e PocketPC, existem o Pocket Browser (<http://www.conduits.com/products/launcher/PocketLauncher.exe>) e o AvantGo (<https://my.avantgo.com/home/index.html?referer=http%3A%2F%2Fwww.avantgo.com%2Ffrontdoor%2Findex.html>), que pode ser encontrado tanto para Windows CE, PocketPC e celulares Symbian quanto para Palm OS. Para este último, existem, inclusive, versões do AvantGo bem mais leves do que as utilizadas nos PocketsPC e que podem ser utilizadas até em um Palm III.

CALENDÁRIOS

Computadores lidam com números. Mesmo que as pesquisas sobre Inteligência Artificial, tanto cognitiva quanto intuitiva, dêem algum resultado, algum dia, o fato de que as máquinas digitais se alimentam de números, assim como as máquinas analógicas vivem do girar das manivelas ou engrenagens, nunca vai mudar.

Máquinas inteligentes

As pesquisas com IA (Inteligência Artificial) foram iniciadas, como a maioria das pesquisas computacionais em larga escala, no final da década de 40, após o término da 2ª Guerra Mundial. Esses esforços iniciais tratavam, principalmente, de medir a reação de computadores e processadores de dados em determinadas situações, como a divisão de números extensos ou o fracionamento de bancos de dados escritos com números binários (formados por seqüências de 0 e 1, os únicos algarismos/valores que o computador entende).

Pesquisas de Inteligência Artificial voltadas para aspectos cognitivos (fazer com que as máquinas aprendam coisas novas a partir de novos elementos ou da correção de erros) ou intuitiva (fazer com que as máquinas se aproximem o máximo possível do pensamento humano, “intuindo” partes que faltam de um problema ou resolvendo um erro de interpretação) só foram possíveis muito tempo depois, com a criação da Lógica Difusa (ou Lógica Fuzzy), que para alguns autores é o marco inicial da Inteligência Computacional por tratar, também, das incertezas e da resolução de problemas incompletos.

Para conhecer mais sobre métodos e projetos de Inteligência Artificial, acesse <http://www.ic.unicamp.br/~ra030014/ic/ia/>.

As pessoas manipulam informações de maneira parecida com a das máquinas, se bem que reconhecem mais valores do que simplesmente 0 e 1 (bom, na verdade, nem todas têm essa capacidade). Para se situar no espaço em que vivem, as pessoas usam números para identificar suas casas, apartamentos, mostrar a que distância está sua casa ou cidade de um determinado lugar etc..

Para se situar no tempo, por sua vez, também utilizam números para representar a contagem de dias, horas e minutos – eles próprios elementos matemáticos. Para, por exemplo, fazer a contagem de horas, pode-se utilizar o relógio, dividido em 12 períodos exatamente iguais.

Para contar períodos de tempo (anos, séculos) mais extensos, inventou-se o calendário, uma representação lógica e gráfica ao mesmo tempo, para “armazenar” os cálculos astronômicos necessários para medir o início e o fim de um ano ou estação do ano (o cálculo não era – e não é – muito simples). Os maias, por exemplo, “imprimiam” seus calendários (um ano de 365 dias, com ciclos de 13 anos para adição de 1 dia suplementar, agrupados, por sua vez, em ciclos maiores, de 52 anos) em folhas feitas de pele humana ou grandes discos de pedra.

O calendário juliano e o Google

Atualmente, todo o mundo, ao menos comercialmente (excetuando-se a Igreja Ortodoxa, na Rússia, e os países muçulmanos, por motivos religiosos) utilizam o calendário gregoriano, que recebe esse nome por ter sido promulgado pelo papa Gregório XIII, em 1682. O calendário gregoriano suprimiu dez dias do calendário anterior – o calendário juliano, inventado por Julio César e que já tinha 1600 anos de atividade. Além disso, organizou os anos bissextos, fazendo com que eles só ocorram quando as duas primeiras cifras são divisíveis por quatro. De acordo com esta norma, os anos de 1600 e 2000 são bissextos, enquanto os anos de 1700, 1800 e 1900 não possuem esse dia adicional.

O calendário juliano tem uma duração exata de 365.25 dias, um arredondamento em relação à duração do ano solar ou equinocial (365.2422). Apesar de mais fácil de calcular (essa era a intenção de Julio César ao codificá-lo), esse arredondamento faz com que o ano perca um dia a cada 128 anos. Com isso, do início de sua vigência até sua substituição pelo calendário gregoriano, o modelo juliano acumulou uma perda de dez dias na contagem dos anos.

Por que falamos tanto de números e cálculos de calendário? Bom, o Google é um conjunto de máquinas (sistema digital que trabalha com números), que roda um software de busca (um software é escrito em linguagem de programação, que nada mais é do que uma ferramenta de transformação de conceitos humanos em linguagem de máquina – 0 e 1), que, por sua vez, indexa páginas localizadas na Web.

Como vimos no **Capítulo 2**, o Google é capaz de procurar uma informação na Internet de diferentes formas. Dentre essas formas, existe a possibilidade de se utilizar a opção **Data** na página **Pesquisas Avançadas** para encontrar informações dos últimos três ou seis meses ou um ano.

Essa opção, contudo, não permite buscas por valores maiores que um ano, nem menores que três meses, nem períodos de tempo com valores intermediários (por exemplo, quatro meses). Para realizar esse tipo de busca, o Google disponibiliza um outro serviço, que é utilizado por meio da área de busca. Pelo comando `datarange`, é possível fazer buscas cronológicas específicas – incluindo dia de publicação da página na Web. A sintaxe do comando é:

```
<termo a ser pesquisado> datarange=<data inicial> -  
<data final>
```

E é aí que está o segredo da coisa. Boa parte do Google é feita em Python, uma linguagem de programação livre e extremamente flexível. Não se sabe bem por que – talvez pelo fato de os números inteiros serem mais fáceis de elencar e processar – o pessoal do Google adotou o calendário juliano como padrão. O script de busca de data do Google (o código-fonte do Google não está disponível para verificação) deve ser muito parecido com o encontrado no box a seguir.

Pensando como o Google

Enquanto o código-fonte do Google não é liberado – existem notícias, e temos esperança de que isso em breve aconteça), podemos imitar ao menos uma das funções do Google: a conversão de datas. O código abaixo foi escrito por Scott Moore e permite a conversão de qualquer data do calendário gregoriano para o juliano.

```
baseConstant = 2440588 (onde 2440588 foi utilizado  
com base de cálculo)  
if theDate == 'now':  
    (year, month, date, _, _, _, _, _, _) = time.localtime()  
    (aqui define-se a entrada de dados no formato de  
    data do calendário gregoriano)  
    theTarget = time.mktime((year, month, date, 19, 0, 0,  
    0, 0, -1)) (aqui define-se a saída de dados no  
    formato juliano)  
    theTargetInDays = int(theTarget / (60 * 60 * 24)) - 1  
    (aqui define-se o formato de saída da base de dados)  
    return theTargetInDays + baseConstant
```

Para “rodar o programa”, basta baixar qualquer editor de código que aceite essa linguagem de programação (você pode baixar o pacote oficial do Python para Windows em <http://www.python.org/ftp/python/2.3.4/Python-2.3.4.exe>) ou utilizar um editor como o Ultraedit (<ftp://ultraedit.com/uedit32.zip>).

Como fazer as pesquisas

Como fazer, então, uma pesquisa por data no Google? Se você não é astrônomo, monge ou simplesmente mais um pedante de plantão, provavelmente não saberá realizar a conversão de datas “de cabeça” – o que, aliás, coloca-o no rol das pessoas sadias e normais.

Felizmente existem várias ferramentas na própria Internet que resolvem esse problema. Uma delas é o site <http://www.measurementsconverter.com/calendars.html>, um endereço no qual é possível fazer a conversão de qualquer data gregoriana para o calendário “antigo”:



1 – Após acessar a página, digite, ao lado da caixa **Month Number**, o número do mês que deseja consultar (1 correspondendo a janeiro, 12, a dezembro). Na caixa **Year**, digite o ano que deseja consultar, no formato de quatro dígitos (por exemplo, 1999, 2004 etc.). Em seguida, clique no botão **Calculate**.

2 – Você terá uma tabela com todos os dias daquele mês, incluindo em que dia da semana eles “caíram”. Nas colunas *JD*, *MJD* e *TJD*, você terá informações relacionadas ao calendário juliano. O formato utilizado pelo Google para as pesquisas é o da coluna *JD*. Assim, 1º de dezembro de 2003 corresponde ao número 52974.

Week	Month	Day	Year	JD	MJD	TJD
Saturday	May	1	2004	2453126	2453126	2453126
Sunday	May	2	2004	2453127	2453127	2453127
Monday	May	3	2004	2453128	2453128	2453128
Tuesday	May	4	2004	2453129	2453129	2453129
Wednesday	May	5	2004	2453130	2453130	2453130
Thursday	May	6	2004	2453131	2453131	2453131
Friday	May	7	2004	2453132	2453132	2453132

3 – Abra a página do Google e digite na barra de buscas o termo de busca e a pesquisa de data, assim:



DATAS E HORÁRIOS NA INTERNET

Como vimos no capítulo anterior, o Google pesquisa por datas exatas de artigos e páginas da Internet, utilizando o calendário juliano como padrão. Além disso, é possível descobrir qualquer data utilizando a Internet. Na página http://www.quediahoje.net/calendario_permanente.htm, por exemplo, existe uma tabela algorítmica que permite saber em que dia da semana caiu uma determinada data.

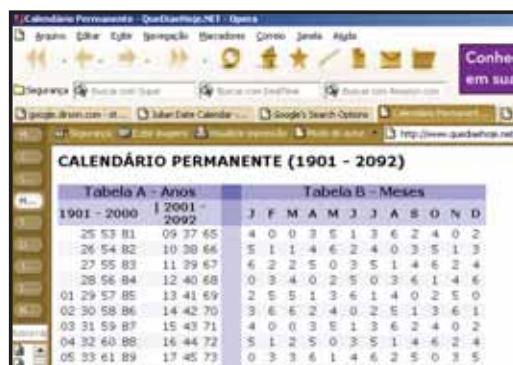


Tabela A - Anos		Tabela B - Meses											
1901 - 2000	2001 - 2092	J	F	M	A	M	J	J	A	S	O	N	D
25 53 81	09 77 65	4	0	0	3	5	1	3	6	2	4	0	2
26 54 82	10 78 66	5	1	1	4	6	2	4	0	3	5	1	3
27 55 83	11 79 67	6	2	2	5	0	3	5	1	4	6	2	4
28 56 84	12 80 68	0	3	4	0	2	5	0	3	6	1	4	6
01 29 57 85	13 81 69	2	5	5	1	3	6	1	4	0	2	5	0
02 30 58 86	14 82 70	3	6	6	2	4	0	2	5	1	3	6	1
03 31 59 87	15 83 71	4	0	0	3	5	1	3	6	2	4	0	2
04 32 60 88	16 84 72	5	1	2	5	0	3	5	1	4	6	2	4
05 33 61 89	17 85 73	0	3	3	6	1	4	6	2	5	0	3	5

Fazer o cálculo com essa tabela é bem simples: suponha que a companhia telefônica está fazendo uma cobrança retroativa de uma conta que deveria ter sido paga no dia 12 de dezembro de 2003, mas foi paga somente no dia 15, no caixa do banco. Você não se lembra bem porque não pagou a conta, mas se recorda de que tem algo a ver com um feriado de sua cidade, que cai justamente no dia 12.

Na verdade existem três tabelas para chegarmos ao resultado final. Primeiramente, procure na **Tabela A** o ano de 2003. A **Tabela B** possui 12 colunas, cada uma correspondendo a um mês. Siga a mesma linha do ano de 2003, mas passando para a **Tabela B**, parando no mês de dezembro. Ao número encontrado (1), adicione o número do dia em questão (12).

Teremos o número 13 como resultado. A partir daí, observe na **Tabela C** que dia corresponde ao número 13. Observamos que dia 12 de dezembro foi uma sexta-feira. Por isso você não pagou sua conta: foi um feriado prolongado...

Operações com datas

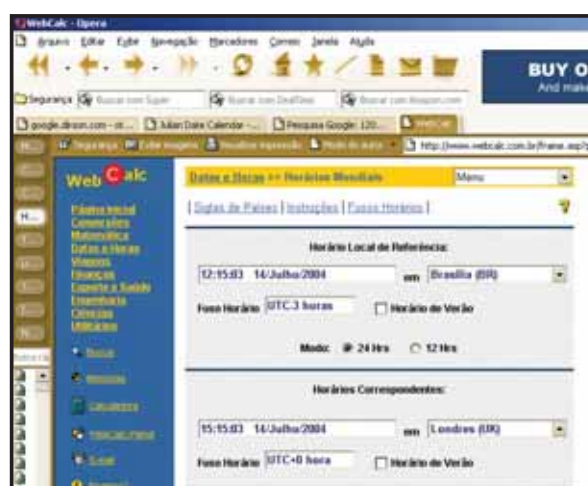
Suponha que você trabalha em uma importadora e precisa saber que horas são em Paris e em Berlim para enviar alguns faxes pela Internet, mas não faz a mínima idéia de como fazer esse cálculo.

Apesar de não ser possível realizar cálculos com datas no Google, existem diversas páginas na Internet que realizam esse serviço gratuitamente, inclusive permitindo que seja impresso um calendário personalizado. Uma delas é a WebCalc (<http://www.webcalc.com.br/frame.asp?pag=http://www.webcalc.com.br/datas/horarios.html>), totalmente em português.

Acesse a WebCalc. Em **Horário Local de Referência**, mantenha as informações do jeito que estão, a não ser que você esteja no Mato Grosso do Sul, na

Amazônia ou em Belém do Pará. Para o caso de o País estar sob vigência do horário de verão, clique na caixa de verificação **Horário de Verão**. Mantenha o modo *24 Hrs* selecionado, já que ele é o padrão utilizado em transações comerciais.

Nos quadros abaixo, é possível localizar a hora atual correspondente em qualquer país do mundo (incluindo os minúsculos Andorra, na Europa, e Samoa Ocidental, na Austrália). Você pode configurar as horas para até cinco relógios simultâneos.



Configure os quadros com as localidades desejadas e não se esqueça de indicar se elas utilizam ou não o horário de verão. As datas serão ajustadas automaticamente.

Observe que as diferenças de fuso horário são iguais para Paris e Berlin – exatas quatro horas. Se você teve curiosidade de observar o fuso horário de Andorra, viu que corresponde ao fuso horário francês: isso acontece porque os dois países são vizinhos muito próximos.

Para que saber de Andorra?

Andorra é um principado (uma espécie de monarquia parlamentar sem rei) situado no interior dos Pirineus Orientais; é limitado a norte pela França e a sul pela Espanha. O país é governado por uma câmara parlamentar formada por homens de Andorra e também pelo presidente da França e pelo bispo de Urgel, que têm atribuições reduzidas em relação à importância de seus cargos no passado, e poderes bem definidos pela Constituição de 1993 – primeira e única do país.

Andorra tem 64 mil habitantes, espalhados pela extensão territorial de 468 km². Apesar desse “aperto” e do tamanho reduzido, talvez interesse muito ir a Andorra. Desde 1950, o país chegou a ser considerado duas vezes a nação com maior crescimento econômico do mundo; já no principado, existem mais de 5 mil lojas e 500 hotéis que garantem aos habitantes uma renda per capita superior à alemã e à japonesa.

Se você deseja adicionar essa página aos Favoritos, antes de tudo, clique em *Salvar Padrão* dentro da própria página. Em seguida, clique em **Favoritos > Adicionar a Favoritos** (se você utiliza o Internet Explorer) ou **Marcadores > Adicionar página aqui** (se você usa o Opera).

Criando um calendário personalizado

O WebCalc também permite a criação de calendários personalizados. Isso é uma boa saída, primeiro, e antes de tudo, porque é de graça. Segundo, porque imprimir os calendários do Outlook ou do Word não é uma tarefa fácil, e nem sempre apresenta bons resultados.

Para começar, acesse o início da página e vá ao **Menu** clicando, em seguida, em **Calendários**.



É possível montar calendários por meses e por ano. Suponhamos, por exemplo, que queiramos um calendário do mês de janeiro de 2005. Em **Calendário Mensal**, clique em **Calendário do Mês** (selecione o mês) **de** (selecione o ano). Clique, em seguida, em **OK**. Agora, basta clicar em **Imprimir**.



O mesmo pode ser feito com um calendário anual. Experimentamos pedir a criação de um calendário do ano 3128, achando que o script “abriria o bico”. Para isso clicamos em **Calendário do ano**, e digitamos 3128. Veja o resultado:



Incluindo datas especiais

A personalização do calendário pode ser feita na área **Inclusão de Datas Especiais**. Na primeira linha, **Data**, basta selecionar o dia que resolveu escolher como especial.



1 – Na linha **Incluir como**, é possível selecionar essa data como um feriado (o WebCalc só reconhece feriados nacionais, portanto, feriados religiosos, municipais ou estaduais ficam de fora).

Se quiser adicionar uma descrição à data, preencha o campo **Descrição da Data**. Por exemplo, é possível adicionar o dia 25 de janeiro como feriado, descrevendo-o como aniversário de São Paulo, ou descrever a **Data Especial** de 10 de janeiro como aniversário da empresa em que você trabalha.

2 – Clique em **Incluir Data** para adicionar a data criada por você em um calendário mensal ou anual.

3 – É possível criar calendários setorizados. Se você precisa de um calendário específico para o setor de contabilidade (pagamentos, vencimentos, leitura de livro caixa etc.), clique no botão **Limpar Tudo** no campo **Inclusão de Datas Especiais**. Em seguida, preencha as datas especiais como desejar.

4 – No campo **Título do Calendário para Impressão**, digite o nome que você deseja dar ao calendário, como *Contabilidade Julho/2004*, e, em **Calendário Mensal**, clique em **OK**.

DOCUMENTOS DINÂMICOS COM O GOOGLE

Copiar e Colar, Ctrl + C e Ctrl + V, opção **Colar Item da Área de Transferência...** Depois do Ctrl + Alt + Del (conjunto de teclas que desliga o computador à força quando o Windows encenra), provavelmente não há comando mais utilizado do que o que permite que informações sejam movidas entre pastas, diretórios e até documentos diferentes.

É claro que essas ferramentas conseguem ser especialmente úteis quando são recortados/copiados dados estáticos, plenamente organizados, e que quase não precisam de retrabalho. É por isso que copiar dados da Internet pode ser tão complicado: nem sempre os problemas com a formatação da página ou a possibilidade de autorizar os dados fazem com que o trabalho valha a pena.

Pedindo socorro ao Google

O Google oferece, ao menos indiretamente, uma solução interessante para os usuários do Microsoft Office: a possibilidade de se procurar arquivos semelhantes aos que estão sendo utilizados, principalmente documentos do Word e planilhas de Excel. Essa ferramenta do Google permite, por exemplo, que procuremos uma planilha na Web relacionada à cotação do Euro e adicionemos valores encontrados nela a uma planilha, localizada em nosso computador.



1 – No Google, digite na barra de busca a expressão **euro**, adicionando o parâmetro *filetype:xls* (documentos do Microsoft Excel. Se quiser ter dados mais recentes, vá a página de Pesquisas Avançadas e, no filtro **Data**, escolha a opção **Nos últimos 3 meses**):



2 – Clique sobre uma planilha e dê uma olhada nos dados, certificando-se de que são os dados de que você necessita. Feito o reconhecimento, volte para a página do Google, clique com o botão direito do mouse sobre o título do link e, em seguida, na opção **Copiar Endereço**.



3 – Abra a sua planilha de trabalho (a que está salva em sua máquina) e clique na(s) célula(s) nas quais deseja inserir o conteúdo da planilha online. Crie uma fórmula que obedeça aos seguintes parâmetros:

= 'http://www.alguembr.homepage.com/[arquivo.xls]
Planilha1'!A1

No caso, estamos trabalhando com uma pasta simples (sem muitas planilhas em seu interior). Portanto, teríamos o seguinte resultado:

	30.09.03	30.09.02	Var. (%)	31.12.02
1 Balanço	345.067,1	335.474,5	2,86	324.208,1
2 Ativo total	168.935,1	164.342,8	2,79	162.973,0
3 Operações de crédito (líquido)	317.231,4	308.735,3	2,75	304.893,0
4 Recursos de clientes administrados	211.375,4	217.729,9	(2,92)	211.555,1
5 No balanço	105.656,0	91.005,4	16,32	93.337,9
6 Fora do balanço	19.256,5	10.000,0	1,20	17.594,2
7 Património líquido	450.923,2	426.479,9	5,73	417.546,0
8 Total fundos administrados				

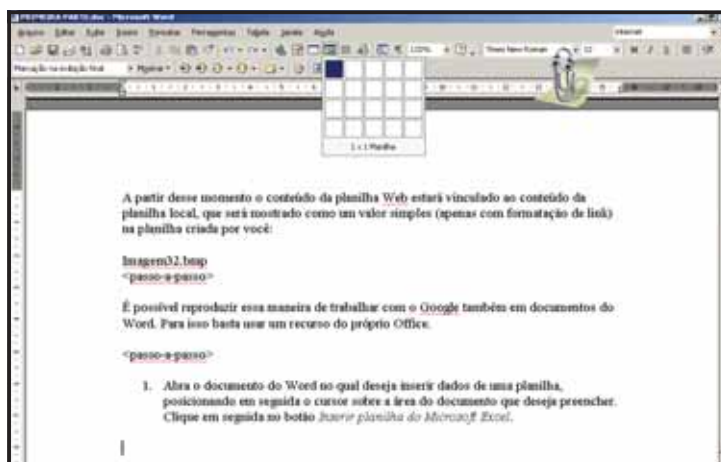
A partir desse momento, o conteúdo da planilha Web estará vinculado ao conteúdo da planilha local, que será mostrado como um valor simples (apenas com formatação de link) na planilha criada por você:

	30.09.03	30.09.02	Var. (%)	31.12.02
1 Balanço	345.067,1	335.474,5	2,86	324.208,1
2 Ativo total	168.935,1	164.342,8	2,79	162.973,0
3 Operações de crédito (líquido)	317.231,4	308.735,3	2,75	304.893,0
4 Recursos de clientes administrados	211.375,4	217.729,9	(2,92)	211.555,1
5 No balanço	105.656,0	91.005,4	16,32	93.337,9
6 Fora do balanço	19.256,5	10.000,0	1,20	17.594,2
7 Património líquido	450.923,2	426.479,9	5,73	417.546,0
8 Total fundos administrados				



É possível reproduzir essa maneira de trabalhar com o Google também em documentos do Word. Para isso, basta usar um recurso do próprio Office.

1 – Abra o documento do Word no qual deseja inserir dados de uma planilha, posicionando, em seguida, o cursor sobre a área do documento que deseja preencher. Clique no botão **Inserir planilha do Microsoft Excel**. Selecione uma única célula para desenhar a planilha.



2 – Copie o link da planilha escolhida no Google nesta célula e, em seguida, formate-a como se estivesse fazendo-o no Excel:

= 'http://www.alguembr.homepage [arquivo.xls] Planilha1' !A1

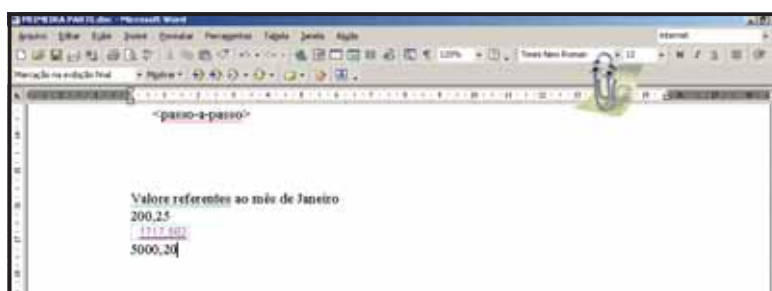
3 – Clique na área de trabalho do Word. O valor se transformará em parte do documento e, ao mesmo tempo, será atualizado toda vez que a planilha original (na Internet) também for.

Valores referentes ao mês de janeiro:

200,25

1717,582

5000,20



COLOCANDO SEU SITE EM PRIMEIRO LUGAR NO GOOGLE

Nos últimos capítulos, abordamos maneiras básicas de utilizar o Google (e outros recursos da Internet) para facilitar sua vida profissional. Este capítulo é um pouco diferente: vamos mostrar uma maneira de usar o Google como instrumento de marketing pessoal.

Se você tem um site pessoal ou comercial e o coloca na Internet, você deseja ser visto ou deseja que seu trabalho o seja. E não há lugar mais democrático – e barato – para fazer isso do que a Internet; afora o pagamento de recursos de hospedagem (veja o quadro a seguir para serviços de hospedagem gratuitos), via de regra ela oferece serviços de e-mail personalizado, transferência de arquivos e contagem de páginas. Se você tem um site hospedado por mais de um ano em um servidor, também sabe que mantê-lo no ar não significa sucesso de audiência – existem casos de páginas que ficam mais de um ano no ar e recebem menos de uma visita por dia. Certificado de fracasso em um universo de mais de 120 milhões de internautas...

E é aí que o Google entra. A partir de sua consolidação, em 2000, estar listado no Google tornou-se um dos melhores meios de propaganda que uma página, empresa ou pessoa (é, tem gente que gosta de aparecer, principalmente depois do advento dos blogs) pode ter. Estar na primeira página de pesquisa que o buscador retorna (a maioria das pessoas, segundo algumas pesquisas, só tem paciência para procurar até a terceira página...) já equivale a ganhar um Prêmio Nobel.

Provedores gratuitos

Às vezes, você pode até acreditar em suas idéias, mas está mesmo sem um centavo furado no bolso. Neste caso, além de não querer gastar com o Google, você provavelmente não irá querer gastar com um provedor de acesso.

Existem várias alternativas de provedores gratuitos na Web. Oferecendo espaços em disco razoáveis (em média 50 MB) e a possibilidade de se utilizar recursos em Flash e ASP, esses provedores são um ótimo paliativo para a falta de recursos:

7Mares

<http://setemares.terravista.pt/>

Provedor português que aceita páginas cadastradas do Brasil. Hospedagem gratuita de sites com espaço de até 12 MB. Possui um assistente automático de criação de páginas.

POP Sites

<http://www.pop.com.br/popsites/>

Hospedagem gratuita com 50 MB de espaço. Você deve utilizar o provedor POP e seu discador para ter direito a essa conta.

Tripod Brasil

<http://www.tripod.com.br>

Oferece assistente para construção de sites e 12 MB de espaço.

Ubbi Home Pages

<http://homepages.ubbi.com.br>

Hospedagem gratuita de sites com até 20 MB. Também permite a hospedagem de álbuns online.

Yahoo! GeoCities

<http://br.geocities.yahoo.com>

Oferece espaço de 15 MB para hospedagem gratuita de sites, assistente para construção, editor de HTML, transferência de arquivos por FTP e gerenciador de arquivos. E-mails do Yahoo! podem ser utilizados para compor o sistema de comunicação da página.

Free Servers

<http://www.freeservers.com>

Serviço internacional de hospedagens gratuitas. Oferece 12 MB de espaço em disco, taxa de transferência de 500 MB, além de gerenciador de arquivos personalizado.

Como tornar-se um campeão

Na verdade, existem duas maneiras de tornar-se líder do Google em sua categoria: uma delas é convencional, rápida, mas dói no bolso; a outra utiliza ferramentas que quase todos conhecem, mas têm preguiça de usar – redes de relacionamento e matemática.

A primeira delas é proporcionada pelo próprio Google. Há cerca de dois anos, o Google disponibilizou um serviço de anúncios em links em suas páginas. Mas, para evitar a polemica, evitou “rifar” os primeiros lugares das buscas para os sites pagos, como costumavam fazer alguns buscadores: os anúncios do Google ficam localizados em uma área destacada da página (a conhecida interface “limpa” do Google permite isso), mas que não se mistura com os resultados tradicionais de busca.



1 – Este serviço é chamado de AdWords e pode ser contratado na página inicial do Google. Clique em **Soluções de publicidade**. Na página Soluções de publicidade do Google, clique no link **Inscriva-se ou saiba mais....**

2 – Na página Google AdWords, clique no botão **Clique para Iniciar**. Na primeira etapa, você deve configurar seu anúncio. Na coluna *a.*, é possível escolher com que idiomas se deseja iniciar o serviço. Se o seu negócio é local (você só pretende atender a uma cidade, região ou Estado do Brasil), selecione apenas a opção **Português**. Se quiser selecionar mais idiomas para seu anúncio, mantenha a tecla Ctrl apertada enquanto clica em outras opções.



3 – Na coluna *b.*, selecione o plano de marketing que você deseja adquirir: existem desde planos globais, cobrindo o mundo inteiro (é claro que, em um caso como esse, você deveria ter escolhido inglês ou espanhol como idioma padrão), até planos personalizados, que atingirão apenas usuários localizados a uma distância específica da sua empresa. Após escolher sua opção, clique em **Salvar e continuar**.



4 – No nosso caso, escolhemos a opção **Personalizada**. Nela, podemos definir a localização de nosso negócio clicando em **Usar o endereço físico da empresa**. A página será expandida e serão mostrados diversos campos de informação que deverão ser preenchidos.



5 – No tópico *b.* A que distância desse endereço estão localizados os seus **clientes-alvo?**, você digitará a distância na qual deseja captar sua audiência usando o Google. Escolhemos a distância de 50 km, maior do que o perímetro da maioria dos municípios brasileiros. Clique em **Salvar e continuar**.



6 – Na próxima tela, o Google pede sua localização exata, tanto em latitude quanto em longitude – uma informação não muito simples de ser obtida. Para nos auxiliar, vamos utilizar uma excelente página de itinerários mundiais, o Maporama (<http://www.maporama.com/share/>).

7 – Acesse esse site sem fechar o seu cadastro do Google. Na área **Mapas** do site, preencha os dados de sua localização e clique no botão **Go**.

8 – Surgirá o mapa com a sua localização. Observe, no canto inferior esquerdo da tela, a coluna **Informações**. Na primeira linha, você encontrará sua posição exata.



9 – Copie os números (não precisa se preocupar com os sinais de coordenadas nem de graus) nos campos **Latitude** e **Longitude** e clique em **Salvar e continuar**.

10 – Agora, você deve situar seu anúncio em um determinado grupo de anúncios. Se você escolher, por exemplo, consultoria de hardware, todos os usuários que procurarem por hardware ou consultoria verão seu anúncio listado na coluna de anúncios pagos. Na linha **A. Criar anúncios**, você irá compor seu anúncio, ao mesmo tempo em que poderá visualizar o resultado final na caixa branca acima do texto. Clique em **Criar anúncio e continuar**.

11 – Na próxima etapa, você deve escolher as palavras-chave para seu anúncio. Utilize termos que tenham a ver com o tema de sua página ou negócio. Inclua também sinônimos ou correlatos (para uma página de Linux, por exemplo, seja esperto e também inclua Windows e XP) além de nomes grafados erroneamente (por exemplo, adicione a *linuxm*, os termos *linus* e *linx*). Clique em **Salvar palavra-chave**.



12 – Está quase acabando... Agora, você deve fazer seu lance pelos serviços do Google. Clique na moeda com a qual quer pagar pelos serviços (no nosso caso, pobres mortais, Reais do Brasil (BRL)). Em seguida, preencha o valor que você deseja oferecer ao Google como custo máximo por clique. O valor mínimo é de R\$ 0,20 (vinte centavos de Real), além dos R\$ 20,00 que serão pagos por sua inclusão na lista. Quanto maior for sua oferta, mais você aparecerá no Google. Clique em **Salvar e continuar...**

13 – Agora, você deve especificar seu orçamento diário. Se você digitar R\$ 1,00 (1 real), o Google mostrará seu anúncio cinco vezes ao dia. Em um mês, portanto, você gastará R\$ 30,00 para aparecer, no máximo, 150 vezes no buscador. Clique em **Salvar e continuar...**



De graça

Se você não achou a barganha muito vantajosa, ou acha que sua idéia não vale a retirada de algumas notas do bolso, existem outras maneiras de se tornar visível ao mundo através do Google. Para praticá-las, você deverá ter alguns amigos, além de um vocabulário um pouco acima da média.

1 – Escolha um termo ou expressão para associar ao seu site. Você pode utilizá-lo tanto na página quanto no link (o ideal é utilizar em ambos). Escolha palavras não muito comuns para não ter concorrência. Se a sua página é sobre sistemas operacionais, escolha algo como NetBSD, “usuário falso” ou “DVI Editor”.

2 – Agora, acione seus amigos. Peça para que eles coloquem em suas páginas as palavras escolhidas por você, além do link de sua página. Por exemplo, associe sua página ao link www.linux.usuariofalso.com.br.

3 – Poste notícias sobre sua página também para listas de discussão e fóruns, não se esquecendo de utilizar o link de sua página, além das palavras-chave.

4 – Espere um tempo (entre um e dois meses) e faça algumas pesquisas no Google para ver se sua página já foi incorporada ao termo. Se você ainda não aparece em primeiro lugar e você tem uma página comercial, tente postar seu endereço para páginas que já estão em primeiro no ranking, como o www.mercadolivre.com.br, www.primeiramao.com.br ou www.classificados.com.br.

Crime e castigo

Além da maneira citada acima (a que, comprovadamente, mais funciona) existem diversos outros recursos para “passar a perna” no Google. Eis alguns deles:

Texto oculto

Mais parecida com artimanhas de piratas e espiões do que com um truque de informática, essa técnica consiste em encher sua página de textos formatados com a mesma cor do fundo da tela, de forma que vários blocos de conteúdo tornem-se invisíveis ao olho humano. Letras minúsculas, escondidas em caixas de cores berrantes que parecem simples erros de diagramação ou animações em Flash com listas de palavras passando em grande velocidade também podem ser utilizadas.

Cross-linking

Suponhamos que você possui diversas páginas, como www.zegoiabab.com.br e www.clubedogoiabinha.com.br, entre outros endereços edificantes e essenciais para o desenvolvimento da cultura humana. Basta colocar, em suas páginas, links para as demais e o trabalho está feito – isso é chamado cross-linking. Se você registrou uma mesma página com vários links redirecionando para ela no Registro.BR, melhor ainda: você setará diversas vezes o mesmo endereço de uma vez só.

Guestbook spamming

Essa técnica consiste em assinar todos os livros de visitas que você possa encontrar. Já que muitos desses livros permitem a adição de endereços Web a comentários, essa é uma maneira rápida de conseguir links sem pedi-los a ninguém – nem mesmo ao webmaster da página em que o livro está.

Fazendinha de palavras

Neste recurso, o webmaster cria uma página, mais ou menos oculta, em seu site, na qual adiciona um verdadeiro dicionário de palavras. Ao criar um site sobre Linux, por exemplo, ele criará uma página html com todos os comandos utilizados na linha de comando desse sistema operacional. Assim, qualquer busca sobre Linux e o respectivo comando cairão em sua página.

É legal?

Depois do crime, chegou a hora do castigo: não, não é legal. Ao perceber que seu site alcança posições de destaque no Google utilizando uma dessas táticas, um usuário descontente ou concorrente poderá denunciá-lo no link <http://www.google.com/contact/spamreport.html>. O Google costuma examinar as denúncias de “picaretagem” e banir os mal-esportistas de sua busca.

A VOLTA AO MUNDO EM 80 DIAS COM O GOOGLE

Quando Julio Verne concebeu seu personagem Phineas Fogg, o protagonista de *A Volta ao Mundo em 80 Dias* (1864), não fazia ainda 30 anos que os barcos a vapor começavam sua carreira comercial, tão cheia de acidentes e atrasos. Os aviões só viriam a ser criados mais de meio século depois, e sua utilização comercial só seria possível após 1920. Como realizar, então, uma volta pelo globo terrestre inteiro, em exatos 80 dias?

Quem leu o livro sabe que Fogg, juntamente com seu criado francês Passepartout, utiliza-se de barcos, jangadas, ferrovias e até de um elefante comprado por uma fortuna para realizar a travessia da Europa até a África, dali até a Ásia e a América, para, por fim, chegar ao destino um dia antes do previsto. O leitor deve se lembrar também de que na viagem ocorrem diversos desencontros, apesar de Fogg dispor de toda a sua fortuna pessoal para realizar sua façanha: falta de hospedagem, tentativas de roubo, atrasos nos transportes e ainda foi confundido com um ladrão de bancos em fuga, por um policial que faz o possível para atrapalhar sua viagem e apanhá-lo.

Uma viagem com o Google

Se na época em que ocorreu a aventura Phineas Fogg possuísse um palmtop e uma conexão à Internet, provavelmente faria sua viagem em muito menos tempo e teria muito menos dissabores. Mesmo sem aviões, trens-bala ou automóveis, bastaria um PDA com o Google para que a viagem se tornasse muito mais fácil. É claro que não teríamos a metade das façanhas e situações eletrizantes do livro, mas isso é uma outra história.

Falta de hospedagem

Mesmo com a utilização de telefones, fazer reservas em um hotel costuma ser uma situação traumática. Alguns administradores de portaria ou gerentes de hotel parecem recusar-se a hospedar pessoas que escolhem seu hotel e têm dinheiro para pagar... Imagine agora que você dispõe de apenas algumas horas para descansar entre uma etapa e outra de sua viagem e descobre que a reserva simplesmente não foi feita; você está em uma cidade que não conhece e simplesmente não tem onde ficar.

Com nosso PDA com o Google, isso seria muito simples. Procurando pelos termos *hotéis* e *mapas*, encontramos o site www.maporama.com, do qual já falamos antes. Essa página é especializada em conseguir mapas, endereços, nomes de ruas, reservas de hotel e aeroportos, não só no Brasil, mas em todo o mundo.

Na área de **Mapas** do site, vamos escolher nosso país de destino. Pode ser a China, por exemplo, e, no nome da cidade, digite *Hong Kong*.



Teremos imediatamente um mapa da cidade. Se conhecêssemos alguma coisa de Hong Kong – um nome de rua, de beco, um CEP – teríamos um mapa específico da região indicada, incluindo hotéis, bares e restaurantes.

Com a página que temos à mão, contudo, já podemos fazer muita coisa. Ao lado do mapa da cidade, temos acesso a uma área chamada **Localizar**. Logo ao lado dessa área temos acesso a vários serviços que nos permitem localizar hotéis na região. Vamos escolher o primeiro, o *Hotels.com*. Selecione-o e clique no botão **Go**.



Veremos automaticamente com quantos hotéis poderemos contar em Hong Kong. Em seguida, uma lista logo abaixo nos levará para os links das páginas dos respectivos hotéis. Agora, basta escolher um hotel, ter um cartão de crédito à mão e fazer a reserva. É claro que realizar reservas em dois hotéis ou refazer a reserva via telefone ou fax também é recomendável.

Procurado

Como dissemos, Fogg tinha um policial – o agente Fisk, da Scotland Yard – em seus calcanhares, mas só vem a saber disso no final da viagem, quando quase perde a aposta por ficar um dia inteiro retido na cadeia até que se verificasse o engano.

Se Fogg (ou até o agente Fisk) possuísse um PDA com a página do Google, esse incidente nunca teria acontecido: Fogg realizaria a viagem tranquilamente, e

Fisk não faria papel de idiota diante de seus superiores, tentando justificar a fortuna que gastou para capturar um homem que, afinal, era inocente. Para o herói do livro, bastaria acessar a página <http://www.met.police.uk/mostwanted>, na qual são listados os homens mais procurados pela Scotland Yard. Depois de levar um susto por ver seu rosto ali ao lado de gente, digamos, não muito simpática, Fogg poderia, com um simples telefonema, limpar seu bom nome e seguir viagem.



Procurando na página informativa para policiais da Inglaterra http://www.met.police.uk/appeals/intro_pages/wanted_current.htm, Fisk saberia, de pronto, que outros policiais estavam na pista do ladrão de bancos que ele procurava e que, alguns dias antes de seu retorno a Londres, o assaltante já havia sido preso.

Segunda Parte:

Google Hack



VOCÊ ESTÁ NO GOOGLE!

O GOOGLE E O ACESSO A INFORMAÇÕES PESSOAIS E PROFISSIONAIS

Números de documentos são as informações mais bem guardadas do mundo real, já que, em alguns sistemas de identificação, como o brasileiro ou o americano, a numeração de certos documentos equivale a ter a posse das impressões digitais ou do rosto de seu proprietário, ou seja, um certificado de autenticidade de que você é quem diz ser.

Ao realizar compras em uma loja no Brasil, por exemplo, você deve apresentar sua Cédula de Identidade (ou RG – Registro Geral de Identificação) ao lado do seu CPF (Cadastro de Pessoa Física). Teoricamente, um prova que você existe, o outro prova que você está apto a comprar. A prática não é bem essa: na verdade, o RG é aceito apenas como um documento “formal”; a emissão de cédulas de identidade é responsabilidade dos Estados da União e de suas respectivas Secretarias de Segurança Pública (SSP), as quais realizam a identificação das pessoas:

- a) pela apresentação do Registro de Nascimento (Certidão de Nascimento);
- b) pela identificação datiloscópica (impressões digitais).

Ocorre, porém, que essas instituições – até o momento – não possuem bancos de dados unificados: o que faz com que os casos de cidadãos com duas ou mais cédulas de identidade, com diferentes nomes, não sejam muito raros.

Já os registros do CPF são centralizados pela Receita Federal, um órgão que está presente em todos os Estados brasileiros. Isso faz com que esse documento seja dificilmente falsificado (se bem que existem crimes desse tipo) e seja, via de regra, reconhecido como prova da identidade de seu portador.

Números de documentos na Web

Na Internet, a maioria dos cadastros é habilitada com a apresentação (digitação) de um número de CPF. No caso dos sites de compras, costuma ser necessário apresentar o número do cartão de crédito, seguido de certos dados pessoais. Para a maioria dos sistemas online de controle de compras ou crédito, basta ter a posse desses dados para assumir uma identidade qualquer.

A partir de 2003, houve um aumento de 275% na utilização da Internet para dar golpes utilizando senhas de contas bancárias e números de cartão de crédito (fonte: <http://www.certisign.com.br>). Além disso, as maneiras de se obter dados tornaram-se mais diversificadas, indo desde “pescaria” de números no Google até a utilização de e-mails falsos de bancos pedindo a confirmação de dados pessoais, também conhecida como *fishing scan*.

O sistema de busca do Google funciona bem... bem demais! Nada que estiver online, mesmo que esteja escondido em várias camadas de subdiretórios, passa despercebido pelo Google. Já mostramos que basta que documentos estejam localizados em compartilhamentos desprotegidos para que eles se tornem públicos na Web – e reclamar por autorização ou pela privacidade na Internet não adianta muita coisa, principalmente se o descuido é seu ou de gente em quem você confiou para guardar suas informações pessoais.

Um hacker que se utilize do Google ou de outros mecanismos de busca à procura de informações pessoais está em busca de documentos compartilhados por engano na Web: planilhas, tabelas, documentos do Word, bancos de dados, todas essas modalidades de arquivo possuem algum tipo de informação relevante. Assim, ele utilizará a **Pesquisa Avançada**, linha **Formato de Arquivo**, para encontrar, por exemplo, planilhas do Excel, associando-as aos termos *cartão*, *crédito* e *despesas*. Para informações mais diretas, um hacker poderia limitar sua busca a datas de até três meses atrás.

Após conseguir uma lista de planilhas que possam conter números de cartão de crédito ou outras informações, ele se depara com um campo fértil: uma planilha localizada no FTP de um escritório de contabilidade.

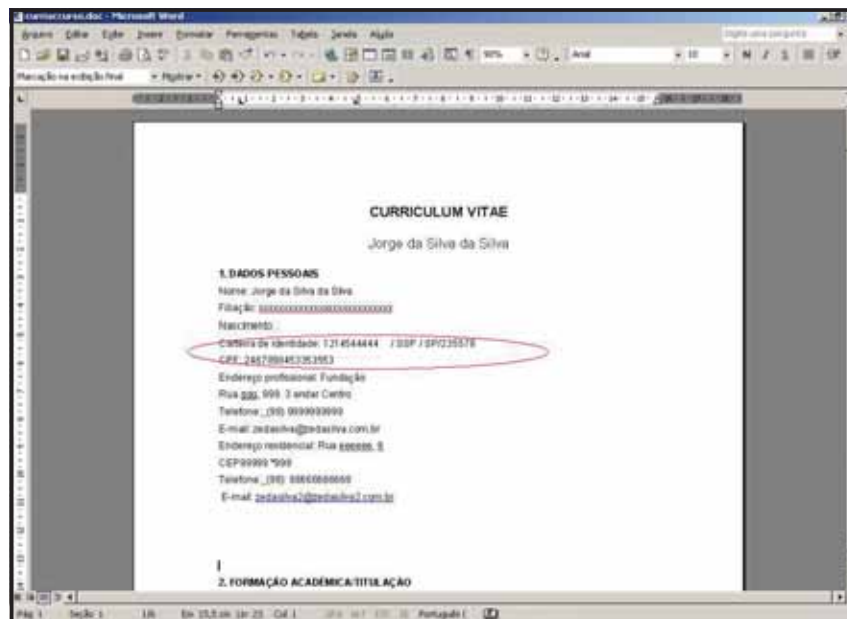
Se não tiver sorte com a primeira planilha, o criminoso já tem uma pista interessante: um FTP totalmente aberto, cheio de informações. E é aí que ele encontra a planilha *Loja de Goiabada.xls*, na qual estão listadas as seguintes informações:

Período	Julho	Agosto	Setembro	Outubro
Despesas Financeiras Liquidas	2.045,00	880,00	1.112,00	44,00
Despesas Bancárias	72,00	72,00	72,00	72,00
Taxa Adm. Cartão Visa Parcelado	240,00	240,00	240,00	240,00
Taxa Adm. Cartão AMEXX	170,00	170,00	170,00	170,00
Taxa Adm. Cartão Credencial	290,00	290,00	290,00	290,00
Taxa Adm. Cartão Bônus	25,00	25,00	25,00	25,00
Cartão de Crédito VISA 2222222222	444,00	444,00	4.000,00	4.000,00
Cartão de Crédito VISA 999913456	622,00	444,00	66,00	6.556,00
Imposto de Renda	33,00	33,00	33,00	33,00
CPMF	187,00	187,00	187,00	187,00
Taxa de Adm. Visa Electron	61,00	61,00	61,00	61,00
REFFI	1,00	1,00	1,00	1,00
Juros Recebidos				
Outras Receitas Operacionais				
Resultado do Exercício				

Primeiro, o hacker terá dois números de cartão de crédito, um pertencente a cada operadora desse serviço. Ao mesmo tempo, ele perceberá valores muito altos em um mesmo mês, que, provavelmente, são correspondentes aos limites dos cartões.

The collage features several documents. On the left, a newspaper clipping from the 'Giant' is visible, with the headline 'Giant' and the sub-headline 'Giant' and 'Giant'. The main text of the clipping is partially obscured but includes the words 'Giant' and 'Giant'. On the right, there is a 'Giant' logo with the word 'Giant' in a stylized font. Below the logo, there is a small text block that reads 'Giant' and 'Giant'.

Assim, ao procurar, na **Pesquisa Avançada**, por arquivos .doc, associados aos termos *currículo*, *Jose*, *Silva*, *CPF*, pode-se encontrar um documento como esse:



Depois de duas buscas pelo Google (e um pouco de engenharia social ao falar com o gerente), um ladrão teria todas as informações necessárias para fechar uma conta bancária e transferir todos os valores para sua conta usando apenas o telefone ou o e-mail. E não seria muito fácil capturá-lo, já que a maioria das contas utilizadas para transferência de valores roubados são frias ou são alugadas de terceiros.

O USO DE CARACTERES CORINGA

Nem todos os documentos estão expostos na Internet com tanto despreendimento: alguns gerentes de tecnologia, secretárias e copeiros (quando não, alguém que exerce as três funções) têm o bom senso de esconder informações confidenciais, tornando-as, se não inacessíveis, ao menos não tão vulneráveis.

Aliás, essa é uma recomendação típica para administradores de sistema, infelizmente nem sempre seguida: arquivos de senhas jamais devem se chamar **logins.doc** ou **senhas.txt**; procedimentos de login de usuário e política de senhas devem ser repassados para o usuário oralmente, e não disponibilizados via e-mail ou documentos compartilhados na rede; por último, abandonando a economia que sai caro, as empresas deveriam possuir dois servidores: um para serviços online (FTP, webmail, página institucional) e outro para o armazenamento de informações confidenciais da empresa.

Descobrimo o que está oculto

Pois lá vamos nós novamente com o Google, esse instrumento de procura de vulnerabilidades e serviços mal-implantados pelo administrador de sistemas. O buscador retorna resultados com base nos termos pedidos, de acordo com os parâmetros ditados: a procura por *senhas*, em documentos ou planilhas, resultará em links que obedeçam exatamente a esses parâmetros. Ocorre que, assim como nos sistemas de busca de códigos de programação e aplicativos em geral (o próprio Office da Microsoft se utiliza desse sistema), existe uma função que aumenta o alcance da busca, ao mesmo tempo em que possibilita a busca de vários termos simultaneamente: os caracteres coringa.

Na maioria dos sistemas, tanto ? (ponto de interrogação) quanto * (asterisco) são caracteres coringa. No Google, apenas o asterisco pode ser utilizado, mas ele funciona praticamente como um canivete suíço: todas as funções dos outros caracteres coringas podem ser implementadas usando o *.

O asterisco (*) serve como um substituto de qualquer palavra, seja ela qual for: ao fazer uma busca, por exemplo, por *curriculum Jose * filetype:doc*, um hacker pesquisador teria os seguintes retornos:



O mesmo vale para validar uma expressão. Suponhamos que um hacker procure pela expressão “5124*1” no Google, que poderia ser um código de validação bancária. A inclusão do asterisco na expressão entre aspas permitiria a captura de todas as páginas em que o código de validação bancária 5142 fosse acrescido de um número de 0 a 9, terminando com 1. O asterisco também poderia possibilitar a descoberta de um possível código especial utilizando letras do alfabeto.



O mesmo truque vale para descobrir contas bancárias em diversos bancos. Ao digitar a expressão “*contas correntes banco **”, o hacker teria como resultado a listagem de todo o conteúdo Web relacionado a contas bancárias de vários bancos.

Para não esquecer dos administradores de sistema

Vale lembrar que a culpa dessas vulnerabilidades e “furos” não é do Google: ele é apenas um instrumento que pode ser utilizado para explorar falhas do sistema operacional, do planejamento de TI ou do próprio administrador de sistemas. É como a velha história da faca: o médico pode usá-la para salvar o paciente operando-o ou, se for louco, para esfaqueá-lo. E é pensando assim que vemos que o Google pode ajudar muito o administrador de sistemas, principalmente por meio dos caracteres coringa.

Um problema muito comum enfrentado pelos membros dessa laboriosa classe é a vigilância do uso que funcionários costumam fazer de recursos da empresa, principalmente e-mails e Internet. Apesar da utilização de serviços de monitoração, como proxys e serviços de log, essa documentação pode ser (e normalmente é) contestada pelos usuários. A solução, então, é procurar os “fujões” com a ajuda do Google – afinal seria ridículo por parte do usuário alegar que o administrador, além dos logs, fraudou conteúdo localizado em servidores externos só para prejudicá-lo (se bem que tal fraude pode, com um pouco de trabalho, ser realizada).

Suponhamos, por exemplo, que o administrador pretende “cortar” a postagem de e-mails para listas de discussão no horário de trabalho, utilizando os e-mails da empresa, mas seus logs são contestados pelos usuários, alegando que eles podem ter sido manipulados. Além disso, “a pescaria pode ser muito maior do que o pescador imagina”, já que funcionários em movimento (laptops, handhelds, *home officers*) também deveriam, dependendo da política de uso da empresa, participar das mesmas sanções – é claro, se fosse possível recolher logs.

Antes de tudo, nosso administrador deveria procurar por contas de e-mail da empresa, perdidas em listas de discussão. Com a ajuda de caracteres curinga, bastaria, por exemplo, fazer a busca via Google utilizando o termo *@digerati.com.br.



Existem diversas outras combinações: fazer a busca de usuários associada a tipos específicos de arquivos (**filetype:doc**, **filetype:xls** etc.), por exemplo, pode revelar desde espões até o manuseio descuidado de documentos confidenciais. A utilização de nomes de usuários, associada à busca de diretórios de servidores “suspeitos” como os utilizados para compartilhamento de MP3 (**\pub\mp3**, **\pub\musik**, **\pub\music**), por sua vez, é capaz de descobrir desde o destino do limite de banda contratado pela empresa até o fato de que alguns servidores podem estar sendo usados, internamente, para compartilhar MP3 com redes peer-to-peer.

Ilegalidade

Apesar de ainda não oficialmente proibido no Brasil, o compartilhamento de arquivos, principalmente músicas e imagens de terceiros, protegidas, portanto, pela Lei de Direitos Autorais, pode tornar-se um problema para empresas nas quais os funcionários se utilizam desses serviços. Na Inglaterra, Nova Zelândia e Austrália (pátria do KaZaA), de um ano para cá, já se tornou prática comum à justiça local vasculhar servidores de empresas privadas em busca de MP3 e filmes ilegais armazenados em servidores Web ou locais e sempre achar alguma coisa...

DESCOBRINDO DONOS DE SITES

A combinação de caracteres curinga – aliada à possibilidade de buscar diretórios que, de um jeito ou de outro, podem trazer informações relacionadas à administração de um servidor –, também pode ser utilizada como ferramenta para a invasão de páginas. Neste caso, a culpa das vulnerabilidades não é nem do Google, nem do administrador de sistemas (pelo menos não toda a culpa).

Aqui, o grande culpado é o FrontPage, o software de editoração de páginas Web da Microsoft. Para páginas simples, como as utilizadas em redes internas de trabalho (Intranets) ou páginas institucionais, o FrontPage pode até ser uma boa solução, já que sua operação é tão simples quanto a do restante dos programas do Microsoft Office (no qual ele vem incluso).

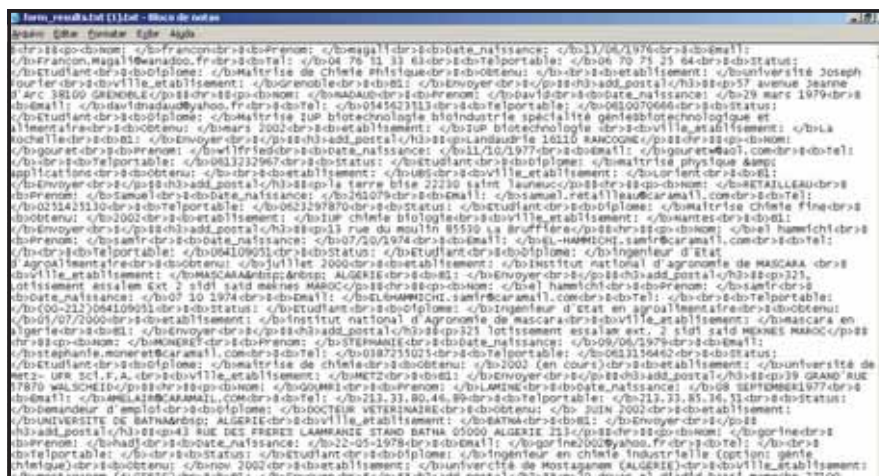
O problema começa quando o usuário utiliza as ferramentas de administração do FrontPage para publicar seu site. Não que elas sejam difíceis de utilizar – pelo contrário. A questão é que, em segurança, já aprendemos que a facilidade pode ser ilusória, e dez minutos ganhos em configurações automáticas podem significar horas perdidas não muito mais tarde.

Veja com os seus próprios olhos. O FrontPage permite adicionar formulários em seu site para capturar um feedback de usuário, e você deve incluir campos para perguntar o nome, endereço, número de telefone, entre outras informações. Por convenção, essas informações são armazenadas em um local denominado **called_private\form_results.txt**, na árvore de diretórios do site criada pelo FrontPage. Atenção: não confunda o funcionamento e a organização do FrontPage com a administração de páginas de outro servidor da Microsoft, o IIS (Internet Information Service), que tem *outras vulnerabilidades* (aliás, muitas), mas não essa.

O arquivo **_private** deveria ter suas permissões configuradas para limitar os acessos a seus dados, mas, algumas vezes, as pessoas se esquecem de fazer isso. E é esse esquecimento que torna alguns diretórios vulneráveis: para ver como isso acontece, visite o Google e entre com a expressão *inurl: "form_results.txt"* na caixa de busca. O resultado será esse:



Você vai encontrar vários sites expostos. Muitos são sites de teste, mas olhe para as páginas com mais de 1 KB de tamanho e você encontrará algumas que exibem os nomes, números de telefone e endereços de e-mail dos seus visitantes.



Outros comandos de busca podem mostrar senhas para quadros de mensagem ou fóruns. Digitando, por exemplo, *inurl: "auth_user_file.txt"* é fácil encontrar arquivos contendo listas de usuários, senha, contra-senha e diretórios de login. Os arquivos são mais ou menos parecidos com este:

A área de texto sublinhada é, na verdade, a senha de acesso ao servidor, que está, porém, criptografada. O que não é empecilho para um cracker ou script kiddie por dois motivos:

2 – Um kiddie, mais paciente, mas também mais inepto, baixaria o arquivo para sua máquina e utilizaria um software de *brute force*, como o Brut_v_24, que pode ser encontrado juntamente com outras ferramentas de teste de senhas em http://www.filescenter.com/navega0/name/_SECA2%20-%20Seca%20Media%20guard/PROGRAMAS%20-%20MOSC.html.

Acabe sozinho com seu e-commerce

Com a ajuda do FrontPage, empresas de e-commerce sem administradores de segurança ou pequenas lojas online também podem causar sua própria falência, com um auxílio mínimo dos crackers utilizadores do Google. Novamente, basta configurar um servidor FrontPage sem atribuir permissões aos diretórios.

Um cracker vasculharia essas informações procurando pelo termo *inurl:"orders.txt"*, e o resultado seria impressionante: é muito fácil encontrar arquivos de texto de mais de 2 MB, reproduzindo bancos de dados com números de cartão de crédito, endereços e senhas de validação de compras. Ao procurar por mais informações, porém, o Google também poderia ajudá-lo: procurar pelo termo *"Index of /admin"* equivale a penetrar no mundo dos servidores desprotegidos e colocados no ar em um dia por equipes apressadas. Uma busca rápida retornou nada mais do que 3.600 resultados...

Outras combinações possíveis são:

"Index of /" +passwd
"Index of /" +password.txt
"Index of /" +.htaccess
index of ftp +.mdb allinurl:/cgi-bin/ +mailto
administrators.pwd.index
authors.pwd.index
service.pwd.index
filetype:config web

Todas elas retornarão algum arquivo esquecido pelas configurações do FrontPage, mas, como dissemos acima, o IIS não está inteiramente livre desses problemas. Basta procurar por *inurl:iisadmin* ou *inurl:"wwwroot/*"* – os dois diretórios que são a base dos servidores baseados em IIS – para ver quantos resultados são retornados. O mesmo vale para *inurl:"ftproot/*"*, que pode servir como uma grande loja de doces na qual um cracker pode encontrar desde bancos de dados em backup (alguns administradores ainda insistem em fazer backups via FTP) até documentos “confidenciais” e softwares ainda em fase de desenvolvimento.

CONSULTANDO A LISTA TELEFÔNICA

Seguindo o que foi exposto nos capítulos anteriores, poderíamos imaginar que consultas “clandestinas” ao Google só poderiam atingir administradores de sistemas ou usuários profundamente viciados em Internet e uso de dispositivos de informática – os autodenominados *hardusers*. Afinal, é necessário ter uma página na Internet, um currículo ou um e-mail de lista de discussão para permitir a entrada de desconhecidos anônimos em sua vida pessoal.

Mas não é isso o que acontece. O Google, em um primeiro momento, seguido pela grande maioria dos serviços públicos, tornou-se um grande repositório de informações pessoais e de maneiras de se aproximar de uma pessoa sem nunca tê-la visto de frente. Basta ter uma conexão à Internet e um pouco de paciência para saber tudo sobre a vida de muita gente.

A lista telefônica do Google

O Google possui um serviço experimental, por enquanto disponível apenas nos Estados Unidos, que permite encontrar informações sobre qualquer pessoa que possua um nome real e endereço fixo no território americano (incluindo Alaska e Havaí) via Internet. Apesar de estar na sua versão Beta, ele permite que, após o encontro da pessoa, seja possível visualizar um mapa dos arredores de sua residência e a melhor maneira de se chegar lá.

Para começar, acesse o endereço <http://local.google.com/lochp>, o endereço do Google Local. Em **Search Terms**, digite o sobrenome da pessoa ou família que deseja encontrar nos Estados Unidos. Na segunda caixa, adicione ao menos duas informações extras, como endereço, cidade, Estado ou Zip Code (o CEP americano). Lembre-se de que o Zip Code é formado por cinco números.

Em nosso caso, fizemos uma busca pela família Klein, de Nova York (não se esqueça de escrever New York). Sabemos que o Zip Code é 02020, mas, como não temos o guia postal americano à mão, não sabemos o nome da rua em que eles residem, nem o número, nem em que bairro a residência se localiza.



Na tela de resultados, temos acesso a diversas informações. Agora, é só aprender a lê-las.



O nome em azul à esquerda (1) é o nome completo da pessoa, comércio ou instituição encontrada com o nome pedido. Logo abaixo (2), temos o número do telefone, incluindo o código de interurbano dos Estados Unidos entre aspas. Ao ligar do Brasil, lembre-se de que a discagem deve ser iniciada com o código internacional dos Estados Unidos (001).

Em **Address** (3), temos o endereço do usuário. E aí descobrimos nosso erro: o código postal 02020 não é de Nova York, mas de Maryland. Talvez por isso não tenha sido retornado nenhum endereço de Nova York... Ao notar esse tipo de “desvio” do Google, o ideal é repetir a consulta sem o Zip Code, quando aparecerem endereços da cidade certa, Estado certo, com os códigos postais certos.

Na coluna **Related Web Page** (4), podemos descobrir se a pessoa que procuramos possui site na Internet, possibilitando a descoberta de outras informações, endereço profissional, endereços de e-mail, entre outras coisas.

Clicando no link **See these results on a map of this region**, acima dos resultados, o usuário é levado para uma página que mostra o mapa da região onde se localiza o endereço consultado, numerando cada um dos resultados no mapa. Clicando ainda sobre **Zoom Out** ou **Zoom In**, é possível ter detalhes dos arredores do endereço, ou descobrir rotas gerais, como estações de metrô ou avenidas, para chegar ao endereço desejado.



Telefones em São Paulo

Enquanto o Google não disponibiliza o mesmo serviço para outros países como o Brasil, as companhias telefônicas já permitem a consulta da lista telefônica via Web. No site <http://www.telefonica.com.br/online/guias>, da empresa que fornece telefonia para o Estado de São Paulo, é possível conseguir telefones de todos os assinantes de linhas de telefonia fixa cadastrados.

Mas existe uma vantagem desse serviço em relação ao Google: pode-se pedir que seu nome não figure na lista, tanto online quanto impressa. Isso pode ser feito no mesmo link <http://www.telefonica.com.br/online/guias>, mas clicando no link **Para sua Casa** e, em seguida, em **Sua Senha/Dados Cadastrais**.

A próxima página acessada fornece um login para a obtenção de diversos serviços da Telefônica sem sair de casa. Para se cadastrar, clique em **1ª Vez** para preencher o formulário. O serviço é relativamente seguro: após responder ao cadastro, você deve esperar alguns dias para que a contra-senha de acesso chegue à sua casa via correio. Ao recebê-la, basta digitá-la no site da Telefônica, clicando em **Desbloqueio de Senha**, juntamente com a senha anteriormente cadastrada, para habilitar o serviço. Cadastrada a senha, vamos ver como fazer com que seu nome desapareça da lista telefônica:



- 1 – Clique no link **Para sua Casa**. Na base da página que será aberta, clique na caixa **Solicitação de serviços** e, em seguida, em **Fig. e Não Fig. na Lista**.
- 2 – Na página seguinte, clique em **Solicite Aqui!** (o serviço é gratuito).
- 3 – Para tirar seu nome da lista telefônica, clique em **Não figuração em lista** e clique, em seguida, em **Confirmar**.

Esse serviço também pode ser obtido no site da Embratel (http://fatura.embratel.net.br/embratel/jsp/Inf102/VipLine_Ent.jsp). Já para fazer buscas em listas telefônicas de todo o Brasil, procure em <http://www.auxilio-a-lista.com.br> e clique na região do mapa onde deseja encontrar o usuário e telefone.



ANALISANDO LOGS VIA GOOGLE

Um administrador de sistemas costuma ter mais trabalho do que braços e mais coisas para lembrar do que cérebro para guardá-las; é por isso que existem os logs. Os arquivos de log são pequenos documentos de texto que são atualizados por scripts, de acordo com a vontade do administrador de sistemas. Existem logs para as mais diversas funções, desde login de usuários em máquinas locais – incluindo tentativas de login sem sucesso –, até logs de acesso a arquivos ou serviços distribuídos por servidores.

Do ponto de vista dos usuários, o log costuma ser a derradeira ferramenta de adestramento – ou melhor, coação – que o administrador de sistema pode apresentar contra suas incursões fora da rede ou da Web, ou suas pretensões de fazer o que quiser. Os usuários costumam resolver esse problema, como já vimos, contestando os logs administrativos e alegando que eles podem ser alterados manualmente para prejudicá-los – apesar de essa defesa, na maioria das vezes, não ser levada em consideração.

Logs remotos

A maioria dos aplicativos de rede gera logs, que podem ser armazenados tanto no servidor em que o aplicativo é rodado quanto em máquinas preparadas especialmente para armazenar logs (os servidores de logs). Alguns aplicativos ainda trazem a opção de envio de arquivos de log ao e-mail do administrador ou a um FTP escolhido para esse fim.

O ideal é que o administrador tenha os logs sempre à mão, mesmo que ele não esteja na empresa – o problema é que nem todos os serviços e softwares disponibilizam logs online. Isto pode ser resolvido de várias maneiras, dentre as quais:

- a) Criar manualmente um serviço de FTP com senhas, armazenando nele os logs.
- b) Criar um diretório escondido, compartilhado para a Web, e utilizar nele todos os diretórios.

Ambas as soluções, na verdade, são muito parecidas e perigosas. Criar um diretório compartilhado na Web, mesmo que encoberto por diversas “camadas” de diretório, pode funcionar para esconder seus arquivos por um curto período de tempo, mas não como solução-padrão. Por mais que os logs pareçam suficientemente ocultos, sistemas de busca como o Google são capazes de devassar servidores, buscando por caminhos de diretório inteiros.

Fazendo a busca digitando *inurl: "/admin/logs"* na barra de buscas do Google, encontraremos diversos servidores – na maioria servidores Web Apache – desprotegidos ou mal configurados, com muitos logs para serem observados. Para ter certeza de que estamos falando de um servidor Apache, observe se o caminho completo é */www/admin/los*. O caminho */wwwroot/logs* ou qualquer um contido no diretório */wwwroot* diz respeito a servidores Windows rodando o IIS (Internet Information Service).

Visível e vulnerável

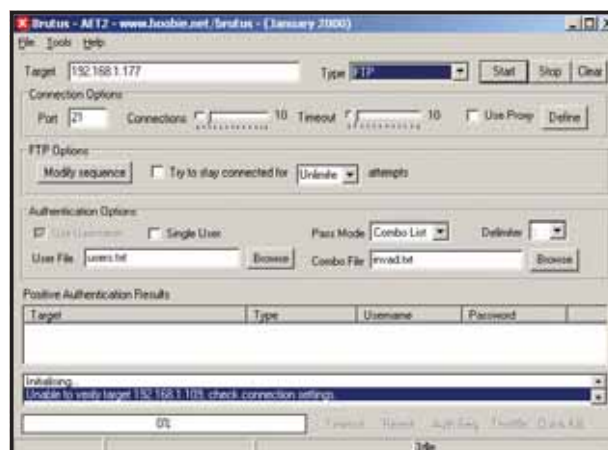
Ao mesmo tempo, servidores FTP contendo arquivos de logs, mesmo que protegidos por senha, podem ser visíveis via Google. Ao fazer uma consulta utilizando o parâmetro *inurl:"ftp/log"*, o Google retorna uma infinidade de links contendo esses diretórios. Ao tentar acessá-los, em muitos deles não é possível entrar: sinal de que o administrador colocou permissões de arquivos nos logs ou inseriu uma senha. Esse último recurso é facilmente contornável: os crackers costumam “reverter” o endereço do log, procurando o endereço que corresponde somente ao FTP.

Um exemplo: ao acessar o link http://www.test.com/test/servidores/headers/ws_ftp.log, o acesso é imediatamente recusado com a exibição de uma mensagem. Mas, abrindo uma nova janela do browser, reverte-se o endereço do link. Apesar de ser um endereço de página Web (www), ele é escrito substituindo o www por FTP. No caso, teríamos <ftp://ftp.test.com>.

Se o FTP não está configurado para receber logins anônimos diretos – a modalidade de logins na qual a tela para inserção da senha nem aparece – ou indiretos – quando basta digitar *anonymous* em usuário e manter a senha em branco – o cracker terá de inserir um nome de usuário cadastrado no sistema de FTP e senha.

Para conseguir acesso a todos os arquivos do FTP – incluindo logs – o cracker costuma utilizar as contas administrativas, que dão poder ilimitado para seu possuidor. Vale até tentar acesso utilizando senhas mal configuradas, esquecidas pelo administrador ou dono do FTP (que nem sempre são a mesma pessoa). Apesar de serem vulnerabilidades antigas e muito exploradas, conhecidas por todos que passaram, no mínimo, por cursos de webmasters, encontrar portas abertas digitando o formato (login/senha) *admin/admin*, *administrador/administrador*, *ftp/ftp* ou *ftp/”senha em branco”* é mais comum do que se imagina.

Quando isso não é possível, o jeito é utilizar técnicas mais agressivas e burras, como os sistemas de *brute force*. Agressivas porque programas deste tipo para o Windows, como o Brute Force AT2 (<http://www.hoobie.net/brutus/brutus-aet2.zip>), acompanhados de um arquivo combo (arquivo de texto que contém milhares de especificações para login e senha que serão lidas pelo software) são capazes de fazer mais de 3000 tentativas de login em uma hora. E isso, utilizando máquinas antigas como 486 e uma conexão de 56 kbps (o equivalente a uma linha telefônica comum).



Burras porque, se você não obtiver sucesso, acabará deixando um log com todas as tentativas de acesso feitas por seu número de IP, que poderá ser facilmente rastreado pelo administrador usando o comando `traceroute` (presente em diversos sistemas operacionais baseados em UNIX) até chegar ao seu provedor e, se você possui IP fixo (que não muda a cada reconexão à Internet), até sua conta de acesso.

Aliás, essa é uma das coisas que a maioria dos crackers e script kiddies aprendem ao invadir um servidor com o sistema de login e conseguir privilégios de administrador: apagar, por via das dúvidas, os arquivos de log. Isso é muito simples de fazer, bastando acessar a linha de comando e programar o seguinte bash:

```
#!/bin/bash
cd /var/log (diretório de armazenamento padrão dos logs em
sistemas baseados em UNIX)
for l in `ls -p|grep '/'`; do
    -n >$l &>/dev/null
    echo Zerando arquivo $l...
done
echo Limpeza dos arquivos de log concluída!
```

Um bom administrador de sistemas, no entanto, nunca deixaria o `/var/log`, sozinho, ser repositório de todos os logs do sistema. Pode-se, por exemplo, copiar todos os logs para um outro diretório:

```
# cp -a /var/log (diretório de destino)
```

Uma esperteza a mais seria enviar o arquivo para um diretório de destino em outra máquina da rede:

```
# cp -a /var/log /192.168.0.1/(diretório de destino)
```

E uma esperteza MUITO maior seria apagar os arquivos originais depois de copiá-los para um repositório de logs:

```
# cp -af /var/log /192.168.0.1/(diretório de destino)
```

RASTREANDO CARTAS E ENTREGAS

A Internet facilitou muito a comunicação entre as pessoas. Cartas, que demoravam dias para chegar ao seu destino, o telégrafo e mesmo os caros envios de fax a filiais de empresas no exterior puderam ser substituídos pelo e-mail. Também conhecido como correio eletrônico, ele é uma forma de comunicação instantânea capaz de enviar, juntamente com as mensagens, várias páginas de documentos, planilhas e até pequenos softwares, dependendo da capacidade da configuração do servidor de e-mails.

Mas, às vezes, é necessário enviar documentos reais, grandes quantidades de arquivos, mídias e até objetos. Nestes casos, não há outra maneira senão apelar para os meios tradicionais: juntar tudo o que deve ser enviado em um envelope ou caixa bem reforçada, ir à agência de correios mais próxima e fazer o envio do pacote – de preferência por meio de serviços confiáveis, como o Sedex. Além disso, é possível fazer um seguro para reembolso do valor nominal do pacote em caso de extravio – se bem que isso raramente acontece, já que os Correios são uma das instituições mais confiáveis que existem.

Há cerca de 6 anos, no entanto, executar essa operação lançava qualquer pessoa em um período de incógnita. Enviar um pacote via Sedex ou um cheque postal, de remetente para destinatário, para dentro da cidade de São Paulo equivalia a esperar cerca de dois dias pela encomenda, sem ter nenhuma notícia dela nesse meio tempo.

Felizmente, o Correio se modernizou – muito mais do que outros serviços de outros países. Como veremos a seguir, podemos rastrear encomendas internacionais via Google. No site dos Correios (www.correios.com.br), é possível rastrear pacotes, vales e cheques postais e até cartas e encomendas registradas, com ou sem declaração de valor.

1 – Na página dos Correios, clique na opção **Rastreamento – T&T**, na barra azul que fica na parte superior da página.

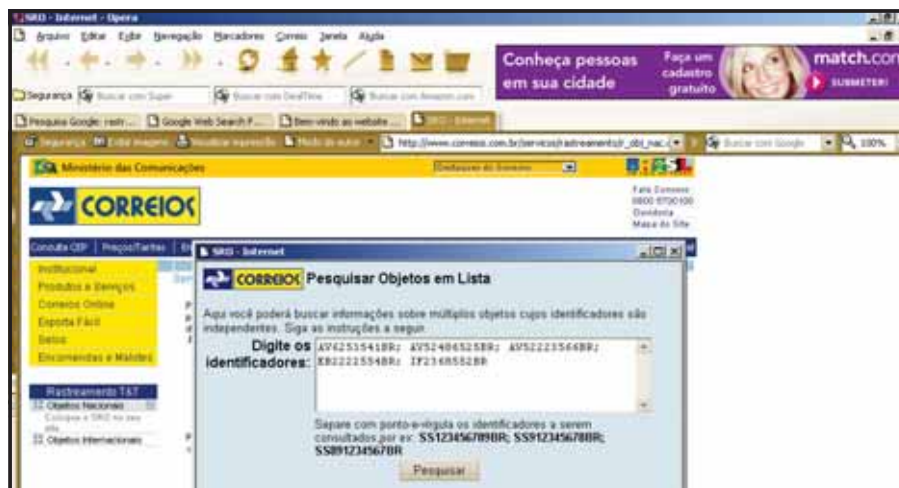


2 – Você será enviado para a página *Rastreamento de Objetos*. Clicando no link **clique aqui** (1), há uma lista com todos os objetos que podem ser rastreados e as siglas que devem ser utilizadas na busca. Essa lista inclui encomendas do DETRAN (documentos e multas) e objetos da Receita Federal.

3 – Para fazer busca por um objeto que já foi enviado pelo correio, dentro do território nacional, clique em **Objetos Nacionais** (2). Na página a seguir, clique em **Um Objeto** para rastrear um único objeto. Para vários objetos do mesmo tipo, enviados em seqüência, clique em **Intervalo de Objetos** e digite os 13 primeiros dígitos recebidos no canhoto do Correio, que estão impressos bem acima dos códigos de barra.



4 – Se você costuma fazer muitas encomendas, enviar muitas cartas e, simultaneamente, comprar ou fazer vários pagamentos pelo correio, a última opção, **Lista de Objetos**, permite a busca de objetos de diferentes tipos. Você pode, por exemplo, rastrear diversos Avisos de Recebimento (AR), ao mesmo tempo em que rastreia um Sedex Internacional (EB) enviado por você ao Suriname, e ver se sua nova cédula de CPF (IF) já foi enviada pela Receita Federal, bastando separar as siglas com “;” (ponto-e-vírgula).



Carta nacional

Outro truque interessante, escondido no site dos Correios, é o que permite que você, mesmo do exterior, envie uma carta para um destinatário do Brasil, pagando a tarifa de postagem nacional. Esse serviço é útil quando o destinatário não possui e-mail ou uma reunião ou conversa deve ser documentada por carta. O serviço está disponível no link http://www.correiosonline.com.br/pt_product.asp?dept_id=1&sku=0&indice=1, e você pode, inclusive, optar por ser avisado sobre o recebimento da correspondência, clicando em **Serviços Adicionais > Aviso de Recebimento**.

Após escrever a carta, basta clicar em **Adicionar à Cesta** para enviar a carta, ao custo de R\$ 5,14. É um pouco caro, mas para quem está em outro país e não sabe como o correio local funciona, ajuda muito. Por último, basta clicar em **Passar no caixa** para pagar a carta (você deve ter um cartão de crédito). A carta será impressa pelo correio, selada e enviada ao destinatário. O Correio garante que a correspondência chega até, no máximo, ao meio-dia do dia seguinte ao envio.

Rastreando FedEx com o Google

Usando o Google, qualquer pessoa é capaz de rastrear a localização atual de uma encomenda da Federal Express (FedEx), o mais famoso serviço de entregas do mundo, capaz de levar encomendas até regiões em que as ruas não têm nome ou as casas não têm numeração, desde que haja um número de referência.

Para fazer isso, vá até a página inicial do Google e digite os 12 dígitos de identificação da encomenda, entre aspas, na barra de busca. Uma boa dica é fazer uma captura dessa tela e guardá-la, caso tenha algum problema de atraso com a encomenda.

GOOGLE AJUDANDO OS CRACKERS

Tudo que vimos até agora mostra o imenso poder que o Google – e os serviços de Internet, por extensão – pode ter para facilitar a vida das pessoas. Coisas como descobrir telefones, endereços ou rastrear encomendas podem, em muitos casos, salvar carreiras, auxiliar no reencontro de parentes, entre outras coisas.

Infelizmente, crackers e aspirantes a hacker também podem utilizar as informações conseguidas pelo Google para se fazer passar por outras pessoas, fazer compras em seus nomes etc.. Além disso, existem – como já mostramos nos capítulos dedicados a logs e vulnerabilidades envolvendo informações sobre páginas e servidores Web –, recursos do Google que podem ser utilizados, **de forma específica**, para abrir ou explorar brechas de segurança.

E, por incrível que pareça, ainda há muitas outras vulnerabilidades e “truques” que podem e são explorados. Basta ter um cracker, um usuário ou servidor, e um browser rodando o Google... tanto do lado do atacante quanto da vítima.

A barra de ferramentas do Google

Desde meados de 2003, o Google disponibiliza para download seu *Google Deskb*ar, uma barra de ferramentas que é adicionada ao Desktop do Windows como um plug-in para trazer todas as funcionalidades do Google à barra de tarefas, além de adicionar mais algumas, como o bloqueio de pop-ups e a procura direta por softwares no Download.com. O programa pode ser baixado em <http://toolbar.google.com/deskbar>. Você deve desabilitar o bloqueio de cookies: no Internet Explorer 6, vá ao menu **Ferramentas > Opções da Internet > Privacidade** e clique em **Aceitar Cookies**.



Apesar de muito útil, essa barra de ferramentas, por estar intimamente ligada ao Internet Explorer, compartilha com ele de todas as suas vulnerabilidades – tanto as velhas conhecidas do universo cracker quanto as mais recentes. Atualmente, são descobertas mais de três vulnerabilidades do Internet Explorer por semana, o

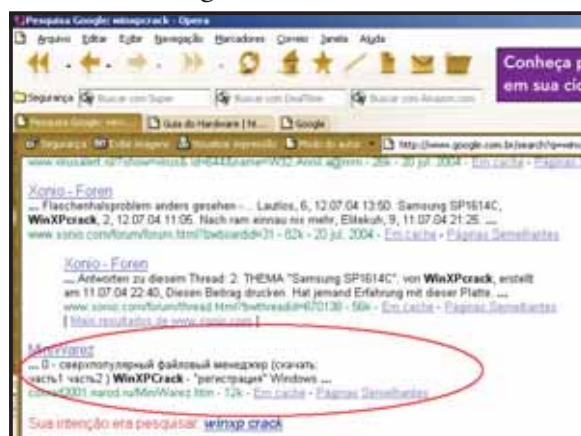
que fez com que o CERT, um dos maiores centros de segurança do mundo, sediado nos Estados Unidos, recomendasse o uso de outros navegadores, até que correções substanciais no código do IE fossem feitas.

Suponhamos que um cracker invadiu um site que possuía uma média razoável de visitantes, conseguiu invadir a conta do administrador e, tendo acesso aos diretórios da página, incluiu vários arquivos executáveis (exe, com) em sua estrutura, dando a eles nomes sugestivos como **cracker_WarcraftIII**, **senhas_porno** ou **hackertools**.

A menos que a página esteja abandonada, o administrador notaria a inclusão de links com código malicioso nas páginas – portanto, links como **eusouumvírus.exe** no meio de uma página de notícias seriam imediatamente eliminados, bastando que o dono da página tenha costume de olhar para ela.

E é aí que entra a arte: com a senha de administrador, qualquer um pode fazer upload de arquivos executáveis, escondendo-os nas árvores de diretórios, sem estabelecer links definidos. Depois, é só contar com o Google.

Os crackers costumam dizer que o Google rasteja (*crawl*) link por link em busca de arquivos e caminhos para um diretório, mesmo que eles sejam invisíveis à primeira vista, e é com isso (mais o Internet Explorer) que eles contam para disseminar arquivos maliciosos. Ao fazer uma busca no Google pelo termo *winxpcrack*, por exemplo, um usuário desavisado estará procurando uma maneira de ativar uma versão ilegal do sistema operacional da Microsoft. Surgirão ao menos 600 resultados, divididos entre os softwares propriamente ditos e links para o aplicativo, como o mostrado a seguir:

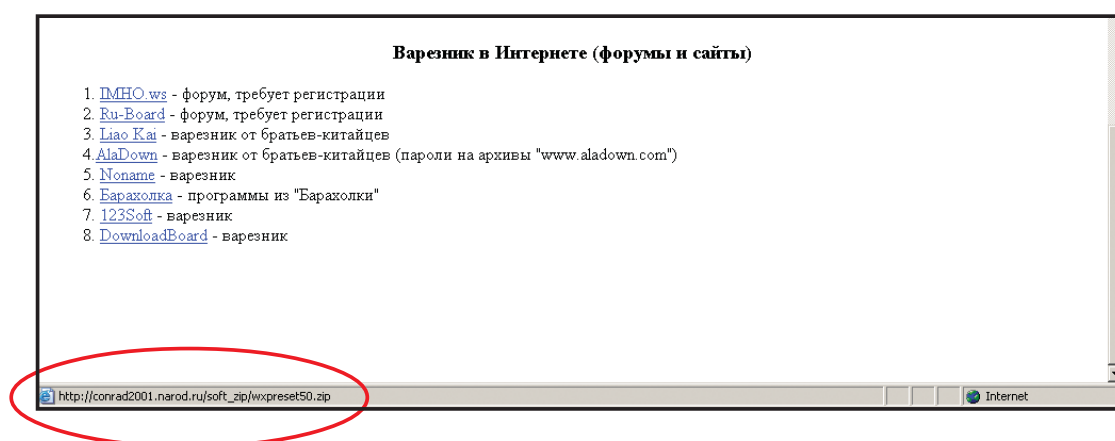


Links e vírus

Com base apenas nos links e suas descrições, um usuário desavisado não é capaz de saber o que está no interior de uma página. No exemplo acima, temos uma página de cracks na qual é possível encontrar o crack que procurávamos: o crack para o windows xp. Mas, ao clicar no link, notamos que ele é uma coleção de links e descrições sumárias de cada crack. Os acostumados a esse tipo de expediente sabem que esses pequenos softwares costumam ser disponibilizados nos arquivos zip, rar, exe ou com. No nosso caso, ao posicionar o ponteiro do mouse sobre o link, descobrimos que o arquivo se chama **wxpreset50.zip**.

O que não significa que ele seja o que diz ser. As configurações-padrão do Windows (Windows Explorer > Ferramentas > Opções de Pasta > Modos de Exibição > Ocultar as extensões dos tipos de arquivos conhecidos) ocultam as extensões de arquivo em todo o sistema, incluindo o navegador. Assim, ao baixar o arquivo **wxpreset50.zip** com o ocultamento de extensões desabilitado, o usuário pode, na verdade, estar baixando o arquivo **wxpreset50.zip.exe**.

É claro que ainda é possível fazer a identificação de arquivos usando a barra de nomes do programa. Veja, na figura abaixo, que, apesar das extensões de arquivos ocultas, ainda é possível visualizar a extensão do arquivo.



Inurl e arquivos maliciosos

Além disso, o Google possui maneiras de construir verdadeiras arapucas, com o auxílio do Internet Explorer. Uma delas é utilizar os filtros *inurl*, associados ao comando *filetype*, para buscar arquivos executáveis de maneira rápida. Ao procurar pelo arquivo **supercrack** utilizando a busca *inurl:supercrack filetype:exe*, o usuário recebe vários links que levam diretamente ao download do arquivo **supercrack.exe**. Ao clicar sobre o link, o arquivo é imediatamente baixado para o diretório local e executado pelo Internet Explorer, mesmo que o usuário não deseje fazer isso. As maneiras de se resolver esse problema são:

a) Não utilizar o Internet Explorer como browser. Existem outras opções, como o Opera e o Mozilla, que não abrem arquivos de forma automática.

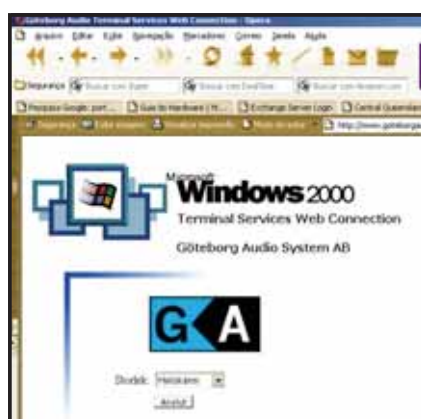
b) Possuir um proxy na máquina ou na rede que, em conjunto com um antivírus, busque certas extensões de arquivo (.exe, .com, .zip) e as varra com um antivírus, procurando por arquivos infectados. Boas duplas são o SQUID (<http://www.squid-cache.org/Versions/v2/2.5/squid-2.5.STABLE6.tar.gz>), agindo em conjunto com o F-Prot (http://www.f-prot.com/download/download_fplinux.html), para Linux; e a dupla WinConnection (<http://www.winconnection.com.br/download.php?versao=S35>) e AVG (http://www.avgbrasil.com.br/br_index.php?obj_id=33).

Descobrendo máquinas vulneráveis

Abandonando uma postura de espera, um cracker também pode utilizar o Google para realizar uma busca agressiva por alvos. Ao procurar, por exemplo, pela entrada de servidores Exchange (o servidor de e-mail corporativo da Microsoft), muito utilizado em empresas e instituições públicas, ele poderia utilizar o comando `intitle:Exchange Server login`, que retornaria todas as páginas com *Exchange Server* no título da página. Uma busca de mais dois minutos, pode retornar páginas como esta, em que basta acertar um nome de usuário ou mailbox para ter acesso aos e-mails:



Outra do comando `intitle` que faz a alegria dos crackers: todos que trabalham com redes sabem que o serviço Telnet (porta TCP 23), ao lado do FTP (porta TCP 21) é um dos serviços mais inseguros da história das redes. Por sua vez, a Microsoft possui um serviço de administração remota, o Terminal Server, que usa muita coisa do Telnet e é utilizado para administrar máquinas em domínios Windows 2000 de redes locais, mas também para administrar hosts remotos (em outras redes, ou ligados diretamente à Internet). Um cracker à procura de servidores específicos, rodando Windows 2000 Server, pode fazer uma busca no Google utilizando o seguinte o parâmetro `intitle:Terminal Server Webs Connection`. O resultado pode ser a descoberta de sites como este:



Tudo bem, nesse caso, o serviço é mais difícil de ser invadido via browser. O que não quer dizer que, ao tentar fazer uma conexão via Telnet ou apontar um scanner de vulnerabilidades para o endereço da página, revelem-se várias brechas no login ou na administração de portas e serviços que permitam a invasão.

PASSWORD GENERATOR COM GOOGLE

Antes do Google, a vida de um cracker não era nada fácil, mesmo para os iniciantes: conseguir exploits, senhas de softwares, senhas de sites, encontrar servidores desprotegidos e até encontrar serviços disponíveis, funcionando atrás de modems desprotegidos era bastante complicado. Essa última modalidade, em especial, era conhecida como *war dialer* e consistia em fazer com que um software de discagem varresse uma faixa de números telefônicos, em busca de modems sem password, com senhas-padrão ou com senhas fáceis de serem quebradas. Normalmente, demorava uns dias para achar um sistema desprotegido dessa maneira, e, além da satisfação pessoal em praticar uma contravenção (naquela época, fazer esse tipo de coisa ainda não era crime) a recompensa vinha na forma de uma conta telefônica que podia mostrar contagens de pulsos de até quatro algarismos.

Somente com a **Página de Pesquisas Avançadas** é possível fazer a substituição de uma série de ferramentas pelo uso de um browser. Geradores de passwords, scanners e testes de vulnerabilidades deixam de ser úteis e fazer com que o antivírus dispare continuamente avisando que há trojans em sua máquina.

Nesta *Segunda Parte* de nosso livro, já vimos, aliás, diversas funcionalidades do Google que permitem encontrar servidores vulneráveis com as brechas mais ridículas. Vamos abordar mais algumas delas, especialmente as relacionadas a senhas, passwords, licenças de software e listagens de acesso, e mostrar porque o administrador de sistemas deve fazer muito mais do que colocar o servidor no ar.

Fontes inesgotáveis

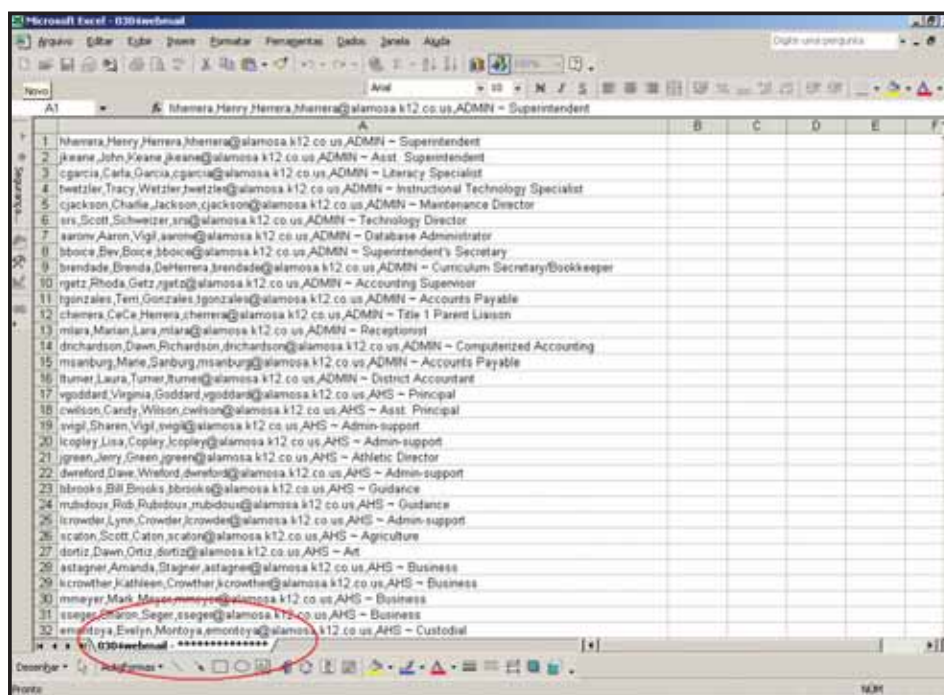
Já mostramos como é fácil encontrar servidores vulneráveis a serviços de login mal configurados ou mesmo com logins inexistentes e também vimos que não é de todo impossível encontrar listas de senhas na Web usando o Google. Mas, e se fosse possível ter uma coleção inesgotável de padrões de logins, senhas, e-mails e seus respectivos endereços de servidores com o auxílio do Google?

Infelizmente, é possível. Por meio de buscas avançadas, pode-se conseguir tantos endereços de e-mail e senhas quantos desejar, montar uma lista de senhas ou uma combo list – basta juntar tudo em um grande arquivo de texto – e gravá-los em CD juntamente com uma lista de servidores e um software de *brute force* para ter uma verdadeira máquina de crackear em menos de 1 hora. Vale lembrar que esse procedimento é mostrado aqui para alertar os administradores de sistema e mostrar como sistemas computacionais podem ser frágeis a ataques feitos com recursos mínimos e por pessoas com conhecimentos exíguos.

Para começar a montar um *password generator*, o cracker pode começar por listas de e-mails reais. Muitos sistemas utilizam endereços de correio eletrônico como forma de login, juntamente com o nome completo do usuário e, algumas

vezes, descrição de seu cargo e departamento. Essas informações são confidenciais, e a maneira como todos os usuários do sistema executam o login – se bem que isso pode ser presumido simplesmente estudando o front-end do serviço de login (os servidores Outlook Mail e Exchange Server são especialmente vulneráveis) – só deveria ser conhecida pelo administrador de sistemas ou alguém autorizado a ter acesso a todas as contas.

Vamos procurar no Google pelo parâmetro *e-mail address filetype:csv csv*, que equivale a dizer que queremos encontrar documentos csv – o padrão utilizado pelo Outlook Express e o Outlook e os serviços de correio padrão do Windows e do Microsoft Office para exportar dados em forma de arquivo-texto. Com o termo *e-mail adress*, mostramos ao Google que queremos apenas arquivos “cheios”, ou seja, que contenham efetivamente, endereços de e-mails e outros tipos de informação. A pesquisa retorna cerca de 500 resultados, em sua maioria planilhas de Excel com listagens de e-mail, nomes completos dos proprietários das contas e até seus cargos. Algumas planilhas, como a mostrada abaixo, mostram até a que instituição e serviço de webmail podem se aplicar os endereços da planilha.



Para serviços de webmail que utilizam o Zip Code (código postal americano) ou o número do usuário como login, também é possível encontrar arquivos csv completos, incluindo data de nascimento e endereço do usuário. Não conseguindo à primeira vista, é possível fazer isso utilizando o parâmetro *residential phone e-mail address filetype:csv csv*.

Outra maneira de conseguir e-mails, utilizada, nesse caso, para “ganhar” muitos logins do Hotmail/MSN Messenger, é utilizar a barra de buscas com o parâmetro *filetype:ctt ctt messenger*. Ctt é a lista de endereços do Hotmail/MSN que, para facilidade do usuário, pode ser exportada no formato XML, mas também pode,

Certificados e chaves PGP

Indo um pouco além de logins, senhas e e-mails, é possível conseguir até chaves PGP e certificados de autenticidade de servidores. Esses elementos podem ser considerados, por um cracker com um mínimo de inteligência, muito mais úteis do que listas de logins ou passwords: logins do administrador aplicados fora de hora e a partir de IPs estranhos são facilmente observáveis pelo próprio administrador, se este tem o costume de ler logs. Já uma chave PGP ou certificado digital podem fazer com que um cracker passe, facilmente, pelo dono da chave, ou que uma rede falsa, montada como armadilha, seja vista por usuários da Web como se fosse a rede corporativa proprietária do certificado.

E, assim como o Google, a chave PGP, que deveria ser uma ferramenta de auxílio ao usuário e, especialmente, ao administrador, se transforma em arma nas mãos do cracker – como salientamos várias vezes, por falhas no sistema operacional, erros em softwares ou falta de aptidão do próprio administrador. Basta fazer a busca com o parâmetro *BEGIN (CERTIFICATE|DSA|RSA) filetype:key* no Google para ver quantas chaves privadas PGP se consegue (algumas até mais elaboradas, como as de 1024 bits).

```
seifried-cert.key - Bloco de notas
Arquivo  Editar  Formatar  Exibir  Ajuda
-----BEGIN CERTIFICATE-----
MIIDtzCCAyCgAwIBAgIQ08AwExKJ74akIjww0x4BrDANBgkqhkiG9w0BAQQFADCB
UDExMBUgA1UEChMOVmVyaVNPZ24sIEIuYy4xXzAdBgNVBASTE1Zlcm1TawduIFRY
dXN0IE51dHdvcm5rXjBEBGqNVBASTPXd3dy52ZXJpc21nb15jb20vcmwvb3NpdG9y
eS9SUEEgSw5jb3JwL1BCeSBzZWYULExJQUIUTFREKGMPotgxNDAYBgNVBAMTK1Zl
cm1TawduIUENSYNXNIZIDigQ0EgLSB3bmrpdm1kdWFsIFNlYnNjcm1hZXIwHhcNOTgx
MDIxMDAwMDAwHhcNOTkxMDIxMjM1OTU5w7CB6TEXMBUgA1UEChMOVmVyaVNPZ24s
IEIuYy4xXzAdBgNVBASTE1Zlcm1TawduIFRYdXN0IE51dHdvcm5rXjBEBGqNVBAST
PXd3dy52ZXJpc21nb15jb20vcmwvb3NpdG9yeS9SUEEgSw5jb3JwL1B1eSBzZWYU
LExJQUIUTFREKGMPotgxJZA1BgNVBASTHkrpZ210YywwgSUQgQ2xhc3MgM1AtIEIu
Y3Jvc29mdDEwMBQGA1UEAxQNS3VydCBTZW1mcm1hZDEKMCIGCSqGSIb3DQEJARYV
```

Servidores FTP

Servidores FTP são especialmente vulneráveis a falhas de login. Os administradores parecem não se preocupar muito com essa classe de servidores: login anônimo, senhas fracas, senhas-padrão ativadas – por exemplo, *login:ftp*, *senha:ftp* – são muito comuns, assim como montagem de servidores com versões instáveis ou com bugs conhecidos, mas não corrigidos, como o *WS_FTP* – que durante muito tempo foi praticamente o padrão para os funcionários de UNIX. Uma busca no Google por *filetype:ini +ws_ftp +pwd* mostra quantos arquivos de senha se consegue capturar de servidores FTP “particulares”.

“FUÇANDO” BANCOS DE DADOS VIA GOOGLE

Existe uma modalidade de cracker que possui, pelas técnicas que utiliza, assim como pelos valores que pode pôr em jogo, um diferencial em relação a tudo o que se costuma falar em matéria de segurança ou de invasões: a exploração de bancos de dados. Na verdade, o cracker de bancos de dados é dividido em diversas modalidades, mais ou menos alinhadas com as fases de uma invasão:

- a) crackers que simplesmente invadem bancos de dados e apagam todos os dados;
- b) crackers que seqüestram bancos de dados e só liberam o acesso mediante um resgate;
- c) crackers que, simplesmente, duplicam os dados localizados nos bancos e os vendem no mercado negro.

O cracker de bancos de dados é tão perigoso que acabou gerando, dentro da área de TI, uma nova profissão: o *Database Security Administrator*, ou, em bom português, a pessoa que deve levar a culpa no lugar do gerente de TI caso dados da empresa apareçam em algum CD pirata ou listagem vendida por hackers. Essa tarefa de proteger bancos de dados, no entanto, não é muito simples, principalmente em servidores de comércio online: os serviços SQL, bancos de dados Access e todos os seus clones teimam em se fazer anunciar – é só perguntar por eles.

Perguntando com o Google

Durante algum tempo, a tarefa era feita por meio de scanners como o nmap, que vasculhavam um range de IPs em busca de resultados que permitissem averiguar que havia um servidor de banco de dados ali. Atualmente, o Google é capaz de encontrar um SQL ou MDB, por mais escondido que ele esteja. Mesmo bancos ocultos por trás de proxys ou firewalls mal configurados podem ser encontrados, já que permitir que uma base de dados receba tentativas de requisição de todas as partes, sem restrição de IP, equivale a tornar o servidor de domínio público.

Para achar um banco de dados, mesmo que oculto por uma simples página HTML, o Google pode utilizar o termo *“Can’t connect to local” intitle:warning*. A maioria dos resultados dessa busca contém bancos de dados SQL/MySQL com falhas de configuração ou atualização que podem ser exploradas. Da mesma forma, o Google pode encontrar servidores IBM DB2 rodando com vulnerabilidades ou serviços defeituosos usando o parâmetro *“detected an internal error [IBM][CLI Driver][DB2/6000]”*.

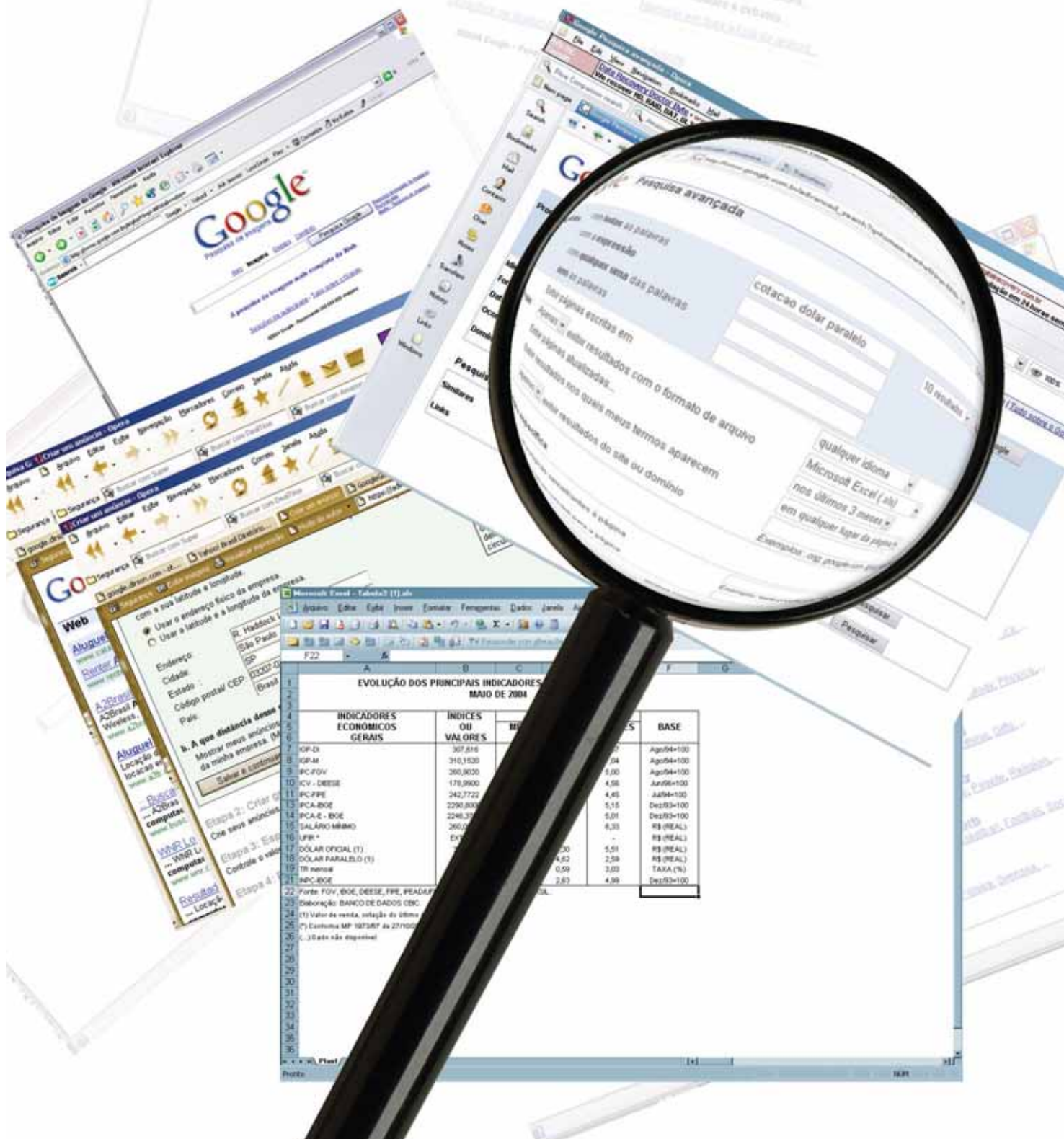
Mas o rei das vulnerabilidades Google é mesmo o SQL. Uma busca por *“Incorrect syntax near”* na barra de busca é capaz de revelar muito mais do que o Nessus e o nmap: caminho do banco de dados, nomes das funções, nomes dos arquivos, detalhes parciais dos códigos, tudo é revelado pelo amigável SQL:

Outras maneiras de buscar bancos de dados usando o Google

- *"ORA-00921: unexpected end of SQL command"* – (a maioria dos servidores Oracle utilizados em base Windows apresenta esse erro em algum momento de sua existência);
- *intitle:"index of" mysql.conf OR mysql_config* – é utilizado para descobrir (e fazer download) de arquivos de configuração MySQL. É claro, se existe um arquivo de configuração, também existe uma base de dados;
- *allinurl: admin mdb* – descobre diretórios de administração de sistemas baseados em Access ou pequenos sistemas SQL. Novamente, vale a regra: onde há um diretório administrativo, há um sistema de banco de dados para ser explorado;
- *filetype:inc intext:mysql_connect* – talvez a mais perigosa das buscas que o Google permite fazer: mostra scripts de login de bases de dados MySQL, detalhando seu processo de execução. Em alguns casos, pode trazer histórico dos últimos logins administrativos efetuados no sistema, revelando nomes de administradores, IPs de estações de trabalho e hosts e senhas-padrão que não foram alteradas no processo de configuração.

Terceira Parte:

Tudo o que você precisa saber



TUDO O QUE EU PRECISO SABER

Os cookies vigiam você

Cookies são biscoitos crocantes, feitos de farinha de trigo, água e leite (não confunda com os *donuts*, a rosquinha frita preferida por nove entre dez policiais americanos). Esses biscoitos são muito quebradiços (não há ovos na receita) e espalham pedaços por todos os lados, mal são pegos nas mãos – talvez por isso, ao projetar a comunicação do protocolo HTTP, os engenheiros tenham resolvido homenagear os famosos biscoitos, dando o nome de cookie aos pequenos arquivos que costumam ser deixados no computador.

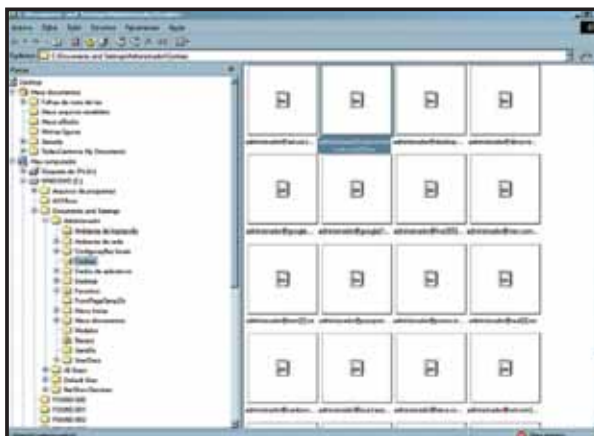
O HTTP, assim como a maioria dos protocolos de comunicação de redes e Web, não utiliza conexões contínuas: isso equivale a dizer que um determinado serviço, como uma página Web ou um servidor de e-mails, só passa a funcionar quando há um pedido de acesso por parte das máquinas-clientes ligadas àquele servidor. Um servidor de e-mail, por exemplo, só recebe as mensagens enviadas para usuários que são cadastrados em seu domínio e só envia as mensagens que esses mesmos usuários cadastrados enviam para outras pessoas – afora isso, o servidor permanece praticamente fechado.

Assim, ao solicitarmos comunicação com qualquer servidor, nossa requisição é recebida, os dados que pedimos nos são enviados (ou não) e, por fim, a conexão é cortada. Esse corte da conexão, porém, faz com que alguns dados, como a configuração do serviço em relação à máquina e dados do usuário (login, senha ou contrasenha), sejam perdidos. Isso impede que páginas possam ser acessadas a partir do ponto em que paramos em nossa última visita, ou que continuemos um curso online a partir do módulo em que paramos, sem precisar fazer todo o curso novamente.

Um “cookie” na minha máquina

Para resguardar a segurança mantendo o sistema de interrupção de conexões, facilitando, ao mesmo tempo, a vida do usuário e tornando a navegação mais intuitiva, a Netscape, ainda na pré-história dos navegadores Web (1995 – 1997), resolveu criar o Magic Cookie (MC), nome mais tarde abreviado simplesmente para cookie.

Um “cookie” é um pequeno pedaço de informação enviado por um servidor Web ao browser do usuário, servindo como uma espécie de “diário de bordo” da navegação por uma determinada página. Ao visitar um servidor que utiliza cookies, este instrui o browser-cliente a criar um desses arquivos, quase sempre um arquivo-texto (extensão txt). Na maioria das vezes, para facilitar a leitura dos dados por parte dos servidores, os arquivos recebem o nome do usuário e o domínio da página à qual o cookie pertence.



Quando o browser está em um computador com mais de um usuário cadastrado (sistemas Windows 2000/XP/2003 e UNIX/Linux com vários usuários), os cookies podem conter informações do usuário local, em vez de informações de login da máquina. Esta modalidade de cookie é utilizada para diferenciar os acessos de diferentes usuários e, ao menos superficialmente, evitar o roubo de senhas.

Independente da forma como o cookie é nomeado, ele costuma conter pequenos pedaços de informação, algumas vezes criptografados: número de IP do usuário, login, senha, último acesso ao serviço, quantidade de dados enviada na última conexão, entre outros.

Por que os cookies são perigosos

Os cookies, inicialmente, deveriam conter somente informações dos sites dos quais se originam. Desse modo, um cookie seria somente um arquivo de texto com algumas informações que plugam o usuário a uma página – nada mais.

Se tudo ocorresse assim, não haveria motivo para ter dois grandes temores:

- a) que um cookie leia informações de outros cookies contidos na máquina ou da árvore de diretórios do sistema operacional;
- b) que um cookie seja utilizado para difundir vírus, trojans ou para espionar a máquina do usuário.

Por que então alguns cookies são identificados pelos antivírus como arquivos maliciosos ou *spywares*? Quando isso ocorre, na verdade, algo mais do que um cookie foi baixado em sua máquina – geralmente aplicativos Java ou pequenos executáveis. Normalmente, isso é feito em segredo, aproveitando a fragilidade das configurações de segurança do Internet Explorer. Nesses casos, o cookie só funciona para avisar ao software espião que o usuário está online, para que as informações possam ser repassadas para o domínio que controla os softwares.

Enquanto o sistema está infectado, alguns destes cookies e softwares podem aproveitar-se, inclusive, de serviços de banda larga e servidores Windows que ficam online 24 horas por dia para se tornarem visíveis na Web, compartilhando o diretório de cookies. Isso visa facilitar a captura dos cookies pelos usuários.

O Google é capaz de rastrear e capturar qualquer um desses cookies colocados na Web. Apesar do perigo não ser muito grande, possuir alguns logins de usuário, endereços de loja eletrônica ou IPs é tudo o que um cracker razoavelmente preparado precisa para utilizar como trampolim, conseguindo mais dados por meio de métodos de engenharia social, *brute force* ou técnicas de SQL Injection (inserção de códigos SQL visando fazer com que bancos de dados SQL retornem seqüências completas de dados).

Ao fazer uma busca por `admin inurl:cookie filetype:txt` na barra de buscas, é possível encontrar resultados listando cookies “perdidos” em servidores de e-mail, arquivos ou pastas do MSN ou ICQ compartilhadas.



No resultado acima, foi possível capturar um cookie, localizado em um servidor doméstico, com dados de uma conversa em grupo em uma sessão do MSN Messenger. Note que os e-mails dos usuários aparecem ao lado das indicações de nickname.

Como evitar indigestão com os cookies

Existem várias maneiras de se livrar, ou pelo menos minimizar os riscos, da presença de cookies em seu computador, sejam eles perigosos ou não. Para os usuários do Internet Explorer, o ideal é acessar o menu **Ferramentas > Opções de Internet > Privacidade** e clicar na opção **Confirmar cookies**. De início, é um pouco chato, mas depois de um certo tempo, além da vantagem da segurança, esta opção começa a mostrar algumas coisas interessantes: você notará, por exemplo, que sites que têm muitos banners (sites pornôns e “hackers”) chegam a colocar 30 cookies em sua máquina, em uma só sessão de navegação.

Outra iniciativa interessante é procurar em **Ferramentas > Opções de Internet > Segurança** a opção **Avisar para download de softwares** – assim, applets Java, arquivos executáveis ou as famosas barras de ferramenta coloridas que aparecem do nada serão bloqueados pelo browser antes mesmo de serem instalados.

Para os mais paranóicos, também é possível bloquear os cookies para que eles possam ser lidos apenas por um domínio. Basta clicar com o botão direito do mouse sobre o cookie, clicar em **Propriedades** e, em seguida, em **Somente Leitura**. Realizar o mesmo processo sobre a pasta **Cookies** de um usuário torna toda a pasta **Somente Leitura**, o que faz com que novos cookies não possam ser adicionados aos já existentes.

PESQUISA DE DOMÍNIOS: COMO SABER TUDO SOBRE UM SITE SEM PRECISAR INVADI-LO

Para que um site possa ser publicado na Internet, ele deve estar localizado em um servidor que obedeça a certos requisitos. O primeiro passo para isso é que ele possua, antes de tudo, um endereço IP válido, que não esteja sendo utilizado por nenhuma outra máquina da Web. O endereço IP, como todos sabem, é formado por quatro seqüências de três números cada, sendo que a primeira série de números mostra a classe da rede à qual a máquina pertence.

A cada número IP corresponde um endereço de domínio, embora vários endereços possam compartilhar um mesmo número IP. Isso ocorre porque os servidores de tráfego da Web não trabalham com nomes: como já dissemos, computadores utilizam a linguagem binária, só trabalham com 0 e 1.

Ao digitar um endereço em seu browser, o endereço é enviado para um servidor DNS, um conjunto de base de dados que contém os nomes dos domínios e as direções IP. Após receber o endereço do domínio, o DNS o envia para o servidor na qual a página está hospedada, que devolve a requisição para o DNS, para que ele a envie ao computador do usuário.

Assim, sem um endereço de DNS, não é possível hospedar uma página em um servidor, nem mantê-la no ar. Mesmo os famosos softwares de programas piratas, crackers e outros tipos de contravenção localizados na Rússia e Romênia precisam ter um servidor DNS habilitado, apontando o endereço de seus servidores para o mundo todo.

Os servidores DNS que sustentam a Internet no mundo todo estão localizados ao lado dos grandes *backbones* – as grandes “estradas” de dados da Web – que utilizam servidores e equipamentos localizados em algumas universidades americanas. Esses servidores DNS são os responsáveis pelo reconhecimento dos domínios .com, .org, .gov, .net, entre outros.

Como a Internet expandiu em muito pouco tempo, foi necessário efetuar a montagem de servidores DNS que hospedassem os endereços de páginas ou domínios que não estivessem localizados nos Estados Unidos. Assim, cada país do mundo organizou seu comitê gestor de Internet, um escritório no qual é possível cadastrar sua página ou servidor (até os provedores têm de se cadastrar nesses domínios). No Brasil, para cadastrar domínios com final .br (com os servidores localizados no Brasil), é necessário fazer a requisição ao Comitê Gestor da Internet Brasileira, que é administrado pela FAPESP (Fundação de Amparo à Pesquisa do Estado de São Paulo).

Como registrar um domínio

Assim, qualquer endereço *.com.br* ou *.org.br* deve passar pelos servidores da FAPESP. Para registrar um domínio, ele não deve estar sendo utilizado por nenhum usuário, deve possuir um endereço IP correspondente a ele e deve estar listado em, ao menos, dois servidores DNS válidos. A página de registro de domínios é www.registro.br, e registrar um domínio por 1 ano custa R\$ 30,00 – um dos valores mais baratos do mundo.

Para que seja possível registrar um domínio, o interessado deve criar uma identidade para si, no endereço <https://registro.br/cgi-bin/nicbr/idmng?stkey=13227292-1792849939>.



Esta página pede vários dados pessoais ao interessado em registrar um domínio. Após informá-los, basta clicar em **Entrar** para que seu pedido de requisição seja feito. Um e-mail será enviado à conta de correio eletrônico cadastrada por você. Basta clicar no link que está no corpo da mensagem para que sua conta no Registro .BR seja ativada.

Aqui, temos a primeira brecha de segurança possível do Registro .BR. Qualquer um pode registrar uma identidade no Registro .BR com o nome de outra pessoa, seja ela quem for. Isso quer dizer que, mesmo que não se consiga registrar nenhum domínio no nome dela (veremos adiante que não é tão impossível assim) é possível “travar” seu registro na FAPESP, fazendo com que a pessoa não possa utilizar seu próprio nome para gerar domínios.

Ao registrar um domínio em definitivo, são exigidos CNPJ – para domínios institucionais –, e CPF – para domínios pessoais – e esse é o único empecilho que o Registro .BR oferece para que alguém registre uma página no lugar de outra pessoa. Como vimos no decorrer de todo este livro, conseguir tais informações não é tão difícil assim, e, basta que elas sejam verdadeiras (o Registro .BR possui um sistema de checagem de registros válidos) para que sejam aceitas em um cadastro.

Todos os domínios registrados devem possuir também um responsável técnico e um responsável pela cobrança da taxa anual. Novamente, não há grandes dificuldades: quem consegue uma identidade falsa, consegue três.

Sites desprotegidos via Registro .BR

A maioria das pessoas, contudo, é honesta, e quer cadastrar um site com seu nome e endereço verdadeiros, além de telefone e e-mail nos quais possam ser encontradas. Infelizmente, os serviços de administração de registros de domínio não guardam as informações pessoais como deveriam: é muito fácil conseguir essas informações pessoais e também várias informações relacionadas aos sites cadastrados.

Na página inicial do Registro .BR, existe uma caixa de verificação que permite a busca por domínios válidos. Se quisermos saber, por exemplo, se o domínio topgames.com.br é um domínio que já está sendo utilizado, basta digitá-lo na caixa de verificação e clicar no botão **Pesquisar**.



O resultado mostra não só que o domínio já está ocupado, como quem é o proprietário, qual é seu endereço e o número do protocolo da assinatura do dono do domínio. Clicando em *documento*, é possível conseguir mais dados sobre a página e seu proprietário, como, por exemplo, todos os domínios que ele cadastrou, outros endereços que, por acaso, tenha declarado e outros endereços DNS para os quais costume direcionar seus domínios.

Toupeira virtual

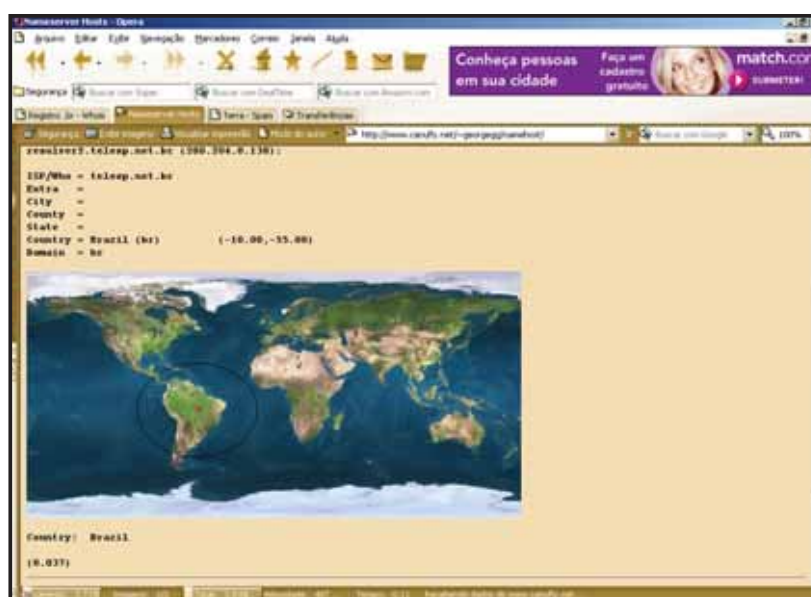
Dá para ver que o Registro .BR compromete um pouco os proprietários de domínios, mas dá poucos detalhes técnicos para quem deseja executar um ataque massivo a uma página. Quanto ao Google, nem adianta tentar: as referências a *Registro .BR* no Google costumam mostrar apenas a maneira de configurar corretamente uma página.

Mas existem outras ferramentas que podem mostrar praticamente tudo sobre um site, domínio ou servidor de DNS. No endereço <http://www.canufly.net/~georgegg/nameshost/> existe uma ferramenta online chamada Nameserver Hosts Utility, que devassa qualquer destino da Web, descendo inclusive à busca de e-mails ativos no DNS e servidores DNS auxiliares.



Procurando, por exemplo, o endereço www.topgames.com.br, descobrimos que, apesar de o domínio já estar registrado, não existe nenhuma página configurada no endereço (1); além disso, o número IP ao qual corresponde a página no servidor será mostrado (2) e teremos todos os endereços do servidor logo abaixo; por último (3), também podemos ter informações sobre o horário no qual o servidor está configurado, bem como que serviço é utilizado para colocar páginas no ar nessa máquina. No caso, como o serviço é o IIS 5.0, conclui-se que a máquina só pode ser um servidor Windows NT/2000. A partir daí, bastaria perder um pouco de tempo à procura de exploits (programas que exploram vulnerabilidades de um determinado sistema operacional ou programa) na Internet para tentar explorar configurações falhas nesse servidor e concretizar uma invasão.

Com este expediente, também é possível conseguir informações sobre servidores DNS. Procurando, por exemplo, pelo DNS do Speedy da Telefônica (200.204.0.138), descobrimos em que domínio ele está contido (telesp.net.com.br), qual sua importância – ele é o resolver2, ou seja, é o segundo DNS da rede etc., e, com um mapa, podemos saber até qual é sua localização geográfica.



AS VÁRIAS FACES DO GOOGLE

Existem diversas formas de aparecer no Google – aliás, para fazer parte das buscas do Google, mesmo que em um ranking muito baixo, basta criar uma página com seu nome e colocá-la na Web, fazendo com que ela tenha ao menos uma visita, nem que seja a sua própria.

Ao divulgar sua página para amigos ou listas de discussão, então, é praticamente impossível ficar invisível. Muitos links significam mais chances de ser encontrado; quanto mais você for encontrado, mais chances terá de aparecer em posições iniciais do ranking do Google.

O que quase ninguém sabe é que o próprio Google possui diversas roupagens, cada uma correspondente a uma maneira de realizar pesquisas ou a uma temática. Apesar de nem todas elas possuírem suas versões em português, é bom conhecê-las: algumas buscas muito, muito específicas só são possíveis com o auxílio desses desvios do buscador do Google.

Google Grupos

(http://groups.google.com.br/groups?group=*&hl=pt-BR)

Com uma interface em português que praticamente não tem diferenças em relação à sua versão em inglês, o Google Grupos reúne praticamente tudo o que foi escrito em grupos de discussão e na própria Internet desde 1995. Estes grupos de discussão (USENET) foram a coqueluche da Web que estava nascendo, muito antes do advento dos browsers gráficos como o Internet Explorer e o Netscape: muitas comunidades científicas, médicas e mesmo relacionadas ao desenvolvimento da própria Internet além de seis serviços nasceram dos grupos da USENET. O Google adquiriu em 2003 os servidores da Deja.com, que era a depositária de boa parte dos grupos de USENET.

Existem representantes de cada tipo de assunto divididos em grupos. Alguns grupos, como o .alt, formado por servidores alternativos, contêm mais de 2471 grupos embutidos. Na interface USENET do Google, é possível visualizar a frequência de cada grupo pelas barras verdes localizadas à esquerda de cada barra: quanto mais verde houver em uma barra, mais atualizado está o arquivo do grupo.



Google Diretórios

(<http://www.google.com.br/dirhp?hl=pt-BR&tab=gd&q>)

O Google Diretórios é uma resposta às críticas em relação à forma como o Google é organizado ou, segundo os partidários de serviços organizados em diretório (como o Yahoo), não é organizado. Muito parecido, aliás, com o serviço de mesmo nome localizado no Yahoo, o Google Diretórios organiza todo o conteúdo da Internet em assuntos e suas ramificações, tratando-os como se cada assunto fosse um diretório diverso em uma grande árvore. Assim, ao procurar no diretório **Games**, podemos ter opções como **Board Games** (Jogos de Tabuleiro) e **Eletronic Games** (Jogos Eletrônicos). Em **Board Games**, podemos encontrar jogos de detetive (**Murder-Mystery and Deduction**); dentro dos jogos de detetive, podemos encontrar, entre diversos diretórios, o **Chinese Detective, The**, um jogo de estratégia ambientado na Pequim Imperial. Resultado desta operação “descasca cebola”: encontramos o link de um fabricante de jogos de tabuleiro de mistério, que fabrica um jogo de detetive chinês, chamado *Pekings Mysterier*.

Google News

(<http://news.google.com>)

O Google News é um dos vários avatares do Google que não possuem semelhantes em português. Ao procurar por um termo no Google News, a busca é realizada apenas em serviços de notícia e, infelizmente, apenas em inglês. Ao mostrar os resultados, observe a indicação de horas ao lado do link da notícia: ela mostra o horário em que a notícia foi colocada na Web, e, por consequência, quão atual ela é.



Froggle

(<http://www.google.com/froogle?hl=en&edition=us&ie=ascii&q>)

Se existe um lugar onde se pode vender alguma coisa, esse lugar é a Internet. Afora as lojas 100% virtuais, que afundaram em finais da década de 90 junto com as empresas .COM, oferecer serviços e mercadorias com um nicho de mercado muito específico na Internet é resultado de sucesso. Mesmo quem nunca comprou nada via Web visitou, ao menos uma vez, sites de leilão como o Mercado Livre

para ter uma idéia de quanto custa um determinado material. O Froggle facilita a vida dos interessados em preços da Web, assim como dos compradores inveterados. Uma busca por Linux, por exemplo, retornou mais de 81.000 resultados entre CDs de distribuições, manuais, livros e cursos.

Google Linux

(<http://www.google.com/linux>)

Como o próprio nome diz, um diretório do Google voltado apenas para a busca de assuntos relacionados ao universo Linux. Apesar de só ser encontrado na versão em inglês do Google, os resultados trazem várias páginas em português. Quando procuramos por *dicas de configuração de rede*, o retorno foi de mais de 32.000 resultados.



Google Microsoft

(<http://www.google.com/microsoft.html>)

O Google Microsoft realiza buscas em todas as páginas relacionadas à empresa de Bill Gates. Assim como seu correlato para Linux, o Google Microsoft é uma boa saída para procurar não só por dicas, mas também por drivers para dispositivos específicos. Em uma procura por drivers para Windows, em específico para dispositivos USB da SiS, encontramos mais de 89.400 resultados.

Cheat Google

(<http://www.cheatoogle.com/index.php>)

Entramos no domínio dos serviços que não são exatamente ligados ao Google, mas se inspiram nele para oferecer serviços específicos. Este site, em específico, traz um verdadeiro objeto do desejo para verdadeiros perdedores: a possibilidade de pular determinadas fases ruins, ganhar muito dinheiro, ou nunca se ferir nem morrer. Pena que isso só diz respeito ao mundo dos joysticks e videogames, mas já é alguma coisa...



Por meio do Cheat Google é possível conseguir dicas para qualquer título em várias plataformas, desde o tradicional Game Boy até o incrementado XBox. Além de códigos (cheats), é possível conseguir roteiros completos para detonar um jogo, reviews de lançamentos feitos em diversos sites e até wallpapers.

Booble

(<http://www.booble.com>)

Um outro clone não autorizado do Google, dessa vez voltado para uma temática, digamos, mais adulta. Nele, todos os resultados remetem a temas eróticos. O serviço é gratuito, mas 90% dos resultados é pago.

CUIDADO COM OS LINKS FALSOS

As modalidades de crime digital tornaram-se muito comuns nos últimos tempos. A fraude online, principalmente a pescaria virtual de senhas retiradas de usuários desavisados ou que estão localizados em redes mal configuradas ou sem uma política de segurança definida é uma constante.

Apesar de, como já vimos, o Google ter o poder de ser utilizado como arma pelos pescadores de senhas das mais variadas maneiras, ele também pode ser usado como ferramenta de segurança. Antes de tudo, no entanto, vamos ver como os links falsos e os e-mails de pescadores de senhas funcionam.

Falsa generosidade

Os e-mails falsos usam nomes de empresas, programas de televisão e, desde o início de 2004 até agora, têm utilizado nomes de instituições bancárias, chegando até a existir um e-mail que utilizava o nome do Banco Central para tentar extorquir informações pessoais do usuário.

No primeiro caso, as mensagens oferecem produtos a preços muito generosos ou prêmios grátis. Já os e-mails que se utilizam de nomes de bancos oferecem seguro grátis, prêmios em dinheiro, cartões de crédito grátis com limites altos ou, em manobras de engenharia social, pedem que os usuários façam um cadastramento de suas contas correntes, cadernetas de poupança ou que enviem seus dados cadastrais para um determinado site, endereço de e-mail, ou que preencham um formulário anexo à mensagem, no qual, entre outras coisas, devem ser declarados os valores existentes na conta corrente, bem como senhas de usuários. Aqui há a primeira coisa a se desconfiar: mesmo que os e-mails venham com endereços que correspondam ao correio eletrônico de um banco, as instituições financeiras possuem regras muito rígidas para o manuseio de senhas e não precisam enviar e-mails para saber quanto cada usuário possui em sua conta corrente.

Além disso, é interessante fazer uma investigação no e-mail. Se você possui Outlook, clique com o botão direito do mouse sobre a mensagem e selecione **Opções**. Em seguida, observe o bloco **Cabeçalhos da Internet**, em especial a linha *Message-Id*. Anote o final do ID da mensagem, por exemplo, *@linux.com*. Em seguida, utilize o prompt de comando e digite `tracert` acompanhado do nome do destinatário do e-mail e do servidor encontrado. No nosso caso, usaríamos:

```
tracert vbanco123.linux.com
```

Observe se um dos últimos endereços corresponde ao nome do banco ou empresa que supostamente lhe enviou o e-mail. Se ele nem passa por um servidor do banco, é certo que você quase caiu no truque da pescaria de senhas.

Links falsos e armadilhas

Você recebe um e-mail do banco Brasileiro e o link aponta para brasileiro1.com ou está radicado no domínio @brasileiro1.com. Só que o domínio conhecido do banco é www.brasileiro.com.br. Se você tem dúvida sobre este link, vá até o Google e mande executar uma busca.

Surpresa: é o site do banco mesmo! Mas, preste mais atenção: o logotipo do banco e as indicações de seus serviços aparecem no meio de uma página branca, ou boiando no meio de uma borda que tenta imitar a cor do banco.

O site é mesmo falso. O cracker utilizou uma vulnerabilidade dos bancos de dados SQL e dos serviços do ISS (o servidor Web da Microsoft), para fazer com que a página inicial do banco seja redirecionada para a página falsa. Não vamos ensinar, obviamente, como isso é feito, mas a técnica é incrivelmente simples.

Se o nome que aparece no e-mail é o nome verdadeiro da instituição ou domínio que ela utiliza, faça mais um teste: instale um outro navegador que não o Internet Explorer (pode ser o Mozilla ou o Opera) em sua máquina e copie o link apresentado pelo e-mail na barra de endereços do navegador. Observe o endereço na barra de endereços abaixo da tela da página. Se o endereço não bater com o apresentado no link, você está diante de uma página falsa.

Por último, desconfie dos links que terminam com extensão .exe, .com ou .bat. Elas contêm arquivos executáveis ou programas que, quando executados, podem fazer downloads muito rápidos e infectar sua máquina com trojans ou keyloggers (softwares que registram tudo o que o usuário digita e enviam para um determinado endereço de e-mail ou endereço de FTP). Para tirar a limpo, experimente copiar a página para a barra de pesquisa do Google e fazer o download do arquivo para o disco rígido sem abri-lo (novamente, tente usar o Opera ou o Mozilla em vez do Internet Explorer). Por último, tente fazer com que o arquivo seja examinado por um antivírus.

DADOS PESSOAIS VIA ICQ E MSN

Depois do Blog, e mais recentemente do Orkut (a rede de relacionamentos criada pelo Google), os *instant messengers* foram, se não uma das maiores revoluções da Internet, ao menos a “febre” que mais conseguiu atrair usuários de todos os tipos e locais do planeta. Principalmente nos escritórios com muitas filiais e para as pessoas que possuem familiares e amigos morando em outros países, os programas de mensagens acabaram por substituir, com vantagens óbvias, até os telefones, já que softwares como o MSN e o ICQ possuem embutidos recursos de voz, que são muito melhorados com a utilização de serviços de banda larga.

Informações pessoais


E estes são, justamente, os *instant messengers* mais populares: o ICQ e o MSN. O ICQ é praticamente o padrão deste tipo de comunicação na Web – o próprio AIM da AOL, por menos que se queira admitir, já era muito parecido com o ICQ, e ficou mais parecido ainda depois que a AOL comprou-o da empresa israelense que o havia desenvolvido. Mas, por outro lado, o MSN vem tomando terreno, principalmente depois das duas últimas versões do software, que trazem a possibilidade de compartilhamento de músicas e imagens – isso sem contar com a tradicional e milionária campanha de marketing promovida pela Microsoft para que as pessoas adotem o produto.

Ao fazer seu cadastro em qualquer um desses serviços, você deve preencher, obrigatoriamente, uma página com informações pessoais. Por exemplo, ao querer utilizar os serviços do MSN, você deve baixar o MSN Messenger (http://g.msn.com.br/7MEPT_BR/PT-BR/SETUPDL.EXE). Além disso, você deve ter uma conta do Hotmail, em cujo formulário é necessário fornecer alguns dados pessoais.

Mas só está aqui a semelhança entre o ICQ e o MSN nesse quesito. Os dados pessoais repassados ao Hotmail não são acessíveis via MSN Messenger, a não ser que você crie um perfil para ser visto por outros usuários. Se este perfil for criado, basta que alguém clique sobre o botão direito no “bonequinho” do MSN que corresponde a você para ser redirecionado para uma página do MSN com várias informações a seu respeito.

É claro que o ideal é não preencher esse perfil – apesar de o Google ainda não retornar resultados guardados no diretório do Hotmail, é óbvio que isso não demorará muito para acontecer. Além disso, o nome que você escolher para exibição é o que será mostrado aos outros usuários. Portanto, se você digitar seu nome verdadeiro na identificação do MSN, é ele que será mostrado.

A situação do ICQ é um pouco mais complexa. Todos podem saber seu nome se você for distraído o suficiente para colocá-lo nas opções de visualização. Além disso, tanto no endereço oficial do ICQ (<http://www.icq.com>) quanto pelo sistema de busca do próprio software, é possível encontrar qualquer pessoa (mesmo que ela não esteja conectada) clicando no botão **Search**. Neste caso, a falta de bom



senso novamente é a maior ajuda para a insegurança: os dados pessoais como endereço, telefone, número de celular e e-mail não são obrigatórios; apesar disso, muitas pessoas preenchem tais informações, que podem ser vistas por qualquer um que tenha acesso à Internet.

Fechando o ICQ

Outra brecha de segurança do ICQ, na qual a maioria dos usuários “escorrega” é deixar sua conta aberta para que qualquer pessoa possa capturá-la e conversar com você, mesmo que você não deseje conversar com ela. Mas isso pode ser facilmente resolvido. Clique no menu **Security & Privacy**, aba **Security** e, em **Change contact authorization**, mude o padrão para *My authorization is required*. Assim, você não poderá ser adicionado a nenhuma lista sem autorização prévia.

Listas de usuário e mensagens guardadas

Fazer backups é uma prática sadia, mas certas coisas não deveriam ser guardadas – ou, pelo menos, deveriam ser muito bem guardadas. No ICQ, por exemplo, é possível guardar toda a lista de contatos, além de todas as conversações realizadas, para evitar possíveis acidentes ou formatar a partição na qual o ICQ está instalado. Para fazer isso, basta copiar a pasta **NewDB**, localizada na pasta do ICQ do diretório **Arquivos de Programas**, para um outro local. O problema é que a maioria dos usuários copia essa lista para seu FTP pessoal ou seu site, esperando para resgatar depois. É só procurar no Google por *newDB* para ver quantos estão à disposição.

Da mesma maneira, é possível ler todas as conversações feitas em um ICQ: para isso, basta clicar com o botão direito do mouse no botão do ICQ e clicar em **Message Archive**. Estas conversas são guardadas em pequenos arquivos de texto, também no diretório do ICQ.

Deixar esses arquivos salvos na Internet ou em pastas compartilhadas em seu computador equivale a publicar sua lista de contatos e suas conversações pessoais ou profissionais. Para evitar isso, salve os backups de sua lista e suas conversações (se você precisa mesmo guardá-las) em um disquete ou CD regravável. Além dos riscos da Internet, qualquer usuário do ICQ mal-intencionado pode, com um programa de *brute force* específico para ICQ, invadir seu ICQ online e roubar tanto sua lista de usuários, quanto seus logs de conversas salvos localmente.

COMPLETE SUA COLEÇÃO

VISITE NOSSA LOJA VIRTUAL EM: [HTTP://WWW.LOJADIGERATI.COM.BR](http://www.lojadigerati.com.br)

PARA MAIS INFORMAÇÕES: (0xx11) 3217-2600

Coleção Segredos



Corel Draw

Um livro que ensina passo a passo, de maneira rápida e descomplicada, como utilizar os principais recursos do software, sendo ideal, portanto, para usuários iniciantes, mas trazendo, certamente, informações valiosas para aqueles que já têm conhecimento prévio desta ferramenta e que podem ser considerados usuários avançados.

Photoshop

Escrito pelos criadores da revista britânica *Computer Arts*, – considerada a melhor publicação de design no mundo –, o livro traz exemplos práticos e dicas de profissionais para explorar todo o potencial da ferramenta. Indispensável para designers.



Coleção Games



Desvendando o Counter Strike

Guia completo com dicas, estratégias, táticas e trapaças para se dar bem no game! No livro você encontra: conceitos básicos, comandos de console, dicas avançadas, tipos de armamento, além de dicas para usar bots para treinar, instalar e configurar suas preferências em um jogo multiplayer, criar e adicionar seu próprio cenário, táticas para contra-terroristas e terroristas, e muito mais...

Counter Strike Detonado

Com este livro é possível, por meio de várias simulações, aprender impressionantes táticas e estratégias de combate e ainda conhecer dicas e truques avançados para ataques fulminantes, defesas eficazes, evoluções em território inimigo e mapas desafiadores.

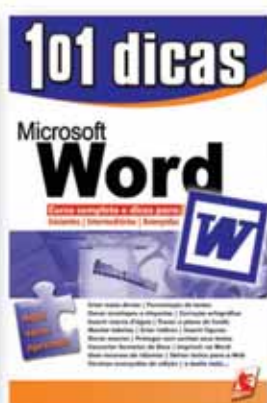


LIVROS EM DESTAQUE



Coleção 101 Dicas

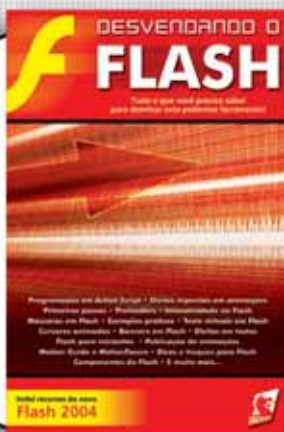
Curso completo e dicas para iniciantes, intermediários e avançados.



Coleção Desvendando

Microsoft Excel

Aprenda diversos recursos como filtragem de dados, tabelas dinâmicas, gerência de cenários, gráficos dinâmicos, filtros avançados, planilhas em 3D, funções lógicas, estatísticas, matemáticas e de textos, campos calculados e muito mais!



Flash

Ideal para quem deseja dominar esta ferramenta. Tudo o que você sempre quis saber, explicado por meio de exemplos práticos para criação de animações. Inclui as novidades do Flash 2004 e um capítulo ensinando ActionScript, a linguagem de programação capaz de adicionar poderosos recursos ao Flash.

Gravação de CDs

Ideal para iniciantes e avançados. Trata-se do maior guia já feito sobre o assunto. Aqui você aprende as principais técnicas para gravar CDs de áudio e dados. Guia completo para escolher o gravador, a mídia e o software de gravação. Inclui especial sobre DVDs, ensinando a escolher modelos de gravadores, tipos de mídia e a fazer autorias em seus DVDs.



Segredos do Google

Desvende os recursos não revelados do poderoso sistema de busca!

Hackeando o Google

- Guia para pesquisas avançadas
 - Uso do Google em celulares e palms
 - Busca de imagens e ferramentas de tradução
 - Uso dos caracteres-coringa
 - Seu site em primeiro lugar no Google
- E muito mais...

Investigação na Web

- Monitoramento de informações pessoais
 - Acesso a banco de dados via Google
 - Rastreamento de cartas e entregas
 - Busca por dados sobre os donos de sites
 - Consulta a dados da lista telefônica
 - Análise de log via Google
 - Password Generator usando o Google
- E muito mais...

E ainda:

- Como saber tudo sobre um site sem precisar invadi-lo
- Cuidado com os links falsos
- Como encontrar dados pessoais via ICQ e MSN



R\$ 9,90

ISBN 858953543-6



9 788589 535434