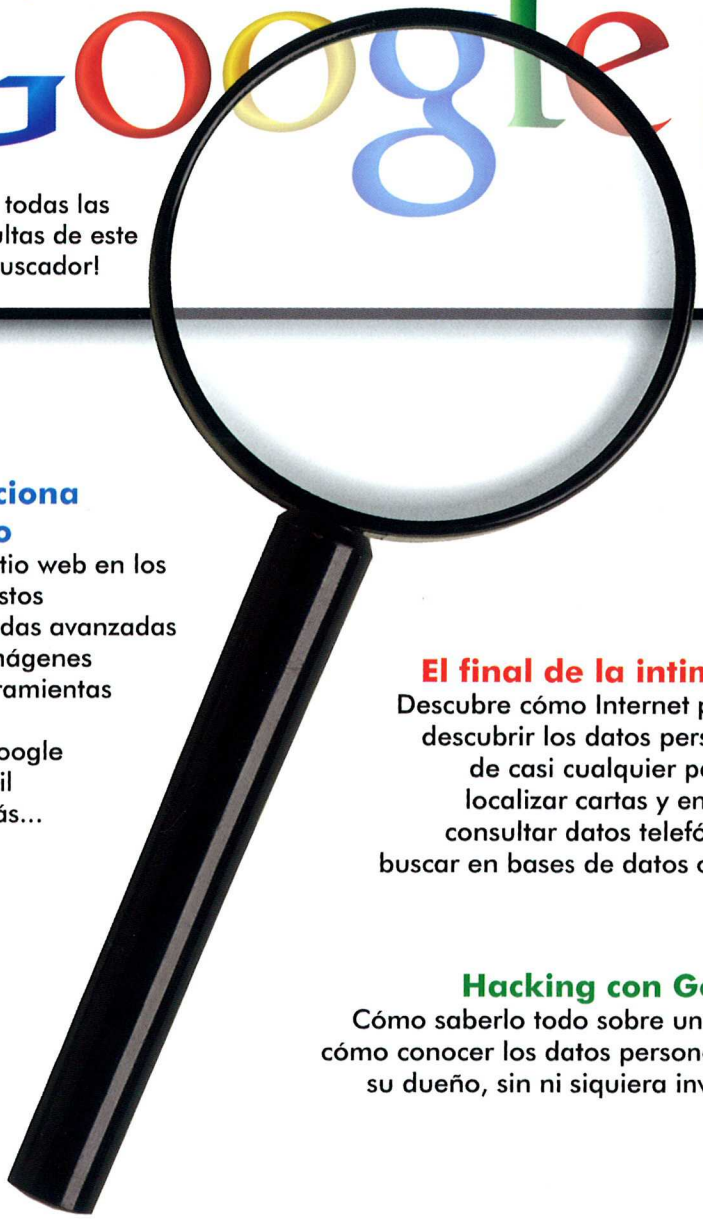
Los secretos de Google

SERIE POCKET 5

¡Descubre todas las opciones ocultas de este potente buscador!



Cómo funciona por dentro

- Coloca tu sitio web en los primeros puestos
- Haz búsquedas avanzadas de textos e imágenes
- Usa las herramientas del idioma
- Accede a Google desde tu móvil
- Y mucho más...

El final de la intimidad

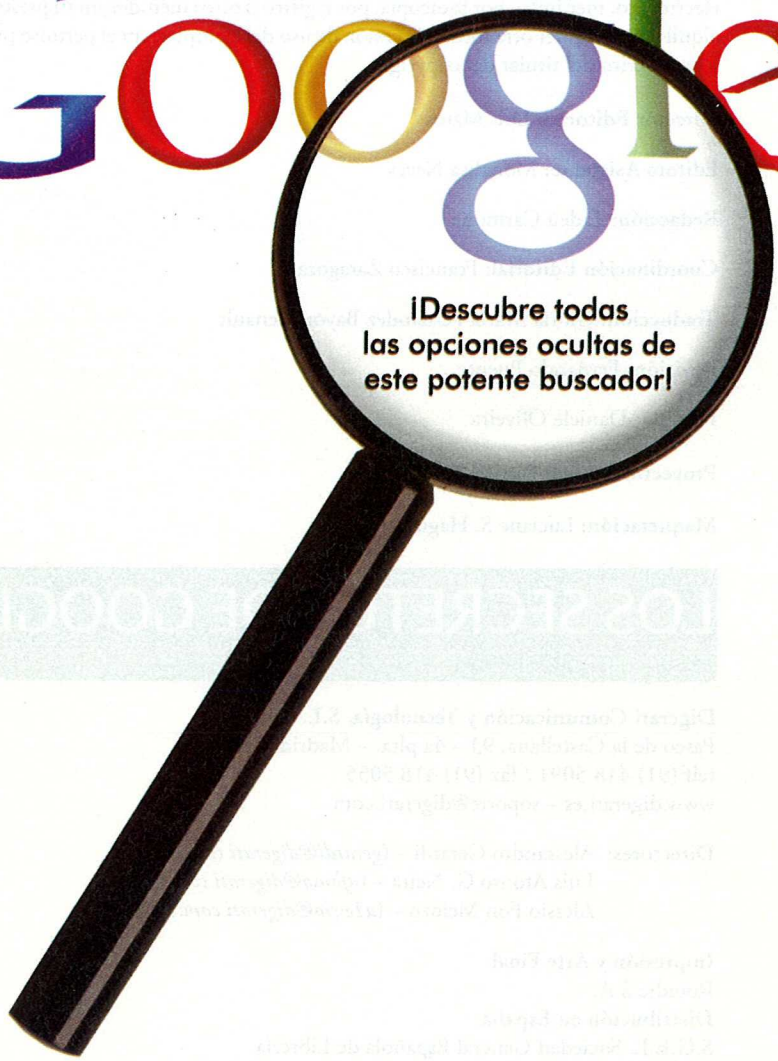
Descubre cómo Internet permite descubrir los datos personales de casi cualquier persona, localizar cartas y entregas, consultar datos telefónicos y buscar en bases de datos ocultas.

Hacking con Google

Cómo saberlo todo sobre un sitio, y cómo conocer los datos personales de su dueño, sin ni siquiera invadirlo.

DIGERATI BOOKS

Los secretos de Google



DIGERATI BOOKS

Director Editorial: Luis Matos

Editora Asistente: Monalisa Neves

Redacción: Tadeu Carmona

Coordinación Editorial: Francisco Zaragoza

Traducción: Gloria María Fernández Bayón Renault

Revisión: Fernando Puente

Portada: Daniele Oliveira

Proyecto Gráfico: Patricia F. Ferrari

Maquetación: Luciane S. Haguihara

LOS SECRETOS DE GOOGLE

Digerati Comunicación y Tecnología, S.L.
Paseo de la Castellana, 93 – 4a plta. – Madrid – España
telf (91) 418 5091 / fax (91) 418 5055
www.digerati.es – soporte@digerati.com

Directores: Alessandro Gerardi – (*gerardi@digerati.com.br*)
Luis Afonso G. Neira – (*afonso@digerati.com.br*)
Alessio Fon Melozo – (*a1essio@digerati.com.br*)

Impresión y Arte Final
Rotedic S.A.
Distribución en España
S.G.E.L. Sociedad General Española de Librería
Avda. Valdeparra 29 – Polígono Industrial
Alcobendas, Madrid 28108

D.L.: M-42885-2004
ISBN: 85-89535-43-6

PREFACIO

Parece un listín telefónico de China. Son nada menos que 1.300 millones de páginas las que están reunidas en un único lugar. El número de consultas que recibe es más de cuatro veces el número de habitantes de España. ¿Difícil de adivinar? Pues aquí tienes una pista más: su nombre se creó modificando las letras de la palabra *googol*, que fue inventada por Milton Sirotta, un sobrino del matemático americano Edward Kasner, para designar a un 1 seguido por 100 ceros. Ya deberías tenerlo chupado...

Su uso está tan difundido que ya ha creado un verbo muy difundido en todo el mundo: "googlear". Sinónimo de investigación, Google es un poderoso sistema de búsqueda en Internet que permite un acceso muy fácil y rápido a cualquier tipo de información que haya disponible en la Gran Red, desde cualquier rincón del mundo.

Es cierto que antes de Google hubo varios y dignos antecesores, que sobreviven hoy, como Altavista, Yahoo, Hispavista y similares. Pero la verdad es que es innegable su superioridad ante toda esa competencia. El fenómeno Google alcanza tales dimensiones que motiva, incluso, la creación de este libro, que hemos enfocado hacia la enseñanza sobre cómo explotar todo el potencial y las peculiaridades que ofrece este potentísimo sistema de búsqueda.

Este y otros buscadores de información en red son tan potentes que también comportan ciertos problemas, ya que gracias a Google y a herramientas similares es casi imposible conservar intacta la privacidad en la Internet. Basta con encontrar el nombre de una persona en la red, y acceder luego a algún servicio como las Páginas Blancas (http://www.paginasblancas.es/), donde metiendo el nombre de esa persona podrás obtener su dirección y su teléfono. Se despiertan así los temores, a veces con justificación, al Gran Hermano de Orwell...

Pero esta misma tecnología de búsquedas masivas que Google simboliza permite cosas muy buenas como, por ejemplo, navegar a través de un auténtico museo que reúna la mayoría de las páginas web desde la fecha de su creación. ¿No te lo crees? Entra en www.archive.org, escribe la dirección del sitio que quieres investigar, y deléitate con toda la colección de páginas y de diseños que se han sido usando a lo largo de sus años de existencia. Sin duda, algo impresionante...

Hemos escrito este libro para quienes quieren conocer los recursos de Google y de otras herramientas similares, y también para quienes quieren investigar en la red o desvelar los secretos de una persona o de una empresa en la Web.

En resumen, el público de la calle, el usuario medio, es el que, como tú, quiere acceder a la velocidad de la luz a los datos, los hechos y las personas. Y es que cada vez está más clara la diferencia entre quienes conseguirán el éxito y los que no: la información. Como verás en las páginas siguientes, el viejo dicho "el conocimiento es poder" nunca ha sido tan cierto...

Luis Matos
luismatos@digerati.com.br
Diretor Editorial

Índice

LOS SECRETOS DE GOOGLE

PRIMERA PARTE

Nociones básicas de Google 06
Hacer búsquedas avanzadas en Google 14
Google fuera de un ordenador 23
Calendarios 27
Fechas y husos horarios en Internet 31
Documentos dinámicos con Google 35
Poner tu sitio en el primer puesto en Google 38
La vuelta al mundo en 80 días, con Google 45

SEGUNDA PARTE

¡Tú también estás en Google! Este buscador da
acceso a miles de datos personales y profesionales 50
Uso de caracteres comodines 52
Descubrir quién es el dueño de un sitio 56
Consultar el listín Telefónico 59
Analizar logs usando Google 62
Rastrear cartas y entregas 65
Google, el amigo de los Crackers 68
Password generator con Google 73
"Husmeando" en bases de datos mediante Google 76

TERCERA PARTE

Conocimientos esenciales 80
Búsqueda de dominios: ¿cómo saberlo
todo sobre un sitio, sin invadirlo? 83
Las mil caras de Google 87
Cuidado con los enlaces falsos 91
Datos personales vía ICQ y MSN 93

Primera Parte:

Bienvenido al Universo Google

NOCIONES BÁSICAS DE GOOGLE

Las herramientas de búsqueda, o buscadores, son algo relativamente antiguo en el mundo de la informática. En el viejo MS-DOS, por ejemplo, bastaba con escribir:

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

para poder ver en pantalla todas las líneas de texto en las que se encontrase la palabra "deuda". La búsqueda se restringía al archivo que se indicaba; en nuestro ejemplo, **banco.rtf.** Nada, aparte del documento que se señalaba, se incluía en la búsqueda. Sin embargo, esto ya era algo, en un tiempo en el que los ordenadores se usaban para poco más que para sustituir a la calculadora y a la máquina de escribir.

Los sistemas de búsqueda de archivos más modernos, como los que se encuentran en los sistemas operativos Windows, MacOS y Linux, son mucho más completos: se pueden buscar datos en todo el ordenador, en una red a la que estemos conectados e incluso en todo un directorio almacenado en la Web, con solamente unos cuantos clics de ratón.

Esta manera de buscar información, ya con cierta utilidad, podríamos compararla con la omnisciencia (la capacidad de saberlo y conocerlo todo), aunque que limitada al interior de tu casa y, como mucho, a la acera de enfrente, lugares en todo caso que ya se supone que conoces bien. Del mismo modo, saber lo que hay almacenado en tu propio disco duro o en tus CDs de copia de seguridad debería ser algo sencillo y natural si la mayor parte de la gente fuera mínimamente organizada, algo que por desgracia no se corresponde ni mucho menos con lo que ocurre en la realidad. Los sistemas de búsqueda tradicionales rumian por tanto los datos que ya están en realidad a tu alcance, o que al menos con un mínimo esfuerzo podrías deducir dónde buscar.

La información está ahí fuera

A mediados de la década de los 80 los servicios "on line" eran ya relativamente conocidos, al menos en los Estados Unidos y entre el público más tecnófilo. Además de Internet (un proyecto en origen militar, pero que empezaba a derivar en una red de servicios entre universidades americanas y europeas), había ya diversos servicios de directorio que permitían el acceso a números de teléfono, a catálogos de bibliotecas y de departamentos públicos, así como a las descargas de archivos. En el caso de Internet, era imprescindible ser universitario, militar, o trabajar sin serlo en uno de esos dos sectores. Los servicios de directorio "alternativos", como el Minitel francés, podían ser subscritos por cualquier persona, siempre y cuando uno se comprometiese a pagar unas tasas mensuales del servicio, que en Francia eran baratas, pero que en el resto de países estaban al alcance de muy pocos.

Fue en 1995 cuando en la Universidad de Stanford se conocieron **Sergey Brin**, entonces con 23 años de edad, un especialista en diseño de aplicaciones Web y

graduado en Ingeniería Electrónica, y **Larry Page**, de 24 años, un especialista en tratamiento de datos y licenciado en Informática y Matemáticas. Ambos estaban muy interesados en acceder al curso de Doctorado en Ciencias Informáticas de esta Universidad, uno de los más conocidos en los Estados Unidos, y tal vez del mundo, y acabaron por descubrir que además de éste tenían otros muchos puntos de interés común, entre ellos, un ambicioso proyecto de crear un algoritmo de extracción de datos que permitiese la recolección y catalogación de grandes volúmenes de información.

Este proyecto se restringía de momento a los sitios y las bases de datos de la Biblioteca Digital de la Universidad de Stanford. Así que para que fuese posible hacer búsquedas fuera de la intranet de Stanford habría que construir un nuevo motor de búsqueda, que tuviese capacidad para leer páginas de toda, absolutamente toda, la Web, rastreando información, y creando después largas tablas de enlaces, acompañadas por el texto de cada una de las páginas que se indexaban.

¿Qué es un motor de búsqueda?

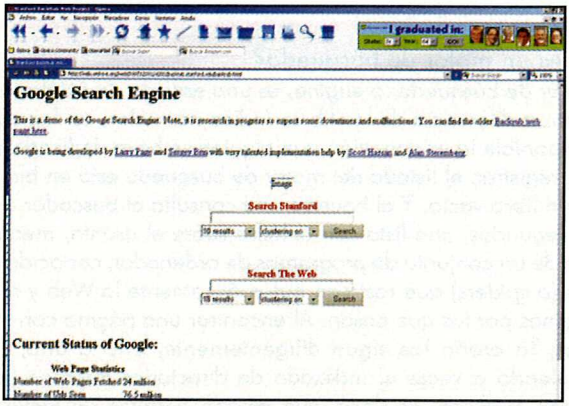
Un motor de búsqueda, o *engine*, es una especie de catálogo mágico. Pero, a diferencia de los libros-índice normales, en los que sólo está disponible la información que alguien se haya dedicado a catalogar y registrar, el listado del motor de búsqueda está en blanco, es como un libro vacío. Y al hacerse una consulta al buscador crea, en pocos segundos, una lista con enlaces sobre el asunto, mediante el trabajo de un conjunto de programas de ordenador, conocidos como arañas (o *spiders*) que rastrean automáticamente la Web y registran las páginas por las que pasan. Al encontrar una página con muchos enlaces, la araña los sigue diligentemente, uno a uno, incluso consiguiendo a veces el indexado de directorios internos (siempre que sean públicos, es decir, que tengan permiso de lectura para usuarios externos), de sitios web en los que aún se está trabajando. Los motores de búsqueda más refinados son incluso capaces de saber letra a letra qué actualizaciones se han producido en un sitio, usando este método de escáner.

Fue así como en 1996 Larry y Sergey lanzaron BackRub. Basado en Java y Phyton (puedes ver un enlace sobre una duda que envió Larry a un newsgroup especializado en esta dirección: http://groups.google.com/groups?hl=en&lr=&ie=UTF-8&safe=off&threa dm=page-0701962007020001@qwerty.stanford.edu&rnum=1&prev=/groups?selm=page0701962007020001@qwerty.stanford.edu) BackRub funcionaba en algunas máquinas Sun e Intel que se encontraban en el interior de la Universidad. Tanta tecnología y potencia (para esa época) marcaba ya la principal diferencia de su programa respecto a otros mecanismos de búsqueda que habían surgido durante ese año y medio en el que el

antecesor de Google estaba siendo incubado: BackRub era capaz de buscar entre los enlaces enumerados dentro de las páginas que se iban buscando, aumentando así considerablemente el número de resultados.

El nombre de Google se adoptaría en 1997, el mismo año en el que el proyecto dejó de utilizar las máquinas de Stanford. Google funcionó con alfileres hasta la primera mitad de 1998 cuando, impulsados por una compra "milagrosa" de varios terabytes de disco a un precio muy bajo, los socios del que sería el buscador número del mundo decidiesen crear el CPD (Centro de Procesamiento de Datos) de la empresa en la casa de Larry (más concretamente, en su dormitorio).

A finales de 1998, Google Inc fue fundada oficialmente, utilizando como capital (y como recursos para saldar las deudas que se habían contraído comprando equipos) 100.000 dólares ofrecidos por Andy Bechtolsheim, uno de los fundadores de Sun, y un millón más donados por una legión de amigos y parientes. En esos momentos, Google era ago así:



¿Qué significa Google?

Al buscar el término "Google", utilizando el propio buscador encontraremos más de 54.200.000 citas páginas. Pero, curiosamente, ninguno de estos enlaces nos explicará lo que quiere decir la palabra Google. Esto se debe a que "Google" es un término creado a partir del vocablo inventado "googol", una creación del Dr. Edward Kasner, de la Universidad de Columbia. El Dr. Kasner quería bautizar, con un nombre sonoro y fácil de recordar, la centésima potencia del número 10, o lo que es lo mismo: un número 1 seguido por 100 ceros. No satisfecho con este absurdo nombre, el científico creó "googol-plex", que equivale a un googol seguido por otro googol de ceros. Sea como sea, la única utilidad conocida del googol, desde el mismo momento de su invención, ha sido la de servir como inspiración para

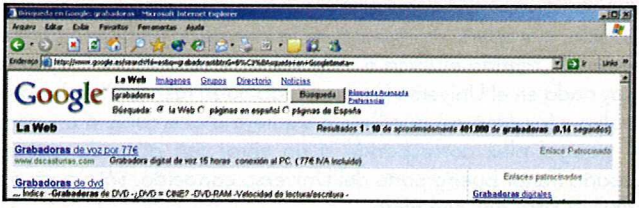
el Google, que quería representar así la idea de un número grandísimo, símbolo del mundo elástico e inagotable de la Web. Al fin y al cabo no hay nada en el Universo (exceptuando quizá las estrellas, los granos de polvo y los átomos) que llegue siquiera a acercarse a un googol. Y el googol-plex corresponde a un valor tan absurdo que sería necesario llenar buena parte del Universo conocido, sólo para poder escribir la cifra por completo.

Cómo funciona Google

Google es, esencialmente, un mecanismo de búsqueda de palabras y enlaces en toda Internet, que utiliza varios recursos de filtrado y de catalogación de resultados. Pero, ¿qué fue lo que garantizó el éxito de Google, habida cuenta de que ya existía una gran variedad de buscadores consolidados cuando salió al mercado? Además de sus algoritmos de búsqueda y extracción de datos, que hacen que su búsqueda sea mucho más rápida que otras realizadas por sus competidores, su sencilla interfaz ayudó mucho a su expansión.

El *front-end* de Google se compone únicamente por texto y por enlaces, y todo en HTML, lo hace que la página que devuelve la búsqueda se cargue casi de inmediato, incluso en combinaciones de equipo-programa muy antiguas, como un 486 con el Internet Explorer 4, o con navegadores muy sencillos, como el navegador en modo texto Links, que aún utilizan varias distribuciones Linux. Esta liviandad, además de hacerle la vida más fácil al usuario, ha permitido que Google no se haya caído nunca, que nunca haya "salido del aire" por problemas de uso, y eso, teniendo en cuenta que recibe unos 200 millones de consultas diarias.

Hacer una búsqueda en Google es tan simple como acceder a su página, tecleando www.google.es. Basta con escribir el término que queremos para la búsqueda: supongamos que quieres buscar alguna cosa sobre *grabadoras*. Escribe ese término en el cuadro de búsqueda y pincha luego en **Búsqueda en Google**, o pulsa el botón Intro.



Como puedes ver en la figura superior, Google devuelve unos cuantos miles de resultados. Este suceso es una consecuencia matemática, muy relacionada con la teoría de conjuntos: cuanto menor es el número de delimitadores en una consulta a un conjunto de elementos, mayor será el número de elementos que veremos en el resultado final. Un ejemplo sencillo: cuanto menos selectivos seamos al elaborar una lista de invitados para una fiesta (los "gordos", los "delgados", y "puede traer un amigo", además de "no necesita traer regalo"), más posibilidades tendremos de acabar invitando a toda la ciudad.

Esta es otra característica que lo diferencia del resto. Su motor de búsqueda es capaz de añadir sin demasiado esfuerzo varios filtros, tanto en nuevas búsquedas como en las que ya hemos realizado. Vamos a conocer algunos de ellos:

La lógica booleana

Otra característica que lo hace sobresalir sobre el resto de buscadores es su "engranaje" lógico, que utiliza la **lógica booleana**. Bautizada así en homenaje al matemático británico George Boole, este sistema permite la posibilidad de hacer una búsqueda de palabras en un texto estableciendo condiciones para la visualización de los resultados en función de una serie de valores lógicos que se sujetan a unas reglas:

– un valor debe ser siempre verdadero o falso,
– un valor no puede ser verdadero y falso al mismo tiempo,
– numéricamente, verdadero puede ser definido como "1", y falso como "0".

Con el fin de no limitar la búsqueda en exceso se pueden utilizar para incrementarla y precisarla varias funciones booleanas, o sistemas de búsqueda, que Google usa, pero que son habituales en los sistemas de búsqueda. De entre ellos, los más conocidos son:

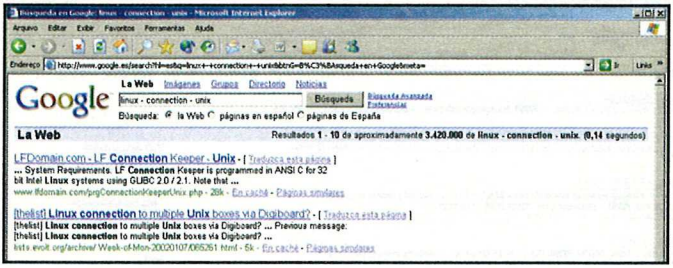
AND – Este operador se utiliza para incluir, en una búsqueda, todos los elementos que van a ser propuestos en una consulta. Este delimitador se utiliza en su forma booleana en la mayoría de los buscadores, como por ejemplo, en Altavista.

Google no utiliza el delimitador AND de una forma explícita, sino que lo da por introducido cuando no se especifica nada más: basta con escribir varios términos separados por un espacio para que sean buscados conforme a esta condición.

OR – Delimitador de variable. El OR (siempre en mayúsculas) se utiliza para encontrar páginas en las que exista uno u otro término (y por tanto, no necesariamente ambos) de la búsqueda. Al introducir, por ejemplo, los términos *libros* y *Digerati* separados por OR, tendremos acceso a todas las páginas en las que aparezca cualquiera de esos dos términos, estén o no ambos incluidos en una misma web. Este delimitador es ideal para términos que son poco comunes, como expresiones científicas o literarias.

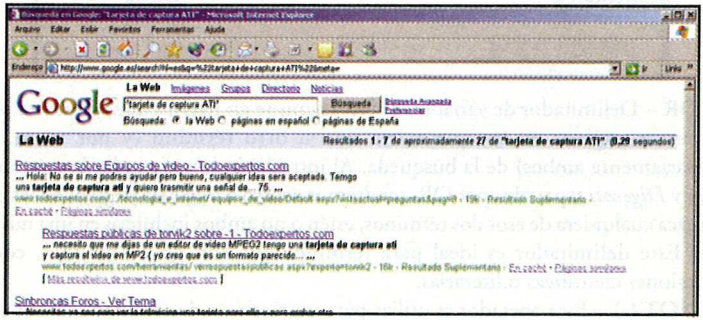
NOT (-) – Este operador se utiliza para suprimir un determinado término de una búsqueda, así que sirve como una especie de filtro de contenido. En Google se utiliza escribiendo el signo negativo o signo "menos", (-) colocado delante de un término. Para buscar, por ejemplo, el término Linux, pero si queremos excluir los términos Conectiva y Unix, utilizaremos el siguiente comando:

y obtendremos el siguiente resultado:



COMILLAS (“ ”) – Las comillas se utilizan, en lógica booleana, para garantizar que una expresión completa (o un conjunto exacto de términos) se incluye en la búsqueda. Es muy útil cuando buscamos expresiones en español. Al buscar, por ejemplo, los términos *tarjeta de captura ATI*, sin comillas, obtendremos sólo como resultado páginas en las que aparezcan tanto la expresión completa como los términos *tarjeta* o *captura sueltos*, *pero* sin ninguna mención a *ATI*.

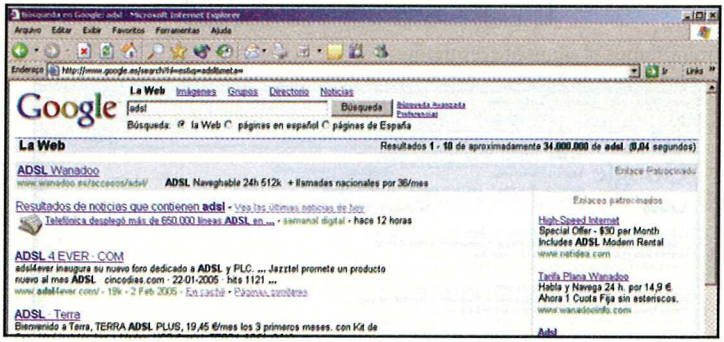
Utilizando, sin embargo, un conjunto de términos entre comillas, como “tarjeta de captura ATI”, crearemos una expresión exacta, que Google utilizará para delimitar el resultado de la búsqueda, como puedes ver debajo:



Otros delimitadores de búsqueda

Hay otros delimitadores de busca muy sencillos, pero que son propios de Google, y que no son frecuentes en otros sistemas. Los puedes encontrar justo debajo de la barra de búsquedas, en la línea que comienza por “**Búsqueda:**”.

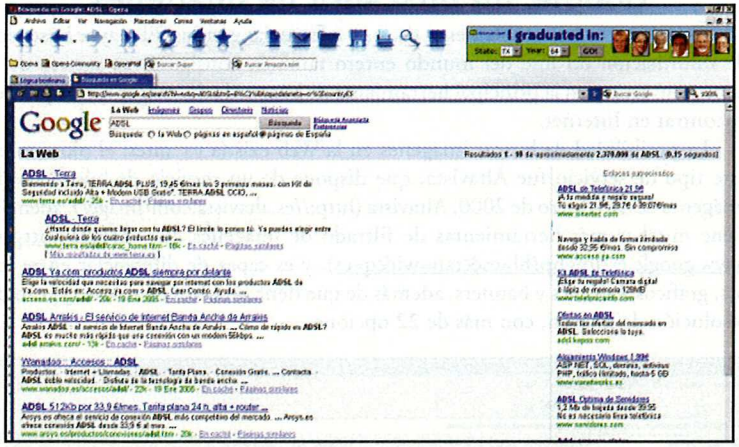
La función de estos campos es elegir entre hacer una búsqueda en toda la Web, buscando datos en todos los servidores que haya disponibles y que contengan páginas activas, o restringirla a un idioma o país determinado. Utilizando la primera opción al buscar términos estandarizados o internacionalizados, como por ejemplo *ADSL*, obtendremos una cantidad absurda de páginas, incluyendo algunas en holandés, ruso y japonés, incluso con caracteres no latinos:



Este diluvio de resultados, sobre todo de lenguas extranjeras, se puede contener. En la página principal de Google haremos clic en la opción *páginas en español*, para recibir solamente resultados de páginas que estén escritas en nuestra propia lengua.

Incluso así, los resultados contendrán aún demasiadas páginas de Latinoamérica (y algunas de otros lugares), así que estaremos perdiendo el tiempo si estamos buscando, por ejemplo, servicios públicos o si buscamos una tienda de nuestra ciudad.

Para resolver este segundo problema, Google dispone de una tercera función: *páginas de España*, que hace que la búsqueda se centre solamente en páginas que estén en servidores españoles o que terminen con el sufijo .es.



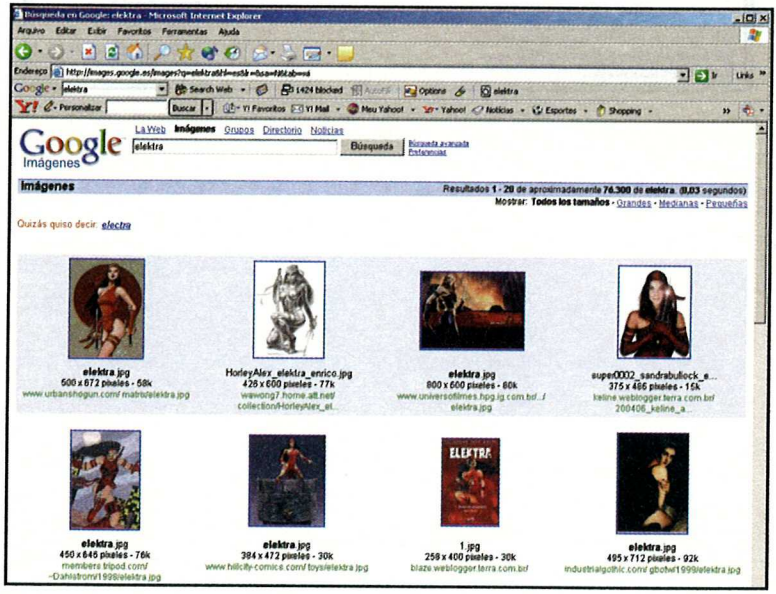
HACER BÚSQUEDAS AVANZADAS EN GOOGLE

Todo lo que se hemos visto en el capítulo anterior no supone ni siquiera el 2% de lo que Google puede hacer. Hay muchas más opciones de búsqueda avanzada (de las cuales, la mayoría no son accesibles desde la página principal del buscador), que pueden ser utilizadas para potenciar y refinar tus búsquedas, y encontrar así muchas más cosas que simples enlaces y páginas Web.

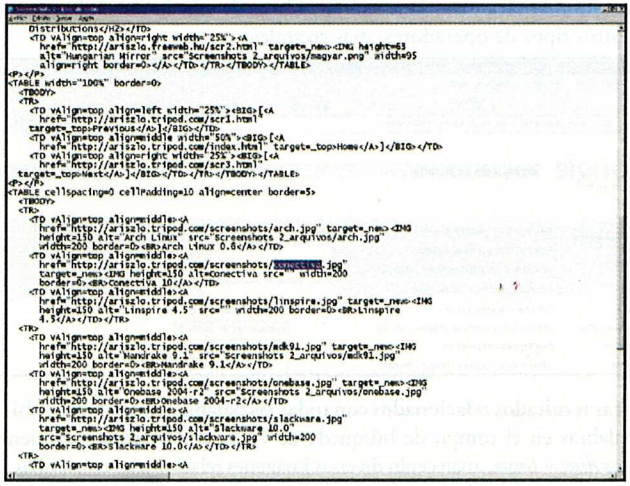
Google: un buscador de imágenes

Google no contiene solamente textos y referencias escritas. El mayor buscador de información on line del mundo entero también consiguió, en menos de 2 años, convertirse en la principal herramienta de búsqueda de imágenes que puedes encontrar en Internet.

La posibilidad de buscar imágenes en la Web existía ya antes: el pionero en este tipo de servicio fue Altavista, que dispone de un servicio de búsqueda de imágenes desde febrero de 2000. Altavista (http://es.altavista.com/image/), además, tiene muchas más herramientas de filtrado de imágenes que Google (http://www.google.es/imghp?hl=es&tab=wi&q=es), y es capaz de diferenciar entre fotos, gráficos, botones y banners, además de que tiene un filtro de colores y otro de resolución de imagen, con más de 22 opciones.



¿Por qué, entonces, al hacer una búsqueda de imágenes relacionadas con el término Conectiva", Altavista encuentra "sólo" 577 imágenes, y Google más de 3.630? La causa no está en el sistema de búsqueda de imágenes, que es muy parecido: en realidad, lo que se busca son referencias textuales a figuras o elementos gráficos contenidos en páginas HTML. Al buscar imágenes JPEG, por ejemplo, Google o Altavista buscan el sufijo *.JPG* en el código de la página. Al buscar, entonces, el término Linux, Google identificará en una página cualquiera una imagen que corresponda a la búsqueda referida, así:

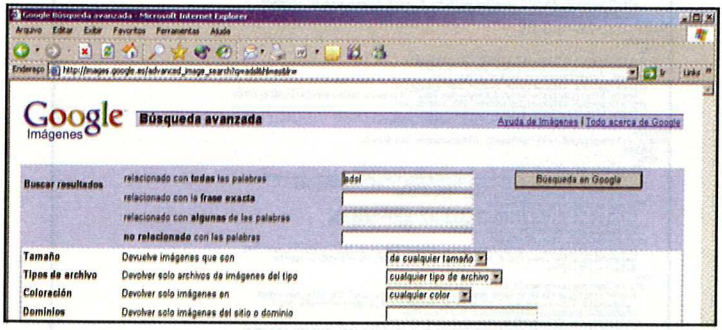


Como esto no ayuda mucho al usuario, Google intenta recopilar los datos necesarios para que el navegador de quien hace la consulta permita visualizar la página de resultados y las imágenes que se han encontrado en ella, como en la imagen del ejemplo inferior:

Búsqueda avanzada de imágenes

Junto al cuadro de búsqueda de Google hay un enlace llamado **Búsqueda avanzada**, que se utiliza para filtrar nuestra búsqueda. Como hemos dicho más arriba, esta búsqueda no es tan específica, por no decir tan perfeccionista, como la que ofrece Altavista, pero puede ayudar a ahorrar mucho tiempo, sobre todo a quienes necesiten un gran volumen de imágenes sobre un mismo asunto, en poco tiempo.

Al hacer clic, el usuario es enviado a la página *Búsqueda avanzada*, que se divide en dos partes: Buscar resultados (en azul), donde se puede refinar la búsqueda con operadores booleanos, y la sección de filtros de relaciones, en la que se pueden utilizar otros tipos de operadores, más complejos.



Buscar resultados relacionados con todas las palabras equivale a escribir sin más varias palabras en el campo de búsqueda de Google. Al buscar, por ejemplo, los comandos *disco* y *linux*, aparecerán diversas imágenes relacionadas con ambas palabras.

Buscar resultados relacionados con la frase exacta equivale a escribir las palabras entre comillas (" ") en la página de inicio de Google. Al buscar, por ejemplo, *Star Wars*, aparecerán todas las imágenes relacionadas con la saga que Google consiga encontrar.

Buscar resultados relacionados con alguna de las palabras equivale a hacer una búsqueda de texto utilizando el operador OR.

Buscar resultados no relacionados con las palabras es una herramienta que puede ser útil a la hora de excluir cierto término de una búsqueda, que podría de otra manera hacerse demasiado genérica. Al buscar *linux*, por ejemplo, podremos eliminar de la consulta todas las imágenes del simpático pingüino, la mascota de ese sistema operativo, poniendo en ese campo la palabra *tux*.

Otra utilidad: para crear un fondo de escritorio se necesitan imágenes con alta resolución, y sin embargo, cuando hacemos una búsqueda en Google, buscando por ejemplo *linux*, puede que nos encontremos desde imágenes gigantescas hasta iconos de 2 KB.

Para resolver este problema la búsqueda avanzada de imágenes dispone de varios filtros. En **Tamaño**, por ejemplo, podemos elegir el peso de las imágenes que queremos encontrar. Eligiendo *grandes* obtendremos imágenes de alta resolución, ideas para fondos de escritorio, diapositivas, etc...

Lo ideal para los fondos es utilizar imágenes JPG: una imagen de estas, en alta resolución, suele ocupar unas pocas decenas de KBs, mientras que su equivalente en BMP podría llegar a ocupar unos cuantos MBs. Para usar figuras solamente en formato JPG o PNG, utiliza en la búsqueda avanzada el campo **Tipos de archivo**. Por otro lado, si lo que buscas es imágenes en tonos de grises o en blanco y negro, puedes utilizar el filtro *Coloración*.

Para terminar, podemos ver todas las imágenes que haya indexadas bajo un determinado dominio. Al escribir, por ejemplo, el dominio *www.digerati.es*, obtendremos todas las imágenes que haya contenidas en la página inicial de este dominio.

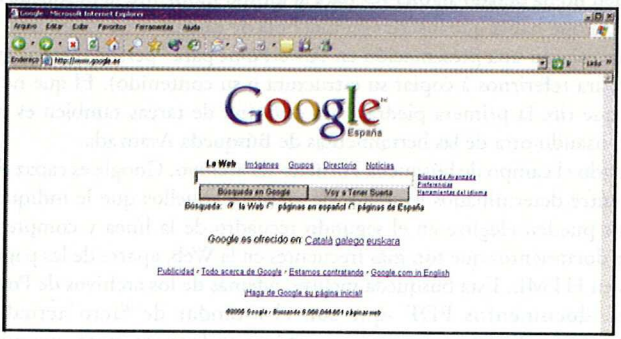
Encontrar imágenes ocultas

A pesar de que sean muy pesados, a veces podríamos necesitar algunos archivos en formato BMP. Aunque en el filtro *Tipos de archivo* no encontremos este formato, también podemos hacer búsquedas bajo esta extensión. Basta con escribir el tema de la búsqueda, por ejemplo *España*, añadiendo en el recuadro de búsqueda el operador *filetype:bmp*. Todos los archivos con esta extensión y que estén relacionados con el término España aparecerán listados en el directorio de imágenes de Google.

Búsqueda avanzada de textos

Hay también varios tipos de búsqueda avanzada para textos. Además de las tradicionales búsquedas mediante términos y páginas Web relacionadas con ellos, se puede buscar por tipo de documentos, como hojas de cálculo de Excel o archivos PDF, además de por títulos de páginas, entre otras muchas cosas.

Esas opciones, así como las que hemos visto en la búsqueda avanzada de imágenes, son accesibles a través de un enlace poco visible, "Búsqueda avanzada", que está a la derecha del logotipo del Google o del campo de búsqueda, según la página.



El área de búsqueda es muy parecida a la que vimos en la búsqueda de imágenes, y como ella, sustituye la utilización de parámetros booleanos por una interfaz gráfica. Para ver la explicación de cómo funciona, relee el punto anterior.

Está claro que hay además otras muchas herramientas que pueden utilizarse para realizar una búsqueda de información en Internet. Estos recursos no utilizan los operadores booleanos "tradicionales", sino algunas incorporaciones que son cosecha propia de Google y su equipo de desarrollo.

Idioma

En idioma podremos unir los elementos de búsqueda tradicionales con la posibilidad de ejecutar la búsqueda en una lengua determinada. Si por ejemplo, quisiéramos ejecutar una búsqueda sobre memorias flash de 32 MB, muy usadas en cámaras digitales baratas, ordenadores de bolsillo y teléfonos móviles, podemos simplemente hacer una búsqueda en **Buscar resultados con todas las palabras** utilizando el término inglés, *Flash card 32 MB*, ya que el idioma del Reino Unido puede considerarse como la lengua "oficiosa" (que no oficial) de la Web.

Supongamos ahora que queremos saber cuánto cuestan este tipo de memorias en Taiwan o en Hong Kong (sobre todo, para compararlo con los precios que te ponen en la tienda de tu barrio). Para hacer una búsqueda usando este parámetro, vete al campo **Idioma** y en **Producir páginas escritas en**, elige **Chino (simplificado)**.

Después de hacer clic en Búsqueda en Google, obtendremos como resultado un inmenso número de páginas en chino (puedes ver los caracteres en mandarín, en el título en cada página). La diferencia está en que al escribir los términos en inglés, has sido enviado a páginas con información en ambos idiomas, y te has librado de tener que aprender chino en una hora. Este procedimiento es válido para todos los idiomas y puede ser útil también para conseguir libros electrónicos (*e-books*) en otros idiomas, escribiendo *books* en la búsqueda y seleccionando el idioma que desees, como el griego o el francés, por ejemplo.

Formato de archivo

Quién no ha dejado alguna vez para el último momento la búsqueda de datos específicos que había que insertar en una hoja de cálculo o en un gráfico, o quién no ha necesitado una presentación en PowerPoint para "personalizarla" (un eufemismo para referirnos a copiar su estructura o su contenido). El que no lo haya hecho, que tire la primera piedra. Para este tipo de tareas también es muy útil Google, usando otra de las herramientas de Búsqueda Avanzada.

Usando el campo de búsqueda **Formato de archivo**, Google es capaz de discriminar, entre determinados formatos de archivo, aquellos que le indiques. Estos formatos pueden elegirse en el segundo recuadro de la línea y comprenden los tipos de documentos que son más frecuentes en la Web, aparte de las páginas web escritas en HTML. Esta búsqueda incluye, además de los archivos de PowerPoint y Excel, documentos PDF -que son el estándar de facto actual para la documentación digital en Internet- y archivos en formato texto que pueden ser colgados en la Web, como archivos DOC o RTF.

Supongamos por ejemplo que quieres hacer una búsqueda entre hojas de cálculo de Excel que detallen los gastos de combustible que realiza una pequeña empresa.

Además de insertar los términos *gastos* y *combustible* en el campo de búsqueda, irás a la línea **Formato de archivo**, seleccionarás la opción **Solamente** y después el tipo de documento que quieres buscar: hojas XLS, en nuestro caso. Al hacer clic en **Búsqueda**, encontrarás todas las hojas que haya publicadas en Internet y que necesites.

Un aviso: por descuido de muchos administradores de redes y sistemas, al implantar sus políticas sobre cómo compartir archivos, muchas hojas o documentos publicados en Web contienen datos que no deberían ser de dominio público. Así que ten cuidado con utilizar hojas, textos o presentaciones que hayas encontrado en Google como base para tus propios documentos, porque puedes tener problemas. Si por ejemplo, al buscar la mejor manera de crear una página de teléfonos para una Intranet (la red de informaciones interna) de tu departamento buscas teléfonos y documentos RTF, y acabas topándote con documentos con listas de teléfonos confidenciales, deja el documento donde lo hayas visto. O mejor aún, borra todas las cookies y el Historial de tu navegador, como si no hubieses utilizado siquiera Internet ese día.

Documentos al aire

Puede que no lo sepas, pero tus documentos privados podrían quedar expuestos por Google. Pasar por una situación de estas es, además, muy fácil, si te limitas a mantener intactos algunos parámetros de red e Internet en Windows después de instalarlo.

Antes de nada, compartir archivos en Windows es muy peligroso, ya que este sistema no diferencia entre compartir un directorio para tu red interna y para toda Internet, y es especialmente grave si el archivo que compartes está en una máquina con acceso directo a la red, tanto telefónico como de banda ancha. Pero si no hay otra manera, lo mejor es que configures las carpetas compartidas de la manera más segura posible.

En Windows 98/ME se recomienda activar una protección con contraseña para los directorios compartidos. Para ello, haz clic con el botón derecho del ratón sobre el directorio que elijas, pinchando después en **Compartir**. Luego haz clic en la opción **Depende de la contraseña** para las opciones del comando compartir. Windows XP, desgraciadamente, no funciona de esta manera. Para habilitar contraseñas en los ordenadores con este sistema operativo se necesita, antes de nada, cambiar algunos parámetros del sistema. Para ello, vete al menú **Inicio > Explorador de Windows** menú **Herramientas > Opciones de carpeta > Ver**, y desactiva la opción **Utilizar uso compartido simple de archivos (recomendado)**. Ahora sólo tienes que compartir el archivo y seleccionar la cantidad límite de usuarios, y también, claro, qué usuarios tendrán permiso para utilizarlos.

Si quieres crear un usuario específico, accede al **Panel de control**, haz clic en el icono **Usuarios** y después pincha en **Crear un usuario nuevo**. Crea un usuario invitado, por ejemplo *Red*. Los servicios del comando compartir de Windows 2000 (mucho mejores que los del XP) funcionan de una manera similar.

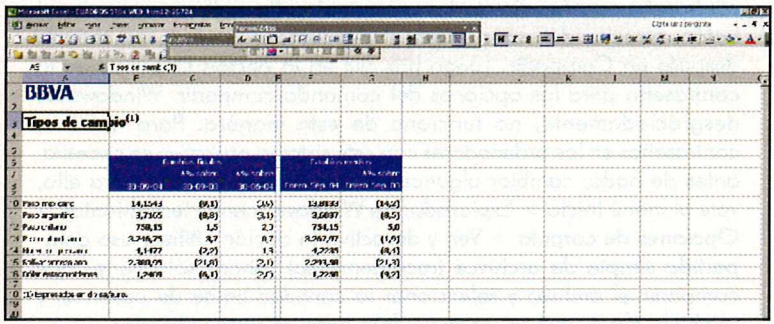
Fecha

El tiempo, después del dinero, es uno de los factores que más pesan en este mundo moderno. Además de hacer todo muy deprisa, a veces tenemos que atenernos a lo último que se ha hecho, o a lo más reciente que haya sobre un determinado asunto, dejando a un lado toda la información anterior.

En este punto, el contenido de Internet tiene algunos inconvenientes: teniendo casi una década de existencia pública, la Web ha conseguido acumular junto al contenido actualizado diariamente, cosas viejas, de sus inicios. Esto ocurre porque muchos servidores (principalmente servidores públicos) no se limpian nunca, o su contenido nunca se revisa. Esto hace que al realizar búsquedas sencillas, por ejemplo, con *Spiderman*, encontremos muchas cosas de la segunda película, que se ha lanzado hace poco, junto a noticias del lanzamiento de la primera parte, que son de hace un par de años.

Google tiene una herramienta que permite el filtrado por fecha. Aunque no tiene opciones para realizar una búsqueda mediante artículos o páginas de fechas definidas (algo que, sin embargo, no creemos que tarde mucho en ocurrir), este filtro puede facilitar mucho el trabajo a un usuario, sobre todo si quiere restringir la línea del tiempo porque... ¡él mismo no tiene tiempo!.

De este modo, cuando buscamos *valor de venta del dólar* en la página de **Búsqueda avanzada**, podemos ir a la línea **Fecha** y hacer un clic en: **en los últimos 3 meses** para obtener los datos más recientes de este tema. También se puede conseguir información de los **últimos 6 meses,** y del **último año.**



Presencia

Google asume que hay varias maneras de anunciar un tema en una página Web. Lo habrás notado al observar que te devuelve una búsqueda sencilla con enlaces tanto a contenidos de páginas, como a títulos de sitios o nombre con el que la página se aloja en el servidor. Basta con buscar términos cotidianos, como por ejemplo, *tartas de limón.*

Así y todo, hay muchos modos de dividir esta búsqueda por localización. En la página *Búsqueda avanzada*, Google presenta la herramienta **Presencia**, que permite la búsqueda tanto por elementos pertenecientes a la propia página, como su título o contenido, como por el servidor en el que la página está hospedada, o por cómo es reconocida por otros servidores de Internet. De este modo, buscar un sitio que venda tartas de limón es mucho más fácil (buscando por *Presencia, en el título de la página*), a no ser que prefieras hacer la receta en casa (*Presencia, en el contenido del texto*).

Dominio

Algunos dominios (la dirección universal que se le da a una página web, y exclusivamente a ella) hacen justicia a ese nombre, consiguiendo ser tan vastos que recorrerlos enlace a enlace en busca de una información podría llevarnos, tranquilamente, toda una vida. Que lo digan los programas de descarga, bien surtidos, pero no siempre muy bien organizados.

Piensa, por ejemplo, que necesites una versión reciente de Linux, como Gentoo. Buscando *Gentoo* o *Linux*, Google te devolverá tantos resultados que al encontrar una página que te sirva, probablemente la distribución ya no te servirá (o te habrás pasado toda una jornada de trabajo buscándola).

Para resolver este problema Google dispone de una herramienta de **Búsqueda avanzada, Dominio**. Para el ejemplo anterior, bastaría con hacer clic en la opción **Solamente** de la línea **Dominio**, escribir la dirección www.isolinux.org (el mayor archivo de sistemas basados en Unix que existe). Así conseguiríamos no sólo una versión sino muchas versiones del sistema concreto que buscamos.

Está claro que podríamos hacer el recorrido contrario, sustituyendo la opción **Solamente** por **No producir resultados del dominio o sitio Web**. Este tipo de busca puede ser utilizado para huir de lugares demasiado comunes, genéricos o incluso peligrosos, evitando que una búsqueda de *cartas* te lleve igual al sitio web de Correos que al de un casino virtual.

Similares en Google

En el capítulo anterior enseñamos varias opciones de búsqueda avanzada de Google, utilizando operadores booleanos diversos. Pero ese no es el único modelo de búsqueda que reconoce esta página: Google también es capaz de reconocer páginas **similares** e incluso de encontrar un "atajo" a una página, que esté dentro de otra página.

Veamos, por ejemplo el recurso **Similares**, que se encuentra en la página de **Búsqueda avanzada**, en el campo **Búsqueda específica a una página**. Buscar por medio de esta herramienta te permitirá encontrar páginas que puedan contener material parecido al de un sitio o dominio que ya hayas sondeado hasta la saciedad, y en el que no encuentres lo que andabas buscando, pero que se parezca bastante a lo que buscas. Si por ejemplo, buscamos libros publicados por Digerati y no conseguimos encontrarlos en una determinada librería, basta con ir a la línea **Similares, Encontrar páginas similares a la página,** escribiendo después la dirección de la librería en la que no encontramos nada. Aparecerá una lista de direcciones con la misma función o asunto del término que hayas buscado.

Esta herramienta también puede ser utilizada por personas que estén de viaje en el extranjero, o que necesiten informaciones sobre un determinado servicio que esté fuera de España. Si por ejemplo, escribimos en esta línea de búsqueda la dirección www.correos.es, Google nos devolverá las direcciones de varios servicios de correos del mundo, además de servicios de paquetería y de comercio internacional.

Rastrear enlaces

La mayoría de los métodos de búsqueda de Google buscan información directa. Ya hemos hablado de ello arriba, al buscar datos sobre la fecha, el título de la página, los tipos de archivos etc...

Pero además de eso, Google también es capaz de buscar informaciones indirectas, sondeando la Web, página a página, hasta encontrar una que enlace a una determinada web, en vez de acceder a ella directamente. Al hacer esto, el buscador realiza dos procedimientos interesantes: primero utiliza el Back Rub (que, como vimos, fue el primer mecanismo de búsqueda de Google). En segundo lugar, en vez de buscar enlaces con sus *arañas de búsqueda* hasta encontrar la página deseada, Google hace el trabajo "desde la mitad", limitándose a crear una lista de los lugares en los que se haya encontrado el sitio.

Esta herramienta puede usarse introduciendo la dirección web en la línea de **Enlaces**, en el campo **Encontrar páginas con enlaces a la página**. Al buscar, por ejemplo enlaces para el sitio del Ministerio de Hacienda (www.minhac.es), tendremos acceso a varias direcciones que se remiten a esta fuente. Otra pista interesante: si eres diseñador gráfico o administrador de Webs, puedes utilizar este recurso para ver quienes están enlazando a tu página desde la suya. A fin de cuentas, si administras de una página sobre joyas, no sería nada agradable ver tu enlace colocado en una página sobre seguridad patrimonial o sobre estadísticas de robos.

GOOGLE FUERA DE UN ORDENADOR

Visto desde fuera, Google es un universo aparte, como su propio nombre indica (mira el recuadro en el primer capítulo, para saber qué significa ese enigmático nombre). A pesar de estar "encerrado" en el mundo virtual de la red y de los ordenadores conectados a ella, Google podría, haciendo un ejercicio de ciencia ficción, verse en sí mismo como una gran entidad, si no real, al menos en construcción, al más puro estilo de The Matrix. A fin de cuentas, los viajes virtuales a través de este buscador por todo el mundo acaban desembocando en lugares reales (servidores administrados por personas que trabajan en periódicos, que después van a su casa, donde viven otras personas que también tienen que usar Google...), que son utilizados por personas reales (que interaccionan con cosas mucho más tangibles que Google) y que por fin, acaban buscando datos que suelen ser sobre cosas reales, como otras personas, casas, mercancías, etc...

Pensar en ello puede ser un poco difícil y abstracto (porque más allá de la pantalla del ordenador, no tenemos ninguna noticia de que Google exista). Pero está claro que necesitamos tener un PC o un ordenador portátil conectado a Internet, que esté anclado en algún lugar, para que Google se haga presente, por no decir real. Bueno, al menos esto era así hasta que despegó la informática móvil...

Google en tu móvil

El teléfono móvil ha sido un gran avance en las comunicaciones. Imagínate que después de casi un siglo de avances, las personas que estaban atadas a sus líneas de teléfono fijas, podían por fin, y a un coste relativamente bajo, ser encontradas en cualquier lugar, no sólo de su barrio o de la ciudad donde viven, sino del mundo entero (aunque es cierto que eso tiene algunos inconvenientes, sobre todo si le debes dinero a alguien...). ¿Un compañero de trabajo ha ido al extranjero y necesitas hablar con él urgentemente? Si se llevó con él el móvil (de la empresa, ¡está claro!), basta con llamarle para que te saque de dudas (sobre la empresa, ¡por supuesto!...).

En 1999, una asociación de grandes empresas de telefonía europeas tuvo una idea brillante: ¿y si además de usarlo como teléfono, también pudiese usarse para buscar o recibir información, como una especie de ordenador de bolsillo conectado a Internet? Fue así como nació el WAP.

WAP

El WAP o *Wireless Access Protocol* (Protocolo de Acceso Inalámbrico) es una tecnología que permite que dispositivos que no utilizan cables, como móviles o PDAs antiguas, como los viejos *Palm Pilot*, accedan a datos que estén en servidores de e-mail, a listas de contactos en servidores corporativos y a sitios Web de Internet. Para que cualquiera de esas modalidades de servicio sea compatible con WAP es necesario crear un servidor WAP y editar las páginas web utilizando el lenguaje WML (Wireless Markup Language), que es una especie de variante del HTML clásico.

Debido a la lentitud de las conexiones desde los teléfonos móviles a finales de los 90, y a la escasez de servicios, el WAP acabó desapareciendo, pero abrió camino al GPRS y a otras formas de acceso inalámbrico aún más sofisticadas.

Acompañando a una lluvia de servicios de e-mail para WAP (la mayoría en realidad aún existen hoy, y por increíble que parezca, hay mucha gente que todavía usa WAP para trabajar con su correo electrónico), Google también quiso lanzar su buscador para móviles, permitiendo que más de 4 millones de resultados pudiesen ser consultados usando los teclados numéricos. Con el "Google de teléfono móvil" se puede hacer todo lo que el Google hace en un ordenador, y tal vez alguna cosa más.

Acceder a Google por teléfono móvil es muy fácil:

1 – Accede al navegador WAP de tu teléfono y conéctate a Internet.

2 – En el campo **Go To URL/Website** o **Ir** (si tu navegador está en español), escribe www.466453.com (el nombre de Google en tu teclado alfanumérico).

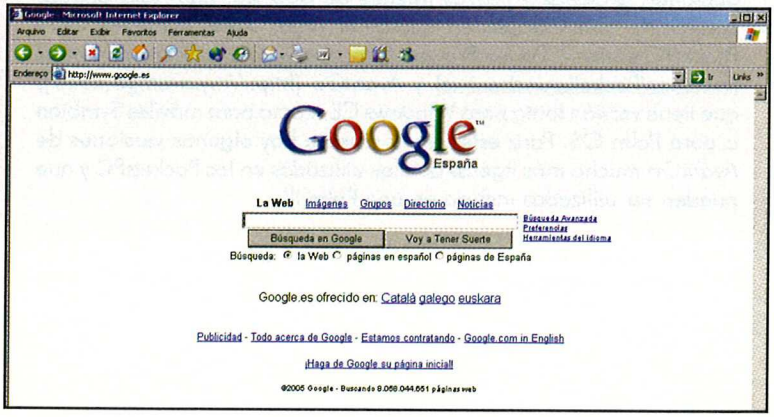
3 – La página de Google que va a aparecer se llama *Google Number Search*. Por último, tienes que utilizar el teclado numérico del móvil para "escribir" letra a letra el término que quieras buscar. También puedes incluir espacios y paréntesis.

Además del WAP, hay otro servicio que permite el acceso a Google vía móvil y que puede ser utilizado tanto en móviles de 2,5G (los famosos GPRS) como en otros modelos más antiguos, pero compatibles con el acceso a la Web. La manera de acceder al servicio es ligeramente diferente:

1 – Realiza la conexión con tu proveedor (para acceder a este tipo de servicio tienes que tenerlo habilitado). Después accede al menú del teléfono móvil.

2 – Selecciona la opción **Acceso a la Web** o equivalente en las opciones del menú. Pulsa el botón **Aceptar**.

3 – Inserta la palabra *Google* cuando se pida la introducción de una URL. Si el navegador no acepta *Google* como dirección, escribe www.google.com/wml. Haz clic en **Aceptar**.



¿Y las PDAs?

Las PDAs (o palmtops) son los sustitutos modernos tanto de los animales de compañía como del genio de la lámpara mágica de Aladino. Incluso quienes creían que era cosa de pijos eso de tener una Palm o un PocketPC y escribir con el inefable lápiz óptico, hoy se ha rendido a los encantos de estos bichitos, que sirven para jugar, para recordar el nombre de personas y recados y, por supuesto, para darse un garbeo por Internet. Para realizar esta última hazaña tienes que tener un módem o una conexión WiFi (o un *smartphone*, un híbrido entre móvil y PDA, todo en uno), una cuenta de correo (puede ser incluso gratuito, como el Yahoo) y por último, un navegador.

El paso siguiente es facilísimo: conectado a internet, escribe la dirección www.google.es/palm. Abrir Google en una Palm III con 8 MB de RAM es casi instantáneo.

Pequeños pero brillantes

Con excepción de los pockets A-20 y E-100 de Casio y de todos los modelos de Palm hasta la Palm IIIe, la mayor parte de las PDAs vienen ya con un navegador para Internet. En el apartado de aplicaciones que ya incorpora el sistema operativo tenemos buenas herramientas, como el tradicional *Pocket Internet Explorer*, de Microsoft, muy parecido a la versión 4.0 de este navegador, que venía con los discos de instalación del Windows 95. Para los usuarios de Palm (principalmente Palm Zire y algunas Tungsten) el acceso es tarea del navegador WebPro, que permite, incluso, la configuración de un proxy.

Si tu PDA o tu *smartphone* no vienen con un navegador, hay diversas opciones, gratuitas y ligeras (menos de 300 KB) para que puedas conectarte con Google en cualquier lugar. Para Windows CE y PocketPC existen el Pocket Browser (http://www.conduits.com/products/launcher/PocketLauncher.exe) y AvantGo (http://my.avantgo.com/), que tiene versión tanto para Windows CE, como para móviles Symbian o para Palm OS. Para este último sistema hay algunas versiones de AvantGo mucho más ligeras que las utilizadas en los PocketsPC y que pueden ser utilizadas incluso en una Palm III.

CALENDARIOS

Los ordenadores trabajan siempre con números. Incluso si las investigaciones sobre Inteligencia Artificial, tanto cognitiva como intuitiva, llegan algún día a buen puerto, hay algo que nunca va a cambiar: el que las máquinas digitales se alimenten de números, mientras que las máquinas analógicas seguirán funcionando a base de de girar manivelas o engranajes.

Máquinas inteligentes

Las investigaciones con IA (Inteligencia Artificial) empezaron, como la mayoría de las investigaciones informáticas a gran escala, a finales de la década de los años 40, justo después de haber terminado la 2ª Guerra Mundial. Esos esfuerzos iniciales intentaban, sobre todo, medir la reacción de los ordenadores y de los procesadores de datos en determinadas situaciones que se les planteaban, como la división de números extensos, o el fraccionamiento de bases de datos escritas en números binarios (formados por secuencias de ceros y unos, los únicos valores que un ordenador entiende).

Pero las investigaciones de Inteligencia Artificial enfocadas hacia aspectos cognitivos (como hacer que las máquinas aprendan cosas nuevas a partir de nuevos elementos o de la corrección de errores) o intuitivos (hacer que la máquinas se acerquen lo más posible al razonamiento humano, "intuyendo" las partes que faltan de un problema o resolviendo un error de interpretación), sólo fueron posibles mucho tiempo después, con la creación de la Lógica Difusa (o Lógica Fuzzy), que para algunos autores es el punto de inicio de la Inteligencia Artificial, por cuanto trata también sobre la interpretación de los errores y sobre la resolución de problemas incompletos.

Para conocer más cosas sobre métodos y proyectos de Inteligencia Artificial, ve a http://aepia.dsic.upv.es/.

Las personas manipulan la información de una manera muy parecida a como lo hacen las máquinas, aunque reconocen más valores que los simples números 0 y 1 (bueno, en realidad, ni siquiera todas tienen esta mínima capacidad). Así, para situarse en el espacio en el que viven, las personas usan números para identificar sus casas y sus pisos, para mostrar a qué distancia está su casa o su ciudad de un determinado lugar, etc...

Para situarse en el tiempo, a su vez, utilizamos números que representan el cómputo de días, horas y minutos (elementos propiamente matemáticos). Por poner un ejemplo, en el cómputo de horas se utiliza el reloj, dividido en 12 periodos exactamente iguales.

Para contar periodos de tiempo más largos (años, siglos) se inventó el calendario, una representación lógica y gráfica, las dos a un tiempo, que sirve para "almacenar" los cálculos astronómicos que eran necesarios para medir el principio y el final de un año o de la estación del año (el cálculo no era muy sencillo, y sigue sin serlo). Los mayas, por ejemplo, "imprimían" sus calendarios (un año de 365 días, con ciclos de 13 años para añadir 1 día complementario, agrupados, a su vez, en ciclos mayores, de 52 años) en hojas hechas de piel humana, o los tallaban en grandes discos de piedra.

El calendario juliano y Google

Hoy en día todo el mundo, al menos el mundo comercial (exceptuando a la Iglesia Ortodoxa, en Rusia y a los países musulmanes, por motivos religiosos) utiliza el calendario gregoriano, que recibe este nombre por haber sido promulgado por el papa Gregorio XIII, en 1682. El calendario gregoriano suprimió diez días del calendario anterior (el calendario juliano, que había sido inventado por Julio César y que acumulaba ese retraso después de 1.600 años de actividad). Además de eso, creó los años bisiestos, y dictaminó que sólo lo serán cuando las dos primeras cifras son divisibles por cuatro. De acuerdo con esta norma, los años 1600 y 2000 son bisiestos, y sin embargo, los años 1700, 1800 y 1900 no tienen ese día adicional.

El calendario juliano tiene una duración exacta de 365,25 días, un redondeo de la duración exacta del año solar o equinoccial (que es de 365.2422 días). A pesar de ser más fácil de calcular (esa era la intención de Julio Cesar al codificarlo), este redondeo hace que el año pierda un día cada 128 años. Por eso, desde el inicio de su vigencia hasta su sustitución por el calendario gregoriano, el modelo juliano había acumulado un adelanto de diez días en el cómputo. Ese desfase afecta también a las máquinas (a todo sistema digital que trabaje con números), que ejecutan un programa de búsqueda (escrito en un lenguaje de programación, que no es más que una herramienta para transformar los conceptos humanos a código máquina –ceros y unos-), y que a su vez, indexa páginas localizadas en la Web: cuando un buscador rastree información sobre fechas, ¿qué calendario debe usar?

Vamos por pasos. Como vimos en el **Capítulo 2**, Google es capaz de buscar una información en Internet de muy diferentes formas. Entre estas muchas formas tenemos la posibilidad de utilizar la **Fecha** en **Búsquedas Avanzadas,** para encontrar datos de los últimos tres meses, seis meses, o un año.

Así y todo, esta opción no permite búsquedas para valores mayores a un año, ni inferiores a tres meses, ni para periodos de tiempo con valores intermedios a los propuestos (por ejemplo, cuatro meses). Para hacer este tipo de búsqueda Google ha dispuesto otro sistema a través del campo de búsqueda. Mediante el comando **daterange** se pueden hacer búsquedas cronológicas específicas, incluyendo el día de publicación de la página en la Web. La sintaxis del comando es:

```
<Término a buscar> daterange=<fecha inicial> -
<fecha final>
```

Y aquí está el secreto de la cosa. Buena parte de Google está escrita en Python, un lenguaje de programación libre y extremamente flexible. No se sabe bien por qué (tal vez por el hecho de que los números enteros son más fáciles de catalogar y procesar), el personal de Google adoptó el calendario juliano como modelo. La función de búsqueda por fechas de este buscador (el código fuente del Google no está disponible para poder comprobarlo) debe ser algo muy similar al que aparece en el siguiente recuadro.

Pensar como Google

Mientras el código fuente de Google no sea liberado (hay noticias que ya se aventuran a afirmarlo, y tenemos la esperanza de que ocurra dentro de poco), podemos imitar por lo menos una de las funciones de Google: la conversión de fechas.

El código que exponemos abajo ha sido escrito por Scott Moore y permite la convertir cualquier fecha del calendario gregoriano al juliano.

```
baseConstant = 2440588 (donde 2440588 ha sido
utilizado como base de cálculo)
if theDate == 'now': (year, month, date, _, _, _, _, _, _)
=time.localtime() (aquí se define la entrada de datos
en el formato de fecha del calendario gregoriano)
theTarget=time.mktime((year,month,date,19,
0,0,0,0,-1)) (aquí se define la salida de datos en
formato juliano)
theTargetInDays = int(theTarget/(60*60*24))-1 (aquí
se define el formato de salida de la base de datos)
return theTargetInDays + baseConstant
```

Para ejecutar este programa sólo hay que bajar cualquier editor de código que acepte este lenguaje de programación (puedes bajar el paquete oficial del Python para Windows en http://www.python.org/download/) o utilizar un editor como Ultraedit (ftp://ultraedit.com/uedit32.zip).

Cómo hacer las búsquedas

¿Cómo hacer entonces una búsqueda por fechas absolutas en Google? Si no eres un astrónomo, un monje o sencillamente, un pedante más de turno, lo más probable es que no sepas hacer la conversión de fechas de memoria (lo que en todo caso te confirma como una persona sana y normal...).

Por suerte para nosotros, hay varias herramientas en la propia Internet que resuelven este problema. Una de ellas está contenida en el sitio http://www.measurementsconverter.com/calendars.html, una dirección en la que se puede realizar la conversión de cualquier fecha gregoriana al formato "antiguo":

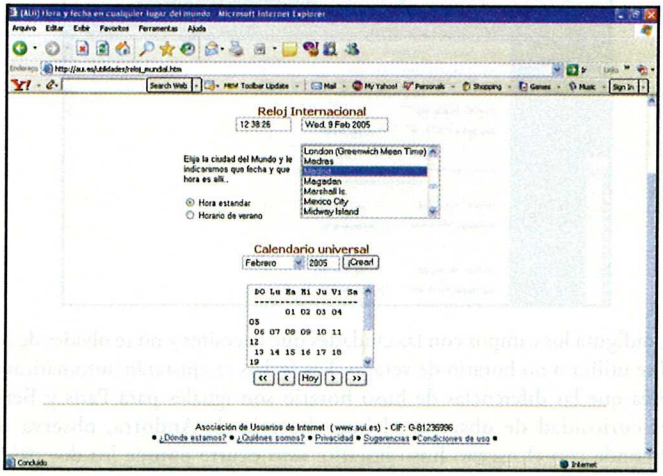
1 – Después de acceder a la página, escribe, junto al recuadro **Month Number**, el número del mes que quieres consultar (1 corresponde a enero, 12 a diciembre). Y en el cuadro **Year**, escribe el año que deseas consultar, en el formato de cuatro dígitos (por ejemplo, 1999, 2004, etc.). Después, haz clic en el botón **Calculate**.

2 – Tendrás una tabla con todos los días de ese mes, incluyendo en qué día de la semana cayeron. En las columnas *JD*, *MJD* y *TJD* tendrás los datos relacionados con el calendario juliano. El formato utilizado por Google para realizar las búsquedas es el de la columna *JD*. Así, el 1 de diciembre de 2003 corresponde al número absoluto 52974.

3 – Abre la página de Google y escribe, en la barra de búsquedas, el término y la búsqueda de fecha, y listo. Así:

FECHAS Y HUSOS HORARIOS EN INTERNET

Tal y como hemos visto en el capítulo anterior, Google busca mediante fechas exactas de artículos y de páginas de Internet, y utiliza para ello como referencia el calendario juliano. Pues bien, para calcular cualquier fecha también podemos usar algunas herramientas de Internet. En la página web http://www.aui.es/utilidades/reloj_mundial.htm, por ejemplo, encontrarás un calendario universal que te permitirá saber en que día de la semana cae una determinada fecha.



Solamente tienes que introducir el mes y el año en el apartado **Calendario universal** y al pulsar en **¡Crear!** aparecerá el mes que has solicitado.

En la parte superior de la página verás también un **Reloj Internacional**, con el que podrás saber cuál es la hora en cualquier parte del mundo, incluso con el horario de verano, simplemente buscando la ciudad de la que quieras saber qué hora es en ese momento.

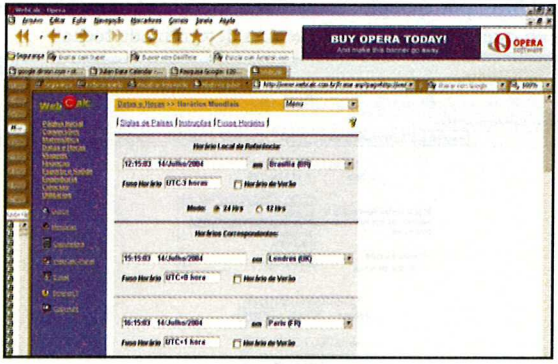
Operaciones con fechas

Supón que trabajas en una empresa de importaciones y necesitas saber que hora es en Washington y en El Cairo, para enviar unos faxes a oficinas en esas ciudades, pero que no tienes ni la más remota idea de cómo saber qué hora es allí.

Aunque no se pueden hacer cálculos con fechas en Google, sí que hay muchas páginas en Internet que te dan este servicio gratuitamente, incluso permitiendo que te imprimas un calendario personalizado. Una de las más completas es WebCalc (http://www.webcalc.com.br/datas/horarios.html), y aunque que está en portugués, es tan sencilla de usar que no encontrarás ningún problema.

Accede a WebCalc, en **Horario Local de Referência**, introduce "Madri" (Madrid) como ciudad de referencia, a no ser que estés en Canarias. Si estamos en horario de verano, pincha en el recuadro Horario de Verano. Mantén el modo *24 Hrs* seleccionado, ya que es el modelo utilizado oficialmente en las transacciones comerciales.

En los campos que hay debajo podrás ir averiguando la hora local que corresponde a cualquier país del mundo (incluyendo los minúsculos Andorra y Samoa Occidental, en Oceanía). Puedes configurar las horas hasta para cinco ciudades simultáneas.



Configura los campos con las ciudades que necesites y no te olvides de indicar si allí se utiliza o no horario de verano. Las fechas se ajustarán automáticamente. Observa que las diferencias de huso horario son iguales para París y Berlín. Si tienes curiosidad de observar el huso horario de Andorra, observa que se corresponde con el mismo huso horario: esto ocurre porque los dos países son vecinos muy próximos.

¿Qué hay que saber de Andorra?

Andorra es un principado (una especie de monarquía parlamentaria sin rey) situado en el interior de los Pirineos Orientales. Está limitado al norte por Francia y al sur por España. El país está gobernado por una cámara parlamentaria formada por personas de Andorra y su jefatura de Estado corresponde teóricamente, y de forma alternativa, al presidente de Francia y al obispo de Urgell en Lérida.

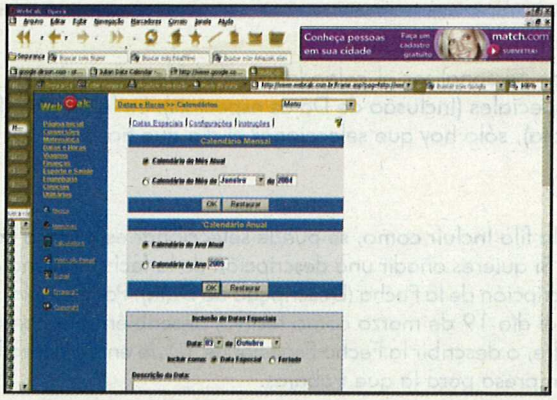
Andorra tiene 64.000 habitantes, repartidos por una extensión territorial de 468 Km^2. A pesar de esta densidad y de su reducido tamaño, tal vez sea muy interesante ir a Andorra. Desde 1950, este país ha llegado a ser considerado por dos veces la nación con más crecimiento económico del mundo. En Andorra hay más de 5.000 tiendas y 500 hoteles que garantizan a sus habitantes una renta "per cápita" superior a la alemana y a la japonesa.

Si quieres añadir esta página a tus favoritos, antes que nada, pincha en *Salvar padrão* (Guardar el modelo), en la parte de abajo de la propia página. Después pincha en **Favoritos > Agregar a Favoritos** (si utilizas el Internet Explorer), **Marcadores > Añadir página aquí** (si usas Opera) o **Bookmarks > Bookmark This Page** (si usas Firefox).

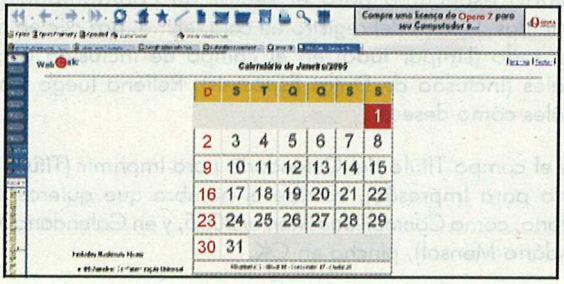
Crear un calendario personalizado

Esta utilidad también te permite crear calendarios personalizados. Esto es una buena herramienta, primero y sobre todo, porque es gratis. Y segundo, porque imprimir los calendarios del Outlook o del Word no es una tarea fácil, y no siempre da buenos resultados.

Para empezar, ve a http://www.webcalc.com.br/frame.asp?pag=http://www.webcalc.com.br/datas/horarios.html y abre el menú desplegable del campo **Menú,** haciendo un clic, y pinchando acto seguido en **Calendários.**



Se pueden crear calendarios mensuales y anuales. Supongamos, por ejemplo, que queramos un calendario del mes de marzo de 2005. En **Calendário Mensal** (Calendario Mensual), pincha en **Calendário do Mes de** (selecciona el mes) **de** (escoge aquí el año). Luego, haz clic en **OK.** Ahora sólo hay que hacer clic en **Imprimir.**



Lo mismo se puede hacer con un Calendario Anual. Vamos a pedir, por ejemplo, que webcalc nos cree un calendario del año 3128, algo que en principio parece complicado. Para ello, haremos clic en Calendario anual (**Calendário do ano**), y escribiremos *3128*. Observa el resultado:

Incluir fechas especiales

Podemos personalizar el calendario en el área de **Inclusión de Fechas especiales (Inclusão de Datas especiais)**. En la primera fila, Fecha (**Data**), sólo hay que seleccionar el día que has elegido.

1 – En la fila **Incluir como**, se puede seleccionar esta fecha como un festivo. Si quieres añadir una descripción de la fecha, rellena el campo Descripción de la Fecha (**Descripção da Data**). Por ejemplo, puedes añadir el día 19 de marzo como festivo, describiéndolo como el día del padre, o describir la Fecha Especial de 10 de enero como la fiesta de la empresa para la que trabajas.

2 – Haz clic en Incluir Fecha (**Incluir data**) para añadir la fecha que has creado en un calendario mensual o anual.

3 – Se pueden crear calendarios sectorizados. Si necesitas un calendario específico para el sector de contabilidad (pagos, vencimientos, lectura del registro de caja, etc...), haz clic en el botón Limpiar todo (**Limpar tudo**) en el campo de Inclusión de Fechas Especiales (**Inclusão de Datas Especiais**). Rellena luego las fechas especiales como desees.

4 – En el campo Título del Calendario para Imprimir (**Título do Calendário para Impresão**), escribe el nombre que quieras darle al calendario, como *Contabilidad Marzo/2005*, y en Calendario Mensual (**Calendário Mensal**), pincha en OK.

DOCUMENTOS DINÁMICOS CON GOOGLE

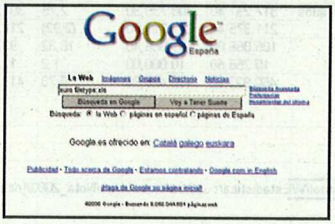
Copiar y pegar, Ctrl + C y Ctrl + V, la opción **Pegar Elemento desde el Portapapeles...** Después de Ctrl + Mayúsculas + Suprimir (el conjunto de teclas que necesitas usar tantas veces, cuando Windows se cuelga), probablemente no exista ningún otro comando más utilizado que el que te permite mover los datos entre carpetas, directorios e incluso entre archivos o documentos muy distintos.

Está claro que estas herramientas acaban siendo especialmente útiles cuando se cortan o copian datos estáticos, que ya están del todo organizados, y que casi no necesitan ser retocados o modificados. Por eso precisamente, copiar datos de Internet puede resultar complicado: los problemas con el formato de la página, su veracidad, o la posibilidad de firmar como propios los datos hacen que muchas veces el trabajo no valga la pena.

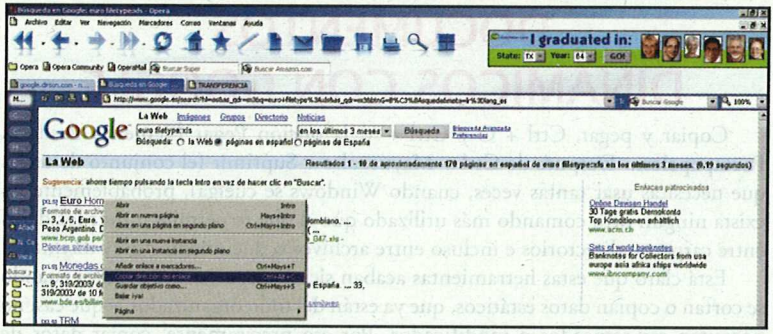
Google: S.O.S.

Google te ofrece, al menos indirectamente, una solución bastante interesante si utilizas Microsoft Office: te permite buscar archivos similares a los que ya estás utilizando, sobre todo documentos de Word y hojas de cálculo de Excel. Esta herramienta de Google te permite, por ejemplo, buscar una hoja de cálculo en la Web que esté relacionada con el valor del Euro, y añadirle los valores que encontremos en ella a una de nuestras propias hojas de cálculo, que estará situada en nuestro propio ordenador.

1 – Escribe en Google, en el campo de búsqueda, la expresión **euro**, añadiéndole el parámetro *filetype:xls* (documentos de Microsoft Excel. Si quieres obtener sólo resultados muy recientes, vete a la página de Búsqueda Avanzada, y en el filtro **Fecha** elige la opción **En los últimos 3 meses**):



2 – Haz clic en una de las hojas de cálculo que hayas encontrado, y échale un vistazo a sus datos, comprobando si son los que necesitas. Si es así, vuelve a la página de Google, pincha con el botón derecho del ratón en el enlace del que quieras obtener los datos y luego en la opción **Copiar Dirección del Enlace.**

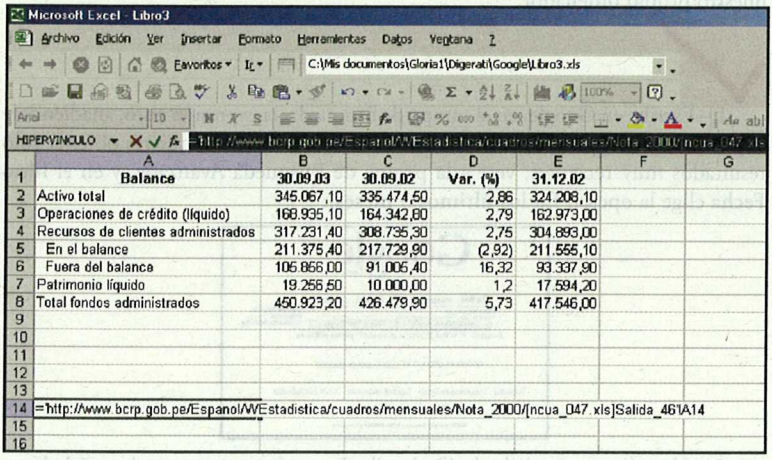


3 – Abre tu hoja de cálculo (la que tienes guardada en tu ordenador) y haz clic en la celda o en las que quieras insertar el contenido de la hoja que está en la Red. Crea una fórmula que contenga los siguientes parámetros:

```
='http://www.alguienes.homepage/[archivo.xls]Hoja1'!A1
```

[archivo.xls] se refiere al nombre que tiene en internet el archivo, es decir, en nuestro caso "[ncua_047.xls], Hoja1'!" se refiere al nombre de la hoja de cálculo en el archivo original, Salida_46'!, es decir, el de la Web.

En este ejemplo estamos trabajando con un archivo o un libro sencillo (sin muchas hojas en su interior). Así que obtendremos el siguiente resultado:



Microsoft Excel - Libro3

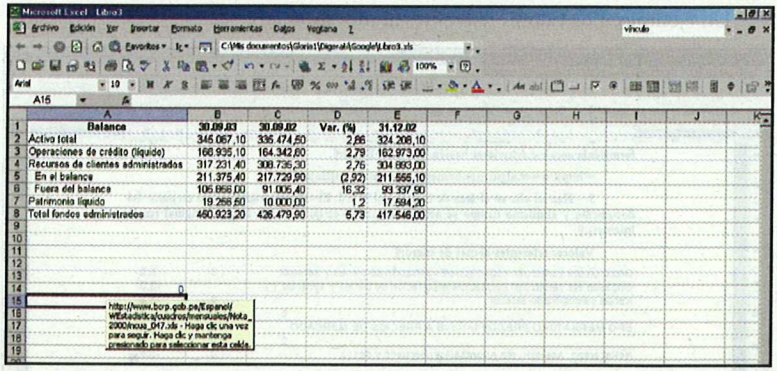
Archivo Edición Ver Insertar Formato Herramientas Datos Ventana ?

C:\Mis documentos\Gloria1\Digerati\Google\Libro3.xls

HIPERVINCULO =http://www.bcrp.gob.pe/Espanol/WEstadistica/cuadros/mensuales/Nota_2000/ncua_047.xls

	A	B	C	D	E	F	G
1	**Balance**	**30.09.03**	**30.09.02**	**Var. (%)**	**31.12.02**		
2	Activo total	345.067,10	335.474,50	2,86	324.208,10		
3	Operaciones de crédito (líquido)	168.935,10	164.342,80	2,79	162.973,00		
4	Recursos de clientes administrados	317.231,40	308.735,30	2,75	304.893,00		
5	En el balance	211.375,40	217.729,90	(2,92)	211.555,10		
6	Fuera del balance	105.856,00	91.005,40	16,32	93.337,90		
7	Patrimonio líquido	19.256,50	10.000,00	1,2	17.594,20		
8	Total fondos administrados	450.923,20	426.479,90	5,73	417.546,00		
9							
10							
11							
12							
13							
14	=http://www.bcrp.gob.pe/Espanol/WEstadistica/cuadros/mensuales/Nota_2000/[ncua_047.xls]Salida_46!A14						
15							
16							

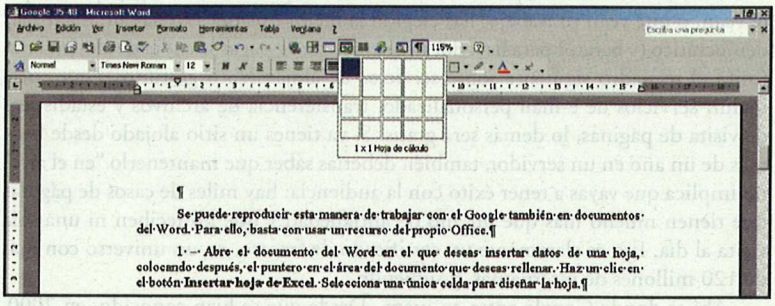
A partir de este momento el contenido de la hoja de cálculo de la Web estará vinulado permanentemente al contenido de la hoja local, y aparecerá como un valor sencillo (con formato de enlace simple) en la hoja que hayas creado:



	A	B	C	D	E
1	Balance	30.09.03	30.09.02	Var. (%)	31.12.02
2	Activo total	346.067,10	336.474,60	2,86	324.208,10
3	Operaciones de crédito (líquido)	166.935,10	164.342,60	2,79	162.973,00
4	Recursos de clientes administrados	317.231,40	308.736,20	2,76	304.893,00
5	En el balance	211.375,40	217.729,90	(2,92)	211.565,10
6	Fuera del balance	105.856,00	91.006,40	16,32	93.337,90
7	Patrimonio líquido	19.266,60	10.000,00	1,2	17.584,20
8	Total fondos administrados	460.923,20	426.479,90	6,73	417.546,00

Puedes hacer lo mismo que hemos hecho aquí con Google en otros documentos de Office, como los de Word. Para ello, sólo tienes que usar una herramienta del propio Office.

1 – Abre el documento de Word en el que quieras insertar los datos de una hoja, colocando después el puntero en la zona del documento que quieras rellenar. Pincha en el botón **Insertar hoja de Excel**. Selecciona una sola celda, para que contenga la hoja.

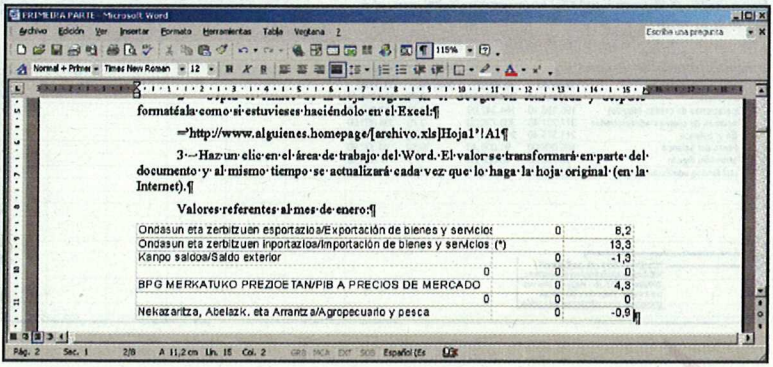


2 – Copia el enlace de la hoja que hayas elegido mediante Google en el interior de esta celda y dale después formato, igual que si estuvieses haciéndolo en Excel:

```
='http://www.alguienes.homepage/[archivo.xls]Hoja1'!A1
```

3 – Pulsa en el área de trabajo de Word. El valor se convertirá en parte del documento, pero además se actualizará automáticamente cada vez que lo haga la hoja de cálculo original (la de Internet).

Un ejemplo. Los valores referentes al mes de enero:



formatéala como si estuvieses haciéndolo en el Excel:¶

=’http://www.alguienes.homepage/[archivo.xls]Hoja1’!A1¶

3.—Haz un clic en el área de trabajo del Word. El valor se transformará en parte del documento y al mismo tiempo se actualizará cada vez que lo haga la hoja original (en la Internet).¶

Valores referentes al mes de enero:¶

Ondasun eta zerbitzuen esportazioa/Exportación de bienes y servicios		0	8,2
Ondasun eta zerbitzuen inportazioa/Importación de bienes y servicios (*)			13,3
Kanpo saldoa/Saldo exterior		0	-1,3
	0	0	0
BPG MERKATUKO PREZIOETAN/PIB A PRECIOS DE MERCADO		0	4,3
	0	0	0
Nekazaritza, Abelazk. eta Arrantza/Agropecuario y pesca		0	-0,9

PONER TU SITIO EN EL PRIMER PUESTO EN GOOGLE

En los capítulos anteriores henos abordado algunas de las formas más básicas de utilizar Google y otros recursos de búsqueda de Internet, para hacer tu vida profesional un poco más fácil. Este capítulo que comienza es un poco diferente: te vamos a mostrar una manera de usar Google como un instrumento de marketing personal.

Si tienes un sitio web, personal o comercial, colgado de Internet, y quieres que te vean, o que vean tu trabajo, has elegido la mejor opción. No hay un lugar más democrático (y barato) para hacerlo, ya que, aparte de pagar la cuota de alojamiento (mira el recuadro siguiente sobre servicios de alojamiento gratuitos), que suele incluir servicios de e-mail personalizado, transferencia de archivos y estadísticas de visita de páginas, lo demás será gratis. Si ya tienes un sitio alojado desde hace más de un año en un servidor, también deberías saber que mantenerlo "en el aire" no implica que vayas a tener éxito con la audiencia: hay miles de casos de páginas que tienen mucho más que un año de antigüedad y que no reciben ni una sola visita al día. Eso es claramente un certificado de fracaso, en un universo con más de 120 millones de internautas, y subiendo...

Ahí es donde Google entra en juego. Desde que se hizo conocido, en 2000, aparecer indexado en Google se ha convertido en uno de los mejores métodos de propaganda para una página web (no sólo de empresas, sino también personales: hay gente a la que le gusta mostrarse, sobre todo después de que se inventaron los blogs). Estar en la primera página de búsqueda cuando Google muestra el resultado (la mayoría de las personas, según algunas estadísticas, sólo tienen paciencia para buscar hasta la tercera página...) es casi como ganar el Premio Nóbel.

Proveedores gratuitos

Puede que tengas unas ideas magníficas, pero que estés sin un céntimo en el bolsillo. En este caso, además de no querer gastar anunciándote en Google, probablemente no querrás tampoco gastar tu dinero en un servicio de alojamiento (hosting).

Hay muchas alternativas de proveedores gratuitos en la Web, que ofrecen espacios bastante razonables en sus discos duros (una media de 50 MB), y que permiten incluso utilizar recursos en Flash y ASP. Estos proveedores serán la mejor solución ante la falta de recursos:

Wanadoo
http://www.personales.wanadoo.es
Alojamiento gratuito de sitios con un espacio de hasta 50 MB. Tiene un asistente para confeccionar páginas y un gestor de ficheros.

Terra
http://www.terra.es
Hosting gratuito para usuarios de Terra, con un espacio de 5 a 30 MB, dependiendo del plan que hayas contratado.

Lycos Tripod
http://www.tripod.lycos.es
Te ofrece un asistente para la construcción de sitios, además de 50 MB de espacio.

Ya.Com
http://espacio.ya.com/
Alojamiento gratuito de sitios de hasta 30 MB, con asistente de creación de páginas, administrador de archivos y acceso ftp.

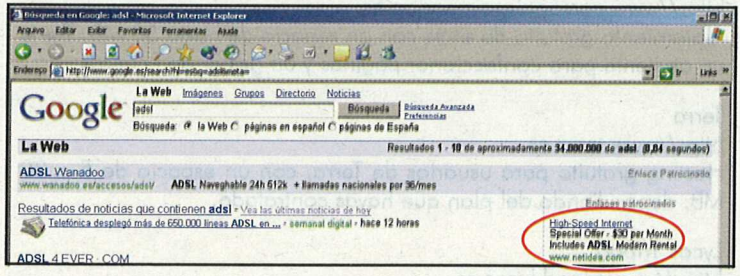
Yahoo! GeoCities
http://es.geocities.yahoo.com
Un clásico. Ofrece espacio de 15 MB para el hosting gratuito de sitios, un asistente para la construcción de webs, un editor de HTML, la transferencia de archivos mediante FTP y un administrador web de archivos. Otra opción interesante es que los emails de Yahoo! pueden usarse para crear el sistema de información de la página.

Free Servers
http://www.freservers.com
Servicio internacional de alojamiento gratuito. Ofrece 12 MB de espacio en disco, una tasa de transferencia de 500 MB para que puedan ver tu página muchos, además de un administrador de archivos personalizable.

Cómo convertirte en un campeón

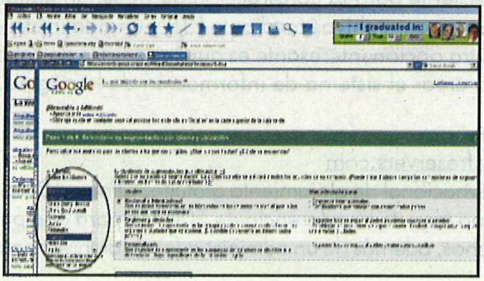
Lo cierto es que hay dos maneras distintas de convertirse en un líder de Google en tu categoría: una de ellas es convencional, rápida, pero duele en el bolsillo, y la otra utiliza herramientas que casi todos conocemos, pero que nos da pereza utilizar: relaciones personales y matemáticas.

La primera de ellas te la proporciona el propio Google. Hace cerca de dos años el buscador puso a disposición de los internautas un servicio de anuncios, mediante enlaces, en sus páginas. Pero para evitar la polémica, evitó subastar los primeros lugares de la búsqueda para los sitios de pago, como solían hacer algunos buscadores: los anuncios de Google están situados en una zona destacada e independiente de la página (el ya mencionado diseño "limpio" de Google lo permite), y que no se mezcla con los resultados tradicionales de búsqueda.

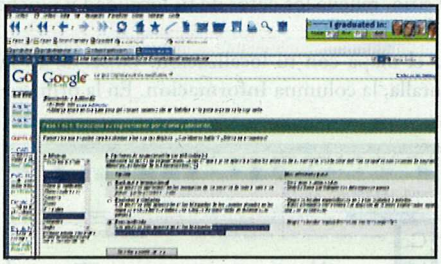


1 – Este servicio se llama AdWords y puede contratarse en la página de inicio de Google. Pincha en **Todo Acerca de Google** y luego en el enlace de **AdWords**.

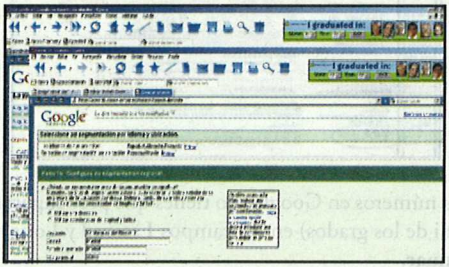
2 – En la página Google AdWords, pulsa en el botón **Regístrese Ahora Empezar**. En la primera fase tendrás que crear tu anuncio. En la columna *a.* se puede elegir con qué idiomas quieres desea iniciar el servicio. Si tu negocio es local (solamente quieres dar servicio a una ciudad, una región o una comunidad autónoma), selecciona sólo la opción **Español**. Si quieres seleccionar más idiomas para tu anuncio, mantén la tecla Ctrl apretada mientras vas pinchando en otras opciones.



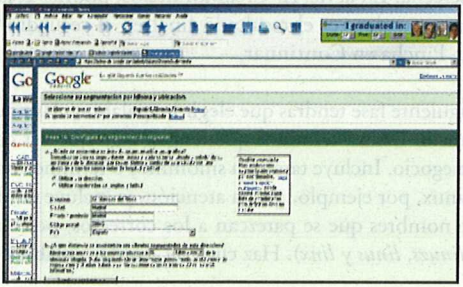
3 – En la columna *b.*, selecciona el plan de marketing que quieres contratar: verás desde planes globales, que cubren el mundo entero (está claro que en un caso como ese, deberías haber elegido inglés o español como idioma estándar), hasta algunos planes personalizados, que llegan sólo a algunos usuarios que estén situados a una cierta distancia de tu web. Después de elegir tu opción, haz clic en **Guardar y continuar.**



4 – En nuestro caso, elegimos la opción **Personalizado.** Con ella podremos definir la situación de nuestro negocio, pinchando en **Utilizar una dirección.** La página se abrirá y se mostrarán varios campos que habrá que rellenar.



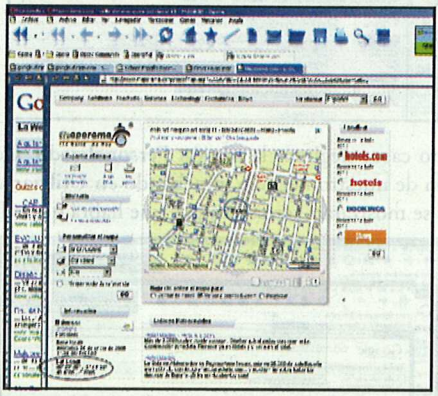
5 – En el apartado **b. ¿A qué distancia de esta dirección están localizados tus clientes segmentados?**, debes escribir la distancia en la cuál quieres captar tu audiencia usando Google. Elegiremos la distancia de 50 Km,, que es un poco mayor que el perímetro de la mayoría de los municipios. Haz clic en **Guardar y continuar.**



6 – En la siguiente pantalla Google pedirá tu localización exacta, tanto en latitud como en longitud (una información que no se obtiene fácilmente). Para ayudarnos, vamos a utilizar una excelente página de itinerarios mundiales, Maporama (http://www.maporama.com/share/).

7 – Accede a este sitio sin cerrar tu formulario de Google AdWords. En el campo **Mapas** del sitio, rellena los datos de tu localidad y pincha en el botón **Go**.

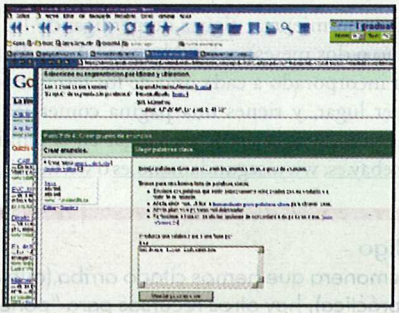
8 – Aparecerá el mapa con tu localización. Observa, en la esquina inferior izquierda de la pantalla, la columna **Información**. En la primera línea encontrarás tu posición exacta.



9 – Copia los números en Google (no tienes que preocuparte de las señales de las coordenadas ni de los grados) en los campos **Latitud** y **Longitud**, y haz clic en **Guardar y continuar**.

10 – Ahora tienes que incluir tu anuncio en un determinado grupo. Esto es delicado. Si eliges, por ejemplo, consultor de sistemas, todos los usuarios que busquen sistemas o consultor verán tu anuncio nombrado en la columna de anuncios de la derecha. En la fila **A. Crear anuncios**, podrías crear tu anuncio, al mismo tiempo que puedes ver el resultado final en el recuadro blanco que hay encima del texto. Pincha en **Continuar**.

11 – En la siguiente fase tendrás que elegir las palabras clave que quieres relacionar con tu anuncio. Utiliza términos que tengan algo que ver con el tema de tu página o con tu negocio. Incluye también sinónimos o términos relacionados (para una página de Linux, por ejemplo, presta atención, ¡e incluye también Windows y XP!), además de nombres que se parezcan a los correctos, pero mal escritos (por ejemplo, añade *linuzs*, *linus* y *linx*). Haz clic en **Guardar palabras clave**.



12 – Ya estás casi terminando... Ahora deberás pujar por los servicios de Google. Haz clic en la moneda con la que quieras pagar por los servicios (en nuestro caso, **Euros**). Rellena después el valor que quieras ofrecer a Google como coste máximo por clic. El valor mínimo es de 0.05 € (cinco céntimos), además de los 5.00 € que vas a pagar por meter tu web en la lista. Cuanto mayor sea tu oferta, más veces aparecerás en Google. Haz clic en **Guardar y continuar...**

13 – Ahora tienes que especificar tu presupuesto diario. Si escribes 1,00 € (1 euro), Google mostrará tu anuncio unas veinte veces por día. En un mes, por lo tanto, gastarás 30,00 € por aparecer, como máximo, 600 veces en el buscador. Pincha en **Guardar y continuar...**

Gratis

Si crees que esta transacción no es demasiado ventajosa, o que tu idea tampoco vale tanto como para que unos cuantos billetes desaparezcan así de tu bolsillo, hay otras maneras de hacerte visible para el mundo mediante Google. Para ponerlas en práctica necesitarás tener algunos amigos, además de un vocabulario un poco superior a la media.

1 – Elige un término o una expresión para asociarla a tu sitio. Puedes utilizarlo tanto en el nombre de la página como en su dirección (lo ideal es utilizar ambos). Elige palabras que no sean muy comunes, para no tener competidores. Si tu página trata sobre sistemas operativos, escoge algo como NetBSD, "usuario falso" o "DVI Editor".

2 – Ahora, activa tus resortes. Pídele a tus amigos que pongan en sus páginas las palabras que has elegido, junto con el enlace de tu página. Por ejemplo, asocia tu página al enlace www.linux.usuariofalso.es.

3 – Coloca noticias sobre tu página también en listas de discusión y en foros, sin olvidarte de incluir el enlace de tu página, además de las palabras clave.

4 – Espera un tiempo mientras sigues desperdigando los datos de tu página por la Web (entre uno o dos meses), y haz algunas búsquedas en Google para ver si tu página ya se ha incorporado a cada uno de los términos elegidos. Si todavía no aparece en primer lugar, y tienes una página comercial, intenta colocar tu dirección en páginas que ya están en primer lugar en el ranking, o sea, que sean famosas, como www.ebay.es, www.segundamano.es o www.elmundo.es/clasificados.

Crimen y castigo

Además de esta manera que hemos citado arriba (que es la que mejor funciona en la práctica), hay otros recursos para "poner la zancadilla" a Google. Aquí tienes algunos de ellos:

Texto oculto

Se parece más a las artimañas de los piratas y los espías que a un truco de informática. Esta técnica consiste en llenar tu página de palabras que tengan el mismo color que el fondo de la página, de manera que varios bloques de contenido se hagan invisibles al ojo humano. Letras minúsculas, escondidas en cuadros de colores chillones que parecen sencillos errores de diagramación, o listas de palabras pasando a gran velocidad.

Cross-linking

Supongamos que tienes diversas páginas, como www.clubdelapera.com.es y www.disfracesdecolores.com.es, entre otras direcciones edificantes y esenciales para el desarrollo de la cultura humana. Basta con que coloques, en tus páginas, enlaces hacia las demás, y parte del trabajo estará hecho. Es lo que se llama cross-linking. Si ya has registrado una misma página con varios enlaces apuntando hacia ella en ESNIC.ES, pues mejor todavía: de un disparo alcanzarás varias veces la misma dirección.

Guestbook spamming

Esta técnica consiste en firmar todos los libros de visitas que puedas encontrar, ya que muchos de estos libros permiten añadir direcciones Web en sus comentarios. Ésta es una manera rápida de conseguir enlaces sin pedírselos a nadie, ni siquiera al dueño de la página (el webmaster) en la que está puesto el libro.

Granja de palabras

En este recurso el dueño del sitio crea una página, más o menos oculta, en la que añade toda una panoplia de palabras.
Al crear un sitio sobre Linux, por ejemplo, se creará una página HTML con todos los comandos que se utilizan en la línea de comando de

este sistema operativo. Así, muchas búsqueda sobre Linux y sus respectivos comandos acabarán en tu página.

¿Es legal?
Después del crimen siempre es hora del castigo: no, la respuesta es que no es legal. Al percibir que tu sitio llega a posiciones destacadas en Google utilizando una de estas tácticas, un usuario descontento o un competidor puede denunciarte en el siguiente enlace: http://www.google.com/contact/spamreport.html. En Google suelen examinar las denuncias de actividades piratas, y suelen cargarse a los "malos deportistas", sin más, de sus búsquedas.

LA VUELTA AL MUNDO EN 80 DÍAS, CON GOOGLE

Cuando Julio Verne concibió a su personaje, Phineas Fogg, el protagonista de *La vuelta al mundo en 80 días* (1864), no hacía ni 30 años que los barcos de vapor habían empezado a cruzar los mares en una carrera loca, llena de accidentes y de pasos en falso. Los aviones no serían inventados hasta medio siglo después, y su utilización comercial sólo seria posible después de 1920. ¿Cómo hacer, entonces, una vuelta al globo terrestre entero, en esos exactos 80 días?.

Si has leído el libro sabrás que Fogg, junto con su criado francés Passepartout, utiliza barcos, canoas, ferrocarriles e incluso un elefante que compra de carambola, para realizar su travesía de Europa a África, desde allí hasta Asia y luego a América, para llegar por fin al destino un día antes de lo previsto. El lector recordará también que en el viaje ocurren varios imprevistos, a pesar de que Fogg dispone de toda su fortuna personal para realizar su hazaña: falta de alojamiento, intentos de robo, retrasos en los transportes y el error de un policía, que lo confunde con un ladrón de bancos fugado, y que hace lo imposible para intentar terminar con su viaje y atraparlo.

Un viaje con Google

Si en la época en la que ocurrió esta aventura Phineas Fogg hubiese tenido un palmtop y una conexión a Internet, probablemente haría su viaje en mucho menos tiempo y tendría muchos menos sinsabores. Incluso sin aviones, sin trenes de alta velocidad o sin coches, bastaría una PDA y Google para que este viaje se hiciese mucho más sencillo. Está claro que no pasaríamos ni por la mitad de las hazañas y situaciones electrizantes del libro, pero eso es otra historia.

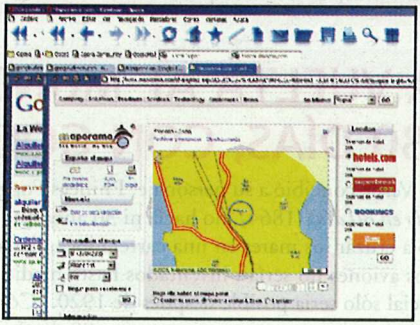
Falta de alojamiento

Incluso con utilizando teléfonos, hacer reservas en un hotel solía ser una situación bastante traumática. Algunos recepcionistas o gerentes de hotel parecen negarse a

alojar a algunas personas que eligen su hotel y que tienen dinero para pagar... Así que imagínate ahora que tuvieses sólo unas horas para descansar entre una etapa y otra de tu viaje, y descubres que esta reserva sencillamente no se ha hecho, y que estás en una ciudad que no conoces, y no tienes ningún sitio donde quedarte.

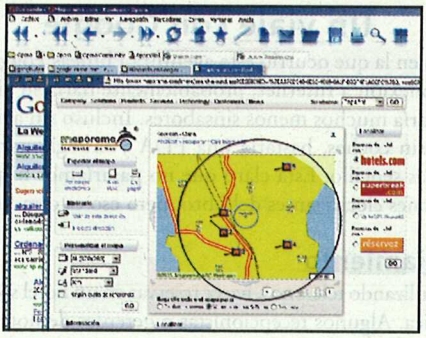
Con nuestra PDA, y con Google, solucionarlo sería sencillo. Buscando los términos *hoteles* y *mapas*, encontraremos el sitio www.maporama.com, del que ya hemos hablado antes. Esta página está especializada en conseguirte mapas, direcciones, nombre de calles, reservas de hotel y aeropuertos, no sólo en España, sino en todo el mundo.

En la zona de **Mapas** del sitio, elegiremos nuestro país de destino. Puede ser por ejemplo China, y el nombre de la ciudad podría ser *Hong Kong*.



Veremos de inmediato un mapa de la ciudad. Si conociéramos alguna cosa de Hong Kong, un nombre de una calle, de un rincón, un código postal, ya tendríamos un mapa específico del barrio en el que estamos, incluyendo hoteles, bares y restaurantes.

Sólo con la página que tenemos en la mano ya podremos hacer muchas cosas. Junto al mapa de la ciudad tendremos acceso a un área titulada **Localizar**. Junto a esta área tendremos acceso a varios servicios que nos permiten encontrar hoteles en la región. Vamos a elegir el primero, *Hotels.com*. Selecciónalo y pincha en el botón **Go**.

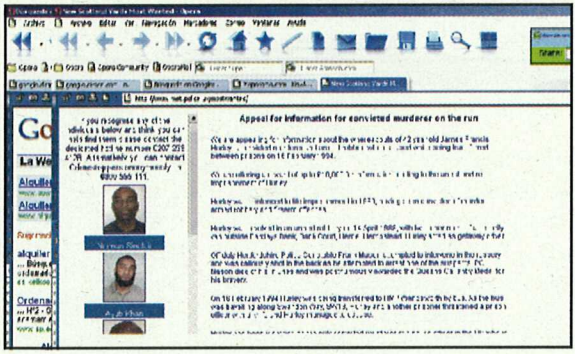


Veremos automáticamente con cuantos hoteles podemos contar en Hong Kong. Después, una lista justo debajo nos llevará hasta los enlaces de las páginas de esos hoteles. Ahora, basta con que elijamos uno, con que tengamos la tarjeta de crédito a mano y podremos hacer la reserva. Está claro que podemos realizar reservas en dos hoteles, o confirmar luego la reserva por teléfono o fax.

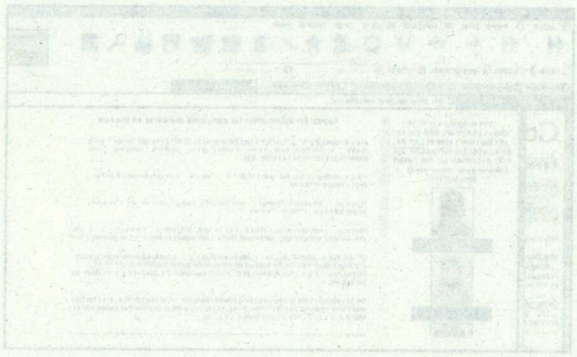
Se Busca

Como ya dijimos, Fogg tenía un policía, el agente Fisk, de Scotland Yard, que le pisaba los talones, pero él sólo se entera al final del viaje, cuando está a punto de perder la apuesta, por quedarse un día entero retenido en la cárcel hasta que se demuestra su inocencia.

Si Fogg (o incluso Fisk) dispusiese de una PDA con Google, este incidente tampoco habría ocurrido: Fogg realizaría el viaje tranquilamente, y Fisk no representaría el papel de idiota delante de sus superiores, intentando justificar la pasta gansa que se había gastado recorriendo medio mundo para capturar a un hombre que resultó ser inocente. El héroe del libro sólo tendría que acceder a http://www.met.police.uk/mostwanted, en la que se enumeran los hombres más buscados por Scotland Yard. Y después de llevarse un susto por ver su cara allí, junto a gente, digamos, no muy simpática, Fogg podría, con una sencilla llamada de teléfono, limpiar su buen nombre y seguir su viaje.

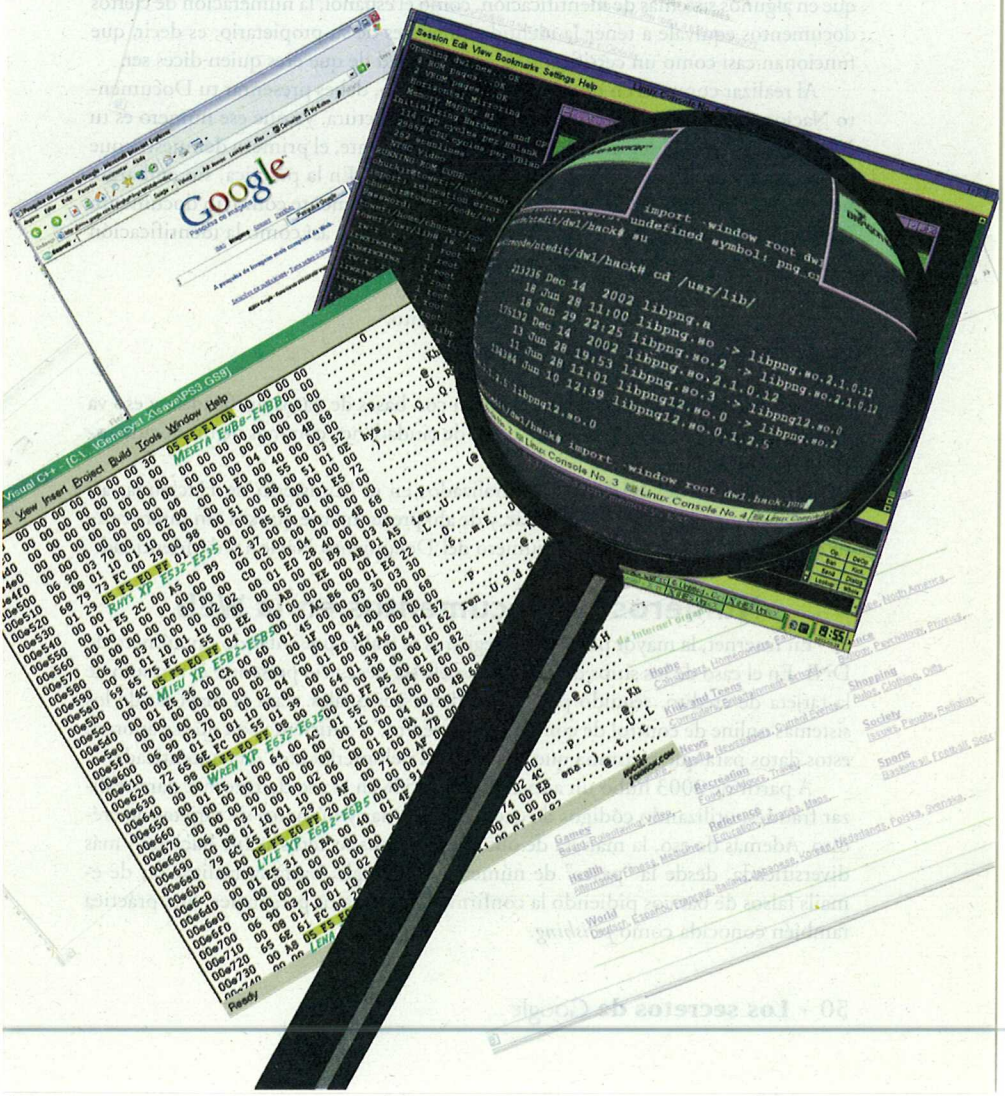


Buscando en la página informativa para policías de Inglaterra http://www.met.police.uk/appeals/intro_pages/wanted_current.htm, Fiks sabría, también, que otros policías estaban siguiendo la pista del ladrón de bancos que él perseguía, y que unos días antes de su vuelta a Londres había sido ya capturado.



Segunda Parte:

Google Hack



¡TÚ TAMBIÉN ESTÁS EN GOOGLE! ESTE BUSCADOR DA ACCESO A MILES DE DATOS PERSONALES Y PROFESIONALES

Los números de documentos son los datos mejor guardados del mundo real, ya que en algunos sistemas de identificación, como el español, la numeración de ciertos documentos equivale a tener la identidad digitales de su propietario, es decir, que funcionan casi como un certificado de autenticidad de que eres quien dices ser.

Al realizar compras en una tienda, por ejemplo, debes presentar tu Documento Nacional de Identidad si quieres que te hagan factura, ya que ese número es tu NIF, tu Número de Identificación Fiscal. Teóricamente, el primero demuestra que existes, y el segundo que pagas impuestos al Estado. En la práctica, sin embargo, no sucede lo mismo: el número del DNI se acepta solamente como un documento "formal", y la emisión de los documentos de identidad, así como la identificación de las personas, se realiza en las comisarías de Policía:

a) presentando el Certificado de Nacimiento,
b) identificando tus huellas digitales.

Las instituciones españolas ya cuentan con bases de datos unificadas, y eso va haciendo casi imposible que se den casos de ciudadanos con más de un documento de identidad.

Los registros del NIF están centralizados en el Ministerio de Hacienda, así que este número es difícil de falsificar, y de ahí precisamente que sea tan interesante su coincidencia como parte del número del Documento Nacional de Identidad.

Números de documentos en la Web

En Internet, la mayor parte de los registros se realiza presentando un número de DNI. En el caso de los sitios de compras, suele ser necesario presentar el número de la tarjeta de crédito, seguido por ciertos datos personales. Para la mayoría de los sistemas online de control de compras o de crédito, es suficiente con que se aporten estos datos para que se asuma que la identidad del interlocutor está autenticada.

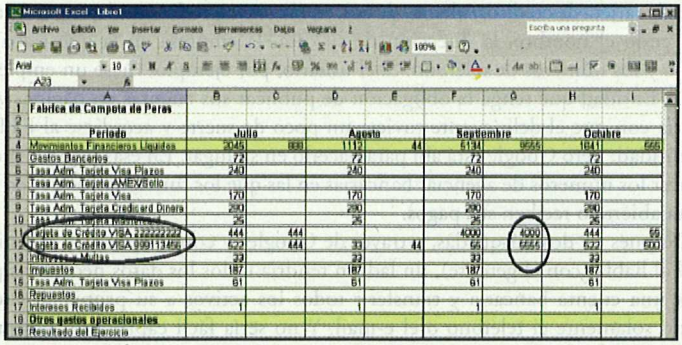
A partir de 2003 hubo un aumento del 275% en el uso de Internet para realizar fraudes, utilizando códigos de cuentas bancarias y números de tarjetas de crédito. Además de eso, la manera de obtener los datos secretos se fue haciendo más diversificada, desde la "pesca" de números en Google, hasta la utilización de e-mails falsos de bancos pidiendo la confirmación de datos personales, una práctica también conocida como *phishing*.

El sistema de búsqueda del Google funciona bien... ¡demasiado bien! Nada de lo que haya en Internet, incluso aunque esté escondido en varias ramas de subdirectorios, pasa desapercibido para Google. Ya te hemos comentado que basta con que los documentos estén situados en carpetas de red compartidas y desprotegidas, para que se hagan públicos en la Web. Así que una autorización, o una supuesta privacidad en la Red no ayudan mucho, sobre todo si el descuido es tuyo o de gente en quien hayas confiado para que vele por tus datos personales.

Un hacker que utilice Google u otros mecanismos de búsqueda rastreando datos personales puede así ir buscando documentos compartidos por descuido en la Web: hojas de cálculo, tablas, documentos de Word, bases de datos... todas estas modalidades de archivo suelen tener algún tipo de información relevante. De esta manera, utilizan la **Búsqueda Avanzada**, en la línea **Formato de Archivo**, para encontrar, por ejemplo, hojas de Excel, asociándolas a términos como *tarjeta*, *crédito* y *gastos*. Para informaciones más directas, un pirata podría limitar su búsqueda a fechas de hasta hace tres meses.

Y después de conseguir una lista de hojas que puedan contener números de tarjetas de crédito u otras informaciones, se podría incluso topar con un campo abonado: una hoja de cálculo localizada en el FTP de un escritorio de contabilidad.

Si no tiene suerte con la primera hoja, el delincuente tiene al menos una pista interesante: ya tienen un servicio FTP totalmente abierto, lleno de datos. Y ahí podrá encontrar una hoja de cálculo llamada **Tienda de Peras.xls,** en la que están catalogados los siguientes datos:



	A	B	C	D	E	F	G	H	I
1	Fabrica de Compota de Peras								
2									
3	Periodo	Julio		Agosto		Septiembre		Octubre	
4	Movimientos Financieros Líquidos	2045	888	1113	44	6134	9555	1841	555
5	Gastos Bancarios	72		72		72		72	
6	Tasa Adm. Tarjeta Visa Plazos	240		240		240		240	
7	Tasa Adm. Tarjeta AMEXBello								
8	Tasa Adm. Tarjeta Visa	170		170		170		170	
9	Tasa Adm. Tarjeta Credicard Dinera	200		200		200		200	
10	Tasa Adm. Tarjeta Diners	25		25		25		25	
11	Tarjeta de Credito VISA 2222222222	444	444			4000	4000	444	55
12	Tarjeta de Credito VISA 9999113456	522	444	33	44	55	5555	522	500
13	Intereses a Multas	33		33		33		33	
14	Impuestos	187		187		187		187	
15	Tasa Adm. Tarjeta Visa Plazos	61		61		61		61	
16	Repuestos								
17	Intereses Recibidos	1		1		1		1	
18	Otros gastos operacionales								
19	Resultado del Ejercicio								

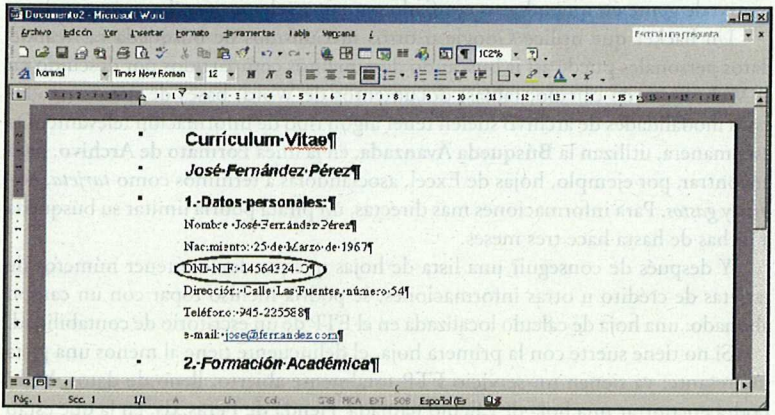
Primero, el hacker tendrá dos números de tarjeta de crédito, uno perteneciente a cada operadora de este servicio, y al mismo tiempo, verá valores muy altos en un mismo mes, que probablemente serán los correspondientes a los límites de las tarjetas.

Hacerse pasar por otra persona

Mucho menos complicado para un pirata es pescar en Google en busca de currículos dejados en Internet en páginas personales, puesto que la mayoría de los profesionales liberales crea este tipo de página precisamente para mostrar su carrera

profesional. El problema es que al hacer esto, pueden aparecer algunos datos del currículum impreso en la versión digital, que no deberían estar ahí, y que por descuido podrían acabar en manos de alguien que ponga en riesgo la seguridad personal de su autor.

De este modo, al pedir en **Búsqueda Avanzada** archivos .doc, asociados a los términos *currículum, José, Fernández, NIF*, se puede encontrar un documento como este:



El usuario insertó en el currículum todos sus números de documentos: sólo le faltaba la tarjeta de crédito. Además de eso, en él constan todos sus datos personales, incluyendo el nombre, la dirección de su casa, la empresa en la que trabaja o trabajó y los números de teléfono. Si el usuario, por ejemplo, fuese un empleado de una Entidad que tenga la obligación de publicar sus cuentas en páginas web, bastaría con que el delincuente tuviese un poco de suerte y accediese al FTP de esta entidad: pero Google está ahí para hacerlo en su lugar. En esas páginas suelen aparecer los números de agencias bancarias en las que los funcionarios tienen una cuenta abierta o reciben sus pagos.

Después de dos búsquedas, a través de Google (y con un poco de ingeniería social al hablar con el gerente), un ladrón tendría todos los datos necesarios para cerrar una cuenta bancaria y transferir todos los activos a su propia su cuenta usando solamente el teléfono o el e-mail. Y no sería fácil capturarlo, porque la mayoría de las cuentas que se suelen usar para transferir activos robados suelen ser falsas, o estar alquiladas por terceros.

USO DE CARACTERES COMODINES

Pero desde luego, no todos los documentos están expuestos en Internet con tanto descaro: muchos administradores de redes, secretarias y proveedores (cuando no se trata de alguien que ejerce varias de estas funciones), tienen el sentido común de esconder datos confidenciales, haciéndolos inaccesibles, o al menos no tan vulnerables.

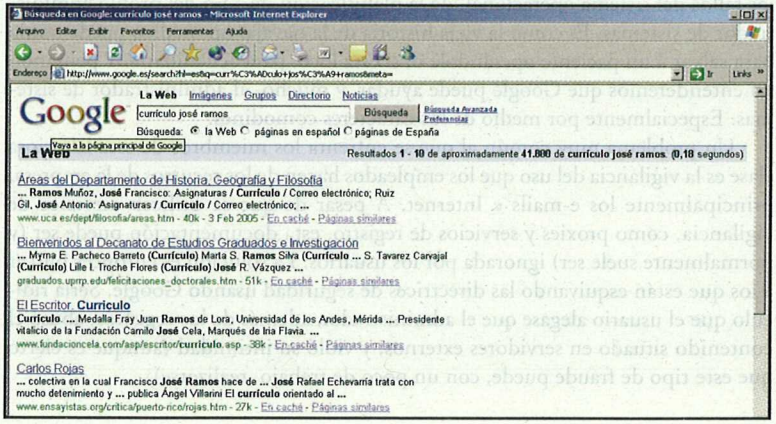
Esta es la recomendación típica de los administradores de sistemas, pero desgraciadamente no suele ser tenida en cuenta: los archivos de contraseñas jamás deben llamarse **logins.doc** o **contraseñas.txt**. Los procedimientos de conexión de un usuario y la política de contraseñas tienen que comunicarse al usuario oralmente, y nunca ser enviados mediante e-mail, y mucho menos los documentos que estén compartidos en la red. Por último, y abandonando los supuestos ahorros, que terminan saliendo caros, las empresas deberían tener dos servidores: uno para servicios online (FTP, e-mail, página institucional) y otro para el almacenamiento de datos confidenciales de la empresa.

Descubrir lo que hay oculto

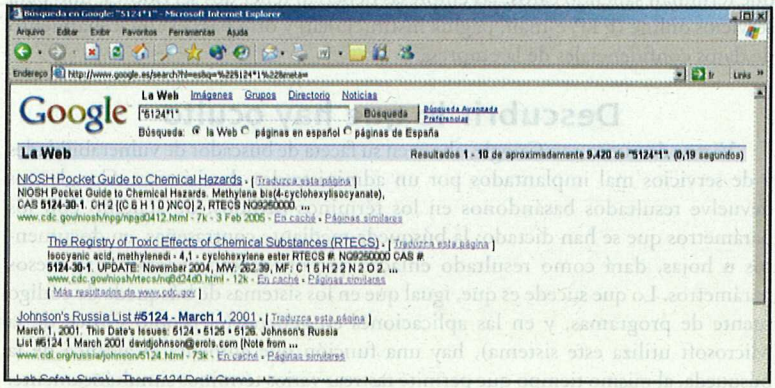
Vamos de nuevo con Google, ahora en su faceta de buscador de vulnerabilidades y de servicios mal implantados por un administrador de sistemas. Google nos devuelve resultados basándonos en los términos pedidos, de acuerdo con los parámetros que se han dictado: la búsqueda mediante contraseñas, en documentos u hojas, dará como resultado enlaces que obedezcan exactamente a esos parámetros. Lo que sucede es que, igual que en los sistemas de búsqueda de código fuente de programas, y en las aplicaciones en general (el propio Office de la Microsoft utiliza este sistema), hay una función que aumenta el alcance de la búsqueda, al mismo tiempo que permite rastrear varios términos simultáneamente: se trata de los caracteres comodines.

En la mayoría de los sistemas, tanto "*?*" (el símbolo de interrogación) como "*" (el asterisco) son caracteres comodines. En Google sólo puede ser utilizado el asterisco, pero funciona prácticamente como una navaja suiza: todas las funciones de los demás caracteres comodines pueden incorporarse usando el *.

El asterisco "*" sirve como un sustituto de cualquier palabra, sea cual sea: al hacer una búsqueda, por ejemplo, con *currículum Ramón* filetype:doc*, un pirata con paciencia obtendría los siguientes resultados:



Lo mismo vale para confirmar una expresión. Supongamos que un hacker busque la expresión *"5124*1"* en Google, que podría ser un código de confirmación bancario. Al incluir el asterisco en la expresión entre comillas permitirá que se recuperen todas las páginas en las que haya un código de confirmación bancaria formado por *5124*, un número de 0 a 9, seguido por un 1. El asterisco también podría permitir el descubrimiento de un código especial, utilizando letras del alfabeto.



Este mismo truco vale además para descubrir cuentas bancarias en diversos bancos. Al escribir la expresión *"cuentas corrientes banco *"*, el pirata tendría como resultado la lista de todo el contenido de la Web que esté relacionado con cuentas bancarias de varios bancos.

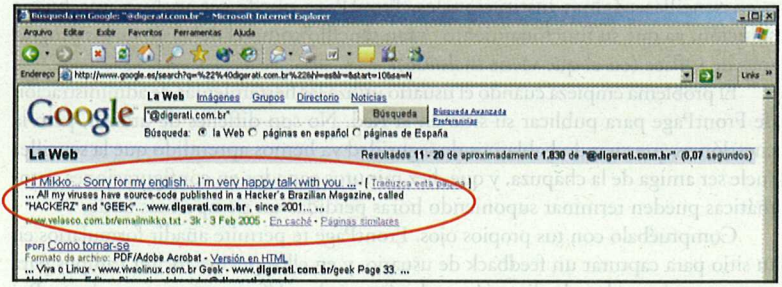
Para no olvidarnos de los administradores de sistemas

Hay que recordar aquí que la culpa de estas vulnerabilidades y agujeros no es de Google: el buscador es sólo un instrumento que puede utilizarse para explorar los fallos del sistema operacional, de la planificación de TI o del propio administrador de sistemas. Es como la vieja historia de la navaja: un médico puede usarla para salvar a un paciente operándole o, si está loco, para acuchillarle. Y pensando así entenderemos que Google puede ayudar, y mucho, al administrador de sistemas. Especialmente por medio de los caracteres comodines.

Un problema muy común al que se enfrenta los miembros de esta laboriosa clase es la vigilancia del uso que los empleados hacen de los recursos de la empresa, principalmente los e-mails e Internet. A pesar de que se utilicen servicios de vigilancia, como proxies y servicios de registro, esta documentación puede ser (y normalmente suele ser) ignorada por los usuarios. La solución es entonces buscar a los que están esquivando las directrices de seguridad usando Google. Sería ridículo que el usuario alegase que el administrador, además de los registros, abusó el contenido situado en servidores externos, y violó su intimidad (aunque es cierto que este tipo de fraude puede, con un poco de trabajo, realizarse).

Supongamos que el administrador quiere cortar el servicio de correo para su uso en newsgroups en el horario de trabajo, utilizando los e-mails de la empresa, pero las pruebas que contienen los registros siempre son rechazados por los usuarios, alegando que pueden haber sido manipulados. La pesca, para colmo de males, puede ser mucho mayor de lo que "el pescador" espera, ya que algunos empleados "en movimiento" (con ordenadores portátiles, PDAs, teletrabajadores) deberían también, dependiendo de la política de uso de la empresa, participar de las mismas reglas y sanciones, siempre que fuese posible registrar sus actividades.

Para empezar a enfrentarse a esta tarea titánica, nuestro administrador deberá buscar cuentas de correo de la empresa, perdidas en newsgroups. Con la ayuda de caracteres comodines, será suficiente, por ejemplo, realizar una búsqueda vía Google, utilizando el término **@digerati.com.br*



Hay otras combinaciones: realizar la búsqueda de usuarios asociándolos a tipos específicos de archivos (**filetype:doc**, **filetype:xls** etc.) por ejemplo, podría revelar desde espías hasta el manejo poco cuidadoso de documentos confidenciales. La utilización de nombres de usuarios, asociada a la búsqueda de directorios de servidores "sospechosos" como los utilizados para compartir MP3 (**\pub\mp3**, **\pub\musik**, **\pub\music**), permitirá descubrir también el destino del ancho de banda que la empresa tiene contratado, e incluso, si hay alguna máquina que esté siendo usada para compartir archivos MP3 en redes peer-to-peer.

Ilegalidad

A pesar de que no esté –todavía– prohibido oficialmente, compartir archivos, sobre todo canciones e imágenes de terceros protegidas por la legislación de Propiedad Intelectual, conductas de este tipo podrían convertirse en un problema para aquellas empresas en las que los empleados utilicen estos servicios. En Inglaterra, Nueva Zelanda y Australia (la patria del KaZaA), desde un año a esta parte, se ha convertido en una práctica común para la policía el ir sondeando servidores de empresas privadas, en busca de MP3 y películas piratas almacenadas en servidores Web o de otro tipo, y casi siempre encuentran alguna cosa...

DESCUBRIR QUIÉN ES EL DUEÑO DE UN SITIO

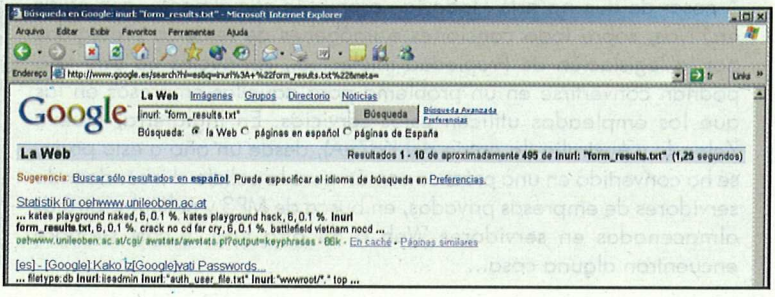
La combinación de caracteres comodines (junto con la posibilidad de buscar los directorios que, de uno u otro modo, pueden contener datos relacionados con la administración de un servidor), también puede utilizarse como una herramienta para invadir páginas. En este caso, la culpa de las vulnerabilidades no es ni de Google, ni del administrador de sistemas (o al menos, no toda la culpa).

Esta vez el gran culpable es FrontPage, el programa de maquetación de páginas Web de Microsoft. Para páginas sencillas, como las que se utilizan en Intranets o en sencillas páginas institucionales, FrontPage puede ser incluso una buena solución, ya que su funcionamiento es tan sencillo como el del resto de los programas de Office (en el que viene incluido).

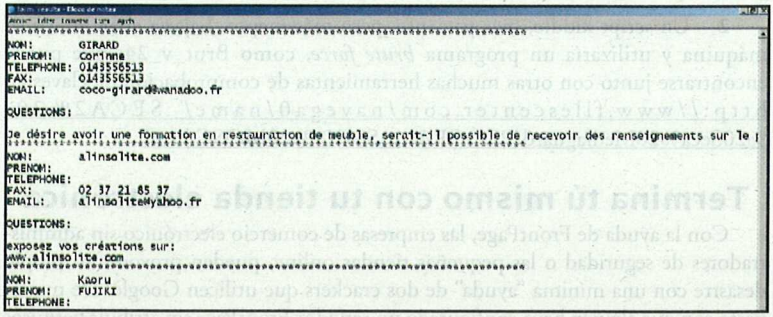
El problema empieza cuando el usuario utiliza las herramientas de administración de FrontPage para publicar su sitio en la Red. No son difíciles de utilizar, pero la cuestión es que cuando hablamos de seguridad ya hemos aprendido que la sencillez suele ser amiga de la chapuza, y que diez minutos ganados en configuraciones automáticas pueden terminar suponiendo horas perdidas un tiempo más tarde.

Compruébalo con tus propios ojos. FrontPage te permite añadir formularios en tu sitio para capturar un feedback de usuario, y en ellos puedes incluir campos para preguntar el nombre, la dirección y el número de teléfono, entre otros datos. Por defecto, esos datos se almacenan en un lugar denominado **called_private\form_results.txt**, en el árbol de directorios del sitio que FrontPage crea. Ten cuidado: no confundas el funcionamiento y la organización de FrontPage con la administración de páginas que se hace mediante el otro servidor de Microsoft, el IIS (Internet Information Service), que tiene *otras vulnerabilidades* (son muchas), pero no precisamente esta.

El archivo **_private** debería tener sus permisos configurados de manera que limitasen los accesos a sus datos, pero algunas veces las personas se olvidan de hacerlo. Y es este olvido lo que convierte en vulnerables a algunos directorios: para ver cómo ocurre esto, visita Google y mete la expresión *inurl: "form_results.txt"* en el campo de búsqueda. El resultado será este:



Vas a encontrar varios sitios expuestos. Muchos serán sitios de prueba, pero observa aquellas páginas que tengan mas de 1 KB de tamaño, y encontrarás algunas que enseñarán los nombres, los números de teléfono y las direcciones de e-mail de sus visitantes.



Otros comandos de búsqueda pueden mostrar contraseñas para cuadros de mensaje o foros. Escribiendo, por ejemplo, *inurl: "auth_user_file.txt"* es fácil encontrar archivos que contengan listas de usuarios, contraseñas y logins. Los archivos son más o menos parecidos a este:

El pedazo de texto que hemos resaltado es, en realidad, la contraseña de acceso a un servidor, que está, sin embargo, cifrada. Lo que no es un impedimento para un buen cracker o incluso para un script kiddie, por dos motivos:

1 – Un cracker guardaría los archivos de contraseñas en su máquina y ejecutaría algún programa que las descodificara, como LophtCrack. Hay versiones antiguas, más funcionales, de esta y de otras herramientas de ruptura de claves en http://www.evadenet.com/downloads/lophtcrack.shtml.

2 – Un script kiddie, más paciente, pero más inepto, bajaría el archivo a su máquina y utilizaría un programa *brute force*, como Brut_v_24, que puede encontrarse junto con otras muchas herramientas de comprobación de claves en http://www.filescenter.com/navega0/name/_SECA2%20-%20Seca%20Mediaguard/PROGRAMAS%20-%20MOSC.html.

Termina tú mismo con tu tienda electrónica

Con la ayuda de FrontPage, las empresas de comercio electrónico sin administrad[illegible]e seguridad o las pequeñas tiendas online, pueden provocar su propio desastre con una mínima "ayuda" de dos crackers que utilicen Google. De nuevo, basta con que alguien haya configurado un servidor FrontPage sin atribuir permisos a los directorios.

Un cracker podría sondear estos datos buscando el término *inurl: "orders.txt"*, y el resultado sería impresionante: es muy fácil encontrar archivos de texto de más de 2 MB, que reproducen bases de datos con números de tarjetas de crédito, direcciones y claves de confirmación de compras. Al buscar más datos, Google también podría ayudarle: buscando a través del término *"Index of /admin"*, que equivale a penetrar en el mundo de los servidores desprotegidos y puestos online por un trabajador con demasiada prisa. Una búsqueda simple da así como resultado más de 3.600 resultados...

Otras combinaciones posibles son:

"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 ellas devolverán algún archivo que esté olvidado en la configuración de FrontPage, pero, como hemos dicho más arriba, tampoco IIS se libra por completo de estos problemas. Basta con buscar por *inurl:iiadmin* o *inurl:"wwwroot/*"* (los dos directorios que son la base de los servidores basados en IIS) para ver cuántos resultados aparecen. Lo mismo vale para *inurl:"ftproot/*"*, que vendría a ser servir como una gran tienda de caramelos en la que un cracker puede encontrar desde bases de datos comprimidas -algunos administradores todavía se empeñan en realizar backups vía FTP- hasta documentos "confidenciales" y programas aún en fase desarrollo.

CONSULTAR EL LISTÍN TELEFÓNICO

Siguiendo lo que hemos expuesto en capítulos anteriores, podríamos imaginar que las consultas "clandestinas" a Google sólo alcanzarían a los administradores de sistemas o a usuarios muy avanzados de Internet. En definitiva, a quienes tengan una página en Internet, un currículum o un mensaje colgado en un newsgroup, que permita que desconocidos anónimos entren en su vida personal.

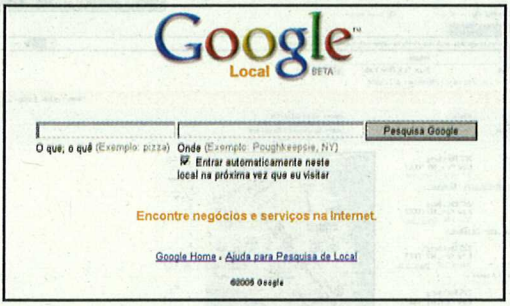
Pero eso no ocurre. Google, en un primer momento, y fue seguido después por la gran mayoría de los servicios públicos, se convirtió en un gran depósito de datos personales y de maneras de aproximarse a una persona sin haberla visto nunca de frente. Basta con tener una conexión a Internet y un poco de paciencia para saber todo sobre la vida de mucha gente, aunque éstos no se hayan conectado en su vida a la Red.

El listín telefónico de Google

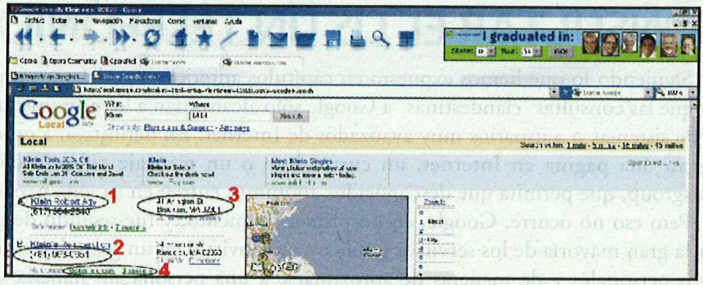
Google tiene un servicio en pruebas que, por el momento, sólo está disponible para Estados Unidos, y que permite encontrar datos vía Internet obre cualquier persona que tenga un nombre real y una dirección fija en el territorio americano (Alaska y Hawai incluidos). A pesar de estar aún en su versión Beta, permite que después de encontrar a esa persona se pueda visualizar un mapa de los alrededores de su residencia, y la mejor manera de llegar hasta allí.

Para empezar, entra en esta dirección http://local.google.com/lochp, la dirección de Google local. En el primer campo, escribe el apellido de la persona o familia que quieras encontrar en Estados Unidos. En el segundo campo, añade por lo menos un dato extra, como la dirección, la ciudad, el Estado o el Zip Code (el código postal). Recuerda que el Zip Code está formado por cinco números.

En nuestro caso hemos hecho una búsqueda de la familia Klein, de Nueva York (no te olvides de escribir "New" York). Sabemos que su Zip Code es 02020, pero, como no tenemos una guía postal estadounidense a mano, no sabemos cuál es el nombre de la calle en la que vivien, ni el número, ni en que barrio tienen su residencia.



En la pantalla de resultados, accederemos a diversas informaciones. Ahora, sólo hay que aprender a leerlas.

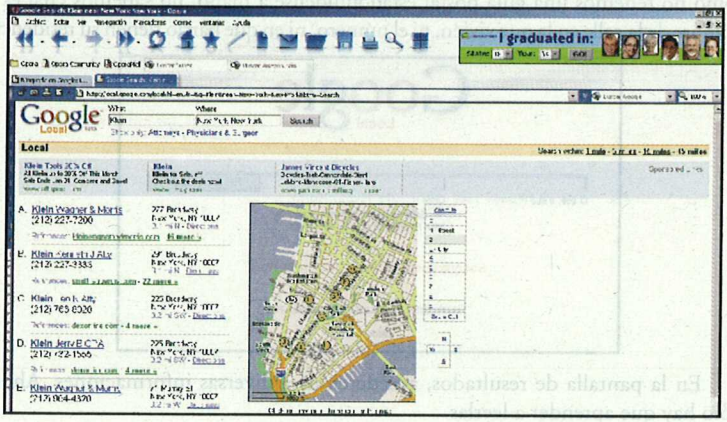


El nombre en azul que ves a la izquierda (1) es el nombre completo de la persona, comercio o institución que coincide con el nombre que has pedido. Justo debajo (2), veremos su número del teléfono, incluyendo el código interurbano de los Estados Unidos entre comillas. Al llamar desde España, recuerda que para el marcado deberías anteponer el código internacional de Estados Unidos (001).

En (3), tenemos la dirección del usuario. Y ahí descubrimos nuestro error, el código postal 02020 no es de Nueva York, pero sí de Massachusetts. Al notar este tipo de error de Google, lo ideal es repetir la consulta sin el Zip Code (ya que éste está equivocado), es decir, usando el nombre de la ciudad, y así veremos varias direcciones con la ciudad, el Estado y los códigos postales que son correctos.

En el apartado (4), podemos descubrir si la persona que buscamos tiene sitio web en Internet, lo que nos permitiría descubrir otras informaciones, como su dirección profesional y sus direcciones de e-mail, entre otras cosas.

Esta página también nos muestra el mapa de la región en la que se encuentra la dirección que hemos consultado, enumerando cada uno de los resultados en el mapa. Haciendo clic en **Zoom Out** o en **Zoom In**, podremos obtener detalles de los alrededores de la dirección o descubrir rutas generales, como estaciones de metro o las avenidas para llegar a la dirección deseada.

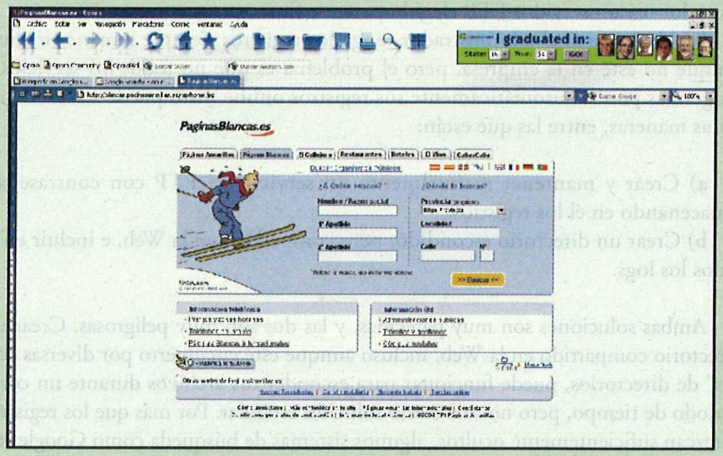


Teléfonos de España

Mientras Google no disponga de ese mismo servicio para otros países, como España, Telefónica sí que permite la consulta del listín telefónico a través de su Web. En el sitio http://www.paginablancas.es, de Telefónica, podemos conseguir los números de teléfono de todos los abonados a líneas.

Hay una ventaja de este servicio respecto al de Google: puedes pedir que tu nombre no figure en este listado, tanto online como impreso. Para hacerlo, vete a http://www.telefonica.es/casa. En la página que aparezca, haz clic en el enlace **Consultar/Modificar Datos**, en donde tendrás que introducir tu **Nombre de usuario** y tu **Contraseña**. Si entras por vez primera en este sitio lo primero que tendrás que hacer es registrarte; para ello, entra en el enlace http://www.telefonicaonline.com/on, y acto seguido en **Regístrese**, que aparece en la parte superior de la pantalla. Serás enviado a otra pantalla en la que te darán las indicaciones sobre lo que tienes que tener para poder registrarte y en donde tendrás que introducir los datos que se te piden: **Teléfono**, NIF, **e-mail** y el **Nombre de usuario** con el que quieres aparecer. Pulsa después en **Continuar**, y aparecerá una pantalla de confirmación de datos. Si todo es correcto, pulsa en **Ok**. La contraseña te será enviada por correo electrónico, para que después puedas entrar y modificar tus datos en el primer enlace.

Para hacer búsquedas en listitas telefónicos de toda España puedes buscar en http://www.paginasblancas.es o en el sitio de QDQ http://blancas.qdq.com.



ANALIZAR LOGS USANDO GOOGLE

Un administrador de sistemas suele tener más trabajo que brazos para hacerlo, y más cosas para recordar que cerebro para guardarlas, por eso precisamente existen los registros o logs. Los archivos de registro son pequeños documentos de texto que se actualizan automáticamente por distintos programas, de acuerdo con las intenciones del administrador. Hay registros para las funciones más diversas, desde el login de usuarios en máquinas locales, incluyendo los intentos de conexión que no han éxito, hasta registros de acceso a archivos o servicios distribuidos mediante servidores.

Desde el punto de vista de los usuarios, el registro suele ser la última herramienta educativa, o mejor dicho, coactiva, de la que el administrador del sistema dispone para evitar las aventuras de los usuarios de la red fuera de la red, y para frustrar sus intenciones de hacer lo que quiera. Los usuarios suelen resolver este problema, como ya hemos visto, rechazando los registros administrativos recordando su intimidad, y aduciendo que de todas maneras pueden ser modificados manualmente para invalidarlos, por lo que no servirían, según ellos, para prevenir los agujeros de seguridad.

Registros remotos

La mayor parte de los programas de la red genera registros, que se pueden almacenar tanto en el servidor en el que se ejecuta el programa como en las máquinas preparadas especialmente para almacenar registros (los servidores de registros). Algunos programas tienen una opción de enviar los logs al correo del administrador o a un servidor FTP elegido para ese fin.

Lo ideal es que el administrador tenga los registros siempre a mano, incluso aunque no esté en la empresa, pero el problema es que no todos los servicios y programas ponen automáticamente sus registros online. Esto puede resolverse de varias maneras, entre las que están:

a) Crear y mantener manualmente un servicio de FTP con contraseñas, almacenando en él los registros.

b) Crear un directorio escondido, pero compartido en la Web, e incluir en él todos los logs.

Ambas soluciones son muy parecidas, y las dos son muy peligrosas. Crear un directorio compartido en la Web, incluso aunque esté encubierto por diversas "capas" de directorios, puede funcionar para esconder tus archivos durante un corto periodo de tiempo, pero no vale como solución estándar. Por más que los registros parezcan suficientemente ocultos, algunos sistemas de búsqueda como Google son capaces de "invadir" los servidores, buscando caminos de directorio internos.

Haciendo una búsqueda con *inurl:"/admin/logs"* en el campo de búsqueda de Google, encontraremos diversos servidores, en su mayoría servidores Web Apache, desprotegidos o mal configurados, con muchos registros que podremos ver. Para asegurarnos de que estamos ante de un servidor Apache, observa si la ruta completa es **/www/admin/logs**. La ruta **/wwwroot/logs** u otra cualquiera, contenida en el directorio **/wwwroot** será una señal clara de que estamos ante servidores que ejecutan IIS (Internet Information Service).

Visible y vulnerable

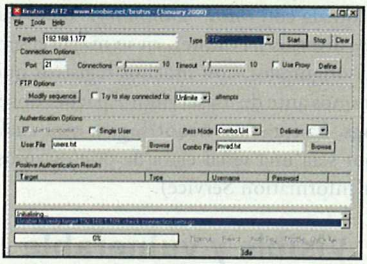
Al mismo tiempo, los servidores FTP que contengan logs, incluso protegidos por contraseña, pueden hacerse visibles mediante Google. Al hacer una consulta utilizando el parámetro *inurl:"/ftp/log"*, Google devuelve una infinidad de enlaces que contienen esos directorios. Al intentar acceder a ellos, en muchos no se puede entrar: eso es señal de que el administrador sí ha restringido los permisos de acceso a los logs o que activó una contraseña. Este último recurso es fácil de ignorar: los crackers suelen invertir la dirección del registro, buscando sólo la dirección que corresponde al FTP.

Un ejemplo: al acceder al enlace http://www.test.com/test/servidores/headers/ws_/ftp.log, el acceso es inmediatamente rechazado, mostrando un mensaje. Pero abriendo una nueva ventana del navegador se puede retocar la dirección del enlace. Si sustituimos www por FTP, obtendremos ftp://ftp.test.com.

Si el servidor FTP no está configurado para recibir conexiones anónimas directas (una modalidad de conexión en la que no aparece la pantalla para la introducción de la contraseña), o indirectas (cuando sólo hay que escribir *anonymous* en el usuario y dejar la clave en blanco) el cracker tendrá que introducir un nombre de usuario que ya esté registrado en el sistema FTP y una contraseña.

Para conseguir acceder a todos los archivos del servidor FTP, incluidos los registros, el cracker puede utilizar las cuentas del administrador si dispone de ellas, que proporcionan a quien las tiene un poder ilimitado. Otra posibilidad es intentar acceder utilizando contraseñas mal configuradas, que han sido olvidadas por el administrador o por el dueño de la cuenta FTP (que no tienen por qué ser la misma persona). Pese a que sean vulnerabilidades antiguas y muy explotadas y conocidas por todos los que pasan por lo menos, por un curso simplón de webmaster, encontrar puertas abiertas escribiendo cosas del estilo de login/clave (conexión/contraseña) en Google, *admin/admin*, *administrador/administrador*, *ftp/ftp* o *ftp/"clave en blanco"* es más común de lo que te imaginas.

Cuando esto no es posible, la única manera es utilizar técnicas más agresivas y burras, como las de los sistemas de *brute force*. Se califican como agresivas porque los programas de este tipo para Windows, como Brute Force AT2 (http://www.hoobie.net/brutus/brutus-aet2.zip), acompañados de un archivo de texto que contiene millares de especificaciones para de usuario y contraseña, y que serán leídas por ellos, son capaces de hacer más de 3.000 intentos de conexión en una hora. Y eso utilizando máquinas antiguas, como 486 y una conexión de 56 kbps, el equivalente a una línea telefónica común.



Y burras, porque si no tienes éxito, van a dejar una enorme huella en el registro con todos los intentos de acceso que has hecho, con tu dirección IP, que podrá ser rastreada fácilmente por el administrador usando el comando traceroute (que incorporan en varios sistemas operativos basados en UNIX) hasta llegar a tu proveedor y, si tienes una IP fija (que no cambia a cada reconexión a Internet), incluso hasta tu cuenta de acceso a internet.

Esta es una de las cosas que la mayoría que los crackers y script kiddies aprenden al invadir un servidor logeándose en él y consiguiendo privilegios de administrador: borrar, por si acaso, los archivos de registro (log). Eso es muy fácil de hacer, sólo hay que acceder al terminal de línea de comando, y programar el siguiente bash (script de línea de comando de Unix):

```
#!/bin/bash
```

cd var/log (directorio de almacenamiento estándar de los logs en sistemas basados en UNIX)

```
for l in 'ls -p|grep '/''; do
-n >$l &>/dev/null
echo haciendo archivo $l...
done
echo Limpieza de los archivos de registro concluida!
```

Un buen administrador de sistemas, sin embargo, nunca dejaría que **/var/log**, fuese el almacén de todos los registros del sistema. Se pueden copiar, por ejemplo, todos los registros en otro directorio:

 # cp –a/var/log (directorio de destino)

Un paso más sería enviar el archivo a un directorio de destino en otra máquina conectada a la red:

 # cp –a/var/log/192.168.0.1/(directorio de destino)

Y un paso mucho mas avanzado sería el de borrar los archivos originales después de copiarlos en un almacén de logs:

 # cp –af/var/log/192.168.0.1/(directorio de destino)

RASTREAR CARTAS Y ENTREGAS

Internet ha facilitado mucho la comunicación entre las personas. Cartas que antes tardaban días en llegar su destino, el telégrafo e incluso los caros envíos de fax a sucursales de empresas en el extranjero... todos han sido sustituidos por el e-mail. El correo electrónico es una forma de comunicación instantánea capaz de enviar, junto con los mensajes, varias páginas de documentos, archivos de todo tipo e incluso pequeños programas, dependiendo de la capacidad del servidor de e-mails y del ancho de banda disponible.

Pero a veces es necesario enviar documentos reales, grandes cantidades de archivos, videos o incluso objetos. En estos casos no hay otra manera de hacerlo que la de usar los medios tradicionales: juntar todo lo que haya que enviar en un sobre o en una caja bien cerrada, ir a la oficina de correos más próxima, y enviar el paquete (preferentemente por medio de servicios fiables, como Postal Exprés). Además de eso, podemos hacer un seguro para que nos reembolsen el valor nominal del paquete en caso de que se pierda, aunque eso ocurre raras veces. Correos suele ir lento, pero seguro.

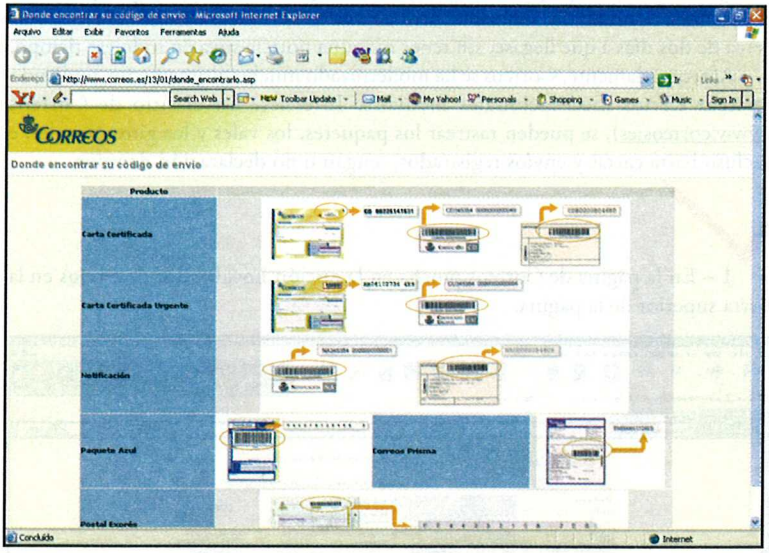
Hace unos años esta operación sumía a quien hacía el envío en un periodo de incógnita. Enviar un paquete mediante Postal Exprés, o mandar un giro postal, de remitente a destinatario, incluso dentro de una misma ciudad, equivalía a esperar cerca de dos días a que llegase, sin tener ninguna noticia suya en todo ese tiempo.

Afortunadamente, Correos se ha modernizado mucho últimamente, pudiendo rastrear envíos internacionales mediante Internet. En el sitio de Correos (www.correos.es), se pueden rastrear los paquetes, los vales y los giros postales, e incluso hasta cartas y envíos registrados, tengan o no declaración de valor.

1 – En la página de Correos pincha en la opción **Localizador de envíos** en la barra superior de la página.

2 – Serás enviado a esta página, *Localizador de envíos.*

3 – Pinchando en el enlace **Dónde encontrar su código de envío** (1) aparecerá una lista con todos los objetos que pueden ser rastreados y los números de los códigos de barras que hay que utilizar en la búsqueda.



4 – Para hacer una búsqueda de un objeto que ya haya sido enviado por correo, vete a la página anterior escribe el código del objeto y pincha en **Consultar estado** (2).

5 – En la página siguiente aparecerá directamente el estado del envío.

Carta nacional

Otro truco interesante, aunque un poco escondido en el sitio de Correos, es el que permite que, incluso estando en el extranjero, envíes una carta a un destinatario de España, pagando la tarifa postal nacional. Este servicio es útil cuando el destinatario no tiene e-mail o hay una reunión o una conversación que debe documentarse mediante carta. El servicio está disponible en el enlace http://www.correosonline.com.br/pt_product.asp?dept_id=1&sku=0&indice=1, y puedes incluso pedir que te avisen electrónicamente cuando se reciba tu envío, pinchando en **Servicios Adicionales > Aviso de Recibo**.

Después de escribir esta carta, sólo tienes que hacer clic en **Añadir a la cesta**, para enviarla, con un coste ligeramente superior. Es un poco caro, pero para quien esté en otro país y no sepa cómo funciona el correo local, ayuda mucho. Por último, sólo hay que pinchar en **Pasar por la caja** para pagar la carta (debes tener una tarjeta de crédito). La carta se imprimirá en Correos, se sellará y se enviará al destinatario. Y Correos garantiza que la correspondencia llegará como máximo al medio día de la jornada siguiente al envío.

Rastrear en FedEx con Google

Usando Google, cualquier persona puede seguir la localización actual de un envío de Federal Express (FedEx), el servicio más famoso de paquetería del mundo, que es capaz de llevar envíos hasta regiones en las que las calles no tienen nombre o donde las casas no tienen numeración, siempre que haya una mínima referencia.

Para hacer esto, vete a la página de inicio de Google y escribe los 12 dígitos de identificación del envío, entre comillas, en el campo de búsqueda. Un buen truco es hacer una captura de esta pantalla y guardarla, por si se produce algún problema de retraso con el envío.

GOOGLE, EL AMIGO DE LOS CRACKERS

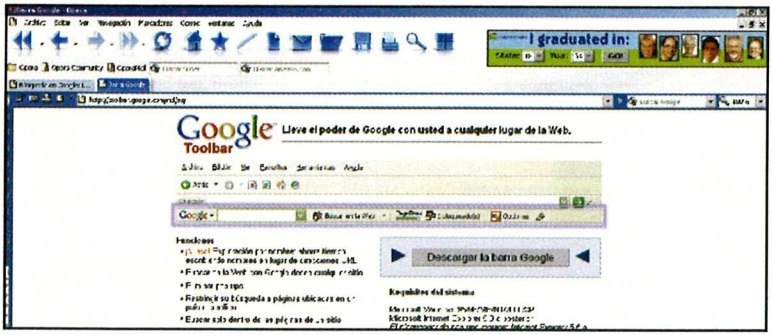
Todo lo que hemos visto hasta ahora nos muestra el inmenso poder que Google, y por extensión, todos los servicios de búsqueda masiva en Internet, tienen para facilitar nuestra vida. Cosas como descubrir teléfonos, direcciones o localizar envíos pueden salvar carreras profesionales en muchos casos, o ayudar al reencuentro de parientes, entre otras cosas.

Desgraciadamente, crackers y aspirantes a hacker también pueden utilizar los datos que consigan mediante Google para hacerse pasar por otras personas, para hacer compras en su nombre etc... Además de esto, hay (como ya hemos visto en los capítulos dedicados a los logs y las vulnerabilidades que afectan a las informaciones sobre páginas y servidores Web), recursos de Google que pueden ser utilizados, **de forma específica**, para abrir o explorar fallos de seguridad.

Y, por increíble que parezca, hay aún muchas vulnerabilidades y trucos por explotar. Basta con que pongamos juntos a un cracker, a un usuario o a un servidor, y a un navegador accediendo a Google... tanto en el lado del atacante como en de la víctima.

La barra de herramientas de Google

A mediados de 2003 Google puso a disposición de los usuarios un programa para descargar, su *Google Deskbar*, una barra de herramientas que se añade al escritorio de Windows como un conector para acercar todas las funciones de Google a tu barra de tareas, además de añadir alguna más, como el bloqueo de los pop-ups o la búsqueda directa de programas en **Download.com**. El programa puede bajarse de esta URL http://toolbar.google.com/intl/es. Para hacerlo, tendrás que deshabilitar el bloqueo de cookies: en el Internet Explorer 6, vete al menú **Herramientas > Opciones de Internet > Privacidad** y haz un clic en **Aceptar siempre las Cookies de sesión.**



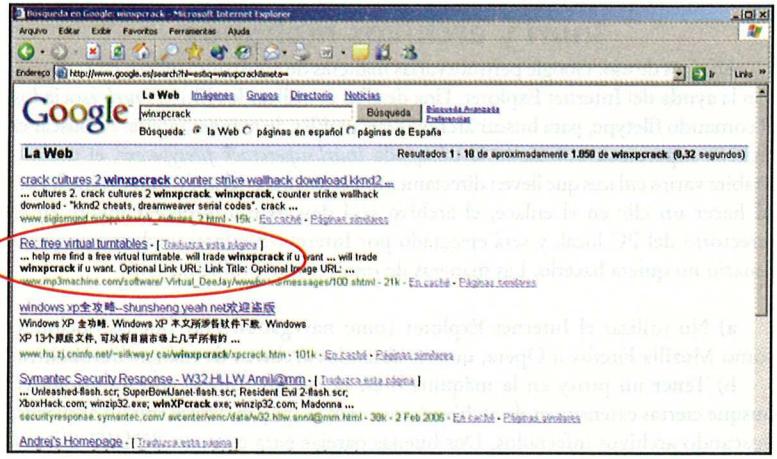
A pesar de que es muy útil, esta barra de herramientas, al estar íntimamente ligada al Internet Explorer, comparte con él todas sus vulnerabilidades, tanto aquellas que son ya viejas conocidas del universo cracker, como las más recientes. Hoy en día se descubren alrededor tres vulnerabilidades del Internet Explorer por semana, lo que hace que el CERT, uno de los mayores centros de seguridad del mundo, situado en los Estados Unidos, haya llegado incluso a recomendar el uso de otros navegadores, mientras se realicen correcciones sustanciales en el código del IE.

Supongamos que un cracker invadiese un sitio web con una media razonable de visitantes, que consiguiese penetrar además en la cuenta del administrador, y teniendo acceso a los directorios de la página, incluyera varios archivos ejecutables (.exe, .com) en su estructura, dándoles nombres sugerentes como **cracker_WarcraftIII**, **claves_porno** o **hackertools**.

A menos que esta página esté abandonada, el administrador notará pronto que se han incluido enlaces en ella con código maligno, así que enlaces del tipo **yosoyunvirus.exe** en todo el medio de una página de noticias serían inmediatamente eliminados, simplemente con que su dueño la revise con cierta frecuencia.

Y aquí entra el arte: con la contraseña del administrador, cualquiera puede subir archivos ejecutables, y esconderlos en los árboles de directorios sin establecer enlaces definidos. Después sólo hay que contar con Google.

Los crackers suelen decir que Google rastrea siguiendo enlaces en busca de archivos y caminos para un directorio, incluso aunque sean invisibles a primera vista, y cuentan con esto, además de con Explorer de Microsoft, para diseminar sus archivos malignos. Al hacer una búsqueda en Google con el término *winxpcrack*, por ejemplo, un usuario que no lo sepa estará buscando una manera de activar una versión ilegal del sistema operativo de Microsoft. Aparecerán al menos 600 resultados, divididos entre los programas propiamente dichos y enlaces hacia el programa, como mostramos a continuación:

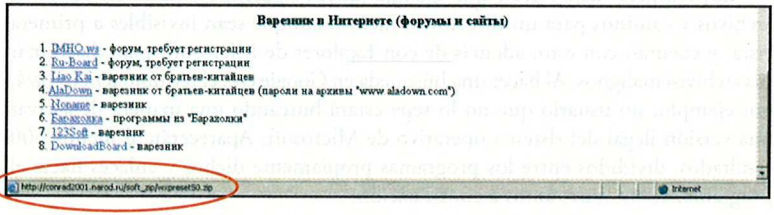


Enlaces y virus

Usando únicamente los enlaces y sus descripciones, un usuario que no sea experto no será capaz de distinguir si está en el interior de una página. En el ejemplo superior, tenemos una página en la que podemos encontrar el crack que andábamos buscando: el crack para Windows Xp. Pero, al hacer clic en el enlace notamos que es una colección de enlaces y descripciones sumarias de cada crack. Los que estén acostumbrados a este tipo de búsquedas sabrán de inmediato que esos pequeños programas suelen estar colgados en formato zip, rar, exe o com. En nuestro caso, al colocar el puntero del ratón sobre el enlace, descubriremos que este archivo se llama **wxpreset50.zip**.

Pero esto no significa que el archivo sea lo que dice ser. Las configuraciones estándar de **Windows** (**Windows Explorer > Herramientas > Opciones de Carpeta > Modos de exhibición > Ocultar las extensiones de los tipos de archivos conocidos**) ocultan las extensiones de archivo en todo el sistema, incluido el navegador. Así, al bajar el archivo **wxpreset50.zip** con el comando de ocultar extensiones desmarcado, el usuario puede, verdaderamente, estar bajando el archivo **wxpreset50.zip.exe**.

Está claro que siempre puedes identificar los archivos usando la barra de nombres del programa. Observa la figura inferior, en la que pese a las extensiones de archivos ocultas, se puede ver la auténtica extensión del archivo.



Inurl y archivos malignos

Además de eso, Google permite varias maneras de construir verdaderas trampas, con la ayuda del Internet Explorer. Una de ellas es utilizar los filtros *inurl*, asociados al comando filetype, para buscar archivos ejecutables de manera rápida. Al buscar el archivo **supercrack** utilizando la búsqueda *inurl:supercrack filetype:exe*, el usuario recibirá varios enlaces que lleven directamente a la descarga del archivo supercrack.exe. Al hacer un clic en el enlace, el archivo será descargado inmediatamente en el directorio del PC local, y será ejecutado por Internet Explorer, incluso aunque el usuario no quiera hacerlo. Las maneras de resolver este problema son:

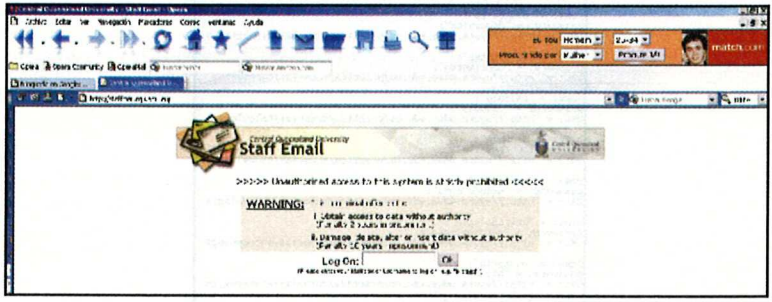
a) No utilizar el Internet Explorer como navegador. Hay buenas opciones, como Mozilla Firefox u Opera, que no abren los archivos de manera automática.

b) Tener un proxy en la máquina o en la red que, junto con un antivirus, busque ciertas extensiones de archivo (.exe, .com, .zip) y las revise con un antivirus, buscando archivos infectados. Dos buenas parejas para esto son SQUID (http://

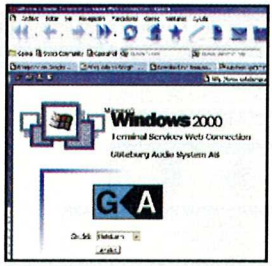
www.squid-cache.org/Versions/v2/2.5/squid-2.5.STABLE6.tar.gz) junto con F-Prot (http://www.f-prot.com/download/download_fplinux.html), para Linux, y WinConnection más AVG (http://www.softonic.com/ie/18522) para Windows.

Descubrir máquinas que son vulnerables

Abandonando una postura pasiva, un cracker también puede utilizar Google para realizar una búsqueda agresiva de posibles objetivos. Por ejemplo, al buscar la entrada de servidores Exchange (el servidor de e-mail corporativo de Microsoft), muy utilizado en empresas e instituciones públicas, podría utilizar el comando `intitle:Exchange Server Login`, que devolvería todas las páginas en las que figuren los términos Exchange Server en el título de la página. Una búsqueda de dos minutos más podría devolver páginas como ésta, en la que basta sólo hay que introducir el nombre de usuario en el buzón para tener acceso a todos sus e-mails:

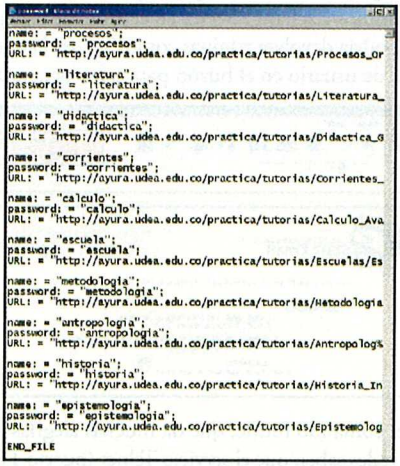


Otra utilidad del comando intitle, que da muchas alegrías a los intrusos: todos los que trabajan con redes saben que el servicio Telnet (puerto TCP 23), junto con el FTP (puerta 21) es uno de los servicios más inseguros de la historia de redes. A su vez, Microsoft dispone de un servicio de administración remota, Terminal Server, que usa muchas funciones de Telnet, y que es utilizado para administrar hosts remotos (aquellos que están en otras redes, o que están unidos directamente a Internet). Un cracker que busque servidores específicos que ejecuten Windows 2000 Server, puede hacer una búsqueda en Google utilizando el siguiente parámetro *intitle:Terminal Server Webs Connection*. El resultado podría ser el hallazgo de sitios como este:

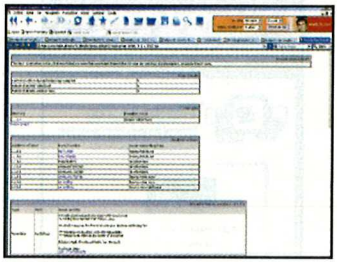


Bien, en este caso, el servicio es más difícil de invadir mediante un navegador. Lo que no quiere decir que intentando hacer una conexión por Telnet o apuntando un escáner de vulnerabilidades hacia la dirección de esta página, no se puedan revelar varias brechas en la conexión o en la administración de puertos y servicios que permitan la invasión.

No es difícil conseguir algunas conexiones mediante Google sin pegar ni robar a nadie. Los administradores tienen una pésima costumbre: anotar usuarios y claves en documentos, como archivos de Word y Excel, lo que ayuda mucho a cualquier invasor si estos terminan en Internet. Al escribir el comando inurl:password.log, es posible que veamos archivos como el que mostramos abajo. Increíble, pero cierto.



¿Quién necesita herramientas de invasión después de tener una lista completa de claves? Y más sencillo aún que esto es regalarle al invasor una lista con todas las vulnerabilidades de la red... para ello sólo hay que dejar en el servidor web informes creados por Nessus, un escáner de agujeros de seguridad. Al buscar: *this file was generated by nessus* en Google se pueden encontrar informes completos de todos los agujeros de un sistema, y eso sin usar ni un solo recurso, casi sin necesidad de saber nada de informática.



PASSWORD GENERATOR CON GOOGLE

Antes de Google, la vida de un cracker era mucho más difícil, sobre todo para los principiantes: conseguir recursos, contraseñas de programas, contraseñas de sitios, encontrar servidores desprotegidos e incluso encontrar servicios disponibles, que funcionasen a través de módems desprotegidos, era bastante complicado. Esta última modalidad era conocida como *war dialing,* y consistía en hacer que un programa de marcado telefónico barriese un rango de números telefónicos, en busca de módems sin claves, con claves estándar o con claves fáciles de romper. Normalmente se tardaban unos cuantos días en encontrar un sistema desprotegido de esa manera, y además de la satisfacción personal de saber que se estaba haciendo algo malo, la recompensa venía en forma de cuenta telefónica, a la que se podían cargar llamadas de hasta cuatro cifras.

Hoy en día, y sólo con la **Página de Búsqueda Avanzada** de Google se pueden sustituir toda una serie de herramientas por un simple navegador. Los generadores de passwords, escáneres y pruebas de vulnerabilidades dejan de ser útiles y provocan que tu antivirus se dispare continuamente, avisando de que hay troyanos intentando dar acceso a un tercero a tu máquina.

En esta *Segunda Parte* de nuestro libro ya hemos visto muchas funciones de Google que permiten encontrar servidores vulnerables, con algunas de las brechas más ridículas que se han visto. Vamos a ver algunas más, especialmente aquellas que están relacionadas con claves, con licencias de programa y con listas de acceso, y demostraremos porqué el administrador de sistemas debe hacer mucho más que poner, simplemente, el servidor en el aire.

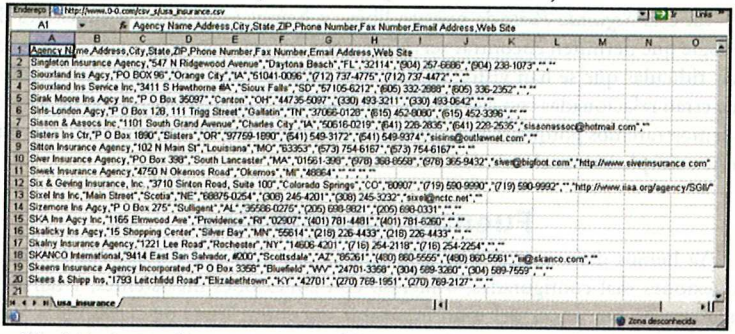
Fuentes inagotables

Ya hemos demostrado lo sencillo que es encontrar servidores vulnerables a conexiones mal configuradas, o incluso con conexiones inexistentes, y también hemos visto que no es del todo imposible encontrar listas de claves en la Web usando Google. Pero... ¿y si fuera posible tener una colección inagotable de conexiones, claves y direcciones de e-mails, además de sus respectivas direcciones de servidores con la ayuda del Google?

Desgraciadamente, y una vez más, esto es posible. Mediante la búsqueda avanzada se pueden conseguir tantas direcciones de e-mail y contraseñas como quieras, y crear una lista de contraseñas y grabarlas en un CD, junto con una lista de servidores y con un programa de ataques *brute force*, para tener nuestra propia máquina de craquear. Todo ello en menos de una hora. Recordemos que este procedimiento lo vamos a enseñar con el fin de alertar a los administradores de sistemas menos atentos, y para demostrar que los sistemas informáticos pueden ser susceptibles de ataques realizados con recursos mínimos, y por personas sin apenas conocimientos, no para alentar este tipo de ataques.

Para empezar a crear un password generator, el cracker puede usar listas de e-mails reales. Muchos sistemas utilizan direcciones de correo electrónico como login, o el nombre completo del usuario y, algunas veces, la descripción de su cargo o de su departamento. Estos datos son en principio confidenciales, y la manera de que todos los usuarios del sistema se logean (aunque esto puede verse ojeando directamente el front-end del servicio de conexión, los servidores Outlook Mail y Exchange Server son especialmente vulnerables), solo debería ser conocida por el administrador de sistemas o por alguien con acceso autorizado a todas las cuentas, además, claro, de por sus propios usuarios.

Vamos a buscar en Google usando el parámetro *e-mail address filetype:csv csv*, que equivale a decir que queremos encontrar documentos csv (el formato utilizado por Outlook Express y por Outlook, y por los servicios de correos estándar de Windows y de Microsoft Office para exportar los datos en forma de archivo de texto). Con el término *e-mail address* indicaremos a Google que queremos solamente archivos "llenos", es decir, que contengan efectivamente direcciones de e-mails y otros datos. La búsqueda nos devuelve cerca de unos 500 resultados, en su mayor parte hojas de cálculo de Excel con listas de e-mails, nombres completos de los propietarios de las cuentas e incluso sus cargos. Algunas hojas, como la que vemos debajo, nos muestran incluso a qué institución y a qué servicio concreto de correo electrónico (webmail) corresponden las direcciones de la hoja.



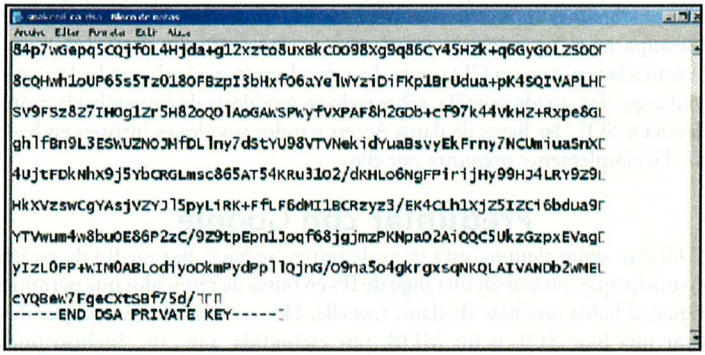
También se pueden encontrar archivos *csv* completos relativos a servicios de correo electrónico vía web que utilicen el código postal o el número del usuario como login, incluyendo la fecha de nacimiento y la dirección del usuario. Si no se consigue a primera vista, siempre es posible hacer esto utilizando el parámetro *residential phone e-mail address filetype:csv csv*.

Otra manera de conseguir e-mails, utilizada en este caso para conseguir muchos nombres de usuario de Hotmail/MSN Messenger, es usar la barra de búsquedas con el parámetro *filetype:ctt ctt messenger*. Ctt es la lista de direcciones de Hotmail/MSN que, para mayor facilidad del usuario, puede exportarse en formato XML. Lo malo es que también puede, por accidente, pasar a ser pública en caso de que se contenga en un directorio compartido de MSN.NET o del Messenger. Los nombres de otros servicios de e-mail y de proveedores también pueden devolver listas.

Certificados y llaves PGP

Yendo un poco más allá de conexiones, contraseñas y e-mails, se pueden conseguir incluso llaves PGP y certificados de autenticidad de servidores. Estos elementos en manos de un cracker con un mínimo de inteligencia pueden ser mucho más útiles que las listas de logins o de passwords: los logins del administrador aplicados a deshora, y desde IPs extrañas son siempre muy fácilmente detectables por el propio administrador, si éste tiene la buena costumbre de ojear los logs. Una llave PGP o un certificado digital pueden hacer que un cracker pase fácilmente por ser el dueño de la llave, o que una falsa web montada como trampa sea vista por los usuarios de la Web como si fuese la red corporativa que es la propietaria del certificado.

Y al igual que ocurre con Google, la llave PGP, que debería ser una herramienta de ayuda al usuario y especialmente, al administrador, se transforma en un arma letal en manos del cracker (como comentamos varias veces, por fallos en el sistema operativo, errores en los programas o falta de capacidad del propio administrador). Basta con hacer una búsqueda con el parámetro *BEGIN (CERTIFICATE\DSA\RSA) filetype:key* en Google para ver cuántas llaves privadas PGP consigues (incluso algunas de las más elaboradas, como las de 1.024 bits).



Servidores FTP

Los servidores FTP son especialmente vulnerables a los fallos de conexión. Los administradores parecen no preocuparse demasiado de este tipo de servidores: login anónimo, contraseñas débiles, contraseñas modelo activadas (por ejemplo, *login:ftp*, *seña:ftp*) son moneda común en la configuración típica. También es común ver servidores con versiones instalables o con errores conocidos pero no corregidos, como el *WS_FTP* (que durante mucho tiempo fue prácticamente un estándar para los usuarios de UNIX). Una búsqueda en Google con *filetype:ini +ws_ftp +pwd* nos mostrará todos los archivos de claves que se puedan recuperar de servidores FTP "particulares".

"HUSMEANDO" EN BASES DE DATOS MEDIANTE GOOGLE

Hay una modalidad de cracker que tiene, por las técnicas que utiliza, así como por los valores que puede poner en juego, marca una diferencia respecto a todos los demás: su capacidad de explotar las bases de datos. En rigor, el intruso de bases de datos se divide en diversas modalidades, más o menos correspondientes con las fases de invasión:

a) crackers que sólo invaden bases de datos, y borran lo que contienen,
b) crackers que secuestran las bases de datos y sólo las devuelven previo pago de un rescate,
c) crackers que sólo duplican los datos que encuentran, y que los venden en el mercado negro.

Un cracker de bases de datos es tan peligroso que su existencia ha terminado por generar, dentro del campo de Técnicos en informática, una nueva profesión: el *Database Security Administrator*, o en castellano normal, la persona que carga con la culpa, en el lugar del jefe de sistemas, en caso de que los datos de la empresa aparezcan a la venta en un CD pirata. La tarea de proteger las bases de datos no es, sin embargo, demasiado sencilla, sobre todo en servidores de comercio electrónico: los servicios SQL, las bases de datos Access y todos sus clones insisten en hacerse notar. Tú simplemente pregunta por ellos.

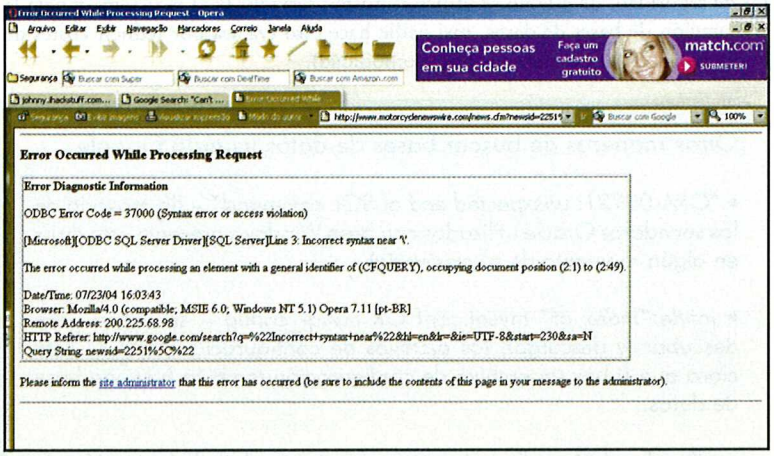
Preguntar con Google

Durante algún tiempo, esta tarea de rastreo se hacía por medio de escáneres como nmap, que sondeaban un rango de IPs en busca de resultados que permitiesen averiguar si había una base de datos tras ella. Hoy en día Google es capaz de encontrar una base SQL o un MDB, por escondida que esté. Incluso pueden encontrarse bases ocultas tras proxies y cortafuegos mal configurados. Permitir que una base de datos reciba peticiones desde todas partes, sin restringir la IP, equivale a convertir el servidor en un dominio público.

Para encontrar una base de datos, incluso aquellas que estén ocultas en una sencilla página HTML, puedes utilizar en Google el término *"Can't connect to local" intitle:warning*. La mayor parte de los resultados de esta búsqueda contienen bases de datos SQL/MySQL que tienen fallos de configuración o actualización que pueden ser explotados. De la misma manera, Google puede encontrar servidores IBM DB2 ejecutándose con vulnerabilidades o servicios defectuosos, usando el parámetro *"detected an internal error [IBM][CLI Driver][DB2/6000]"*.

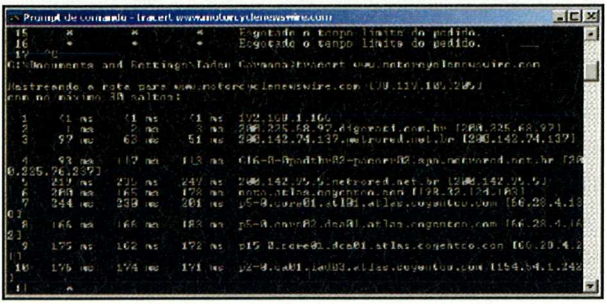
Pero el rey de las vulnerabilidades en Google es SQL. Una búsqueda por *"Incorrect syntax near"* puede revelar muchos más agujeros que Nessus o nmap:

ruta de la base de datos, nombres de las funciones, nombres de los archivos, detalles parciales de los códigos... todo esto es revelado por el simpático SQL:



Google también se lleva bastante bien con las bases de datos Oracle. Preguntando por ellas en Internet como el parámetro "ORA-00933:SQL command not properly ended", es posible encontrar varios. Si no funciona o no es suficiente, usa *"ORA-12541: TNS:no listener" intitle:"error occurred"* para encontrar otras bases de Oracle con errores de recepción de datos.

Si hay dudas sobre la localización de una dirección IP, sólo hay que utilizar herramientas sencillas como el tracert, una herramienta de redes de Windows, prima hermana del traceroute de Unix. Al encontrar una base de datos desprotegida en la página http://www.motorcyclenewswire.com/news.cfm?newsid=2251%5C%22, por ejemplo, podría ser necesario descubrir la IP del servidor en el que se aloja la base de datos. Acceder al Prompt de comando del Windows y escribir tracert www.motorcyclenewswire.com nos daría el siguiente resultado:



Que probablemente es el servidor en el que se aloja la base de datos. Pese a la recomendación de las empresas desarrolladora de bases de datos y de las consultoras de seguridad, de que no se utilice el mismo servidor para el sistema y para la elaboración de bases de datos, casi nadie hace esto (ya sea por reducir costes de equipos y licencias, o bien por pura comodidad).

Otras maneras de buscar bases de datos usando Google

- *"ORA-00921: unexpected end of SQL command"* – (la mayoría de los servidores Oracle utilizados con base Windows presenta este error en algún momento de su existencia).

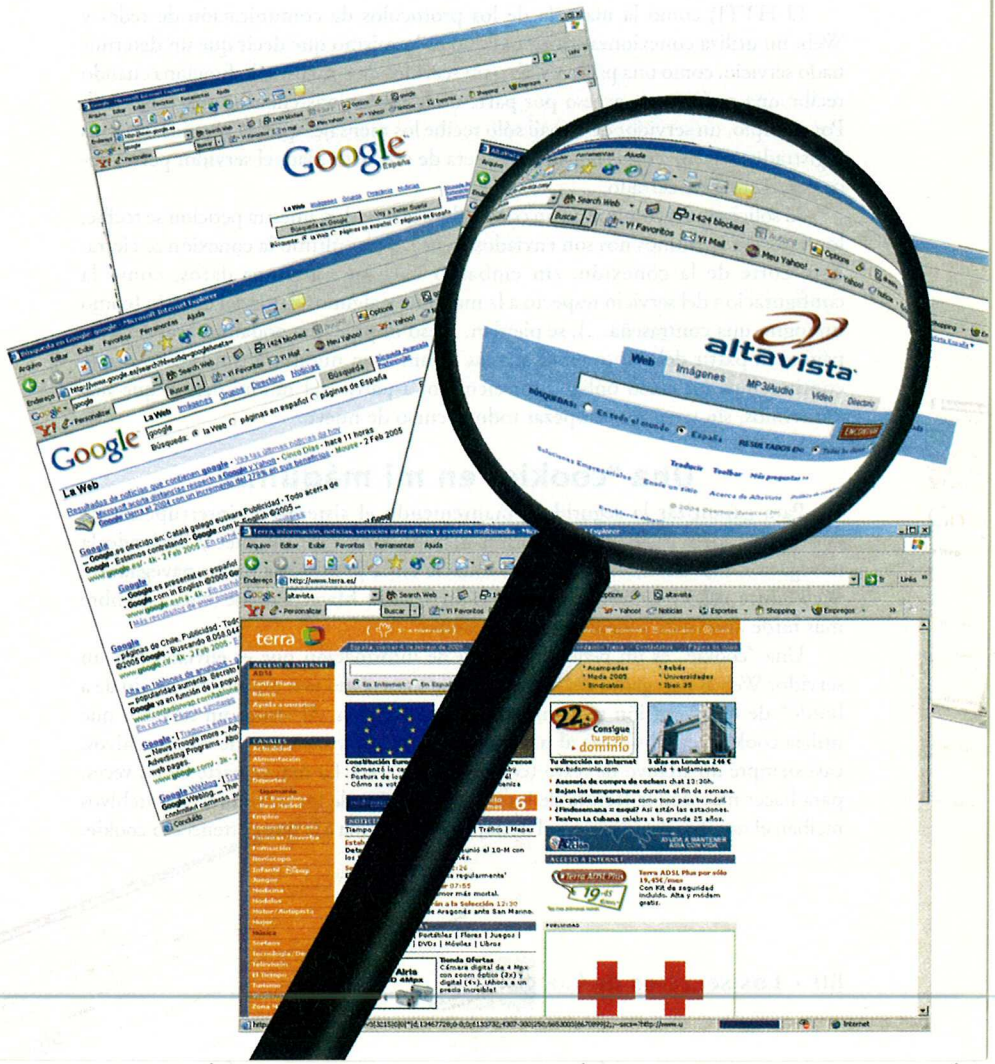
- *intitle:"index of" mysql.conf OR mysql_config* – se utiliza para descubrir (y descargar) los archivos de configuración MySQL. Está claro que si hay un archivo de configuración también hay una base de datos.

- *allinurl: admin mdb* – descubre directorios de administración de sistemas basados en Access, o pequeños sistemas SQL. De nuevo nos sirve la regla: donde hay un directorio administrativo, hay una base de datos a explotar.

- *filetype:inc intext:mysql_connect* – es tal vez la más peligrosa de las búsquedas que Google permite hacer: muestra scripts de logins de bases de datos MySQL, detallando su proceso de ejecución. En algunos casos puede devolver incluso el histórico de las últimas conexiones del administrador que se han hecho en el sistema, revelando los nombres de administradores, las IPs de sus estaciones de trabajo y su servidor, y contraseñas por defecto que no fueron tocadas tras el proceso de instalación.

Tercera Parte:

Todo lo que necesitas saber



CONOCIMIENTOS ESENCIALES

Las cookies te vigilan

Las cookies son literalmente galletas crujientes, hechas de harina de trigo, agua y leche (no las confundas con los *donuts*, esas rosquillas que son las preferidas de nueve de cada diez policías estadounidenses). Estas galletas, decimos, son muy quebradizas (no hay huevos en la receta) y sueltan migas por todas partes, y apenas si se pueden coger con las manos. Tal vez por esto, al ir implementando utilidades para HTTP, los ingenieros hayan querido homenajearlas, dándole el nombre de cookies a los pequeños archivos que va dejando en el ordenador.

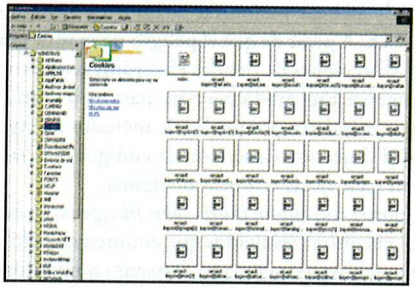
El HTTP, como la mayoría de los protocolos de comunicación de redes y Web, no utiliza conexiones continuas: eso es lo mismo que decir que un determinado servicio, como una página Web o un servidor de e-mails, sólo funciona cuando recibe una petición de acceso por parte de las máquinas-cliente conectadas a él. Por ejemplo, un servidor de e-mail sólo recibe los mensajes que los propios usuarios registrados envían a otras personas, y fuera de esta actividad, el servidor permanece prácticamente cerrado.

Al solicitar una comunicación con cualquier servidor, nuestra petición se recibe, los datos que pedimos nos son enviados (o no), y por último la conexión se cierra. Este corte de la conexión, sin embargo hace que algunos datos, como la configuración del servicio respecto a la máquina y algunos datos del usuario (como un login, una contraseña...), se pierdan. Y eso impide que podamos acceder a las páginas a partir del punto en el que las dejamos en nuestra última visita, o que continuemos un curso online, por ejemplo, a partir del módulo en el que nos detuvimos, sin tener que empezar todo el curso de nuevo.

Una "cookie" en mi máquina

Para garantizar la seguridad, manteniendo el sistema de interrupción de conexiones, pero facilitando al mismo tiempo el uso de Internet y haciendo la navegación más intuitiva, Nestcape, todavía en la prehistoria de los navegadores Web (entre 1995 y 1997), tomó la decisión de crear Magic Cookie (MC), nombre más tarde abreviado sencillamente en cookie.

Una "cookie" es un pequeño pedazo de información que es enviado por un servidor Web al navegador del usuario, y que sirve como una especie de "diario de a bordo" de la navegación por una determinada página. Al visitar un servidor que utiliza cookies, éste informa al navegador del cliente y crea uno de estos archivos, casi siempre un archivo de texto (con extensión .txt). La mayor parte de las veces, para hacer más fácil la lectura de los datos por parte de los servidores, los archivos reciben el nombre del usuario y el dominio de la página a la que pertenece la cookie.



Cuando el navegador está en un ordenador con más de un usuario registrado (sistemas Windows 2000/XP/2003 y UNIX/Linux con varios usuarios), las cookies pueden contener datos del usuario local, en vez de informaciones de login de la máquina. Esta modalidad de cookie se utiliza para discriminar entre los accesos de diferentes usuarios y, al menos en un primer nivel, para evitar el robo de contraseñas.

Independientemente del contexto en el que se use una cookie, suele contener pequeños pedazos de información, en ocasiones cifrados: el número de IP del usuario, su login, su clave, su último acceso al servicio y la cantidad de datos enviada en la última conexión, entre otros datos.

Por qué son peligrosas las cookies

Las cookies, en principio, deberían contener sólo datos de los sitios en los que se originan. De este modo, una cookie sería solamente un archivo de texto con algunos datos que relacionan a un usuario con una página, y nada más.

Pero si todo fuese así, no habría ningún motivo para estos dos grandes temores:

a) que una cookie lea informaciones de otras cookies contenidas en la máquina, o del árbol de directorios del sistema operativo,

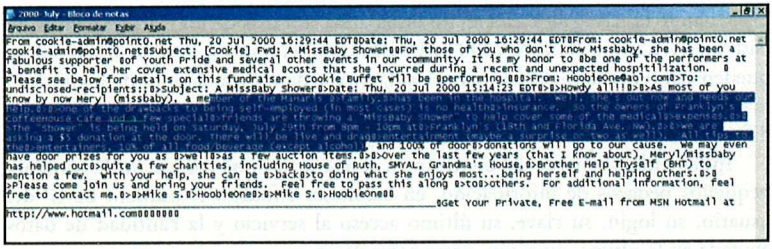
b) que una cookie sea utilizada para difundir virus, troyanos, o para espiar la máquina del usuario.

¿Por qué si no algunas cookies son identificadas por los antivirus como archivos malignos o *spywares*? Cuando ocurre esto, hay algo más que una cookie en tu máquina, generalmente será un programita en Java, o pequeños ejecutables. Su descarga suele producirse en secreto, aprovechando la fragilidad de las configuraciones de seguridad de Internet Explorer. En estos casos, la cookie sólo funciona para avisar al programa espía de que el usuario está online, para que sus datos puedan ser enviados al "maestro", que controla la conexión y programa.

Mientras el sistema esté infectado, algunas de estas cookies y de los programas que contengan pueden aprovecharse, incluso, de tus servicios de banda ancha y de servidores Windows que están conectados 24 horas al día, y hacerlos visibles en la Web, compartiendo el directorio de cookies. Esto tiene por objetivo el facilitar a otros usuarios que las capturen a su vez.

Google es capaz de rastrear y capturar cualquiera de estas cookies que hayan sido colocadas en la Web. A pesar de que el peligro no es muy grande, tener algunos nombres de usuario, direcciones de tiendas electrónicas o IPs, es tener todo lo que un cracker medianamente preparado necesita para usar como trampolín, ya que luego podrá conseguir más datos por medio de métodos de ingeniería social, *brute force* o técnicas de SQL Injection (inserción de códigos que hacen que las bases de datos SQL devuelvan secuencias erróneas de datos).

Al hacer una búsqueda con *admin inurl:cookie filetype:txt* en el campo de búsqueda de Google se pueden encontrar resultados que enumeren cookies "perdidas" en servidores de correo, en archivos o incluso en carpetas compartidas de MSN o ICQ.



En el caso superior, ha sido posible capturar una cookie que se encuentra en un servidor doméstico, con los datos de una conversación en grupo en una sesión del MSN Messenger. Observa que los e-mails de los usuarios aparecen junto a las indicaciones de sus apodos.

Cómo evitar una indigestión de cookies

Hay varias maneras de librase, o por lo menos de minimizar, los riesgos de la presencia de cookies en tu ordenador, sean o no peligrosas. Para los usuarios de Internet Explorer, lo ideal es acceder a menú **Herramientas > Opciones de Internet > Privacidad** y pinchar en la opción **Pedir Datos**. Al principio resulta un poco pesado, pero después de un cierto tiempo, además de la ventaja de que aumentará tu seguridad, esta opción revelará algunas cosas interesantes: notarás por ejemplo que los sitios que tienen mucha publicidad (como los sitios pornos y los de hackers) llegan a colocar 30 cookies en tu máquina en una sola sesión de navegación.

Otra iniciativa interesante es la de buscar en **Herramientas > Opciones de Internet > Seguridad** la opción **Desactivar descargas de archivos**. Así, applets de Java, archivos ejecutables o las famosas barras de herramientas coloridas que surgen de la nada serán bloqueadas pos el navegador antes incluso de que sean instaladas.

Para los más paranoicos, también se pueden bloquear las cookies, de modo que puedan ser leídas solamente desde un dominio concreto. Sólo hay que pinchar con el botón derecho del ratón sobre la cookie, hacer un clic en **Propiedades**, y después en Sólo Lectura. Hacer el mismo procedimiento en la carpeta de Cookies de un usuario convertirá toda la carpeta al modo Sólo Lectura, algo que impedirá que las nuevas cookies puedan añadirse a las ya existentes.

BÚSQUEDA DE DOMINIOS: ¿CÓMO SABERLO TODO SOBRE UN SITIO, SIN INVADIRLO?

Para que un sitio pueda ser publicado en la Internet, tiene que estar alojado en un servidor que a su vez cumpla ciertos requisitos. El primer paso para esto es, antes de nada, que tenga una dirección IP válida, y que no esté siendo utilizada por ninguna otra máquina de la Web. La dirección IP, como es notorio, se compone de cuatro secuencias, de tres cifras cada una (entre el 000 y el 255). La dirección IP sirve para determinar la categoría de la red a la que pertenece la máquina (de más extensa a menos: A, B, C y D).

A cada nombre de dominio le corresponde una dirección, y varias máquinas pueden compartir una misma dirección IP. Esto ocurre porque los servicios de tráfico de la Web no trabajan con nombres: como ya dijimos, sino que utilizan el lenguaje binario, y sólo usan unos y ceros.

Al escribir una dirección normal, con letras, en tu navegador, ésta se envía a un servidor de DNS, una red de enormes bases de datos que contienen una correspondencia mundial entre los nombres de los dominios y las direcciones IP a las que dirigen. Después de recibir una petición desde el dominio, el DNS la reenvía al servidor en que se aloja la página, que devuelve la solicitud al DNS, para que envíe al ordenador del usuario a la dirección correcta.

De este modo, y sin una dirección de DNS, no se pueden alojar páginas en un servidor, ni mantenerlas en el aire. Incluso los famosos crackers y demás fauna ilegal, con base en Rusia y otros países del este, necesitan estar enganchados a un servidor DNS activo, que apunte a la dirección de sus servidores, desde todo el mundo.

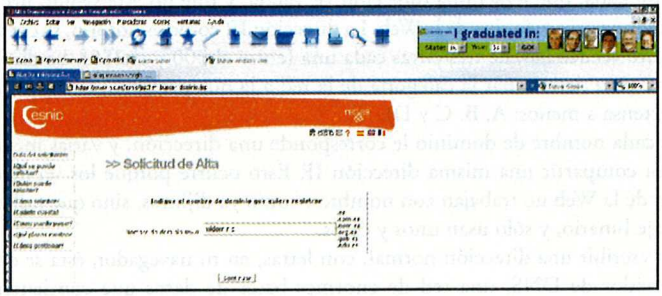
Los servidores DNS que sustentan Internet en todo el mundo están localizados junto a los grandes *backbones* (las grandes autopistas de datos de la Web), y utilizan a su vez servidores y equipos situados en algunas de las grandes universidades americanas. Estos servidores DNS son los responsables de que se puedan reconocer los dominios de primer nivel más conocidos, como .com, .org, .gob y .net, entre otros.

Como Internet se expande a demasiada velocidad, se ha hecho necesario el montaje de servidores DNS que traduzcan las direcciones de páginas y dominios que no se encuentren situados en los Estados Unidos. Por eso, cada país del mundo ha organizado su propio comité administrador de Internet, un directorio en que puedes registrar tu página o servidor (incluso los proveedores tienen que registrarse en estos dominios). En España, para registrar dominios con final en .es se necesita hacer una solicitud al ente RED.ES (adscrito al Ministerio de Industria, Turismo y Comercio).

Cómo registrar un dominio

Cualquier dirección del tipo *.es, .com.es, .nom.es, .org.es, .gob.es* y *.edu.es* debe pasar por los servidores de la RED.ES. Para registrar un dominio, éste no puede estar siendo utilizado por ningún otro usuario, y es obligatorio proporcionar unos cuantos datos personales y un e-mail para contacto. La página de registro de dominios es www.esnic.es y registrar un dominio por un año puede costarte entre 25 € y 200 € dependiendo del nivel del dominio que quieras tener.

Para poder registrar tu dominio, lo primero tienes que hacer es escribir su nombre en https://www.nic.es/esnic/jsp/frm_buscar_dominio.jsp



Pulsando en "Continuar" verás una página con los dominios que hay disponibles, y es ahí donde tienes que elegir el dominio que quieras, si está disponible, y donde debes seleccionar cómo quieres hacer el registro: a través de una empresa –el registrador-, o directamente desde esta web. En la siguiente página tendrás que realizar la autenticación del usuario. Si es la primera vez que accedes a este sitio, tendrás que registrarte en **Registrarse en ESNIC**. En ésta página te van a pedir varios datos personales. Después de proporcionarlos, sólo hay que hacer clic en **Crear contacto** para que tu solicitud se complete. Aparecerá una página en la que te anunciarán que tu usuario ha sido de alta, cuál es tu identificador de usuario, que tendrá el formato xxxx-ESNIC-F4 y te recordarán que tu clave de acceso llegará a la cuenta de correo electrónico que hayas indicado.

Aquí tenemos la primera brecha de seguridad, ya que cualquier persona puede registrar una identidad en ESNIC.ES con el nombre de otra, sea quien sea. Esto quiere decir que incluso, aunque no se consiga registrar ningún dominio en su nombre (veremos más adelante que tampoco es difícil) se puede bloquear su registro en RED.ES, haciendo que esa persona no pueda utilizar su propio nombre como dominio.

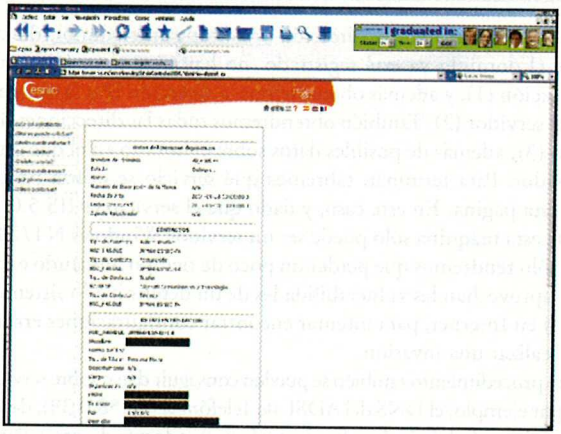
Al registrar un dominio definitivo, te van a pedir un CIF, para dominios de empresas o el NIF, para dominios personales. Este es el mayor impedimento que RED.ES ofrece a quienes registren una página en nombre de otra persona. Como hemos visto a lo largo de todo el libro, conseguir estos datos no es demasiado difícil, ya que basta con que sean datos auténticos (RED.ES utiliza un sistema de confirmación de que los registros de nombre y CIF/NIF coinciden), para que acepten un registro.

Todos los dominios que hayan sido registrados tienen que tener también un responsable técnico y un responsable de los cobros de la factura anual. Una vez más no hay demasiadas dificultades: quien puede conseguir una identidad falsa, puede conseguir tres.

Sitios desprotegidos, via ESNIC.ES

Sin embargo, la mayor parte de los usuarios son honestos, y registrarán un domino con su nombre auténtico y con una dirección verdadera, además de con un teléfono y un e-mail en los que se les pueda encontrar. Por desgracia, los servicios de administración de registros de dominio no almacenan estos datos personales de una manera demasiado restrictiva: es tremendamente sencillo conseguirlos, así como otras informaciones útiles sobre los dominios que se han registrado.

En la página inicial de Red.es hay un campo de búsqueda que te permite rastrear los dominios válidos. Si queremos saber, por ejemplo, si ya está registrado el dominio digerati.es, sólo tendremos que escribirlo en dicho campo de búsqueda, y hacer clic en el botón **Buscar**.

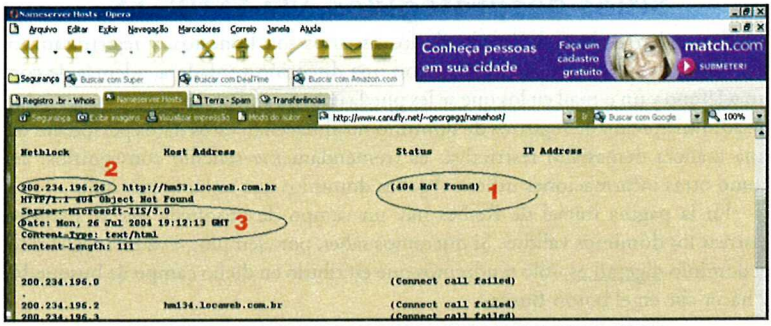


Con el resultado averiguamos no sólo que el dominio ya está ocupado, sino también quién es su propietario, cuál es su dirección, y cuál es el número de registro de la firma del dueño del dominio.

Madriguera virtual

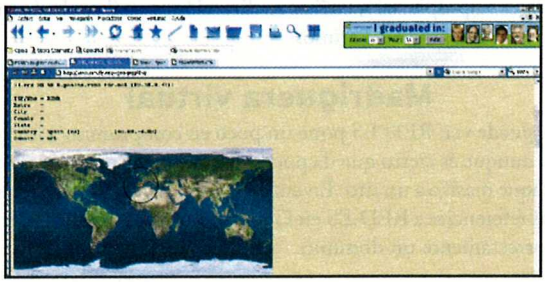
Como se puede ver, RED.ES pone un poco en compromiso a los propietarios de dominios, aunque es cierto que da pocos detalles técnicos para quienes quieran realizar un ataque masivo a un sito. En cuanto a Google, no merece la pena siquiera intentarlo: las referencias a RED.ES en Google sólo muestran cuál es la manera de configurar correctamente un dominio.

Pero existen otras herramientas que pueden mostrar prácticamente todo sobre un sitio, dominio o servidor de DNS. En la dirección http://www.canufly.net/~georgegg/namehost/ podrás encontrar una herramienta llamada Nameserver Host Utility, que sirve para escanear cualquier dirección de la Web, y con la que puedes buscar incluso si hay e-mails activos en el servidor DNS y en los servidores DNS secudanarios.



Buscando, por ejemplo, la dirección www.topgames.com.br, descubriremos que aunque el dominio ya está registrado, no hay ninguna página configurada para esa dirección (1), y además obtendremos la dirección IP a la que corresponde la página del servidor (2). También obtendremos todas las direcciones del servidor justo debajo (3), además de posibles datos sobre el horario en el que está configurado el servidor. Para terminar, sabremos qué servicio se utiliza para colgar las páginas en esta página. En este caso, y dado que el servicio es IIS 5.0, podremos concluir que esta máquina sólo puede ser un servidor Windows NT/2000. A partir de aquí, sólo tendremos que perder un poco de tiempo buscando exploits (programas que aprovechan las vulnerabilidades de un determinado sistema operativo o aplicación) en Internet, para intentar encontrar configuraciones erróneas en ese servidor, y realizar una invasión.

Con este procedimiento también se pueden conseguir datos sobre servidores DNS. Buscando, por ejemplo, el DNS del ADSL de Telefónica (80.58.0.33), descubriremos en qué dominio está alojado (pooles.rima-tde.net), además de otros datos y con la ayuda de un mapa, podremos saber incluso cuál es su localización geográfica.



LAS MIL CARAS DE GOOGLE

Hay mil maneras de aparecer en Google. En realidad, para ser parte de cualquier búsqueda de Google, incluso figurando muy abajo en el ranking, sólo hay que hacer una página con tu nombre y colocarla en la Web, y hacer que tenga al menos una visita, aunque sea la tuya propia.

Al dar a conocer tu página a amigos o a listas de discusión, te harás visible, y será casi imposible permanecer oculto. Tener muchos enlaces significa tener más oportunidades de que te encuentren, cuanto más te encuentren, más oportunidades tendrás a su vez de aparecer en mejores posiciones en el ranking de Google.

Lo que casi nadie sabe es que el propio Google tiene múltiples fachadas, y que cada una de ellas responde a una determinada manera de buscar, o a determinado tema. Y aunque no todas ellas dispongan de versiones en español, es bueno conocerlas, ya que algunas búsquedas muy específicas sólo son posibles gracias a la ayuda de estos caminos alternativos de Google.

Google Grupos

(http://groups.google.es/grphp?hl=es&tab=wg&q=)

Con una interfaz en español que casi no se diferencia de su versión en inglés, Google Grupos reúne prácticamente todo lo que se ha escrito en la historia de los grupos de discusión (newsgroups) de Internet desde 1995. Estos grupos de discusión (USENET) fueron en su día el último grito en una Web que estaba naciendo, mucho antes de que llegasen los navegadores gráficos, como Netscape y más tarde Internet Explorer: muchas comunidades científicas, médicas e incluso aquellas relacionadas con desarrollo de la propia Web, además de otros seis servicios distintos, nacieron de los grupos de USENET. Google compró en el 2003 los servidores de Deja.com, que era la depositaria de buena parte de la información que se había enviado a grupos de USENET.

Hay categorías en función de cada asunto, divididas en grupos. Algunos grupos, como .alt, que está formado por servidores alternativos, contienen a su vez más de 2471 subgrupos. En la interfaz de Google Grupos se puede visualizar la actividad de cada grupo con barras verdes que verás a la izquierda: cuanto más verde sea una barra, más activo será ese grupo.

Google Directorio

(http://www.google.es/dirhp?hl=es&tab=wd&q=)

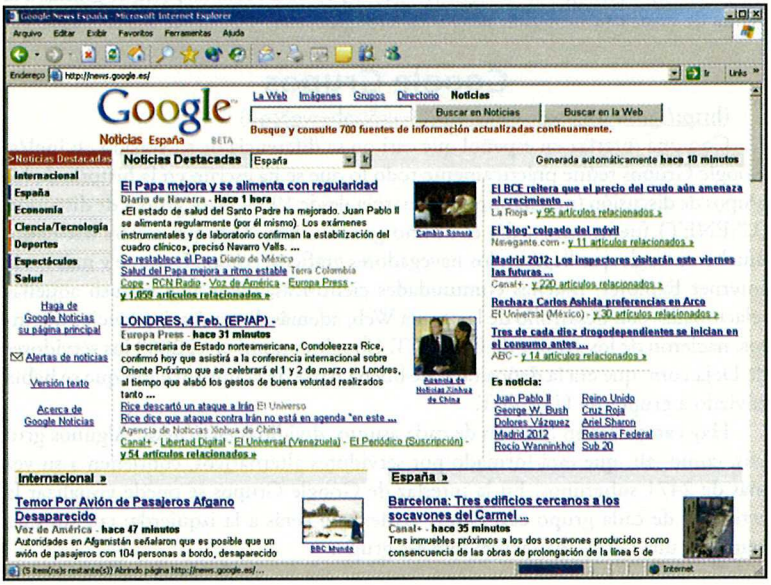
El Directorio de Google es una respuesta a algunas de las críticas que surgieron respecto a la forma que tiene Google de organizar la información, o según los partidarios de aquellos servicios organizados como un directorio (o sea, Yahoo), es desordenado. Aunque es muy parecido al servicio de Yahoo, el Directorio de Google organiza todo el contenido de Internet en asuntos y ramas, tratándolos como si cada uno de ellos fuese un directorio, dentro de un gran árbol. Así, al buscar en el directorio **Games**, podremos ver opciones como **Board Games** (Juegos de tablero) y **Electronic**

Games (Juegos electrónicos). En **Board Games**, podemos encontrar a su vez juegos de detectives (**Murder-Mystery and Deduction**), y dentro de los juegos de detective, podemos encontrar entre otros directorios, como **Chinese Detective**, un juego de estrategia ambientado en el Pequín Imperial. El resultado de esta operación es como pelar cada una de las capas de una cebolla: encontramos el enlace de un fabricante de juegos de mesa de misterio, que fabrica un juego de detective chino.

Noticias Google

(http://news.google.es/nwshp?hl=es&gl=es)

El servicio Noticias Google es otra de las estrellas del buscador. Al pedir un término en Google Noticias, nuestra búsqueda se hará sólo en servicios de noticias y podremos ver, junto a los resultados, la indicación del horario en el que esta noticia ha sido colgada en la Web, y por tanto, cómo es de reciente.



Froogle

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

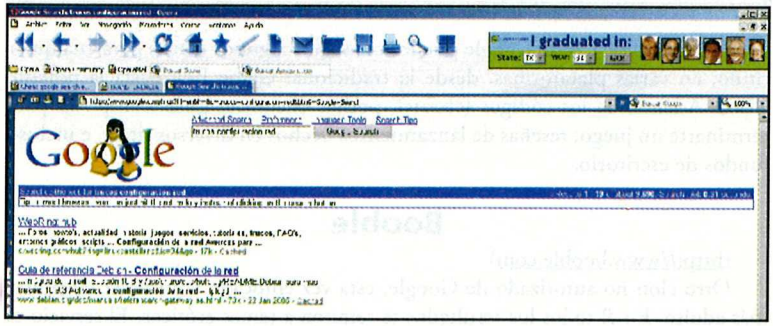
Si hay un lugar donde se puede vender una cosa, ese lugar es Internet. Fuera de las tiendas 100% virtuales, la mayoría de las cuales se hundieron con la crisis de las empresas .COM, ofrecer servicios y mercancías en un nicho de mercado muy específico a través de Internet suele tener éxito. Incluso quien nunca haya comprado nada a través de Internet puede que, al menos una vez, haya visitado sitios de subastas,

como ebay.es, para tener una idea de cuánto cuesta un objeto determinado. Froogle hace más fácil la vida a los que buscan precios por la Web, y también a los adictos a las compras. Una búsqueda de Linux, por ejemplo, nos devuelve más de 81.000 resultados, entre CDs de distribuciones, manuales, libros y cursos.

Google Linux

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

Como su propio nombre indica, es un directorio de Google enfocado únicamente a la búsqueda de asuntos relacionados con el mundo Linux. Pese a que sólo tiene versión en inglés, sus resultados pueden incluir traen varias páginas en español. Cuando buscamos *trucos de configuración de red*, obtendremos más de 9.000 resu[illegible].



Google Microsoft

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

El Google Microsoft hace búsquedas en todas las páginas que estén relacionadas con la empresa de Bill Gates. Al igual que su correlativo para Linux, Google Microsoft es una buena salida para buscar, no sólo trucos, sino también unidades para dispositivos específicos. En una búsqueda de unidades para Windows, más concretamente, de dispositivos USB de la SiS, encontramos más de 89.400 resultados.

Cheat Google

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

Esto ya es terreno de aquellos servicios que no están exactamente ligados a Google, pero que se inspiran en él para ofrecer ciertos servicios. Este sitio, concretamente, es un verdadero objeto de deseo para los jugadores patológicos: aquí podras encontrar la manera de saltarte ciertas fases difíciles, de ganar mucho dinero, o evitar que puedan herirte o matarte. La lástima es que esto sólo sirva para el mundo de los joysticks y los videojuegos, pero al menos ya es algo...

Por medio de Cheat Google puedes conseguir algunos trucos para cualquier título, en varias plataformas, desde la tradicional Game Boy hasta la potente XBox. Además de los códigos (cheats), podrás encontrar guías completas para terminarte un juego, reseñas de lanzamientos hechos en diversos sitios, e incluso fondos de escritorio.

Booble

(http://www.booble.com)

Otro clon no autorizado de Google, esta vez enfocado a un tema, digamos, más adulto. En él todos los resultados te remiten a temas eróticos. El servicio es gratuito, pero el 90% de los resultados no lo es.

CUIDADO CON LOS ENLACES FALSOS

Las variantes de delitos en el mundo digital se están haciendo cada vez más frecuentes. El fraude online, sobre todo la "pesca" virtual de claves sin cambiar, de usuarios despistados, o cuentas que están situadas en redes mal configuradas o sin una política de seguridad definida, es constante.

Pese a que, como ya hemos visto, Google puede ser utilizado como un arma por los pescadores de contraseñas, de mil maneras distintas, también podemos usarlo como una herramienta de seguridad. Antes de nada, sin embargo, vamos a ver cómo funcionan los enlaces falsos y los e-mails de los pescadores de contraseñas.

Falsa generosidad

Los e-mails falsos usan nombres de empresas, programas de televisión y, desde comienzos del año pasado, nombres de instituciones bancarias (incluso se ha detectado un e-mail que utilizaba el nombre del Banco de España para intentar robar datos personales del usuario.

En el primer caso, los mensajes ofrecen productos a precios muy jugosos, o directamente premios. Los e-mails que utilizan nombres de bancos ofrecen un seguro gratis, premios en dinero, tarjetas de crédito gratis con límites altos o, en maniobras de ingeniería social, piden que los usuarios hagan un registro de sus cuentas corrientes, libretas de ahorro o que envíen sus datos de registro hacia una determinada página, dirección de e-mail, o que rellenen un formulario que va adjunto al mensaje, y en el que entre otras cosas hay que declarar el saldo de la cuenta corriente, así como contraseñas de usuarios. Ésta es la primera señal para desconfiar: incluso aunque los e-mails vengan con direcciones que correspondan al correo electrónico de un banco, las instituciones financieras tienen reglas muy estrictas sobre el manejo de contraseñas, y no necesitan enviar un e-mail para saber cuánto guarda cada usuario en su cuenta corriente.

Además de eso, también es interesante realizar una investigación del e-mail. Si tienes Outlook, pincha con el botón derecho del ratón sobre este mensaje y selecciona **Propiedades**. Acto seguido, observa el bloque de **Encabezados de Internet**, en especial la línea *Message-Id*. Anota el final de la ID del mensaje, como por ejemplo, @linux.com. Después, utiliza la línea de comando y escribe tracert, seguido del nombre del destinatario del e-mail y del servidor que hayas encontrado. En nuestro caso usaríamos:

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

Observa si una de las últimas direcciones se corresponde con el nombre del banco o con la empresa que supuestamente te ha enviado el correo. Si no se parece siquiera a un servidor del banco, habrás estado a punto de caer en un *phising*.

Enlaces falsos y trampas

Imagina que recibes un e-mail de un banco imaginario, el Español, y que el enlace apunta a espanol1.com, o está situado en el dominio @espanol1.com. Pero observa: el dominio que es conocido del banco es www.espanol.com.es. Si tienes alguna duda sobre este enlace, vete hasta Google y ejecuta una búsqueda.

¿Sorpresa?: ¡Es el mismo sitio, el del Banco! Ahora presta más atención: el logotipo del banco y las indicaciones de sus servicios aparecen en el centro de una página en blanco, o flotando junto a un reborde que intenta imitar el color del banco, por ejemplo.

Este sitio es falso. El cracker utilizó una vulnerabilidad de las bases de datos SQL y de los servicios de ISS (el servidor Web de Microsoft), para hacer que la página de inicio del banco redirija a la página falsa. No te vamos a enseñar, obviamente, cómo se hace esto, pero créenos: la técnica es increíblemente sencilla.

Si el nombre que aparece en el e-mail es el nombre verdadero de la institución o del dominio que utiliza, haz una comprobación más: instala otro navegador que no sea Internet Explorer (mejor si es Mozilla Firefox u Opera) en tu máquina, y copia el enlace que te mandan en el e-mail en la barra de direcciones del navegador. Observa si la dirección, en la barra de de estado, debajo de la pantalla de la página coincide o no con la del enlace que has copiado y pegado arriba. Si no lo hace, estás ante de una página falsa.

Por último, desconfía de los enlaces que terminen con la extensión .exe, .com o .bat. porque contienen archivos ejecutables o programas que, cuando son ejecutados, pueden realizar descargas muy rápidas e infectar tu máquina con troyanos o Keylogers (programas que registran todo lo que el usuario escribe y lo envían después a una determinada dirección de e-mail o FTP). Para estar tranquilo, prueba a copiar la dirección de la página en la barra de búsqueda de Google e intenta hacer la descarga del archivo para el disco duro sin abrirlo (de nuevo, usa Opera o Mozilla, en lugar de Internet Explorer). Por último, intenta hacer que un antivirus examine el archivo.

DATOS PERSONALES VÍA ICQ Y MSN

Después del Blog, y más recientemente del Orkut (la red de relaciones que ha creado Google), los *instant messengers* fueron, si no una de la mayores revoluciones del Internet, al menos la "fiebre" que más usuarios de todos los tipos y lugares del planeta ha enganchado. Sobre todo en los organismos con muchas sedes, y para aquellas personas que tienen familiares y amigos viviendo en otros países, estos programas de mensajes han llegado a sustituir, y con ventajas obvias, incluso a los propios teléfonos, ya que programas como el MSN y el ICQ llevan recursos de voz, que además han mejorado mucho con el uso de los servicios de banda ancha.

Datos personales

Y son estos, precisamente, los *instant messengers* más populares: el ICQ y el MSN. ICQ es prácticamente el estándar de este tipo de comunicación en la Web (el propio AIM de AOL, aunque no lo quieran admitir, era una copia de ICQ, y se parece aún más desde que comprase a la empresa israelí que lo había desarrollado). Por otro lado, el MSN viene robándole terreno desde hace tiempo, sobre todo con las dos últimas versiones del programa, que permiten de compartir canciones e imágenes (y eso sin contar con la tradicional y millonaria campaña de marketing promovida por Microsoft para que las personas instalen sus productos).

Al hacer tu registro en cualquiera de estos servicios tienes que rellenar, obligatoriamente, una página con datos personales. Por ejemplo, si quieres utilizar los servicios del MSN, tendrás que bajarte el MSN Messenger (http://messenger.msn.es/). Además de eso, deberías tener una cuenta en Hotmail, en cuyo formulario hay que escribir algunos datos personales.

Pero ésta es la única similitud entre ICQ y MSN. Los datos personales que se entregan a Hotmail no son accesibles mediante el MSN Messenger, a no ser que crees un perfil que quieras que sea visto por otros usuarios. Si creas este perfil, basta con que alguien haga clic sobre el botón derecho del "muñequito" del MSN que te representa, para que le redirijan a una página del MSN, con datos sobre ti.

Está claro que lo ideal es que no rellenes este perfil (a pesar de que Google no devuelve los resultados que hay guardados en el directorio de Hotmail, es obvio que eso no tiene por qué tardar mucho en ocurrir). Además, el nombre que elijas para mostrar es el que verán los demás usuarios. Así que si escribes tu nombre verdadero en la identificación del MSN, será ese el que se muestre.

La situación de ICQ es un poco más compleja. Todos pueden saber tu nombre si eres lo suficientemente distraído como para colocarlo en las opciones de visualización. Además de eso, tanto en la dirección oficial de ICQ (http://www.icq.com), como en el sistema de búsqueda del propio programa, puedes encontrar a cualquier persona (incluso aunque no esté conectada) pinchando en el

botón **Search**. En este caso, la falta de sentido común es de nuevo la mejor ayuda para la inseguridad: los datos personales como tu dirección, tu teléfono, tu número de móvil y tu e-mail no son obligatorios. Y sin embargo, muchas personas rellenan esos campos, que pueden ser vistos por cualquiera con acceso a Internet y un poco de paciencia.

Cerrar ICQ

Otra brecha de seguridad de ICQ, en la que la mayoría de los usuarios "cae", es dejar la cuenta abierta para que cualquier persona del mundo pueda charlar contigo, incluso aunque no quieras hablar con ella. Esto puede resolverse fácilmente. Haz clic en el menú **Security & Privacy**, en la pestaña **Security**, y en **Change contact authorization**, cambia la configuración que hay por defecto y pon *My authorization is required*. Así, no podrás ser añadido a ninguna lista sin tu autorización previa.

Listas de usuario y mensajes guardados

Hacer copias de seguridad es una práctica sana, pero ciertas cosas no deberían guardarse nunca (o por lo menos, deberían guardarse bajo siete llaves). En ICQ, por ejemplo, puedes almacenar toda la lista de contactos, además de todas las charlas que has realizado, con sólo copiar la carpeta **NewDB**, que se encuentra en la carpeta de ICQ, del directorio **Archivos de Programa**, en otro lugar. El problema es que la mayor parte de los usuarios copia esta lista en su servidor FTP o en su website personal, esperando para recogerla después. Ahora sólo hay que buscar en Google el término *newDB* para ver cuantos de estos archivos están a tu disposición.

De la misma manera, también se pueden leer todas las charlas que se han hecho con ICQ: para ello sólo tienes que hacer clic con el botón derecho del ratón en el botón de ICQ y pinchar luego en **Message Archive**. Estas conversaciones se guardan en pequeños archivos de texto, también en el subdirectorio de ICQ.

Dejar estos archivos guardados en Internet o en carpetas compartidas en tu ordenador equivale a publicar tu lista de contactos y tus conversaciones personales o profesionales. Así que, para evitar esto, guarda las copias de seguridad de tu lista y tus conversaciones (si es que realmente necesitas guardarlas) en un disquete o en un CD regrabable. Además del riesgo que supone Internet, cualquier usuario de ICQ mal intencionado puede, con un programa de ataque *brute force* específico para ICQ, invadir tu ICQ online y llegar a robar tanto tu lista de usuarios como tus registros de conversaciones guardadas.

Los secretos de Google

¡Descubre todas las opciones ocultas de este potente buscador!

Hacking con Google

- Guía para hacer búsquedas avanzadas
- Entrar en Google desde móviles y PDAs
- Buscar imágenes y usar las herramientas del idioma
- Usar los caracteres comodín
- Poner tu sitio en el primer puesto en Google

Y mucho más...

Investigación en la Web

- Descubrir datos personales
- Acceder a bases de datos usando Google
- Localizar tus cartas y entregas
- Buscar datos sobre los dueños de sitios
- Consulta los datos del listín telefónico
- Analizar logs vía Google
- Hacer un Password Generator con Google

Y además:

- Cómo saberlo todo sobre un sitio, sin necesidad de invadirlo
- Descubrir y evitar enlaces falsos
- Cómo encontrar datos personales usando ICQ y MSN

www.digerati.es

Año 01 – Número 5 – 3,90 €

ISBN 858953543-6

9 788589 535434 00005