

Hack x Crack

Hacking con buscadores
(Google, Bing, Shodan)



Hack X Crack

Diseño hecho por: Hacker Fashion

www.hackxcrack.es

Cuaderno creado por Stakewinner00

Índice

- 1 - ¿Que buscador es mejor?
- 2 - ROBOTS.TXT
- 3 - Técnicas de Hacking con Buscadores
 - 3.1 - PreIndexación
 - 3.2 - PosIndexación
- 4 - Hacking con Google
 - 4.1 - ¿Qué es Google?
 - 4.2 - Operadores lógicos con Google
 - 4.3 - Operadores en Google
 - 4.4 - Ventajas de Google
 - 4.5 - Google Dorks
- 5 - Hacking con Bing
 - 5.1 - Operadores en Bing
 - 5.2 - Ventajas de Bing
 - 5.3 - Bing Dorks
- 6 - Hacking con Shodan
 - 6.1 - ¿Qué es Shodan?
 - 6.2 - Operadores en Shodan
 - 6.3 - Ejemplos de dorks en Shodan
- 7 - Programas de recopilación de información
 - 7.1 - ¿Para que sirven?
 - 7.2 - Foca
 - 7.3 - SearchDiggity
- 8 - Otros usos para los buscadores
 - 8.1 - Utilización de Google como proxy
 - 8.2 - BlackSeo
 - 8.2.1 ¿Qué es?
 - 8.2.2 Engañando a los buscadores
 - 8.2.3 GoogleBomb
 - 8.2.4 Spamindexing
 - 8.2.5 Spam

1 - ¿Que buscador es mejor?

No hay ningún buscador que sea mejor que otro simplemente son buscadores con diferentes características las cuales aprovechamos los hackers para investigar y buscar vulnerabilidades en diferentes webs.

Existen decenas de buscadores distintos pero solo nos vamos a centrar en los mas 4 mas importantes y de más uso, los conocimientos aquí aprendidos también son aplicables a otros buscadores con mínimas diferencias .

Existen programas especializados en realizar búsquedas automatizadas en buscadores para conseguir archivos de los cuales se puedan extraer meta-datos que se podrían utilizarse en un una futura intrusión. Esto lo veremos más adelante.

2 – ROBOTS.TXT

Muchas webs tienen un archivo llamado “robots.txt” este archivo ayuda a las webs a evitar que los buscadores indexen directorios o otro contenido que no tendríamos que ver. Como veremos más adelante hay buscadores que indexan contenidos por mucho que la web no quiera y esté el archivo “robots.txt”.

Por ejemplo en la dirección de <http://www.google.com/robots.txt> encontramos un archivo en texto plano. Si nos fijamos en la parte principal hay el esto

```
User-agent: *  
Disallow: /search  
Disallow: /sdch
```

el **User-agent: *** indica que este archivo sea reconocido para todos los buscadores. Después tenemos **Disallow: /search** el disallow evita la indexación de una carpeta o archivo, en este caso no indexara los contenidos del directorio search ni sdch.

Este es el típico archivo “robots.txt” con el que nos encontraremos la mayoría de veces.

Como vemos el archivo “robots.txt” también es una fuente de información ya que si lo bloquean sera por que hay algo interesante ¿no?.

3 – Técnicas de Hacking con Buscadores

Con ayuda de los buscadores podríamos crear dos categorías generales de ataques, los que se dan antes de que una web sea indexada y los que se dan después.

1) PRE INDEXACIÓN

En este caso el atacante tiene que descubrir una vulnerabilidad tipo RFI que sea explotable con una sola petición inyectando el exploit en el parámetro GET. Luego de descubrir la vulnerabilidad el atacante es listo y no quiere dejar huellas de su ataque.

Como no quiere dejar huellas le pide a su amigo Google o su otro amigo Bing que hagan ellos le trabajo sucio.

El atacante para completar la intrusión simplemente tendrá que pedir a sus amigos que indexen una URL maliciosa, y al visitar la página para poder indexarla en condiciones cometerán el ataque a la web.

Luego si la policía va a investigar la web atacada la policía comenzaría a mirar por quien cometi6 el ataque y verían que los que cometieron el ataque fueron sus amigos, mientras la policía esta entretenida buscando un nuevo sospechoso el atacante ha podido atrasar un poco su captura.

Si bien que en este caso el archivo "robots.txt" seria una buena medida de protección los buscadores no siempre cumplen al pie de la letra el archivo "robots.txt"

A parte de este ejemplo esta técnica se puede usar en ataques XSS, con lo que podríamos dejar un sitio inaccesible o hacer phishing gracias a Google y otras muchas cosas interesantes.

2) POST INDEXACIÓN

El segundo tipo de ataques son los que se dan después de indexar alguna web, aquí entrarían todos esos dorks que sirven para buscar mensajes de error y que nos aportan información.

Esta parte no la voy a explicar ya que más adelante explicaremos los dorks de Google y ejemplos de como podemos usar Google y Bing para recopilar información.

4 - Hacking con Google

4.1 - ¿Qué es Google?

Google es un buscador muy usado. XD

4.2 - Operadores l6gicos en Google

Al igual que los lenguajes de programación con Google se pueden utilizar operadores l6gicos para poder realizar búsquedas más certeras.

Por ejemplo si utilizamos el operador (-) se suprimirá ese argumento de la búsqueda. Ej: **hacking -etico** que como resultado devolvería una búsqueda sin el termino ético.

Otro operador importante es el or o (|) este operador hace que en una búsqueda haya más de un termino.

Ej: **Hacking | Hacker** como resultado obtendríamos una búsqueda con el termino hacker o hacking.

4.3 - Operadores en Google

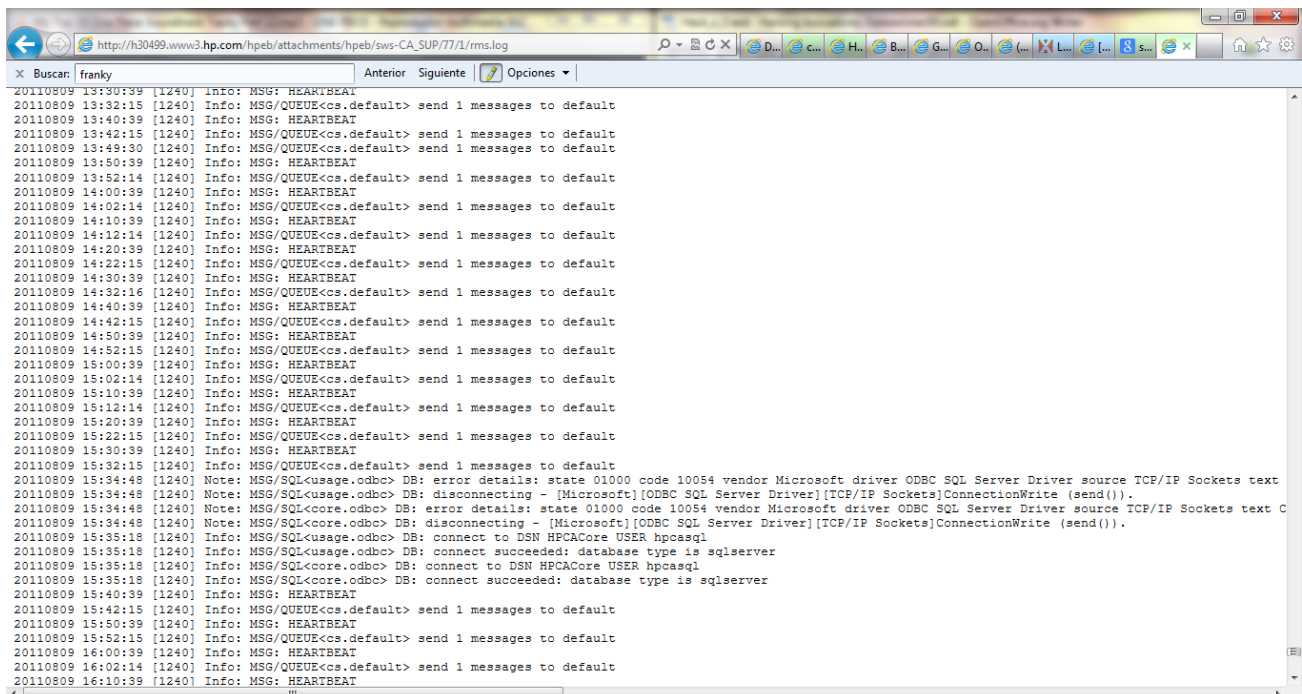
Google también nos da la posibilidad de utilizar ciertos operadores que son los que realmente nos interesan y que luego utilizaremos.

Uno de ellos muy usado es **site:dominio.com** este operador filtra la búsqueda a un dominio concreto, esto se puede utilizar por ejemplo si queremos auditar un dominio en concreto y no queremos que salgan más dominios podríamos poner "site:google.com".

Filetype:extensión es otro muy importante, su función es filtrar la búsqueda para que solo nos salgan archivos con esa extensión, esto puede resultar de utilidad si queremos comprobar si una web tiene guardado algún archivo log del firewall o alguna hoja de excel con nombres y passwords.

Por ejemplo si solo queremos ver archivos log de un sitio en concreto podríamos buscar por **site:hp.com filetype:log**

Un ejemplo de uno de esos archivos.



```
20110809 13:30:39 [1240] Info: MSG: HEARTBEAT
20110809 13:32:15 [1240] Info: MSG/QUEUE<cs.default> send 1 messages to default
20110809 13:40:39 [1240] Info: MSG: HEARTBEAT
20110809 13:42:15 [1240] Info: MSG/QUEUE<cs.default> send 1 messages to default
20110809 13:49:30 [1240] Info: MSG/QUEUE<cs.default> send 1 messages to default
20110809 13:50:39 [1240] Info: MSG: HEARTBEAT
20110809 13:52:14 [1240] Info: MSG/QUEUE<cs.default> send 1 messages to default
20110809 14:00:39 [1240] Info: MSG: HEARTBEAT
20110809 14:02:14 [1240] Info: MSG/QUEUE<cs.default> send 1 messages to default
20110809 14:10:39 [1240] Info: MSG: HEARTBEAT
20110809 14:12:14 [1240] Info: MSG/QUEUE<cs.default> send 1 messages to default
20110809 14:20:39 [1240] Info: MSG: HEARTBEAT
20110809 14:22:15 [1240] Info: MSG/QUEUE<cs.default> send 1 messages to default
20110809 14:30:39 [1240] Info: MSG: HEARTBEAT
20110809 14:32:16 [1240] Info: MSG/QUEUE<cs.default> send 1 messages to default
20110809 14:40:39 [1240] Info: MSG: HEARTBEAT
20110809 14:42:15 [1240] Info: MSG/QUEUE<cs.default> send 1 messages to default
20110809 14:50:39 [1240] Info: MSG: HEARTBEAT
20110809 14:52:15 [1240] Info: MSG/QUEUE<cs.default> send 1 messages to default
20110809 15:00:39 [1240] Info: MSG: HEARTBEAT
20110809 15:02:14 [1240] Info: MSG/QUEUE<cs.default> send 1 messages to default
20110809 15:10:39 [1240] Info: MSG: HEARTBEAT
20110809 15:12:14 [1240] Info: MSG/QUEUE<cs.default> send 1 messages to default
20110809 15:20:39 [1240] Info: MSG: HEARTBEAT
20110809 15:22:15 [1240] Info: MSG/QUEUE<cs.default> send 1 messages to default
20110809 15:30:39 [1240] Info: MSG: HEARTBEAT
20110809 15:32:15 [1240] Info: MSG/QUEUE<cs.default> send 1 messages to default
20110809 15:34:48 [1240] Note: MSG/SQL<usage.odbc> DB: error details: state 01000 code 10054 vendor Microsoft driver ODBC SQL Server Driver source TCP/IP Sockets text
20110809 15:34:48 [1240] Note: MSG/SQL<usage.odbc> DB: disconnecting - [Microsoft][ODBC SQL Server Driver][TCP/IP Sockets]ConnectionWrite (send()).
20110809 15:34:48 [1240] Note: MSG/SQL<core.odbc> DB: error details: state 01000 code 10054 vendor Microsoft driver ODBC SQL Server Driver source TCP/IP Sockets text
20110809 15:34:48 [1240] Note: MSG/SQL<core.odbc> DB: disconnecting - [Microsoft][ODBC SQL Server Driver][TCP/IP Sockets]ConnectionWrite (send()).
20110809 15:35:18 [1240] Info: MSG/SQL<usage.odbc> DB: connect to DSN HPCACore USER hpcaasql
20110809 15:35:18 [1240] Info: MSG/SQL<usage.odbc> DB: connect succeeded: database type is sqlserver
20110809 15:35:18 [1240] Info: MSG/SQL<core.odbc> DB: connect to DSN HPCACore USER hpcaasql
20110809 15:35:18 [1240] Info: MSG/SQL<core.odbc> DB: connect succeeded: database type is sqlserver
20110809 15:40:39 [1240] Info: MSG: HEARTBEAT
20110809 15:42:15 [1240] Info: MSG/QUEUE<cs.default> send 1 messages to default
20110809 15:50:39 [1240] Info: MSG: HEARTBEAT
20110809 15:52:15 [1240] Info: MSG/QUEUE<cs.default> send 1 messages to default
20110809 16:00:39 [1240] Info: MSG: HEARTBEAT
20110809 16:02:14 [1240] Info: MSG/QUEUE<cs.default> send 1 messages to default
20110809 16:10:39 [1240] Info: MSG: HEARTBEAT
```

NOTA

Para conseguir más buenos resultados es importante ir siempre a la última página, clicar en [repetir la búsqueda e incluir los resultados omitidos](#). Y después tienes que buscar archivos que tengan abajo un texto como el siguiente *No hay disponible una descripción de este resultado debido al archivo robots.txt de este sitio.* [Más información](#).

De esta forma encontrareis esos archivos que la empresa no quiere que veas. Por ejemplo en la screenshot anterior vemos como tuvieron un error con la base de datos.

Hay otros tipos de operadores pero son de menos uso.

LISTA DE OPERADORES DE GOOGLE

Allintext:texto → Este operador busca una cadena de texto dentro de una pagina web y no dentro de una URL. (No se puede utilizar junto a otros)

Allintitle:texto → Busca una cadena de texto solo dentro del titulo de una web. (No se puede utilizar junto a otros)

Intitle:texto → Busca una cadena de texto dentro del titulo de una web. (Se puede utilizar junto a otros)

Allinurl:texto → Busca una cadena de texto solo en la url. (No se puede utilizar junto a otros)

Inurl:texto → Busca una cadena de texto en la url. (Se puede utilizar junto a otros)

Author:texto → Busca artículos o noticias escritos por el nombre o la dirección de correo indicada. (Se puede utilizar junto a otros)

Cache:dominio.com → Con este operador accedemos a la web que Google tiene en su cache. Útil para cuando borraron un tema y no ha pasado mucho tiempo (No se puede utilizar junto a otros)

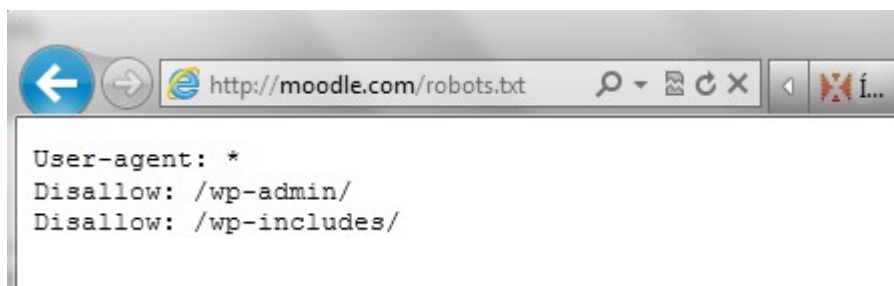
Link:dominio.com → Este operador se utiliza para buscar enlaces que apunten a un determinado sitio web. (No se puede utilizar junto a otros)

Related:dominio.com → Busca paginas relacionadas. (No se puede utilizar junto a otros)

4.4 – Ventajas de Google

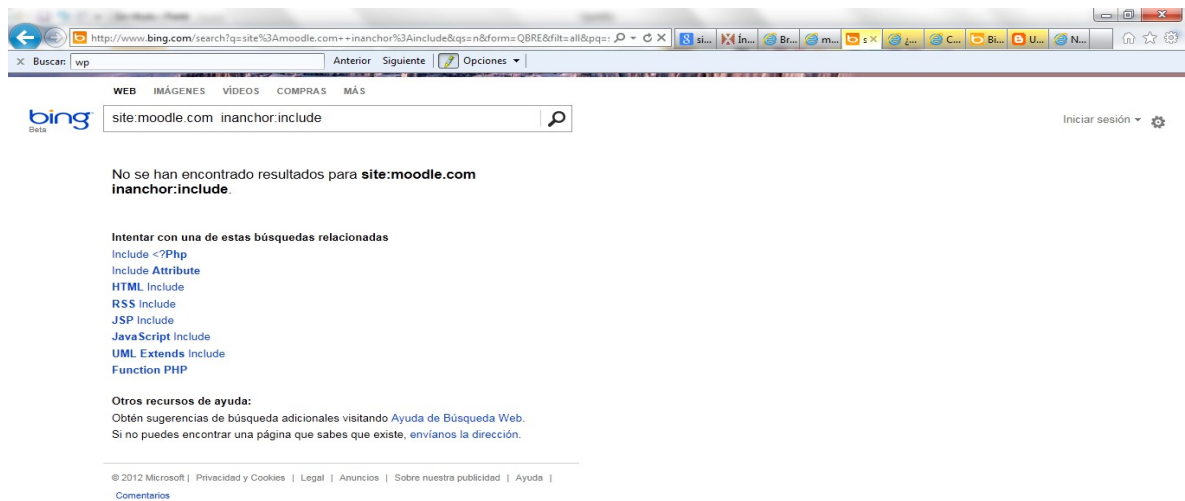
Después de este repaso vayamos ya al grano. A diferencia de Bing que si interpreta bien el archivo "robots.txt" Google no lo interpreta bien lo que significa que para los que nos interese hacer una auditoría Google jugara en nuestro favor a la hora de saltarse los archivos "robots.txt"

Como veis en las screen shoots el archivo "robots.txt" es bastante explícito.

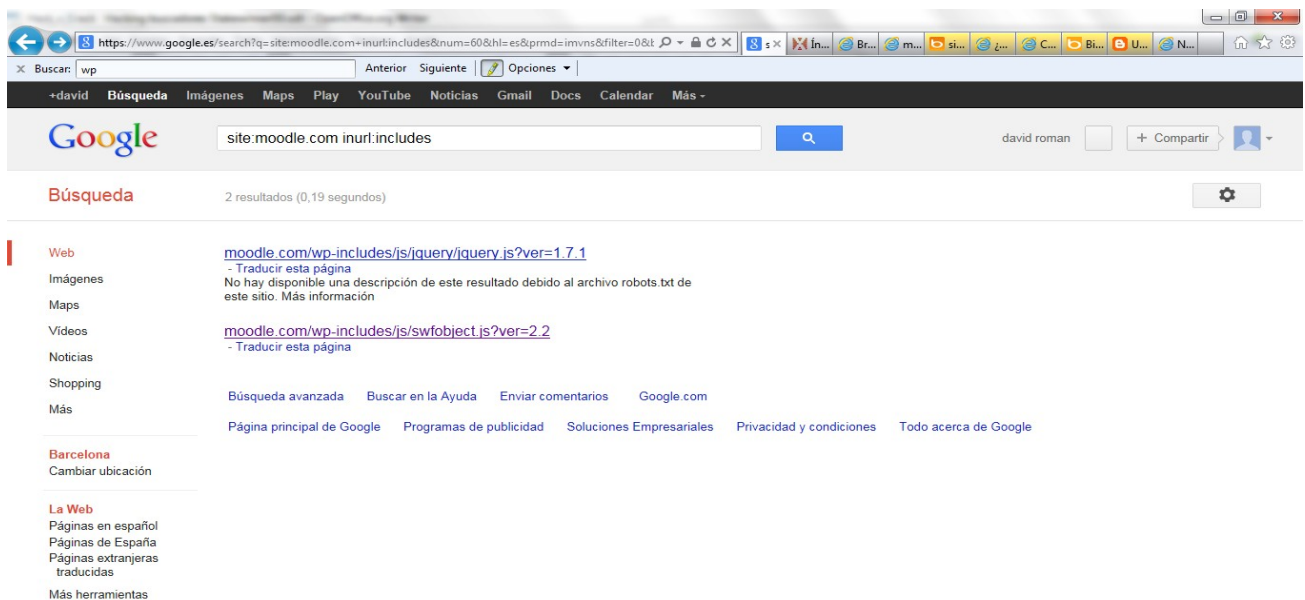
A screenshot of a web browser window. The address bar shows the URL "http://moodle.com/robots.txt". The main content area displays the text of the robots.txt file:

```
User-agent: *  
Disallow: /wp-admin/  
Disallow: /wp-includes/
```

Y Bing si que lo sigue al pie de la letra



Ejemplo de la búsqueda en Bing



Ejemplo de la búsqueda en Google

En cambio Google se lo salta por el forro e indexa esas dos carpetas.

Como vemos Google tiene la ventaja (o desventaja para algunos) que nos permite recolectar información muy interesante, en este caso un archivo javascript que nos podría ayudar a comprender el funcionamiento de la web para una futura intrusión.

Otra ventaja de Google que olvide mencionar es su cache. Que se puede usar para poder ver ataques a webs de hace meses o paginas ya inexistentes.

4.5 – Google Dorks.

Utilizando los operadores descritos y pensando un poco se pueden lograr grandes cosas. De todos modos hay una pagina web donde nos facilitan mucho este trabajo ya que están todos los tipos de búsquedas que podríamos usar clasificados.

<http://www.exploit-db.com/google-dorks/> en esa web esta la GHDB (GoogleHackingDataBase) es la base de datos de diferentes búsquedas que se podrían utilizar para diferentes objetivos.

Voy a explicar un poco los diferentes categorías que hay.

USUARIOS → Buscando por Google podemos encontrar desde una lista de usuarios asta entrar en una web donde al entrar ya eres administrador. Una búsqueda que podríamos utilizar seria `filetype:xls "username | password"` esta búsqueda nos devolvería una lista documento excel (filetype:xls) donde encontraríamos o nombres de usuarios o passwords ("username | password). Puede ser que la lista este obsoleta y muchos nombres de usuario ya no existan o la contraseña no coincida pero siempre hay alguna web con la que funciona.

Link con dorks similares: <http://www.exploit-db.com/google-dorks/2/>

PASSWORDS → Con el ejemplo anterior también nos valdría para encontrar contraseñas. Pero daremos otro ejemplo para poder aprender un poco como va todo.

`inurl:"passes" OR inurl:"passwords" OR inurl:"credentials" -search -download -techsupt -git -games -gz -bypass -exe filetype:text @yahoo.com OR @gmail OR @hotmail OR @rediff` como vemos esta búsqueda ya es un poco más compleja. Bueno vamos explicar un poco que es esto y como funciona.

El `inurl:"passes"`, `inurl:"passwords"` y `inurl:"credentials"` buscara una url que contenga la palabra passes, passwords o credentials,, a esto le podríamos añadir `inurl:"usernames"` lo que nos permitiría encontrar también listas de usuarios.

El OR es el operador lógico que nos permite decirle a Google que queremos una de esas condiciones y no todas ellas.

`-search -download -techsupt -git -games -gz -bypass -exe` eliminando estos términos (con el -) obtendremos una búsqueda más limpia. En este caso excluirá webs de juegos,descargas, etc.

El `filetype:text` que nos buscara archivos con la extensión .txt.

Los `@gmail.com`, `@yahoo.com`, etc buscara archivos de texto que contenga algún correo de esos dominios.

Esta búsqueda nos devolverá muchas webs de phishing por el simple echo que le hemos especificado que busque archivos con la extensión ".txt", y archivos que contengan passwords, y tengan esa extensión normalmente son de webs de phishing.

Link con dorks similares: <http://www.exploit-db.com/google-dorks/9/>

DETECCIÓN DE SERVIDORES → Por mucho que esta función la cumple mejor SHODAN con Google también podemos buscar servidores web por ejemplo servidores que usen apache. Para ello buscaremos en el título de la página usando el operador **intitle:** la cadena quedaría algo así **intitle:"Apache Status" "Apache Server Status for"**

Link con dorks similares: <http://www.exploit-db.com/google-dorks/4/>

MENSAJES DE ERROR → Otra función es la de buscar mensajes de error estos mensajes nos podrían indicar que el programador que diseñó la web no tiene muchos conocimientos o no se dedicó a arreglarlos por lo que es probable que tenga muchas otras vulnerabilidades para explotar.

Un ejemplo para buscar mensajes de error de mysql: **"Warning: mysql_query()" "invalid query" -foro -help -ayuda -como** es importante eliminar términos que puedan molestarnos como "foro" "ayuda" o "como" ya que muchas veces cuando buscamos mensajes de error suelen salir páginas como "Como arreglo este error Warning: mysql_query()", ayuda en este error, y otros similares.

Link con dorks similares: <http://www.exploit-db.com/google-dorks/7/>

CARPETAS SENSIBLES → Una función muy interesante es la de buscar carpetas y archivos sensibles como correos carpetas de configuración y otros archivos por el estilo.

Por ejemplo si queremos buscar emails podemos usar **intitle:index.of /maildir/new/** lo que nos devolvería un listado de correos, hoy mientras estaba probando vi correos de el gobierno de Venezuela donde salían usuarios y contraseñas.

Link con dorks similares: <http://www.exploit-db.com/google-dorks/3/>

ESCANEO DE ARCHIVOS VULNERABLES → Google lo podemos usar también como un escaner de vulnerabilidades web. Por ejemplo si queremos buscar webs con la vulnerabilidad LFD podríamos usar el siguiente texto **allinurl:"forcedownload.php?file="** con esto encontraríamos muchas webs vulnerables a LFD. También entraría dentro de esta categoría si buscamos nombres de servicios vulnerables por ejemplo si sabemos que los creadores de **my little forum** tienen una vulnerabilidad de SQL INJECTION podríamos buscar por **"powered by my little forum"**.

Link con dorks similares: <http://www.exploit-db.com/google-dorks/5/>

ESCANEO DE SERVIDORES VULNERABLES → En esta categoría entrarían todos los servidores con backdoors y otras cosas por el estilo, y hoy estaba intentando encontrar un server con una backdoor y en la primera página había unas 10 webs en el mismo servidor al final conseguí subir el nc y crear mi propia shell, quede alucinado con lo fácil que es hackear una web con ayuda de Google XD

En Google me puse a buscar esta shell **filetype:php inanchor:c99 inurl:c99 intitle:c99shell** como podéis observar el nombre de la shell es c99 para los que quieran jugar un poco les dejo el link de una de esas webs, supongo que al momento de publicar la revista aun funcionara <http://eezeelive.com/files/news/c99.php?act=ls&d=C:\HostingSpaces\sembcl7\2lucks.com\wwwroot&sort=0a> para los que quieran jugar les aconsejo que en vez de meter un mensaje de "hacked by ..." arregléis las webs que estén desfasadas borrando el index.html que moleste a la web principal, hoy conseguí reparar 2 webs que estaban desfasadas y jugué un poco con una tercera que estaba vacía.

Link con dorks similares: <http://www.exploit-db.com/google-dorks/6/>

BUSQUEDA DE INFORMACION JUGOSA → En esta categoría entraría todo ese material que no nos sirve para mucho pero puede ser material muy interesante. Por ejemplo D.N.Is documentos no públicos del gobierno etc.

Por ejemplo buscando **"not for public release" inurl:gob OR inurl:edu OR inurl:mil -.com -.net -.es**

encontraríamos toda esa información secreta, o que como mínimo no es para uso publico, de los dominios del gobierno de educación y dominios militares.

Link con dorks similares: <http://www.exploit-db.com/google-dorks/8/>

PAGINAS DE LOG IN → La ultima categoría que voy ha mostrar es la de descubrir páginas de login. Estas paginas podrían servirnos para poder lograr entrar en estas paginas (con ayuda de fuerza bruta) y poder tener acceso a toda la web. Un simple: **intitle:"Log In" "Access unsecured content without logging in"**

5 - Hacking con Bing

5.1 – Operadores en Bing

Los operadores de Bing son muy parecidos a los de Google pero con unas cuantas excepciones que comentare a continuación.

En Google si queremos buscar archivos log de un servidor ftp tenemos que usar el operador **filetype:log "ftp"**, en cambio en Bing para buscar los mismos ficheros debemos usar el operador **ext:log "ftp"** si usáramos el operador "filetype:log" nos devolvería un mensaje de error.

Para buscar archivos log de un servidor ftp también podemos usar "filetype:" pero tenemos que hacerlo de la siguiente manera. **Filetype:txt "ftp.log"**.

El "filetype:txt" indica que es un archivo de texto plano y luego buscamos esta cadena de texto "ftp.log", utilizando esta característica podemos buscar archivos jar con **filetype:txt ".jar"** y nos buscara un archivo plano .jar.

Un operador especial de Bing es el "feed:", este operador busca fuentes RSS

Otro operador interesante es el "contains:" este operador nos sirve para buscar paginas web que tengan un enlace a archivos zip,rar,bak,tmp o lo que nos parezca. Claro que para encontrar el archivo que queramos encontrar tendremos que buscar en el código fuente

Un tercer operador interesante el "ip:" este operador nos muestra todas las webs que pertenecen a una misma ip. Por ejemplo si queremos mirar las webs que forman parte de un subdominio y ya tenemos una web con ese subdominio podríamos hacer un nslookup al dominio y luego tendríamos que buscar por ip:31.170.160.169 (subdominio netne.net). Esta utilidad es muy interesante a la hora de saber como esta formada una empresa con diferentes dominios alojados en una misma ip.

El operador "loc" permite buscar paginas alojadas en un país concreto pero eso no quiere decir que la pagina este en español. Su uso es simple "loc:es" para paginas españolas "loc:ru" para paginas russas etc.x

5.2 – Ventajas de Bing.

Las ventajas de Bing básicamente son los operadores dichos anteriormente y que Google no implementa.

Utilizando "filetype:txt" utilizado para buscar archivos de texto plano podemos buscar en todos los textos planos en busca de contraseñas y usuarios. Con filetype nos ahorramos tener que estar enumerando todo los tipos de archivos que queremos buscar.

5.3 – Bing dorks.

En la siguiente URL se encuentran muchos otros dorks que se pueden usar, ya no me parare a explicarlos ya que son muy similares a los de la GHDB.

<http://pastebin.com/d2WcnJqA>

6 – Hacking con Shodan

6.1 - ¿Qué es Shodan?

Shodan es un buscador creado por John Matherly especializado en buscar en los banners de las webs, de este modo podemos buscar webs que usen el IIS, APACHE o otro software.

Actualmente SHODAN soporta los siguientes protocolos: *HTTP, HTTP Alternate (8080), HTTPS, RDP, SNMP, MySQL, MongoDB, Oracle Web, Synology, NetBIOS, UPnP, Telnet, SSH, Redis y FTP.*

La parte mala es que te has de registrar para poder utilizar SHODAN al 100%.

Link a entrevista a John Matherly: <http://www.elladodelmal.com/entrevista-john-matherly.html>

6.2 - Operadores en Shodan.

Algunos operadores importantes de SHODAN son:

after:dd/mm/yyyy → Este operador filtra las búsquedas y nos muestra un listado de servidores actualizados después de la data especificada.

before:dd/mm/yyyy → Este operador hace lo contrario a "After", en este caso nos mostrara los resultados previos a la data especificada.

os:windows → Con este operador limitamos la búsqueda a un sistema operativo.

port:21 → Busca servidores de un servicio específico.

net: → Busca en rango de ip's

hostname:dominio.com → Filtra los resultados por nombres de dominios.

6.3 - Ejemplos de dorks en SHODAN.

iis/7.5 200 → Con este dork buscaríamos dominios o ips que tuvieran el servidor web IIS 7.5 y la devolviese un mensaje que nos indique que esta disponible. Sino pusiéramos el 200 (OK) algunos de los servicios que queremos buscar no estarían disponibles como vemos en las imágenes siguientes.

The screenshot shows a Shodan search result for the query 'iis/7.5 200'. The page is divided into several sections:

- Services:** A list of services with their respective counts: HTTP (1,076,882), HTTP Alternate (17,705), Synology (508), Oracle iSQL Plus (268), and MongoDB (34).
- Top Countries:** A list of countries with their respective counts: United States (441,921), China (82,207), United Kingdom (68,263), Germany (49,521), and Canada (37,250).
- Top Cities:** A list of cities with their respective counts: Chaoyang (31,448), Beijing (16,030), Dallas (15,762), Houston (12,045), and Phoenix (11,068).
- Top Organizations:** A list of organizations with their respective counts: HiChina Web Solutions ... (47,413), Amazon.com (10,200), Comcast Business Commu... (7,3059,005), Road Runner (5,968), and Comcast Cable (5,968).

The main content area displays three search results, each with a highlighted status code:

- 401 - Unauthorized: Access is denied due to invalid credentials.** (HTTP/1.0 401 Unauthorized) - Server: Microsoft-IIS/7.5
- Document Moved** (HTTP/1.0 302 Redirect) - Server: Microsoft-IIS/7.5
- 403 - Forbidden: Access is denied.** (HTTP/1.0 403 Forbidden) - Server: Microsoft-IIS/7.5

At the bottom, there is a result for **Default Parallels Plesk Pane** (HTTP/1.0 200 OK) - Server: Microsoft-IIS/7.5.

Como vemos en las imágenes anteriores todas tienen el servidor IIS 7.5 pero algunas de esas ips no están disponibles. Esta característica de shodan nos permite buscar webs que respondan a la manera que nosotros queramos. En el listado siguiente aparecerán los tipos de códigos de respuesta más usados (aplicables a casi todos los comandos de shodan), os recomiendo que vayáis probando con shodan los distintos códigos para practicar y además puedes toparte con alguna sorpresa.

TIPO	Número	Mensaje	Descripción
Succes	200	OK	Correcto
Succes	201	Created	Se recibe después de un comando POST.
Succes	202	Accepted	Se ha aceptado la solicitud.
Succes	203	Partial information	Respuesta a un comando GET indicando que la respuesta no está completa
Succes	204	No response	Se ha recibido la solicitud pero no hay información de respuesta
Succes	205	Reset content	El servidor le indica al navegador que borre los valores de los campos de un formulario
Redirection	301	Moved	Los datos solicitados han sido movidos a una nueva ubicación
Redirection	302	Found	Los datos se encuentran en una nueva dirección URL
Redirection	304	Not modified	El documento solicitado no ha sido modificado.
Client error	400	Bad request	La sintaxis de la solicitud no es correcta
Client error	401	Unauthorized	La solicitud ha sido denegada por no estar autorizado
Client error	402	Payment required	El cliente debe indicar una petición con datos de envió correctos
Client error	403	Forbidden	Acceso denegado
Client error	404	Not found	No existe
Server error	500	Internal error	El servidor web encontró una condición no controlada que le impide continuar

Server error	501	Not implemented	El servidor web no implementa el servicio solicitado
Server error	502	Bad gateway	El servidor web que actúa como puerta de enlace o proxy ha recibido una petición errónea.
Server error	503	Service unavaliabe	El servidor se encuentra saturado
Server error	504	Gateway timeout	La respuesta al servidor excedió el tiempo máximo.

Os:cisco after:1/01/2011 → Este dork buscara los sistemas operativos cisco de después del 1/01/2011, este valor lo podemos modificar para coger muchas ips anteriores a la reparación de un bug, por lo que podríamos probar exploits con webs abandonadas y que nadie se cuida de ellas para poder practicar sin perjudicar a nadie.

Server: SQ-WEBCAM → Con esto podríamos encontrar cámaras web en todo el mundo, yo buscando conseguí webcams de escocia, una tienda de china, un campo de entrenamiento militar, y una oficina de lo que parecía unos grandes almacenes ya que se veía los pies de la gente que pasaban. Como veis se pueden encontrar cosas muy interesantes cuesta un poco de encontrar webcams sin passwords complejas o públicas pero es bastante divertido.

"HP-ChaiSOE" → Este dork como vemos en la imagen anterior busca impresoras Hp Color LasserJet, es un poco inútil pero va bien por si te falta tinta pues imprimes tu documento en china por ejemplo XD. Como vemos estas impresoras tienen un panel de control con los colores etc.

HP Color LaserJet 3800 / 140.109.48.90
HP Color LaserJet Impresor. 3800

Información | Configuración | Redes

Estado del dispositivo

Modo de reposo activado

● Pausa/Reanudar ● Continuar

Consumibles

Tóneres: (% restante)

Cartucho negro: 27%	Cartucho cian: 62%	Cartucho magenta: 28%	Cartucho amarillo: 68%
---------------------	--------------------	-----------------------	------------------------

[Detalles sobre consumibles](#)

Soporte impres.

Entrada/salida	Estado	Capacidad	Tamaño	Tipo
BANDEJA 1	<input type="checkbox"/> Vacío	100 hojas	CUALQUIER TAMAÑO	CUALQUIER TIPO
BANDEJA 2	<input type="checkbox"/> Aceptar	250 hojas	A4	NORMAL
BAND. SUP. EST.	<input type="checkbox"/> Aceptar	N/D	N/D	

[Cambiar configuración](#)

Posibilidades

Número de serie de la impresora: CNYLH66590
 Código fecha del firmware: 20061009 46.038.2
 Impresión dúplex: DESACTIVADA
 Capacidad almac. en RANURA TARJETA: 2 MB
 Capacidad almac. en DISCO RAM: 160 MB
 Ranura DIMM 1: Vacío
 Ranura 2: Ranura DIMM 160 MB

A parte de estos ejemplos hay muchos otros ejemplos de dorks similares, aquí dejo una url con diferentes dorks interesantes.

<http://www.shodanhq.com/browse>

7 – Programas de recopilación de información

7.1 - ¿Para que sirven?

Existen programas que nos pueden ayudar a recopilar información. Estos programas solo pueden buscar información de un dominio en concreto, cosa que nos puede ayudar a conocer el funcionamiento interno, pero no sirven para ataques masivos.

7.2 - Foca

Foca es un software para fingerprinting y information gathering creado por [informatica64](#). Este software se utiliza para auditorias a webs.

Foca nos sirve sobretodo para poder ver meta-datos interesantes de un dominio, para ello foca realiza una búsqueda con diferentes buscadores para hallar los ficheros que pudieran contener meta-datos luego esos archivos se deberán descargar para poder extraer los meta-datos.

Foca es un software muy intuitivo y de fácil uso, que nos puede ayudar a encontrar SQLi, información jugosa, nombres de usuarios, listados de directorios, y otras cosas muy interesantes.

Para descargar FOCA pueden hacer [click aki](#), también hay [un video sobre foca](#) donde Chema Alonso y José Palazón presentan a FOCA en la defcon.

También existe una [versión de foca online](#) que soporta los siguientes archivos: .doc .ppt .pps .xls .docx .pptx .ppsx .xlsx .sxw .sxc .sxi .odt .ods .odg .odp .pdf .wpd .svg .svgz .jpg

Hasta el momento FOCA siempre ha dado muy buenos resultados.

7.3 - [SearchDiggity](#)

SearchDiggity es una serie de herramientas creadas por la empresa STACH&LIU dedicadas a la seguridad ofensiva.

SearchDiggity también fue presentado en defcon y [aquí esta el video](#). SearchDiggity a diferencia de foca esta mas dedicado a la explotación y no tanto al fingerprinting. Para descargarlo hacer click [aquí](#).

8 - Otros usos para los buscadores

8.1 - Utilización de Google como proxy

Google tiene su propio traductor, este traductor es bastante famoso y como muchos sabréis se le puede pedir que traduzca una URL.

Cuando traducimos una URL a través de Google, la ip que se muestra es la IP del Google y no nuestra IP, aprovechándonos de esto podríamos hacer ataques de SQLi a través de el traductor de Google

8.2 - Black Seo

8.2.0 - ¿Qué es?

Para las empresas si su web no esta en la primera pagina ya casi no les sirve de nada estar indexados en los buscadores por ese motivo hay empresas dedicadas especialmente a aumentar la relevancia de las webs en los buscadores.

Estas empresas no siempre usan métodos legales, y algunas veces puede ser que un buscador como Google deje de mostrar tus resultados por hacer "trampas", como es el caso de BMW de Alemania.

8.2.1 - Engañando a los buscadores.

Esta técnica lo utilizaban los diseñadores flash, ya que les era muy difícil obtener relevancia con sus swf, y en general por todas las webs que querían tener mas relevancia. Esta técnica consistía en dar una pagina web falsa.

En estos casos la web detectaba el user-agent, y en caso de que el user-agent fuera de un buscador la web les entregaba una pagina totalmente distinta, destinada única y exclusivamente a tener mas relevancia.

La web de BMW de Alemania hizo esta trampa y Google borro su web de la base de

datos hasta que no cambiase su comportamiento.

8.2.2 - Google Bomb.

El Google bomb es una técnica utilizada para mejorar la relevancia de una web, algunas personas también usaban el Google bomb para asociar términos concretos a webs que no tenían nada que ver, por ejemplo la SGAE con ladrones como se ve en la siguiente imagen.



Antes del día 25 de enero de 2007, día en que Google incorporo un nuevo algoritmo para reducir los Google bombs, Google tenia un algoritmo con el que en una pagina obtenía un lugar superior si estaba enlazada por otras paginas ya conocidas.

Con este algoritmo era muy fácil conseguir un Google bomb, ya que solo era necesario meter un enlace del tipo `Palabra`. Aprovechándose de este error algunas personas consiguieron relacionar la palabra LADRONES con la web del SGAE.

Para protegerse del Google bombing se podía incluir un código html parecido al siguiente.

```
<meta name="googlebot" content="noindex, nofollow" />
```

Para mas información pueden visitar el artículo en wikipedia en español sobre [Google Bomb](#) o el artículo sobre [Google Bomb](#) en inglés con más información.

8.2.3 - Spamindexing.

El Spamindexing tiene varias variaciones, como este no es el tema del que trata el cuaderno solo explicare un poco por encima una de las más usadas, el texto oculto.

El texto oculto es otra técnica que se usa para obtener mayor relevancia. Esta técnica consiste en ocultar un texto en el código fuente para que el buscador lo detecte y se crea que en nuestra web hay ciertas palabras clave que en la mayoría de casos no tiene nada que ver con lo que trata la web.

Esto se puede lograr con CSS, modificando el tamaño de la letra a 0, modificando el color de la letra, etc.

Si intentamos utilizar esta técnica lo mas probable es que los buscadores lo detecten y nos echen fuera, por eso recomiendo tener cuidado al usar estas técnicas ya que en la mayoría de los casos nuestra web seria expulsada.

8.2.4 - Spam.

Otra forma muy común de hacer publicidad conseguir visitas y de paso tener más relevancia es el SPAM .

Esta técnica consiste en postear un link en muchos lugares distintos de ese modo cuando el buscador rastree esa web y vea nuestra link mejoraremos nuestra relevancia.

Existen ciertos programas que se pueden usar para este fin, la mayoría de ellos también nos permiten saltarnos los captchas débiles, crear usuarios,etc. Por eso cuando queramos proteger nuestra web sera necesario implementar un captcha robusto, como siempre cuanto mas seguridad más incomodo resulta para el usuario final y cuanto más cómodo es menos seguridad hay.

Para mas información aquí hay un [power point online](#) sobre captchas. Y aquí un post sobre [como saltarse un captcha de voz](#).