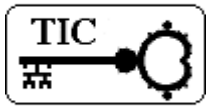


Hacking con Google

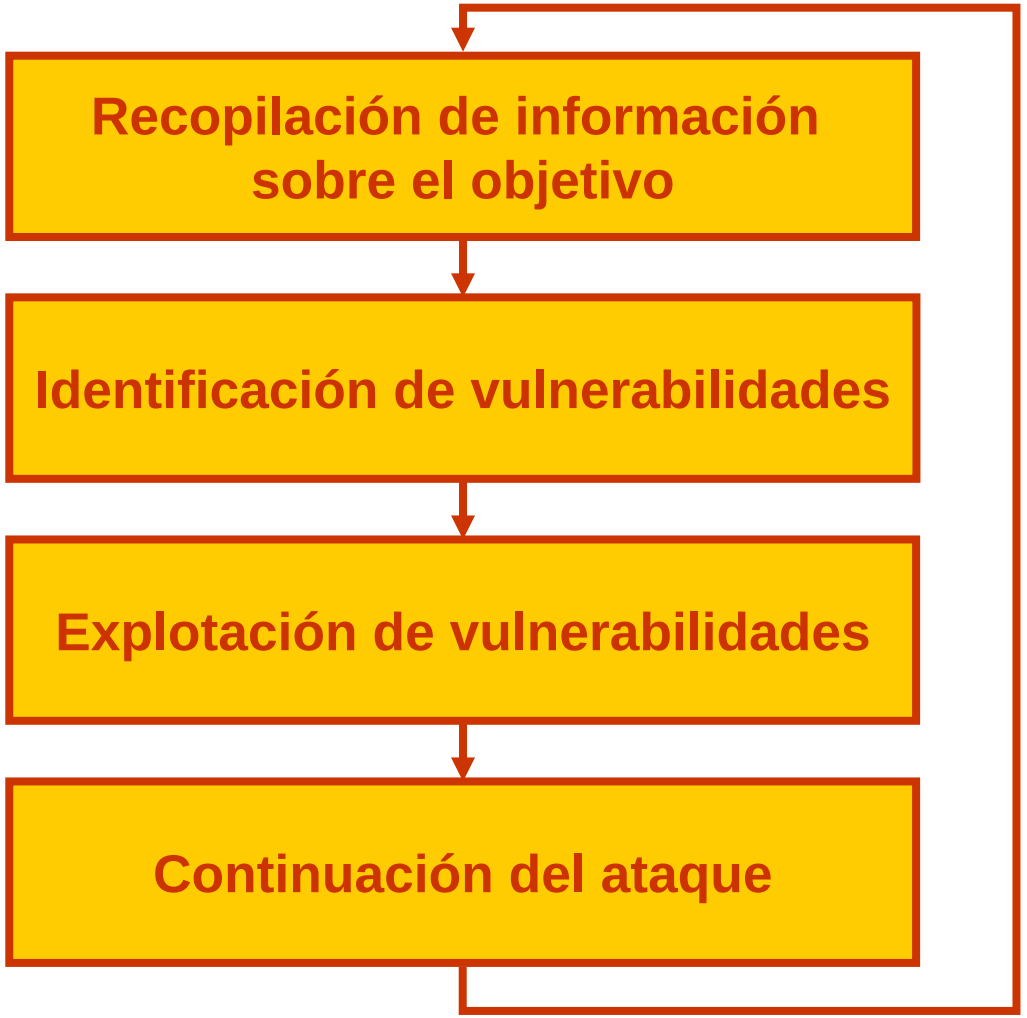
Dr. Gonzalo Álvarez Marañón

Presentación

- Introducción
- El ABC de Google
- Técnicas básicas de hacking con Google
- Búsqueda de información con Google
- Automatizando a Google
- Cómo evitar a Google

Introducción

Proceso de ataque




Abc de Google

Google es mucho más que un simple buscador

Google Services

- Alerts — Receive news and search results via email
- Answers — Ask a question, set a price, get an answer
- Catalogs — Search and browse mail-order catalogs
- Directory — Browse the web by topic
- Froogle — Shop smarter with Google
- Groups — Create mailing lists and discussion groups
- Images — Search for images on the web
- Labs — Try out new Google products
- Local — Find local businesses and services
- Mobile — Use Google on your mobile phone
- News — Search thousands of news stories
- Scholar — Search scholarly papers
- Special Searches — Search within specific topics
- University Search — Search a specific school's website
- Web Search — Search over 8 billion web pages
- Web Search Features — Do more with search

Google Tools

- Blogger — Express yourself online
- Code — Download APIs and open source code
- Desktop Search — Search your own computer
- Hello — Instant message your pictures to friends
- Keyhole — Explore the world from your PC
- Picasa — Find, edit and share your photos
- Toolbar — Add a search box to your browser
- Translate — View web pages in other languages

Consultas básicas

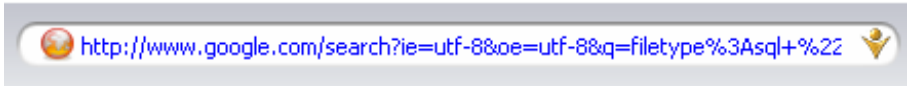
- Consulta de palabras
- Consulta de frases
- Operadores booleanos
  - AND
  - OR
  - NOT
- Caracteres especiales:
  - +
  - -
  - .
  - *

Consultas avanzadas

- Intitle, Allintitle
- Allintext
- Inurl, Allinurl
- Site
- Filetype
- Link
- Inanchor
- Cache
- Numrange
- Daterange
- Info
- Related
- Author, Group, Insubject, Msgid
- Stocks
- Define
- Phonebook

El URL de Google

- q
- start
- num
- filter
- restrict
- hl
- lr
- ie
- Etc.




Hacking básico
con Google

Navegación anónima

- El caché de Google almacena copias de las páginas indexadas
- Consulta:

  site:sitio.com texto
- Las imágenes y otros objetos no están en el caché
- Utilizando la propiedad “texto guardado en el caché” no se cargan las imágenes
- Se necesita utilizar el URL

Servidor proxy

- Un servidor proxy hace las peticiones en lugar del usuario
- Se utiliza la capacidad de traducción de páginas de Google y otros servicios similares:

  www.google.es/translate?u=http://www.playboy.com&langpair=en|en
- No es un proxy anónimo
- Puede saltarse filtros de contenidos

Cabeza de puente

- Google indexa todo lo que se le diga
- ¿Qué pasa si se crea una página con URLs de ataque y se le dice a Google que la indexe?


Buscando
información

Google Hacking Database

- Mantenida en http://johnny.ihackstuff.com/
- Almacena cientos de **googledorks**
- Actualizada continuamente
- Clasificada en varias categorías:
  - Vulnerabilidades
  - Mensajes de error
  - Archivos con contraseñas
  - Portales de entrada
  - Detección de servidor web
  - Archivos sensibles
  - Detección de dispositivos

Listados de directorios

- Contienen información que no debería verse
- Consulta:

  intitle:index.of “parent directory”
- Búsqueda de directorios/archivos específicos

  intitle:index.of inurl:admin

  intitle:index.of ws_ftp.log
- Búsqueda de servidores
  - Apache: intitle:index.of “Apache/*” “server at”
  - IIS: intitle:index.of “Microsoft-IIS/* server at”
- Técnicas de navegación transversal

Reconocimiento de red

- Araña: site:sitio.com
- Objetivo: conocer subdominios
- Consulta:

  site:sitio.com -site:www.sitio.com
- Herramientas automatizadas:
  - SP-DNS-mine.pl www.sensepost.com/restricted/SP-DNS-mine.pl

SO y servidor web

- Listado de directorios
- Aplicaciones/documentación de ejemplo del servidor web
- Mensajes de error del servidor web
- Mensajes de error de aplicaciones web

Búsqueda de sitios vulnerables

- Páginas de demostración del fabricante
- Código fuente
- Escaneo CGI

Información de base de datos

- Nombres de usuario

  “acces denied for user” “using password”

- Estructura de la base de datos

  filetype:sql “# dumping data for table”

- Inyección de SQL

  “Unclosed quotation mark before the character string”

- Código fuente SQL

  intitle:"Error Occurred" "The error occurred in” “incorrect syntax near”

- Contraseñas

  filetype:inc intext:mysql_connect filetype:sql "identified by" -cvs

- Detección de archivos la base de datos
  - filetype:mdb inurl:users.mdb filetype:mdb inurl:email inurl:backup

Nombres usuario/contraseñas

intitle:index.of passwd

intitle:"Index.of..etc" passwd

intitle:index.of pwd.db passwd

intitle:index.of ws_ftp.ini

intitle:index.of people.lst

intitle:index.of passlist

intitle:index.of .htpasswd

intitle:index.of “.htpasswd” htpasswd.bak

filetype:reg reg intext:"internet account manager“

filetype:mdb inurl:profiles

“http://*:*@www”

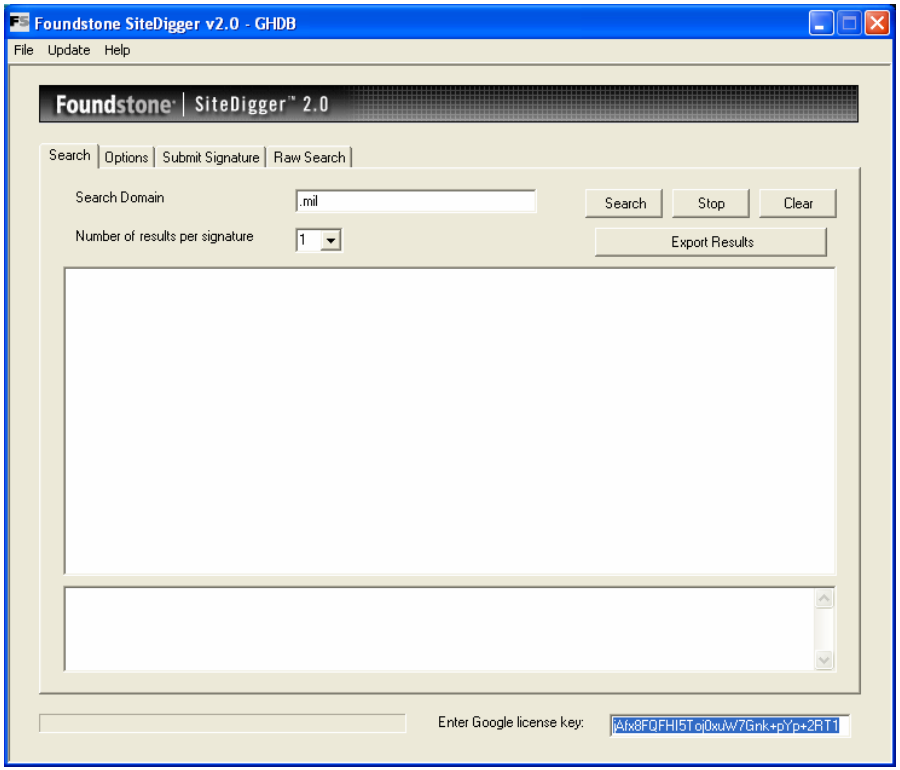

Automatizando
a Google

Google API

- Servicio web gratuito para consulta remota a Google
- Incorpora la potencia de Google a cualquier programa
- Soporta las mismas funciones que el URL
- Limitaciones:
  - Exige registrarse
  - 1000 consultas/día
  - 10 resultados/consulta

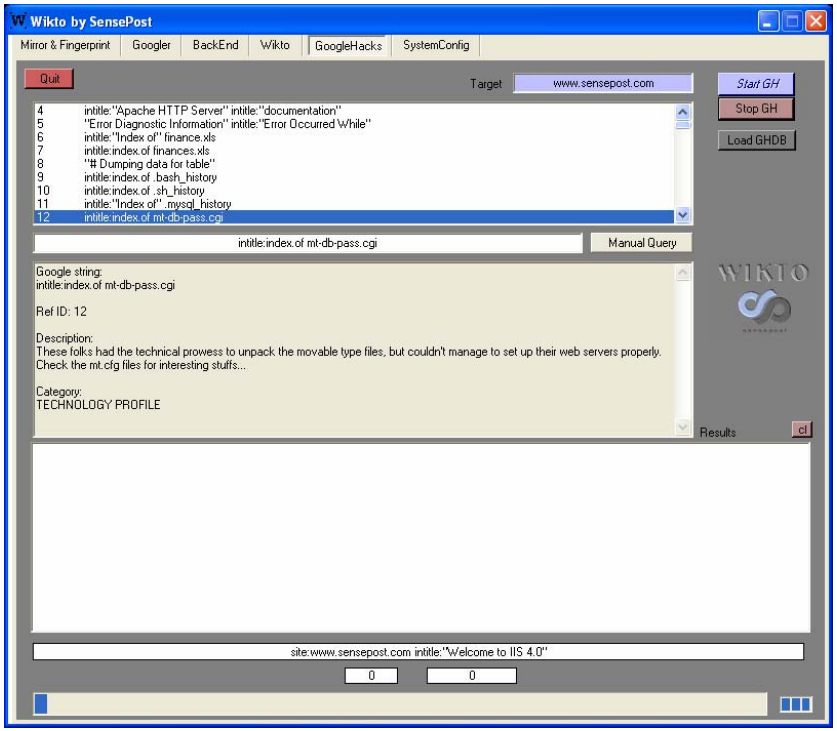
SiteDigger

- Creado por Foundstone www.foundstone.com
- Utiliza la Google API
- Exige la instalación de .NET Framework 1.1
- Utiliza la base de datos FSDB o GHDB
- Genera bonitos informes en HTML



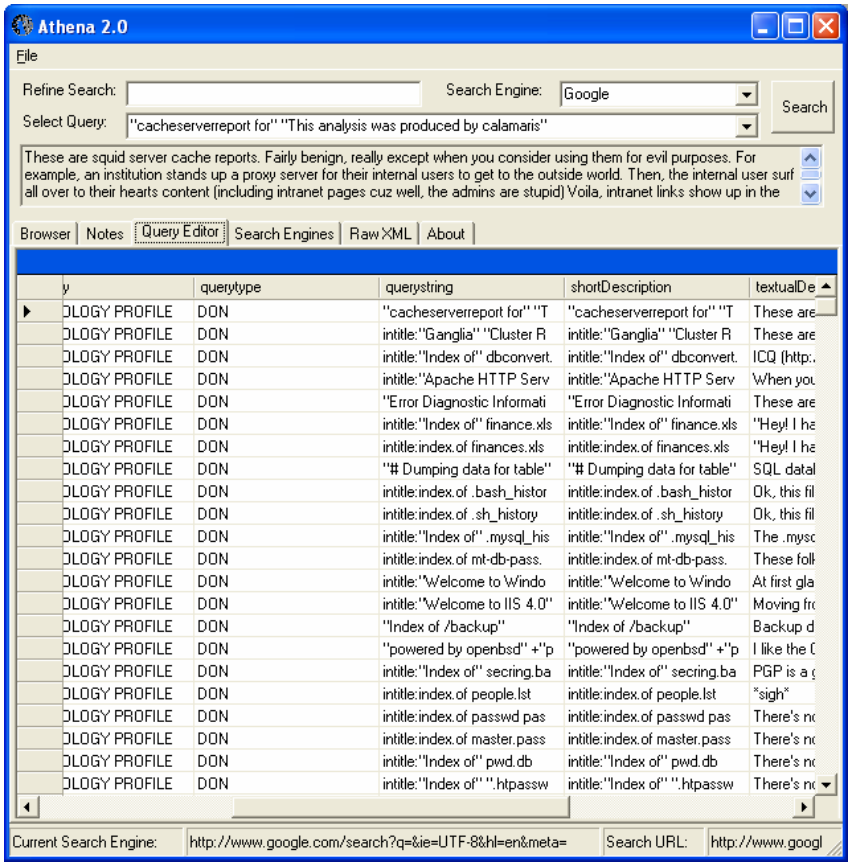
Wikto

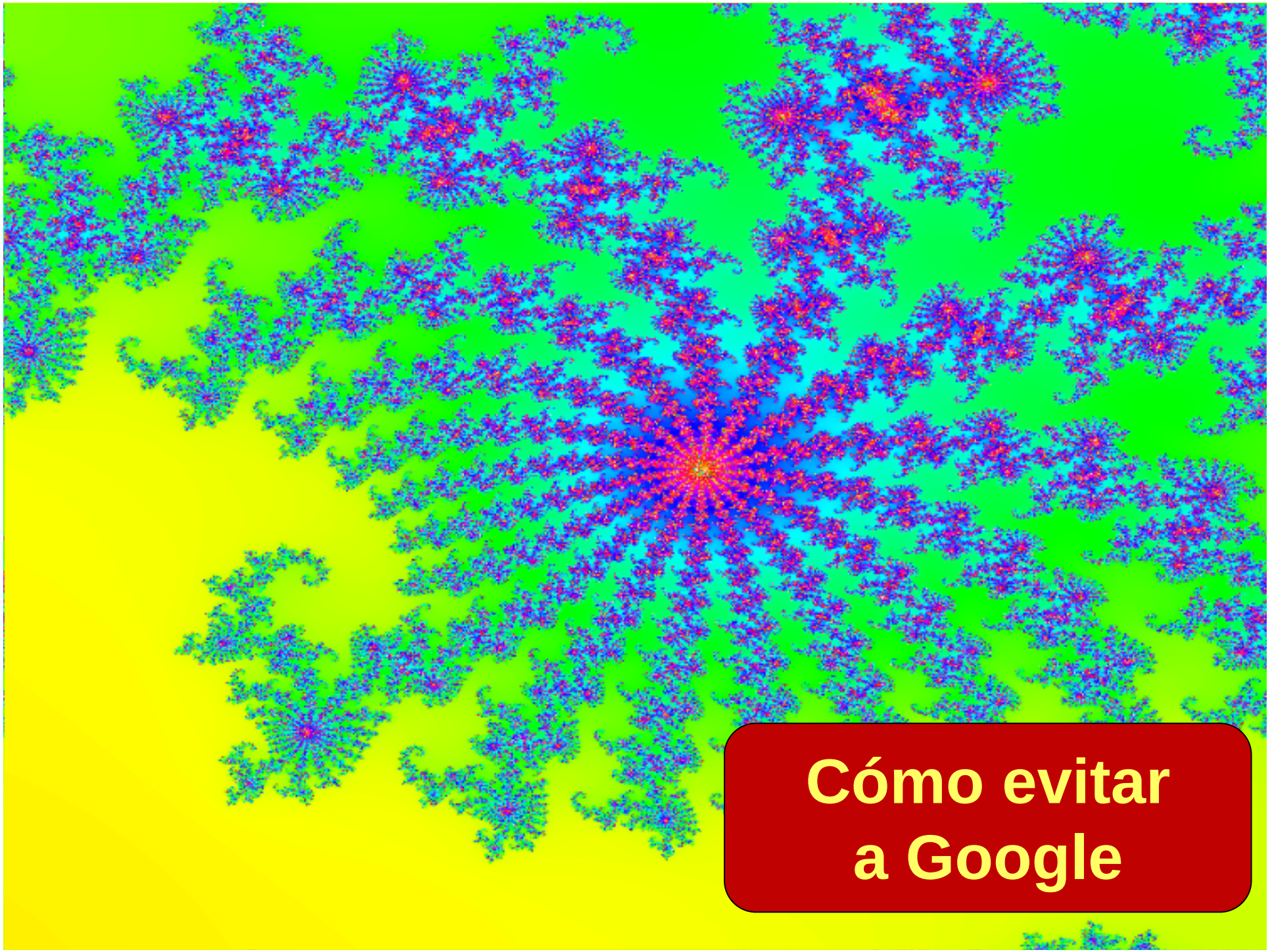
- Creada por Sensepost www.sensepost.com
- Herramienta de análisis de seguridad web de propósito general
- Módulo de hacking Google similar a Sitedigger
- Requiere Google API
- Utiliza GHDB
- Modo de operación automático o manual



Athena

- Creada por SnakeOil Labs snakeoillabs.com
- No utiliza la Google API, por lo que viola los términos de uso de Google
- Requiere el .NET Framework
- Utiliza archivos de configuración XML




Cómo evitar
a Google

Proteger el servidor web

- Bloqueo de buscadores mediante robots.txt
  - User-agent: googlebot
  - Disallow: /directorio/archivos
- No indexar:

  <META NAME="GOOGLEBOT" CONTENT="NOINDEX, NOFOLLOW">
- No almacenar en caché:

  <META NAME="GOOGLEBOT" CONTENT="NOARCHIVE">

  <META NAME="GOOGLEBOT" CONTENT="NOSNIPPET">

Eliminación de páginas de Google

- Eliminación de la caché de Google desde la página de eliminaciones services.google.com/urlconsole/controller
  - Dirigirle al archivo robots.txt
  - Utilizar una directiva META
  - Opción de eliminación de enlaces antiguos

http://www.iec.csic.es/~gonzalo/

Preguntas