Artículo publicado en la revista @RROBA # 97
Suplemento *"Hack Paso a Paso"* #28 – Diciembre 2005
(Material Sin Editar)

Hacking Google

Introducción

Sin lugar a dudas, si en los últimos años ha habido "algo" que ha cambiado el modo en que el común de la gente interactúa con Internet, este ha sido la aparición de "Google". En lo personal, debo admitir que desde un principio, Google me ha parecido un proyecto sumamente interesante. Su particular forma de trabajo, su filosofía, su relación con la comunidad de codigo abierto y el software libre, así como también la innovación constante a la que su gente nos tiene acostumbrada, no hacen mas que mostrarnos el verdadero motivo por el cual, Google se ha convertido en un emprendimiento fascinante y exitoso.

Probablemente, ni siquiera Sergey Brin y Larry Page (23 y 24 años en aquel entonces, y actuales Presidente y CEO de Google) imaginaban el modo en que millones de personas se verían afectadas por el proyecto en el que comenzaban a trabajar aquel otoño de 1995 en la Universidad de Stanford, cuando decidieron desarrollar una tecnología que llamaron "PageRank", la cual inmediatamente fue integrada en "BackRub" (Enero 1996), el primer buscador por ellos desarrollados.

Ya hacia 1997 'Backrub' se transformaba en "Google", nombre otorgado a partir del parecido a la palabra 'googol', que en inglés es el nombre que suele darse a la cifra *"10 elevado a la 100"*.

Pero el tiempo no ha hecho más que traer mejoras en este fantástico buscador y promover nuevos proyectos de interés general tales como "Google Earth" (Luego de conocerlo... créeme que te sentirás observado...), "Google Moon", "Google Maps" y "Google Video Viewer", entre tantos otros, todos ellos salidos de su genial incubadora: "Google Labs" (**Figura 1**). Claro que también existen sitios asociados de gran éxito tales como "Orkut" y "Gmail", los cuales al igual que el buscador, suelen ser utilizados a diario por millones de personas.

Pero yendo al tema que nos interesa... me gustaría contarte en esta nota, de que modo este fantástico suceso llamado Google, a menudo es utilizado por los profesionales relacionados con la seguridad de la información, como una herramienta o recurso fundamental e imprescindible a la hora de llevar a cabo algunas de las tareas que componen sus practicas habituales de evaluación de la seguridad.

Definición, Propósito, Ventajas y Utilidad

Cuando hablamos de "Google Hacking", básicamente nos referimos a la utilización de motores de búsqueda públicos, con el objeto de hallar información sensible del sitio objetivo a través de búsquedas específicamente construidas. Ok... probablemente a esta altura no alcances a ver la potencialidad detrás de un concepto tan sencillo como el que deja entrever esta definición, pero te prometo que al finalizar este artículo tu opinión

respecto del inofensivo y apacible buscador, cambiara rotundamente... o en rigor de verdad, la utilización que comenzarás a hacer del mismo será donde notarás el verdadero cambio!

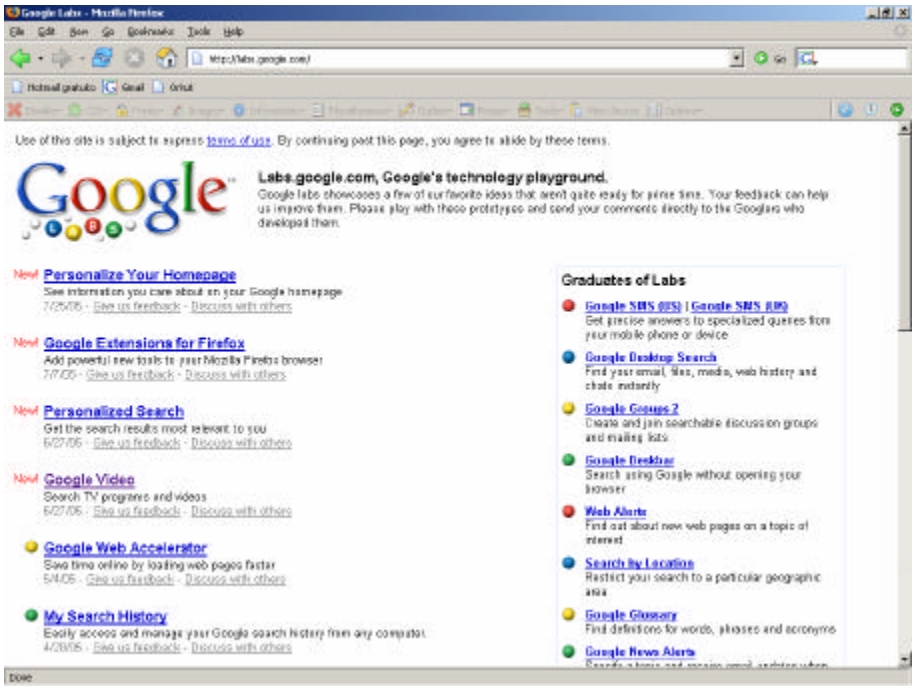

Figura 1

Ahora bien, que hay de la verdadera utilidad? cuales son las etapas, funciones o tareas que siendo realizadas con frecuencia por las personas relacionadas con la seguridad informática, se verán beneficiadas por el uso de las técnicas que describiremos a lo largo de este artículo? Precisamente a los efectos de responder esta pregunta, me permitiré enunciar a continuación, tan solo algunas de las más importantes:

- Obtención de Información Sensible.
  - Archivos de Office.
  - Nombres de Usuarios y Contraseñas.
  - Listado de Directorios.
  - Cuentas de Correo.
  - Etc.. etc... etc..
- Localización de Dispositivos On-Line
  - Localización de Impresoras
  - Cámaras Web
    - Cámaras Públicas
    - Cámaras Privadas
    - Cámaras de Vigilancia
- Localización de Sitios Web Vulnerables
  - Vulnerabilidades Específicas.

  - Servidores de una plataforma determinada.
  - Archivos vulnerables.
- Utilización de Google como Proxy.
- Utilización de Google como Escáner CGI.
- Utilización de Google como herramienta de Salteado de la Autenticación.
- Muchas otras cosas mas!! solo limitadas por la inventiva del Atacante/Auditor!

La utilización de google por parte del profesional de seguridad o del eventual atacante, brinda entre otras ventajas, la posibilidad de recolectar todo tipo de información del sitio objetivo, haciendo de esta una herramienta esencial en la etapa comúnmente referida como "Information Gathering" en la metodología utilizada a la hora de realizar un test de intrusión, pero lo que es aún mas importante, es que este objetivo se logrará sin siquiera enviar un solo paquete en forma directa a la red o host objetivo! logrando de este modo evitar la utilización de otras técnicas o herramientas que podrían ser menos sigilosas.

Para Comenzar...

Antes de comenzar a practicar aquellas búsquedas concretas, que podrían ayudarte en tu próximo test de intrusión controlado, me gustaría recordarte aquellos principios básicos que tienen que ver con operadores booleanos, caracteres especiales y el aprovechamiento de los mismos.

Las búsquedas por palabras clave o frases completas, sin lugar a dudas representan el tipo de consultas mas frecuentes que a diario se realizan sobre buscadores. Tal vez un aspecto poco conocido por los frecuentes usuarios de google, sobre todos aquellos viejos usuarios de Internet y sus primeros buscadores, es el hecho de que Google sólo muestra aquellas páginas que incluyen todos los términos de la búsqueda y que por lo tanto, no es necesario incluir "and" entre los términos como solíamos hacer en otros buscadores. En cambio, para acotar la búsqueda un poco más en el ambiente Google, probablemente resulte más efectivo el agregar más términos a la misma.

Por otra parte, Google hace uso de una serie de "caracteres especiales" a fin de permitirnos interactuar con nuestras búsquedas. Sin ir mas lejos, todos probablemente hemos utilizado el recurso de encerrar entre comillas, una frase de uso común, para encontrar su utilización "exacta" como parte de documentos que la contengan; aunque probablemente pocos conocerán que del mismo modo, es posible (y a veces de suma utilidad) excluir una palabra de la búsqueda colocando un signo menos ("-") inmediatamente antes del término que requiera ser excluido.

El signo "+" por su parte, no solo nos permite unir palabras de una frase (Funcionalidad que generalmente también automatiza google en forma exitosa), sino que en algunas ocasiones puede ser de suma utilidad a fin de exigir al motor de búsqueda que distinga por ejemplo una palabra de otra similar que posea acento, diéresis o la letra eñe (ñ) tan común en nuestro idioma. Por ejemplo, [Hernán] y [Hernan] tipiados indistintamente en

el campo de búsqueda devolverán el mismo resultado. Si nuestra intención fuera distinguir estas dos palabras, podríamos utilizar el signo +, escribiendo [+Hernan] en vez de [+Hernán].

Bien, puesto que sin dudas estarás esperando otra cosa de este artículo, no planeo seguir aburriéndote con conceptos básicos, solo me gustaría invitarte a que investigues por tus propios medios los diferentes tipos de caracteres especiales, que junto a los mencionados, podrían ser de tu interés. Del mismo modo, probablemente no sea mala idea el que puedas interiorizarte de la sintaxis utilizada por Google al momento de conformar las URLs o URIs de búsqueda! te sorprenderás! Y como no podía ser de otro modo, Google será un buen lugar donde comenzar a investigar respecto de estos aspectos :D ¡!

Operadores Avanzados

Ok, crees que lo has visto todo, crees que los buscadores no tienen secretos para ti y que Google ya te ayuda bastante a diario como para exigirle más... mmmm bien... espera a ver lo que podemos hacer tan solo yendo un poco mas allá por medio de la utilización de lo que suele ser conocido como "Operadores Avanzados"!!!

Puesto que todo es cuestión de práctica, en esta sección tan solo presentaremos la sintaxis requerida y su funcionalidad asociada, dejándote a ti en libertad de cargar tu browser y comenzar a realizar tu propia experiencia! En todos los casos, para ver los resultados, bastará con que ingreses el "string" de búsqueda en el campo correspondiente del sitio de Google, presiones el botón mencionado como "Búsqueda en Google" y observes lo que sucede. Y por si acaso no tuvieras suerte armando tu string de consulta con estos operadores o tu imaginación no estuviera en su mejor día, la sección "Ejemplos: Jugando con Google" al final de este articulo, sin lugar a dudas servirá para aclarar algunas ideas y despertar aún mas tu curiosidad por este tema!!!

A continuación, alguno de los operadores avanzados:

Link: El operador "Link" nos permite básicamente, buscar páginas que *linkeen* a otras páginas, buscando el término indicado, pero solo cuando el mismo aparezca en un link.

Ejemplo: Link: www.hernanracciatti.com.ar

Tip: Este operador por ejemplo, podría ser utilizado en la etapa de "Information Gathering" a fin de conocer todos aquellos sitios que se relacionan de algún modo con el sitio objetivo.

Cache: Un aspecto de Google muy discutido pero de suma utilidad, es aquel relacionado con la conservación en la Cache, de *snapshots* de páginas que habiendo sido indexadas por el buscador, continúan almacenadas en su estado original, aún cuando las mismas hayan cambiado, desparecido o se encuentren actualizadas. El operador

"Cache", precisamente busca el termino indicado como argumento, en la versión *cacheada* del site almacenada en Google.

Ejemplo: Cache: http://www.servidor-objetivo.com

Nota: *Eventualmente este operador podría terminar en resultados erróneos. Si este fuera el caso, aún tienes la posibilidad de hacer una búsqueda tradicional y luego escoger en la página de resultados abrir la información almacenada en la cache, tal como podrás observar en la* **Figura 2**.

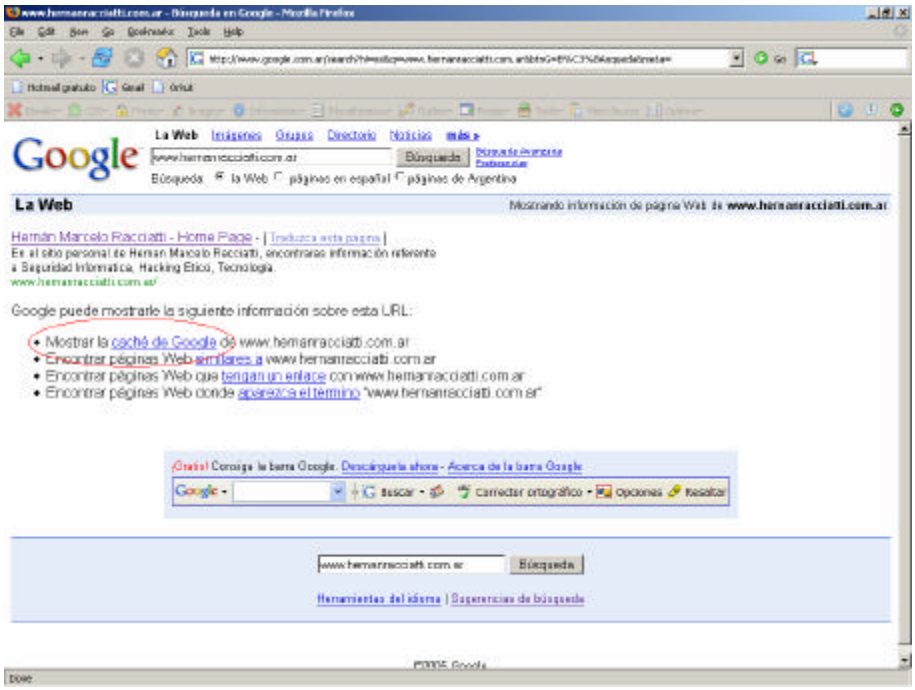

Figura 2

Tip: La utilización de la cache de Google, puede eventualmente significar un gran riesgo en sitios mal administrados o descuidados, a la vez que podría permitir al profesional en función de "tester", obtener algunos beneficios adicionales tales como el salteado de la autenticación sin esfuerzo alguno!!! Sucede que en determinadas oportunidades/configuraciones, contenido privado normalmente accedido en forma on-line previo ingreso de "credenciales validas" (Usuario y Contraseña), al ser indexado y luego almacenado en la cache de Google, termina por encontrarse fuera del ámbito de protección del servidor original, permitiendo de este modo que un eventual intruso logre acceder a información off-line (O cacheada) que inicialmente fue pensada como privada. (**Figura 3**)

Intitle: Este es uno de los operadores mas utilizados por los atacantes. Al indicar la palabra clave "Intitle", en el campo de búsqueda de Google, obtendremos como

resultado, precisamente aquellas páginas en donde el termino a buscar se encuentre contenido en el titulo mismo del Site.

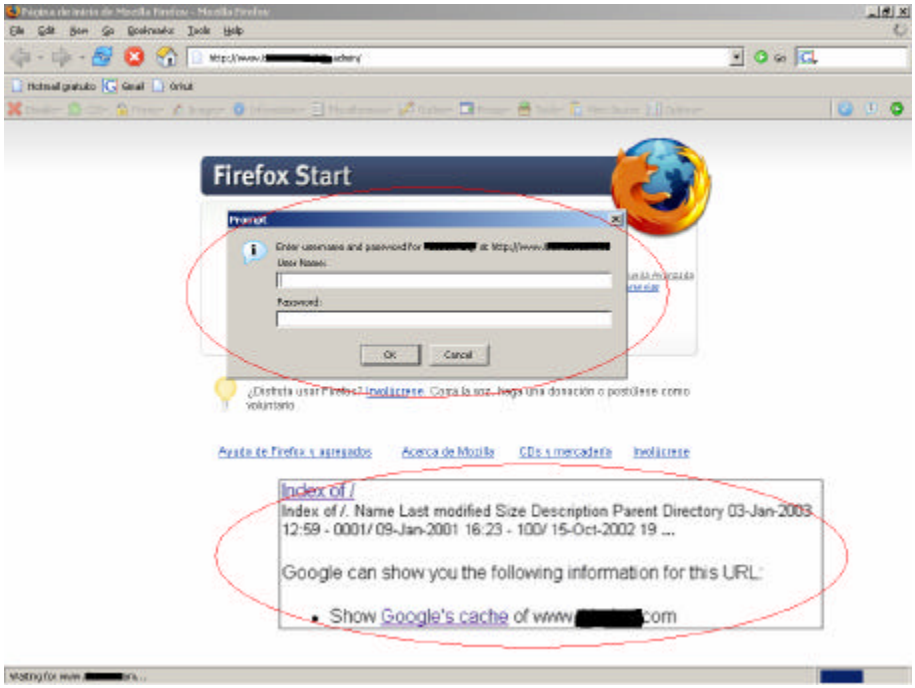

Figura 3

Inurl: Básicamente, con este operador, estamos instruyendo a Google que busque el término o parámetro indicado, solo cuando este se encuentre contenido en la dirección web o URL. Quizás a simple vista este pueda ser un operador sin mucha utilidad, pero lo cierto es que mediante este, podremos ser capaces incluso de utilizar Google como un sigiloso "Escaner CGI", por medio de la solicitud de archivos específicos, conocidos como vulnerables!!!

Filetype: Este operador, nos abre una gran puerta al permitirnos buscar y encontrar algunos de los formatos de archivos más populares. Aunque... probablemente no existan archivos de Word o Excel con información confidencial en la web... o estaré equivocado??? :)
Tip: Como veremos en nuestra sección de ejemplos, gran parte de las técnicas de Hacking Google mas utilizadas, se encuentran relacionadas con los operadores "Intitle","Inurl" y "Filetype".

Ejemplos: Jugando con Google

Bien... hemos dado mucha lata hasta aquí y estimo que estarás deseoso por dejar de leer y comenzar a jugar con los operadores en Google!!, Me parece perfecto! de este modo, probablemente entiendas porque se ha estado hablando tanto de "Hacking Google" en los últimos años.

A continuación, encontraras una serie de ejemplos y utilidades, a los cuales a menudo solemos referirnos como "***googledorks"***, con los que podrás comenzar a probar y de este modo evaluar de qué modo el resultado obtenido en cada caso, puede colaborar con tus tareas al momento de realizar un "Test de Intrusión Controlado":

Servidores Windows 2000 con IIS, Instalación por default... muy posiblemente vulnerables.
intitle:"Welcome to Windows 2000 Internet Services"

Servidores Windows NT 4.0 con IIS, Instalación por default... muy posiblemente vulnerables.
intitle:"Welcome to IIS 4.0"

Servidores en estado "Under construction"... muy posiblemente vulnerables.
intitle:"Under construction"

Servidores con Terminal Services... existen ataques para Terminal Services? :)
intitle:"Terminal Services Web Connection"

Archivos conteniendo Passwords...
intitle:"Index of" config.php
intitle:index.of.etc
inurl:passwd filetype:txt (**Figura 4**)

Copias de Resguardo
intitle:index.of./ backup

Posible Información Sensible
allintitle:private filetype:doc
allintitle:private filetype:xls
inurl:admin

CGI Scanner (Búsqueda de archivos conocidos por sus vulnerabilidades)
inurl:/random_banner/index.cgi
inurl:/cgi-bin/mailview.cgi
inurl:/cgi-bin/cgiemail/uargg.txt

Dispositivos On-Line (Camaras? Dispositivos con Passwords por defecto?)
allintitle:Brains, Corp. camera
intitle:Cisco Systems, Inc. VPN 3000 Concentrator
camera linksys inurl:main.cgi

Consultas Simples: SQL Injection
"Unclosed quotation mark before the character string"

Curiosidades
filetype:reg reg +intext:"internet account manager"

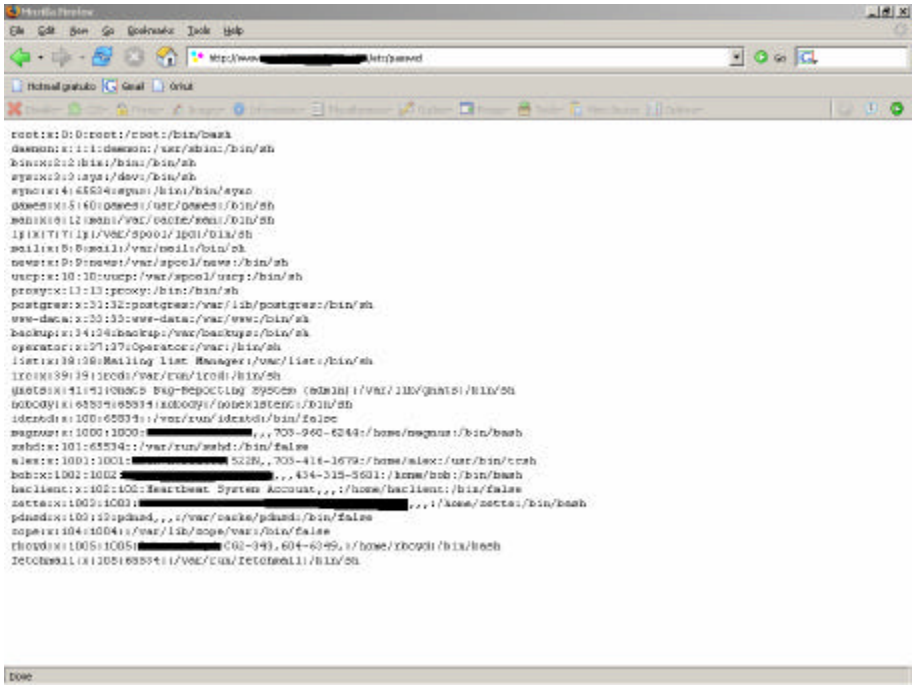

Figura 4

El predicador

Si bien es cierto que la utilización de buscadores por parte de los profesionales en seguridad, es tan antigua como la aparición misma de los buscadores (o casi…), en verdad existe una persona en particular, que con su trabajo y estilo particular, ha sabido llevar un paso mas allá aquellos aspectos relacionados con Google y la utilización de *"Googleforks"*: Johnny Long!!

Johnny se ha convertido desde hace un tiempo, en un referente respecto de la utilización de Google como herramienta de hacking (o algo así…). A través de sus presentaciones, sus trabajos de investigación y el sitio que mantiene en Internet (http://johnny.ihackstuff.com) (**Figura 5**) en donde es posible encontrar entre otras interesantes cosas, una base de datos repleta de "Googleforks" denominada simplemente "Google Hacking Database (GHDB)".

Si eres de los que prefieren tener un libro entre sus manos, sin lugar a dudas "Google Hacking for Penetration Testers", un excelente trabajo de Johnny Long, que cuenta con la participación especial de Pete Herzog (ISECOM) y Matt Fisher (SPI Dynamics) entre otros, significara un recurso de valor incalculable.

Figura 5

Google API y Herramientas Automatizadas

Bien… si has llegado hasta esta sección, probablemente te estés preguntando si hay algún modo de automatizar algunas de las técnicas, búsquedas o "Googleforks" que hemos tenido oportunidad de conocer. La respuesta es afirmativa y a continuación haremos un resumen muy escueto de cada una de ellas a la vez que incluiremos los links desde donde podrás descargarlas y echar manos a la obra.

Pero antes de esto, es necesario que conozcas que Google, no solo permite la realización de consultas o búsquedas a través de su interfaz web, sino que por el contrario existen otros métodos. Uno de los mas utilizados y en regla con los "Términos de Uso de Google" es la "Google API", una completa interfaz de aplicación, a partir de la cual se alienta a la comunidad a desarrollar herramientas que puedan potenciar el motor de búsqueda de Google a través de distintos tipos de interfaces.

Básicamente, "Google API" es un servicio web gratuito para ejecutar consultas remotas a Google. De este modo, es posible incorporar la potencia de Google a cualquier programa (Si… a cualquiera, por ejemplo los de la Intranet de tu compañía, o aquellos que desarrollas para tus clients!). Si bien la utilización de esta API, soporta las mismas funciones que las existentes a través de la URL, también es cierto que posee algunas limitaciones a saber:

- Exige registrarse (Y obtener de este modo una clave de utilización)
- Permite hasta 1000 consultas/día

- Devuelve 10 resultados/consulta (?)

Ok, ahora que conoces un poco mas de "Google API" y su aprovechamiento, veamos rápidamente que tal lucen algunas de las herramientas que aprovechándose de de esta nos permiten agilizar nuestro trabajo.

SiteDigger (**Figura 6**)
Ya en su versión 2.0, esta herramienta creada por Foundstone, sin dudas es una de las más utilizadas. Su utilización es muy sencilla. Requiere .Net Framework 1.1 para funcionar, así como también tu "Google License Key" o clave de licencia de Google, la cual podrás obtener sin mayor complicaciones previo registro en la sección APIS de Google (http://www.google.com/apis). Site Digger utiliza la base de datos FSDB (Foundstone Database) o GHDB y nos brinda unos informes sumamente útiles y bonitos en formato HTML.

http://www.foundstone.com/resources/proddesc/sitedigger.htm

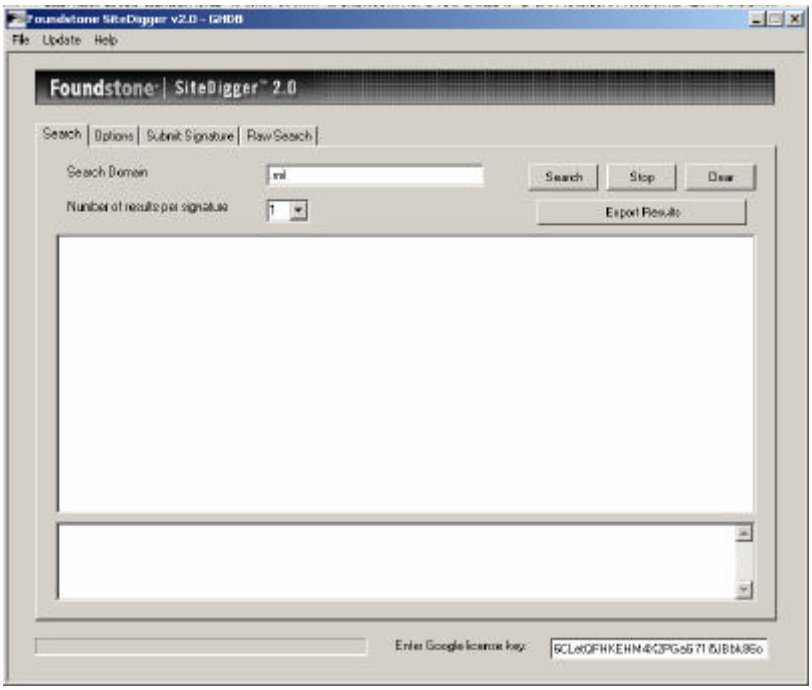

Figura 6

Wikto
Wikto (**Figura 7**) es una excelente herramienta de análisis de seguridad de servidores web de propósito general creada por Sensepost (http://www.sensepost.com), la cual incluye como funcionalidad adicional un modulo o plugin de hacking google (GoogleHacks) algo similar a SiteDigger. Al igual que la mayoría de las herramientas de este tipo, para utilizar esta funcionalidad en Wikto necesitaras ingresar tu "Google License Key". Con modos de operación tanto manual como automático, Wikto sin lugar

a dudas puede ser tu elección, debido a que probablemente sea una herramienta que al margen de esta capacidad, ya tengas instalada en su sistema… o no la tienes? Que esperas!!

http://www.sensepost.com/research/wikto

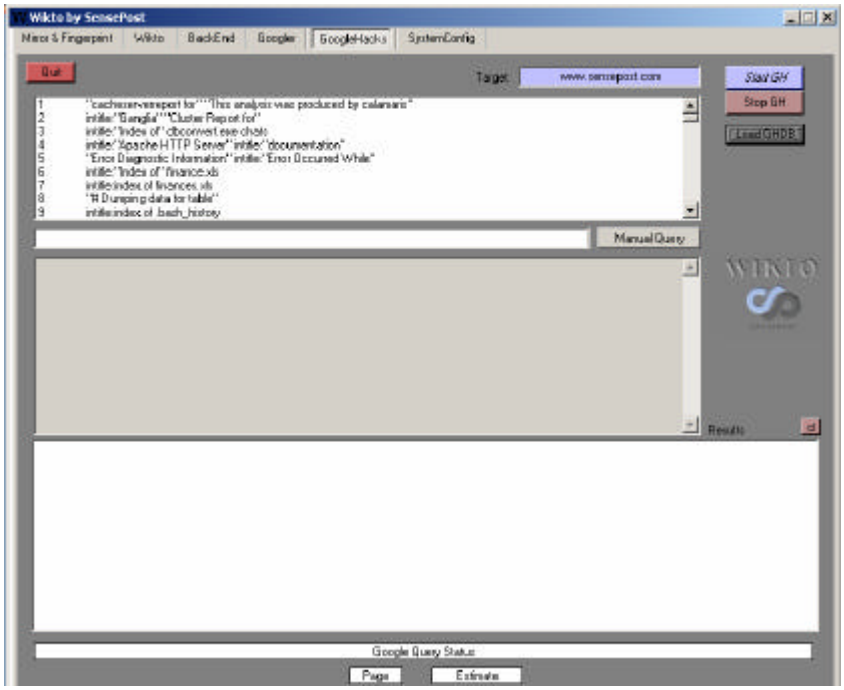

Figura 7

Athena (Figura 8)

Desarrollada por la gente de SnakeOilLabs, esta herramienta es bastante diferente al resto. Por empezar, Athena 2.0 NO utiliza la Google API, por tanto viola los términos de uso de Google!!! Si bien es cierto que este hecho podría ser ampliamente criticado por algunos, yendo a lo estrictamente técnico y dejando de lado aspectos éticos y/o legales, Athena es sencillamente una herramienta asombrosa!

Para ponernos a trabajar con ella, requeriremos tener instalado el .Net Framework. Athena hace un excelente uso de XML, y sus archivos de configuración precisamente se encuentran en dicho formato. A diferencia de lo que ocurre con el resto de las herramientas comentadas, Athena requerirá de la descarga y lectura del manual de usuario que puede ser encontrado en el mismo site. Sucede que la herramienta requiere de ser comprendida conceptualmente para luego comenzar su customizacion a fin de lograr el tuning necesario para hacer de esta una herramienta realmente infalible. Por cuestiones de espacio, se hace difícil en este artículo explicar su uso, pero créeme que el tiempo invertido en su manual, te será devuelto en prestaciones.

http://www.snakeoillabs.com

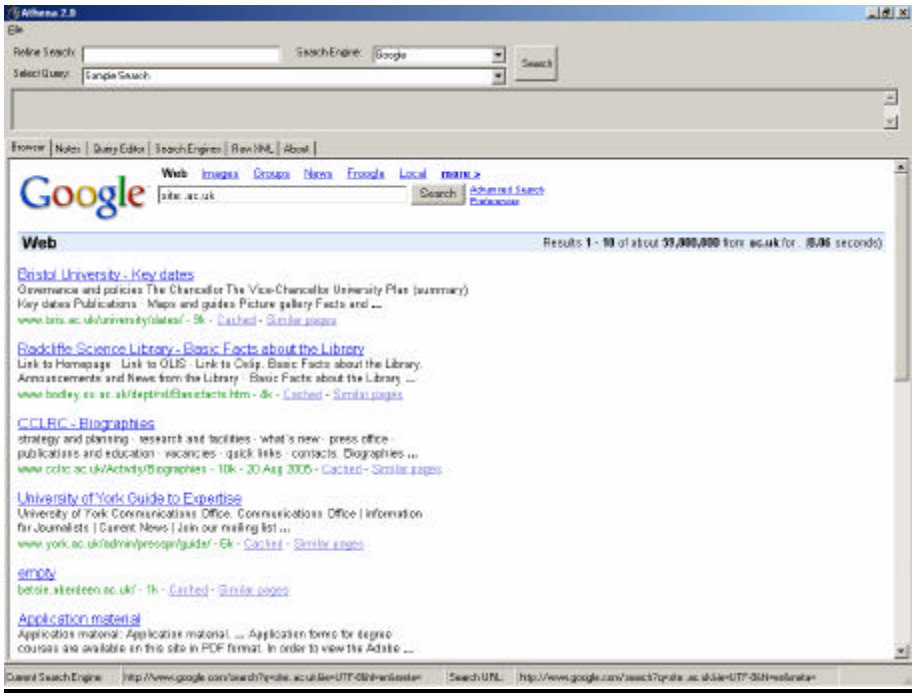

Figura 9

Conclusión

Si este articulo logra generar en ti la curiosidad suficiente como para que tomes tu computador y comiences el verdadero aprendizaje de las técnicas de Hacking Google, haciendo tu propia experiencia, sin lugar a dudas mi objetivo primordial se habrá cumplido!! Muchos aspectos han quedado fuera de este artículo de carácter introductorio, aspectos que probablemente puedas investigar y practicar por tu cuenta.

Es que como siempre digo, el Hacking es algo practico y si bien es cierto que a menudo este requiere de fuertes fundamentos, también lo es el hecho de que nada se logra aprender realmente si no es con practica y experiencia propia.

Por ultimo, déjame decirte que la utilización de Google como herramienta, puede ser sumamente divertida, espero puedas tomar lo comentado en estas paginas para que junto a tu inventiva, seas capaz de colaborar con la comunidad, generando nuevos *Googledorks* que permitan desarrollar aun mas, la base de datos que Johnny Long mantiene en su sitio!!

Hernán Marcelo Racciatti
http://www.hernanracciatti.com.ar

Referencias

Imágenes del almacenamiento construido a base de "Lego's" en los inicios de Google.
http://www-db.stanford.edu/pub/voy/museum/pictures/display/0-4-Google.htm

http://www.google.com
http://moon.google.com
http://earth.google.com
http://maps.google.com
http://video.google.com
http://labs.google.com
http://www.orkut.com
http://www.gmail.com
http://www.snakeoillabs.com
http://www.sensepost.com/research/wikto
http://www.foundstone.com/resources/proddesc/sitedigger.htm
http://johnny.ihackstuff.com
http://www.amazon.com/exec/obidos/ASIN/1931836361/ihackstuff-20/102-0373644-9893725?creative=327641&camp=14573&link_code=as1
http://www.amazon.com/exec/obidos/ASIN/0596004478/ref=pd_sxp_elt_l1/102-0373644-9893725