

Guía de referencia SEO

Javier Casares García

www.javiercasares.com

www.ojobuscador.com

Guía de referencia SEO

Guía de referencia SEO
© 2007-2008 Javier Casares
© 2007-2008 OJO internet, S.L.

Apartado de Correos 2017, 08917 Badalona (Spain)
CIF: ESB64307036

“OJObuscador” y “Mascota OJO” son marcas registradas

1ª edición: Mayo 2007
2ª edición: Junio 2008 (ampliada)

Maquetación: David Ordóñez Rodríguez
Edición: OJO internet, S.L.
Imprime: Bubok Publishing, S.L.

Impreso en España - Printed in Spain

guía SEO

guía de referencia SEO

Tabla de contenidos

0 Antes de empezar a leer	7	2 Factores Internos (on-page)	17
0.1 Historia del posicionamiento en buscadores.....	7	2.1 Dominio.....	17
0.2 Por qué aparecer primero, segundo o tercero.....	8	2.1.1 Territorialidad.....	17
0.3 Nadie puede asegurar tu lugar.....	9	2.1.2 Edad.....	18
0.4 Lo importante son los contenidos.....	9	2.1.3 Renovación.....	18
0.4.1 Contenidos Originales.....	9	2.1.4 Dominios Aparcados.....	18
0.5 Qué debes saber antes de ser.....	10	2.2 Programación.....	19
0.6 ¿Por dónde empezar?.....	10	2.3 La cabecera de la página web.....	20
0.6.1 A quién va dirigido el sitio web.....	10	2.3.1 El título de cada página.....	21
0.6.2 Qué buscan para encontrarte.....	10	2.3.1.1 La calidad de un buen título.....	22
0.6.3 Errores comunes.....	11	2.3.2 La descripción de la página.....	22
0.7 Rankings: Factores.....	12	2.3.3 Las palabras clave de la página.....	22
0.7.1 Navegabilidad.....	12	2.4 Los Enlaces.....	22
1 Situación actual	14	2.5 Cabeceras, Negritas, Cursivas, Listados.....	24
1.1 Páginas indexadas.....	14	2.6 Las imágenes.....	25
1.2 Enlaces entrantes.....	15	2.7 Flash.....	26
1.3 Herramientas para Webmasters.....	16	2.8 Los contenidos.....	26
		2.8.1 Contenido fresco.....	27
		2.8.2 Frases de búsqueda y palabras clave.....	28

2.8.2.1 Densidad de contenidos.....	29	5.2 Google.....	46
2.8.3 Contenidos atractivos.....	30	5.3 Yahoo!.....	47
2.9 Estructura del sitio.....	30	5.3.1 Yahoo! SearchScan.....	47
2.10 Mapa del Sitio.....	31	5.4 Windows Live.....	47
2.11 Dirección URL.....	32	6 Los pequeños detalles.....	48
2.12 robots.txt.....	34	6.1 Densidad de Palabras.....	48
2.12.1 Dirección URL del protocolo Sitemaps.....	35	6.2 Alta en motores de búsqueda.....	48
2.12.2 Crawl-delay.....	35	6.3 Sitemaps.....	49
2.13 Página no encontrada (Error 404).....	35	6.4 Aviso manual de actualización.....	49
2.14 Hosting / Alojamiento Web.....	36	6.5 Panel de gestión.....	50
3 Factores Externos (off-page).....	37	6.6 Aumentar la velocidad de indexación.....	51
3.1 La fuerza de la IP.....	38	6.7 Cambiar la dirección web o dominio.....	51
3.1 Conseguir enlaces de sitios con autoridad.....	38	6.8 Sitio “en mantenimiento”.....	52
3.3 Directorios.....	39	6.9 Enlaces y Pagerank.....	53
3.4 Intercambio de Enlaces.....	39	6.9.1 Valor del Pagerank.....	53
3.5 Alquiler de Enlaces.....	40	6.9.2 Pagerank como filtro AntiSpam.....	54
3.6 Notas de Prensa.....	40	6.10 Tiempo de permanencia.....	55
3.7 Foros.....	40	6.11 ¿Cómo ve mi sitio un buscador?.....	55
3.8 Páginas Prepagadas.....	40	6.12 ¿Cuándo saldré posicionado?.....	55
3.9 Comprar Artículos.....	41	6.13 Indexación selectiva.....	56
3.10 Link Baiting.....	41	6.13.1 Google Search Appliance.....	56
3.11 No Follow.....	42	6.13.2 Yahoo!.....	56
3.12 La Competencia.....	42	6.14 Buscadores de noticias.....	56
3.13 Envíos Automáticos.....	42	6.14.1 Google Noticias.....	56
4 Técnicas Penalizables.....	43	6.14.1.1 Sitemap para Google Noticias.....	57
4.1 Cloaking.....	43	6.14.2 Yahoo! Noticias.....	57
4.2 Texto Oculto.....	43	7 Búsquedas y etiquetas.....	58
4.3 Granja de Enlaces.....	44	7.1 Buscador en el sitio.....	58
4.4 Páginas Puerta.....	44	7.2 Las etiquetas.....	59
4.5 Contenido Duplicado.....	44	8 SEO para Blogs.....	60
4.6 Redirección HTML / Javascript.....	44	8.1 Dominio y alojamiento.....	60
4.7 NoScript / Nolframe.....	45	8.2 Los contenidos.....	60
4.8 Sitios de contenido dudoso.....	45	8.2.1 Frecuencia.....	61
5 Salir de una penalización.....	46	8.3 ¿Categorías o Etiquetas?.....	61
5.1 Ask.....	46	8.4 Enlaces.....	61
		8.5 Socialización.....	62

9 SEO para Wordpress	63	10.4 Microformato Geo, hCard y hCalendar	74
9.1 Instalación.....	63	10.4.1 Geo.....	74
9.2 Configuración.....	63	10.4.2 hCard.....	75
9.2.1 Creación del .htaccess.....	63	10.4.3 hCalendar.....	75
9.2.2 Administración.....	63		
9.2.2.1 General.....	63	11 SERM Search Engine Reputation Management	76
9.2.2.2 Escritura.....	64	11.1 Detección.....	77
9.2.2.3 Privacidad.....	64	11.2 Comienza el movimiento.....	77
9.2.2.4 Permalinks.....	64	11.3 Ocupar los resultados.....	77
9.3 Gestión de Plugins.....	64		
9.3.1 All in One SEO Pack.....	65	12 SMO Social Media Optimization	79
9.3.1.1 Opciones del plugin.....	65	12.1 Algunos detalles.....	79
9.3.2 Askimet.....	66	12.1.1 Favicon.....	79
9.3.3 Feed Statistics.....	66	12.1.2 Versión para imprimir o descargar.....	79
9.3.3.1 Feed.....	66	12.2 Marcadores sociales.....	80
9.3.3.2 Top Feeds.....	67	12.3 Sitios de recomendación.....	80
9.3.3.3 Feed Readers.....	67	12.4 Preguntas y Respuestas.....	80
9.3.3.4 Post Views.....	67	12.5 AJAX.....	81
9.3.4 Google XML Sitemaps.....	67	12.5 Páginas personalizadas.....	81
9.3.4.1 Basic Options.....	67		
9.3.4.2 Post Priority.....	68	13 SEO para vídeos	82
9.3.4.3 Sitemap Content.....	68	13.1 Tipos de buscadores.....	82
9.3.4.4 Change frequencies.....	68	13.2 MetaDatos.....	82
9.3.4.5 Priorities.....	69	13.2.1 MetaDatos Internos.....	82
9.3.5 Pagebar2.....	69	13.2.2 MetaDatos Externos.....	83
9.3.5.1 Basic settings.....	69	13.2.3 Transcripción y Subtítulos.....	83
9.3.6 WordPress Related Posts.....	70	13.3 Promoción.....	83
9.4 Plantillas (Diseño).....	70	13.3.1 Media RSS.....	84
9.4.1 Encabezado (header.php).....	70	13.3.2 Sitemaps de vídeo.....	84
9.4.2 Pie de página (footer.php).....	71	13.3.3 Alta en buscadores.....	84
9.5 WYSIWYG o no WYSIWYG.....	71		
9.6 Blogroll.....	72	14 SEO para móvil	85
		14.1 XHTML Mobile.....	85
10 GeoSEO (SEO Local)	73	14.2 Alta en buscadores.....	86
10.1 Dominios, SubDominios o Carpetas.....	73	14.3 robots.txt.....	86
10.1.1 Dominios.....	73		
10.1.2 SubDominios.....	73	15 SEO para PPC (Pago por Clic)	87
10.1.3 Carpetas.....	73		
10.2 El dominio y la IP.....	74		
10.3 Mapas y Negocios.....	74		

16 SEO para PDF (Portable Document Format)	89	19 Historia de los Buscadores	98
17 Cómo funciona un buscador	91	19.1 De 1993 a 2008.....	98
17.1 Creando un índice.....	91	19.2 Tabla Resumen de la Aparición de los Buscadores.....	105
17.2 Creando unos resultados.....	92	20 Diccionario	106
17.3 Qué tendrán los buscadores.....	93	21 Sentido Común	110
18 Herramientas	94	21.1 Convertirse y entrenar a SEOs.....	110
18.1 Los buscadores.....	94		
18.1.1 Google Webmaster Tools.....	94		
18.1.2 Live Search Webmaster Tools.....	95		
18.1.3 Yahoo! Site Explorer.....	95		
18.2 Palabras Clave.....	96		
18.3 Ranking en Internet.....	96		
18.4 Seguimiento del usuario.....	96		
18.5 Notas de prensa.....	97		

A

CAPÍTULO

0

Antes de empezar a leer

Si estás leyendo esta guía es probablemente porque estás interesado en mejorar el posicionamiento natural de tu sitio. Antes de empezar a leer te recomiendo que te hagas la pregunta fundamental:

¿Por qué quieres que la gente visite tu sitio web?

Con esta pregunta simplemente quiero hacerte ver que el posicionamiento tiene un fin y que salir primero en los buscadores ha de tener alguna utilidad... sino no valdría la pena molestarse en salir primero... ¿verdad?

Ahora, si estás convencido de querer saber los consejos para salir bien posicionado en los buscadores, sólo has de leer con calma y aprender.

0.1 Historia del posicionamiento en buscadores

El SEO no es algo que se haya iniciado hace un par de años como muchos piensan, existe desde mucho antes de que apareciera Google en escena. En sus inicios la única forma de aparecer en los buscadores era gracias a un botón *Add URL* que había al final de las páginas de muchos buscadores como Altavista.

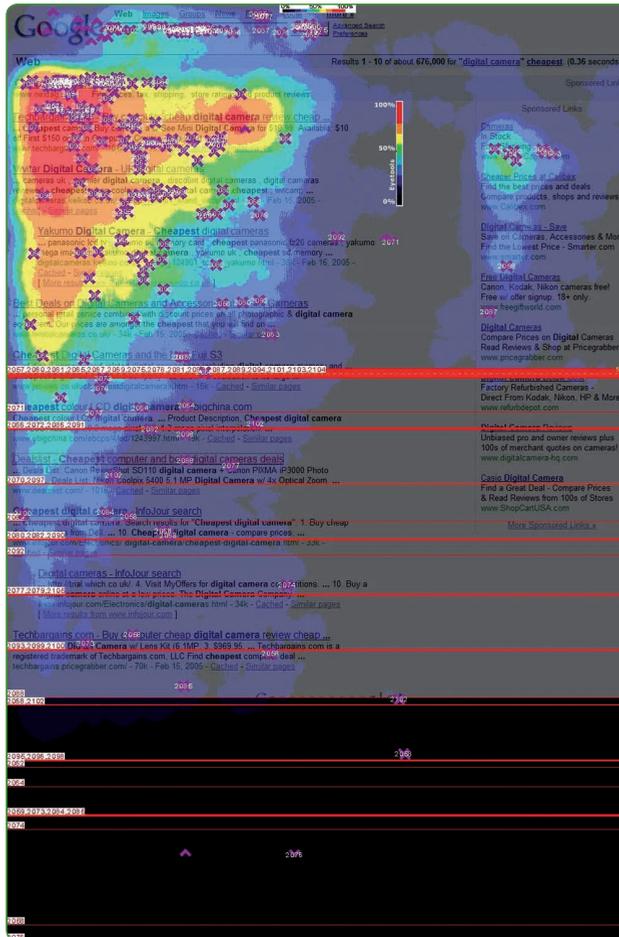
En el año 2000, debido al boom que sufrió Internet, tanto en usuarios como por la burbuja que se produjo en aquellos años, muchas empresas se dieron cuenta de que invertir en publicidad a grandes precios no era rentable y comenzaron a investigar cómo aparecer entre los primeros resultados de los buscadores, ya que todos aquellos buscadores en los que se podía comprar los primeros puestos estaban dejando de funcionar. Ahora todo era algoritmos automáticos y, como cualquier fórmula matemática, se tenía que poder realizar ingeniería inversa y sacar cuáles eran los factores para aparecer.

Hay que tener en cuenta que, según los últimos estudios, un 95% de los usuarios de Internet en España buscan... ¿pero realmente encuentran?. Además, un 4% de las visitas que llegan por los buscadores se convierten en clientes fieles y 1/3 de los que buscan creen que si estás arriba en los resultados de búsqueda es que eres relevante o líder del sector.

Eso sí, hoy en día si sumamos el tráfico que genera Ask¹, Google², Exalead³, Live⁴ y Yahoo!⁵ tenemos el 99% de la cobertura en la mayoría de países, por lo que el posicionamiento hay que pensarlo para esos 5 motores de búsqueda ya que el resto de los que haya utilizará técnicas de indexación y ranking similares o estos mismos buscadores.

0.2 Por qué aparecer primero, segundo o tercero

Uno de los estudios que se hacen sobre los resultados de búsqueda son los de *eye-tracking*, basándose en los puntos en los que una persona mira la pantalla cuando realiza una consulta.



Como puedes ver, tampoco hay mucho más que decir que aparecer entre los 3 primeros resultados te asegura la mayor parte del tráfico de esa consulta.

- (1) <http://www.ask.com/>
- (2) <http://www.google.com/>
- (3) <http://www.exalead.com/>
- (4) <http://www.live.com/>
- (5) <http://www.yahoo.com/>

0.3 Nadie puede asegurar tu lugar

Muchas empresas que se dedican al posicionamiento te aseguran que son capaces de situar tu sitio y mantenerlo allí por una cuota mensual. Otras te ofrecen un servicio mediante el cuál añades un código en el sitio y automáticamente sales posicionado... ¡CUIDADO!. Lo mejor es que el posicionamiento lo haga el propietario del sitio web con su equipo de desarrollo propio o de confianza, aunque, lógicamente, esto es sólo una recomendación para que no tengas problemas futuros.

Existe una desconfianza en el mundo SEO debido a aquellos que han querido aprovechar muchos rumores y desconocimiento, por lo que la primera fase antes de comenzar cualquier estrategia es aplicar una mínima serie de conceptos básicos que te ayudarán a dar los primeros pasos.

0.4 Lo importante son los contenidos

Un sitio web ha de ser dinámico, ha de cambiar cada cierto tiempo... La época en la que únicamente había páginas corporativas que no se actualizaban en meses ha finalizado. Ahora lo que se llevan son los contenidos dinámicos, la actualidad y sobretodo los contenidos originales...

Ten en cuenta que aunque pueda parecer algo obvio, los motores de búsqueda quieren tener contenidos que no estén duplicados, por lo que el factor más importante es el que puedas aportar a la red como generador de contenidos, ya sea en formato textual, de imágenes, vídeos...

0.4.1 Contenidos Originales

El secreto de un sitio web son sus contenidos y principalmente que estos contenidos sean originales, básicamente porque ¿de qué me sirve tener contenidos que otros ya han publicado? Además los buscadores saben encontrar cuál de todos es el contenido original y, posiblemente, llegarán a penalizarte por haberlo copiado.

Un sitio se vuelve importante debido a este tipo de contenidos, a generar información que, al fin y al cabo, es lo que los buscadores quieren: organizar la información mundial (tal y como la filosofía de Google⁶ se presenta).

Conseguir esto es muy sencillo. Un sitio se puede hacer por muchas razones... puedes tener la web de tu empresa que, lógicamente, tiene sus propios contenidos ya que ofrecerá los productos o servicios propios; podrías tener tu medio de comunicación que puede dar información de otros pero siempre adecuándola a tu manera de ver lo que ocurre y lo que quieres informar; también puedes hacer un blog y en este caso lo más probable es que expliques tus experiencias y lo que te ocurre, es decir, tus propios contenidos... Ahora párate a pensar... ¿de qué sirve explicar algo que ya han explicado otros y encima más tarde sin aportar nada nuevo? Los usuarios no son tontos y acaban dándose cuenta de que no están en la fuente original...

(6) <http://www.google.com/intl/es/corporate/tenthings.html>



0.5 Qué debes saber antes de ser...

Antes de comenzar a ponerse a estudiar todo lo referente al posicionamiento en buscadores es bueno tener una formación previa en:

▶ **HTML**

Tener conocimientos avanzados de HTML (es decir, conocerlo completamente) y saber editarlo desde un editor de texto plano es básico.

▶ **Programación**

Aunque no es indispensable, hoy en día es básico saber PHP o ASP y conocer algún tipo de base de datos como MySQL o SQL Server.

▶ **Buscadores**

Hay que saber cómo funciona un buscador de forma general: indexación, recuperación de información, ordenación, algoritmos...

▶ **Redacción**

Es importante saber escribir bien y tener unos mínimos conocimientos editoriales (es muy recomendable tener un blog personal y actualizarlo).

0.6 ¿Por dónde empezar?

Antes de comenzar a optimizar el sitio web en cuestión, debes aplicar una serie de pasos y estados para saber a qué te enfrentas, qué es lo mejor para el sitio, ver la globalidad del sitio en Internet.

0.6.1 A quién va dirigido el sitio web

Lo primero que has de hacer antes de empezar a plantearte el sitio u optimizarlo es definir elementos varios como qué formación, idiomas o edad tienen los usuarios potenciales, si quieres realizar B2B (vender a empresas) o B2C (vender a clientes finales), geolocalizar el sitio (ámbito internacional, nacional, regional, local...), si vas a aplicar un único o varios idiomas...

0.6.2 Qué buscan para encontrarte

¿Te has preguntado alguna vez qué busca la gente para encontrar tu sitio web? Te recomiendo que sigas unos pasos para saberlo:

▶ **Listado propio de palabras**

Que la persona que solicita el estudio realice una lista de palabras por las que cree que debe aparecer. Esta lista puede incluir palabras genéricas, productos, etc...

▶ **Listado de amigos y conocidos**

Por ejemplo, si en la empresa hay varios trabajadores solicítales una lista de las palabras clave por las que cree que debe aparecer. Igualmente a clientes, amigos y conocidos también se les podría pedir la lista.

▶ **Listado de herramientas de palabras clave**

En la red hay herramientas en las que les puedes dar una página web y te sacan un listado o les das una palabra genérica y te devuelven listas de palabras relacionadas.

▶ **Logs del servidor + estadísticas de referentes**

Los servidores web guardan todo lo que ocurre en los sitios web y ahí se puede comprobar habitualmente las palabras clave por las que los usuarios llegan actualmente al sitio.

▶ **Qué palabras usa la competencia**

Si la competencia está usando determinadas frases de búsqueda hay que pensar si lo está haciendo a conciencia o sin querer, y tener en cuenta las mismas.

▶ **El “long tail”**

La larga cola es aquella lista de palabras (de combinaciones) que si las sumamos, aunque sean pocas búsquedas, pueden suponer una gran cantidad del total de las búsquedas realizadas frecuentemente.

El realizar este listado de palabras es para darte una idea de qué conceptos son los que la gente relaciona con tu sitio y no palabras por las que debes posicionarte... no existe el posicionamiento de palabras clave, existe el posicionamiento de sitios web.

Hay que tener en cuenta que elegir una combinación de palabras u otra puede llevar a tener hasta un 20% más de tráfico, así que cuida mucho las que decidas utilizar.

0.6.3 Errores comunes

Habitualmente cuando se quiere posicionar un sitio es porque ya tenemos uno hecho. Con frecuencia los sitios son complicados de modificar pero vale la pena aplicar ciertos cambios contra los errores más comunes:

▶ **Poner en la página principal un Flash**

Los robots de búsqueda no saben indexar Adobe Flash⁷ (al menos no como quisiéramos) y por ello no es bueno comenzar una página web con esta tecnología. Esto no implica que no se pueda utilizar para determinados detalles, como algún botón o similar dentro de la web y a modo decorativo, pero hay que recordar que la web es hipertextual y el Flash es multimedia.

▶ **Banderas de idiomas**

Algo que no le gusta a los motores de búsqueda es que la página principal sea una lista de enlaces a varios idiomas. Lo ideal es que una página esté en un idioma y, si hay que añadir más, se haga en subdominios con el código del idioma. Por ejemplo, nuestro www.pagina.dom podría ser en español y podríamos crear ca.pagina.dom para catalán, eu.pagina.dom en euskera, en.pagina.dom en inglés... Otra opción es tener un dominio específico para cada idioma / país (www.pagina.es, www.pagina.fr...).

▶ **Sitios web con frames (o marcos)**

Antiguamente, debido a que no había lenguajes de programación web avanzados, era muy común realizar páginas con “frames” (o marcos) de manera que siempre teníamos una parte del sitio estático y otro que mediante enlaces variaba... Estas páginas son complicadas de posicionar ya que se indexan por separado, cosa que complica el trabajo a los motores de búsqueda.

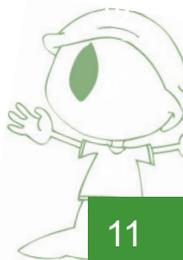
▶ **Sitios web pesados**

Con esto me refiero al tamaño de las páginas (al HTML). Los robots de búsqueda sólo quieren la “chicha”, los contenidos, por lo que es interesante aplicar nuevas tecnologías al mostrar las páginas por pantalla, como el XHTML y CSS, en contra de recurrir al anticuado HTML. Estas nuevas tecnologías suelen dar mejor resultado a la hora de indexarse.

▶ **La página sólo tiene navegación por imágenes**

Es muy habitual que, por efectos especiales, se hayan incluido en las opciones y navegación imágenes en contra de texto. Esto hace que una web sea más complicada de indexar ya que, como comentaba antes, la web es hipertextual. Otra opción que suele haber es la del menú hecho con Javascript que tampoco ayuda a la indexación.

(7) <http://www.adobe.com/es/products/flash/>



▣ Ventanas emergentes (o popups)

Los robots no suelen ser capaces de leer la información que hay en páginas que se abren en ventanas emergentes. No hay que confundirlo con abrir en una ventana nueva, que sí sería indexable.

▣ Dejar las páginas con un mismo título o sin él

Teniendo en cuenta que en los resultados de búsqueda es donde el usuario pulsa, es importante que sea correcto.

Este es un primer resumen de las cosas que no se deberían hacer en sitios si queremos que sea muy indexable.

0.7 Rankings: Factores

En la actualidad cada motor de búsqueda tiene sus propios factores de posicionamiento en sus resultados. Los primeros buscadores se basaban en el contenido de los meta-tags, más adelante en el contenido de la página y ahora se tienen en cuenta tanto los factores internos como los factores externos.

Los primeros son los que tenemos bajo control ya que responden a lo que podemos hacer en el sitio web... títulos, hosting, contenidos, tecnología...

Los segundos corresponden a los que no tenemos bajo control total ya que se encuentran fuera de nuestro sitio y que principalmente son los enlaces externos que otros sitios puedan hacer sobre nuestras páginas.

0.7.1 Navegabilidad

Un detalle que hay que tener en cuenta, aunque no afecte al posicionamiento, es el de la usabilidad del sitio. Cuando alguien llega desde un buscador hay que darle la mayor cantidad de facilidades para que se quede y no cierre el navegador o se vaya a otro resultado de búsqueda. Por esto hay que hacer hincapié en la usabilidad del sitio.

Lo primero que hay que revisar es la página principal. Ésta no ha de ser una página de esas con un Flash y un continuar porque perdemos navegabilidad y, lo peor, si un usuario llega buscando algo y en esa página no está lo que busca se marchará, lo que significa que los esfuerzos de posicionamiento no habrán servido de nada.

Lo ideal es que la navegación esté en un lugar “habitual” para los usuarios, es decir, en los laterales (visible en el primer “pantallazo”) o en la parte superior de forma horizontal. Los enlaces deberían ser como ya he comentado: hipertextuales.

Otro tema que hay que tener siempre en cuenta al hacer un sitio web (tanto para el usuario como para los robots de búsqueda) es que no sea necesario el uso de las “flechas” de adelante y atrás que tienen los navegadores. Hay que tener en cuenta que un robot de búsqueda no tiene esos botones y que un “back” en Javascript tampoco sirve, lo que nos da pie a que sea “casi obligatoria” una forma sencilla de volver a la sección “padre” de la que nos encontramos. Por ejemplo los blogs que de forma habitual, cuando entras en un artículo, siempre tienen enlaces a las categorías en donde está además de tener el logo o título enlazado a la página principal.

Estos detalles son importantes ya que si un buscador o un usuario llega a una página concreta del sitio, que no la principal, ha de tener todos los medios necesarios para poder navegar el sitio de la mejor forma posible y de estar enlazada a toda la estructura del mismo.

De todo esto se extrae un conclusión importante: el sitio web ha de estar muy bien estruc-

turado, en forma de árbol principalmente, para que todo esté organizado de la mejor forma posible: hay que organizar la información (página principal, categorías, subcategorías, contenidos...).





CAPÍTULO

1

Situación actual

Antes de empezar a optimizar el sitio web necesitas saber cómo se encuentra en los motores de búsqueda. Para ello debes ver varios factores y utilizar las herramientas para webmasters de los buscadores.

1.1 Páginas indexadas

La cantidad de páginas indexadas (o saturación en los buscadores) se refiere a la cantidad de páginas que hay en el buscador en relación a un dominio en particular.

El posicionamiento en buscadores es un juego de números en el cual cada página es potencialmente una página de aterrizaje para el tráfico que proviene de los buscadores. El volumen de tráfico es directamente proporcional a la cantidad de páginas indexadas por los buscadores. Por ende, la cantidad de palabras clave por las cuales los usuarios pueden encontrar un sitio web dependerá de la saturación misma.

También hay que tener presente un detalle que quizás es el más importante para el que usaremos estos valores: cada buscador tiene una cantidad de páginas indexadas, y tú puedes calcular aproximadamente la cantidad de páginas reales que tiene tu dominio... si las cifras varían mucho, es probable que haya un problema grave a la hora de revisar tu sitio, por lo que será un factor a tener en cuenta más adelante. Puede que con la lista de problemas habituales de la que hablé antes hayas detectado los elementos que hacen que no se indexe completamente.

Por estos motivos, la saturación es un elemento a tener en cuenta a la hora de estimar el desarrollo de una campaña de optimización y deberás trabajar en ello para seguir aumentándola con el tiempo.

▶ Ask

<http://es.ask.com/web?q=site:dominio.ext+dominio.ext>

▶ Exalead

<http://www.exalead.es/search/results?q=site:dominio.ext>

▶ Google

<http://www.google.es/search?q=site:dominio.ext>

▶ Live

<http://search.live.com/results.aspx?q=site:dominio.ext>

▶ Yahoo!

<http://search.yahoo.com/search?p=site:dominio.ext>

Uno de los errores habituales al hacer un cálculo de estas páginas es la diferenciación entre sitios “con www” y sitios “sin www”. Más adelante trataré de forma más detallada como hacer que esta diferencia no se produzca y así evitar duplicidad de páginas y posibles penalizaciones. Recuerda que las páginas “con www” o “sin www” pueden ser completamente distintas, pero si son iguales hay que “decírselo” a los buscadores para que no las tengan duplicadas.

Un detalle que te puedes encontrar es que al hacer la consulta en Google (principalmente), te responda al pie de página de los resultados más avanzados (página 5-10) con el mensaje: *Para mostrarle los resultados más pertinentes, omitimos ciertas entradas muy similares a las que ya hemos mostrado*. Esto es debido principalmente a que los títulos del sitio no están optimizados y la forma de solucionarlo es que cada una de las páginas tenga su propio título, como verás más adelante, o que haya encontrado contenido duplicado.

Un primer salto para que un sitio se considere lo suficientemente importante es llegar a las 1.000 páginas indexadas de contenido, momento en el cual notarás un salto significativo en la recepción de tráfico desde los buscadores.

1.2 Enlaces entrantes

Los enlaces entrantes (o popularidad de los enlaces) ha evolucionado hasta ser uno de los factores más importantes para determinar hasta dónde puede posicionarse un sitio, aunque últimamente está volviendo a perder valor, influyendo principalmente en la posibilidad de que tu sitio se convierta en fresco cuantos más enlaces se consigan, algo que hará que los robots te visiten con más frecuencia e indexen mayor cantidad de páginas. Los buscadores se encuentran constantemente intentando eliminar resultados no válidos de sus índices para proveer al usuario con resultados más relevantes.

Históricamente los buscadores ordenaban sus rankings según lo que se encontraba en cada página, es decir, los factores internos. Los metadatos resultaban ser lo más importante. Pero debido a la explotación por parte de *spammers*, su importancia ha disminuido y los buscadores comenzaron a darle más importancia a factores externos, es decir, a los enlaces y la información que estos proporcionan, aunque, como digo, ahora se está volviendo a ir hacia atrás, ya que los sistemas antispam de contenidos funcionan mucho mejor y los de enlaces peor.

La popularidad de enlaces se basa en la cantidad de enlaces que apuntan a un sitio web en particular y que son encontrados e indexados por los buscadores. En general los sitios populares tienden a tener contenido de calidad lo que genera resultados relevantes para los buscadores. Con este fin los buscadores asignan mejores posiciones a los sitios que gozan de mayor popularidad de enlaces. Sin embargo es importante aclarar que el hecho de tener muchos enlaces apuntando a una web no es una garantía. Los buscadores analizan también la importancia relativa de cada uno de estos enlaces. Por lo tanto no vale lo mismo un enlace desde, por ejemplo, un sitio de una universidad que de una página personal.

▶ Exalead

<http://www.exalead.es/search/results?q=link:dominio.ext>

▶ Google

<http://www.google.es/search?q=link:dominio.ext>

▶ Live

<http://search.live.com/results.aspx?q=link:dominio.ext>



▣ Yahoo!

<http://search.yahoo.com/search?p=linkdomain:dominio.ext>

Como puedes observar, la cantidad de enlaces que registra Google para cada uno de los sitios varía desproporcionadamente con respecto a lo que indica Yahoo!, que habitualmente es el que indexa mayor cantidad de información. La razón es que Google no indica con certeza el número exacto de enlaces, por el valor competitivo que estas estadísticas significan, aunque se estima que el valor que muestra es cercano al 10% del total.

Si la cantidad de enlaces es inferior a 5.000, el sitio deberá continuar atrayendo enlaces, ya que mejorar la popularidad es una tarea constante. Se podría estimar que la popularidad de enlaces se presenta como un problema principal para el sitio, por lo que la optimización de los factores internos (diseño, código, contenido, etc) no es una garantía para obtener un mejor posicionamiento.

1.3 Herramientas para Webmasters

Los 3 grandes buscadores disponen de unas herramientas que nos pueden proporcionar una información muy interesante, incluso avisarte si tu sitio se encuentra penalizado, de forma que puedes resolver la situación gracias a sus sistema de contacto.

▣ Google

<http://www.google.es/webmasters/>

▣ Live

<http://webmaster.live.com/>

▣ Yahoo!

<http://siteexplorer.search.yahoo.com/>

En cada una de las herramientas puedes añadir los sitios web que quieras y tras eso cada sistema te hará validar que realmente te pertenecen para poder controlar que los datos que se ofrecen sean realmente a los propietarios.

Además, lo interesante de estas herramientas es que dan las cifras exactas de los valores de páginas indexadas y enlaces entrantes, lo que hará que no sea necesario recurrir a las direcciones web que he puesto antes.



Factores Internos (on-page)

CAPÍTULO

2

Tu sitio web es el escaparate de lo que eres y ofreces a millones de internautas alrededor del mundo, y como alrededor del mundo hay muchas formas de ver las cosas, lo que has de hacer es tener en cuenta eso y plantear un sitio web que sea comprensible para todos, que utilice tecnologías lo más estándar posible siempre pensando en el usuario tipo, además de las muchas excepciones que pueda haber, ya que esas excepciones son las que harán del sitio el mejor, entre los cuales hay que incluir personas de visibilidad reducida, motores de búsqueda, navegadores de todo tipo...

Si lo importante son los contenidos, aquí tienes la forma de ponerlos en el sitio web de forma óptima. Hay muchos factores a tener en cuenta pero sobretodo has de aplicar el sentido común ya que es imposible tener una página optimizada al 100% y hay que sopesar qué es mejor aplicar en cada momento. No hay que olvidar que tampoco debes sobreoptimizar tu sitio ya que podría ser contraproducente.

2.1 Dominio

Uno de los elementos más importantes de un sitio web es su dominio. Si quieres que tenga algún peso en posicionamiento es interesante que contenga la palabra clave principal del proyecto. Esto no significa que sea imprescindible, pero ayudará a mejorar la posición en la mayoría de motores de búsqueda.

Ante la elección entre un dominio con o sin guiones, lo mejor es no usar ningún tipo de elemento “extraño” en la dirección, y en caso de ser necesario, sólo tenga como máximo un guión.

2.1.1 Territorialidad

Los buscadores, cada vez más, están dando más importancia a la geoposición de un dominio, es decir, que si la información de un dominio está enfocada a un país en concreto, este dominio aparecerá mejor posicionado en la edición local de ese buscador.

Esto significa que si el contenido de un dominio está enfocada a España, lo mejor es usar un dominio .ES, que si está enfocada para Estados Unidos usemos un dominio .US o si está enfocada para Grecia usemos un .GR.

Esto no significa que no podamos usar los dominios .COM o .NET para incorporar

información, pero en estos casos los haremos de una forma más genérica.

Al fin y al cabo... ¿para qué existen dominios territoriales si luego no tienen ninguna influencia en los usuarios o los motores de búsqueda?

2.1.2 Edad

Que un dominio lleve mucho tiempo registrado y en línea es un buen elemento a tener en cuenta para que se indexe de forma correcta y se tenga en cuenta como lugar destacado de internet.

Aún así, aquellos que compran un dominio pensando en mantener su edad histórica, han de saber que hoy en día la mayor parte de buscadores también son registradores de dominios, de forma que tienen control absoluto sobre los cambios de titularidad o que haya en los mismos.

Fuentes internas de Google dicen que si un dominio cambia de titular, ese dominio queda re-seteado en su edad y el motor lo tratará como uno nuevo. Aún así, los resultados de búsqueda por ahora no indican eso.

En caso de que tengas la posibilidad de adquirir un dominio antiguo, es muy interesante ver qué había en él históricamente y para ello puedes usar la herramienta Wayback Machine¹ donde verás de forma temporal el contenido de un sitio a modo de caché.

2.1.3 Renovación

De la misma forma que los motores saben controlar la edad de un dominio, también saben la fecha de registro y de caducidad del mismo.

Hoy en día, y por el precio que tienen, tener un dominio que se vaya a utilizar de una forma segura y con un producto de calidad es interesante mantenerlo registrado a 5-10 años vista, dándole pie al motor de búsqueda que sepa que es un proyecto estable y con previsión de futuro.

2.1.4 Dominios Aparcados

En la actualidad cuando se compra un dominio que no está pensado desarrollarse es posible establecerlo en sitios donde, al entrar, aparece una página en la que únicamente hay enlaces de pago.

Aunque este sistema puede ser rentable, en el momento en el que se decida lanzar un proyecto bajo este dominio es probable que te encuentres con que los robots de búsqueda no pasan a indexarlo y mucho menos posicionarlo.

En caso de encontrarte en esta situación lo mejor es comenzar a generar contenido y cuando hayan pasado al menos 6 semanas dar de alta el dominio en los buscadores de nuevo, crear Sitemaps y comenzar un proceso de reinclusión², como el que permite Google.

(1) <http://web.archive.org/>

(2) <http://www.google.com/webmasters/tools/reinclusion>

2.2 Programación

Cuando se va a empezar un sitio web lo primero que hay que tener claro, a parte de su finalidad, es cómo lo vamos a hacer. No es lo mismo hacer una página web con Microsoft Expression³, con Adobe Dreamweaver⁴ o con un editor de texto plano, básicamente porque cada uno de esos programas deja el código estructurado de una forma diferente y el hecho de que facilite la estructura a los buscadores va a ayudar a que nos considere un sitio amigo.

Otro tema interesante es que da igual que el sitio esté desarrollado en HTML plano, o que tenga programación en PHP, ASP o JSP por ejemplo, ya que, al final, el contenido que aparece por pantalla sigue siendo HTML y no programación, por lo que harás el sitio web siempre con la herramienta que te funcione mejor o consideres más adecuada para cada ocasión, ya que un sitio puede tener desarrollos en varios lenguajes de programación.

Una de las cosas que agradecen los buscadores es que en una página web haya mucha información (sin pasarse, claro), y por ello no tiene ningún sentido que hagamos una página con 3 tablas una dentro de otra si al final tenemos realmente un 70% de código en HTML y un 30% de información. Lo que has de conseguir es que haya un 30% de código y un 70% de información. Otra cosa que es de agradecer es que las páginas se puedan ver bien en cualquier navegador y que pasen (al menos de forma razonable) los validadores de HTML, de CSS...

Hoy en día hay 3 formas de mostrar la información por pantalla: HTML, XHTML y XML.

El HTML es la base de una página web, de eso no hay duda, y es como se han estado haciendo los diferentes sitios web desde que apareció el HTTP, al menos hasta que empezaron a salir las primeras cosas en XHTML y los CSS. Hoy en día la mayoría de proyectos nuevos se plantean en XHTML, a menos que sea un proyecto de futuro que se puede plantear en XML.

¿Por qué es importante elegir uno u otro? Muy simple: con cada uno de ellos se puede estructurar la información de una manera u otra. Digamos que en HTML tienes unas etiquetas predefinidas y no puedes jugar casi nada con la información... sólo la puedes estructurar. En cambio el XHTML te da algo más de juego... te permite, con muchas menos etiquetas, organizar algo más la información y con la ventaja de que el código suele ocupar mucho menos que con HTML.

Claro está que lo ideal sería poder hacer un sitio web en XML. No todos los sitios lo necesitarían pero al poder inventarte tus propias etiquetas ganarías en la llamada Web Semántica. Puedes decir realmente qué es lo que hay en un contenido, agruparlo, estructurarlo... Eso sí, en la actualidad, el XML es mucho más complejo de indexar y almacenar por los motores de búsqueda que el XHTML ya que al no tener etiquetas establecidas, no puede asignarles un valor a cada una de ellas.

Lo mejor es optimizar los recursos y aplicar XHTML 1.1 y CSS 2.1, soportada por todos los navegadores, y comenzar a plantearte el soporte futuro a HTML 5.0, XHTML 2.0 y CSS 3.0, sobretodo en el momento en que se consideren estándar.

(3) <http://www.microsoft.com/expression/>

(4) <http://www.adobe.com/products/dreamweaver/>



2.3 La cabecera de la página web

Uno de los puntos interesantes de un sitio web es su título y descripción. Entre 1995 y 1998, cuando los robots de búsqueda sólo eran capaces de leer las cabeceras de un sitio web, era muy importante poner todo tipo de metatags. Se ponía prácticamente de todo aunque hoy en día ya no es necesario porque los robots leen el contenido de toda la página.

La cabecera de cada página es lo primero que leen los robots de búsqueda y que van a decidir la estructura de la página y su configuración, por lo que es muy importante dedicarle una gran atención.

El código HTML de una cabecera debe ser similar a éste:

```

1.  <!DOCTYPE html PUBLIC "-//W3C//DTD XHTML 1.1//EN"
    "http://www.w3.org/TR/xhtml11/DTD/xhtml11.dtd">
2.  <html xmlns="http://www.w3.org/1999/xhtml">
3.    <head>
4.      <title>Titulo del Sitio Web</title>
5.      <meta http-equiv="Content-Type" content="text/html;
    charset=UTF-8" />
6.      <meta name="description" content="Explicación de lo que vamos a
    encontrar en la página web..." />
7.      <meta name="keywords" content="algunas palabras, palabras clave,
    contenidos importantes" />
8.      <link rel="stylesheet" href="style.css" type="text/css"
    media="screen" />
9.      <link rel="alternate" type="application/rss+xml" title="feed"
    href="feed.xml" />
10.   </head>
11.   [...]
12. </html>

```

Si revisas una a una las líneas de código:

- ▣ El **DOCTYPE** es la explicación de que tipo de código HTML / XHTML vas a encontrar en el sitio. Con esto ayudas al navegador a que muestre el sitio de una forma correcta. Es un elemento obligatorio en los sitios web. Además puede llegar a indicar si un sitio está pensado para verse en un tipo de plataforma u otra (como por ejemplo el especificar el formato XHTML para móviles).
- ▣ El **TITLE** contendrá el título de la página. Más adelante hay un capítulo que únicamente trata este elemento.
- ▣ La línea del **META Content-Type** indica la codificación del texto. Habitualmente UTF-8 o ISO-8859-1, teniendo en cuenta que si quieres realizar una web internacional y con soporte de futuro es mejor tenerla optimizada para UTF-8.
- ▣ La línea del **META Description** tendrá una pequeña descripción de lo que vas a encontrar en esa página (en esa en concreto y no en toda la web...). Este texto ha de ser, por tanto, único en cada una de las páginas de todo el sitio. Lo ideal es que sea breve... entre 25 y 150 caracteres.
- ▣ En la línea del **META Keywords** pondrás una relación de 4 a 6 grupos de palabras separados por comas, que tengan relación con el contenido de esa página. Al igual que la descripción, deberían ser únicas para esa página.
- ▣ El **LINK Stylesheet** indica la dirección URL del CSS (Hojas de Estilo). Este parámetro es opcional, aunque se recomienda hacer uso de CSS.

- ▣ El **LINK Alternate** indica la dirección XML del feed (sindicación de contenidos). Este parámetro es opcional, aunque si la página dispone de noticias o información actualizada se recomienda hacer uso del mismo ya que ayudará a los buscadores específicos de blogs.

En principio esto es suficiente para que la cabecera del sitio sea lo más estándar para los robots de búsqueda, navegadores y usuarios.

Gran parte de estos elementos son los que aparecen en los resultados de búsqueda (título, descripción...) y que hay que tenerlos muy estudiados ya que son los que los usuarios encuentran:

título	Museo Nacional del Prado: Bienvenida
descripción	15 Obras Maestras · Visita el Museo · Página principal · Benvinguts · 15 obres mestres · Visita el Museu · Benvido · 15 obras maestras · Visita o Museo ...
dirección URL	www.museodelprado.es - En caché

2.3.1 El título de cada página

Al hacer una consulta en cualquier buscador lo que ves en los resultados son los títulos de las páginas web juntamente a su descripción (o a un fragmento de la página). Por eso el título de cada página debe ser el que necesita y no otro. El título de la página es el lugar donde los usuarios de los buscadores pulsan, por eso se hace el elemento más importante (visible) frente al resto de resultados.

Si comienzas por la página principal del sitio deberías indicar el nombre del sitio y, por ejemplo, el lema de la empresa o una frase que exprese lo que vendes. ¿Por qué? Muy sencillo. Habitualmente si aparece nuestro sitio entre los resultados con la página principal es porque se ha hecho una búsqueda por el nombre de la empresa o de alguno de los productos.

La página principal no es la página que ha de aparecer siempre en los resultados de búsqueda, un error muy común, que intenta considerar aparecer por muchas frases de búsqueda, algo que es complejo debido a la densidad de los contenidos posibles.

Otro tipo de página que puedes encontrar en tu sitio son las secciones o categorías. El título debería ser también muy informativo. Por ejemplo puedes indicar como primera palabra el nombre o lo que hay en esa categoría seguido de “en Nombre de la empresa / Web”. Con esto indicarás el contenido y que lo ofreces tú.

Estas páginas generales son las que hay que intentar posicionar para frases de búsqueda más generales, lo que hace muy interesante optimizar sus contenidos y su título.

Por último tienes la página “final”, la página donde la persona que busca ha de encontrar el producto, servicio o contenido que está buscando realmente. Por eso el título de esta página ha de indicar únicamente el contenido. En este caso no es necesario indicar el sitio web o la empresa donde está ya que se puede ver fácilmente mirando la URL del sitio, o una vez dentro si le interesa ya se informará.

Piensa que cuando alguien busca tu nombre, marca o empresa, lo que ha de encontrar es la página principal, y ese título se lo indicará; si busca algún producto general, un contenido genérico, deberá llegar a la página de la sección o categoría, que, como decía antes, llevará el nombre del sitio al lado para generar *branding*; finalmente, aquel que busque algo muy concreto buscará independientemente del sitio, de forma que cuanto más claro sea el título de los contenidos, mejor.

Si te das cuenta no he comentado en ningún momento que en el título haya que poner una larga lista de palabras clave, y es que el título de la página ha de indicar precisamente eso: qué hay en esta página.



2.3.1.1 La calidad de un buen título

Aunque ya he comentado qué forma debe tener el título, también hay que pararse a pensar el contenido del mismo. Hay que tener en cuenta que las palabras que aparecen primero en el título tiene más valor que las del final, por lo que si quieres darle más fuerza a unas palabras has de intentar ponerlas al principio.

Por ejemplo, podrías plantearte un título que fuera *Añaden nuevos tipos de techo al Mini Cooper*. En este caso lo más probable es que la frase de búsqueda o las palabras más importantes sean *Mini Cooper*. ¿No sería mejor utilizar un título más parecido al siguiente?: *Mini Cooper añade nuevos tipos de techo*. Al fin y al cabo significa lo mismo y dejamos al principio las palabras de más valor.

Tampoco es lo mismo *yo me río en el baño* que *yo me baño en el río*. Con esto quiero decir que no es lo mismo poner Mini Cooper que Cooper Mini, porque los resultados de los buscadores serán diferentes aunque haya algunos resultados comunes.

2.3.2 La descripción de la página

Como decía antes, la línea del *META Description* ha de tener una pequeña descripción de lo que vas a encontrar en esa página (en esa en concreto y no en todo el sitio web). Este texto ha de ser, por tanto, único de cada una de las páginas de todo el sitio. Lo ideal es que sea breve... entre 25 y 150 caracteres, lo suficiente para que aparezca en los resultados de búsqueda, ya que no olvidemos que este elemento es exclusivamente para este uso.

Este elemento vendría a ser como las pequeñas descripciones que aparecen en los periódicos de papel. Primero está el gran titular y luego hay una pequeña frase que describe y resumen el contenido de la página.

2.3.3 Las palabras clave de la página

Un buen META Keywords viene dado de la relación de las palabras con el contenido que hay en la página. Si esa lista de palabras coincide con las que el buscador considera más importantes, le dará un valor superior al sitio. Las palabras clave que hemos de indicar no han de exceder en su cantidad por lo que lo mejor sería una media de 5 grupos de palabras por página.

Hay que tener en cuenta que debes tratar el sitio mediante herencia, o sea, si una palabra clave está en la página principal ya es válida para todo el sitio; has de elegir palabras de más general a más concreto según vayas entrando en secciones o páginas del sitio.

2.4 Los Enlaces

Internet es como una biblioteca gigante y la gran diferencia con una biblioteca de verdad es que en la red los saltos de “un libro a otro” se hacen a través de los hipervínculos o enlaces. Es por esto que si hay algo a tener en cuenta dentro de un sitio web es que sea fácilmente indexable y saber cómo crear un buen mapa de enlaces internos, ya que eso es en lo que se basan los motores de búsqueda, en encontrar enlaces para poder indexar toda la red.

Lo primero de todo es ver cómo hay que crear un enlace. Según los estándares del W3C⁵ se indica que un enlace tiene un posible atributo (opcional) que es el del “title”. Este atributo, como recomendación, es que tenga el título <title> de la página de destino.

(5) <http://www.w3.org/>

Este podría ser un enlace habitual...

```
<a href="http://www.mundijuegos.com/">mundijuegos</a>
```

Aunque, ¿no sería mejor uno así?

```
<a title="Mundijuegos: Juegos MultiJugador" href="http://www.mundijuegos.com/">juegos multijugador</a>
```

El uso de los títulos (que, personalmente, a mi me gusta poner antes de la URL y “lejos” del texto del enlace, precisamente para ayudar a la densidad de palabras clave) se supone que es para ayudar a las personas invidentes y a determinados navegadores. No hay que olvidar que un robot de búsqueda se comporta como un invidente ya que no es capaz de ver el resultado del sitio web sino que ha de imaginárselo y por ello le has de facilitar la ayuda. Esto no supone que haya que hacer webs sólo texto, pero sí que hay que ayudar a cualquier persona a que la navegación sea correcta.

Volviendo a los enlaces, una vez sabemos cómo hay que construirlos, hay que dejar claro que la navegación es fundamental en un sitio web. Nadie ha de tomárselo literalmente pero queda prohibido el uso de “popups” o “JavaScript” dentro de la parte donde se pone la URL. Además, hay que dar facilidades al usuario para que siempre pueda tener una especie de “línea de navegación” de forma que pueda volver a la categoría padre de la sección del sitio donde se encuentre, y así facilitar una estructura del sitio.

Es muy habitual en algunos sitios ver esto:

```
Inicio -> Noticias -> Buscadores -> Google -> Imágenes
```

Cada una de estas opciones o palabras deberían estar enlazadas a su categoría o sección correspondiente para que el usuario siempre sepa en todo momento dónde está y cómo poder volver “atrás” sin necesidad de usar el botón del navegador... Recuerda que los robots que rastrean la web no tienen ese botón y no pueden llegar a una página en la web en la que no se pueda volver atrás sin un enlace.

El hecho de tener unos enlaces lo más cuidados y trabajados posibles van a hacer que la indexación del sitio sea óptima.



2.5 Cabeceras, Negritas, Cursivas, Listados...

No hay que olvidar nunca que los motores de búsqueda, por ahora, lo que priorizan son los contenidos diferentes de cada página de un sitio web. Por esto, además de tener buenos contenidos, hay que pensar en darles un buen formato y aplicar el sentido común, ya que si ofrecemos a los buscadores un contenido estructurado y con unas etiquetas correctas, se priorizarán las palabras o frases marcadas.

Para empezar, todo texto debería tener un título. Este título, al ser lo más importante de la página (del texto ó contenido que incluyamos), debería ser el tradicional <h1> del HTML. Las cabeceras van del <h1> al <h6> pero habitualmente sólo se usan del <h1> al <h3>.

Teniendo en cuenta se según los estándar de usabilidad sólo puede haber un único <h1> por página, lo añadiríamos junto a una breve entrada o bloque resumen y el resto se puede intentar “seccionar” con bloques <h2> y <h3> como si fueran subsecciones del texto principal, ya que estos no tienen límites. No es bueno abusar de esto (como de nada) por lo que harás un uso responsable y sólo lo usarás si es necesario.

De la misma forma que pones títulos a tus textos, también harás uso de las negritas, cursivas... en HTML hay muchos sistemas de formateado de texto (se supone que el HTML se hizo para eso) por lo que deberías también utilizarlos si se diera el caso.

- ▶ **<abbr>**
indica que el texto es una abreviatura (WWW, HTTP, URI...).
- ▶ **<acronym>**
indica que el texto es un acrónimo (WAC...).
- ▶ **<address>**
indica una dirección postal.
- ▶ ****
es la etiqueta que indica fuerza en un texto (negrita). Se usa también .
- ▶ **<big>**
marca que el texto ha de tener un tamaño mayor.
- ▶ **<blockquote>**
se utiliza de forma muy similar al <cite>.
- ▶ **<cite>**
contiene una cita o referencia de otro texto.
- ▶ **<code>**
indica que el texto es un fragmento de código informático.
- ▶ ****
marca que el texto ya no tiene valor. Suele acompañar a <ins>.
- ▶ **<dfn>**
marca la definición de una palabra o frase.
- ▶ ****
es la etiqueta para indicar énfasis en un texto (cursiva). Se puede usar también <i>.
- ▶ **<j>**
es la etiqueta para indicar énfasis en un texto (cursiva). Se puede usar también .
- ▶ **<ins>**
marca que el texto es nuevo con respecto al original. Suele acompañar a .
- ▶ **<kbd>**
indica texto que el usuario debe introducir en su ordenador.
- ▶ **<pre>**
para mostrar texto sin formato.

- ▶ **<q>**
muy similar a **<cite>**, pero para fragmentos muy cortos (normalmente en una misma línea).
- ▶ **<samp>**
código que daría un ordenador.
- ▶ **<small>**
marca que el texto ha de tener un tamaño menor.
- ▶ ****
es la etiqueta que indica fuerza en un texto (negrita). Se usa también ****.
- ▶ **<sub>**
indica que el texto está por debajo del normal.
- ▶ **<sup>**
indica que el texto está por encima del normal.
- ▶ **<tt>**
marca que el contenido tiene formato teletipo.
- ▶ **<var>**
marca el texto como si fuera una variable.

Es muy interesante que hagas un uso responsable de cada uno de ellos en la medida de las necesidades del texto, ya que si damos la importancia a una palabra o fragmento del texto, estará mucho mejor valorado y puntuado por las arañas de indexación sobretodo si la búsqueda puede tener alguna relación con los contenidos que allí se muestran.

De igual forma, también se pueden realizar listas de elementos, ya sean listas ordenadas ****, desordenadas **** o de definición **<dl>**.

2.6 Las imágenes

Lo mismo que el texto, una imagen es un contenido y también es importante. Hoy en día la mayoría de motores de búsqueda tienen un buscador específico de imágenes que siempre vienen bien para atraer tráfico, ya que se estima que entre un 5% y un 10% llega desde este tipo de búsquedas.

En las imágenes debes tener dos detalles siempre presentes: el nombre de la imagen y describirla correctamente.

El nombre de la imagen es el nombre que se le da al fichero que se publica en la red... no es lo mismo que una imagen sobre la reproducción de las ranas se llame "imagen1.jpg" que "reproduccion-ranas.jpg". Aunque no lo parezca, además del nombre, que es mucho más significativo y tiene incluidas las palabras clave, es mucho más "razonable" para que una persona vea el nombre del fichero y sepa de qué trata, lo que hará que el usuario tienda a pulsar sobre la imagen.

La descripción viene dada para el cumplimiento del HTML y de la usabilidad, sobretodo para personas invidentes que utilizan navegadores poco habituales. Toda imagen ha de llevar, en la actualidad, un texto alternativo. Este texto alternativo es el que aparece en el sitio web si la imagen, por la razón que sea, no puede aparecer. En este caso lo interesante es incorporar palabras clave, al igual que en el nombre del fichero, pocas palabras pero que describan claramente qué es la imagen. Por otro lado, es muy recomendable incluir en cada imagen un título descriptivo -que explique realmente en qué consiste la imagen y que es lo que aparecerá si alguien deja el ratón encima de ella-.

Esto se da con dos elementos: "alt" y "title", que son los mandatarios en estas cuestiones. Los



resultados de búsqueda de imágenes muestran, además de la imagen, una breve descripción de lo que incluye dicha imagen o un fragmento de texto, que suele ser el “title”, y que ayuda a subir posiciones entre los diferentes resultados que puedes encontrar.

```

```

Al igual que en los enlaces el hecho de poner el “title” delante de la dirección de la imagen y tras ella el texto alternativo “alt” ayuda a mejorar la densidad de información, ya que ayuda a recordar que los motores de búsqueda utilizan la información existente alrededor de una imagen para posicionarla, ya que no pueden saber qué contenido tiene la misma (aunque se está trabajando en ello).

2.7 Flash

Puedes encontrarte con sitios web en Flash, no cabe duda, y si bien es cierto que los buscadores pueden leer el contenido de determinados textos dentro de los ficheros SWF, por normal general no sirve de mucho y es difícil de posicionar.

Si de una forma más o menos obligatoria hemos de dejar los contenidos en Flash, lo mejor es que las primeras capas del mismo sean las de texto, y así podrán leerse.

Con respecto a la página que aloja el contenido en Flash, el título y la dirección URL se vuelven mucho más importantes que el resto de elementos. Incluso hay que plantearse separar algunos elementos en diferentes páginas para poder optimizar mejor. Además, a parte de crear enlaces dentro de los SWF, será interesante que los enlaces entre páginas estén disponibles desde el exterior de la página también. Así conseguirás que indexe el robot, de forma más sencilla, todas las páginas del sitio.

Por esta misma razón, el conseguir enlaces ha de basarse en el contenido de los títulos y las direcciones URL. Si las palabras de los enlaces coinciden con los títulos, como he comentado anteriormente, habrá mayores posibilidades de éxito. Eso sí, en estos casos lo mejor es conseguir enlaces de sitios con mucho peso, ya que sino será muy complejo el aparecer entre los resultados de búsqueda.

Aún así, lo más recomendable sigue siendo tener un sitio web en el que no haya Flash. Una posibilidad para hacer sitios es usar SWFObject()⁶ que permite disponer de un bloque de contenido textual y ser sustituido por un objeto Flash. Esta solución sería la óptima a nivel SEO si se utiliza correctamente.

2.8 Los contenidos

El contenido es el elemento de mayor importancia para cualquier sitio web. No importa lo sofisticados se vuelvan los buscadores o cuántos factores externos se agreguen a los algoritmos, en definitiva, lo que juega un rol vital a la hora de los rankings es el contenido textual principalmente, junto al resto de contenidos que se puedan ir añadiendo (imágenes, videos...).

Cada página de un sitio es una potencial página de aterrizaje para los usuarios provenientes de los buscadores. Por ello es esencial que cada página tenga las palabras claves más relevantes para esa página y que no las obviemos.

Es muy habitual querer que la página inicial de nuestro sitio consiga atraer el 100% del tráfico del sitio, pero es mucho mejor distribuir la recepción de tráfico entre varias páginas o

(6) <http://code.google.com/p/swfobject/>

categorías que tengan una temática más específica. Es crucial que el contenido no sólo sea amistoso con los motores de búsqueda, sino que también sea provechoso para la audiencia, es decir el usuario.

Hay que tener en cuenta que si un usuario llega desde un buscador a una página y no encuentra el contenido que busca, automáticamente cierra esa ventana y se vuelve al buscador... en cambio, si encuentra lo que busca y, además, le damos más información relacionada, es probable que se quede un rato navegando por el sitio.

2.8.1 Contenido fresco

Uno de los factores más importantes en la actualidad es el factor frescura ya que juega un rol muy importante que se manifiesta en una patente que le ha sido otorgada a Google y que trata sobre los factores temporales a la hora de medir la actualización de una página. Se puede decir, como regla básica, que cuanto más se actualiza el contenido de una página, mayor es la probabilidad de que un buscador la considere un punto extra positivo a la hora de los rankings.

Además, gracias a los nuevos sistemas de indexación, los robots de búsqueda son capaces de llegar a encontrar una página a los pocos minutos de haber aparecido y ponerla en buenos puestos de los resultados de búsqueda, de forma que si los usuarios entran en esta nueva página, el buscador la mantendrá o incluso subirá de puestos.

Sin embargo, agregar elementos de frescura a la página puede ser un tarea que consuma tiempo. Hay diferentes maneras para generar contenido fresco, sin la necesidad de consumir demasiados recursos.

▣ Notas de prensa

son algunos los casos en los que se dispone de una agencia de prensa... lo más interesante es que esas notas de prensa estén en una sección del sitio y así mantenemos esa sección con cierta frecuencia actualizada.

▣ Preguntas frecuentes

en la mayoría de sitios hay una zona de contacto donde se pueden añadir una serie de preguntas y respuestas habituales que suelen realizar los visitantes al llegar al sitio.

▣ Blogs

fácilmente con un sistema de blog se puede integrar contenido más o menos actualizado al sitio web en cuestión que haga que parte del material que haya no sea estático y esté actualizado.

▣ Directorios

crear un pequeño directorio en el propio sitio web ayudará a que generemos contenidos temáticos y que enlacen a sitios de calidad, algo que valoran positivamente los robots de búsqueda.

▣ Glosarios

un pequeño listado de palabras con sus respectivas definiciones, enlazando a las secciones de la web, puede ayudar a que los usuarios encuentren lo que buscan cuando no tienen claros los conceptos.

Los buscadores intentan entregar al usuario contenido relevante para sus búsquedas, por lo que las páginas que están enfocadas concretamente a una temática tendrán mejores posiciones en los resultados. Lo ideal es organizar el sitio web de temas muy genéricos a temas más concretos, de forma que cuanto más profundidad tenga el sitio más concretas sean las temáticas que allí se tratan.



Pero no sólo es importante mantener el contenido fresco, sino que la cantidad de contenido y páginas han de ir aumentando con el tiempo. Los sitios que tienen pocos contenidos pueden ser planteados por los buscadores como páginas que no se actualizan o no evolucionan y eso no es bueno para subir posiciones, de forma que si puedes generar contenidos con frecuencia, ganarás en calidad y reconocimiento (además de tener más contenido por el que ser encontrado).

La implementación de un blog corporativo que se actualice con cierta frecuencia no sólo es una excelente herramienta de marketing y comunicación sino que además es beneficioso para el posicionamiento por su naturaleza propia.

2.8.2 Frases de búsqueda y palabras clave

Una cosa que sí debes tener en cuenta al crear contenidos es cuidar cómo lo haces. Sí, puede parecer que crear contenidos es muy sencillo, pero si realmente lo fuera no existiría la carrera de Periodismo, ni por supuesto, los periodistas, principales generadores de contenidos.

Por eso cuando hagas un contenido, y sobretodo los textuales, hay que tener muy claro qué es lo que queremos destacar del texto. Hasta ahora se hablaba del concepto de “palabras clave”, pero cada vez más los usuarios que buscan utilizan más palabras a la hora de realizar sus consultas, lo que hace que no te centres tanto en palabras clave como en frases de búsqueda o conceptos.

Y es que cada vez más hay de tener en cuenta que tipo de perfil es el que queremos conseguir atraer al sitio, ya que sabiendo este perfil puedes preveer cómo busca y por ende el lenguaje que va a utilizar.

Un ejemplo de artículo podría ser éste (como ejemplo de una supuesta web de ciencia con un nivel medio-bajo):

La vida empezó “en el fondo del mar”

Una nueva y controvertida teoría sobre el origen de la vida en la Tierra ha dado que hablar en la comunidad científica.

La nueva concepción sostiene que los organismos vivos se originaron en los que llama “incubadores inorgánicos”. Se trata de pequeños nidos formados en rocas de sulfuro de hierro.

La teoría, propuesta por los profesores William Martin, de la Universidad de Düsseldorf (Alemania), y Michael Russell, del Centro Escocés de Investigación Ambiental en Glasgow, cuestiona de la raíz ideas convencionales.

Hasta ahora se ha sostenido que primero se formaron los ladrillos elementales de la vida, los cuales se fueron juntando para generar una estructura celular básica.

La nueva propuesta señala, en cambio, que la célula se formó primero y que después se rellenó con moléculas de vida.

Una de las implicaciones de la teoría es que existen más probabilidades de que se desarrollen organismos en planetas donde antes se pensaba que sería imposible.

Tras la lectura de este fragmento de noticia podrías plantearte cuáles son los conceptos que tal vez interesaría posicionar: “dónde comenzó la vida” y “origen de la vida”. Estas dos frases pueden ser una búsqueda que la gente haga y por tanto es algo que debes plantearte posicionar. Además podrías destacar los nombres de los científicos.

Así pues, puedes dejar redactado el artículo de la siguiente manera:

El origen de la vida en el fondo del mar

Una nueva y controvertida teoría sobre el origen de la vida en la Tierra ha dado que hablar en la comunidad científica.

La nueva concepción sostiene que los organismos vivos se originaron en los

que llama "incubadores inorgánicos". Se trata de pequeños nidos formados en rocas de sulfuro de hierro.

La teoría sobre dónde comenzó la vida, propuesta por los profesores William Martin, de la Universidad de Düsseldorf (Alemania), y Michael Russell, del Centro Escocés de Investigación Ambiental en Glasgow, cuestiona de la raíz ideas convencionales.

Hasta ahora se ha sostenido que primero se formaron los ladrillos elementales de la vida, los cuales se fueron juntando para generar una estructura celular básica.

La nueva propuesta señala, en cambio, que el origen de la vida está en la célula que se formó primero y que después se rellenó con moléculas de vida. Una de las implicaciones de la teoría es que existen más probabilidades de que se desarrollen organismos en planetas donde antes se pensaba que sería imposible.

Una vez tienes el texto completamente redactado (con las principales frases con las que conseguir tráfico en párrafos con cierta distancia, una combinación de palabras cada 2 ó 3 párrafos), debes plantear si debe quedar como texto plano o marcando algunas palabras de alguna forma.

Con este nuevo título y el haber introducido un par de veces los conceptos es muy probable que el posicionamiento mejore en muchos buscadores. Eso sí, hay que tener en cuenta una cosa: **no abusar**. La lectura para una persona es muy similar en los dos textos, aunque en el segundo, hacemos que la lectura para los buscadores mejore significativamente.

El origen de la vida en el fondo del mar

Una nueva y controvertida teoría sobre el origen de la vida en la Tierra ha dado que hablar en la comunidad científica.

La nueva concepción sostiene que los organismos vivos se originaron en los que llama "incubadores inorgánicos". Se trata de pequeños nidos formados en rocas de sulfuro de hierro.

La teoría sobre dónde comenzó la vida, propuesta por los profesores William Martin, de la Universidad de Düsseldorf (Alemania), y Michael Russell, del Centro Escocés de Investigación Ambiental en Glasgow, cuestiona de la raíz ideas convencionales.

Hasta ahora se ha sostenido que primero se formaron los ladrillos elementales de la vida, los cuales se fueron juntando para generar una estructura celular básica.

La nueva propuesta señala, en cambio, *que el origen de la vida está en la célula que se formó primero y que después se rellenó con moléculas de vida*. Una de las implicaciones de la teoría es que existen más probabilidades de que se desarrollen organismos en planetas donde antes se pensaba que sería imposible.

En este último fragmento, el haber añadido conceptos en negritas y cursivas hará que el robot de búsqueda se quede atento a esos conceptos y les de cierta importancia por encima del resto del documento.

2.8.2.1 Densidad de contenidos

Cada buscador tiene sus propias fórmulas para calcular y encontrar resultados de búsqueda relevantes. La ubicación de los conceptos y su densidad con respecto a la cantidad de texto que se encuentra en un página, puede afectar los resultados, ya que es lógico que si un concepto aparece una única vez en un contenido de 5 páginas no tenga mucho valor con respecto al global.

Ubicar conceptos al comienzo de cada página asegura que el enfoque de la página sea leído rápidamente. El título y los metadatos forman parte de la información disponible en la cabecera de la página, pero si se tienen palabras al comienzo del contenido, esto será beneficioso también.



La densidad se refiere a la cantidad de veces que se repite una palabra cada 100 palabras. Es aconsejable mantener un ratio de 2%-3%, tratando de que los conceptos aparezcan en intervalos regulares entre el texto del contenido, lo que significaría que cada 2 o 3 párrafos has de incluir de alguna forma los temas principales que se están tratando.

Habiendo dicho esto, crear páginas pensadas solamente para los buscadores no es una buena idea. El texto de la página debe poder ser leído correctamente y debe persuadir al visitante a tomar una acción. Esto significa que si ves que incluir los conceptos no es posible es mejor no forzar y aplicar otro sistema que encuentres más conveniente.

Una forma de verlo es con la versión “solo texto” que tiene guardada Google de las páginas:

```
http://www.google.com/search?strip=1&q=cache:www.dominio.ext/dir/pagina.html
```

2.8.3 Contenidos atractivos

Pero no todo es cuestión de añadir u optimizar contenidos, sino que la forma de hacerlo también es uno de los secretos. No basta con atraer a los usuarios desde los motores de búsqueda sino que además hay que conseguir que si llegan a la página la utilicen como fuente y se queden el mayor tiempo posible.

Es muy importante que la página cargue rápidamente, que lo que aparece en una primera impresión sea claro y fácil de leer y que sea sencillo navegar de un sitio a otro por la misma. Además, hay que aportar elementos que nos distinguan, como son enlaces relacionados a un contenido, noticias de interés, ofertas y promociones...

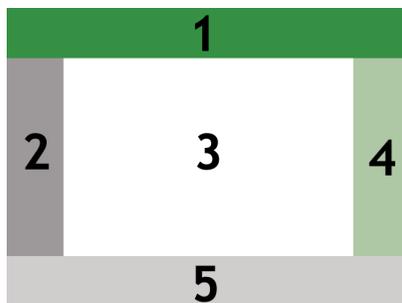
2.9 Estructura del sitio

La forma en que se organiza la información (o sea, dónde están los textos, imágenes y enlaces dentro del código fuente) puede afectar a cómo los buscadores la leen. Los robots de búsqueda leen los sitios de izquierda a derecha y de arriba a abajo (si no se les indica lo contrario, como puede pasar en países con otros alfabetos) y consideran lo que aparecería en “en primer pantallazo” (lo que se ve por pantalla nada más entrar en la página) lo más importante.

Gracias a los CSS, si reorganizas el sitio web de una manera más estructurada, los datos se pueden colocar de manera que el contenido principal de la página quede lo más arriba posible, manteniendo la apariencia general del sitio. Como consecuencia, el contenido de la página podrá conseguir una clasificación superior, lo que permitirá mejorar el resultado global del sitio.

A continuación te planteo algunas plantillas habituales de sitios web:

Este podría ser el planteamiento de una estructura completa y compleja en la que nos encontramos las 5 zonas primarias de contenido.



La zona (1) es aquella en la que habitualmente encontramos el nombre del sitio, la cabecera, junto a un buscador, una zona de publicidad u otros elementos similares.

Las zonas (2) y (4) suelen incluir los menús de navegación o enlaces a ciertos contenidos destacados del sitio.

La zona (3) incorporaría el contenido principal de la página, donde encontraremos el grueso del contenido de la página y donde deberán estar los contenidos principales a posicionar.

Para acabar, la zona (5) es el pie de página, donde habitualmente están los datos de acceso a la propiedad del sitio y los enlaces a datos legales correspondientes a cada país.

Cada sitio web tiene unas necesidades concretas y no se puede generalizar, pero dándose el caso, lo ideal es que se indexe la página en 4 bloques:

1. Cabecera
2. Contenidos
3. Navegación
4. Pie de página

Esta estructura es la ideal en el caso de que en los menús de navegación no haya enlaces externos directos a otras páginas que no sean las propias, ya que en el caso contrario, el de haber una gran cantidad de enlaces externos al final de la página, lo ideal sería cambiar el formato a la siguiente estructura:

1. Cabecera
2. Navegación
3. Contenidos
4. Pie de página

Como digo, cada caso es un mundo y hay que estudiar bien si un motor de búsqueda podría considerar ese menú de navegación como algo “negativo” en el conjunto del sitio y por tanto será bueno estudiar la posición del mismo.

Aún así y gracias a las tecnologías XHTML y CSS podemos permitir el lujo de que el código esté ordenado de la manera óptima SEO, pero que aparentemente se vea de otra forma.

2.10 Mapa del Sitio

Un mapa del sitio es una sección donde encontrarás una especie de índice de todo lo que podemos encontrar. Lógicamente no hace falta incluir todos los contenidos, sólo la estructura principal (secciones y subsecciones) de forma que el robot de búsqueda, si entra en este mapa, tenga accesibles todas las secciones principales donde encontrar información.

Como beneficio adicional, los mapas de los sitios ofrecen la oportunidad de incluir páginas que normalmente no podrían formar parte de la estructura normal de navegación del sitio. Como ejemplo, se podría encontrar una breve lista de los contenidos más destacados para que estén a un clic y no haya de buscar el robot a través de las diferentes categorías.

Aún todo esto, hay que tener en cuenta que no puedes excederte en la cantidad de enlaces de este mapa. Si el sitio es muy grande deberás elegir las páginas que quieras destacar.

Esta página sirve de ayuda a la indexación pero no implica que si una página no está en el mapa no se indexe... el mapa es una vía para facilitarle la ayuda a los robots de indexación. Además, en algunas ocasiones, si el usuario no tiene mucha idea de dónde encontrar un contenido es probable que vaya a este mapa del sitio para hacerse una idea más clara de su estructura y así poder llegar a su cometido.



2.11 Dirección URL

Este capítulo tiene contenido altamente técnico, por lo que se recomienda la lectura a todo el mundo, pero la aplicación del mismo sólo a aquellos con conocimientos técnicos suficientes como para aplicarlos con seguridad.

Como ya he comentado en varias ocasiones, la estructura del sitio es uno de los elementos que hace que la información quede bien organizada y de esa forma sea mucho más sencillo llegar a ella.

Al igual que el título de la página web, la dirección URL es uno de los elementos que se muestran en los resultados de búsqueda, algo que la convierte en uno de los elementos importantes.

Habitualmente los robots de búsqueda suelen indexar los sitios web en base a su profundidad. La profundidad suele venir dada por la cantidad de carpetas, archivos o parámetros de las URLs y por esto mismo hay que tener muy en cuenta su funcionamiento y su corrección (en la medida de lo posible).

Una dirección web puede ser esta:

<http://www.sitiodeejemplo.com/index.php?s=1234>

O puede ser esta:

<http://www.sitiodeejemplo.com/contenidos/pagina-de-contenido/>

Está claro que la segunda opción es mucho más intuitiva que la primera ya que el usuario que ve los resultados de búsqueda encuentra que la dirección tiene relación con lo que busca, además de ser mucho más sencilla de recordar que la primera.

La pregunta es simple... ¿cómo se pueden conseguir estas direcciones de una forma sencilla? En este caso no es tan sencillo como incluir una negrita o una cursiva en un texto, ya que lo más probable es que si tienes estas direcciones tan complejas es porque usas algún sistema de base de datos. Si simplemente tienes una página en “html” lo que hay que hacer es cuidar muy bien cómo llamas a los ficheros.

Se podría hacer una clasificación simple de dos tipos de direcciones web: las que llevan “identificadores” y las que no. Las que llevan identificador son aquellas que a un mismo fichero, como es el caso del primer ejemplo, le pasamos un número o nombre que va cambiando y muestra una página u otra.

También has de tener en cuenta que es muy diferente el sistema de cambios de URLs para Internet Information Server (IIS)⁷ como para Apache HTTP Server⁸, dos de los servidores web más frecuentes en Internet, ya que existe un sistema que permite hacer estas mejoras de forma sencilla que viene implementado por defecto en Apache pero no en IIS, aunque para ello dispones de Isapi Rewrite⁹ que será 100% funcional con la versión Lite (gratuita) para las necesidades básicas.

El funcionamiento del Mod_Rewrite¹⁰ es muy sencillo, aunque hay que llevar mucho cuidado con él ya que una mala configuración podría hacer que no funcione nada. Para la versión del Isapi Rewrite es recomendable la lectura de su funcionamiento. Para la versión de Apache que viene por defecto simplemente hay que crear un fichero en el directorio raíz de nuestro sitio web llamado “.htaccess” (vigila el punto al principio del nombre), aunque también es

(7) <http://www.iis.net/>

(8) <http://httpd.apache.org/>

(9) <http://www.isapirewrite.com/>

(10) http://httpd.apache.org/docs/2.2/mod/mod_rewrite.html

recomendable darle una ojeada a su documentación. En el caso de los ejemplos siguientes, todos se basan en Apache.

Como casos de ejemplo podrías encontrarte lo siguiente: tenemos unas páginas creadas por nosotros que son medianamente estáticas.

```
http://www.sitiodeejemplo.com/index.php?pagina=productos
http://www.sitiodeejemplo.com/index.php?pagina=ayuda
```

Quizá para estas páginas sea mucho más simple aplicar un cambio con Mod_Rewrite para que queden así:

```
http://www.sitiodeejemplo.com/pagina/productos/
http://www.sitiodeejemplo.com/pagina/ayuda/
```

¿Cómo lograr esto? Con un fichero “.htaccess” que incluya lo siguiente:

```
1. RewriteEngine on
2. RewriteRule pagina/([A-Za-z0-9]*)/ index.php?pagina=$1
```

En este caso, la primera línea (RewriteEngine) lo que indica es que activamos el Mod_Rewrite y la segunda le indica la regla: todas las páginas que tengan el primer formato (pagina/([A-Za-z0-9]*)/), internamente llamarán al segundo formato (index.php?pagina=\$1), es decir, que una página que llegue en formato “/pagina/loquesea/” internamente llamará a “/index.php?pagina=loquesea”.

Esto vendría a ser lo más sencillo, que se puede complicar como queramos... por poner un ejemplo diferente y algo más complejo, utilizando una base de datos, el sistema funciona con identificadores, como por ejemplo:

```
http://www.sitiodeejemplo.com/articulos.php?cat=novedades&id=548
http://www.sitiodeejemplo.com/articulos.php?cat=soluciones&id=32
http://www.sitiodeejemplo.com/soporte.php?ayuda=tecnologia
```

En este caso quizá estaría mejor que las direcciones fueran como:

```
http://www.sitiodeejemplo.com/articulos/novedades/548/
http://www.sitiodeejemplo.com/articulos/soluciones/32/
http://www.sitiodeejemplo.com/soporte/tecnologia/
```

Un fichero “.htaccess” para aplicar el Mod_Rewrite quedaría como:

```
1. RewriteEngine on
2. RewriteRule articulos/([A-Za-z0-9]*)/([0-9]*)/ articulos.
php?cat=$1&id=$2
3. RewriteRule soporte/([A-Za-z]*)/ soporte.php?ayuda=$1
```

La primera regla serviría para los dos primeros ejemplos, ya que la estructura de ambas es muy similar...

Lógicamente puedes hacer combinaciones de muchos tipos y para ello existe multitud de documentación sobre Mod_Rewrite, como la documentación de Isapi Rewrite o la documentación de Apache Mod_Rewrite. Una búsqueda en la red te arrojará multitud de opciones con las que trabajar.

Un detalle a tener en cuenta es que, a nivel posicionamiento, las sesiones y las cookies no tienen validez por lo que no se recomienda el uso de ellas. En el caso de necesitarse lo más recomendable es que no se indexen esas zonas en las que se utilicen ya que se podría llegar a



convertir en una penalización al duplicar el buscador todas las sesiones en páginas idénticas que lo único que cambia es el identificador de sesión.

2.12 robots.txt

Cuando en el año 1994 comenzaron a aparecer los primeros robots de búsqueda, o simplemente unos robots que investigaban Internet para calcular su tamaño, los administradores de sitios web se encontraban con problemas porque estos robots entraban en las páginas web y las saturaban. Es por eso que se desarrolló un protocolo para que los robots de búsqueda los cumplieran creando así los ficheros “robots.txt”.

El fichero “robots.txt” es muy importante en lo que se refiere a los robots de búsqueda, ya que es lo primero que van a visitar de un sitio web, antes incluso que la página principal del dominio. Este fichero hay que colocarlo en la carpeta raíz del dominio (o subdominio)... no sirve de nada que esté en una carpeta interior ya que no será leído nunca.

Este fichero, un documento de texto (TXT) normal y corriente, incluye directrices propias a cada uno de los buscadores que permiten indicarle que partes del sitio pueden o no pueden indexar, lo que te ayudará a indicarle si hay contenido que quieras que quede excluido o incluido.

El archivo “robots.txt” más simple utiliza varias reglas:

- ▣ **User-Agent**
el robot al que se aplica la regla.
- ▣ **Disallow**
las páginas que desea bloquear.
- ▣ **Allow**
las páginas que desea indexar.

Se pueden incluir todas las entradas que se desee así como varias líneas “Disallow” y varios “User-Agent”.

El fichero que, por defecto, debería haber en todos los dominios si no quieres aplicarles ninguna regla sería:

```
1. User-Agent: *
```

Con esto le indicas que todos los robots de búsqueda tienen vía libre para indexar todos los contenidos que haya en el sitio web sin excepción, siempre y cuando sean capaces de rastrearlos.

Un ejemplo sencillo puede ser el siguiente, en el que todo el sitio quedaría excluido por los robots de búsqueda, es decir, que no se indexaría ni debería aparecer en los resultados:

```
1. User-Agent: *
2. Disallow: /
```

Un ejemplo algo menos sencillo es el siguiente, en el que te encontrarás con un directorio al que los robots no han de acceder, pero donde le aplicas una excepción con una página concreta del mismo:

```
1. User-Agent: *
2. Disallow: /directorio/
3. Allow: /directorio/fichero.html
```

Aunque también se puede jugar con los robots de cada buscador:

1. `User-Agent: Google`
2. `Disallow: /`
3. `User-Agent: MSNBot`
4. `Disallow: /fichero_prohibido.html`
5. `User-Agent: *`
6. `Allow: /`

En este caso le estás diciendo al robot de Google que no puede indexar nada, al robot de Microsoft que puede indexar todo excepto el fichero prohibido y al resto de robots que pueden indexar lo que quieran.

Existe una base de datos con la lista de robots¹¹ que mantiene el sitio oficial de robots.txt¹².

2.12.1 Dirección URL del protocolo Sitemaps

Aunque más adelante hablaré sobre el protocolo Sitemaps¹³, al tener este punto relación directa con el robots.txt, comentarte que desde el pasado mes de abril de 2007, los buscadores que apoyan el protocolo Sitemaps decidieron ampliar la autodetección para que, sin necesidad de registrarse en los paneles de los buscadores uno a uno, el fichero de Sitemap sea autodescubrible gracias al robots.txt.

Para ello sólo has de añadir una línea en cualquier parte del fichero (se recomienda al inicio o al final) que incluya la dirección completa (no sirve poner el fichero únicamente):

1. `Sitemap: http://www.ejemplo.ext/sitemap.xml`

2.12.2 Crawl-delay

Una de las opciones que soportan algunos de los motores de indexación (la mayoría de los importantes lo soportan) es el parámetro Crawl-delay que indica el tiempo en segundos que ha de pasar entre una petición y otra del buscador. Esto sólo debemos utilizarlo si vemos que el servidor recibe excesivas peticiones y afecta a la estabilidad del sitio.

Las cifras estándar a utilizar son 0.5 o 1:

1. `Crawl-delay: 1`

Aunque si ves que el servidor sigue respondiendo mal puedes ampliar esa cifra para que tarde más entre una petición y otra.

1. `Crawl-delay: 5`

Si pones cifras muy elevadas el robot decidirá que no puede pasar tanto tiempo allí y acabará dejando el sitio, sin indexar muchas de las páginas que tendría a la espera.

2.13 Página no encontrada (Error 404)

Una de las cosas que no suele tenerse muy en cuenta cuando se da de alta un sitio nuevo es el control de las páginas de error. Por defecto, los servidores web como Apache o IIS suelen mostrar una página en blanco con un mensaje de error que dice algo como “Error 404 - Página no encontrada”.

Hay que decir que estas páginas no son indexadas por los robots de búsqueda (y que a la larga, si estaban indexadas irán desapareciendo de los resultados de búsqueda), pero sí que llevan a

(11) <http://www.robotstxt.org/db.html>

(12) <http://www.robotstxt.org/>

(13) <http://sitemaps.org/>



un problema y es que, si el robot está indexando una lista de páginas y llega a una 404 automáticamente “se le corta el rollo” y puede llegar a marcharse del sitio web.

La opción por defecto y que deberías usar habitualmente para tratar este tipo de errores es aquella en la que se devuelve un error 404 del servidor pero, en vez de utilizar una plantilla básica que diga que esa página no se ha encontrado, configuremos el servidor web para que, por ejemplo, muestre la página del mapa del sitio. De cara a los buscadores da igual qué página mostremos, ya que al ser un error no se indexará ni se tendrá en cuenta, pero de cara al usuario es mucho mejor mostrarle el mapa del sitio o una página similar en la que se le indique que esa página que buscaba ya no existe pero que pueda seguir navegando por el sitio sin problema.

En algunas ocasiones podrías plantear crear redirecciones o mostrar por defecto la página principal del sitio, pero esto puede suponer a corto plazo que los robots indexen páginas que no existen con el mismo contenido de la página principal y que podrían suponer una penalización por tener muchas páginas iguales.

En cualquier caso, tener unas páginas de error para todos los errores posibles¹⁴ es siempre interesante, ya que aunque el Error 404 es el más habitual existen muchos otros.

2.14 Hosting / Alojamiento Web

Tener una web alojada en un sitio gratuito sirve de poco. Puede haber miles de webs, el rendimiento puede ser bajo y tu sitio se merece algo mejor. Por eso has de buscar un alojamiento de calidad, en cuya IP hayan pocos sitios web, donde puedas tener el control absoluto del dominio y que permita tener todo lo que necesitas. No es necesario tener un alojamiento dedicado -que puede ser algo caro- pero siempre hay que buscar la calidad.

Los robots de búsqueda van pasando e indexando tu sitio a diferentes horas del día y si en ese momento el rendimiento del servidor es bajo lo más probable es que el robot se marche y no indexe los contenidos actualizados. Si el problema es recursivo puede llegar a producirse la pérdida completa del posicionamiento.

Además, uno de los factores que los robots de indexación tienen en cuenta es la velocidad de descarga del sitio, por lo que si el rendimiento es bajo se considerará un sitio de baja calidad y ello podría suponer pérdida de puestos en los resultados de búsqueda.

(14) http://es.wikipedia.org/wiki/HTTP#C.C3.B3digos_de_respuesta



Factores Externos (off-page)

CAPÍTULO

3

Desde que apareció el algoritmo de ordenación Pagerank¹, los sitios web no sólo ganan importancia por sus propios contenidos y desarrollo, sino que entran en juego los enlaces externos que puedes conseguir a tu sitio. Qué el origen sea el Pagerank no significa que esto sea sólo válido para Google ya que en la actualidad todos los motores de búsqueda lo tienen en cuenta porque sus algoritmos funcionan de una forma similar.

Para conseguir posicionar un sitio web es importante que la cantidad de enlaces sea alta y de calidad. Un sitio web suele conseguir enlaces de forma natural, sobretodo si tiene contenidos interesantes, pero no suele ser suficiente y has de dedicar tiempo y esfuerzo a conseguir otros enlaces que tengan utilidad para los lectores de los sitios.

Una lista de factores que hay que tener en cuenta sobre estos enlaces es:

1. Los enlaces han de “parecer” lo más naturales posibles. De esta forma si pides a alguien que ponga un enlace en otro sitio, no utilices las mismas palabras sino que hay que ir variándolas para que los enlaces no parezcan forzados y por tanto no sean considerados fraudulentos en vista a los buscadores.
2. Una forma de conseguir calidad en los enlaces es que las páginas de origen tenga alguna relación con la temática de tu sitio, contenidos similares y relacionados (complementarios).
3. Intentar buscar sitios web que lleven mucho tiempo online. Se priorizan los sitios web antiguos frente a los nuevos, sobretodo si se siguen actualizando con frecuencia. Aún así, hay que buscar páginas de cualquier tipo, ya que internet está lleno de sitios de todo tipo.
4. El idioma es un detalle muy importante. Si tu sitio está en español, se recomienda que los enlaces de origen también sean de sitios en español; si está en francés, se recomienda que los enlaces también vengan de sitios en francés.
5. Hay que ser progresivos al enlazar. No es normal conseguir muchos enlaces de golpe de forma continuada. Lo mejor es hacerlo poco a poco y de forma incremental. Aún así, a veces, por algún tipo de contenido que puedas publicar, puede llegarte una avalancha inesperada de enlaces (Link Baiting), algo que también es interesante ya que suele conducir a que otros te enlacen.

(1) <http://www.google.com/intl/es/technology/>

6. Los enlaces no deben estar siempre en el mismo lugar de la página de origen (hablo de los diferentes sitios web), es decir, debes conseguir que haya enlaces en la parte superior, en los laterales, en la parte inferior (siempre hablo del código fuente, ya que las posiciones se pueden intercambiar gracias a los CSS). De esta forma tu enlace no será siempre considerado “del montón”.
7. Los enlaces no sólo han de estar en la página principal (pero tampoco es necesario ni recomendable que esté en todas las páginas de un sitio); hay que recibir enlaces desde páginas iniciales, secciones que a lo mejor tienen más relación con nuestro sitio, artículos en páginas interiores...
8. Los enlaces que llevan mucho tiempo en un sitio suelen tener más valor. Es interesante que si se llega a un acuerdo de intercambio o alquiler dure lo máximo posible con un mínimo de 6 meses para que haga un efecto sobre los motores de búsqueda.

3.1 La fuerza de la IP

Uno de los elementos más interesantes de la red es su clasificación por rangos de IP. Hoy en día todavía se trabaja con el IPv4 y hasta que llegue el IPv6 los dominios de Internet pueden compartir una misma IP.

Es muy interesante tener en cuenta varios factores sobre las IPs, como por ejemplo conseguir enlaces de sitios de IPs diferentes (de diferentes clases) y que estas IPs no estén penalizadas.

Con respecto al primer punto es interesante saber que las direcciones IP se dividen en 4 bloques: xAx.xBx.xCx.xDx. Cuando se habla de conseguir enlaces de diferentes clases C significa que has de intentar conseguir enlaces de sitios que no estén cerca del tuyo, es decir, que no sean de tu propia red de sitios. Eso no significa que no pueda haberlos ya que un primer paso para conseguir enlaces es comenzar a enlazar desde todos tus sitios para que aparezcan rápidamente en los buscadores. Si ya aparecen les darás pie a que este nuevo sitio sea encontrado más fácilmente. Aún así, esto es una primera fase ya que estos enlaces no tienen todo el valor que quisiéramos. Por eso, en muchas ocasiones, te podría ser interesante tener alojamiento web en varios sitios diferentes con IPs diferentes, incluso en países diferentes.

Con respecto a la penalización de IPs, en Internet hay listados públicos de sitios, y más concretamente de IPs, desde donde se ha realizado envíos masivos de correo no solicitado. Estas listas son públicas y son revisadas periódicamente, por lo que es interesante estar atento a no aparecer en ellas, y si se aparece, intentar salir de ellas demostrando que la penalización que se indica no es correcta (en caso de ser así).

3.1 Conseguir enlaces de sitios con autoridad

Antes de ponerse a conseguir enlaces sin control vale la pena plantearse cómo conseguir enlaces de sitios de calidad ofreciendo contenidos de calidad. Si es lo que quieres (deberías quererlo) lo primero que has de hacer es una lista de los sitios donde deberías aparecer, ya sean en tu idioma o en otro.

Una vez elegidos estos sitios has de ver si tienes contenido o herramientas que sean fácilmente de interés para los usuarios de esos sitios web de referencia. Recuerda que tras un sitio web hay una o varias personas y si vas a contactar con ellos has de conocer mínimamente con quién vas a tratar y no hacerlo “porque sí”. Intenta ofrecerle algo que sea de ayuda para estas personas.

Recuerda que no es de buena educación asaltar a alguien, por lo que primero intenta acceder

a estos usuarios y después de un cruce de correos ya te puedes comenzar a plantear mostrarles contenidos de tu sitio y la promoción de algún artículo o herramienta.

3.3 Directorios

En internet hay dos grandes directorios que son utilizados por los buscadores para ampliar su información. Uno de ellos es el Directorio de Yahoo!² que, lógicamente, utiliza Yahoo! para extraer datos. El otro es el Directorio Abierto DMOZ³ que mantienen usuarios y que también es utilizado por buscadores como Google o Windows Live para ampliar sus resultados.

El hecho de aparecer en estos directorios no mejora el posicionamiento de por sí, pero cuando son indexados por los buscadores, ayudan a conseguir enlaces de calidad (al estar en categorías con otras páginas de temática relacionada).

No hay que descartar añadirse en otros directorios, aunque has de vigilar que sean sitios en los que haya una calidad, tanto de contenidos como que no haya posibilidad de ser tratado de Granja de Enlaces ni que esté penalizado.

Por ejemplo, es interesante que al lanzar un nuevo sitio (con algo de contenido) se utilice este método primero. Los buscadores al encontrar los enlaces puede que no los tengan en cuenta en una primera fase pero nos ahorrará problemas en un futuro en caso de penalizaciones de otro tipo.

Es importante dar de alta los sitios de forma manual, sin utilizar herramientas automáticas. Cada directorio debería incluir tu sitio con unas palabras clave y textos adaptados y que sean diferentes con respecto a otros, para tener variedad.

3.4 Intercambio de Enlaces

Una manera simple de conseguir enlaces es la de intercambiarlos. Mucha gente intercambia enlaces según el Pagerank que tienen sus páginas intentando conseguir enlaces de páginas con Pagerank superior o de uno igual. Según esta tecnología esto no debería hacerse así, ya que si entramos en una cantidad de intercambios elevada de similar Pagerank, se puede confundir esos intercambios con una Granja de Enlaces.

Lo recomendable al hacer intercambios de enlaces es hacerlo a 3 bandas, es decir, el sitio A enlaza al sitio B, el sitio B al sitio C y el sitio C al sitio A, aunque esto es sólo una sugerencia, se pueden hacer anillos mucho mayores, intercambios cruzados, etc...

Otra posibilidad es hacerlo encadenado y escalonado:

A -> B -> C -> D

A -> C

B -> D

Lo mejor en los intercambios de enlaces es hacerlos con otros sitios de igual o similar temática. Como esto a veces es complejo, ya que puede entrar en medio el concepto de competencia, tal vez haya que plantearse hacer intercambios con sitios que aporten contenidos complementarios, como por ejemplo un sitio que venda vehículos y un taller de reparación, o una tienda de música y una de películas.

También has de plantearte hacer intercambio en diferentes secciones de tu sitio, en los que los temas puedan tener mayor relevancia.

(2) <http://dir.yahoo.com/>

(3) <http://dmoz.org/>



No hay que excederse con los intercambios, intentando no tener en un sitio más de 5 a 10 enlaces, ya que los buscadores los podrían tratar de “enlaces vendidos” y penalizarte por ello.

3.5 Alquiler de Enlaces

También conocido como “compra de enlaces”. En ocasiones puedes encontrarte un sitio web que, después de estudiarlo con detenimiento, puede interesarte mucho conseguir un enlace desde allí. Por ello podrías alquilar un enlace al propietario del sitio pagándole una cuota mensual. El interés de estos enlaces es que estén como mínimo 6 meses, tiempo suficiente para que haga efecto.

Si el sitio es muy interesante, has de tener en cuenta que este alquiler no sólo mide la ayuda a posicionarnos sino la cantidad de visitas que nos lleguen, porque, no has de olvidar que una de las razones finales para posicionarnos en los buscadores es la de atraer visitas y que vuelvan a visitarte.

Estos enlaces han de estar muy estudiados, ya no sólo por el dinero que se invierte, sino por la posibilidad de que alguien llegue a denunciar a los buscadores sobre ello. Aún así no hay que preocuparse si es un enlace que parece natural.

Este sistema, al fin y al cabo, es lo mismo que el intercambio de enlaces pero unidireccional.

3.6 Notas de Prensa

Una forma de conseguir enlaces es también la de usar las notas de prensa de forma moderada. Puede que no aparezcan en grandes medios pero seguro que si la información es de interés le llega de una u otra forma a un sitio que está interesado.

Hay multitud de sitios en la red que permiten el envío, como podrás ver en el capítulo de herramientas.

3.7 Foros

Los foros siempre han sido lugares participativos de los que sacar ideas para artículos y para ofrecer información y enlaces a contenidos que dispones, además de opiniones personales.

Habitualmente los participantes en foros están acostumbrados a participar en estos sitios, por lo que si en alguna de las direcciones a las que enlazamos tenemos un formulario para comentar es probable que hagan uso de ellos. Además, los foros son sitios en los que la principal finalidad que nos encontramos es la de crear contenido, así que si vas a enlazar a alguno de los tuyos, lo ideal es que primero aportes la información y a continuación enlaces al punto original.

3.8 Páginas Prepagadas

En inglés conocidas como “presell pages”, consiste en alojar una página con un contenido que quieras en el dominio de otro sitio web. Páginas de contenido, con enlaces a tu sitio, alojadas en un sitio “antiguo” o “con autoridad”, de forma que envía tráfico a través de enlaces “de valor” y creando más resultados de búsqueda.

De esta forma el sitio contiene una página posicionable por algunas palabras, que está en un

dominio importante y que te enlaza, de forma que se pueden conseguir enlaces de calidad en un contexto cuidado, ya que tú eres el que aporta el contenido de la página donde están los enlaces.

3.9 Comprar Artículos

Un sistema para ganar enlaces, que últimamente se está poniendo de moda, es “comprar artículos” de bloggers (principalmente) para que hablen de ti o tus productos. Pongo entre comillas el “comprar artículos” porque no ha de ser obligatoriamente con dinero, sino que algunas empresas lo que hacen a veces es enviar productos o dejar probar servicios para que los receptores acaben publicando algo en su sitio web. En este caso se trabaja el *que hablen de uno aunque sea bien*. Aún así, también se puede llegar a un acuerdo económico con esa persona para que hable de ti.

Hay servicios como Bloguzz⁴ o Sync⁵ que permiten gestionar de una forma sencilla esta actividad.

3.10 Link Baiting

Esta técnica consiste en conseguir muchos enlaces en muy poco espacio de tiempo (dos o tres días).

Habitualmente esto ocurre por el lanzamiento de un artículo que haga “pensar” a la gente, o por el lanzamiento de algún servicio o herramienta útil, o simplemente por sitios web curiosos, como el que hace un tiempo apareció que pretendía vender un millón de píxeles.

Algo que ayuda a conseguir este tipo de situaciones es la aparición en sitios como Barrapunto⁶ o Menéame⁷, en los que un enlace que aparece allí hace que de repente llegue mucho tráfico y por ello otra gente enlace a tu sitio web porque le resulte interesante, de forma que de un día para otro muchos enlazan.

Existen algunas estrategias sencillas para crearlo:

▶ Premios a los sitios que más tráfico mandan

Una estrategia curiosa que consiste en regalar algún tipo de premio a aquel que a lo largo de un espacio de tiempo (normalmente un mes) te manda más visitas, controladas por *referers*.

▶ Intercambio de comentarios

Consiste en ofrecer la posibilidad de poner un enlace en un artículo de un sitio si te enlazan comentando sobre tu sitio.

▶ Memes

Son artículos en los que, siguiendo una plantilla de preguntas o peticiones, te invitan a contestarla o a hacer una lista de elementos y a enviársela a otros sitios.

▶ Opinión sobre tu sitio

Similar al intercambio de comentarios, pero en este caso el planteamiento es que uno enlace al sitio y a cambio el otro hace un pequeño comentario sobre el sitio de origen, dando consejos u opinando sobre él.

(4) <http://www.bloguzz.com/>

(5) <http://www.sync.es/>

(6) <http://www.barrapunto.com/>

(7) <http://www.meneame.net/>



3.11 No Follow

A principios de 2005 se planteó el microformato `nofollow`⁸ que indica a los robots de búsqueda que ese enlace no ha de ofrecer ningún tipo de valor, voto o similar de la página de origen a la página de destino. La propuesta fue ideada principalmente por Google y no todos los robots la siguen.

Si en un sitio web te encuentras que tu enlace tiene un formato similar al siguiente, ese enlace no tiene valor como enlace entrante a tu sitio web (aparentemente, es como si no existiera).

De esta misma forma, cuando hagas intercambios de enlaces o compra de enlaces, asegúrate que no tienen este microformato para que el enlace tenga valor a la hora de contabilizar y no ser tratado de enlace vendido o de enlace basura.

```
<a href="http://www.red.es/" rel="nofollow">Red.es</a>
```

Incumpliendo toda filosofía, ya que “nofollow” significa que ese enlace no se debe seguir, la mayoría de enlaces lo siguen para conseguir enlazar contenidos a los que tal vez no llegarían de otra manera.

3.12 La Competencia

Para salir primero en un buscador has de salir mejor que tu competencia. Es por esto que es interesante hacer un estudio de quién enlaza a tu competencia para intentar saber porqué les enlazan a ellos y no a ti y, claro está, conseguir un enlace desde esos sitios.

Para esto sólo has de mirar y mirar enlaces de Google y Yahoo!. En sus herramientas encontrarás la lista de enlaces que llegan a estos sitios web en concreto y de esa forma poder analizar y conseguir la lista.

3.13 Envíos Automáticos

Desde 1995 hasta el año 2000 el posicionamiento en buscadores se basaba principalmente en el uso de programas que hacían envíos masivos a determinados sitios webs que eran frecuentemente actualizados por los buscadores.

Los FFA's (Free for All) tenían dos grandes inconvenientes: el de que frecuentemente había que registrarse y dejar datos (lo que significaba que luego te llegaba spam) y que la mayoría de esos sitios ya han sido penalizados por los motores de búsqueda.

Por eso lo mejor es realizar cambios en tu sitio y conseguir enlaces de forma natural y manual, como vengo diciendo desde el principio, intentando que cada enlace sea tratado con una atención especial y no de forma masiva, ya que si los algoritmos detectan que puede haber fraude te penalizarán sin preguntar.

(8) <http://www.ojomicroformatos.com/microformatos/rel-nofollow/>



CAPÍTULO

4

Técnicas Penalizables

Haz páginas para los usuarios, no para los buscadores

Esta frase podría resumir fácilmente todo lo que un buscador necesita para posicionar tu sitio web y no tener que penalizarla por alguna trampa que hayas hecho -queriendo o sin querer-.

Una cosa está clara: si las cosas se hacen bien y con tiempo, salen rentables y los buscadores lo tienen en cuenta, por eso te recomiendo que si no es estrictamente necesario, no utilices ninguna de las técnicas que te voy a comentar a continuación.

4.1 Cloaking

Esta es la forma de llamar al sistema por el cual se muestra a los usuarios una información por pantalla y se muestra a los buscadores otra información.

Para ser más concreto, se utiliza para mostrar la página normal a tu usuario y si es un buscador (detectado por la IP o por el User-Agent) le muestras otra, que suele estar llena de palabras clave y de frases de búsqueda. Esta técnica está ya muy detectada y penalizada (ya que los buscadores tienen IPs de rangos alternativos y Agentes normales que comparan los resultados de ambas indexaciones y las penalizan).

Aún así, los buscadores pueden llegar a recomendar su uso en el caso de zonas de la web que estén bajo clave o zonas privadas de forma que sean capaces de saltarse esa limitación y poder indexar zonas ocultas.

4.2 Texto Oculto

En una página hay mucho sitio para ocultar texto. Se puede ocultar de muchas maneras: con texto a tamaño de letra 6 pixels, poniéndolo del mismo color que el fondo... al fin y al cabo, son “truquitos” que los buscadores conocen y penalizan.

Los buscadores son capaces de leer parte de JavaScript y CSS, por lo que algunos sistemas de ocultación que pueden hacerse servir con esta tecnología también son detectables.

4.3 Granja de Enlaces

Éste es quizá de los temas más peligrosos, ya que no sólo puedes crearlo tú para favorecerte, sino que pueden incluirte para penalizarte.

Una granja de enlaces, como su nombre indica, es un sitio web en el que hay muchísimos enlaces que se supone que hacen que los buscadores entren y visiten tu sitio web muchas veces. El problema está en que si tu sitio web entra en una página (o redes de páginas) que están penalizadas, pueden llegar a hacerlo con la tuya, por lo que siempre es importante hacer un seguimiento para saber desde dónde te enlazan.

Aún así, todos los buscadores tienen un margen de control de estas técnicas por si la competencia pone enlaces a tus sitios, presuponiendo que no todos los enlaces que se hacen contra un sitio web han de ser de calidad. El problema es que si la cantidad de enlaces “malos” supera al de “buenos” tu sitio quedará penalizado y probablemente tendrás que plantearte hasta dejar de usar ese dominio.

4.4 Páginas Puerta

Las páginas sin contenido y que sirven para redirigir el tráfico a otra parte de la web, al no aportar nada, pueden quedar penalizadas. Tal vez no todo el sitio pero sí esa página en concreto, donde, habitualmente, también se utiliza texto oculto o algunas palabras o frases para posicionar de forma indebida.

Al no aportar ningún tipo de información y que no tienen ningún sentido para el usuario, si se usan con frecuencia, podrían llegar a penalizar todo el sitio web.

4.5 Contenido Duplicado

En general en el mundo de la informática se tiende a no duplicar o hacer dos veces una misma tarea. Los buscadores también intentan no tener dos veces un mismo texto o contenido, ya que no aporta nada a la calidad de resultados.

Esto significa que sitios que generan sus contenidos gracias a contenidos de otros sitios no tienen aportación a los motores y pueden llegar a ser considerados sitios duplicados.

Un detalle a tener en cuenta sobretodo son los feeds (RSS ó Atom) que permiten ofrecer contenido a terceros de una forma sencilla y que estos otros sitios la acaben incorporando entre sus contenidos. En estos casos deberás intentar filtrar que estos sitios no roben los contenidos para uso propio.

4.6 Redirección HTML / Javascript

En algunos casos, al igual que el cloaking, hay sitios que pueden plantear mostrar un contenido y, que se cargue un JavaScript o un metatag de actualización para reenviar a otra página web.

Aunque en principio los motores de búsqueda no saben JavaScript si que son capaces de leer e interpretar parte del código, por lo que no es recomendable usar este tipo de redirecciones.

En caso de ser necesaria una redirección, te recomiendo que hagas una Redirección 301, como más adelante podrás comprobar.

4.7 NoScript / Nolframe

Como muchos piensan que los robots de búsqueda no saben leer Javascript (cosa que no es cierto), utilizan las zonas de <noscript> para escribir o indicar hacer ciertas cosas que sólo los robots verán y que los usuarios no, ya que se suelen usar navegadores lo suficientemente modernos para que todos interpreten lenguajes de este estilo.

Habitualmente en estas zonas se ponen listas de enlaces a otros sitios... pero, si la lista es extensa y no parece ser lo que pueda haber en la zona del “iframe” o “script” puede hacer sospechar al algoritmo y penalizar esa página.

4

4.8 Sitios de contenido dudoso

Desde hace un tiempo los buscadores intentan hacer uso de sus propios filtros para detectar posible *malware* en los sitios web, lo que significa que si un usuario llega a ese sitio podría llegar a descargarse algún tipo de programa o acabar yendo a algún sitio web en el que pueda haber virus o similares.

Si en tu sitio los robots de búsqueda encuentran enlaces a sitios que ellos consideran dudosos lo pueden indicar en los resultados de búsqueda y hacer que antes de entrar en ese resultado aparezca un mensaje de aviso alertando de que tu página no es segura, algo que hará perder una visita en usuarios que sigan las recomendaciones del buscador.





CAPÍTULO

5

Salir de una penalización

A veces, por la razón que sea, es posible que tu sitio web no aparezca entre los resultados de búsqueda. En este caso te recomiendo que te asegures de que no incumples ninguna de las normas para los webmasters que dicta cada buscador (que puedes encontrar en el capítulo de *Técnicas Penalizables* de esta guía) ya que es el primer paso para poder solicitar una reconsideración de tu sitio.

Cada motor de búsqueda tiene su propio sistema para revisar los contenidos, y suele hacerse de forma manual, así que es mejor que, en cuanto te des cuenta de que no apareces, revises los contenidos y lo solicites.

5.1 Ask

1. Entra en el formulario de contacto¹.
2. Rellena todos los datos.
3. Selecciona en el *Asunto* la opción *Petición de reinclusión de páginas*.
4. Explica con detalle los siguientes elementos:
 - Los términos de búsqueda que has utilizado para generar los resultados.
 - La dirección URL de tu sitio.
 - Algún detalle que hayas corregido de algo que estuvieras haciendo de forma incorrecta.
5. Finaliza el formulario y envíalo.

5.2 Google

1. Entra en solicitar reconsideración² dentro de Webmaster Tools³.
2. Selecciona el sitio de la lista. Has de tener en cuenta que has de tenerlo añadido anteriormente.
3. Explica con detalle los siguientes elementos:
 - Algún detalle que hayas corregido de algo que estuvieras haciendo de forma incorrecta.Si usabas una empresa SEO, indicar cuál es y las acciones que han realizado.
4. Finaliza el formulario y envíalo.

(1) http://about.es.ask.com/es/docs/about/customer_service.shtml

(2) <https://www.google.com/webmasters/tools/reconsideration?hl=es>

(3) <http://www.google.com/webmasters/tools/>

5.3 Yahoo!

En la actualidad Yahoo! no dispone de un formulario en español para solicitar reinclusiones, aunque sí dispone uno en inglés⁴.

5.3.1 Yahoo! SearchScan

Los resultados de Yahoo! también se revisan a nivel de seguridad para el usuario, ya sea por descargas indebidas, posible envío de spam... en el caso de que tu sitio esté marcado y no dispongas de ninguna de estas inseguridades, es recomendable seguir los siguientes pasos.

1. Entra en el formulario de contacto⁵.
2. Rellena los datos.
3. En el *Tema de la pregunta* selecciona *Discusión de clasificación*.
4. Explica con detalle los elementos necesarios:
 - Los términos de búsqueda que has utilizado para generar los resultados.
 - Algún detalle que hayas corregido de algo que estuvieras haciendo de forma incorrecta.
5. Finaliza el formulario y envíalo.

5

5.4 Windows Live

1. Entra en el formulario de contacto⁶.
2. Rellena los datos y tu dirección de correo electrónico.
3. En el *Tipo de problema* selecciona *Mi sitio ya no aparece en la clasificación o el índice*.
4. Explica con detalle los siguientes elementos:
 - Los términos de búsqueda que has utilizado para generar los resultados.
 - La dirección URL de tu sitio.
 - Algún detalle que hayas corregido de algo que estuvieras haciendo de forma incorrecta.
5. Finaliza el formulario y envíalo.

(4) <http://help.yahoo.com//us/yahoo/search/urlstatus.html>

(5) http://help.yahoo.com//es/yahoo/search/search/security/search_scan.html

(6) https://support.live.com/eform.aspx?productKey=wsearch&page=wlsupport_home_options_form_byemail&ct=eformts





Los pequeños detalles

Los buscadores son una fórmula matemática y para que no se descubra cómo funciona su sistema, la modifican cada cierto tiempo pero, aunque se modifiquen ciertos aspectos, si sigues los consejos de esta guía tendrás muchas probabilidades de llegar al éxito de una forma eficaz por muchos cambios de algoritmo que se hagan.

6.1 Densidad de Palabras

Uno de los secretos al redactar un texto es la densidad de palabras. Introducir en un texto varias veces y de forma distribuida ciertas palabras harán que esos contenidos tengan valor por sí mismos, de forma que podrás conseguir la densidad de palabras correcta.

Los buscadores centran muchos esfuerzos en valorar los contenidos mediante enlaces externos, pero debido a la gran cantidad de resultados y contenidos no deseados se está volviendo a dar mucha importancia a los contenidos originales bien redactados y formateados.

Aunque en algunos casos se habla de que para aparecer bien posicionado la densidad de palabras ha de estar entre el 2% y el 3%, estas cifras nunca han de tomarse al pie de la letra, ya que puede darse el caso de que en un texto tengamos que redactar más una palabra o ciertos grupos de palabras y quedaría algo poco natural si te ciñes a esas cifras.

6.2 Alta en motores de búsqueda

Darse de alta en los buscadores es muy simple y se hace en pocos minutos. De todas formas mi mayor consejo no es darse de alta en ellos (que es bueno, no cabe duda) sino el conseguir enlaces, es decir, que otras páginas pongan (por méritos propios si puede ser) un enlace a tu sitio web. Lógicamente si tus contenidos son de calidad nadie dudará en enlazar tu sitio web.

También has de tener una cosa clara, el 95% de las búsquedas que se hacen en internet se centran en 4 o 5 buscadores, por lo que son absurdos esos anuncios que prometen poner tu sitio web en más de 100 buscadores e incluso en más de 1.000...

Recuerda que al dar de alta un sitio, sólo has de hacerlo con la página principal y no con todas las páginas existentes en tu sitio web.

- ▶ **Exalead**
<http://www.exalead.es/search/submitYourSitePage>
- ▶ **Google**
<http://www.google.es/addurl/>
- ▶ **Live**
<http://search.live.com/docs/submit.aspx>
- ▶ **Yahoo!**
<http://es.siteexplorer.search.yahoo.com/free/request>

Eso sí, has de tener paciencia, ya que aunque des de alta de forma manual el sitio web es posible que tarden en visitarla entre una semana y un mes hasta aparecer los resultados.

6

6.3 Sitemaps

Una de las características que ha aparecido hace poco en los buscadores es la posibilidad de informarles de qué direcciones tenemos en el sitio. Son un punto de partida ya que una vez allí seguirán buscando más páginas.

El protocolo Sitemaps¹ se ha convertido en un estándar para Ask, Exalead, Google, Live y Yahoo!.

```

1.  <?xml version="1.0" encoding="UTF-8"?>
2.  <urlset xmlns="http://www.sitemaps.org/schemas/sitemap/0.9">
3.    <url>
4.      <loc>http://www.example.com/</loc>
5.      <lastmod>2005-01-01</lastmod>
6.      <changefreq>monthly</changefreq>
7.      <priority>0.8</priority>
8.    </url>
9.  </urlset>

```

Cada uno de los elementos de dicho protocolo y su correcta configuración han de revisarse correctamente. Por ejemplo, la página principal de un sitio debería tener como frecuencia de actualización *always* y una prioridad de *1*, los contenidos finales una actualización *weekly* y una importancia de *0.8* y las categorías o páginas de navegación una actualización *monthly* y una prioridad de *0.5*, por ejemplo.

6.4 Aviso manual de actualización

Aunque los propios buscadores donde los sitemaps están dados de alta ya se encargan de revisar cada cierto tiempo su actualización, no está de más que, si no puedes avisar de forma automática cuando se actualice tu sitio, lo hagas a través de las diferentes direcciones que han puesto a disposición los buscadores:

- ▶ **Ask**
<http://submissions.ask.com/ping?sitemap=URLsitemap>

(1) <http://sitemaps.org/es/protocol.php>



▶ **Google**

<http://www.google.com/webmasters/sitemaps/ping?sitemap=URLsitemap>

▶ **Yahoo!**

<http://search.yahooapis.com/SiteExplorerService/V1/updateNotification?url=URLsitemap>

▶ **Windows Live**

<http://webmaster.live.com/ping.aspx?siteMap=URLsitemap>

En este caso tendrás que sustituir “URLsitemap” por la dirección completa de tu fichero de Sitemap.

6.5 Panel de gestión

Los 3 grandes buscadores disponen de sus herramientas para webmasters que permiten gestionar los Sitemap.

▶ **Google Webmaster Tools**

<http://www.google.com/webmasters/tools/>

Google Herramientas para webmasters javier@casares.org | Mi cuenta | Ayuda | Salir

Panel > Sitemaps

Sitemaps
javiercasares.com ▾

Estadísticas de Sitemaps:
Total de URL: 995
URL que se han indexado: 926

Añadir un Sitemap 1 - 1 de 1

Nombre de archivo	Formato	Tipo	Enviado	Descargado por última vez	Estado	URL enviadas
<input type="checkbox"/> sitemap.xml.gz		Sitemap web	14-nov-2007	Hace 5 horas	OK	995

Eliminar seleccionados | Volver a enviar selección

[Descargar esta tabla](#)
[Descargar los detalles de Sitemap de todos sus sitios web como archivo csv](#)

▶ **Live Search Webmaster Tools**

<http://webmaster.live.com/>

Live Search | Webmaster Tools

Site summary for - javiercasares.com

Your websites
Click to expand

Tools
Validate robots.txt

Help
Forum
Blog
Webmaster Help

FAQ
See all your websites

Summary Profile Keywords Top Outbound Links Top Backlinks **Sitemap**

Add a sitemap for better results.
Sitemaps help the Live Search robot find all of the files to be indexed. (more at <http://www.sitemaps.org/protocol.php#informing>). You'll get the best indexing results by using robots.txt autodiscovery.

Other ways to add a sitemap:
Sitemap ping service. Notify the Live Search crawler if you change an existing sitemap, or add a new one. Use this web address format:
[http://webmaster.live.com/webmaster/ping.aspx?siteMap=\[Your sitemap web address\]](http://webmaster.live.com/webmaster/ping.aspx?siteMap=[Your sitemap web address])
Submit service. You can also submit a sitemap directly.

Sitemap web address:
javiercasares.com/sitemap.xml.gz
Submit

▶ **Yahoo! Site Explorer**

<http://siteexplorer.search.yahoo.com/>



6.6 Aumentar la velocidad de indexación

Google dispone de uno de los robots de indexación más rápido, aunque ello no significa que suponga directamente posicionarse por los contenidos de esa información. Aún así, el hecho de aparecer de forma fresca por algunos contenidos puede ser muy interesante de cara a ser el primero en hablar de un tema o actualidad.

Para conseguir aumentar la velocidad de indexación lo mejor es seguir una frecuencia de actualización (incluso varias veces al día), tener tráfico en tu sitio (un sitio que tiene pocas visitas al día es más complicado que sea actualizado con frecuencia) y estar posicionado en palabras que tienen cierta competencia.

En algunos casos puedes encontrarte frases de búsqueda en las que aunque no haya muchas búsquedas hay mucha competencia. Es en estos casos en los que habitualmente te encontrarás con sitios que se actualizan con mucha frecuencia.

6.7 Cambiar la dirección web o dominio

Una de las cosas que puedes necesitar es cambiar tu sitio web de una dirección de algún servidor gratuito a otro o que un cambio de marca o de producto haga que tengas que cambiar su dirección o incluso todo el sitio, dominio incluido.

Esto lo puedes hacer mediante una redirección 301. Para hacer una redirección de este tipo necesitas modificar las cabeceras de todas las páginas web de forma que hagan algo de este estilo:

Ejemplo en PHP:

```

1. <?php
2. header("HTTP/1.1 301 Moved Permanently");
3. header("Location: http://www.nuevodomino.ext/");
4. ?>
    
```

Ejemplo en ASP:

```

1. <%
2. Response.Status="HTTP/1.1 301 Moved Permanently"
3. Response.AddHeader "Location", "http://www.nuevodomino.ext/"
4. %>
    
```

Con esto conseguirás dos cosas. Las páginas antiguas se desindexarán de los buscadores poco a poco ya que le estás diciendo que esa dirección se ha movido a otra nueva y le das la nueva



dirección donde encontrará los contenidos e indexará en poco tiempo.

Pero lo que harás realmente es reenviar todas las páginas del sitio a la página principal. En el caso en el que mantuvieramos todos los contenidos y sólo cambiase el dominio, puedes aplicar algo como esto:

Ejemplo en PHP:

```
1. <?php
2. $url = "http://www.nuevodomino.ext" . $_SERVER["REQUEST_URI"];
3. header("HTTP/1.1 301 Moved Permanently");
4. header("Location: $url");
5. ?>
```

6

En este caso el sistema tomará las direcciones antiguas y de forma transparente para el usuario y los robots de búsqueda lo mandará al nuevo dominio.

Además, de una forma similar podrías eliminar los problemas de tener contenido duplicado en un sitio “con WWW” o “sin WWW”. En este ejemplo dejarías que la web siempre tenga las WWW:

Ejemplo en PHP:

```
1. <?php
2. if($_SERVER["HTTP_HOST"] != "www.dominio.ext") {
3.     $url = "https://www.dominio.ext" . $_SERVER["REQUEST_URI"];
4.     header("HTTP/1.1 301 Moved Permanently");
5.     header("Location: $url");
6. }
7. ?>
```

Otra posibilidad que te da este tipo de redirección es la de eliminar los ficheros del estilo index.php o index.html. Si queremos que se eliminen esas páginas del índice podemos aplicar esta configuración en el fichero .htaccess de Apache.

Ejemplo en .htaccess:

```
1. RewriteCond %{THE_REQUEST} ^GET \./index.(php|html|html) HTTP
2. RewriteRule (.*)index.(php|html|html)$ /$1 [R=301,L]
```

6.8 Sitio “en mantenimiento”

Puedes necesitar dejar fuera de línea un sitio web por una actualización o por reformas que no te llevarán mucho tiempo, pero que si un robot de búsqueda entrase podría rastrear páginas mal formadas o encontrarse con errores que no ha de encontrar. A veces se suele poner una página inicial que dice que la web está en mantenimiento o en construcción, con la consecuencia de que esta página se indexa y deja de lado el contenido original. Además, los buscadores pasan por allí y, como ha dado una página correcta (Código 200) pues volverá de nuevo cuando toque habiendo indexado y mostrando ese mensaje de error. Si lo que quieres es que la araña no la indexe y vuelva a pasar en un rato y volver a intentarlo hay que devolver un Código 503 (Network Unavailable).

Ejemplo en PHP:

```
1. <?php
2. header("HTTP/1.1 503 Service Unavailable");
```

```
3. header("Retry-After: 3600");
4. ?>
```

Ejemplo en ASP:

```
1. <%
2. Response.Status="HTTP/1.1 503 Network Unavaible"
3. Response.Status="Retry-After: 3600"
4. %>
```

6.9 Enlaces y Pagerank

Sobre 1993 se empezaron a ver los primeros documentos en una especie de HTML. Había dos cosas que diferenciaba estos documentos del resto: los podría ver cualquiera y se podrían enlazar entre ellos.

Los buscadores siguen la filosofía de Internet... poder tener información y poder ir de un lugar a otro. Por ello siempre se dice que si quieres tener una buena web has de seguir la filosofía de Internet.

Con esto no quiero más que decir que uno de los detalles más importantes en tu sitio web son los enlaces que puedas tener en él. Es importante recibir enlaces, pero también lo es dar, porque no querrás convertirte en la última web de internet ¿verdad?, un lugar donde puedes llegar, pero de donde no puedes irte.

Google revolucionó los rankings de búsqueda con su tecnología Pagerank. Ésta consiste en un cálculo matemático que puntúa de 0 a 10 una página web según su importancia y uno de los factores más influyentes para calcularla es la cantidad de enlaces que recibe.

Por eso, si quieres que tu sitio salga bien posicionado tienes que conseguir que los diferentes sitios de Internet te enlacen. Lo ideal sería dejar que una web evolucione por sí misma, es decir, comentársela a amigos y conocidos, ponerla en las tarjetas de visita... haciendo que la gente que te rodea realmente vea que es interesante ya que, si lo es, seguro que por el “boca-oreja” se extiende.

Como ves, en ningún momento he comentado realizar el intercambio de enlaces, es decir, yo pongo un enlace en mi sitio web y tú pones un enlace en el mio. Se puede hacer, no cabe duda, aunque hacerlo “a lo loco” puede acarrear problemas. Si un sitio penalizado te enlaza es posible que te penalice también. Además, los enlaces importantes son aquellos que vienen de sitios que pueden tener una temática similar a la tuya sin ser competencia (o siéndolo). A este sistema de clasificación se le puede llamar por “clusters”, es decir, un grupo de páginas similares o con temática similar a la tuya y por ello darle mayor importancia.

Además, otra opción para ganar visitas es dar el sitio de alta en los directorios de enlaces. ¡Cuidado! No en cualquier directorio, sino en unos específicos. En la red hay dos grandes directorios importantes, DMOZ y Yahoo!. Estos dos directorios son revisados por “seres humanos” y por tanto tienen cierta importancia y credibilidad, de forma que cuando los buscadores los indexen sabrán que hay páginas de calidad.

6.9.1 Valor del Pagerank

Mucha gente todavía cree que el Pagerank tiene algún tipo de valor (en este caso me refiero al valor que aparece en la barrita verde de los navegadores). Lo cierto es que desde que se saben técnicas para modificarlo ya no sirve de mucho. Además, por mucho que Google diga que se utiliza para ordenar los resultados de su búsqueda cualquiera puede darse cuenta de que eso no es así.



Se supone que el Pagerank se calcula en base a los enlaces que un sitio web recibe, pero eso no se refleja en los backlinks del propio buscador, por lo que no tiene mucho sentido el que se siga utilizando el Pagerank como herramienta de cálculo, ya que además de actualizarse cada 3 meses (de media) a estas alturas no es indicador de nada; con una simple redirección se ha podido hacer subir a Pagerank 10 un sitio web.

Es muy similar a lo que ocurre con aquellos que utilizan Alexa para calcular la posición mundial de un dominio en base al tráfico que genera... También es muy fácilmente manipulable por lo que no tiene ningún sentido utilizarlo.

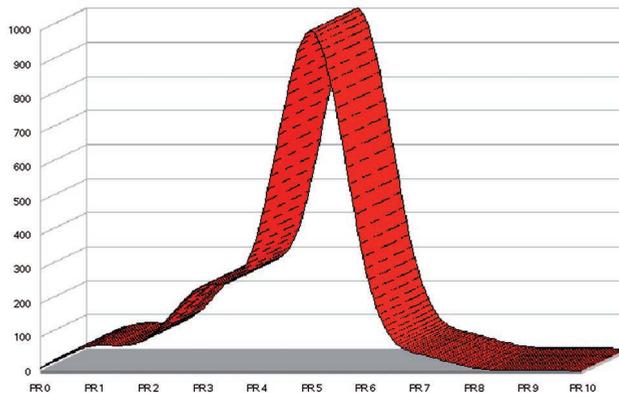
Lo que has de mirar es si la página que te enlaza es una web que hace las cosas bien... si es así, tendrá mucho más valor.

6

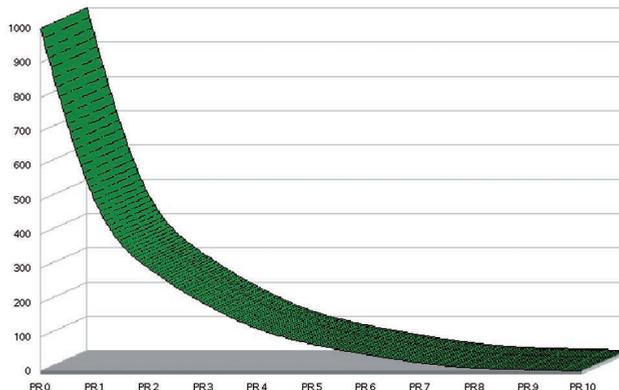
6.9.2 Pagerank como filtro AntiSpam

Uno de los problemas del algoritmo del Pagerank es que no es bueno detectando los enlaces que son spam y los que no. Aunque hay que tener una cosa en cuenta y es que, como algoritmo que es, o sea, fórmula matemática, si encuentra algún tipo de plantilla en la forma de enlazar puede penalizar.

Un ejemplo claro es que muchas veces se buscan enlaces de un mismo Pagerank, algo que llevaría a crear que nuestros enlaces tengan una forma similar a esta:



En este caso consigues muchos enlaces desde sitios con un mismo o similar Pagerank, lo que significa que se crea una campana que no es lo más lógico, ya que en Internet hay muchas páginas de poca importancia y pocas de mucha, por lo que el gráfico debería parecerse más a este:



6.10 Tiempo de permanencia

Uno de los factores clave a la hora de decidir si un sitio ha de estar por encima de otro en caso de que los dos sean similares es el de permanencia en el sitio. Esto significa, a grandes rasgos, que tiene más posibilidades de aparecer mejor un sitio en el que los usuarios pasan un tiempo mayor y no “entran y se marchan”.

Los buscadores tienen formas de hacer este tipo de mediciones, algunas más sencillas que otras.

La más sencilla es la de los propios resultados de búsqueda. Esto significa que si tú realizas una consulta y pulsas en el primero, pero no te gusta, vuelves a la lista y pulsas sobre el segundo y ya no sigues viendo más resultados, lo más probable es que el primer resultado no te interesaba y al no haber un clic en el tercero significa que el segundo resultado te ha significado la respuesta a tu consulta. ¿Por qué no cambiar ahora el primer por el segundo resultado si la calidad es mejor?

Otros métodos son las barras de herramientas. Cuando haces una consulta algunos buscadores hacen una petición de información sobre ese sitio. Esto significa que, siguiendo con el método que antes he explicado, la propia barra de herramientas puede saber en qué sitios vas navegando y el tiempo entre una petición y otra, es decir, el tiempo que pasas en el sitio.

Un último sistema de medición son los sistemas de estadísticas relacionados con los buscadores. Al fin y al cabo les estás dando todo tipo de información sobre el sitio y pueden hacer los cálculos necesarios para poder comparar dos sitios y elegir si uno es “más querido por los usuarios que otro”.

6.11 ¿Cómo ve mi sitio un buscador?

Es probable que quieras ver realmente cómo un buscador ve tu sitio web, el resultado de tanto esfuerzo. Te recomiendo el uso del navegador Lynx². Así podrás comprobar realmente lo que se está viendo, qué se indexa, y cómo mejorar la navegación del sitio web.

Otra opción es utilizar un lector -en audio- de escritorio, como el que utilizan los invidentes habitualmente, llamado JAWS³. Este lector, visita una web y la lee tal y como lo haría un robot, lo que puede dar pie también a saber cómo es tu sitio.

Como curiosidad, si visitas Google desde cualquier navegador, incluidos los de texto o audio, podrás comprobar que se ve igual que si navegas con uno moderno.

6.12 ¿Cuándo saldré posicionado?

Salir primero en un buscador no es tarea de un par de días... incluso puede llevar años...

Lo más probable es que aparezcas en los primeros puestos de un buscador y no de otros, porque cada uno tiene su forma de entender los contenidos, la estructura del sitio web, los enlaces... lo mejor es insistir, probar, probar, probar y tener mucha paciencia.

Una vez se aparece en un primer puesto lo que hay que hacer es un seguimiento de esa posición, ver sus fluctuaciones una o dos veces por semana (por no decir cada día), ver que hacen los que aparecen por encima y debajo nuestro, en definitiva, estudiar esa posible competencia.

Así que, como he dicho desde el principio, el posicionamiento no es una ciencia exacta, de forma que no existe un tiempo en el que aparecer... simplemente aparece.

(2) <http://lynx.isc.org/current/>

(3) <http://www.freedomscientific.com/products/fs/jaws-product-page.asp>



6.13 Indexación selectiva

Una de las necesidades en muchas ocasiones es la de que no se indexe determinadas partes de un sitio. Aunque no todos los robots lo permiten, sí que hay algunos de los importantes que actúan frente a algunos elementos.

6.13.1 Google Search Appliance

Aunque no se utiliza para los resultados de búsqueda normales, es interesante conocer que Google Search Appliance⁴ (su buscador para empresas) permite algunas directivas con las que filtrar determinados elementos de un sitio.

Para ello sólo hay que indicarlo de la siguiente manera:

```

Lorem ipsum dolor sit amet, <!--googleoff:all-->consectetur adipisicing elit,
<!--googleon:all-->sed do eiusmod tempor incididunt ut labore et dolore magna
aliqua.

```

En este caso, el contenido que hay entre ambos elementos no se utilizaría por Google Search Appliance en ninguno de los elementos de los resultados de búsqueda.

Aún así, el parámetro *all* puede cambiarse por *anchor*, *snippet* ó *index*.

6.13.2 Yahoo!

Yahoo! desarrolló el microformato robots-nocontent⁵ que indica que todo el contenido que haya en un bloque no sea considerado por el buscador.

```

Lorem ipsum dolor sit amet, <span class="robots-nocontent">consectetur
adipisicing elit, </span>sed do eiusmod tempor incididunt ut labore et
dolore magna aliqua.

```

Este sistema, al ser un microformato público, podría usarse en el futuro por más robots de indexación, algo que puede ser muy interesante a la hora de optimizar los sitios.

6.14 Buscadores de noticias

Parte de los buscadores tienen una sección específica de búsqueda de noticias. En el caso de que tengas una página en la que haya noticias de actualidad no está de más intentar que te den de alta en su plataforma ya que con esto conseguirás algunas cosas como que el sitio aparezca en más resultados de búsqueda y que la indexación del sitio sea más rápida.

6.14.1 Google Noticias

Para darse de alta como fuente en Google Noticias⁶ hay que rellenar el formulario⁷ habilitado para ello.

A parte de disponer de la versión web de este servicio también está disponible la posibilidad de añadir el sitio de noticias específicamente para Google Noticias Móvil⁸ que tendrá que darse de alta a través de una dirección de correo electrónico⁹.

(4) <http://www.google.com/enterprise/gsa/>

(5) <http://www.ojmicroformatos.com/microformatos/robots-nocontent/>

(6) <http://news.google.es/>

(7) http://www.google.es/support/news_pub/bin/request.py?contact

(8) <http://www.google.es/m/news>

(9) <mailto:mobilenews-source@google.com>

6.14.1.1 Sitemap para Google Noticias

Una vez hayas sido aceptado en el servicio, en la herramienta para webmasters te permitirán enviar un Sitemap para Noticias, que es igual que el Sitemap normal sólo que ha de cumplir algunos requisitos específicos¹⁰, como que las direcciones sólo pueden ser de noticias, pueden tener un máximo de 3 días de antigüedad y ha de actualizarse frecuentemente.

6.14.2 Yahoo! Noticias

En el caso de Yahoo! Noticias¹¹ es muy similar, aunque en este caso lo único que hay que hacer es enviar un correo electrónico¹² explicándoles el sitio que tienes y la información que ofreces.

(10) http://www.google.com/support/news_pub/bin/answer.py?answer=74287

(11) <http://es.noticias.yahoo.com/>

(12) <mailto:noticias@es.yahoo-inc.com>



B

CAPÍTULO

7

Úsqueadas y etiquetas

Uno de los factores clave para saber lo que más interesa a los visitantes de un sitio es saber lo que buscarían en él. Es por esto que es muy importante, no sólo tener un buscador propio, sino registrar todas las búsquedas que se realizan, con el seguimiento más completo posible (por ejemplo, todas las búsquedas de un mismo usuario).

Otro de los elementos es la forma de organizar la información. Hasta ahora la mayor parte de gestores de información sólo permitían establecer unas categorías fijas, pero ahora, con el concepto de las etiquetas todo queda en manos de la imaginación.

7.1 Buscador en el sitio

Los buscadores internos son muy interesantes por la información que pueden ofrecerte, sobretodo en sitios con mucha información, pero hay que tener mucho cuidado con el tratamiento que se les da.

Aunque, como decía antes, es muy interesante poder sacar información sobre las búsquedas que realiza la gente y poder sacar ideas, temas y mucha más información, también hay que tener en cuenta el contenido duplicado que puede generarse sin querer.

En estos casos hay que estudiar mucho este concepto y plantearse seriamente si merece la pena la indexación de este tipo de páginas o no. Aún así, el que no sean indexables no significa que no se puedan usar para ayudarnos a posicionar elementos que sería, en ocasiones, difícil de que aparezcan.

Para esto lo mejor es añadir, a las páginas de resultados de búsqueda, un metatag como el siguiente:

```
<meta name="robots" content="noindex, follow" />
```

Con esto conseguirás que esa página en concreto, que puede contener material duplicado, no se indexe, pero sí que se sigan los enlaces que aparezcan en ella para poder encontrar la información que sí es importante.

7.2 Las etiquetas

Una de las cosas buenas que tienen los resultados de búsqueda local y las etiquetas (en general conocidos como *tags*) es que al agrupar información de un mismo tema se genera el concepto de *cluster*, que en el mundo de la informática es la forma de agrupar contenidos.

Es por esto que una de las razones de utilizar etiquetas es su mayor facilidad de posicionarse en los buscadores, ya que al agrupar muchos contenidos de un mismo tema consiguen que esos enlaces internos hagan importantes a esas páginas de etiquetas, algo que puede ayudar a posicionarse por palabras genéricas.



SEO

para Blogs

CAPÍTULO

8

Los *blogs* se están presentando como una herramienta útil, rápida y eficaz para gestionar contenidos de una forma sencilla y que, a su vez, sean muy sencillos para los lectores.

8.1 Dominio y alojamiento

Si te planteas tener un blog profesional lo primero que has de hacer es tener tu propio dominio y un alojamiento web razonable. El dominio lo mejor es que lo registres y gestiones tú mismo en sitios como Godaddy¹ o Nic.es², donde tendrás control absoluto del mismo.

Con respecto al alojamiento intenta no utilizar uno gratuito. Hoy en día hay servicios de alojamiento realmente baratos y que por poco dinero al año te permiten una gran maniobrabilidad. Has de tener en cuenta que una de las cosas que priman los buscadores es que el blog esté siempre en línea y no de errores por saturación. Además, algunos servidores gratuitos o de baja calidad impiden que algunos de los robots de búsqueda más importante rastreen tu sitio.

Además, para que la mayoría de gestores funcionen a su mejor capacidad lo mejor es que el servidor sea compatible con Apache.

8.2 Los contenidos

Mantener un blog no es tarea sencilla sobretodo porque uno de sus secretos es el contenido original. Mantener una bitácora sencillamente para copiar contenidos de otros no tiene mucho sentido, por lo que lo ideal al montar una es explicar experiencias personales o comentar sobre los conocimientos profesionales o intereses que puedas tener. Al fin y al cabo, un usuario sigue un blog porque lo que se muestra allí le es interesante.

Un detalle a tener en cuenta es la longitud de los contenidos. Es muy probable que haya entradas que puedan ser muy breves porque no hay más que hacer, pero en otros casos seguro que te encuentras con artículos muy largos. ¿Por qué no separar las entradas en varios capítulos, separándolas por temas? Con esto ganarás en contenidos relacionados y seguro que si invitas a que tus lectores te sigan leyendo, lo harán.

(1) <http://www.godaddy.com/>

(2) <http://www.nic.es/>

Otro elemento a tener en cuenta es la actualización de los artículos. Aunque los lectores del día a día tal vez no lo consigan leer, sí que aquellos que lleguen desde motores de búsqueda o a través de enlaces directos sí que verán que vas manteniendo actualizada la información.

Y si lo que quieres es salir primero lo mejor es que comiences por cosas que aún no han ocurrido... o sea, haz un poco de vidente y escribe artículos en los que hables sobre eventos que van a ocurrir en un espacio breve de tiempo, pero con la suficiente antelación como para aparecer primero durante un tiempo significativo antes del mismo.

8.2.1 Frecuencia

Los blogs se rigen por tener contenidos más actualizados que muchos sitios web. Esto supone una exigencia personal pero que no necesariamente sea obligatoria.

El disponer de entre 3 entradas diarias y 3 entradas semanales es más que suficiente para que una bitácora se considere activa y sea revisada por los motores de búsqueda de forma frecuente.

¿Cuál es la mejor hora para publicar? Pues justo antes de las comidas... es decir, antes de desayunar, antes de comer y antes de cenar, ya que justo después, son los momentos en los que los lectores son más activos.

8

8.3 ¿Categorías o Etiquetas?

Una de las preguntas más frecuentes cuando se mantiene un blog es si organizar la información sólo por categorías, sólo por etiquetas o combinando ambas opciones. Como es normal, cada situación y blog tiene unas necesidades aunque el uso de etiquetas da una flexibilidad que hasta ahora no daban las categorías, mucho más estáticas.

Por norma general si en la bitácora se tratan pocos temas, lo mejor es hacer unas categorías generales y gestionar el resto de la navegación y organización creando etiquetas, mucho más dinámicas y adaptables a cada uno de los artículos.

8.4 Enlaces

Uno de los secretos de los blogs es el de los enlaces. En más de una ocasión habrás escuchado que conseguir enlaces es bueno, pero los blogs también funcionan por enlazar a sitios muy variados y a sitios muy relevantes, como es la Wikipedia³.

Un detalle importante al enlazar en los blogs es que los enlaces sean cortos e identificativos con el contenido de destino, ya que si pones, por normal general, un enlace de más de 5 palabras, puedes confundir al lector.

A parte, un elemento muy importante es citar las fuentes donde encuentras la información, ya que ese cruce de enlaces al final generará una mayor veracidad a tus contenidos. Intenta usar varias fuentes e incluso enlazar a varias en un mismo artículo. Puedes seguir sitios poco conocidos y usarlos como fuentes en caso de necesidad.

Para mejorar los enlaces internos también es interesante que hagas un listado de los artículos que más te haya gustado escribir y de los que te sientas orgulloso y los añadas de forma visible.

Algo que no debes olvidar nunca es enlazar a artículos propios que hablen de temas relacionados con lo que comentas en nuevos artículos. Estos enlaces, a parte de proporcionar impor-



(3) <http://es.wikipedia.org/>

tancia a los artículos, pueden hacer que los usuarios comiencen a navegar de forma profunda por el blog saltando de uno a otro.

8.5 Socialización

Un elemento de socialización para crear comunidad que, aún no está muy extendido al no estar todavía en español, es MyBlogLog⁴, una herramienta de Yahoo! mediante la cual podrás saber qué usuarios registrados visitan tu blog y cómo interactúan con él.

Otras herramientas útiles pueden ser AddThis⁵ que te permitirá mediante un botón enviar tus textos a multitud de sitios, además de indicar claramente los enlaces a los feeds.

(4) <http://www.mybloglog.com/>

(5) <http://www.addthis.com/>

SEO

para Wordpress

CAPÍTULO

9

9.1 Instalación

Para la instalación no hay nada más que hacer que seguir las instrucciones¹. Si vas a actualizar toma las medidas necesarias para no perder información y, sigue también las instrucciones².

Recuerda también instalar y configurar el idioma³ de Wordpress, para que antes de acceder a él por primera vez te aparezca todo en tu idioma.

9.2 Configuración

Aunque no voy a entrar en la configuración general de Wordpress, sí que voy a repasar todas aquellas opciones que, por defecto, ha de haber en el sistema para que puedas obtener una estrategia de posicionamiento favorable.

9.2.1 Creación del .htaccess

Para empezar a funcionar uno de los elementos que por defecto no viene con Wordpress es el fichero `.htaccess`. Lo mejor es crearlo, subirlo al servidor y, aunque esté vacío, darle permisos de escritura (CHMOD 666) para que el sistema pueda modificarlo a su gusto.

9.2.2 Administración

En el panel de Administración te encontrarás con algunas configuraciones que, por defecto, no vienen correctamente.

9.2.2.1 General

Aunque seguramente ya habrás puesto el título del blog al darlo de alta, no habrás puesto el “tagline”, que, tal y como él mismo dice, se trata de poner en pocas palabras de que trata el blog. Esta frase puede ser útil como frase de búsqueda que informe del creador del blog, de la temática, pero que no ha de exceder en 5-10 palabras.

Con respecto a la hora, recuerda elegir la Zona Horaria que corresponda (y que

(1) http://codex.wordpress.org/Installing_WordPress

(2) http://codex.wordpress.org/Upgrading_WordPress

(3) http://codex.wordpress.org/WordPress_in_Your_Language

cuando haya cambios horarios tendrás que modificar), además de dar un formato de fecha razonable (como puede ser “j F Y”) y un formato de hora también razonable (como puede ser “H:i”).

9.2.2.2 Escritura

Una de las opciones que has de tener activadas siempre es la de “WordPress corregirá de forma automática el HTML incorrectamente anidado”. Hay que tener en cuenta que si activas esta opción el código fuente de tus artículos o textos validarán, algo que siempre va bien para el posicionamiento.

Una lista que tendrás que ampliar es la de los “Servicios de Actualización”. Por defecto viene uno, aunque no está de más avisar directamente a los servicios más importantes, dejando una lista más parecida a la siguiente:

1. <http://rpc.pingomatic.com/>
2. <http://api.my.yahoo.com/rss/ping>
3. <http://blogsearch.google.com/ping/RPC2>
4. <http://ping.bitacorras.com>
5. <http://ping.feedburner.com>
6. <http://rpc.technorati.com/rpc/ping>
7. <http://rpc.weblogs.com/RPC2>

No puedes pasarte poniendo muchas direcciones ya que sino cuando publiques alguna novedad la página se puede saturar, ir muy lenta, y acabar fallando el envío.

9.2.2.3 Privacidad

Aunque al dar de alta el blog ya lo pide, no está de más comprobar que está activada la opción “Me gustaría que mi Blog sea visible por todos, incluyendo motores de búsqueda (como Google, Sphere, Technorati) y archivadores”.

9.2.2.4 Permalinks

Uno de los secretos de Wordpress son sus enlaces permanentes, es decir, la dirección que aparece en el navegador cuando entras en un artículo o página.

Por defecto Wordpress no puede configurarse si no es por parámetros, y aunque no es una mala opción, no es la mejor. Es por esto que anteriormente has creado el fichero `.htaccess`, para ahora poder hacer que la estructura de la dirección sea la correcta.

En las opciones generales elegirás la “Estructura personalizada” con un formato como el siguiente:

`/%postname%/`

En el caso de “Categoría base”, que será la dirección que aparezca al acceder a una categoría, puede tener un formato como:

`/tema`

Y en el caso de una “Etiqueta base” puedes utilizar el formato:

`/etiqueta`

En caso de tener correctamente todo, al actualizar no deberías de tener ningún tipo de mensaje.

9.3 Gestión de Plugins

Uno de los elementos más potentes de Wordpress es la cantidad de *plugins* que tiene y la sencillez de instalación.

Voy a comentarte algunos de los que pueden ser más interesantes y que, lógicamente, puedes decidir si instalar y activar, ya que aunque pueden ser recomendables para el correcto funcionamiento del blog, siempre hay que ver si son útiles para lo que necesitas. Aún así, los que te voy a comentar siempre irán enfocados a un planteamiento para el posicionamiento.

9.3.1 All in One SEO Pack

All in One SEO Pack permite gestionar de forma sencilla la estructura de títulos de toda la bitácora además de los títulos y metatags de cada una de las entradas en el sitio.

- ▶ Sitio de descarga:
<http://wp.uberdose.com/2007/03/24/all-in-one-seo-pack/>
- ▶ Versión probada: 1.4.6.8

La configuración consta de dos partes, una general y otra particular para cada artículo. En el caso de los artículos verás que aparece un menú en el que podrás añadir el título, descripción y palabras clave si no quieres que, en ese caso, se generen de forma automática.

9.3.1.1 Opciones del plugin

- ▶ **Título de la Página Principal**
Sustituye al que se le ha puesto por defecto, así que puedes dejarlo en blanco.
- ▶ **Descripción de la Página Principal**
Es interesante rellenar este campo con entre 50 y 160 caracteres. Cuando alguien busque el nombre de tu bitácora, será el texto que aparezca bajo el título en los resultados de búsqueda.
- ▶ **Palabras Clave de la Página Principal**
Separadas por comas, una lista de 6-10 palabras que describan los temas de la bitácora
- ▶ **Reescribir los Títulos**
Activando esta opción permites que el sistema use el título de cada artículo o el de este sistema si se le añade.
- ▶ **Formato del Título de Post**
`%post_title%`
- ▶ **Formato del Título de Página**
Si tienes páginas con mucho contenido puedes usar simplemente `%page_title%`. En el caso de usarla para contactos y similares, mejor ampliarlo con `%page_title% en %blog_title%`
- ▶ **Formato del Título de Categoría**
`%category_title% en %blog_title%`
- ▶ **Formato del título de Página de Archivo**
`%date% en %blog_title%`
- ▶ **Formato de Página de Tag**
`%tag% en %blog_title%`
- ▶ **Formato del Título para Búsquedas**
`%search% en %blog_title%`
- ▶ **Description Format**
`%description%` es el “excerpt” que genera automáticamente Wordpress
- ▶ **404 Title Format**
`Página no encontrada en %blog_title%` teniendo en cuenta que esto sólo lo verán usuarios, ya que la página no será indexada por los buscadores.



- ▶ **Paged Format**
 - [página %page%](#) se añade al final de los artículos que estén paginados.
- ▶ Usar Categorías como Palabras Clave
- ▶ No Indexar Páginas de Categoría
- ▶ No Indexar Páginas de Archivo
- ▶ No Indexar Páginas de Tag
- ▶ Generar Descripciones Automáticamente

No he marcado las opciones de no indexar ya que esta gestión la explicaré directamente en las plantillas, ya que el contenido duplicado va mucho más allá de lo que puede gestionar este plugin; si marcas las opciones los metatags que genera se podrían duplicar.

9.3.2 Askimet

Akismet revisa tus comentarios para comprobar si son o no spam. De esta forma evitarás de una forma sencilla el que aparezcan direcciones o contenidos no deseados por los usuarios o buscadores.

- ▶ Sitio de descarga:
<http://akismet.com/>

- ▶ Versión probada: 2.1.4

Su configuración es sencilla, ya que únicamente tendrás que activarlo y usar la clave que podrás conseguir junto a la API de Wordpress⁴.

9.3.3 Feed Statistics

Feed Statistics es un sistema sencillo para saber los usuarios suscritos a tu feed, al igual que hace Feedburner⁵, pero sin necesidad de externalizar el tráfico a terceros.

- ▶ Sitio de descarga:
<http://www.chrisfinke.com/wordpress/plugins/feed-statistics/>

- ▶ Versión probada: 1.3.1

Habitualmente sólo necesitas saber la cantidad de usuarios que están suscritos a nuestra bitácora y, si quieres, los artículos más leídos. La configuración por defecto ya es bastante correcta, aunque la voy a revisar.

9.3.3.1 Feed

Deja la cantidad de lectores marcada a los últimos 2 días.

En caso de querer saber si los usuarios pulsan en alguno de los enlaces del feed para ir a la bitácora activaremos “Track which links your subscribers click” aunque no es una buena idea ya que las direcciones de los artículos se volverán complejas y no ayudarán al posicionamiento. Además, el saber este dato, en principio, no aporta nada importante.

Un elemento que sí puedes dejar activo es “Track individual post views”, que inserta una pequeña imagen de seguimiento al final de cada artículo para saber cuánta gente lo ha visto, y de esa forma poder saber lo más leído a través de los feeds.

(4) <http://wordpress.com/api-keys/>

(5) <http://www.feedburner.com/>

9.3.3.2 Top Feeds

Wordpress deja por defecto elegir entre varios sistemas de feed, ya sea RSS o Atom. En esta opción podrás ver lo que tus usuarios eligen.

Your most popular feeds

	Feed URL	Subscribers
1.	/feed/atom/	168
2.	/feed/	33

9.3.3.3 Feed Readers

¿Desde dónde leen tus usuarios los feeds? Aquí tienes una lista con los lectores más habituales.

Top Feed Readers

	Reader	Subscribers
1.	Feedfetcher-Google	70
2.	Mozilla Firefox 2.0.0.14	33
3.	Netvibes	11
4.	Internet Explorer 7.0	8
5.	Bloglines	8

9.3.3.4 Post Views

Los artículos más leídos en los últimos 30 días y la cantidad de lecturas que ha tenido.

Your most popular posts (last 30 days)

	Post Title	Views
1.	Posicionamiento y CMS	205
2.	Cuidado con tus dominios y alojamiento web	203
3.	Cómo saber la IP real del usuario con PHP	201
4.	Feedburner y el peligro del contenido duplicado	191
5.	Futurama, pero de verdad	184

9.3.4 Google XML Sitemaps

Google XML Sitemaps genera de forma automática el fichero Sitemap compatible con los grandes buscadores y que facilita el encontrar las páginas del blog y, de esta forma, su indexación.

▣ Sitio de descarga:

<http://www.arnebrachhold.de/redir/sitemap-home/>

▣ Versión probada: 3.0.3.3

Aunque es básico tener este sistema, hay que tener cuidado sobretodo en blogs muy grandes (con muchos artículos).

Antes de configurarlo, tendrás que crear un fichero llamado “sitemap.xml.gz” (puede no contener nada) en la carpeta raíz de la instalación y darle permisos de escritura (CHMOD 666). Además, también deberías darle permisos de escritura al fichero “robots.txt” que estará en la carpeta raíz del dominio.

9.3.4.1 Basic Options

▣ Sitemap files

[] Write a normal XML file (your filename)

[X] Write a gzipped file (your filename + .gz)



- ▣ **Building mode**
 - Rebuild sitemap if you change the content of your blog
 - Enable manual sitemap building via GET Request
- ▣ **Update notification**
 - Notify Google
 - Notify Live Search
 - Notify Ask
 - Notify Yahoo!
 - Modify or create robots.txt file
- ▣ **Advanced options**
 - Limit the number of posts in the sitemap 1000
 - Enable MySQL standard mode.
 - Build the sitemap in a background process

9.3.4.2 Post Priority

A menos que tu blog tenga muchos comentarios y quieras priorizar estas entradas, lo mejor es no priorizar nada.

- ▣ **Post Priority**
 - Do not use automatic priority calculation
 - Comment Count
 - Comment Average

9.3.4.3 Sitemap Content

Para evitar contenidos duplicados, lo mejor es no avisar de categorías, archivos y autores.

- ▣ **Sitemap Content**
 - Include homepage
 - Include posts
 - Include static pages
 - Include categories
 - Include archives
 - Include tag pages
 - Include author pages

9.3.4.4 Change frequencies

- ▣ **Change frequencies**
 - always Homepage
 - daily Posts
 - weekly Static pages
 - monthly Categories
 - never The current archive of this month
 - never Older archives
 - daily Tag pages
 - monthly Author pages

9.3.4.5 Priorities

▣ Priorities

- [1.0] Homepage
- [0.8] Posts
- [0.6] Minimum post priority
- [0.8] Static pages
- [0.2] Categories
- [0.2] Archives
- [0.6] Tag pages
- [0.2] Author pages

9.3.5 Pagebar2

Pagebar2 mejora la navegación de las páginas del blog.

▣ Sitio de descarga:

<http://www.elektroelch.de/hacks/wp/pagebar>

▣ Versión probada: 2.20

La configuración es muy sencilla ya que se puede hacer de forma casi automática. Lo bueno de este sistema es que reduce la cantidad de niveles de navegación de los buscadores para encontrar la información necesaria.

9

9.3.5.1 Basic settings

▣ Basic settings

- Left [3]
- Center [5]
- Right [3]
- Leading text [Páginas:]
- Standard page: [{page}]
- Current Page [{page}]
- First page [{page}]
- Last page [{page}]
- Connector [...]
- Previous [<<]
- Next [>>]
- Tooltips [Página {page}]
- Insert pagebar automatic into blog
- Front of postings
- Behind postings
- Footer
- Remove standard navigation

Los valores se pueden retocar en base a la cantidad de artículos que haya en el blog... si hay menos de 100 artículos se pueden reducir los valores a la mitad (por ejemplo).



9.3.6 WordPress Related Posts

WordPress Related Posts permite de forma automática mostrar otras entradas relacionadas con la actual de forma automática, para así mejorar los enlaces internos del blog.

- ▣ Sitio de descarga: <http://fairryfish.net/2007/09/12/wordpress-23-related-posts-plugin/>
 - ▣ Versión probada: 0.7
- La configuración es muy sencilla
- ▣ Related Posts Title: **Contenidos Relacionados**
 - ▣ When No Related Posts, Display: **Text No Related Posts**
 - ▣ No Related Post's Title or Text: **No hay contenidos relacionados**
 - ▣ Limit: **5**
 - ▣ Other Setting:
 - Auto Insert Related Posts
 - Related Posts for RSS
 - Display Comments Count
 - Display Post Date

Una vez tengas esto activado tendrás que añadir a la plantilla (single.php) el siguiente código:

```
<?php wp_related_posts(); ?>
```

9.4 Plantillas (Diseño)

Cada una de las plantillas de Wordpress puede ser un mundo y como hay tantas voy a centrarme en explicar la plantilla “por defecto” (default) que trae la plataforma. Lógicamente todo lo que voy a explicar se puede llevar a cualquier otra plantilla.

Uno de los detalles importantes es aplicar todo lo que he comentado en los capítulos de factores internos y externos, como por ejemplo el de ponerle título a los enlaces e imágenes (que por defecto no suele venir). Estos detalles no los comentaré de nuevo en esta parte, ya que me centraré principalmente a las mejoras en la estructura.

9.4.1 Encabezado (header.php)

Lo primero que has de revisar en la plantilla es la cabecera. Hay que aplicar lo que ya comenté páginas atrás sobre la forma que ha de tener cualquiera de las cabeceras, aunque en este caso hay que adaptarla a algunas peculiaridades de Wordpress.

```
1. <!DOCTYPE html PUBLIC "-//W3C//DTD XHTML 1.0 Transitional//EN" "http://
www.w3.org/TR/xhtml1/DTD/xhtml1-transitional.dtd">
2. <html xmlns="http://www.w3.org/1999/xhtml" <?php language_attributes();
?>>
3. <head profile="http://gmpg.org/xfn/11">
4. <title><?php wp_title(); ?></title>
5. <meta http-equiv="Content-Type" content="<?php bloginfo('html_type');
?>; charset=<?php bloginfo('charset'); ?>" />
6. <link rel="alternate" type="application/atom+xml" title="Atom de
<?php bloginfo('name'); ?>" href="<?php bloginfo('atom_url'); ?>" />
7. <link rel="alternate" type="application/rss+xml" title="RSS de <?php
bloginfo('name'); ?>" href="<?php bloginfo('rss2_url'); ?>" />
```

```

8. <link rel="stylesheet" href="<?php bloginfo('stylesheet_url'); ?>"
   type="text/css" media="screen" />
9. <link rel="pingback" href="<?php bloginfo('pingback_url'); ?>" />
10. <?php
11. wp_head();
12. if(is_paged() or is_search() or is_archive()) {
13. ?>
14. <meta name="robots" content="noindex, follow" />
15. <?php
16. }
17. ?>
18. </head>
19. <body>

```

Como he comentado anteriormente en el plugin *All in Pack SEO* he introducido directamente código de programación mediante el cuál no dejaremos que se indexen páginas de resultados paginados, de resultados de búsqueda o de archivos, de forma que eliminamos páginas poco útiles a la hora de posicionar, pero sí dejamos que los enlaces que hay en ellas sean seguidos por los robots de búsqueda, ya que pueden tener los contenidos que sí son importantes.

9.4.2 Pie de página (footer.php)

El pie de página incluye algunos elementos que no deben repetirse en todas las páginas. Además incluyen algunos comentarios que es mejor eliminar. De esta forma se simplifica mucho el resultado.

```

1. <?php
2. wp_footer();
3. ?>
4. <!-- Si has de poner el código de Google Analytics insértalo aquí -->
5. </body>
6. </html>

```

9.5 WYSIWYG o no WYSIWYG

Uno de los problemas habituales a la hora de introducir contenidos en Wordpress es el uso del editor WYSIWYG que lleva activado por defecto el gestor. Lo mejor para todos es que el editor no esté activado, de forma que se pueda copiar y pegar cualquier tipo de código sin necesidad de más opciones. Además, el código será mucho más limpio y el HTML sencillo es comprensible para cualquier persona.

Un detalle con el que te puedes encontrar es que, si dejas activado el editor y copias por ejemplo desde otro sitio web o desde un documento de Word, podrás encontrarte código fuente absurdo que sólo se convertirá en un impedimento para una buena edición y contenido de calidad.



9.6 Blogroll

Una de las funcionalidades de los blogs es el Blogroll, que es la lista de enlaces que aparece habitualmente en el menú lateral. Un detalle que hay que tener en cuenta es que esta lista no debería aparecer nada más que en la página principal, ya que sino distorsionaremos de formas muy variadas los resultados de los buscadores de blogs.

Para que esto sea así, puedes ir a la plantilla **Barra Lateral (sidebar.php)** y sustituir por lo siguiente:

```
1.  <?php
2.  if(is_home()) {
3.      wp_list_bookmarks();
4.  }
5.  ?>
```

GEOSEO

(SEO Local)

CAPÍTULO

10

Uno de los elementos más importantes a la hora de posicionar es tener en cuenta si el sitio está enfocado a un país, a un idioma, en definitiva, a un territorio.

10.1 Dominios, SubDominios o Carpetas

La forma de organizar la información es muy importante. Cada situación es un mundo y tendrás que elegir la opción más adecuada para tu caso.

10.1.1 Dominios

Los dominios se utilizan para agrupar la información principal. En caso de tener productos o servicios enfocados a un país concreto lo mejor es usar los dominios territoriales. De esta forma tendrás un dominio principal con la información general y corporativa, y en cada uno de los dominios territoriales los servicios y productos de este país en concreto.

10.1.2 SubDominios

Los subdominios han de utilizarse en dos casos completamente diferentes.

Por un lado, en caso de no tener disponibles los dominios territoriales por la razón que sea, se pueden crear los contenidos en subdominios haciendo referencia al TLD de cada país, es decir, *ar.dominio.ext* para Argentina o *fr.dominio.ext* para Francia.

Por otro lado, en caso de enfocar tu sitio a un idioma y no a un país, lo mejor es usar los subdominios, pero en este caso con los códigos de idioma. Por ejemplo, un *es.dominio.ext* nos serviría para España, Argentina o México.

Esto hay que hacerlo así porque, por defecto, a un dominio se le asigna un idioma y todas las carpetas del mismo tendrán ese idioma por defecto. Eso no significa que con el paso del tiempo el buscador detecte el idioma del que se trata, pero es mucho más sencillo gestionarlo de esta forma y facilitarle el trabajo a los robots de indexación.

10.1.3 Carpetas

Las carpetas han de utilizarse para agrupar contenidos, es decir, dentro de cada carpeta podremos agrupar contenidos de una misma temática o tipo de contenidos.

10.2 El dominio y la IP

Uno de los elementos más importantes para posicionar un sitio en una edición local y no en otra es, o que el dominio que se utilice haga referencia al país, es decir, para aparecer bien en Google España debes usar un dominio .es y para aparecer mejor en Yahoo! Francia debes usar un dominio .fr.

Pero como esto no siempre es posible, otra posibilidad de conseguir este mismo efecto de una forma automática es alojar el sitio en una IP de ese país. Esto puede ser más costoso pero permite, por ejemplo, que en caso de no tener los dominios territoriales puedas aplicar, a cada subdominio, una IP diferente en un país diferente.

En algunos casos, si realmente necesitas posicionar al máximo la información, puedes plantearte tener los servidores DNS en la IP del mismo dominio, es decir, que hasta la DNS esté geolocalizada.

10.3 Mapas y Negocios

Cada vez más los buscadores están geolocalizando las búsquedas desde su lugar de origen, es decir, que si realizas una búsqueda desde una ciudad te muestra contenidos de esa ciudad, y alguien que viva en la ciudad de al lado verá resultados de esa ciudad. Esto es posible gracias a sistema de geolocalización de IP.

Es muy importante aparecer en los directorios locales de los grandes buscadores. Por ejemplo, Live¹ y Yahoo!² utilizan los datos de Páginas Amarillas³, y aunque Google⁴ también lo hace, está comenzando a utilizar sus propios datos gracias al Local Business Center⁵, una herramienta con la que podrás dar de alta tu negocio de forma gratuita.

10.4 Microformato Geo, hCard y hCalendar

Los buscadores, encabezados por Yahoo!, están comenzando a utilizar los microformatos para ampliar la información de sus resultados de búsqueda.

10.4.1 Geo

Con este microformato⁶ puedes indicar la localización de un elemento. Tan sólo necesitarás usar la *clase geo*.

1. `<div class="geo">Lugar:`
2. `41.405318,`
3. `2.176859`
4. `</div>`

(1) <http://local.live.com/>

(2) <http://es.search.yahoo.com/search/yp>

(3) <http://www.paginasamarillas.es/>

(4) <http://local.google.es/>

(5) <http://www.google.es/local/add/>

(6) <http://www.ojomicroformatos.com/microformatos/geo/>

10.4.2 hCard

Con este microformato⁷ puedes indicar la información de una persona o empresa. Tan sólo necesitarás usar la *clase vcard* combinada con la *clase geo*.

```

1. <div class="vcard">
2.   <a class="url fn" href="http://javiercasares.com/">Javier Casares</a>
3.   <div class="adr">
4.     <span class="type">Work</span>:
5.     <div class="street-address">Ap. Correos 2017
6.     <span class="locality">Badalona</span>,
7.     <abbr class="region" title="Barcelona">BCN</abbr>
8.     <span class="postal-code">08917</span>
9.     <span class="country-name">Spain</span></div>
10.  </div>
11.  <div class="tel">
12.    <span class="type">Work</span> +34.555.123.456
13.  </div>
14.  <div class="geo">Lugar:
15.    <span class="latitude">41.40370</span>,
16.    <span class="longitude">2.17433</span>
17.  </div>
18. </div>

```

10

10.4.3 hCalendar

Con este microformato⁸ puedes indicar la información de un evento. Tan sólo necesitarás usar la *clase vevent* combinada con la *clase geo*.

```

1. <div class="vevent">
2.   <a class="url" href="http://forumdominios.com/congreso/">http://
forumdominios.com/congreso/</a>
3.   <span class="summary">Congreso ForumDominios</span>
4.   <abbr class="dtstart" title="2008-05-12">12 de Mayo de 2008</abbr>-
5.   <abbr class="dtend" title="2008-05-13">13 de Mayo de 2008</abbr>,
6.   en <span class="location">Citilab, Cornellà de Llogregat, Barcelona</
span>
7.   <div class="geo">Geo:
8.     <span class="latitude">41.40370</span>,
9.     <span class="longitude">2.17433</span>
10.  </div>
11. </div>

```

(7) <http://www.ojomicroformatos.com/microformatos/hcard/>

(8) <http://www.ojomicroformatos.com/microformatos/hcalendar/>



SERM

Search Engine Reputation Management

CAPÍTULO

11

La gestión de la reputación en los buscadores es una especialidad muy concreta del posicionamiento orgánico que consiste en tener controlados todos los resultados de búsqueda de los diferentes buscadores al realizar una consulta sobre el nombre de una persona, una empresa o una marca, intentando que todos los resultados tengan comentarios o información correcta y positiva.

Uno de los casos más conocidos en España es el de IKEA. Si realizas la búsqueda [ikea]¹ en Google encontrarás que, bajo el sitio de esta empresa, aparece un artículo de Microservios² en el que explican una situación que tuvo lugar en uno de estos centros.



La Web

[IKEA | IKEA España | muebles complementos salones dormitoriosos ...](#)

Cadena de tiendas de muebles y complementos de decoración a precios asequibles.
www.ikea.com/es/es/preindex.html - 5k - [En caché](#) - [Páginas similares](#) - [Anotar esto](#)

[IKEA | Página principal IKEA, muebles, decoración, cocinas ...](#)

[IKEA](#) Página principal [IKEA](#), muebles, decoración, cocinas, dormitorios, salones, oficinas, textiles, niños, baños, sofás, sillas, mesas, librerías, ...

www.ikea.com/es/es/ - 31k - [En caché](#) - [Páginas similares](#) - [Anotar esto](#)

[Más resultados de www.ikea.com >](#)

[Ikea: cómo mienten a los clientes | Microservios \(MundoReal™\)](#)

Mi mujer me obliga a publicar esto, que conste por adelantado. Esta mañana se indignó por cómo nos trataron en [Ikea](#), la tienda de muebles que hay en San ...

www.microservios.com/archivo/mundoreal/ikea-miente.html - 203k -

[En caché](#) - [Páginas similares](#) - [Anotar esto](#)

[IKEA - Wikipedia, la enciclopedia libre](#)

[IKEA](#) es una empresa sueca multinacional dedicada a la venta minorista de muebles y objetos para el hogar y decoración, a bajo precio y diseño contemporáneo. ...

es.wikipedia.org/wiki/IKEA - 36k - [En caché](#) - [Páginas similares](#) - [Anotar esto](#)

En este caso, el Search Engine Reputation Management lo que intentará es desban-car de los resultados de búsqueda este resultado añadiendo otros en los que no se hable negativamente.

(1) <http://www.google.es/search?q=ikea>

(2) <http://www.microservios.com/archivo/mundoreal/ikea-miente.html>

11.1 Detección

A parte de hacer una búsqueda por tu nombre, el de tu empresa o el de tus productos en los principales buscadores, hay muchas formas de estar alerta sobre lo que comentan sobre ti o tus relacionados.

La más sencilla es utilizar alguno de los servicios que los propios buscadores utilizan para crear alertas:

- ▶ Google Alertas³
- ▶ Windows Live Alertas⁴
- ▶ Yahoo! Alertas⁵ (sólo en inglés)

11.2 Comienza el movimiento

Hay muchas formas de atacar un problema como este. El primero es, nada más detectar el problema, intentar solventarlo contactando con el sitio web, ya no para que retire el artículo sino para intentar solventar el problema que lo ha generado. De esta forma, al menos, podrías llegar a conseguir que el artículo introduzca una nota que diga que se ha solventado el problema, algo que reducirá la mala imagen que pueda generar esa entrada mientras intentas desbancarla de los resultados.

Una vez hayas cumplido este primer trámite hay que considerar la cantidad de resultados y el esfuerzo que supondrá mejorar otros sitios para desbancar otros. Has de tener en cuenta que hay un factor en contra que juega en este campo, y es el factor tiempo, ya que si tardas mucho en darte cuenta de que existe un elemento distorsionados en los resultados luego costará más sacarlo de allí.

11

11.3 Ocupar los resultados

Una vez sepas la cantidad de resultados a eliminar la situación es clara: hay que poner esa misma cantidad de contenidos en los primeros puestos. Para ello hay que plantear estrategias como las de crear un blog o utilizar dominios territoriales. En este último caso creando subdominios o usando dominios territoriales no hay que superar el 60% de los resultados, es decir, 6 de los 10 primeros, ya que podría suponer una mala experiencia para el usuario y un cambio en los resultados por parte del buscador.

Si tienes detectadas algunas reseñas sobre tu sitio o tus productos que hablen bien, una opción para hacer que suban en los resultados de búsqueda y que se pongan por encima de otros resultados no deseados es enlazarlos desde el propio sitio web, ya que eso les dará más credibilidad.

Además, al ser el poseedor de la máxima información sobre ti o tus productos, podrías plantearte crear dominios separados con micrositiros en los que ofrecer un tipo de información algo diferente a la que puedes tener en el sitio web original. De esta forma podrás crear más sitios y resultados relacionados con tu marca o productos.

Otra posibilidad es la de acabar creando un concepto de **Brand Evangelist** con algunos de tus productos. ¿Qué es un Brand Evangelist? Es un usuario que ha creado un sitio que habla sobre

(3) <http://www.google.es/alerts>

(4) <http://alerts.live.com/>

(5) <http://alerts.yahoo.com/>



ti o sobre tus productos y, en este caso, has de cuidarlo, ya sea enviándole información, contestando a sus dudas, con productos y detalles exclusivos... Si tratas bien a los usuarios lo más probable es que los usuarios te traten bien a ti, así que podrás aumentar las posibilidades de que estos sitios queden entre los primeros resultados de búsqueda.

Otras opciones que puedes tener en cuenta son:

- ▶ Crear tu perfil en LinkedIn⁶, MySpace⁷, XING⁸...
- ▶ Crear vídeos optimizados y promocionarlos en Youtube⁹, Metacafé¹⁰, Dailymotion¹¹... y también de fotos en Flickr¹²...
- ▶ Crear un buscador personalizado con tus sitios en Rollyo¹³.

(6) <http://www.linkedin.com/>

(7) <http://www.myspace.com/>

(8) <http://www.xing.com/>

(9) <http://www.youtube.com/>

(10) <http://www.metacafe.com/>

(11) <http://www.dailymotion.com/>

(12) <http://www.flickr.com/>

(13) <http://www.rollyo.com/>

S M O

Social Media Optimization

CAPÍTULO

12

Internet cada vez más se está volviendo un medio social, en el que los internautas interactúan entre sí, ya sea mediante invitaciones, colegueo...

Es por esto que este componente social y tan “2.0” ha de tener un momento de dedicación aunque no afecte directamente al posicionamiento en buscadores, ya que hoy en día los buscadores son poco sociales de por sí, basándose en ofrecer contenidos sin la posibilidad de interactuar con ellos de una forma proactiva.

Aún así, hay que tener en cuenta que el objetivo final de un sitio y el posicionamiento de los mismos es para conseguir tráfico y usuarios fieles, por lo que si podemos aplicarle un componente social en el que los usuarios del sitio puedan interactuar con los contenidos mejoraremos su satisfacción además de incrementar la posibilidad de enlaces y de *buzz*.

12.1 Algunos detalles

12.1.1 Favicon

Algunos de los elementos que hoy en día hacen social a un sitio son los feeds. Muchos internautas utilizan servicios web para leerlos o programas de escritorio, programas que permiten, en muchos casos, la opción de ver el *favicon* del sitio. Este icono es el que aparece en la barra de direcciones de la mayor parte de navegadores.

Además de ser un buen sistema para crear un elemento distintivo del sitio, es tan sencillo como crear un fichero .ICO y añadir un metatag como el siguiente:

```
<link rel="shortcut icon" href="http://www.misitio.ext/favicon.ico" />
```

12.1.2 Versión para imprimir o descargar

Uno de los elementos sociales más interesantes son las versiones de impresión o la descarga en PDF de un contenido, algo que no supone esfuerzo en su creación pero que hemos de controlar para que no sean indexadas por los robots de búsqueda porque podrían encontrar varias veces el mismo contenido en el sitio en formatos muy diferentes.

12.2 Marcadores sociales

Los marcadores sociales o la posibilidad de enviar a sitios de puntuación son interesantes si, como todo, se trata de forma estudiada y no masiva.

Hoy en día hay varios sitios de importancia en la red a los que añadir marcadores como del.icio.us¹ o Mister Wong². Estos lugares permiten tener tu lista de sitios favoritos y son muy útiles para encontrar información clasificada por otros usuarios.

Añadir un botón a tu sitio web es muy sencillo, ya que tan sólo necesitarás una dirección como esta:

Dirección para del.icio.us:

```
http://del.icio.us/post?title=Titulo&url=http://misitio.ext/
```

Dirección para Mister Wong:

```
http://www.mister-wong.es/?action=addurl&bm_url=http://misitio.ext/&bm_description=Titulo
```

12.3 Sitios de recomendación

Los sitios de recomendación son una especie de redes sociales en las que los usuarios votan direcciones que hacen referencia a artículos, noticias y, en el fondo, cualquier tipo de contenido. Los usuarios votan de forma positiva o negativa y acaban apareciendo en la página principal de estos sitios, página que suele tener muchas visitas y que los robots de búsqueda indexan varias veces al día.

Los sitios para promocionar artículos de interés más importantes pueden ser Menéame³ o Fresqui⁴ en español y Digg⁵ o Yahoo! Buzz⁶ en inglés.

12

12.4 Preguntas y Respuestas

Desde hace ya un tiempo tanto algunos sitios como los propios buscadores están poniendo de moda las Preguntas y Respuestas. Estas páginas funcionan habitualmente por puntos y regalos y se basan en que un usuario que tiene una necesidad hace una pregunta y otros le dan una respuesta.

Lo interesante de estos sitios es que si eres un concededor de ciertos temas puedes comenzar a contestar y a usar contenidos de tu sitio como enlaces de referencia que van a implicar dos cosas: al ser sitios muy frecuentados por los buscadores, serán enlaces de gran valor y, además, conseguirás tráfico de calidad ya que serán usuarios que tienen dudas sobre esa temática.

Sitios a tener en cuenta son Yahoo! Respuestas⁷ y Windows Live QnA⁸, donde encontrarás multitud de categorías e ideas sobre temas a publicar.

(1) <http://del.icio.us/>

(2) <http://www.mister-wong.es/>

(3) <http://meneame.net/>

(4) <http://www.fresqui.com/>

(5) <http://digg.com/>

(6) <http://buzz.yahoo.com/>

(7) <http://es.answers.yahoo.com/>

(8) <http://qna.live.com/>

12.5 AJAX

Últimamente se está utilizando mucho la tecnología AJAX en los sitios web. Tal y como está desarrollado AJAX los buscadores no pueden indexar la información que se genera ya que esta se carga y descarga “al vuelo” y los robots no pueden llegar a ella.

La única posibilidad en estos momentos es la posibilidad de que el usuario pueda generar una dirección única para el contenido que está viendo en ese momento, tal y como hacen servicios como Windows Live Maps⁹ que permiten navegar por el sitio con estas tecnologías pero, que te ofrecen la posibilidad de darte una dirección única que, si utilizas directamente, mostrará los contenidos que tenías en ese momento en pantalla.

12.5 Páginas personalizadas

Gracias a la popularización de algunos sitios que permiten crearte una página personalizada es posible tener hojas en las que incluir tus sitios de noticias desde un mismo lugar.

Para ello puedes plantearte 2 opciones: crear un enlace directo en tu sitio que añada el feed de noticias o crear una página en cada uno de los servicios y compartir la dirección en tus sitios.

Los sitios más habituales y usados en la red son:

- ▶ **Google Personalizado**
<http://www.igoogle.com/>
- ▶ **netvibes**
<http://www.netvibes.com/>
- ▶ **Pageflakes**
<http://www.pageflakes.com/>
- ▶ **Windows Live**
<http://my.live.com/>
- ▶ **Mi Yahoo!**
<http://es.my.yahoo.com/>

(9) <http://maps.live.com/>



SEO

para vídeos

CAPÍTULO 13

Los formatos de vídeo no tienen un estándar claro como podría ser el MP3 de audio y esto complica bastante la posibilidad de que los buscadores de vídeo consigan indexar todos los formatos en todos los sistemas.

Aún así, optimizar un vídeo es factible si se siguen una serie de medidas para ello.

13.1 Tipos de buscadores

Existen 2 tipos de buscadores de vídeo según la tecnología que utilizan.

1. Primera generación

Son aquellos en los que la forma de organizar los resultados se basan en los metadatos del vídeo, es decir, los datos que puede incorporar el propio vídeo junto al contenido que hay en la página donde se encuentra (título, descripción, etiquetas, resumen, transcripción, resolución...)

2. Segunda generación

Aunque también incorporan los elementos anteriores, llevan la tecnología un paso más allá siendo ellos mismos los que utilizan reconocimiento de voz, análisis visual, reconocimiento del texto que puede aparecer por pantalla...

La mayoría son del primer tipo y los nuevos, con tecnología más avanzada, del segundo.

13.2 MetaDatos

En los vídeos existen dos tipos de metadatos, los que podemos integrar en el propio código del vídeo y los externos de la página donde se encuentra el vídeo. En ambos casos podemos optimizarlos para que sean los idóneos a la hora de ser indexados.

13.2.1 MetaDatos Internos

Todos los formatos de vídeo, en mayor o menor medida, incorporan sus propios metadatos. El título, descripción, calidad de codificación o incluso la transcripción del contenido son posibles.

Para revisar estos datos tendrás que hacerlo con el programa de edición de vídeo

con el que lo crees, o con programas como FLV MetaData Injector¹, Sorenson Squeeze² o Castfire³. Cada uno tiene unas utilidades y especificaciones especiales.

Pero cuidado, recuerda que si conviertes de un formato a otro podrías perder muchos de estos datos, así que, una vez los tengas en todos los formatos posibles, intenta comprobar que la información de cada uno es correcta.

Hay que tener en cuenta que un elemento interesante es tener el vídeo en cuantos más formatos posibles mejor, ya que así te asegurarías que todos los robots de búsqueda son capaces de encontrarlo e indexarlo.

13.2.2 MetaDatos Externos

En la mayor parte de casos hay que aplicar los mismos mecanismos que en el posicionamiento de una imagen o de un contenido textual.

▣ Título y descripción

Son los factores más importantes a la hora de posicionar un vídeo, ya que ahí encontrarán la mayor parte de la información necesaria. Los títulos han de ser claramente descriptivos y hay que eliminar toda palabra o frase que no sea útil.

▣ Nombre del fichero

Ha de ser descriptivo, incluyendo las palabras clave que describan al vídeo sin necesidad de ver su título o descripción. Lo ideal es separar esas palabras por guiones “-”.

▣ Etiquetas (tags)

En la página donde se encuentre el vídeo debe haber una lista de palabras clave enlazadas con el microformato rel-tag⁴. Estas palabras serán las que los buscadores utilicen como base a la hora de organizar su información.

13.2.3 Transcripción y Subtítulos

Aunque formaría parte de los elementos externos, son quizá dos elementos lo suficientemente importantes como para dedicarles especial atención. Aunque no son completamente un estándar, hay 2 formatos en los que guardar la información con un componente temporal en el que se especifica, uno para subtítulos y otro para transcripciones, que también puede usarse para los subtítulos.

Los formatos son SRT o SUB para los subtítulos y los TTXT para las transcripciones. Para gestionar este tipo de formatos es interesante el uso del programa MP4Box⁵ con el que se puede conseguir este tipo de ficheros y formatos.

13.3 Promoción

A parte de la indexación que puedan hacer los robots de búsqueda, puedes dar unos pasos para ofrecer de forma sencilla a algunos buscadores la lista de vídeos además de alguna información añadida sobre los mismos.

(1) <http://www.buraks.com/flvmdi/>

(2) <http://www.sorensonmedia.com/products/>

(3) <http://www.castfire.com/>

(4) <http://www.ojomicroformatos.com/microformatos/rel-tag/>

(5) http://gpac.sourceforge.net/doc_ttxt.php



13.3.1 Media RSS

El formato Media RSS⁶ propuesto por Yahoo! es un módulo de ampliación del formato RSS en el que se incluyen algunas etiquetas en las que dar información sobre elementos multimedia de páginas web.

Los elementos principales indican la dirección, tamaño, formato, duración, tamaño, idioma... Además, se pueden incluir otros elementos como un *thumbnail*, descripción, título, palabras clave, créditos e incluso la transcripción.

13.3.2 Sitemaps de vídeo

Al igual que existen los Sitemaps para indicar direcciones dentro de un sitio web, también hay una especificación que, por ahora sólo soporta Google, en la que se puede informar de los vídeos existentes. En este caso, el Sitemaps de vídeo⁷ es similar al Media RSS pero pensado para Sitemaps y no RSS.

Aquí puedes indicarle la dirección, la dirección con el reproductor, un *thumbnail*, título, descripción o duración.

13.3.3 Alta en buscadores

Al igual que hay estos dos formatos, el Sitemaps de vídeo por ahora sólo es soportado por:

▣ **Google Vídeo**⁸: Hay que dar de alta el Sitemap⁹

En cambio el formato Media RSS es soportado por otros como:

▣ **Truveo**¹⁰: hay que darlo de alta como Director¹¹

▣ **blinkx**¹²: hay que dar de alta el MediaRSS¹³

▣ **Yahoo!**¹⁴: hay que dar de alta el MediaRSS¹⁵

(6) <http://search.yahoo.com/mrss/>

(7) <http://www.google.com/support/webmasters/bin/topic.py?topic=10079>

(8) <http://video.google.com/>

(9) <http://www.google.com/webmasters/tools/>

(10) <http://www.truveo.com/>

(11) <http://developer.truveo.com/DirectorAccountsOverview.php>

(12) <http://www.blinkx.com/>

(13) <http://www.blinkx.com/rssupload>

(14) <http://video.search.yahoo.com/>

(15) <http://search.yahoo.com/mrss/submit>

SEO

para móvil

CAPÍTULO

14

El futuro de Internet pasa por la navegación desde dispositivos portátiles o móviles, y los grandes buscadores ya están trabajando en unas ediciones especiales de los mismos que organizan la información de una forma diferente a los resultados habituales.

¿Qué buscadores están comenzando a usar búsquedas específicas en sitios para móviles? Pues los principales ya lo tienen adquirido, e incluso algunos de ellos ya tienen una opción para sólo mostrar resultados adaptados a terminales móviles:

- ▣ Google Mobile¹: dispone de una opción para filtrar contenidos sólo para móviles.
- ▣ Live Search Mobile²
- ▣ Technorati Mobile³
- ▣ Yahoo! Mobile⁴: dispone de una opción para filtrar contenidos sólo para móviles.

Además, desde hace poco existen los dominios .MOBI, que obligan a utilizar una especificación concreta y seguir una normativa restrictiva en cuanto a la forma de mostrar los contenidos. Así que es recomendable utilizar este tipo de dominios a la hora de posicionar un sitio móvil.

14.1 XHTML Mobile

Uno de los primeros elementos a tener en cuenta es que el formato del sitio no puede ser igual que el normal. Existe un XHTML pensado para los terminales móviles en los que no se permiten algunos elementos frecuentes en otras versiones. Para conocer todo sobre este tipo de sitios puedes visitar dev.mobi⁵, el sitio para desarrolladores.

Algunas de estas limitaciones son:

- ▣ Ancho de pantalla: mínimo 120 pixels
- ▣ XHTML Basic 1.1 (application/xhtml+xml)

(1) <http://www.google.es/m/>

(2) <http://m.live.com/>

(3) <http://m.technorati.com/>

(4) <http://es.m.yahoo.com/>

(5) <http://dev.mobi/>

- ▣ Codificación: UTF-8
- ▣ Formato de imágenes: JPEG y GIF 89a
- ▣ Peso máximo de la página: 20 kilobytes
- ▣ 256 Colores (mínimo)
- ▣ CSS 1 y 2
- ▣ HTTP/1.0 y HTTP/1.1
- ▣ No se permiten scripts

En este caso es muy importante que el sitio web valide correctamente, ya que aunque los navegadores para escritorio permiten saltarse algunas de estas normas, los navegadores para móviles son mucho más estrictos.

Para comprobar que un sitio es compatible y para ver cómo se ve en algunos terminales móviles es recomendable usar el validador ready.mobi⁶.

14.2 Alta en buscadores

Tal y como comentaba antes, hay buscadores que ya soportan este formato completamente y, por eso, permiten dar de alta los sitios web para móviles:

- ▣ **Alta en Google**
Ha de hacerse desde sus herramientas y mediante un Sitemap Móvil⁷.
- ▣ **Alta en Yahoo!**
Puede hacerse de dos maneras diferentes: introduciendo la dirección del sitio web y dejando que el robot lo indexe o mediante feeds o Sitemaps Móvil. En ambos casos Yahoo! ofrece un formulario para el alta de sitios móviles⁸.

14.3 robots.txt

Uno de los problemas más importantes a los que te tendrías que enfrentar al tener un sitio móvil es la de generar contenidos duplicados, ya que si no lo impides te encontrarás con que los robots de búsqueda indexarán, por defecto, el sitio en todos sus posibles sitios de resultado. Para que esto no ocurra, has de indicarle al fichero robots.txt unas directivas que sólo permitan la indexación por parte de los robots específicos para móvil.

```

1.  User-Agent: *
2.  Disallow: /
3.  User-Agent: Googlebot-Mobile
4.  User-Agent: YahooSeeker/M1A1-R2D2
5.  Allow:
    
```

Por el momento los motores de búsqueda que tienen agentes son:

- ▣ **Google:**
Googlebot-Mobile
- ▣ **Yahoo!:**
YahooSeeker/M1A1-R2D2

(6) <http://ready.mobi/>

(7) <http://www.google.com/support/webmasters/bin/topic.py?topic=8493>

(8) <http://search.yahoo.com/free/mobile/request>

SEO

para PPC (Pago por Clic)

CAPÍTULO 15

Aunque en España existen otros servicios, uno de los más automatizados para hacer campañas PPC es Google Adwords¹, que permite realizar pujas por palabras en los resultados de búsqueda de Google.

Este servicio utiliza algunos factores para reducir el precio de los clics además de medir la relevancia de los anuncios, muy relacionados con el SEO de las páginas de destino de estos anuncios.

Es por esto que hay que plantearse, para reducir y optimizar las campañas, una optimización de estas páginas de destino.

Como primer paso y consejo, cuando crees una campaña nueva no lo hagas enviando el tráfico de la misma a la página principal o a alguna de las páginas ya existentes de tu sitio web, sino que es muy recomendable crear una página en HTML y estática para la misma. Esto significa crear una página en la que haya contenido relacionado directamente con el producto que estés ofreciendo.

¿Qué debe tener esta página en especial para que se cumplan las expectativas de la campaña?

▣ Título

El título de la página ha de ser el mismo que el título del anuncio. Con esto harás saber al sistema que estás anunciando lo mismo que hay en el contenido de la página, algo que hará aumentar su calidad.

▣ Dirección

Que la dirección de la página contenga las palabras claves del título del anuncio. De esta forma si el anuncio es *Venta de pisos en Madrid* el fichero podría llamarse **venta-pisos-madrid.html**.

▣ Velocidad

La página ha de cargar muy rápido. De ahí el crearla en un sencillo HTML ya que su tiempo será mayor, y hoy en día uno de los factores que se tienen en cuenta es el de la velocidad.

▣ Contenido

Dentro del contenido de la página deberían estar las diferentes palabras clave por las que se puja en la campaña. Esto no significa que haya que ponerlas una tras otra, sino que en el contenido y descripción de la misma estén bien integradas, incluso si es posible, usando negritas y cursivas.

(1) <http://www.adwords.com/>

Hay que tener en cuenta que para que una campaña funcione has de estudiar muy bien si funciona o no, y en estos casos es muy útil usar una herramienta de seguimiento como puede ser Google Analytics², que se integra perfectamente con tus campañas y podrás saber si tu inversión en publicidad está generándote beneficios.

(2) <http://www.google.es/analytics/>

SEO

para PDF (Portable Document Format)

CAPÍTULO 16

Los documentos PDF¹ de Adobe también aparecen entre los resultados de búsqueda, y aunque no son la primera opción al mostrar un tema, sí que pueden ayudar a que el conjunto del sitio aparezca en mejores posiciones.

Hay que recordar que este tipo de documentos son un estándar internacional y que se pueden generar de forma rápida y sencilla desde aplicaciones como OpenOffice², de forma que puedes convertir cualquier tipo de documento ofimático a este formato.

En el fondo, el tratamiento de estos documentos ha de ser similar a la de cualquier sitio web, aunque en este caso se encuentre encapsulado en un fichero. Aún así te dejo con una lista de sencillas sugerencias:

▣ Contenido textual

Hay que tener en cuenta que el principal contenido que se espera en un documento PDF es el textual, de forma que si has de generarlo intenta que sea así, y no sea una captura en imagen. En caso de tener un texto en una imagen, lo mejor es aplicarle un OCR que reconozca el texto y convertirlo a tal. Al igual que ya he comentado anteriormente, un buen uso de negritas, cursivas y subrayados ayudará a optimizar esas palabras o frases.

▣ Títulos y secciones

De la misma forma que un sitio web tiene títulos y secciones, en el caso de este tipo de documentos también ha de hacer uso de ellos. Además, uno de los detalles interesantes es que pueden enlazarse e ir de unos a otros con un simple clic, algo que ayudará al sistema de enlaces internos del propio documento.

▣ Documentos grandes

En caso de que el documento tenga capítulos muy grandes es recomendable plantearse dividirlo en otros más pequeños, ya que la indexación de tanta información no será posible.

▣ Metadatos

Una de las posibilidades de los documentos PDF es su capacidad de tener metadatos como la fecha de creación, palabras clave, empresa y autor... el disponer de una mayor cantidad de metadatos ayudará a que el documento tenga más información útil para el usuario y los robots.

(1) <http://www.adobe.com/es/products/acrobat/>

(2) <http://es.openoffice.org/>

▣ Enlaces

Al igual que se pueden poner enlaces entre una sección y otra, los documentos PDF también permiten imágenes y enlaces, lo que significa que estas direcciones también serán rastreadas por los robots y servirán para sumar puntos como enlaces entrantes a tu sitio web.

▣ Contenidos duplicados

No has de olvidar que, aunque un contenido sea HTML y el otro PDF, si son el mismo contenido se puede considerar como duplicado, por lo que es mejor en este caso decidirse por uno de los dos y el otro intentar quitarlo del rastreo de los robots de búsqueda.



ómo funciona un buscador

CAPÍTULO

17

Para saber posicionar en los buscadores hay que saber cómo funciona y los tipos que hay. Principalmente hay 3 tipos de buscadores:

- ▶ **Recuperación de información**
Los resultados te indican contenidos donde puede haber información relevante, tal y como hacen Google, Yahoo!...
- ▶ **Extracción de información**
En vez de mostrar resultados generan contenidos con los datos más importantes que encuentran. Hoy en día no existe ningún sistema funcional para Internet.
- ▶ **Búsqueda de respuesta**
El usuario realiza una pregunta y el motor es capaz de extraer la respuesta con una frase concreta. Hasta hace poco Hakia¹ utilizaba este sistema de búsqueda semántica, aunque ya lo dejó de usar. Puede verse en Powerset².

A continuación haré un resumen muy sencillo para que te hagas una idea del funcionamiento de cualquiera de los motores de búsqueda.

17.1 Creando un índice

Lo primero que hacen los buscadores es **leer toda la información de Internet**. Esto lo hacen con los robots de búsqueda que primero leen y guardan la información. De esta forma, cuando alguien realiza una búsqueda lo hace con información que puede tener cierto tiempo y puede ser antigua. Por eso en la actualidad uno de los trabajos más importantes de los grandes buscadores es la frescura de sus resultados.

Una vez tienen la información podrían realizar búsquedas directamente sobre estos contenidos cacheados. Este tipo de búsqueda secuencial sería bastante lento, por lo que en vez de utilizarse esta técnica, la información se trata y procesa, **creando la indexación**. Este sistema lee toda la información almacenada y crea unas tablas que relacionan palabras con contenidos, de forma que al realizar una consulta en el buscador, en vez de revisarse todos los contenidos indexados se van a estas tablas y extraen el listado de contenidos que lo contienen.

En este punto, cuando hacemos una consulta el buscador no revisa los contenidos

(1) <http://www.hakia.com/>

(2) <http://www.powerset.com/>

sino que revisa unas tablas (**índice**) en la que están todas las palabras clave posibles y de esta forma se ha reducido la cantidad de información a consultar, ya que no tienen repeticiones.

17.2 Creando unos resultados

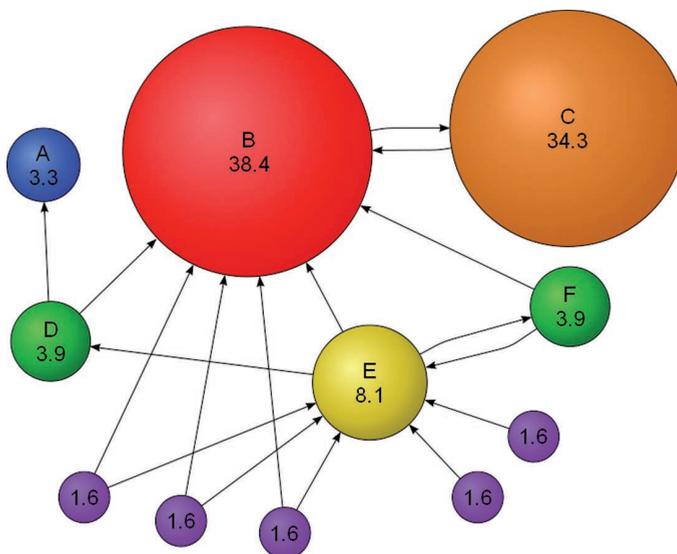
Hasta ahora tienen los contenidos, pero al hacer la consulta en el buscador ¿cómo los ordena?. Existen cientos de algoritmos de ordenación de resultados...

El más sencillo es el de “cuantas más veces aparece una palabra en el contenido, más arriba aparece”. Así funcionaban muchos de los primeros buscadores de Internet, pero no es la mejor forma, ya que también se debería tener en cuenta la cantidad de palabras que contiene el documento...

¿Qué es más importante, un documento que tiene 10.000 palabras y sólo aparece 1 vez la búsqueda o un documento en el que sólo hay 100 palabras y aparece la consulta 5 veces? Está claro que la segunda opción tiene más opciones de ser la mejor.

Al fin y al cabo, aquí lo que se está mirando es la **densidad de palabras**. Pero, tampoco sería el método más correcto, ya que podría ampliarse a hacer este mismo cálculo por párrafo o frase de un documento de texto, y luego haciendo un cálculo de la media de cada uno.

Aunque si un buscador se basase únicamente en estos factores seguiríamos estando en tecnologías de búsqueda de la primera generación, por lo que hay que introducir otros factores, como puede ser el de la cantidad de enlaces que se reciben en esa página web.



De ahí viene la frase de que *cuantos más enlaces tengamos, en mejor puesto saldremos*. Esta frase es parcialmente correcta, pero no es cierta completamente, ya que hay muchos métodos de corrección, como se está pudiendo comprobar.

El algoritmo PageRank de Google viene a introducir este sistema de mejora en la calidad de los resultados y es por eso que se le da tanta importancia a lo que hay “alrededor” de un sitio web, nunca olvidándonos de tener siempre optimizada el propio sitio.

17.3 Qué tendrán los buscadores

Intentar saber que tendrán los buscadores en un futuro es una tarea sencilla y a la vez complicada, ya que, aunque los algoritmos se vayan mejorando van apareciendo nuevos tipos de contenido que han de tratarse de formas bien diferentes.

Algunos de los elementos que iremos viendo poco a poco en todos los motores de búsqueda son:

- ▶ **Personalización**
de forma que los resultados se adapten más a tus necesidades.
- ▶ **Socialización**
de manera que los propios usuarios sean capaces de mejorar los resultados añadiéndoles información.
- ▶ **Semántica**
tanto a los contenidos como a las propias búsquedas, un trabajo de los creadores de contenidos y de los buscadores.



H

erramientas

CAPÍTULO

18

18.1 Los buscadores

Los principales buscadores de Internet tienen herramientas que permiten la gestión de los sitios una vez comprueban la legitimidad de la solicitud. Desde ellas podrás comprobar los enlaces, la indexación y algunos otros elementos.

18.1.1 Google Webmaster Tools

<http://www.google.com/webmasters/tools/>

Indexación Consultas de búsqueda más habituales >		
Rastreo de la página principal:	✓	Googlebot accedió correctamente a su página principal por última vez el 03-may-2008 .
Estado del índice:	✓	Algunas páginas de su sitio están incluidas en el índice de Google. Consulte las estadísticas del índice . ⓘ
	✓	Algunas páginas de sus Sitemaps están incluidas en el índice de Google. Consulte Visión general de Sitemaps .
Errores de rastreo web		
Errores de HTTP	✓ 0	--
No se encuentra	⚠ 27	Detalles >
No se han seguido estas URL	✓ 0	--
Todos los errores de direcciones URL en Sitemaps	⚠ 2	Detalles >
URL caducadas	✓ 0	--
URL inaccesibles	⚠ 2	Detalles >
URL restringidas por robots.txt	⚠ 5	Detalles >
Total:	36	

Ofrece una serie de herramientas generales y otras particulares para cada una de las direcciones que añadidas. Entre las generales está la posibilidad de denunciar sitios fraudulentos, notificar enlaces de pago o solicitar reconsideraciones con respecto a un sitio.

En las particulares nos encontramos con listado de errores a la hora de indexar, análisis de los contenidos (títulos, metatags, contenidos no indexables...), consultas más habituales (filtradas por país o tipo de buscador), los enlaces externos e internos, gestión de los “sitelinks”, mantenimiento de los Sitemaps y algunas herramientas como la de gestión de los robots.txt, frecuencia de rastreo, orientación geográfica, dominio preferido o búsqueda mejorada de imágenes.

18.1.2 Live Search Webmaster Tools

<http://webmaster.live.com/>

The screenshot shows the 'Summary' tab of the Live Search Webmaster Tools interface. It displays the following information:

- Website status:** Website address: javiercasares.com, Indexed pages: 3.970, Last crawled: 01/05/2008, Blocked: No, Domain rank: [Progress bar]
- Top 5 pages:** A table listing the top 5 pages with columns for Web page, Rank, Language, Country/region, Last crawled, and Blocked.

Web page	Rank	Language	Country/region	Last crawled	Blocked
Javier Casares	[Progress bar]	es	es	01/05/2008	No
posicionamiento	[Progress bar]	es	es	22/04/2008	No
Citilab	[Progress bar]	es	es	01/05/2008	No
Amsterdam - Javier Casares	[Progress bar]	es	es	23/04/2008	No
manuales	[Progress bar]	es	es	30/04/2008	No

Aunque la cantidad de herramientas que ofrece no es muy amplia, puede dar algunas ideas sobre la cantidad de páginas indexadas, si hay alguna de ellas penalizada, permite realizar consultas para encontrar información sobre contenidos, muestra el listado de las 10 mejores páginas a las que enlaza y que enlaza, además de la gestión del Sitemap.

Además tiene una herramienta general para analizar el fichero robots.txt.

18.1.3 Yahoo! Site Explorer

<http://siteexplorer.search.yahoo.com/>

The screenshot shows the Yahoo! Site Explorer interface. The search bar contains 'http://javiercasares.com'. Below the search bar, it displays the following information:

- Results:** Pages (1,630) | Inlinks (7,061) | No Subdomains. Results 1 - 10 of about 1,630 - 0 sec.
- Show pages from:** All subdomains | Only this domain
- Result Details:** [Icons]
- Top 3 results:**
 - flv**: Language: es, Last crawled: 04/26/2008 03:10:50 UTC, <http://javiercasares.com/wp-content/joanplanas.flv> - 513k
 - DivX**: Language: es, Last crawled: 04/29/2008 13:50:56 UTC, <http://javiercasares.com/wp-content/joanplanas.divx> - 513k
 - Javier Casares**: Language: es, Last crawled: 05/03/2008 03:57:32 UTC, text/html <http://javiercasares.com/> - 33k - [Cached](#)

De todas las herramientas la más sencilla en utilidades ya que únicamente permite la gestión de Sitemaps y la gestión de direcciones dinámicas, para reducir elementos duplicados como pueden ser identificadores de sesión o cambiar parámetros a otros por defecto. Además permitirá ver los enlaces internos y externos del sitio.



18.2 Palabras Clave

- ▶ **Google Keyword Tool**
<https://adwords.google.com/select/KeywordToolExternal>
 Es la herramienta que utiliza Google para los listados de Adwords pero que puede ser útil para saber lo más buscado y búsquedas relacionadas dada una dirección web o una frase de búsqueda.
- ▶ **Overture Keyword Selector Tool**
<http://inventory.overture.com/>
 Es la herramienta de selección de palabras de Overture (Yahoo! Search Marketing). Da una cifra aproximada de las búsquedas mensuales de una frase de búsqueda.
- ▶ **MIVA Generador de palabras clave**
<https://account.es.miva.com/advertiser/Account/Popups/keywordgenbox.asp>
 Es la herramienta de selección de palabras de MIVA.
- ▶ **Google Trends**
<http://www.google.com/trends>
 Permite comparar varias frases de búsqueda entre sí y ver su tendencia histórica.
- ▶ **Google Sets**
<http://labs.google.com/sets>
 Genera listados de palabras relacionadas dadas unas palabras.

18.3 Ranking en Internet

- ▶ **Alexa**
http://www.alexa.com/data/details/traffic_details/dominio.ext
- ▶ **Compete**
<http://siteanalytics.compete.com/dominio.ext/>
- ▶ **Quantcast**
<http://www.quantcast.com/dominio.ext>
- ▶ **Ranking**
<http://scripts.ranking.com/data/details.aspx?theurl=dominio.ext>

18.4 Seguimiento del usuario

- ▶ **ClickHeat**
<http://www.labsmedia.com/clickheat/>
 Permite generar mapas de calor de tu sitio web, de forma que puedas saber en todo momento en qué lugares de la página pulsa el usuario. (Descarga Gratuita)
- ▶ **ClickDensity**
<http://www.clickdensity.com/>
 Igual que el anterior. (Versión Gratuita Limitada)
- ▶ **CrazyEgg**
<http://crazyegg.com/>
 Igual que el anterior. (Versión Gratuita Limitada)

▶ **PHP User Track**

<http://ctxt.media.mit.edu/php5usertrack/>

Permite hacer un seguimiento de los movimientos del ratón por el sitio web y ver el mapa de calor generado. (Descarga Gratuita)

▶ **ClickTale**

<http://www.clicktale.com/>

Igual que el anterior. (Versión Gratuita Limitada)

▶ **FormAlive**

<http://www.formalive.com/>

Permite estudiar los campos de un formulario y saber si un usuario ha rellenado un campo o lo ha dejado, pudiendo detectar el punto de salida.

18.5 Notas de prensa

▶ **Comunicae**

<http://www.comunicae.com/>

▶ **Notas Online**

<http://www.notas-online.com/>

▶ **Notas de Prensa**

<http://www.notasdeprensa.com.es/>

▶ **Comunicados de Prensa**

<http://www.comunicadosdeprensa.com/>

▶ **Articulo.org**

<http://www.articulo.org/>

▶ **Yugler**

<http://www.yugler.com/>



H

istoria de los Buscadores

CAPÍTULO

19

19.1 De 1993 a 2008

La web se comenzó a idear en 1980 cuando **Tim Berners-Lee**¹ se plantea lo que más adelante, 10 años después, sería lo que hoy conocemos como sitios web. A finales de 1990 comenzó a aplicar sus ideas creando el primer servidor web en **NeXT**², el primer navegador web llamado **WorldWideWeb**³ (que también era editor HTML) y la primera página web⁴...



Una vez comenzó la web a tener contenido, en junio de 1993 y desde el MIT⁵ (con **Matthew Gray** a la cabeza), se desarrolló **World Wide Web Wanderer**, un robot de búsqueda creado en Perl⁶ que pretendía medir el tamaño de la red. Ese robot se amplió pudiendo leer direcciones URL creándose así **Wandex**, el que se podría considerar el primer buscador de internet, y que tuvo grandes problemas de infraestructura y velocidad cuando alcanzó los cientos de visitas diarias, ya que su robots conseguía “tirar” los sitios que indexaba.

El siguiente buscador (quizá mejor llamarlo directorio) fue **Aliweb**⁷ (*Archie Like Indexing on the Web*), también apareció en octubre de 1993 y todavía está en marcha. Creado por **Martijn Koster**⁸, lo que hacía era indexar los metatags de las páginas que se le daban a su índice, es decir, que no tenía un robot de búsqueda que consumiera gran cantidad de ancho de banda como Wandex.

Tras estos primeros procesos de rastreo en la red, **Martijn Koster** propuso unas sugerencias⁹ para lo que sería el fichero **robots.txt**¹⁰ que limita la acción de los robots de búsqueda en los sitios web. En ese momento se comenzaba a poder de-

(1) http://es.wikipedia.org/wiki/Tim_Berners-Lee

(2) <http://es.wikipedia.org/wiki/NeXT>

(3) <http://en.wikipedia.org/wiki/WorldWideWeb>

(4) <http://www.w3.org/History/19921103-hypertext/hypertext/WWW/TheProject.html>

(5) <http://www.mit.edu/>

(6) <http://www.perl.com/>

(7) <http://www.aliweb.com/>

(8) <http://www.greenhills.co.uk/mak/mak.html>

(9) <http://www.robotstxt.org/wc/guidelines.html>

(10) <http://www.robotstxt.org/>

tectar cada uno de los robots mediante su agente, se sugería usar siempre una DNS única, filtrar determinados lugares del sitio para que no fuera accesible...

Aquí comenzaron a desarrollarse los primeros robots (arañas) como **Jumpstation** que indexaba el título, URL y cabecera del sitio, al igual que **World Wide Web Worm**, creado por Oliver Mc.Bryan en 1994 (y comprado en 1998 por Goto.com¹¹) que funcionaba de la misma manera. Aunque era interesante que indexaran, el problema de estos motores era la forma de mostrar resultados, ya que lo hacían sin aplicar ningún algoritmo, simplemente mostrando los resultados según la fecha de indexación. Más adelante, en diciembre, también lo hizo así el **RBSE** (*Repository-Based Software Engineering*) comenzando a aplicar un primer ranking en base a la relevancia de la palabra dada.

De forma paralela iban apareciendo algunos directorios como EInet Galaxy, que en enero de 1994 podría considerarse el primer directorio tal y como los conocemos en la actualidad. Pero, fue en abril de 1994 cuando **David Filo**¹² y **Jerry Yang**¹³ crearon **Yahoo!**¹⁴ (anteriormente conocido como *Jerry's Guide to the World Wide Web*), una colección de las páginas web favoritas. El gran problema de Yahoo! era que comenzó siendo un directorio hecho por personas y eso llevaba mucho tiempo, por lo que tuvo que evolucionar incorporando un buscador para ese directorio. Es curiosa la historia de su nombre debido a que muchos le otorgan el nombre de *Yet Another Hierarchical Officious Oracle*, aunque sus fundadores insisten en que se basa en los personajes de unas *aventuras de Gulliver*.

Hasta aquí podríamos hablar de la *versión "beta" de los buscadores*, la versión que no era aún ni la primera y en la que todo eran experimentos.



Así hasta que el 20 de abril de 1994 **Brian Pinkerton**¹⁵, desde la Universidad de Washington¹⁶, presentase **WebCrawler**¹⁷. En realidad este buscador era de escritorio y nació el 27 de enero de 1994 pero en 3 meses se convirtió

en un robot de la red. La gran diferencia y paso diferencial que podríamos llamar la *versión 1 de los buscadores* era que indexaba las páginas de forma completa y que buscaba información en ellas, al contrario de sus antecesores, que sólo buscaban en la dirección web, título o metatags. Esto hizo que la relevancia de los resultados fuera mucho mayor. Además, tenía la peculiaridad de poder ver las búsquedas en tiempo real con su *Webcrawler Search Voyeur*. Hay que tener en cuenta que al cabo de 7 meses llegó a su millón de consultas, que en 1995 fue comprado por AOL¹⁸, en 1997 por Excite¹⁹ y que en 2001 pasó a ser parte de InfoSpace²⁰.

InfoSeek²¹ también apareció a principios de 1994, y aunque no llegó a tener nada especial, tuvo un gran salto en diciembre de 1995 cuando fue el motor de búsqueda por defecto en Netscape²².

(11) <http://www.goto.com/>

(12) <http://docs.yahoo.com/docs/pr/executives/filo.html>

(13) <http://docs.yahoo.com/docs/pr/executives/yang.html>

(14) <http://www.yahoo.com/>

(15) <http://www.thinkpink.com/bp/>

(16) <http://www.washington.edu/>

(17) <http://www.webcrawler.com/>

(18) <http://www.aol.com/>

(19) <http://www.excite.com/>

(20) <http://www.infospace.com/>

(21) <http://www.infoseek.com/>

(22) <http://www.netscape.com/>



Tras la aparición de Webcrawler hay que destacar la aparición de Lycos²³ el 20 de julio de 1994, creado por Michale Mauldin en la Universidad de Carnegie Mellon²⁴, con un algoritmo interesante que incluía el concepto de proximidad entre palabras. Eso sí, no indexaba de forma completa las páginas, sólo las 20 primeras frases, las 200 primeras de la cabecera y un grupo de las 100 más relevantes de todo el documento. Aunque si hay que destacar a

LYCOS

meet you there



Lycos es por la cantidad de documentos indexados: se lanzó con 54.000 documentos; en agosto de 1994 tenía 394,000 documentos; en enero de 2005, 1,5 millones; y en noviembre de 1996 llegó a los 60 millones, convirtiéndolo en el motor de búsqueda más destacado. La palabra que ofrecía más resultados era [surf].

En esta época comenzaron también a aparecer los primeros metabuscadores. Este sistema lo que hace realmente es unificar los resultados de varios motores de búsqueda para ofrecer los resultados mezclados. En 1995 apareció el primero de ellos llamado MetaCrawler²⁵ creado por Erik Selberg y Oren Etzioni en la Universidad de Washington (como Webcrawler). En este caso devolvía resultados de Lycos, Altavista, Yahoo!, Excite, Webcrawler e Infoseek. El problema era su velocidad. Como curiosidad, en noviembre de 1996 ya tenía una nueva versión, en beta, para probar.



En diciembre de 1995 seis estudiantes de Stanford²⁶ lanzaron Excite²⁷ gracias al proyecto Architext (iniciado en 1994) que introdujo uno de los conceptos base de las búsquedas. El complicado algoritmo intentaba crear un sistema parecido a los sinónimos mediante estadísticas entre las relaciones de palabras, de forma que se podía

realizar una búsqueda obteniendo resultados aunque la misma no existiera en la página (si tenía alguna relación, claro). En 1996 compró Magellan y Webcrawler y creó su propio directorio. Fue el buscador de referencia hasta mediados de 1999 y en 2002 perdió su sistema de búsqueda pasando a ser un metabuscador.

El siguiente gran lanzamiento fue AltaVista²⁸. Fue en Diciembre de 1995 cuando hizo su aparición en escena y es muy importante este lanzamiento por las mejoras que proponía: tenía ancho de banda casi ilimitado, permitía consultas en lenguaje natural²⁹ (las que utilizamos habitualmente para hacer búsquedas), consultas avanzadas mediante operadores lógicos (AND, OR...), añadir o eliminar direcciones web en 24 horas, comprobar los enlaces entrantes a un sitio web e incluso permitía hacer búsquedas en los nombres de imágenes y algunos ficheros multimedia. No sólo era grande en resultados sino veloz al entregarlos. Además, ofrecía una serie de “ayudas / trucos” para mejorar la calidad de las consultas.

A finales del 95 apareció Ozú³⁰ como directorio y buscador de la mano de Advernet³¹ (todavía funciona su primera dirección³²). La Base de Datos del buscador y del directorio estaba íntegramente gestionada y actualizada por personas (con ayuda de herramientas de rastreo y

(23) <http://www.lycos.com/>

(24) <http://www.cmu.edu/>

(25) <http://www.metacrawler.com/>

(26) <http://www.stanford.edu/>

(27) <http://www.excite.com/>

(28) <http://www.altavista.com/>

(29) http://es.wikipedia.org/wiki/Lenguaje_natural

(30) <http://www.ozu.es>

(31) <http://www.advernet.es/>

(32) <http://www.advernet.es/ozu/>

gestión de calidad).

Durante 1997 y hasta el 2000 la empresa tuvo un largo litigio contra un empleado que se robó el dominio [ozu.com](http://www.ozu.com)³³ y que estuvo explotando el buscador paralelamente hasta que salió el juicio a favor. Mientras se usaba [ozu.es](http://www.ozu.es) que es el sitio web que ha perdurado.

A partir del 2001 las cosas evolucionaron ya que Ozú fue comprado por el Grupo Vocento³⁴ y se llega a un acuerdo con Google para usar sus resultados, priorizando los resultados de Ozú, hasta que finalmente se dejó de usar el contenido original y, en 2006 comenzó a utilizar Noxtrum³⁵, ya desaparecido, y su actual vuelta a Google.

También a finales de 1995 apareció un nuevo e importante directorio: **LookSmart**³⁶. Creado por un matrimonio australiano, Evan Thornley y Tracey Ellery, y tras muchos problemas financieros, en 1997 se trasladaron a San Francisco y en 1998 llegaron a un acuerdo con Microsoft³⁷ para ser su proveedor.

Pocos meses después, el 20 de mayo de 1996 Paul Gauthier y Eric Brewer, desde la Universidad de Berkeley³⁸, lanzaban **Hotbot**³⁹, que con su motor **Inktomi**⁴⁰, llegaron a un acuerdo con el sitio web de **Wired**⁴¹ que fue el que le ayudó a darse fama. Se consideró el primer motor de búsqueda capaz de indexar los millones de sitios web que había en ese momento. Tras la burbuja del 2001, perdió muchos de sus usuarios y en 2002 fue comprado por Yahoo!.

El 1 de Junio de 1996 se lanzaba el primer directorio español universitario, **Dónde? Directorio Online de España**⁴² que incorporaba 1.157 resultados en su base de datos, según explica su cómo se hizo⁴³. A principios de 1997 estaba en 5.000 registros y 3,5 millones de accesos al mes. Después el número de consultas creció más lentamente. El tope se alcanzó en junio de 1998, dos años después de inaugurar el servicio, con 29.000 registros y casi 5 millones de consultas en un mes: una media de más de 100 consultas por minuto, momento en el que se produjo un estancamiento. El número de registros siguió aumentando, pero las consultas no. La razón era el colapso del servidor: no daba más de sí, no admitía más usuarios. Con 44.563 registros en la base de datos, dónde anunció que cerraba sus puertas el 1 de junio de 1999⁴⁴.



El siguiente de la lista es **Ask Jeeves**⁴⁵, lanzado en 1996 también; su idea era la de poder contestar preguntas de forma natural, tal y como las hacemos habitualmente. Como peculiar, su mayordomo, basado en Jeeves⁴⁶ de P.G. Wodehouse⁴⁷. En 1999 compró una empresa llamada **Direct Hit** (creado por Gary Culli) y aplicó su tecnología en su motor de búsqueda. Apareció el verano de 1998 y mucha gente ya lo utilizaba a finales del mismo debido a la alta relevancia que ofrecía, principalmente porque basaba sus resultados de búsqueda en los clicks que hacían los usuarios (técnica que utilizan algunos en la actualidad).

(33) <http://www.ozu.com>

(34) <http://www.vocento.es>

(35) <http://www.noxtrum.com/>

(36) <http://www.looksmart.com/>

(37) <http://www.microsoft.es/>

(38) <http://www.berkeley.edu/>

(39) <http://www.hotbot.com/>

(40) <http://www.inktomi.com/>

(41) <http://www.wired.com/>

(42) <http://dónde.uji.es/>

(43) <http://dónde.uji.es/themakingof.html>

(44) http://dónde.uji.es/adios_amigos.html

(45) <http://www.ask.com/>

(46) <http://en.wikipedia.org/wiki/Jeeves>

(47) <http://www.pgwohousebooks.com/>



El proyecto **Google**⁴⁸ comenzó a desarrollarse en enero de 1996 por **Sergey Brin**⁴⁹ y **Larry Page**⁵⁰ en la Universidad de Stanford, llamándose **BackRub**⁵¹ debido a la tecnología que utilizaba, que calculaba la importancia de un sitio web en base a los enlaces que recibía. En esa época fue cuando Page fabricó un ordenador con piezas de Lego⁵² y con tecnología antigua que más adelante se convertiría en lo que hoy es Google. El 15 de septiembre de 1997 el dominio *google.com* era comprado y el 7 de septiembre de 1998 se creaba *Google*⁵³ *Inc.* Una peculiaridad de Google es que en momentos especiales cambia su logo⁵⁴ para adaptarlo a esa ocasión.

Hay que destacar dos razones por las que Google se hizo muy interesante: una interfaz muy clara y sencilla (como la de Altavista en sus inicios) y unos resultados muy relevantes. El secreto de los resultados, la tecnología **PageRank**⁵⁵ (patentada⁵⁶ el 4 de septiembre de 2001) hizo que el mundo de los motores de búsqueda cambiase completamente dando por iniciado lo que se puede llamar la *versión 2 de los buscadores*. Page lo que implementó fue un sistema mediante el cuál no sólo se tenían en cuenta los factores de la propia página en la que se buscaba información, sino que se tenían en cuenta otros factores externos que daban mayor o menos importancia al sitio web. Hoy en día todos los motores de búsqueda utilizan una tecnología similar (o que, al menos, tiene en cuenta estos factores).

Uno de los primeros motores de búsqueda con enfoque profesional fue **Norther Light**, creado en agosto de 1997 disponía ya de *resultados en clustering* de forma que daba sugerencias de búsqueda muy interesantes, pero no llegó a ser uno de los más utilizados por el público general. En 2002 dejó de dar servicio y hoy en día vende su tecnología.

Otro de los motores más importantes del mundo es **Yandex**⁵⁷, lanzado el 23 de septiembre 1997, el principal buscador ruso que lleva su nombre desde “Yet Another iNDEXer”, además de ser muy similar al primer buscador: Wandex. Además, la R invertida se pronuncia l en inglés, por lo que su nombre se pronuncia “Index”. En el año 2000 se fundó la compañía como tal. Uno de los elementos más interesantes de este motor es su especialización en la gramática rusa, lo que lo hace único en el mundo.

En 1998 apareció **MSN Search**⁵⁸, de la mano de **Microsoft**, utilizando los datos de Inktomi y también apareció el **Open Directory Project**⁵⁹ (**DMOZ**), que, aunque no era el primer directorio hecho por personas, sí que era el primero en hacerlo de forma colaborativa. Creado por Rich Skrenta y Bob Truel y llamado inicialmente **Gnuhoo**, pasó a llamarse **Newhoo** el 5 de junio de 1998 y finalmente fue adquirido por **Netscape** en octubre de 1998 cuando pasó a ser el ODP, momento en el cual ya disponía de 100.000 direcciones y cerca de 4.500 editores.



A mediados de 1999 apareció en el mercado **AllTheWeb**⁶⁰. Utilizaba la tecnología de **Fast**⁶¹,

(48) <http://google.stanford.edu/>

(49) http://es.wikipedia.org/wiki/Sergey_Brin

(50) http://es.wikipedia.org/wiki/Larry_Page

(51) <http://backrub.stanford.edu/>

(52) <http://www-db.stanford.edu/pub/voy/museum/pictures/display/0-4-Google.htm>

(53) <http://www.google.com/>

(54) <http://www.google.com/holidaylogos.html>

(55) <http://www-db.stanford.edu/~backrub/google.html>

(56) <http://patft.uspto.gov/netacgi/nph-Parser?patentnumber=6,285,999>

(57) <http://www.yandex.com/>

(58) <http://search.msn.com/>

(59) <http://directory.mozilla.org/>

(60) <http://www.alltheweb.com/>

(61) <http://www.fastsearch.com/>

una empresa noruega que venía de la Norwegian University of Science and Technology⁶². Este buscador ofrecía algunas mejoras con respecto a Google como por ejemplo una base de datos más actualizada, una búsqueda avanzada mucho más amplia búsquedas clusterizadas, pero nunca llegó a tener tanto éxito. En febrero de 2003 fue comprado por **Overture**⁶³ y ésta, a su vez, en marzo de 2004 por Yahoo! que redujo algunas de sus funcionalidades. Hay que tener en cuenta que su base de datos pasó de 80 millones a finales de 1999 a 200 millones a principios de 2000 llegando a los 2.000 millones en junio de 2002 quedando en 3.300 millones cuando fue adquirida por su actual propietario. Ahora la empresa Fast ha sido comprada por Microsoft.

En 1999 también aparecía un gigante de la red: **Baidu**⁶⁴. El motor de búsqueda chino sería un punto de referencia hasta la actualidad debido a la presión que mantiene el gobierno chino sobre Internet. Es curioso que la mayor parte de la inversión que tiene viene dada de empresas estadounidenses.

En el año 2000 se lanzó el motor de búsqueda **Teoma**⁶⁵ de mano de **Apostolos Gerasoulis** en la Universidad de Rutgers⁶⁶. Utilizaba un sistema de clustering para organizar los sitios en base al **Subject-Specific Popularity** (actualmente **Expert Rank**) que, al contrario del Pagerank de Google, analizaba los enlaces en un contexto en el que se daba un ranking a una página web según el tema tratado. El 11 de septiembre de 2001 fue comprado por **Ask Jeeves**.

El 15 noviembre de 2003, Google implementó uno de los primeros grandes cambios⁶⁷ en su motor de búsqueda añadiendo mejoras en la búsqueda semántica. Los motores de búsqueda tenían unas necesidades básicas: *clustering* y *semántica*.

En diciembre de 2003 se creó la empresa de **Seekport**⁶⁸, funcional a partir de enero de 2004. La empresa comenzó asociándose con **Arexera**, que tenía tratos con la versión alemana de Infoseek. La peculiaridad de Seekport es que dispone de una versión local para cada país con un índice independiente preparado para cada una de las necesidades locales.

En 2004, cuando MSN Search dejó de recibir los datos de Looksmart pasó a utilizar los resultados del motor de Inktomi. De esta misma forma, Yahoo! también dejó los datos de Google para unificar los motores de Alltheweb, Inktomi y Altavista. En marzo de 2004 esos buscadores comenzaron a utilizar la base de datos de Yahoo!.

En esta época también comenzaron a aparecer nuevos proyectos muy interesantes. Uno de ellos es **Nutch**⁶⁹, un motor de búsqueda en código abierto y creado en Java⁷⁰. Aunque su desarrollo es costoso, ha conseguido el apoyo de Yahoo!.

El 30 de septiembre de 2004 se lanzó **Clusty**⁷¹ de la mano de **Vivisimo**⁷². Entre sus peculiaridades hemos de destacar su filosofía completamente basada en el clustering, lo que da pie a que el idioma inglés tenga resultados razonables, pero falle en muchos otros idiomas. Además fue el primer gran buscador que ofrecía *búsquedas en Blogs* o la *Wikipedia* en una de sus opciones.



En noviembre de 2004, MSN Search y de la mano de Christopher Payne y Oshoma Momoh pu-

(62) <http://www.ntnu.no/>

(63) <http://www.overture.com/>

(64) <http://www.baidu.com/>

(65) <http://www.teoma.com/>

(66) <http://www.rutgers.edu/>

(67) <http://www.webworkshop.net/florida-update.html>

(68) <http://www.seekport.es/>

(69) <http://www.nutch.org/>

(70) <http://java.sun.com/>

(71) <http://www.clusty.com/>

(72) <http://vivisimo.com/>



sieron en marcha una primera fase pública del motor de Microsoft, que se hizo pública el 20 de enero de 2005. El 1 de noviembre de 2005 se presentaba la plataforma **Windows Live**⁷³ que sería la nueva interfaz del motor de búsqueda.

También hay que hacer referencia a buscadores como **Noxtrum**⁷⁴ del que se comenzó a saber a finales de Junio de 2005, haciendo referencia a estar enfocado principalmente en sitios en español y portugués, además de contener toda la información de **Páginas Amarillas**⁷⁵. El 30 de Noviembre de 2005 se lanzaba la versión beta, y el 22 de Abril de 2006 era presentada la primera versión final del buscador. El 1 de Abril de 2008 dejaba de estar en línea este proyecto.

Y, por ahora la última presentación, la de **Quaero**, el buscador europeo que se presentaba los primeros días de 2006, impulsado principalmente por los gobiernos de Francia y Alemania (que ya se ha retirado del mismo lanzando el proyecto **Theseus**) y potenciado por grandes empresas tecnológicas europeas. Finalmente a principios de Abril, se lanza como **Exalead**⁷⁶ (buscador que llevaba años en funcionamiento con los resultados de AOL, y que tras el acuerdo de éste último con Google, decidió usar el índice de Quaero). Desde Septiembre de 2006 su índice cuenta con 8.000 millones de resultados.

En los últimos meses han aparecido dos motores centrados en las búsquedas semánticas. El primero es **Hakia**⁷⁷, que, aunque lleva desarrollándose desde 2004, se ha lanzado en beta en 2006 y se está dando a conocer desde 2007. Su planteamiento es el lanzamiento oficial a lo largo de 2008. El segundo es **Powerset**⁷⁸, lanzado en beta privada el 17 de septiembre de 2007 bajo el nombre de Powerlabs, comienza a hacer sus primeras pruebas con la Wikipedia y utiliza un algoritmo desarrollado por Xerox PARC⁷⁹.

(73) <http://www.live.com/>

(74) <http://www.noxtrum.com/>

(75) <http://www.paginasamarillas.es/>

(76) <http://www.exalead.es/>

(77) <http://hakia.com/>

(78) <http://www.powerset.com/>

(79) <http://www.parc.com/>

19.2 Tabla Resumen de la Aparición de los Buscadores

1993	Wandex Aliweb robots.txt
1994	Elnet Galaxy JumpStation Yahoo! (directorio) WebCrawler WWWorm InfoSeek Lycos RBSE
1995	MetaCrawler Excite LookSmart Altavista Ozú
1996	Hotbot / Inktomi Dónde? Ask Jeeves Backrub / Google
1997	Norther Light
1998	MSN Search ODP / DMOZ
1999	AllTheWeb Baidu
2000	Teoma
2003	Seekport
2004	Yahoo! (buscador) Nutch Clusty MSN Search (buscador)
2005	Windows Live Noxtrum
2006	Exalead (motor Quaero) Hakia
2007	Powerset





iccionario

CAPÍTULO

20

A

Adcenter

Herramienta de Microsoft para gestionar la inversión de los anuncios en su buscador y red de sitios.

Adsense

Herramienta de Google para gestionar los anuncios que se insertan en sitios web y otros soportes y que dan beneficios a los webmasters.

Adwords

Herramienta de Google para gestionar la inversión de los anuncios en su buscador y red de sitios.

Algoritmo

Es la fórmula mediante la cual los buscadores organizan sus resultados de búsqueda.

Anchor Text

Texto visible que hay en un enlace. Por defecto el texto que aparece en azul y subrayado en una página web y que permite visitar otra página.

Araña

Una forma de nombrar a los robots (ordenadores) de búsqueda que indexan la red.

Autoridad

Una página recibe enlaces de muchos nodos, y cuando esta página contiene contenido relevante para un tema se convierte en una “autoridad” en la materia.

B

Backlink

Ver *Enlaces entrantes*.

Banear

(del inglés: ban) Penalización consistente en no aparecer en los resultados de un buscador por ninguna búsqueda.

Blacklist

Las black list (listas negras) son unos listados en los que apareces por alguna razón, habitualmente siendo penalizado en un futuro por los que las consultan.

Blog

Sitio web con un formato peculiar: artículos ordenados cronológicamente inverso y que tienen direcciones permanentes (cada artículo tiene una URL concreta).

Blogroll

Lista de enlaces amigos que podemos encontrar habitualmente en el menú de navegación de un Blog.

C

Caché

En un buscadores aquella copia de la página web original que almacena el sistema para poder realizar consultas.

Cloaking

Técnica penalizada que consiste en enseñarle al buscador una página y al usuario otra con el fin de engañar y mejorar el posicionamiento.

Cluster

Forma de agrupar la información o resultados de búsqueda en base a una temática o palabra clave.

Consulta

Cada una de las palabras o frases que se utilizan para encontrar en un motor de búsqueda.

CPA

Coste por acción. En publicidad es el tipo de anuncio en el que se paga cuando el usuario realiza una acción en el sitio del anunciante, habitualmente el rellenar un formulario.

CPC

Coste por clic. En publicidad, es el precio o sistema de anuncios en el que se paga cuando el usuario pulsa en el anuncio.

CPM

Coste por mil. En publicidad, sistema de anuncios en el que se paga una cantidad por cada 1.000 impresiones.

Crawler

Una forma de nombrar a los robots (ordenadores) de búsqueda que indexan la red.

D

Densidad

En SEO, es el porcentaje de aparición de unas palabras determinadas en una página que permiten ser lo óptimo para que un buscador la posicione.



Directorio

Lugar web en el que, en algunos casos de forma manual, se incorporan otras páginas web organizadas por categorías.

DMOZ

Directorio gestionado por personas y que utilizan algunos buscadores grandes para rastrear la red.

E

Edad

Es el tiempo que lleva un sitio en línea o el tiempo que lleva un dominio registrado.

Enlaces absolutos

Son aquellos hipervínculos en los que aparece la dirección completa y no parcial, incluyendo el dominio y las páginas y carpetas de la dirección.

Enlaces entrantes

Son aquellos hipervínculos que enlazan a nuestro sitio web. Como la mayoría de factores de posicionamiento se basan en los enlaces que llegan a nuestra web, es un factor siempre a tener presente.

Enlaces profundos

Son aquellos hipervínculos que hacen referencia a páginas muy alejadas en clics de la página principal.

Enlaces recíprocos

Son aquellos hipervínculos que enlazan de un sitio A a un B, y desde el mismo B al A.

Enlaces salientes

Son aquellos hipervínculos a los que enlazamos desde nuestro sitio web.

F

Findability

Se podría traducir como “encontrabilidad” y viene a ser la facilidad para que un buscador encuentre tu sitio web...

L

Landing Page

Página de entrada o de recepción de una búsqueda o campaña.

Link Bait

Acción de conseguir gran cantidad de enlaces en poco tiempo... un bombardeo de enlaces.

Long Tail

A nivel estadístico, se utiliza para denominar al conjunto de búsquedas que si bien no suponen una gran cantidad, la suma de ellas es razonablemente importante para competir con búsquedas que atraen mucho tráfico.

M

Metabuscador

Buscador que utiliza y mezcla los resultados de motores de búsqueda que tienen su propio índice para generar una lista de resultados nueva.

N

Nodo

Son aquellos sitios que disponen de una autoridad mínima como para ser puntos de importancia de Internet. Por ejemplo DMOZ o Wikipedia.

P

Pagerank

Algoritmo utilizado por Google para organizar los resultados de búsqueda en una consulta.

Pago-Par-Inclusión (o PPI)

Sistema utilizado por algunos buscadores y directorios que permite aparecer en su índice a cambio de una cuota.

Puesto (o Posición)

En los resultados de búsqueda, el lugar en el que aparece un resultado concreto.

Q

Query

Ver Consulta.

R

Ranking

Ver Puesto.

Redirección

Sistema para enviar a un usuario de una página a otra.

robots.txt

Fichero de control de los robots de búsqueda.

ROI

Return On Investment (Retorno de la Inversión). Ganancia sobre el capital invertido. Es el factor que indica si una campaña de cualquier tipo nos es rentable o no, ya sea SEO, SEM, etc...

S

SEM

Search Engine Marketing (Marketing en Motores de Búsqueda), hace referencia, principalmente, a los sistemas de publicidad en buscadores, aunque realmente hace referencia a todo lo que tenga que ver con ampliar el negocio de los sitios web con el tráfico generado desde los buscadores.

SEO

Search Engine Optimization (Optimización en Motores de Búsqueda), se utiliza para designar las técnicas de posicionamiento natural y a las personas que se dedican a realizar estas acciones.

SERP

Search Engine Results Page (Páginas de Resultados de Buscadores), es la página de los resultados, aunque frecuentemente se hace referencia a la lista de resultados.





CAPÍTULO

21

entido Común

Aunque lo que indico en esta guía es lo que se debería aplicar de forma ideal a nivel de posicionamiento y estrategia de visibilidad en Internet, siempre hay que utilizar el sentido común al crear un sitio web. Hay que plantearse el funcionamiento de la red de redes y pensar cómo utilizarla, en el buen sentido de la expresión, a nuestro favor.

Seguramente muchas veces te encontrarás en situaciones a las que no creías poder llegar, con sitios que copian un contenido y aparecen por encima nuestra... pero siempre debes mirar la parte positiva de este tipo de situaciones.

Si bien es cierto que aparecer entre los primeros resultados de búsqueda nos hará la mejor promoción posible a nuestro sitio, y con ello todo lo que conlleva, no hay que olvidar que hay vida más allá de los buscadores, que “el SEO ha muerto” y que ahora hemos de ser más ambiciosos para pasar de ese SEO a una IVS (Internet Visibility Strategy), una estrategia que nos permita estar en boca de todos.

21.1 Convertirse y entrenar a SEOs

Por suerte o desgracia de muchos, los conocimientos sobre posicionamiento en buscadores no se pueden adquirir leyendo artículos o con lo que te ha podido decir alguien. El SEO es un trabajo que hay que aplicar y experimentar día a día si realmente quieres dedicarte a esto.

Aunque hay detalles que no cambiarán nunca, en un tono general, sí que se van haciendo ajustes en los algoritmos de los motores de búsqueda para mejorar su calidad y penalizar aquello no válido. Es por eso que el entrenamiento de un SEO es una tarea nada sencilla, pero que se puede plantear de forma segura.

Algunos consejos que te puedo dar son:

- ▣ La persona que se vaya a dedicar ha de tener conocimientos razonables a nivel técnico y mínimos de marketing, aunque se recomienda haber tratado estos temas sobretodo en Internet con cierta visión. No hay que olvidar que el SEO es una de las herramientas de marketing más potente, complementaria a muchas otras y que para poder realizar trabajos de mejora en los sitios es necesario conocimiento técnico. Al final esta persona ha de tener un espíritu bastante lanzado y emprendedor.

- ▶ Escribir bien, tener buena ortografía es indispensable para poder convertirse en un buen SEO. No hace falta ser periodista, pero sí alguien que sepa jugar con el idioma.
- ▶ “El que mucho abarca poco aprieta”. Intentar que una persona pueda acumular todo el conocimiento es bastante complejo, por lo que lo mejor es crear un equipo de varias personas (2 ó 3) en el que cada una controle un tema específico, que se desenvuelva bien y se sienta agusto.
- ▶ Hay que ser progresivo. Primero comenzar colaborando con un SEO en algún proyecto, luego gestionando otro bajo supervisión. La tarea de formación puede durar meses o años.
- ▶ Estar al día es básico... no sirve de nada saber lo que ocurría hace un par de meses, porque ese conocimiento ya se habrá desfasado.
- ▶ Paciencia. Al igual que el SEO es una estrategia a medio-largo plazo, la formación también lo ha de ser.
- ▶ Aunque, quizá lo más importante es no intentar convertirse en un SEO, sino en alguien que piense globalmente y que vea Internet como una herramienta global.

