

Posicionamiento web: conceptos y herramientas

Por Lluís Codina y Mari Carmen Marcos

Resumen: Se presentan la terminología y conceptos relacionados con el posicionamiento web, los criterios que se deben tener en cuenta para planificar una campaña de posicionamiento, su ciclo de vida y una selección de herramientas de coste cero que ayudan a analizar el posicionamiento de un sitio y dan claves para mejorarlo. Se proponen y analizan los elementos que componen una metodología de trabajo en el ciclo de vida completo de una campaña de posicionamiento. Partiendo del esquema general de análisis, diseño e implantación, se describen y articulan los componentes y fases de la campaña. Finalmente, se presentan consideraciones sobre aspectos éticos en el posicionamiento web.

Palabras clave: Posicionamiento web, Posicionamiento en buscadores, Optimización de sitios web, Palabras clave, Campaña de posicionamiento, Herramientas de posicionamiento, Motores de búsqueda, Optimización para motores de búsqueda.

Title: Web positioning: concepts and tools

Abstract: The article presents terminology and related web positioning concepts, features to be considered in planning a positioning campaign, its life cycle, and a selection of free tools that help to analyse a site's position and provide keys for improving it. Elements making up a working methodology for a positioning campaign's life cycle are proposed and analysed. Beginning with a general schema for analysis, design, and implantation, the authors describe components and phases of the campaign. Finally, ethical aspects of web positioning are discussed.

Keywords: Web positioning, Search engine positioning, Web site optimisation, Keywords, Positioning campaign, Positioning, Positioning tools, Search engines, Search engine optimisation.

Codina, Lluís; Marcos, Mari Carmen. "Posicionamiento web: conceptos y herramientas". En: El profesional de la información, 2005, marzo-abril, v. 14, n. 2, pp. 84-99.



Lluís Codina es doctor en Ciencias de la Información y profesor de la sección de Ciencias de la Documentación de la Universidad Pompeu Fabra de Barcelona. Coordina el grupo de investigación DigiDoc del Instituto Universitario de Lingüística Aplicada (IULA) de esta misma universidad. Participa como docente en programas de postgrado universitarios, como el Máster Online en Documentación Digital del que es co-director, y en doctorados de la UPF, Universitat de Barcelona y Universidad del País Vasco. Asimismo, es co-director del Anuario Hipertext.net. Es especialista en sistemas de información documental, investigación en línea y documentación digital.

Mari Carmen Marcos es doctora y licenciada en Documentación, y profesora en la Sección de Ciencias de la Documentación de la Universidad Pompeu Fabra de Barcelona. Desarrolla labores de investigación en el grupo DigiDoc del Instituto Universitario de Lingüística Aplicada (IULA) de esta misma universidad. Participa como docente en programas de postgrado universitarios, de entre los cuales destaca el Máster Online en Documentación Digital, del que además de profesora es Coordinadora Académica. Colabora como consultora en los estudios de Documentación de la Universitat Oberta de Catalunya. Es coordinadora editorial del Anuario Hipertext.net.



POSICIONAR es colocar alguna cosa en su lugar óptimo. En el ámbito de la world wide web, posicionar un sitio significa optimizarlo para que aparezca en las primeras posiciones de las páginas de resultados de los motores de búsqueda. Así mismo, podemos definir posicionamiento web como el conjunto de procedimientos y técnicas que tienen como

finalidad dotar a un sitio o a una página web de la máxima visibilidad en Internet.

Existen dos factores que han hecho necesaria la disciplina emergente del posicionamiento: en primer lugar, el uso de los motores de búsqueda como fuente de información privilegiada sobre la Web por parte de los internautas; en segundo lugar, el hecho reiterada-

mente constatado de que los usuarios de motores de búsqueda apenas consultan más allá de los primeros veinte o treinta resultados, con una abrumadora mayoría que únicamente revisa los diez primeros.

La combinación de factores anteriores (uso intensivo de los motores como vía previa para acceder a un sitio y no visualización de los sitios listados más allá de la segunda o tercera página de resultados) ha generado la legítima preocupación por parte de los responsables de sitios web porque los mismos aparezcan bien posicionados en los motores de búsqueda (y directorios, aunque en menor medida). Una prueba directa de esta preocupación es el enorme número de resultados que obtenemos por ejemplo en *Google* por el término “posicionamiento en buscadores” vinculados con empresas que ofrecen esta clase de servicios.

«Visibilidad es el término acuñado, entre otros, por los ingenieros de *Google* para referirse a una medida que indica el número de enlaces que un sitio web recibe de otros sitios y la calidad de los mismos»

En este artículo nos proponemos presentar algunos conceptos de base relacionados con el posicionamiento web. Estos conceptos pueden ser clave para la toma de decisiones en el diseño y mantenimiento de sitios web y ser útiles tanto a profesionales de la biblioteconomía y documentación como a responsables del mantenimiento y la promoción de sitios web en general. En concreto, nos proponemos presentar un conjunto de criterios que pueden optimizar un sitio de cara a su posicionamiento, esto es, de cara a su visibilidad en la Web.

Como veremos, estos criterios serán tanto intrínsecos (elementos internos del propio sitio) como relacionales (o de relación con otros sitios web). Presentaremos también en qué puede consistir el ciclo de una campaña de posicionamiento, con sus diversas etapas y algunas herramientas de software gratuitas que nos pueden ayudar a analizar el posicionamiento de un sitio y/o nos darán claves para mejorarlo. Para preparar este trabajo nos hemos basado en estudios previos de posicionamiento de sitios web vinculados con los autores, hemos analizado la producción científica sobre la cuestión disponible sobre todo (pero no únicamente) en la propia web. Por último, y dado que el posicionamiento además de una disciplina científica emergente es ya un sector económico, hemos analizado documentación sobre estrategias de posicionamiento en diversas empresas del sector. De esta batería de estu-

dios hemos podido derivar los conceptos que se presentan a continuación.

En primer lugar, vamos a (re)definir algunos conceptos que, aunque nos son familiares por otros ámbitos de la biblioteconomía y la documentación, al trabajar con ellos desde el punto de vista del posicionamiento web ven algo cambiada su definición tradicional (Codina, 2004).

1. Conceptos sobre posicionamiento web

1.1. Palabra clave

Es el término a través del cual se realiza o bien el análisis o bien la campaña de optimización de una página web (o ambas cosas). Es decir, cuando nos proponemos analizar (u optimizar) el posicionamiento de un sitio siempre lo hacemos a través de una o más palabras que nos sirven para comprobar, por ejemplo, cuáles son los sitios mejor situados. A estas palabras las denominamos con el mismo término que se utiliza en el mundo de la indización: palabras clave. No debemos olvidar que los motores de búsqueda no indizan conceptos (por mucho que pretendan hacérselo creer), sino cadenas de caracteres, que aquí llamaremos *palabras clave* de acuerdo con la terminología habitual tanto en el mundo de la indización como en el del posicionamiento. Por tanto, lo que debe preguntarse quien desee mejorar la visibilidad de un sitio es: “para qué palabra (o palabras) clave queremos que nuestro sitio web obtenga altas posiciones en las páginas de resultados de los buscadores”. Naturalmente, una palabra clave también es el término con el que se representa una necesidad de información; por lo tanto, al optimizar un sitio para una palabra clave, lo estamos haciendo de cara a las preguntas que harán los futuros usuarios o el futuro público del sitio.

La palabra clave para la cual deseamos optimizar un sitio puede ser una palabra única, como “automóviles”, o una frase, como “automóviles de ocasión”. Es más fácil optimizar para un conjunto de palabras o para una frase que para una sola palabra. Por ejemplo, siempre será más fácil posicionar para *vehículos and Barcelona*, que para cada una por separado. Además, se ha comprobado empíricamente que es muy difícil posicionar a la vez un mismo sitio para más de tres o cuatro palabras clave (cada una de ellas por separado). También es mucho más difícil optimizar un sitio por una palabra clave cuanto más competitiva sea la palabra clave, pues habrá muchos sitios web ya posicionados para ella, y ser uno de los primeros costará más que si elegimos una palabra para la que pocos se han posicionado. Sin duda es mucho más difícil posicionar un sitio para la palabra clave “web” que para la palabra clave “metamorfosis”. La primera tiene mucha

presión comercial, con muchos sitios queriendo posicionarse, a diferencia de la segunda. Existen herramientas que sugieren términos idóneos para posicionarse (**Gonzalo**, 2004).

1.2. Metadatos

Al igual que en otros ámbitos, en el contexto de la visibilidad los metadatos son datos sobre datos. En nuestro caso son datos sobre sitios o sobre páginas web que ayudan a su indización tanto como a su presentación (*ranking*) en las páginas de resultados de los motores. Por tanto, los metadatos ayudan a la visibilidad de los sitios. Ahora bien, en posicionamiento sucede que algunos elementos que técnicamente son metadatos (como las etiquetas “meta”) son ignorados por los motores de búsqueda; mientras que elementos de sitios web que, en rigor no son metadatos, los motores de búsqueda los utilizan *de facto* como tales. En la disciplina del posicionamiento se suelen considerar al menos cinco formas de metadatos:

1. Etiqueta <title> en la sección <head>, que corresponde al título de la página web que el usuario visualiza en la parte superior del navegador. A efectos de los motores de búsqueda, posiblemente es uno de los componentes de metadatos más importante.

2. Etiquetas <meta> en la sección <head>, entre ellas *keywords* para los descriptores, *description* para la descripción de la página web y *author* para indicar la autoría del sitio (las tres más comunes). De éstas, las dos últimas probablemente son las más importantes, siempre hablando en términos de visibilidad.

3. Atributo *title* en las etiquetas de los enlaces (<a>); genera el efecto adicional de que al pasar el cursor sobre un enlace se muestre el título del mismo en una pequeña ventana.

4. Atributo *title* en etiquetas de imágenes . Al igual que el anterior, genera el efecto adicional de que al pasar el cursor sobre una imagen se muestre el texto que se haya incluido dentro del atributo *title*.

5. Atributo *alt* en etiquetas de imágenes , que es un texto alternativo a una imagen en caso de que ésta no se visualice, ya sea por un problema en el servidor de la página web o porque se utilice un navegador no gráfico.

6. Contenido de las etiquetas <h1>,...<h6>. En general, los motores de búsqueda interpretan como indicadores valiosos para sus inferencias el contenido de las mencionadas etiquetas, en particular, las que indican mayor rango jerárquico en la estructura de la página (h1 y h2).

¿Por qué en posicionamiento consideramos que los anteriores elementos son metadatos? Se trata de una

convención útil porque los términos que aparecen en los mismos actúan *de facto* como contenedores de palabras clave que los motores de búsqueda toman como indicadores fiables del contenido de un sitio web.

1.3. Visibilidad y PageRank

Es el término acuñado, entre otros, por los ingenieros de *Google* para referirse a una medida que indica a) el número de enlaces que un sitio web recibe de otros sitios y b) la calidad de los mismos esto es, si esos sitios tienen a su vez una alta popularidad hipertextual, es decir, si esos sitios a su vez son muy enlazados. Por lo tanto, se trata de un rasgo que en parte es hereditario: un sitio web citado por un sitio web muy popular tendrá mejor posicionamiento que un sitio web citado (enlazado o *linkado*...) por uno poco popular.

La medida numérica que ha establecido *Google* relacionada con la visibilidad o popularidad hipertextual se denomina *PageRank*. El *PageRank* de un sitio es un número que será mayor cuantos más enlaces reciba de páginas web que, a su vez, tengan un buen *PageRank*. En su página para webmasters *Google* indica lo siguiente:

“*PageRank* interpreta un enlace desde la página A a la página B como un voto para la página B por parte de la página A. *PageRank* determina entonces la importancia de una página por el número de votos que recibe. También considera la importancia de cada página que emite el voto, por lo que los votos de algunas páginas se consideran de mayor valor, dando así a la página enlazada mayor valor. Las páginas más importantes tienen un mayor *PageRank* y aparecen en la parte superior de los resultados de búsqueda”.

<http://www.google.com/corporate/tech.html>

Existen al parecer dos índices de *PageRank*: uno de uso interno, que *Google* no da a conocer y al que denominaremos *PageRank Auténtico*, y otro que se presenta a través de una barra de herramientas del propio *Google*:

<http://toolbar.google.com>

Este *PageRank* al que denominaremos *PageRank Externo* consiste en una escala de 0 a 10, de manera que un sitio con una puntuación de 0 indica que se trata de un sitio que recibe muy pocos o ningún enlace, mientras que un sitio con una puntuación de 10 recibe al menos varios miles de enlaces de sitios que a su vez poseen índices de *PageRank* relativamente altos.

Ahora bien, más allá de la definición “oficial” que proporciona *Google* sobre el *PageRank*, se conocen los siguientes aspectos del mismo:

1. *Carácter logarítmico*. Los analistas (véase por ej., **Kent**, p. 220) creen que el *PageRank Externo*

(*PRE* a partir de ahora), es decir, el que da a conocer *Google* de cada página a través de su barra de herramientas, es de carácter logarítmico, probablemente con base 5. Esto significa que para un *PR* de 0, el *PageRank Auténtico* (*PRA* a partir de ahora) puede ser cualquiera entre 0 y 5; mientras que un *PRE* de 4, puede ser cualquier *PRA* de entre 625 y 3.125 y uno de 9 en *PRE* corresponde a un *PRA* que puede ir desde los nueve millones a más de cuarenta y ocho millones, etc. Así, pues, en realidad el *PRE* constituye un dato sumamente impreciso, ya que dos páginas con el mismo *PRE* pueden tener un número de enlaces (y un *PR* auténtico) muy distinto. Además, cuanto mayor es el *PRE* más impreciso es.

2. *Carácter “de adorno”*. Existe también un cierto consenso entre algunos analistas (por ej. el mencionado **Kent**, pero también los analistas de la empresa *Axandra*) sobre el hecho aparente de que el *PR* tiene actualmente muy poca relación con el grado de relevancia que utiliza *Google* para ordenar los resultados de las búsquedas, al menos el *PRE*, mientras que persiste la duda sobre cómo y qué medida afecta a los resultados el *PRA*. Es posible, que el *PRA* sí posea un auténtico valor para *Google* a la hora de determinar la relevancia o la importancia de una página. Diversas pruebas realizadas por la empresa *Axandra* sugieren que en determinadas listas de resultados el *PR* (ya sea el auténtico o ya sea el externo) no parece jugar actualmente un papel tan decisivo en su algoritmo de ordenación. Es decir, últimamente, se ha detectado que un sitio web puede obtener las primeras posiciones con un *PR* muy bajo.

A pesar de todo, *PageRank* sigue ofreciendo una información que permite realizar comparaciones en relación con la evolución de un mismo sitio web y, hasta cierto punto, entre distintos sitios web. Por supuesto no acaban aquí los aspectos de interés de *PageRank*, ni mucho menos. Con el tiempo, el estudio de qué es el *PageRank*, cómo se calcula, etc., se ha convertido en toda una especialidad propia con un fuerte carácter detectivesco. Remitimos al lector interesado a la propia página de *Google* para obtener más información así como a las referencias indicadas en la bibliografía de este artículo y de otros publicados en este número monográfico de *EPI* y en el anterior.

1.4. Relevancia

Es la capacidad que tiene un documento (un sitio o una página web) de satisfacer una necesidad de información de un usuario. Es decir, el grado de utilidad de un documento en relación con una pregunta concreta.

Los motores de búsqueda ordenan los documentos presentados en la página de resultados utilizando algoritmos que tratan de inferir la relevancia de cada página

en relación con cada pregunta, de modo que entregan los resultados ordenados por el grado de probabilidad de resultar útiles para la necesidad de información expresada por el usuario. Para realizar ese cálculo cada motor de búsqueda combina diversas medidas, normalmente consideran la frecuencia y la densidad de las palabras clave en cada página web recuperada, en cuanto a criterios internos a la página, así como algún otro indicador externo como los enlaces entrantes o el tráfico de la página (ver a continuación).

1.5. Popularidad y *Traffic rank*

La popularidad es el número de visitas que recibe un sitio o un dominio y, según algunas versiones, incluye también el número de páginas servidas por dicho sitio o dominio. Una de las medidas de popularidad más conocidas es *Traffic rank*, que determina *Alexa* — una empresa que realiza análisis y mediciones sobre tráfico en internet y que proporciona también una base de datos y un directorio de sitios web—. Este cálculo combina tanto el número de usuarios que visitan un mismo dominio web como el número de páginas vistas por los usuarios en ese dominio.

<http://www.alexa.com>

En su sitio web *Alexa* explica que su sistema de medición de la popularidad combina dos medidas que denomina como *reach* y *page views* respectivamente. http://pages.alexa.com/prod_serv/traffic_learn_more.html

Reach se expresa en porcentaje y mide el número relativo de usuarios de un dominio: el *reach* de *Alexa* indica el número de usuarios de un sitio por cada millón. *Page views* indica el número de páginas vistas de un dominio. Con los resultados acumulados de tres meses, *Alexa* combina ambas medidas en una sola en forma de un número que va de 1 a varios millones.

«PageRank es un número que es mayor para un sitio web dado cuantos más enlaces reciba de páginas web que, a su vez, tengan un buen PageRank»

Traffic rank sirve a *Alexa* para realizar *rankings* de páginas (del tipo “Los 40 principales”) en su directorio de sitios web por categorías, de manera que los sitios con más tráfico y mayor número de páginas vistas obtienen los primeros puestos en cada categoría temática. Las páginas con mayor número de visitantes y mayor número de páginas servidas muestran los números de *Traffic rank* más bajos. Por ejemplo, en diciembre de 2004, el sitio de la *CNN* obtenía un *Traffic rank* de 27, mientras que el *El país* era 1.546, lo que indica que de los dos sitios, el de la *CNN* era mucho

más popular (aunque un *Traffic rank* de alrededor de 1.500 sería envidiable para cualquier sitio).

<http://www.cnn.com>

<http://www.elpais.es>

Puede conocerse el *Traffic rank* de un dominio de dos formas: o bien consultando el directorio de sitios web de *Alexa* o bien utilizando la barra de herramientas de *Alexa*

<http://download.alexa.com>

A través del sitio de *Alexa* es posible saber por separado el *reach* y el *page view* de un dominio, mientras que con la barra de herramientas únicamente se puede conocer la medida combinada.

«La popularidad es el número de visitas que recibe un sitio y según algunos incluye también el número de páginas servidas»

La principal crítica que se puede hacer al *Traffic rank* es que se basa en los datos que proporcionan los usuarios que tienen instalada la barra de navegación de *Alexa*. En su web, la propia *Alexa* alerta sobre las posibles limitaciones de su sistema de medición. Entre otras cosas, advierte que cuantas menos visitas tiene un sitio, menos fiables pueden ser los datos que proporciona sobre el mismo, y al revés: cuantas más visitas tiene un sitio, mayor es la probabilidad de que los datos sean ajustados. No obstante, como sucede con el *PageRank*, no deja de ser un elemento de referencia que puede tomarse como medida de la evolución de un mismo sitio en el tiempo o para comparar diversos sitios entre sí.

Por último, es importante señalar que los datos de *Alexa* se refieren a un dominio y no a un sitio o a una página. Por ejemplo, sitios como

<http://www.msn.com>

<http://carpoint.msn.com>

<http://slate.msn.com>

son tratados por *Alexa* (según indican ellos mismos) como parte del mismo sitio, ya que todos son parte del mismo dominio *msn.com*.

1.6. Spam (malas prácticas)

En el ámbito del posicionamiento se denomina así a las malas prácticas destinadas a forzar una alta posición de una página web para una o más palabras clave, sin que tal posición vaya acorde con la relevancia real de la página web en relación con esas palabras clave.

Una de las técnicas primitivas más usadas para realizar *spam* en este contexto es poner repetidas veces un término como descriptor en los metadatos, de manera que el indicador del motor de búsqueda conside-

re que trata sobre ese tema, aunque en realidad trata de otro. Un ejemplo: si palabras como “música” o “dvd” son muy utilizadas por los internautas, una técnica de *spam* consistiría en poner esos términos como metadatos, a pesar de que el sitio web no trate en absoluto sobre ellos.

Otro tipo de *spam*, en este caso realizado de manera externa al sitio web, es lo que se viene denominando *Google bombing*, que consiste en enlazar numerosas veces, desde distintas páginas, a un sitio web utilizando para ello como texto de anclaje un término o expresión que no corresponde al contenido del sitio al que se está enlazando. Un ejemplo ficticio: imaginemos que queremos posicionar nuestro sitio web <http://www.ejemplo.com>, que está dirigido a la venta de libros de segunda mano, para la frase “Bill Gates”, de manera que cuando un usuario busque en *Google* esta frase obtenga nuestro sitio web en lugar de la página personal de **Bill Gates** o el sitio de *Microsoft*. Bastará con que consigamos que la frase “Bill Gates” sea el texto de enlace de salida en diversas páginas web hacia el sitio <http://www.ejemplo.com>

El *Google bombing* se puede realizar también de forma agresiva contra un sitio al que los autores del *Google bombing* desean denigrar. Veremos más adelante dos casos reales de este fenómeno. Como es natural, los administradores de motores de búsqueda consideran fraudulentas las prácticas de *spam*.

1.7. Anclaje

En terminología (x)html (y también en terminología de hipertextos) un anclaje es el punto de inicio o de destino de un enlace. Este punto de inicio puede ser un texto (ya sea una sola palabra o una frase) o una imagen. En la convención más habitual un texto de anclaje se muestra en color azul y subrayado. En términos de código fuente, un anclaje es el texto que aparece entre las etiquetas de inicio y de cierre de un enlace (<a> y). En posicionamiento es importante tener presente el concepto de anclaje (y no sólo el de enlace) porque, como veremos después, el texto que forma parte de las etiquetas de los anclajes tiene influencia en la visibilidad de un sitio ya que los motores de búsqueda lo interpretan como una forma de metadatos.

1.8. Enlaces recibidos, de salida e internos

Como es sabido, un enlace es la unión entre dos secciones de un mismo sitio web o entre dos sitios web. En posicionamiento los enlaces juegan un papel muy importante, ya que los motores de búsqueda utilizan el análisis de enlaces (o análisis hipertextual) para tratar de inferir no solamente los temas, sino la calidad o la posible condición de “autoridad” de un sitio web.

Los enlaces recibidos son los enlaces que un sitio recibe de otras sedes webs. Por ejemplo, si el sitio A es enlazado por los sitios B, C y D decimos que A tiene tres enlaces de entrada. Los enlaces de salida son los enlaces que salen de un sitio hacia otras webs. Por ejemplo, si el sitio A enlaza a los sitios P, Q, R, S, decimos que A tiene cuatro enlaces de salida. Por último podemos considerar los enlaces internos: son aquellos que permiten la navegación en el interior de un sitio. A efectos de visibilidad los más importantes son los enlaces de entrada.

Tanto la teoría hipertextual como la cibermetría aportan análisis mucho más detallados sobre las características de los enlaces. Remitimos al lector interesado a la literatura sobre hipertextos y cibermetría para conocer otras formas más complejas de interpretar y estudiar el mundo de los enlaces. Sin embargo a efectos de posicionamiento la breve referencia presentada aquí es suficiente para entender los aspectos básicos de los procesos de optimización de un sitio.

1.8. Grupo de referencia

Según sean las fuentes consultadas, existen en la Web desde unos pocos centenares <http://www.searchenginewatch.com> hasta unas cuantas decenas de miles de motores de búsqueda y directorios especializados <http://www.completeplanet.com>

Sin embargo es sabido que solamente existe un número muy pequeño de motores de búsqueda y directorios significativo. A nivel internacional y en el ámbito de los motores y directorios generalistas, apenas cuentan cinco buscadores. La inmensa mayoría del tráfico dirigido por motores (y directorios) está actualmente en manos únicamente de cinco grandes servicios; son los que llamaremos aquí el *grupo de referencia*. Aunque son bien conocidos, no está de más recordar cuáles son:

Google

Motor de búsqueda
<http://www.google.com>

Yahoo Búsquedas

Motor de búsqueda
<http://search.yahoo.com>

MSN Search

Motor de búsqueda
<http://search.msn.com>

DMoz

Directorio
<http://www.dmoz.org>

Yahoo

Directorio
<http://www.yahoo.com>

Existen otros motores, como *AskJeeves*, *Gigablast* y *Wisnut*, así como centenares de directorios locales y *web rings*, pero su relevancia es muy relativa debido al peso del grupo de referencia indicado.

<http://www.aj.com>

<http://www.gigablast.com>

<http://www.wisnut.com>

Se calcula que los tres primeros motores de búsqueda (*Google*, *Yahoo Search*, *MSN Search*) reciben el 90 por ciento de todas las consultas en la Web, con una clara mayoría para *Google* y *Yahoo Search* y un papel (de momento) menor de *MSN Search*. Por su parte, los dos directorios mencionados (*Yahoo* y *DMoz*) desempeñan un papel semejante en lo tocante a uso de directorios.

Dejando de lado por el momento la obvia preocupación por el hecho de que tan reducido número de actores tenga tanta influencia en el tráfico de información en la Web, la consecuencia pragmática desde el punto del posicionamiento es que ninguna actividad de planificación o de estudio de visibilidad puede ignorar al grupo de referencia. A partir de ahora, cuando hablemos de motores de búsqueda y directorios, salvo indicación en contrario, nos estaremos refiriendo precisamente a este “grupo de referencia”.

2. Criterios de optimización de páginas web

Con los conceptos anteriores podemos entrar a considerar ahora algunos criterios de optimización de un sitio web. Puesto que el objetivo del posicionamiento es lograr que un determinado sitio web aparezca entre los primeros lugares de los resultados cuando se efectúa una consulta en un buscador, evidentemente la clave de la optimización está en conocer lo mejor posible qué criterios están considerando los motores de búsqueda para organizar su *ranking* de resultados.

Ahora bien, los motores guardan en secreto sus algoritmos concretos, ya que poseen un valor industrial y comercial cada vez mayor. Además, suelen ir haciendo variaciones sobre los mismos, por lo que estos criterios no se conocen con exactitud, sino que se deducen o se infieren mediante pruebas y estudios de resultados. De tales estudios no se pueden deducir los algoritmos exactos, pero sí se puede llegar a saber cuáles son los factores que tienen alguna influencia.

De hecho, es sabido que tanto las características internas de las páginas web como las externas tienen efecto en los resultados. Veamos cuáles son las más importantes según un cierto consenso existente en la mayoría de analistas del mundo del posicionamiento web. Para esta exposición nos hemos basado en **Nobles** y **O’Neil** (2000), **Marckini** (2001), **Thurrow**

(2003) y Kent (2004), así como en los numerosos informes publicados en sitios web de reconocido prestigio (varias veces citados en este trabajo) como *Search Engine Watch* y *Axandra*.

2.1. Criterios intrínsecos (internos de la página web)

En este sentido hay que señalar que dependiendo de si el posicionamiento se está haciendo para una sola palabra clave o para varias palabras, cambiarán en parte los criterios que hay que considerar.

2.1.1. Para una sola palabra clave

Los elementos de un sitio o de una página web que influyen en su visibilidad para una sola palabra clave (es decir, si la pregunta contiene una sola palabra) son:

I) Frecuencia

Número de veces (sumatorio) que aparece la palabra clave en el documento. Cuanto mayor es la frecuencia, más relevante es la página. Si la frecuencia es demasiado alta es posible que los motores “sancionen” a la página por considerar que está realizando *spam*. Esto último sucederá si la frecuencia es mucho más alta de la que puede esperarse de acuerdo con las frecuencias típicas de la lengua. Por ejemplo, es estadísticamente improbable que en un texto en lenguaje natural de 100 palabras en total aparezca una palabra determinada un total de 50 veces. Típicamente, ningún término posee una frecuencia de aparición tan alta, un 50%.

II) Posición (ubicación)

Lugar donde aparece el término de búsqueda. Una página web donde el término figure en el título será considerada más relevante que si figura en el cuerpo. Los lugares considerados de mayor importancia son la etiqueta <title> de la sección <head>, las etiquetas <description> y <keywords> también de la cabecera y de los primeros y los últimos párrafos de la página. Debido al *spam* los motores están dejando de considerar algunos metadatos de la sección <head> como información fiable (en particular los metadatos de la etiqueta <meta> *keyword*). También son importantes las palabras incluidas en los enlaces de salida y los enlaces internos.

III) Emergencia

Número de orden de aparición de la palabra dentro de una sección. Si el término figura al inicio del título es más importante que al final. Si está al inicio del cuerpo del texto será más importante que si está en la parte media, etc. Por ejemplo, en un título como el siguiente: <title>Bienvenidos a la página de la Universidad ZXY sobre historia del cine</title>, los términos “historia” y “cine” tienen una mala posición al final

del título. La posición óptima debería ser ésta: <title>Historia del cine: bienvenidos a la página de la Universidad XYZ</title>. El orden de aparición de una palabra (o emergencia) como indicador de relevancia puede computarse por parte de los motores con relativa facilidad debido, entre otras cosas, a las características del índice que utilizan, denominado fichero invertido.

IV) Densidad

Frecuencia dividida por el número de palabras de la página. Cuanto mayor es la densidad mayor es la relevancia, siempre que se mantenga en unos márgenes estadísticos normales, pues como se ha dicho los motores de búsqueda pueden penalizar frecuencias muy altas por considerarlas *spam*.

2.1.2. Para más de una palabra clave

En este caso se tendrán en cuenta además estos dos criterios adicionales:

V) Diversidad

Número de palabras clave de la consulta de búsqueda presentes en el documento.

VI) Proximidad

Número de palabras entre dos (o más) términos de búsqueda. En general, cuantas menos palabras separen a los dos términos de búsqueda en el documento, mayor es su relevancia. La adyacencia entre las palabras sería el caso óptimo de proximidad.

Además de los anteriores los analistas han identificado una serie de criterios adicionales que al parecer observan los motores de búsqueda a efectos de sus *rankings*, por ejemplo, si la palabra clave aparece marcada con etiquetas como , <i>, tipo <h1>, etc.

2.2. Criterios relacionales (externos a la página web)

Básicamente hay dos factores externos al sitio web que tienen influencia en la visibilidad: los enlaces de entrada y el tráfico del sitio.

Entre los criterios basados en análisis de enlaces destacan, según la mayoría de los observadores, los siguientes:

VII) Número de enlaces recibidos por un sitio

A igualdad de otros factores, un sitio web será más relevante cuantos más enlaces reciba de otras páginas web.

VIII) Peso de los enlaces recibidos por un sitio

No todos los enlaces de entrada otorgan el mismo valor para calcular la relevancia desde el punto de vista de los motores. Las páginas que a su vez son muy



Figura 1. Ejemplo de Google bombing

enlazadas otorgan más valor que las páginas poco enlazadas. Dicho de otro modo: el enlace procedente de una página personal otorga menos valor que el enlace procedente de un sitio como *Yahoo!*, o de un sitio como una universidad, por mencionar dos ejemplos fáciles. En general, los sitios que tienen muchos enlaces de entrada se consideran “autoridades”, y tiene mayor influencia un enlace de entrada procedente de una “autoridad” que un enlace de un sitio “estándar”.

IX) Texto del anclaje

Algunos buscadores, y en particular *Google*, consideran el texto que sirve de anclaje de inicio como una pista o una inferencia para calcular la relevancia de la página de destino. En casos extremos, si muchas páginas web contienen, por ej., el texto “biología molecular” enlazado hacia un mismo sitio web, es posible que ese sitio web llegue a ser considerado como una “autoridad” por *Google* para ese término, incluso aunque el sitio web considerado no contenga el término “biología molecular”.

Este criterio está siendo utilizado para hacer el tipo de *spam* que comentábamos anteriormente, el *Google bombing*. Veamos dos ejemplos reales de este fenómeno. Si buscamos en *Google* “*miserable failure*” (figura 1) o “ladrones” (figura 2), el primer resultado respectivamente nos va a sorprender, pues ninguna de las dos sedes web contienen esas palabras.

Por su parte, los criterios que tienen en cuenta el tráfico de visitas son dos:

X) Número de visitas que recibe la página

Esta medida se refiere al número de visitantes que tiene un sitio web y la ha desarrollado, entre otras empresas, *Alexa* bajo la denominación de *Traffic rank*, que ya hemos explicado.

XI) Número de páginas visitadas



Figura 2. Ejemplo de Google bombing

Además del número de visitas, se suele considerar el número de páginas vistas en un sitio. Forma parte también del indicador *Traffic rank*.

3. El ciclo de vida de una campaña de posicionamiento

Cuando se acomete una estrategia de posicionamiento es conveniente seguir una serie de pasos que aquí hemos clasificado en cuatro fases (Codina, 2004):

1. Análisis del sitio que se quiere posicionar.
2. Diseño de las estrategias de posicionamiento que se van a emplear.
3. Implementación de las técnicas apropiadas para conseguir el objetivo.
4. Seguimiento de los resultados.

Tras la cuarta fase, y en función del posicionamiento adquirido por el sitio, deberán aplicarse de nuevo las tres primeras hasta alcanzar los resultados esperados.

Dados los ritmos de cambio de los índices de los motores de búsqueda, es aconsejable repetir el ciclo (análisis, diseño, implantación, seguimiento) al menos una vez cada seis meses, e incluso mensualmente según la importancia que otorguemos a la visibilidad del sitio objeto de la campaña.

Por otro lado es importante destacar que, dependiendo del tipo de sitio web (dimensiones y variedad de sus secciones), deberán aplicarse los análisis y las medidas de posicionamiento indicados para tantas secciones o subsecciones de interés estratégico como tenga. La razón es la siguiente: en primer lugar, dada la naturaleza de la Web, no hay por qué suponer que los visitantes accederán a través de la página principal del sitio. En segundo lugar, si éste es rico en contenidos será necesario posicionar cada una de las secciones que

los tengan, y no solamente la página principal. Con mayor motivo deberá aplicarse este procedimiento distribuido a empresas o corporaciones que posean diversos dominios.

Una vez presentada esta visión global de una campaña de posicionamiento, vamos a examinar con más detalle cada una de sus fases principales.

3.1. Fase de análisis

Su objetivo es determinar la palabra o palabras clave para las cuales se desea optimizar el sitio, así como identificar los posibles *partners* que pueden enlazarlos de forma justificada (con enlaces recíprocos). De forma justificada significa que deben escogerse sitios que por su contenido, público visitante u objetivos tengan alguna relación real con el nuestro.

Las palabras clave serán elegidas igualmente en función de los tres elementos mencionados: contenido, objetivos y público del sitio, haciendo énfasis en este último elemento: el público al que está destinado nuestra web.

Algunas veces la palabra clave será obvia. Por ejemplo, el sitio web del ayuntamiento de un municipio debería optimizarse para al menos el topónimo de ese municipio; el de una marca comercial para la denominación de esa marca, etc. No obstante, incluso en casos obvios, probablemente deberemos pensar en una o más palabras clave adicionales. En el caso de la web de un ayuntamiento seguramente desearemos optimizarla por otros términos, por ejemplo, “turismo” (suponiendo que la web del municipio contenga información turística local de calidad, etc.). Ahora bien, en todo caso lo que obtendremos como resultado del análisis serán conceptos o ideas pero no necesariamente palabras clave.

Por tanto, en una segunda fase necesitaremos convertir esos conceptos en palabras clave concretas. De nuevo, en algunos casos el paso del concepto a la palabra clave puede ser obvio, pero en otros no. Así, el sitio de un museo de pintura de la ciudad X puede desear posicionarse justo para esas palabras clave: “pintura” y “X”. Pero en el caso del sitio de una empresa, los responsables de posicionamiento en estrecha relación con la gerencia y el departamento comercial deberán decidir en qué palabras concentrar su estrategia. Una web de comercio electrónico en artículos de informática deberá decidir si opta por términos como “ordenadores”, “software”, etc., o más bien por “multimedia”, “ocio digital”, etc., por citar un ejemplo habitual. Necesitará para ello identificar cuáles son las palabras clave que suelen utilizar sus clientes o su público potencial.

La determinación de las mejores palabras clave candidatas para optimizar un sitio es una especialidad completa dentro de lo que podemos llamar la industria del posicionamiento. Para ayudar en esta tarea existen, al menos, cuatro tipos de herramientas:

1. Análisis del sitio por un experto humano y/o con programas específicos como *IBP4* o *Arelis*, que comentaremos más adelante.

2. Análisis de frecuencias de la palabra clave y de sus diversos sinónimos utilizando los propios motores de búsqueda.

3. *Benchmarking* por comparación de la mejor competencia, bien realizado por un experto humano, bien mediante un programa específico o más habitualmente combinando ambas cosas. A pesar de las frecuentes críticas a tales mediciones, este análisis contempla, típicamente, el *PageRank* y el *Traffic rank* (ver más arriba) como medio de disponer de elementos de comparación relativa.

4. Programas de ayuda a la selección de palabras clave de los motores de búsqueda vinculados con sus sistemas de anuncios (*pay per click*).

El resultado de los análisis es un conjunto de datos que convendrá que organicemos o anotemos en forma de fichas (se muestran algunos ejemplos en el apartado 5):

1. Ficha de identidad del sitio (y de cada sección principal si es el caso), incluyendo:

- Título
- Frase descriptiva (*tagline*)
- Metadatos-Description
- Metadatos-Keywords

2. Ficha de propuesta de recomendaciones de optimización para el sitio en cuanto al contenido del mismo, así como del texto que debe figurar en los metadatos y en el cuerpo, a poder ser, inmediatamente después de la etiqueta <body>, así como en lugares sensibles de la página como los atributos “title” de las imágenes y de los enlaces y atributos “alt” de las imágenes.

3. Lista inicial de *partners* que pueden enlazar hacia nuestro sitio.

3.2. Fase de diseño

Su objetivo, tal como ya hemos señalado, es la *optimización específica* de: la página principal, cada una de las páginas principales de las diversas secciones y subsecciones de la organización, así como la *optimización genérica* de todas las páginas de la organización sobre la base de una política de metadatos.



Figura 3. Toolbar de Google

Para llevar a cabo la fase de diseño algunos programas de edición de páginas web o sistemas de gestión de contenidos pueden automatizar en parte esta optimización, aunque nunca sin una planificación previa.

En todo caso el resultado de esta fase debería ser una propuesta de contenidos y de código fuente de las diversas secciones del sitio web optimizado en los apartados de:

- Metadatos (en el sentido amplio indicado en 1.2.)
- Contenido del sitio
- Navegación (rótulos de las secciones del sitio)
- Enlaces de navegación internos
- Enlaces de salida

3.3. Fase de implantación

Una vez realizadas las modificaciones en los contenidos y el código fuente del sitio, puede comenzar la fase de implantación, en la que se desarrolla finalmente la campaña. Para ello se utilizan formularios de alta de los distintos directorios y motores, fichas de seguimiento, etc., con el fin de obtener básicamente dos resultados:

1. *Aparición en la primera página de resultados* de los motores de búsqueda del denominado “grupo de referencia” (*Google, Yahoo, Dmoz*, etc.). Se recomienda empezar dando altas en directorios, dejar pasar dos semanas y dar de alta en los motores de búsqueda. Los directorios y motores de búsqueda actualizan sus índices entre una y cuatro veces al mes. Por tanto, conviene dejar pasar al menos un mes después de una campaña de altas antes de empezar a analizar resultados. Es aconsejable dar de alta el sitio en el denominado grupo de referencia por razones obvias, pero, dependiendo de sus características, convendrá darlo de alta en directorios temáticos específicos y en motores y directorios regionales si es el caso (así, el sitio web de un museo puede darse de alta también en directorios especializados en arte; si el museo es de tipo regional o local, deberá hacerlo también —en caso de existir— en algún directorio local, etc.).

2. *Captación de enlaces hacia el sitio*. Además de los enlaces que nuestro sitio puede obtener de los directorios en los que se haya dado de alta, sin duda podremos obtener enlaces de otras webs vinculadas con

nuestra institución, de organismos con intereses similares (siempre que podamos alegar que nosotros los hemos enlazado también) y, por supuesto, de los citados directorios temáticos y regionales.

4. Algunas herramientas de posicionamiento

En este apartado hemos hecho una selección de algunas de las herramientas más utilizadas. Comenzamos presentando algunas de coste cero que nos ayudan a determinar el posicionamiento de una página web y dan facilidades para mejorarlo, y pasamos a continuación a comentar algunas de pago, que en ocasiones disponen de una versión gratuita.

4.1. Toolbar de Google: PageRank

<http://toolbar.google.com>

Se trata de una barra de herramientas que se integra en el navegador —por el momento sólo en *Internet Explorer*, pero se espera una nueva versión para *Firefox*— y permite realizar las búsquedas en *Google* sin necesidad de acudir a la página del buscador, directamente desde la barra. Las funcionalidades son las mismas que las del buscador: buscar en toda la web, en el sitio de la página actual (equivale a la búsqueda *site*, buscar en el directorio, en las *news*, en las imágenes... Además ofrece otros servicios como destacar (*highlight*) los términos de búsqueda en el interior de la página o bloquear *pop-ups*.

Pero hay un elemento del que ya hemos hablado al definir el concepto de popularidad: el *PageRank* de la página visitada. Como sabemos, su valor puede oscilar entre 0 y 10. La forma de indicarlo en la barra es mediante un diagrama similar a un termómetro: cuanto mayor sea la línea, mayor es el *PageRank*, y por lo tanto mayor será el número de páginas que tienen un enlace dirigido a ella y/o mayor será la visibilidad de esas páginas que la enlazan. Si mantenemos el cursor sobre el diagrama, aparecerá una ventana con indicación del valor numérico exacto.

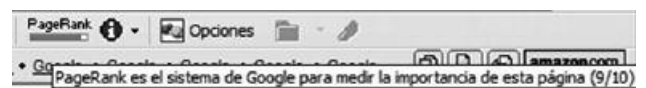


Figura 4. PageRank de la Toolbar de Google

Para que la *toolbar* pueda ofrecernos la información de *PageRank* es necesario que conozca en qué página estamos navegando, pues si no es así no podría

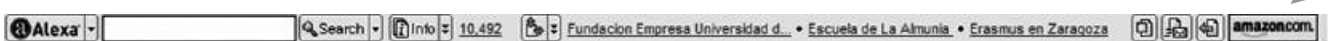


Figura 5. Toolbar de Alexa

darnos ese dato. Por eso, la *toolbar* de Google se considera un caso de *spyware*; esto es, programas que “vigilan” la navegación del usuario, en este caso con la intención de proporcionarle un indicador de visibilidad del sitio web donde se encuentra.

4.2. Toolbar de Alexa: Traffic rank
<http://www.alexa.com>

Alexa, de la empresa Amazon, proporciona una barra de herramientas que también queda integrada en el navegador y que permite efectuar búsquedas en Google, en el tesoro y en el diccionario Merriam-Webster, y en la web de Amazon. Otras funcionalidades que tiene son la lista de enlaces a sitios relacionados con el que tenemos en pantalla, el número de sitios web que tienen un enlace a esa página, la velocidad de descarga del sitio y la posibilidad de dirigirnos a Internet Archive Wayback Machine, donde se muestra la evolución gráfica de cada sitio web. Pulsando en *Site info* se obtiene un resumen de datos relacionados con el sitio web, se puede pedir más información y ver los gráficos de tráfico de visitas de los últimos meses.

Además de esa integración para las búsquedas, lo que da valor a la barra de Alexa es la información añadida que nos da de los sitios web por los que navegamos: a veces proporciona una captura de pantalla y



Figura 6. Alexa provee información sobre el Traffic rank, el número de sitios que citan cada página, la velocidad de descarga, la fecha desde la que se tiene constancia de que existe la página, las páginas más visitadas por los usuarios que han entrado en este sitio, una captura de pantalla de la home, y gráficos estadísticos del tráfico de visitas en los últimos meses

siempre nos da información sobre el tráfico de visitas del sitio, a lo que llama *Traffic rank*.

«Traffic rank sirve a Alexa para realizar rankings de páginas en su directorio de sitios web por categorías, de manera que los que tienen más tráfico y mayor número de páginas vistas obtienen los números más bajos en cada categoría temática»

Como decíamos, se trata de una medida estadística calculada según dos variables: el número de usuarios que visitan un sitio web y el número de páginas que han visto. El valor del *Traffic rank* se interpreta como un ranking en el que los mejores obtienen los primeros puestos, por lo que el valor máximo es 1. Si miramos el *Traffic rank* de una página cualquiera ¿qué nos indica la cifra que obtenemos?, ¿es un buen o un mal *Traffic rank*? Depende. La mejor forma de apreciar si es un buen número es buscar el *Traffic rank* de las sedes web de instituciones del mismo tipo. No tiene sentido comparar el *Traffic rank* de una página personal con el de Yahoo! o con el de Hotmail, pero sí lo tiene comparar el de varias universidades, varios ayuntamientos, varias empresas que se dedican a la misma actividad, etc.

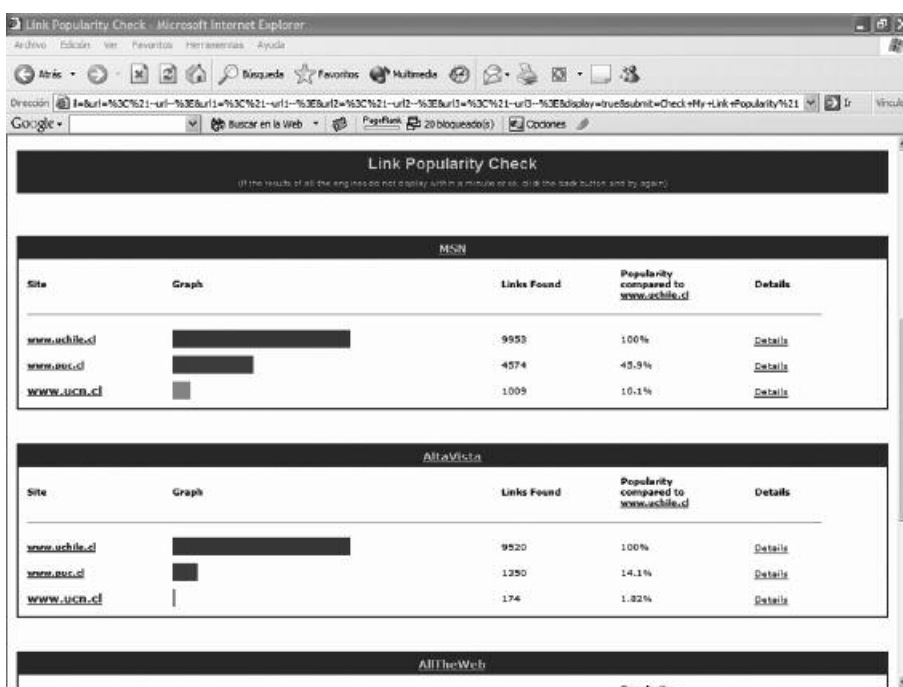


Figura 7. Link Popularity Check

Domain	Total	AllTheWeb	AltaVista	HotBot/Google	HotBot/Inktomi	MSN Search
www.upf.edu	3.974	2.486	4	304	582	598
www.ub.edu	25.306	18.749	0	3.560	1.454	1.543
www.uab.es	29.555	8.481	68	2.270	9.303	9.433
www.upc.es	38.593	16.924	142	2.390	9.432	9.705

Figura 8. Check your Link Popularity

Al igual que la barra de *Google*, para poder ofrecer esta información sobre cada sitio web por el que se navega es necesario que *Alexa* sepa en qué página nos encontramos en cada momento, por lo que este programa también entra en el grupo de los *spyware*.

4.3. Link Popularity Check

<http://www.linkpopularitycheck.com>

Esta herramienta indica cuántos sitios web tienen un enlace dirigido al que estemos analizando y cuáles son. Esta opción no es novedosa, muchos buscadores la tienen contemplada desde hace años con el limitador “site:”, por ejemplo *Altavista* y *Google*. La ventaja de usar *LPC* es que permite, en una sola vez, conocer la popularidad (que sería la visibilidad según las definiciones dadas en este artículo) de hasta tres páginas web según los datos de varios motores de búsqueda: *MSN*, *Lycos* y *Altavista*.

4.4. Check your Link Popularity

<http://www.checkyourlinkpopularity.com>

Proporciona las mismas funcionalidades que la anterior: envía un test de popularidad (como decimos, la tendencia es llamarlo visibilidad) de una o más páginas web a los motores de búsqueda *AllTheWeb*, *Altavista*, *Hotbot* y *MSN*. La diferencia está en que en este caso es necesario instalar un software y permite que se introduzca una gran cantidad de sitios web para analizar, mientras que *Link Your Popularity Check* funciona directamente online y permite hacer el análisis en un máximo de tres.

4.5. IBP4 (Instant business promoter)

<http://www.axandra.com>

Axandra es una empresa alemana especializada en servicios de posicionamiento web. Desde su sitio web se pueden descargar las versiones demo de *IBP4* y *Areliis*, dos softwares que permiten automatizar algunas tareas de ayuda al posicionamiento.

IBP4 realiza el análisis de posicionamiento a partir de una palabra clave elegida y de un sitio web. Para ello cuenta con las siguientes funciones:

—Chequeo de la posición en que se encuentra el sitio web analizado —y otros sitios que de los que nos interesa conocer ese dato— en varios buscadores cuando se consulta por el descriptor elegido. Esta opción facilita la tarea de realizar la búsqueda en varios motores y revisar en qué posición está nuestro sitio y los de la “competencia”. La versión demo permite usar

esta opción con un solo buscador cada vez, en cambio la versión de pago no tiene limitación en el número de los consultados simultáneamente.

—Lista de los 10 sitios mejor posicionados para ese descriptor según un motor de búsqueda. Si se quiere obtener el resultado de varios motores habrá que repetir la operación, una vez para cada motor, en la versión *lite*.

—Densidad de palabras en distintas partes del documento. Hace un análisis de la cantidad de veces que aparecen las palabras en relación con el número de palabras total del sitio web. La versión demo realiza esta tarea en el título (etiqueta *title*) y en los descriptores (etiqueta *keywords*), la versión de pago amplía el análisis a las demás partes del sitio web. Se puede hacer de forma comparativa con otros sitios que aparezcan mejor posicionados que el nuestro, así conoceremos la densidad en ellos y podremos mejorar la nuestra. El resultado se presenta en forma de informe.

—Generación de descriptores. *IBP4* consulta las herramientas de ayuda para la generación de descriptores (por ejemplo, para España ofrece información de *Spotting* y *Overture*). Son servicios especializados en anuncios por palabras, que en el caso de *IBP4* se utilizan para indicar cuántas veces se realizaron búsquedas en los últimos 30 días (*Spotting*) o en el último mes (*Overture*) por las palabras que le indicamos. De esta forma podemos saber si el término para el que nos queremos posicionar es muy demandado por los usuarios de los buscadores. Según *Spotting*, la búsqueda “posicionamiento buscadores” fue introducida 36.286 veces, mientras que “posicionamiento web” se

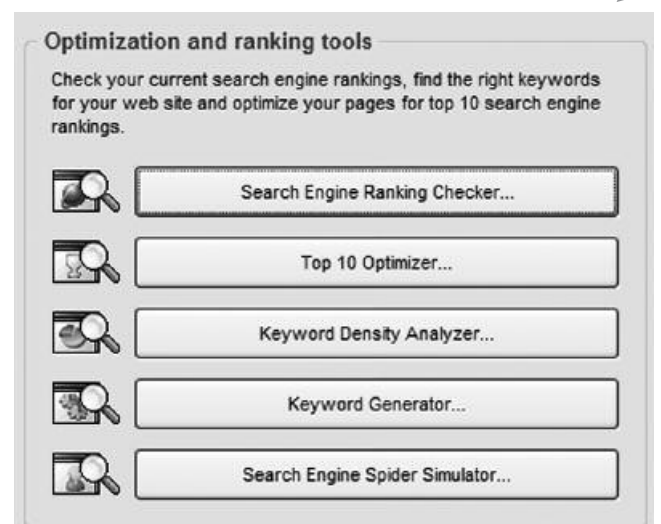


Figura 9. IBP3

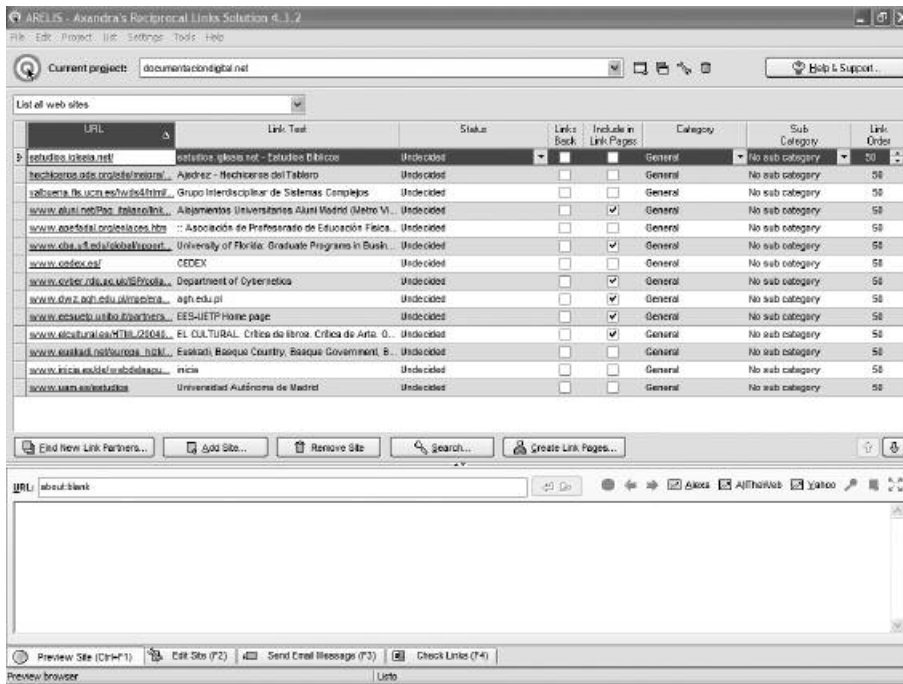


Figura 10. Arelis

éstos. Este dato nos indica qué sitios sería interesante que nos citaran a nosotros, esto es, que tuvieran un enlace hacia nuestra web.

2. Indicar una palabra clave para la que posicionarnos. *Arelis* hace una búsqueda en varios motores para saber qué sitios web están mejor posicionados para esa palabra y proporciona un listado de esas webs, ya que son consideradas como posibles socios también a los que citar y de los que recibir citas.

3. *Arelis* analiza desde qué sitios web existen enlaces al nuestro, pues sería conveniente citarlos también a ellos en nuestro sitio.

usó en 65 ocasiones (datos tomados en diciembre de 2004).

<http://es.espotting.com/home.asp>

<http://www.content.overture.com/d/?mkt=es>

—Simulación de robot de búsqueda. Esta opción realiza un informe sobre cómo vería un robot o *spider* nuestro sitio web, es decir, una vez eliminadas las imágenes, los códigos JavaScript y las CSS, pues los robots sólo pueden leer el código html.

La otra gran funcionalidad de *IBP4* es la posibilidad de dar de alta automáticamente el sitio en multitud de motores. Esta es una opción poco aconsejable, pues algunos motores detectan esta alta masiva y la penalizan, por lo que no la recomendamos.

4.6. Arelis (Axandra's reciprocal links solution)

<http://www.axandra.com>

Este otro software tiene como objetivo ayudar a crear una “red de negocios” (*business network*), lo que complementa a *IBP4* centrándose sobre todo en el *benchmarking*, pues localiza posibles *partners* o socios a los que citar (enlazar) y —lo más importante— por quien ser citados (enlazados). Para ello se deben seguir estos pasos:

1. Indicar cuáles son nuestros competidores. *Arelis* proporciona una lista de los sitios web que enlazan a

Head	
Metadatos	
Texto	
Navegación	
Enlaces	
...	...

Ficha 1

La versión *trial*, que es gratuita, limita este listado a 100 socios, hace un envío simulado de emails a estos contactos, pues lo hace a nuestra propia dirección, así como una simulación de directorio, pues los links que pone son a direcciones genéricas. Las versiones de pago superan estas limitaciones y ofrecen algunos servicios más.

4.7. Otros programas comerciales

Recientemente ha ido saliendo al mercado un gran número de programas que ofrecen servicios en esta línea.

Por ejemplo, *Zeus Internet Marketing Robot* ayuda a generar links recíprocos entre varios sitios web; *Active WebTraffic*, *TrafficSeeker Website Submission Software*, *Search Engine Seeker*, *Super Submission Software* y muchos otros ayudan a dar de alta el sitio web en motores y directorios, incluyendo la gestión de la información de las etiquetas meta.

La mayoría de ellos tienen una herramienta que proporciona el ranking de la posición que ocupa la página web en los motores más importantes según las palabras clave seleccionadas; incluyen un generador de etiquetas meta, así como información de la densidad de las palabras clave en la página web.

Título	
Tagline	
Descripción	
Palabras Clave	

Ficha 2

Motor de Búsqueda/ Directorio	Alta	Modificación	Fecha
...
...

Ficha 3

Indicadores	Resultados	Fechas
Traffic Rank		
PageRank		
Número de enlaces en Motor 1		
Número de enlaces en Motor 2		
...

Ficha 4

Motor	PC1	PC2	PC3	PC4
Google				
AltaVista				
MSN				
...

Ficha 5

4.8. Generación de palabras clave y sistemas de pago por clic

Existen programas, en muchas ocasiones gratuitos, que suministran una lista de palabras a partir de las palabras clave facilitadas por quien desea posicionarse; esta lista se puede generar basándose en variaciones de la palabra (singular/plural, palabras derivadas, de la misma familia...) o incluso utilizando software específico basado en ontologías.

«En el posicionamiento ético solamente se toman medidas que ayuden a los motores a indizar mejor el sitio web. El responsable del posicionamiento es consciente del modo en que actúan los motores de búsqueda y optimiza su sitio para que lo lean de la forma más adecuada posible»

El objetivo de estas herramientas es facilitar la tarea de seleccionar las palabras clave que serán más efectivas, para lo cual comprueban los resultados que se obtienen con las diferentes palabras en los buscadores. Algunos motores como *Altavista*, *Alltheweb* o *Teoma* disponen de una opción de refinar búsquedas basada en estas técnicas de palabras similares.

También se hallan programas específicos como *WordTracker* que aportan una lista de palabras variantes o relacionadas con las que se le indican y dice cuántas veces han sido buscadas en los últimos dos meses; o webs que brindan listados de las palabras más buscadas, por ejemplo *MetaSpy* y *Search Engine*

Watch. Pero quien más rendimiento está sacando de la generación de palabras clave son las empresas de *pay-per-click* como *Google*, *Overture* y *Espotting*, que realizan de forma gratuita un servicio similar al descrito con el objetivo de ofrecer un servicio de publicidad online, que describimos a continuación.

Como se viene observando en los dos últimos años, algunas empresas han encontrado en los sistemas de anuncios por palabras una forma de publicidad en los buscadores muy efectiva. *Overture* (ahora propiedad de *Yahoo!*) fue la primera en ofrecer este servicio de publicidad online, ahora seguido por algunas otras como *Google*. Este sistema de pago consiste en que el anunciante paga por cada clic que un usuario realiza en el anuncio, lo que le garantiza que pagará sólo por el tráfico que llegue por este sistema. Todos los sistemas de pago por clic cuentan con una herramienta de generación de palabras clave que sugieren al anunciante qué términos relacionados con el que él propone son los idóneos para publicitarse, es decir, qué palabras clave de consulta en los buscadores son las que más le conviene tener vinculadas para que el anuncio aparezca cuando un usuario realice esa consulta en el motor de búsqueda.

En este tipo de herramientas el cliente decide cuánto quiere pagar por cada clic realizado a su página. Cuanto más alto sea este precio mayor probabilidad habrá de que el anuncio aparezca.

Como hemos señalado, *Overture* y *Espotting* son los dos servicios más populares dentro de esta categoría, junto con *Google AdWords*. Esta última herramienta de *Google*, a diferencia de las anteriores, permite elegir el idioma y el país donde interesa hacer publicidad. En este caso, para hacer una lista de términos

sugeridos hace uso de una ontología propia desarrollada a partir de la compra por parte de *Google* de la empresa *Applied Semantics*, dedicada al desarrollo de tecnología lingüística.

5. Modelos de fichas de trabajo para posicionamiento

5.1. Recomendaciones de optimización de página principal o de secciones principales del sitio

Como ya hemos explicado, el paso previo en una campaña de posicionamiento web es la optimización del código fuente de la página. En una ficha (o base de datos) como la que indicamos en la ficha 1 recogeremos las recomendaciones principales. Dependiendo del contexto, deberemos anotar tantas fichas como secciones o subsitios (o sitios relacionados) tenga la organización. Una vez que el sitio esté optimizado de acuerdo con las recomendaciones de la ficha, podremos proceder a la campaña de altas.

5.2. Identidad del sitio o de cada sección principal del sitio

Esta es una ficha o documento (ver ficha 2) que debe rellenarse inmediatamente antes de empezar una campaña de altas (e inmediatamente después de haber optimizado el código fuente del sitio web). Naturalmente, el título, la *tagline* (frase simple de explicación), etc., deben consensuarse con el equipo de dirección del sitio web puesto que los datos de esta ficha son un condensado de su identidad. Esta identidad no puede ser distinta de la visión que tiene del sitio la dirección del mismo. La función de esta ficha, además de disponer de una identidad eficaz, consiste en asegurarnos que introducimos los mismos datos en los distintos directorios y motores donde pensemos dar de alta el sitio web.

5.3. Altas o Modificaciones

El objetivo de esta ficha (ficha 3) es mantener un registro o recordatorio de las diversas acciones de alta en directorios y motores.

5.4. Análisis de resultados: I. Indicadores estadísticos

Mediante esta ficha (ficha 4) podemos realizar un seguimiento cronológico de los indicadores más importantes en el posicionamiento: el *Traffic rank*, el *PageRank* y el número de enlaces encontrados en diversos motores. Este análisis debe realizarse antes, durante y después de una campaña de posicionamiento.

5.5. Análisis: II. Indicadores de posicionamiento

Al igual que la ficha anterior, nos permite realizar un seguimiento de la ubicación de nuestro sitio en diversos motores para cada una de las palabras clave pa-

ra las cuales deseamos posicionarnos de manera óptima (ficha 5).

6. Conclusiones

Hemos examinado los conceptos básicos y algunas herramientas y procedimientos del posicionamiento web. No obstante, podríamos preguntarnos qué sucedería si un sitio web de calidad no aplicara ningún procedimiento especial de optimización para buscadores.

Para responder nos serán de utilidad dos parejas de conceptos: *espontáneo* vs. *planificado* de un lado; y *ético* vs. *fraudulento* por otro. Comencemos por la primera pareja. El posicionamiento *espontáneo* es el que obtiene de forma “natural” un sitio web que no ha sido optimizado de forma expresa o consciente. Podemos decir que es, sin duda, el que prefieren los motores de búsqueda (es decir, los propietarios de los motores de búsqueda). La razón es que de este modo todo el poder está del lado del motor de búsqueda y, además, así no corren los riesgos del posicionamiento fraudulento.

Sin duda sería también el ideal para los propietarios de páginas web, porque tendrían una tarea menos a realizar. Sin embargo, confiar en un posicionamiento espontáneo no es realista. En un mundo ideal tal vez funcionaría, pero no en el nuestro. Sin una campaña *planificada* de posicionamiento puede que nuestro sitio web nunca sea indizado por un motor de búsqueda y puede que nunca aparezca en ningún directorio de referencia. Además, aun en el caso de que el sitio apareciera en un motor de manera espontánea, probablemente quedaría mal indizado, esto es, mal representado y no aparecer (o aparecer en posiciones muy retrasadas para la palabra que resulte *clave* para nosotros). Por tanto la conclusión es clara: las tareas de diseño y gestión de un sitio web deben incluir aspectos de posicionamiento web.

Veamos ahora la segunda pareja. Es evidente que siempre ha existido la posibilidad de llevar a cabo un posicionamiento *fraudulento*. Un ejemplo grosero es la repetición de términos en los metadatos para obtener mejores puntuaciones (cosa que ha llevado a los motores a ignorar los metadatos), así como otras prácticas muy variadas que no comentaremos aquí. Es obvio que los propietarios de los motores están preocupados con el posicionamiento debido a esta posibilidad de fraude. Intentan detectarlo y constantemente revisan sus algoritmos. No obstante, creemos que debe existir (y de hecho existe y seguramente es el más frecuente) un posicionamiento honesto o *ético*. En el posicionamiento *ético* solamente se toman medidas que ayuden a los motores a indizar mejor el sitio web. Es decir, el responsable del posicionamiento es consciente del modo en que *piensan* los motores de búsqueda y optimi-

za su sitio para que los motores lo *lean* de la forma más adecuada posible.

¿Dónde está la frontera? Existe una regla muy clara y muy fácil de enunciar: cualquier medida de optimización de la página destinada solamente a los motores de búsqueda es, probablemente, posicionamiento fraudulento.

La razón es simple: si deseamos que nuestro sitio alcance un buen posicionamiento por el término, por ejemplo, “biblioteconomía y documentación”, una de dos, o lo llenamos de contenidos de calidad sobre ese tema, o renunciamos a ese posicionamiento. La tercera posibilidad sería simple fraude: intentar que alcance buenas posiciones, por ejemplo, a través de mentir en los metadatos y a través de conseguir muchos enlaces, pero sin que el sitio contenga auténtica información de calidad sobre el tema.

No está de más señalar que existen otros factores, difícilmente planificables a lo largo de la duración típica de una campaña de posicionamiento, que tienen indudable influencia en el posicionamiento de un sitio, como son las actualizaciones frecuentes del mismo, el mero resultado de una larga permanencia en la web, el mantenimiento del mismo dominio a lo largo del tiempo y una política de revisión y actualización frecuente de los enlaces.

¿Cuál es el papel de los profesionales y estudiosos de la documentación en este juego? Creemos que es doble. Por un lado, no podemos dejar de estudiar y de analizar cómo se producen los resultados en los motores de búsqueda, para criticar a los peores y ensalzar y recomendar los mejores. Por otro lado, estamos obligados a completar nuestra cartera de servicios en el mundo digital proponiendo medidas eficaces de posicionamiento a las instituciones en las que desarrollamos nuestra labor. Esto sí, no hace falta decirlo, siempre dentro de una ética rigurosa.

7. Referencias bibliográficas relacionadas

Abadal, Ernest. *Sistemas y servicios de información digital*. Gijón: Trea, 2001, 147 p.

Aguillo, Isidro. «Información científica en la web: retos y tareas para los documentalistas del siglo XXI». En: **Fuentes, M. E.** (dir.). *Anuario de biblioteconomía, documentación e información*. Barcelona: COBDC, 2001, pp. 33-50.

Azlor, S. (2003). *Posicionamiento en buscadores: guía básica*. Consultado en: 18-01-04.
<http://www.guia-buscadores.com/posicionamiento/>

Bruce Clay Inc. *Search engine optimization*. Consultado en: 18-01-04.
http://www.bruceclay.com/web_rank.htm

Berners-Lee, T.; Hendlar, J.; Lassila, O. «The semantic Web: a new form of Web content that is meaningful to computers will unleash a revolution of new possibilities». En: *Scientific American*, May 2001. Consultado en: 18-01-04.
<http://www.sciam.com>

Calishain, T.; Dornfest, R. *Google: los mejores trucos*. Madrid: Anaya, 2004, 415 p.

Codina, Lluís. «Web semántica: una mirada crítica». En: *El profesional de la Información*, 2003. n.º 12, vol 2, pp. 149-152.

Codina, Lluís. «Posicionamiento web: conceptos y ciclo de vida». En: *Anuario Hipertext.net*, mayo 2004. Consultado en: 18-01-04.
<http://www.hipertext.net>

Gonzalo, Carlos. «La selección de palabras clave para el posicionamiento en buscadores: conceptos y herramientas de estudio» En *Anuario Hipertext.net* Mayo 2004. Consultado en: 18-01-04.
<http://www.hipertext.net>

Kent, Peter. *Search engine optimization for dummies*. Wiley, 2004, 354 p.

Nobles, Robin; O'Neil, Susan. *Maximize web site traffic*. Holbrook: Adams, 2000, 417 p.

Sherman, Chris. «The future of web search». En: *Online*, v. 23, n. 3, May/June 1999, p. 54-61

Sherman, Chris. «The future revisited: what's new with web search». En: *Online*, May 2000. Consultado en: 18-01-04.
<http://www.onlineinc.com/onlinemag/OL2000/sherman5>

Thurrow, Shari. *Search engine visibility*. Indianapolis: New Riders, 2003, 297 p.

Tramullas, Jesús; Olvera, M. Dolores. *Recuperación de la información en internet*. Madrid: Ra-Ma, 2001, 232 p.

Webs y blogs:

XeoWebLog
<http://xeoweb.bitacorras.com/>

Search Engine RoundTable
<http://www.seroundtable.com/>

Seo Book, blog de **Aaron Wall**
<http://www.seobook.com/>

Search Engine Journal
<http://www.searchenginejournal.com/>

Search Engine Low Down, de **Andy Beal**
<http://www.searchenginelowdown.com/>

Search Engine News
<http://www.searchvisibilityreport.com/>

Cre8asite Blog
<http://blog.cre8asite.net/>

Search Engine Watch
<http://www.searchenginewatch.com/>

Foros:

Foro Seo
<http://www.xeoweb.com/foro/>

Foro Buscadores
<http://www.forobuscadores.com/foros/>

Internet Marketing Research Forums
<http://www.internet-marketing-research.net/forums/>

High Rankings Forums
<http://www.highrankings.com/forum/>

WebWorkShop Forums
<http://webworkshop.net/seoforum/>

Lluís Codina y Mari Carmen Marcos, Sección de Documentación, Universitat Pompeu Fabra; Grupo DigiDoc, Instituto de Lingüística Aplicada, Universitat Pompeu Fabra.

lluis.codina@upf.edu
mcarmen.marcos@upf.edu