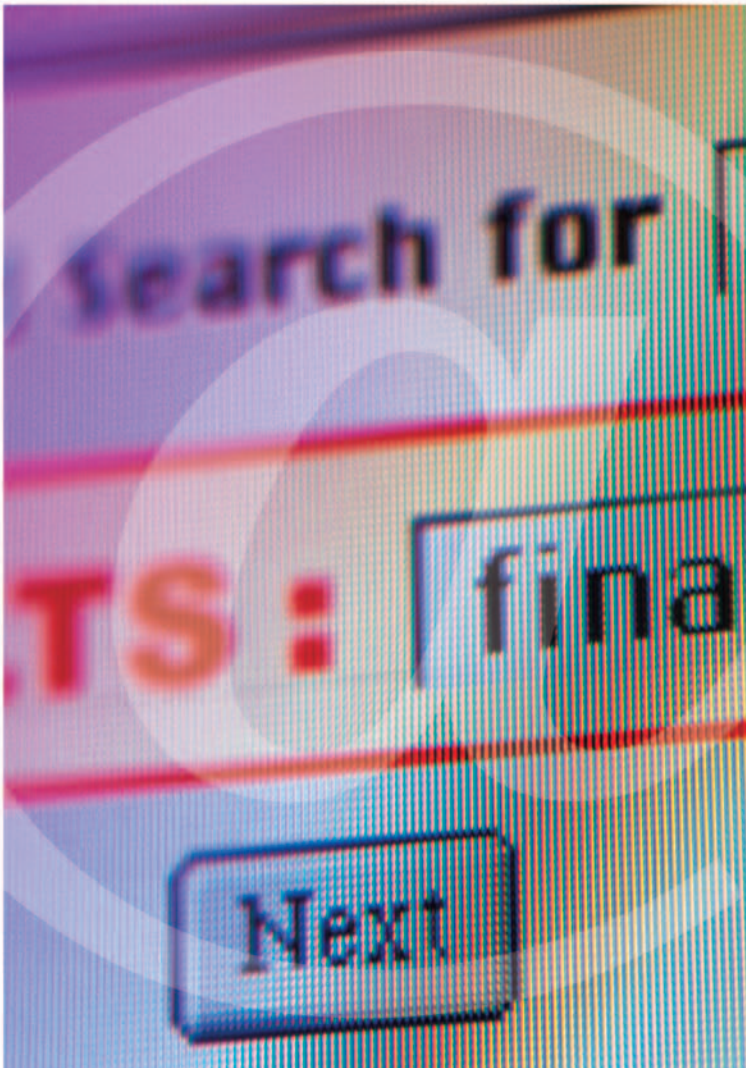
Marketing rentable: Buscadores



anetcom

Marketing rentable: Buscadores

anetcom

Edita:
Anetcom

Creación de contenidos:
Antonio Más
Search Media

Coordinación:
José Luis Colvée

Revisión:
Inmaculada Elum
Creu Ibáñez

Diseño y composición:
Integral Comunicación

Impresión:
Gráficas Marí Montañana

Depósito legal:
V-3667-2006

Índice

Prólogo 7

1. El mercado de los buscadores 11
- 1.1. De buscadores a portales a buscadores 11
- 1.2. En España 12
- 1.3. Comportamientos 13
- 1.4. Visión general de GYM (Google, Yahoo, MSN) 15
- 1.5. Diferencia entre posicionamiento natural y publicitario 16

2. Posicionamiento natural 19
- 2.1. Funcionamiento de un buscador 19
  - 2.1.1. Spiders y el proceso de indexación 20
  - 2.1.2. Detalle del algoritmo de Google. Resumen de su patente 21
  - 2.1.3. Criterios de resultados: Google PageRank, Yahoo WebRank o el concepto de relevancia basado en las citaciones o enlaces 26
- 2.2. Conseguir ser indexado 27
  - 2.2.1. La organización interna del sitio 28
  - 2.2.2. Criterios de organización de directorios 28
  - 2.2.3. Dos usuarios: el lector y el spider 29
  - 2.2.4. Enlaces y textos de enlaces (anchor text) 31
  - 2.2.5. Páginas estáticas vs páginas dinámicas 32
  - 2.2.6. Concepto de densidad de sitio 34
  - 2.2.7. Indexando bases de datos 34
  - 2.2.8. Creando mapas de sitio 35
- 2.3. Conseguir ser presentado 38
  - 2.3.1. Relevancia del resultado 38
  - 2.3.2. Presentación del contenido. La importancia de todos los detalles 39
    - a) Título, descripción, keywords y otros metas 39
    - b) Densidad de página y Densidad de keyword 41

c) Utilización de etiquetas de clasificación y énfasis **43**
d) Equilibrio código/contenido. Maquetación con capas y css **46**
e) Alts en imágenes **46**
f) Links internos **47**
g) Relevancia de contenido **48**
2.3.3. Eligiendo palabras objetivo **50**
a) Herramientas de sugerencias de palabras clave **50**
2.4. El concepto de popularidad **51**
2.4.1. La importancia de los enlaces entrantes **52**
2.4.2. Presencia en directorios **53**
2.4.3. Técnicas para lograr ser enlazado **54**
2.4.4. La importancia de las palabras de enlace **55**
2.5. Normas de calidad de contenido **56**
2.5.1. Normas de calidad de contenido de Google y Yahoo **56**
2.6. Errores más comunes **59**
2.6.1. Frames **59**
2.6.2. Redirects **60**
2.6.3. Enlaces Javascript y Flash **60**
2.6.4. Cloacking **61**
2.6.5. Textos ocultos **62**
2.6.6. Sobreoptimización y Contenido duplicado **62**

3. Posicionamiento publicitario o pago por click **65**
3.1. Publicidad basada en resultados **65**
3.2. Breve revisión de los sistemas de apuestas de AdWords y Yahoo Search marketing **66**

4. El futuro de los buscadores **69**
4.1. Especialización **70**
4.2. Buscadores sociales: Digg, Meneame, etc. **70**
4.3. La importancia creciente de los Blogs: Technorati **71**
4.4. Etiquetas o Tags. Búsquedas basadas en experiencia de usuario. El ejemplo español de Gennio.com. **72**

5. El mundo editorial y los buscadores **75**
5.1. La relación entre editores y buscadores **75**
5.2. El ejemplo de 20minutos.es **76**

Prólogo

Durante mucho tiempo las empresas de todo el mundo han venido practicando el marketing, apoyándose en acciones de promoción y publicidad que dirigían a grupos cuya definición resultaba harto difícil en muchos casos, y no digamos la capacidad de predeterminar qué personas responderían positivamente a dichas acciones. La consecuencia era que no se aprovechaban debidamente los recursos disponibles.

Con esto, no pretendo descalificar ni a las empresas, ni a los profesionales del marketing. Existen innumerables casos de iniciativas geniales y campañas coronadas con éxitos rotundos.

Lo que quiero constatar es que las empresas operan con las herramientas propias de cada época y las técnicas de marketing, si bien han conocido un desarrollo espectacular en el último medio siglo, tenían mucho donde avanzar. El hecho de dirigirse a un sólo target o grupo diferenciado es un paso que se ha ido alcanzando en las últimas décadas.

En Marketing digital avanzado, Antonio Más aborda la "revolución", como él la califica, que se está abriendo para el mundo de las empresas, con independencia de su tamaño y actividad. Este último aspecto cobra enorme importancia porque tradicionalmente las campañas de marketing estuvieron reservadas a las grandes empresas.

El gran público, hoy como antes, o mejor, hoy mucho más que antes, es muy heterogéneo. La diferencia de hoy respecto de ayer es que las tecnologías digitales ponen a nuestro alcance una serie de herramientas de acción y precisión que harían las delicias de los gurús del marketing de las últimas décadas. ¿Es posible captar uno, cien, mil o un millón de clientes, de forma consecutiva, incluso simultánea y conocer en tiempo real dónde, cómo y cuándo deciden comprar tal y cual producto o servicio? La respuesta es sí. Con el marketing digital avanzado.

En el nuevo marketing ya no hay lugar a la intuición. El marketing digital avanzado se rige por criterios mucho más precisos. Manda el criterio inapelable de la rentabilidad, porque el nuevo marketing permite conocer, en cada acción que emprende una empresa, qué clientes responden positivamente a su iniciativa y cuál es el retorno de la inversión de esa iniciativa concreta.

Google está desempeñando un papel estelar en este nuevo escenario. Por un lado, ayudando a ordenar la información en la red para facilitar su uso y, por otro, acercando a las empresas que quieren hacer negocios en y con la red, a los millones de usuarios satisfechos con sus excelentes productos de acceso gratuito.

Si nos circunscribimos a España, el potencial en la red es ingente y comparable a la enorme oportunidad de Google en nuestro mercado. Tres razones avalan esta afirmación. De una parte, España es, hoy por hoy, el mercado en el que Google tiene un más alto porcentaje de mercado entre buscadores, alrededor de un 97%. Lógicamente este porcentaje puede, y diría que debe, bajar, pero a la vez es el país en donde la facturación de Google por ingresos publicitarios está más descompensada con esa cuota de mercado. Digamos que es donde cada punto de porcentaje de mercado representa una menor facturación, no ya en Europa, sino en todo el mundo.

Por otro lado, recientes estudios y también nuestra propia experiencia personal, demuestran que el alcance de la audiencia target más codiciada entre los medios está cambiando de hábitos rápidamente. En la actualidad, con ligeras variaciones de un país a otro, la TV tiene aproximadamente un 30%, internet otro 30% y el otros medios como prensa escrita y radio se reparten el resto.

Basta observar en nuestras propias familias cómo está cambiando nuestro comportamiento: cuántas horas a la semana dedicamos nosotros y nuestros hijos a la radio, la prensa, la TV o a internet. Ya no tenemos tiempo de plantarnos como setas delante de la TV o ya no nos merece la pena, por las razones que sea, y ,sin embargo, como ese codiciado target, poder adquisitivo medio, medio-alto, de 25 a 45 años etc. no dejamos de navegar y de usar, por diferentes razones, de manera creciente y gracias al servicio de Google, la red en nuestro quehacer diario y en nuestro ocio.

Pues bien, aceptando sólo ese redondeo a la baja de un 30% de los medios, sabemos que sólo se dedica a la publicidad en internet, el negocio de Google, un 2% del presupuesto de marketing de las grandes compañías… ¡qué gran coste oportunidad!, porque, por otro lado, el retorno a la inversión en internet se puede medir, rastrear y corregir, y demuestra ser en todos los casos, reconocido por los propios clientes, altamente satisfactorio. Imagínense, si acercáramos ese 2% al 30% cuánto se incrementaría la eficiencia del marketing en nuestro país, por no añadir el potencial crecimiento del comercio electrónico.

El último factor que me hace ser optimista con respecto a la aceleración que nuestro país va a realizar más pronto que tarde con respecto al comercio electrónico es la apuesta definitiva por nuestro mercado de todas las empresas de la red. No sólo Google ha decidido apostar fuertemente por el mercado español, triplicando su presencia, sino también nuestra competencia, Microsoft, sacando mejores productos —sí, no lo lamento, creo que una buena competencia es una oportunidad para los usuarios y para que todos nos esforcemos al máximo—. Igualmente, Telefónica, los otros operadores y los proveedores de equipamiento haciendo ofertas cada vez más asequibles, y todo el tejido empresarial, oyendo y queriendo oír la propuesta de valor que Google puede ofrecer a sus negocios.

Creo que en este momento estamos en el buen camino. Que Internet y los negocios a su alrededor (publicidad on line, comercio electrónico, administración electrónica), y todos los españoles, estamos de enhorabuena en la red.

Dª. Isabel Aguilera
Directora General de Google España y Portugal

1. El mercado de los buscadores

◆ 1.1. De buscadores a portales a buscadores

En 1995 dos jóvenes estadounidenses se dedicaban a clasificar temáticamente y de manera manual direcciones de Internet en una universidad americana. Su trabajo, terminó siendo tan popular que pasó de ser un entretenimiento, al germen de una de las mayores empresas de comunicación del planeta.

Nacía Yahoo!

Y con ella uno de los grandes desafíos de Internet: clasificar toda la información disponible en la red de manera sencilla, relevante y eficaz.

Con Yahoo! surgieron una serie de buscadores, tanto basados en sus propios directorios – listados de webs organizados manualmente en categorías preconcebidas- como en algoritmos de búsqueda basados en los contenidos de las páginas web sin tener en consideración su relación con el resto de Internet.

Estos buscadores se convirtieron en los webs más populares de la red y pronto perdieron el foco de su principal negocio, empezando a ampliar servicios de todo tipo: webmail, noticias, foros, listas de correo, espacio web, mensajería instantánea, centros comerciales… Nacía así el concepto de portal, un lugar de entrada a Internet donde enganchar al usuario y tratar de ofrecerle todo lo que podía esperar de Internet.

Pero en esta obsesión olvidaron que la búsqueda de información era el segundo uso de Internet después del correo electrónico y la facilidad de manipular los algoritmos, al estar basados sólo en los contenidos, facilitaron la aparición de spammers o webmasters especializados en posicionar –a cualquier precio– sus webs en las primeras posiciones, aunque éstos no resultasen rele-

vantes para las búsquedas, creando una gran frustración en el usuario. Eran los tiempos de los keywords, las palabras repetidas, la abundancia de contenido duplicado…

Y esta falta de relevancia en los resultados llevó a cierta decadencia en el uso de los buscadores hasta que otra vez dos jóvenes americanos crearon un nuevo estilo de buscador que se centraba en una sola acción, la búsqueda, y que tomaba del mundo científico uno de los conceptos que hacían más importante un documento o ensayo: la citación o referencia por otros científicos. Cuantas más veces era citado o referenciado un documento en el mundo científico, más importante era su autor y dicho documento.

Era 1998. Y nacía Google.

Sin marketing. En silencio. Con un crecimiento orgánico basado en la experiencia de sus usuarios, Google empezó la más importante revolución de Internet que le ha llevado a convertirse en la mayor empresa de medios de comunicación del mundo y conseguir que su nombre sea sinónimo de Internet.

Mientras, en el camino quedaban Excite, Altavista, HotBot, Ask y muchos más que se compraron, vendieron, fusionaron y cambiaron sin mejorar mucho su posición. Sólo MSN gracias a la base de usuarios de potencia de Microsoft y a su interminable tesorería aguantó, junto a Yahoo! el arrollador empuje de Google.

Y esa es la situación de nuestros días. Un triunvirato donde hay un rey claro y dos aspirantes a constantes que intentan mantener el ritmo innovador y defenderse de la voracidad creciente de Google.

Así, hoy, en 2006, la guerra de los buscadores sólo tiene 3 jugadores: Google, Yahoo! y MSN.

◆ 1.2. En España

La pelea patria en todo esto fue mucho más modesta y se limitó a un jugador y varios intentos sin éxito. Era un jugador basado en el modelo directorio de

Yahoo! se llamó Ole! y fue adquirido por Terra siendo la base del portal que les lanzaría a Bolsa dando comienzo a nuestra particular burbuja.

Terra, Eresmas o Ya.com lanzaron su propio buscador que eliminaron en poco tiempo rindiéndose a la evidencia y tomando las herramientas de Google o Fast. Otros como Ozu o Hispavista también siguieron parecida trayectoria con menor notoriedad y penetración que la adquirida por Terra.

No hubo sorpresas en el mercado nacional y las iniciativas fueron portales que asumieron su posición como espectadores que ganaban gracias a su condición de proveedores de acceso gratuito primero y ADSL después.

La particularidad de nuestro mercado se basa en la elevada penetración de Google donde es el nº1 mundial y el reducido número de usuarios que utilizan un segundo buscador.

	Audiencia		Páginas vistas	
	2005	2004	2005	2004
Google	11.149	7.475	3.123.406	1.506.095
MSN Search	4.391	2.771	69.508	31.444
Yahoo!	2.313	1.402	57.773	31.388
Wanadoo	1.008	562	16.787	5.920

Fuente: Oct-Dic 2005 Nielsen

◆ 1.3. Comportamientos

Es importante hacer alguna puntualización sobre algunos comportamientos de usuario y su forma de interactuar con los buscadores. Y estos dos comportamientos definen cuán estratégico es aparecer en los primeros resultados de búsqueda:

La atracción del primer resultado

La mayoría de usuarios, el 42%, hace click en el primer resultado de los resultados de búsqueda y sólo el 8% lo hace en el segundo resultado. (De hecho, se realizó un experimento en el que se intercambió el primer resultado por el segundo y el 32% de los usuarios seguía haciendo click en el primer resultado. Según los usuarios, éstos afirman que:

- Los buscadores son tan buenos juzgando la relevancia que casi siempre colocan al mejor en el primer puesto.
- Los usuarios son vagos. Hacen click en el primer resultado no porque hay alguno mejor, sino simplemente porque está primero. Además consideran que el primer resultado es el más relevante el 36% de las veces, el segundo el 24% de las ocasiones y en un 40% de las ocasiones los juzgaron como parejos. El usuario piensa entonces que los buscadores no fallan.

Los usuarios no pasan de la primera página

El 62% no pasa de la primera página. Estás en ella o no estás. Y el 90% de los usuarios no pasa de las 3 primeras páginas.

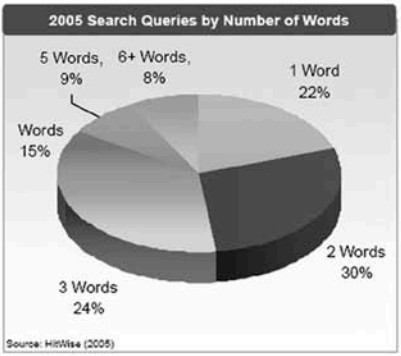
Y además es una tendencia creciente:

Año	2002	2004	2006
Sólo la página 1	48%	60%	62%
Hasta la página 3	81%	87%	90%

Conclusión: los usuarios son más vagos y, por otro lado, consideran que los buscadores son más eficaces.

Número de palabras usadas en la búsqueda

Las búsquedas se hacen cada vez más sofisticadas y complejas. La larga cola tiene muchas más búsquedas que las palabras que podemos considerar "claves".



◆ 1.4. Visión general de GYM (Google, Yahoo, MSN-Live)

Así, con esta idea general del escenario, trataremos de dar una visión general de los 3 grandes jugadores actuales.

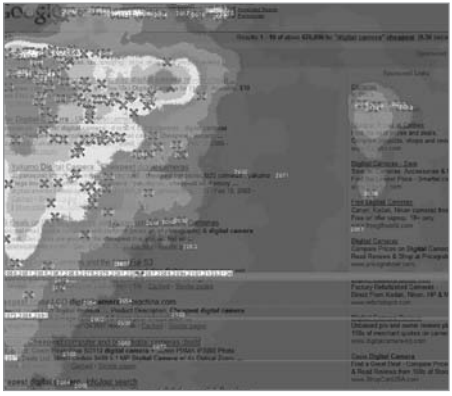
Google
Alcanza el 92% de las búsquedas del mercado nacional y su relevancia no para de crecer.

Con su fuerza es objeto de constante estudio e imitación por sus competidores que han copiado la forma de presentación de sus resultados y la base de su algoritmo centrada en la popularidad.

Es importante conocer detalles sobre su uso como cuáles son sus zonas calientes visuales donde queda clara la importancia de los 3 primeros resultados y la ausencia de diferenciación entre resultados patrocinados y naturales por parte del usuario.

En la actualidad es el motor de búsqueda de Terra, Ya.com, Wanadoo o el mundo.es, lo que le da una capilaridad total.

Bajo el dominio .es se realizan el 98% de sus búsquedas en España.



Mapa visual de Google obtenido de un estudio de eyetracking.
Esta figura es lo que se llama Triángulo de Oro.

Yahoo.es
La tecnología mejor integrada y aquel donde los usuarios pasan más tiempo. Fue el buscador líder y, en su obsesión por lanzar nuevos y más servicios y convertirse en portal, perdió esta posición.

Hoy es uno de los 3 grandes jugadores que está haciendo una importante apuesta por retomar el liderazgo y que está integrando servicios de gran popularidad entre los usuarios al adquirir empresas como Kelkoo, Flirck o del.icio.us y su gran apuesta al comprar en su momento Overture, el sistema de enlaces patrocinados más importante después de AdWords.

MSN.es
Pasará a llamarse Live.es en breve y están realizando un gran esfuerzo para ponerse al día y ganar competitividad en este mercado. De hecho, parte de la estrategia del propio Microsoft pasará por el buscador que aglutinará el uso de otros servicios en línea como Office.

Alguno de sus servicios como Hotmail o Messenger son líderes de manera aplastante sobre cualquiera de los otros jugadores, pero no tienen la posibilidad de monetizarlos de la misma manera, ni parecen ser capaces de llevar a sus usuarios hacia su buscador.

En los primeros días de mayo de 2006 existían rumores sobre la posibilidad de que MSN compre Yahoo! como única forma de tratar de hacer frente al poder creciente de Google.

◆ 1.5. Diferencia entre posicionamiento natural y publicitario

Un punto obvio que no obstante tratamos rápidamente. Pero primero recordemos la actitud de un usuario que busca, que es la que le da un enorme valor publicitario.

Una persona que busca está haciendo una segmentación de sus intenciones que supera la de cualquier otro medio. No hay criterios sociodemográficos o económicos, ni cobertura, ni target… Si busco "Alquiler de coches en Alicante" puedo ser un conductor de Burgos de 25 años que

quiere un fin de semana o uno de Madrid de 70 años para una semana completa. Da igual el perfil. Lo que importa es la necesidad.

Y la actitud de dicho usuario es para resolver en ese momento una necesidad concreta. Justo en ese momento. Y la quiere resolver a un click de distancia.

Esos dos usuarios pueden ser lectores del mismo periódico, pero es en ese momento cuando quieren un "Alquiler de coches en Alicante". No cuando están leyendo el periódico, ni viendo la TV u oyendo la radio. En ese momento tienen la actitud más positiva que un anunciante puede desear.

Por esto el mercado de los enlaces patrocinados crece de esta manera. Un visitante que ha hecho click en un resultado de búsqueda cuya palabra hemos elegido por considerarla relevante para nuestro negocio y por la que hemos decidido cuánto queremos pagar, será una visita que pase de 6 a 8 minutos en nuestra página en el momento en que andaba buscando nuestro servicio en un entorno diseñado por nuestra empresa donde le presentamos exactamente aquello que andaba buscando.

Sin duda es el impacto publicitario de más calidad que existe en la actualidad.

Y lo sorprendente de todo esto es que la posición en la pantalla por la que estamos dispuestos a pagar está justo al lado de otra igual de relevante que no pagará nada.

El mismo usuario con la misma actitud puede ser captado pagando o totalmente gratis.

Los resultados naturales son aquellos que los buscadores presentan en respuesta a una búsqueda del usuario. Acorde con su algoritmo y el país del usuario, entregan una lista de web sites que consideran que serán relevantes y por tanto útiles para la búsqueda realizada.

Y a la vez, y bajo la etiqueta de "Enlaces patrocinados" presentan aquellas empresas que están dispuestas a pagar por un visitante que ha realizado esa búsqueda.

La clave para el buscador es que, tanto uno como otro resultado, sea eficaz y relevante para su usuario. La respuesta que obtenga debe ser eficaz y garantizar la satisfacción para que la próxima búsqueda se haga desde esa misma página.

La línea que distingue el espacio pagado o gratuito es hoy más delgada que nunca y tanto los espacios de pago como gratuitos son limitados.

Y visualmente iguales.

2. Posicionamiento natural

◆ 2.1. Funcionamiento de un buscador

Un buscador almacena una base de datos de páginas web con una serie de datos asociados a las mismas y que responde a una búsqueda de un usuario con un conjunto de resultados obtenidos de aplicar un algoritmo de búsqueda al criterio solicitado por el usuario.

Eso es una definición directa de lo que ocurre cada vez que alguien hace click en "Buscar". El buscador dice: "Para una determinada palabra estos son mis resultados acorde con lo que entiendo que es más relevante".

Y para llegar a este punto ha debido haber dos procesos paralelos:

- Una definición de un algoritmo de búsqueda con decenas de posibles criterios.
- La creación de una base de datos a los que aplicar dicho algoritmo para una determinada búsqueda.

Es importante anotar la diferenciación entre buscador y directorio.

Un directorio es un conjunto de páginas ordenadas normalmente por humanos en categorías y subcategorías. En ellos el usuario busca navegando a través de las mismas o a través de un buscador que sólo presenta los resultados de los webs incluidos en el mismo. Por ejemplo, dmoz.org o el directorio de Yahoo.com.

En el otro lado, el buscador, aplica su algoritmo de búsqueda contra su base de datos.

2.1.1. Spiders y el proceso de indexación

Para construir sus bases de datos los buscadores han desarrollado arañas o spiders. Se trata de programas que navegan por Internet siguiendo enlaces y almacenando los contenidos que van encontrando.

Los spiders aprovechan la propia estructura entrelazada de Internet para descubrir sites y dentro de cada site su propia estructura de directorios para sacar a la luz su contenido. Visualmente sería como una serpiente que se va dividiendo en dos cada vez que encuentra un nuevo camino. Un organigrama inacabable que salta de un sitio a otro hasta encontrar todo aquello que éste le permite descubrir.

Así, suelte un spider en Yahoo.com y seguirá cada enlace de su directorio. Y dentro de cada site que descubra seguirá de nuevo cada uno de sus enlaces, tanto internos como externos y así repitiendo ese ciclo constantemente.

Esos spiders nos visitarán con diferente frecuencia dependiendo de las características de nuestro site. Si el buscador percibe que actualizamos contenidos varias veces al día – por ejemplo, un periódico – , posiblemente nos visite con esa frecuencia, mientras que si lo hacemos cada semana – por ejemplo un catálogo o tienda online – lo hará con una frecuencia similar, sin que ello indique que uno es mejor que otro.

Con esa visita, los spiders almacenan el contenido de nuestro web site, que alimentará los resultados de búsqueda posteriores.

De aquí surge uno de los puntos básicos de la optimización.

Si no está almacenado en la base de datos del buscador, no podrá ser presentado como resultado.

"Si no estás indexado no serás presentado"

Para poder aparecer en los resultados de búsqueda lo primero es conseguir ser indexado. Esto es, lograr que nuestro contenido esté almacenado dentro de la base de datos del buscador.

La primera obsesión que deberemos tener será mostrar todo nuestro contenido a los buscadores, facilitando la labor de los spiders, abriéndoles nuestras puertas y dándoles un listado de enlaces que seguir.

Para ello, hay que mostrar todo nuestro contenido a través de enlaces de texto en un conjunto de directorios bien organizado y con formatos de páginas que el buscador pueda leer y almacenar.

¿Cuál es la mejor manera de abrir nuestro contenido y garantizar que nos indexan?

- Estar enlazados. Saber que llega a nosotros a través de otros enlaces.
- Tener páginas de contenido estático.
- Crear un mapa con todo nuestro contenido pensado en el usuario y, que a la vez, sea útil para que el spider siga los enlaces. Un modelo basado en organización alfabética, temática o por fechas es el ideal. Los blogs consiguen buenos resultados en buscadores, entre otras cuestiones, por la facilidad con la que su contenido es indexado al estar clasificados por categorías y fechas y a la vez tener profusión de enlaces internos y externos. Como ejemplos de mapas: Kelkoo o Idealista que vemos más adelante.
- Crear un Google SiteMaps.

¿Cómo comprobar si nuestro contenido está indexado?

- Hacer la búsqueda con el comando site. Por ejemplo site:www.laverdad.es nos dirá el número de páginas que tiene indexadas de dicho dominio.
- Comprobar los logs con las visitas de GoogleBot o Yahoo Slurp, los nombres de los dos spiders de ambos buscadores. Así comprobaremos además la frecuencia de visita de los mismos.
- Chequear nuestra cuenta de Google SiteMaps.

2.1.2. Detalle del algoritmo de Google. Resumen de su patente

Google hizo una solicitud de registro de patente (http://appft1.uspto.gov/netacgi/nphParser?Sect1=PTO2&Sect2=HITOFF&p=1&u=%2Fnetahtml%2FPTO%2Fsearch.bool.html&r=1&f=G&l=50&co1=AND&d=PG01&s1=20050071741&OS=20050071741&RS=20050071741) donde se encuentran parte de sus conceptos clave de funcionamiento.

Es un documento fundamental que, leído con atención, permite conocer las bases sobre las que se asienta su funcionamiento. Por buscar un símil, sería como conocer los ingredientes de una receta, pero no el orden ni las cantidades.

Voy a tratar de hacer un resumen amplio que permita su comprensión agrupada por los diferentes conceptos.

Google se define como un sistema que identifica un documento o conjunto de documentos (un sitio web) y almacena datos históricos del mismo para generar una valoración basada en una o más partes de dichos datos conjuntamente con el contenido del mismo.

La definición que facilita de documento es muy amplia:

"Un documento, como término usado de ahora en adelante, debe ser interpretado como cualquier contenido legible y almacenable por un ordenador. Un documento puede ser un e-mail, un sitio web, un archivo, una combinación de archivos, uno o más archivos que incluyan links a otros archivos, un mensaje en un grupo de noticias, un blog, una web publicitaria, etc. En el contexto de Internet, el documento más común es una página web. Una página puede ser un documento o sólo una parte del mismo. Por tanto, las palabras "página" y "documento" pueden ser intercambiadas en algunos casos."

Los criterios que utiliza para valorar dicho documento – o lo que es lo mismo, asignarle un pagerank o relevancia – pueden ser agrupados en varios conceptos:

Tiempo

- La fecha de inicio del documento (esto es, cuándo se indexó por vez primera o cuándo se encontró un enlace hacia dicho documento).
- El promedio de las fechas de inicio de todos los documentos indexados en su base de datos. Contra éste se medirá cualquiera de los documentos. (Digamos que si un documento tiene 2 años, pero la media de los documentos de Google tiene 4, tendrá un valor u otro).
- La antigüedad de un dominio y la fecha de expiración del mismo. Y sobre este punto debo traducir literalmente este párrafo:

"Ciertas señales pueden ser usadas para distinguir entre dominios legítimos o ilegítimos. Por ejemplo, los dominios pueden ser renovados hasta por 10 años. Los dominios con valor o legítimos suelen pagar varios años por adelantado, mientras que los dominios usados como "doorways" – puertas de entrada – rara vez son usados por más de un año. Por tanto, la fecha en la que un dominio expira puede ser usada como un factor para predecir la legitimidad del mismo y, por tanto, de los documentos o páginas que contienen".

Actualización de los contenidos

- El cambio o actualización de contenidos del documento, incluyendo diferentes variables sobre:
  - Cuántas veces cambia.
  - Promedio de cambios periódicos, esto es, con qué frecuencia.
  - Comparación de los ratios de cambios sobre periodos previos (quiero entender que un periódico tendrá una dinámica de cambio diferente a un catálogo y que lo considera al valorar dicho web site. No por que un periódico cambie más puede ser considerado más importante).
  - Cuánta parte del contenido cambia.
  - Cuánto cambian los documentos relacionados (o que enlazan) al mismo.
  - Cada cuanto cambian los documentos enlazados
- El deterioro del documento a través del paso del tiempo (imagino enlaces rotos, imágenes rotas, etc.) implica que la novedad del documento es menor.
- Se marca la diferencia entre contenidos que al ser cambiados, tienen poco peso o son completamente ignorados – como Javascript, comentarios de código, publicidad, elementos de navegación, líneas de hora o fecha... – con respecto a los contenidos realmente relevantes y útiles para el usuario.

Resultados de búsqueda

- Cuántas veces el documento es elegido cuando es presentado en un resultado de búsqueda (esto es, el ratio de click que presenta una página sobre una búsqueda concreta).
- Si la elección es continuada dentro de un periodo de tiempo (si ese ratio de click permanece en el tiempo parece ser considerado positivo).

• Comparación de dicha selección con los otros documentos, aportando una valoración superior o inferior acorde con el promedio.
• La asociación entre los términos de búsqueda cuya frecuencia de búsqueda se incrementa con el tiempo y los que se encuentran en el documento (esto es, si un término se busca cada vez más o menos o lo que es lo mismo, la tendencia de dicha búsqueda por tratarse de un tema más o menos actual).

Enlaces (PageRank)

• El comportamiento de los enlaces que tiene un documento desde Internet a lo largo del tiempo (parece indicar que un enlace puede que no sea válido hasta que pase un plazo determinado), así como su aparición y desaparición.
• La novedad de los enlaces genera una valoración que incluye:
  - Determinar la novedad de los enlaces (esto es, cuando es un enlace nuevo o viejo).
  - Asignar pesos a dichos enlaces basándose en la novedad y valorar el documento acorde con el peso asignado a los links. En resumen: nuevos enlaces aportan importancia al documento y cada nuevo enlace tiene un peso diferente. Cuantos más enlaces, más importantes y más actuales mejor.
• La novedad de un enlace está asociada a la fecha en la que éste aparece o cambia el texto que utiliza.
• El peso de un enlace se basa en la calidad y confianza de la fuente (web que enlaza), si el web es una "autoridad" en una materia y en la propia novedad de dicho enlace.
• La manera en cómo cambia en el tiempo el texto utilizado para enlazar.
• La asociación entre el texto utilizado para enlazar y la página o documento al que enlaza. Por ejemplo, si una página enlaza con la palabra "Nokia" debe llevar a otra con contenido de "Nokia" y no de "Siemens".
• La utilización del mismo texto de enlace desde diferentes procedencias. Igualmente considera si todas las palabras para enlazar son las mismas (puede hacerlo de manera positiva o negativa – ¿relevancia para un término o spam? –)
• La calidad de los dominios asociados al documento generando la valoración a partir de la asociación a lo largo del tiempo, la legitimidad de dichos dominios, la fecha de expiración del dominio y el histórico del servidor asociado al dominio.

- El incremento de uso en las palabras de enlace a lo largo de un periodo de tiempo.
- La duración de los enlaces penalizará el documento en caso de ser corta y lo apoyará si se prolonga, considerando además que, si el contenido de las páginas que enlazan se actualiza en el tiempo, será mejor valorado que si no lo hace. Esto es, un enlace que dura poco o mucho y si lo hace desde una página que cambia o no en el tiempo (lo que implica contenido fresco y preocupación por parte de los webmasters).

Dentro de este tema es destacable este párrafo que traduzco literalmente:

"Considérese el ejemplo de un documento que es dado de alta ayer y es referenciado con 10 backlinks. Este documento puede ser valorado más positivamente que uno dado de alta hace 10 años y que es referenciado por 100 backlinks por que el ratio de crecimiento de enlaces del nuevo es relativamente superior que el más antiguo. Mientras que este crecimiento de backlinks puede ser un factor usado para valorar el documento también puede ser una señal de spam de buscadores".

Tráfico y conducta del usuario

- Dentro de los datos históricos, Google almacena información relacionada con el tráfico recibido por un documento y las características del mismo, elaborando posteriormente patrones de tráfico y analizando los cambios durante periodos de tiempo.
- Igualmente almacena la conducta del usuario asociada al documento, incluyendo información relativa a cuántas veces es seleccionado el documento de un conjunto de resultados de búsqueda y cuánto tiempo están los usuarios dentro del mismo. En el fondo se trata de una medida de ratio de click pero sobre los resultados gratuitos. Cuantos más clicks para una búsqueda dada, mejor será el resultado.
- La valoración de un documento también se basa en la historia previa del ranking del mismo, incluyendo en la misma la cantidad y amplitud de los movimientos en el ranking, basando la historia de dicho ranking en la cantidad de veces que es elegido dentro de un conjunto de resultados de búsqueda a lo largo del tiempo o estacionalidad de dicha selección.

• La valoración de un documento se podrá basar en información generada por el usuario donde indique el grado de interés que le ha producido. Esta información puede provenir de favoritos, archivos temporales, caches o basarse en los sistemas de votos de la barra de Google. Esta valoración incluirá la relación o presencia de dicha información a lo largo del tiempo y cómo ésta cambia o se modifica.
• La relevancia del documento para una búsqueda dada generando una valoración de la relevancia para dicha búsqueda y la historia de dicha valoración a lo largo del tiempo.
• La valoración de un documento incluye su identificación, la obtención de sus diferentes datos históricos – incluyendo la trayectoria de su valoración, los enlaces entrantes, el texto y tiempo de los mismos, la actualización o renovación de dichos enlaces – y la combinación de los mismos para asignar el valor.

"Las bases del algoritmo como norma de actuación"

Creemos que dos factores pueden llevar al éxito de la indexación:
- Pensar en el usuario.
- Interpretar adecuadamente la solicitud de patente de Google. Su lectura entre líneas da las claves de su forma de operar.

La lectura de los puntos clave de la patente del algoritmo de Google revelan sus pautas de actuación con todo detalle.

2.1.3. Criterios de resultados: Google PageRank, YahooWebRank o el concepto de relevancia basado en las citaciones o enlaces

Como hemos visto en el punto anterior, el concepto de citación o popularidad marcó un antes y después. E hizo que tanto Yahoo como MSN (LIVE en los próximos meses) utilizasen la misma tecnología.

Y al igual que Google tenía su PageRank, Yahoo lanzó su WebRank que indica también la relevancia de cada web site. Existen herramientas para poder chequear dichas posiciones:

- Para Google: Utilizar la barra de Google o http://www.prchecker.info/
- Para Yahoo: http://www.site-build-it-solutions.com/yahoo-page-rank-web-rank.htm

La relevancia basada en enlaces consiste en el número de links que desde otras webs apuntan hacia otra y de qué manera lo hacen. Cada link es considerado un voto y cada voto tiene un peso diferente dependiendo de la calidad y de la autoridad del emisor sobre el tema que se está tratando.

Los conceptos clave de dicha relevancia ya aparecen indicados en la solicitud de patente del algoritmo de Google y, en general se resumen en:

- Número de enlaces que recibe desde otras páginas.
- Texto que utilizan los enlaces para referir a esas páginas.
- Calidad del sitio que enlaza.
- Autoridad del sitio que enlaza.
- Tema de la página que enlaza.
- Título de la página que enlaza.

"La relevancia de una página viene dada por su relación con otros web sites y las páginas de su propio web "

Una página es más relevante dependiendo del número de enlaces que recibe tanto internamente como externamente y de si los orígenes de dichos enlaces tienen calidad y relevancia en el tema que tratamos.

◆ 2.2. Conseguir ser indexado

"El buen paño en el arca no se vende"

O en otras palabras, si no es posible conocer el producto no se puede vender. O en lo que nos atañe, si no eres indexado no puedes ser presentado.

Ser indexado es el primer paso para ser presentado. Si no estás en la base de datos no puedes ser presentado, por lo que organizar el sitio para que éste sea almacenado en la base de datos del buscador es el primer objetivo que debemos conseguir.

Debemos pues recordar que tenemos dos tipos de lectores en nuestro web. El usuario y el spider.

2.2.1. La organización interna del sitio y de las páginas

Es sorprendente la gran relación que hay entre usabilidad del web site y la facilidad de indexación. Normalmente, si el web site está bien organizado para el usuario, suele estarlo para los spiders.

Esta organización debe de:

- Estar basada en tecnologías que puedan ser indexadas. Las páginas dinámicas con más de un parámetro en su URL (por ejemplo el parámetro de sesión de usuario) no suelen ser guardadas por los spiders.
- Pensar en la usabilidad o accesibilidad del usuario. Si la misma está clara, con menús que permiten navegar de manera categorizada, con enlaces directos y una adecuada clasificación en árbol, se cumplirá el doble objetivo de lograr una mejora para el usuario y conseguir que el spider encuentre todo nuestro contenido.
- La mejor manera de estructurar un web site es crear un árbol de contenidos clasificado temáticamente por cada una de sus secciones y servicios y enlazando en las diferentes secciones o subsecciones de manera jerárquica y ordenada.

2.2.2. Criterios de organización de directorios

Los directorios deben tener criterios humanos y no de programador.

Los programadores olvidan que esos directorios son presentados dentro de las URLs que aparecen en el navegador y que los nombres de los mismos y su organización pueden servir de gran ayuda tanto al usuario como al spider.

Un ejemplo dibuja perfectamente el caso. Estos dos enlaces llevan a la misma página. Mientras uno indica claramente una lógica de usuario y organización de la información, el otro son parámetros y organización propia de programador.

- http://www.movizonia.com/moviles/Nokia_6280/539.htm
- http://www.movizonia.com/product.asp?catgid=2&list=0&prodid=539&shop=1&typeproduct=1&lang=ES

El primero indica al usuario dónde está y en qué modelo. El segundo no indica nada.

El primero ofrece una clasificación lógica de información. El segundo no transmite nada más que una "visión de programador".

Al decidir la estructura de directorios debe de:

- Cuidarse el nombre que se elije pensando en el usuario y en el resultado final en la barra del navegador.
- Utilizar nombres de directorios que puedan encajar con las palabras clave sobre las que tratan los contenidos que en dicho directorio se presenta.
- Crear tantos subdirectorios como sean necesarios para que la información esté clasificada adecuadamente.

2.2.3. Dos usuarios: el lector y el spider

Lo hemos indicado anteriormente pero debo repetirlo. Cualquier decisión en la construcción de un site debe estar tomada pensando en los dos usuarios: lector y spider. Los lectores serán miles o millones, el spider son solo 3 ó 4, pero son la puerta de entrada de los anteriores.

Siempre existe un equilibrio para que sea operativo para ambos. Y programadores y responsables deben trabajar para ambos tipos de visitantes.

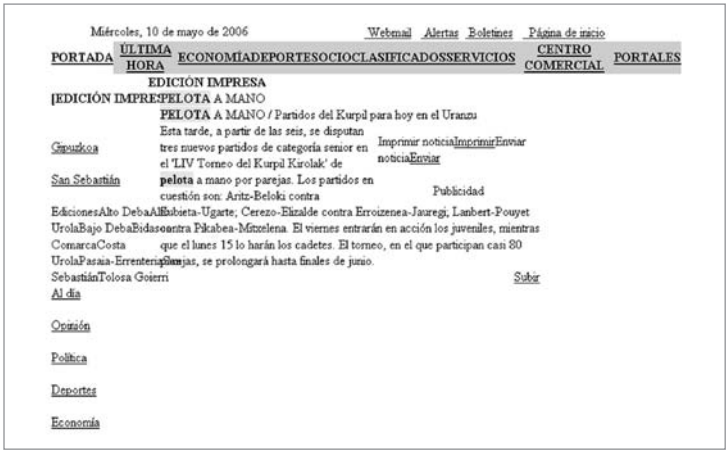
Para ello es imprescindible ver cómo nos siente o percibe el spider, el cual no percibe si somos bonitos, feos o usamos el flash más espectacular. El spider busca contenido de calidad, bien estructurado y referenciado por otros en sites con autoridad.

La mejor manera de analizar cómo nos "ve" es chequear la versión de texto del caché de Google.

Aquí está un ejemplo de ello.

Así es como está cacheada una página del diario vasco:

Y este es el caché de texto que realmente tiene de la misma página:



Este ejercicio es clave para descubrir la percepción de nuestra información del buscador, además de saber cuál es la frecuencia de actualización de nuestros contenidos.

2.2.4. Enlaces y textos de enlaces (anchor text)

Dentro de la fórmula mágica de Google parece que el anchor text, o texto con el que se enlaza hacia una determinada página, tiene un peso muy elevado. De hecho, tiene tal importancia que ha generado una actividad frecuente llamada Google Bombing que consiste en que muchos webs se pongan de acuerdo en enlazar a una página con la misma palabra para que ésta sea considerada relevante para la misma.

Ejemplos:

- "Ladrones" presenta en primer lugar la página de la SGAE.
- "Miserable Failure" presenta la biografía del presidente Bush (y en segundo lugar la página de Michael Moore).

Este texto del enlace o anchor text puede ser interno – con qué palabras se enlaza el site entre sus páginas – o externo – con qué palabras nos enlazan otros web sites –.

Así, es muy común el error de utilizar palabras como "Click aquí" o "Ver más" para ir a enlazar en vez de hacerlo con las palabras clave o utilizar la etiqueta "title" como texto alternativo en dicho enlace.

En los textos de enlace o anchor text los factores que influyen son los siguientes:

- El texto del enlace.
- El texto anterior y posterior al enlace.
- Si después del enlace hay otro enlace, o si lo hay antes, o si lo hay antes y después.
- Si el enlace está al final del código de la página (que le aporta menos valor).
- Si el enlace está en la cabecera (que le aporta más valor).
- Si el enlace procede de la misma IP (que implica que podemos estar haciendo ciertas trampas).
- Si viene de una página afiliada.
- Si el enlace es de una autoridad en nuestra temática (esto es, un web site especializado en el tema sobre el que se está enlazando).

- Si el enlace es simplemente de una página temática (esto es, una página que trata sobre el tema sobre el que se está enlazando).
- El título de la página que nos apunta (si coincide con la temática sobre la que trata el enlace).
- Si nos enlazan páginas que compiten por la misma búsqueda.
- Si el enlace viene de una página que está escrita en el mismo idioma (lo contrario puede considerarse extraño).
- Si hay más enlaces iguales desde el mismo dominio (lo que indica posibles trampas en dicho sentido).
- La extensión del dominio (si es genérico o territorial).

El peso de cada concepto no se conoce con detalle, pero es el cuidado de todos los detalles lo que consigue que nuestra política o estrategia de enlaces pueda ayudarnos a mejorar resultados

2.2.5. Páginas estáticas vs. páginas dinámicas

Por páginas estáticas entendemos aquellas que están construidas sin atacar bases de datos mientras que las dinámicas se ejecutan "al vuelo" dependiendo de las solicitudes del usuario o del gestor de contenidos que las administra.

Tanto Google como Yahoo indexan páginas dinámicas (que son construidas dependiendo de parámetros que consultan a bases de datos). Las páginas dinámicas son claves en cualquier gestión de contenido avanzado y 100% necesarias para poder construir web sites de cierta complejidad.

Ahora bien, es necesario hacer una serie de consideraciones sobre este tema:

- Los robots almacenan páginas dinámicas, pero siempre que éstas no tengan más de dos parámetros dentro de sus URLs y que no utilicen variables de sesión de usuario.
- Uno de los criterios que ayudan en la optimización es la inclusión de palabras clave dentro de las URLs, bien directorios o el propio nombre del archivo. Y esto es más difícil de conseguir en direcciones dinámicas.

Así, analicemos una prueba concreta. Si buscamos "Nokia 6600" veremos que los 6 primeros resultados incluyen la palabra dentro de la URL y que todas ellas no tienen ninguna variable.

Por tanto, el uso de páginas estáticas – o de camuflar las mismas como tal – es una decisión que debe considerarse seriamente.

Lógicamente, usar páginas estáticas es inaceptable en webs de la complejidad actual, pero sí se pueden emplear tecnologías que transforman dichas direcciones en estáticas logrando así los beneficios de ambos mundos.

Para ello existen aplicaciones de servidor que definen cambios de las variables por nombres de directorios construyendo nuevas direcciones más fáciles de interpretar.

Así presento algunos ejemplos de medios que ya usan técnicas de este tipo donde se encuentran todas las palabras clave en las URL:

http://www.elpais.es/articulo/portada/Xbox/PlayStation/entablan/pugna/formatos/carrera/alta/definicion/elpportec/20060518elpcibpor_1/Tes/

http://www.20minutos.es/noticia/120724/0/abbas/gaza/hamas/
http://www.laverdad.es/murcia/pg060518/deportes/200605/18/barcelona-celebracion-champion-arsenal.html
http://www.diariovasco.com/pg060518/actualidad/espana/200605/18/inmigrantes-canarias-patera.html

2.2.6. Concepto de densidad de sitio

Cuanto mayor es el número de páginas que tiene un web site, más importancia y peso tiene.

Google considera que el esfuerzo de mantener un web site con 20 páginas indexadas no puede ser el mismo que el de mantener 200.000 páginas y que para ello debe haber una empresa, estructura y profesionales que lo soporten, indicando en general mayor compromiso y posible calidad de los resultados.

Y esta densidad de sitio se consigue abriendo nuestro contenido y presentando los resultados de nuestras bases de datos que, en muchas ocasiones están ocultos pendientes de una acción de búsqueda o clasificación del usuario (por ejemplo, una búsqueda de hoteles de Madrid de 3 estrellas no existe hasta que se ha hecho una elección de la variable Madrid y 3 estrellas del correspondiente menú.)

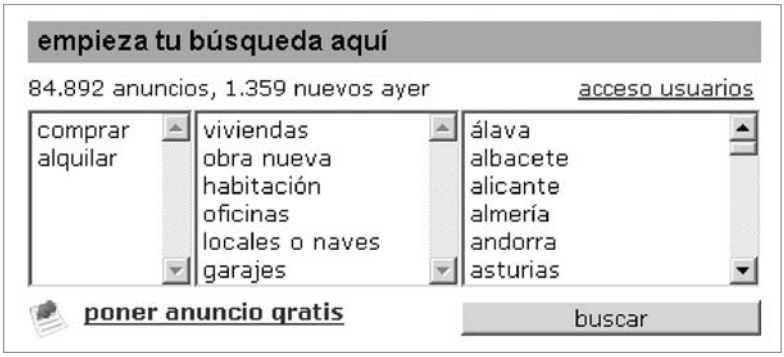
Así, la densidad del sitio tiene relación directa con el siguiente punto: indexando bases de datos.

2.2.7. Indexando bases de datos

La mayoría de las páginas de los webs actuales no existen como tal hasta que se ejecuta una acción. Una búsqueda, una clasificación, una opción. Son sites basados en bases de datos que generan páginas acorde con las elecciones de los usuarios.

Un ejemplo vale más que cualquier explicación.

El web site www.idealista.com basa la búsqueda de sus contenidos en este menú central:



No existe ninguna página hasta hacer las correspondientes selecciones y se hace click en buscar. Y esa acción no la hará jamás un spider que sólo sabe seguir enlaces.

Así, las inmensas posibilidades de este site están encerradas a no ser que se generen dichas páginas para poder ser indexadas, extrayendo todas las variaciones que ofrece esta base de datos. Y esas variaciones son el resultado de calcular las opciones de compra / venta (2) por los tipos de propiedad (6) por las provincias (52) por las poblaciones (n) hasta alcanzar más de 700.000 páginas indexadas que cubren toda búsqueda posible de propiedad en España.

Resultados 1–10 de aproximadamente 743.000 en el dominio www.idealista.com

Y poder presentar el contenido de nuestras bases de datos nos obliga a la creación de mapas de sitio.

2.2.8. Creando mapas de sitio

Primero quiero diferenciar entre los dos tipos de mapas de sitio.

Por un lado está, el mapa de web para el usuario, donde éste puede navegar para encontrar las páginas más habituales y que le permite una clasificación "más ordenada" del site. En general si un usuario tiene que usar esta opción es que el site está lo suficientemente mal organizado para llegar a esta página.

Por otro – el que nos incumbe – es el mapa para buscadores. En este mapa ejecutamos las diferentes opciones y variables que nos permite nuestra base de datos para convertirlas en páginas indexables – con una URL adecuada para ser guardada – y, por tanto, pasen a formar parte de la base de datos del buscador.

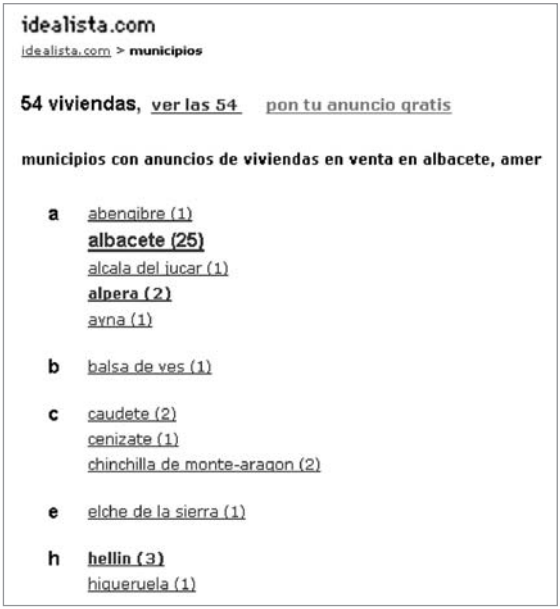
Volvamos al ejemplo de Idealista como mapa y así ver cómo han resuelto la presentación de su contenido. Esta es una imagen de su mapa inicial que presenta las primeras variables como alquiler, compra, etc.:

Al hacer click en cualquiera de las opciones de "tipo de inmueble" nos redirige hacia una búsqueda por provincias:

Sobre las 9 opciones anteriores de viviendas en venta, alquiler, obra nueva, etc., tenemos la primera multiplicación de resultados y además estamos realizando los enlaces internos con las palabras clave que pueden interesarnos (el nombre de la provincia).

Y al hacer click sobre cada opción pasamos a cada uno de los municipios de la misma. El camino que hemos seguido es, por ejemplo, alquiler vivienda/Madrid/Las Rozas, logrando hacer pública toda esa información.

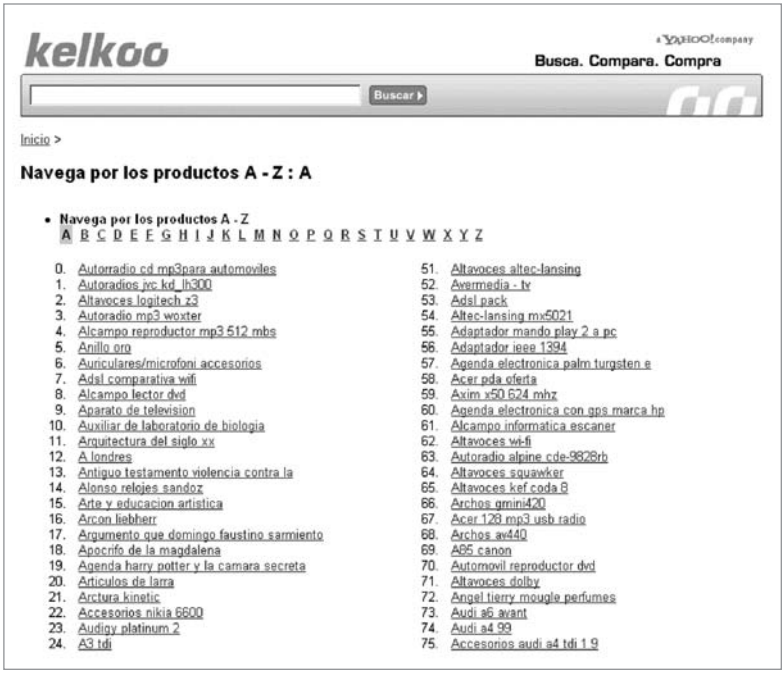
Esta es la pantalla que entra ya en el detalle de poblaciones:



Miles de páginas que no existían quedan ahora públicas y presentadas a través de enlaces que llevan nuestras palabras clave y en menos de 4 niveles de directorios.

Un mapa que permite convertir un site que, de otra manera, no existiría en un gigante de Internet con casi un millón de páginas indexadas de gran relevancia para el usuario.

Otro caso de alta calidad es el mapa de Kelkoo. Kelkoo en vez de publicar su directorio con sus combinaciones, transforma los resultados de búsqueda en páginas estáticas y las añade a un mapa de manera automatizada con una clasificación por letras y paginación de las mismas, generando centenas de miles de páginas que de otra manera no existían.


kelkoo

a YAHOO! company

Busca. Compara. Compra

Buscar

Inicio >

Navega por los productos A - Z : A

- **Navega por los productos A - Z**

A B C D E F G H I J K L M N O P Q R S T U V W X Y Z

0. Autorradio cd mp3para automoviles
1. Autoradios jvc kd_lh300
2. Altavoces logitech z3
3. Autoradio mp3 woxter
4. Alcampo reproductor mp3 512 mbs
5. Anillo oro
6. Auriculares/microfoni accesorios
7. Adsl comparativa wifi
8. Alcampo lector dvd
9. Aparato de television
10. Auxiliar de laboratorio de biologia
11. Arquitectura del siglo xx
12. A londres
13. Antiguo testamento violencia contra la
14. Alonso relojes sandoz
15. Arte y educacion artistica
16. Arcon liebherr
17. Argumento que domingo faustino sarmiento
18. Apocrifo de la magdalena
19. Agenda harry potter y la camara secreta
20. Articulos de larra
21. Arctura kinetic
22. Accesorios nikia 6600
23. Audigy platinum 2
24. A3 tdi
51. Altavoces altec-lansing
52. Awermedia - tv
53. Adsl pack
54. Altec-lansing mx5021
55. Adaptador mando play 2 a pc
56. Adaptador ieee 1394
57. Agenda electronica palm turgsten e
58. Acer pda oferta
59. Axim x50 624 mhz
60. Agenda electronica con gps marca hp
61. Alcampo informatica escaner
62. Altavoces wi-fi
63. Autoradio alpine cde-9828rb
64. Altavoces squawker
65. Altavoces kef coda 8
66. Archos gmini420
67. Acer 128 mp3 usb radio
68. Archos av440
69. A95 canon
70. Automovil reproductor dvd
71. Altavoces dolby
72. Angel tierry mougle perfumes
73. Audi a6 avant
74. Audi a4 99
75. Accesorios audi a4 tdi 1 9

Para encontrar este mapa debe hacerse click en la clasificación alfabética del pie de página descrita como A-Z.

◆ 2.3. Conseguir ser presentado

2.3.1. Relevancia del resultado

Nuestro web está listo y ha sido indexado.
Toda nuestra base de datos está presente y actualizada.

El googlebot pasa con una frecuencia aceptable y comprobamos que nuestros cachés son actuales.

Ahora es necesario aparecer en las posiciones que realmente funcionan tanto para las principales palabras clave como para la "larga cola" de búsquedas que hacen los usuarios relacionadas con nuestro contenido.

Para ello debemos trabajar en dos grandes áreas:

- La manera en la que presentamos nuestro contenido.
- La relación que el mismo tiene con el resto de Internet y dentro de nuestro propio sitio.

Y con ello conseguir que nuestro contenido sea considerado relevante para las búsquedas objetivo.

Un contenido relevante es aquel que, tanto por su propio contenido como por su relación con el resto de Internet, es considerado una de las primeras opciones para ser presentado en resultados de búsqueda de determinados conceptos.

2.3.2. Presentación del contenido. La importancia de todos los detalles

La estructura en la que se presenta un determinado documento o site tiene una importancia clave y es el cuidado de todos los detalles el que hace que se pueda pasar de la posición 3 a 1.

Cuando hablamos con técnicos sobre estos puntos les transmitimos la sensación de obligarles a hacer cambios sin aparente sentido o importancia, basados en criterios que no tienen una base 100% garantizada y que, además, son en ocasiones tan nimios que no pueden tener importancia.

Pero lo son. Una etiqueta, un título o un enlace puede hacer la diferencia.

a) Título, descripción, keywords y otros metas

Hagan cualquier búsqueda y siempre la palabra clave aparecerá en el título.

La etiqueta <title> tiene un peso extraordinario tanto en el documento en sí como en la relación con los otros.

Un título no debería tener más de 60 caracteres y tratar de ser lo más descriptivo posible y que incluya la palabra clave lo más cercana al principio que sea posible. En prensa los titulares suelen ser buenos títulos dado que ya reflejan por sí mismos los principales conceptos. Igualmente, al tratarse de productos el nombre de éste con combinaciones sobre los conceptos que usualmente se asocian al mismo resultan muy eficaces.

Por ejemplo, buenos títulos serían:

- El Barcelona gana la final de la Champions – La Verdad – Deportes
- El gobierno aprueba un Trasvase de emergencia a la cuenca del Segura – La Verdad – Región
- Nokia 6600. Ofertas de móviles para comprar al mejor precio en tu tienda Movizonia.com (nótese que incluye nombre con ofertas, comprar, precio y tienda)

Recordemos que es el título de la página el que Google presenta como enlace para llegar a la misma por lo que debe de invitar también al click y transmitir interés al usuario.

La etiqueta <description> tiene tanta importancia por posicionamiento como por ser en muchas ocasiones, el texto que presenta el buscador conjuntamente con el enlace, ayudando al usuario a tomar la decisión de hacer click.

Description debe ser una ampliación del título – pudiendo incluir de nuevo el mismo – pero ampliando información que ayude en la decisión de click y, en lo posible, que repita de 2 a 3 veces la palabra o frase clave para la que hayamos centrado el documento.

Siguiendo el ejemplo de Nokia 6600 antes expuesto una descripción adecuada sería:

- Nokia 6600. Descubre las mejores ofertas y precios para comprar tu Nokia 6600 en Movizonia, tu tienda de Móviles. Todo sobre teléfonos Móviles y accesorios de las principales marcas en sólo 48 horas.

La etiqueta <Keywords> es una de las grandes denostadas por el abuso del que fue lugar en su momento. Parece no ser tomada en cuenta y no hay cambios significativos por le hecho de usarla. No obstante, sí existe y como etiqueta aprobada por el W3C debe de ser considerada.

Por norma general no debe de incluir más de 10 ó 12 términos separados por comas y deben ser reiterativos con respecto a los presentados en el título y la descripción.

En el head podemos definir otros metas importantes como:

- Idioma. Es positivo indicar que se trata del español, pues ayudará en la búsqueda que puedan competir con conceptos internacionales desde la versión nacional de Google.
- Revisit. Independientemente del caso que pueda hacerle, indicar un periodo adecuado acorde con la frecuencia de cambios.
- Robots. Aunque la ausencia del mismo indica a los robots que tienen camino libre si además encuentran una etiqueta que invita a ello será mejor. No es obligada pero si existe y está validada por el W3C debe de utilizarse.

Como resumen y ejemplo del conjunto de metas:

<META NAME="Title" CONTENT="Nokia 6600. Ofertas de móviles al mejor precio para comprar en tu tienda : Movizonia.com">
<META NAME="Description" CONTENT="Nokia 6600. Descubre las mejores ofertas y precios para comprar tu Nokia 6600 en Movizonia, tu tienda de Móviles. Todo sobre teléfonos Móviles y accesorios de las principales marcas en sólo 48 horas.">
<META NAME="Keywords" CONTENT="Nokia 6600, nokia, móviles, ofertas, tienda, oferta, teléfonos móviles, accesorios, comprar, venta, precio, precios">
<META NAME="Language" CONTENT="Spanish">
<META NAME="Revisit" CONTENT="7 day">
<META NAME="Robots" CONTENT="All">

b) Densidad de página y densidad de keyword (Keyword density)

Me permito utilizar este concepto para dos temas diferentes pero relacionados.

Por un lado, entendemos densidad de página como la cantidad de texto visible que tiene una determinada página. A los buscadores les gustan las páginas con texto. Si estas páginas tienen la construcción adecuada – de la que hablamos más adelante – y tratan de manera consistente sobre un determinado tema.

¿Cuánto texto es mucho, poco o correcto? Depende de cada buscador, pero parece que Google prefiere páginas que tienen entre 600 y 1.500 palabras.

Y por otro lado, la densidad de palabras clave en una determinada página. La presencia de la palabra clave dentro de una página en varias ocasiones es denominada keyword density y se define como un porcentaje de veces que aparece la palabra clave sobre el total de palabras. Si nuestra palabra clave aparece 3 veces en un texto de 100 tendremos una densidad del 3%.

Teóricamente cuanto mayor sea dicho índice más relevante se considerará el contenido para dicha palabra, pero existe un límite que, al ser sobrepasado, convierte la página en spam.

Como en muchos otros casos la naturalidad debe ser la política a seguir. Se puede forzar un determinado texto para incluir alguna palabra clave de más pero el usuario debe ser el primer beneficiario en este punto. De hecho, una de las actuales discusiones es el posible empobrecimiento en la redacción de textos por el condicionante del posicionamiento en buscadores. La influencia de Google está resultando tan aplastante que está provocando cambios en los estilos de redacción para su adecuada indexación. El posicionamiento es clave, pero jamás debería de influir en nuestra forma de comunicarnos con nuestros usuarios o clientes.

"Redacción periodística versus redacción para buscadores"

La enorme influencia que Google presenta en el tráfico de los webs está llevando a que se haya creado un estilo de redacción adaptado a los criterios de posicionamiento, dejando de lado al usuario. Esta discusión está ahora presente en la red y obliga a los profesionales a mantener un criterio de calidad por encima de los resultados. O al menos, a buscar el adecuado equilibrio entre ambos temas con el usuario como primera preocupación.
Un buscador no debería condicionar marcar el estilo o de redacción de un medio.

c) Utilización de etiquetas de clasificación y énfasis

Las etiquetas validadas por el W3C esconden buena parte de la presentación de contenidos acorde con lo que espera un buscador. Y refleja el sentido común que se debe seguir al presentar el mismo.

Un buen contenido periodístico – llevado al esquema de contenido de una página web – debería tener:

- Titular: la etiqueta <title>. Y como el titular la más importante.
- Entradilla: la etiqueta description. Avanza de manera general el contenido de toda la página y debe ser capaz de dar un resumen que invite a seguir leyendo.
- Subtitulares o destacados: las etiquetas de <Heading> o encabezados <H1> hasta <H6>. Ayudan a ordenar adecuadamente la información, abriendo párrafos y destacando ideas generales.
- Negritas o destacados: como si de entresacados, negritas o declaraciones, aparecen las etiquetas de o <emphasis> que aportan mayor importancia a una frase o palabra.
- Cuerpo de texto: que facilita toda la información y es la mayor parte del contenido.

Este mismo ejercicio es el que debemos ejecutar al usar nuestro código HTML y las posibilidades que nos ofrece el W3C.

Debemos usar las etiquetas <H> de la siguiente manera:

- Siempre en el titular.
- De manera escalonada. Esto es, si en el titular es una etiqueta H2 en un subtitular deberá ser H4 y si existiese otro subtítulo de nivel inferior sería H6.
- Nunca iniciar con una etiqueta H inferior a otra que pueda venir a continuación.
- Nunca usar una etiqueta H en un párrafo de cierta longitud. Como bien dice su nombre es un Heading o cabecera.

La etiqueta y <emphasis> se utiliza para destacar un determinado contenido y darle importancia sobre el resto. Habitualmente los programado-

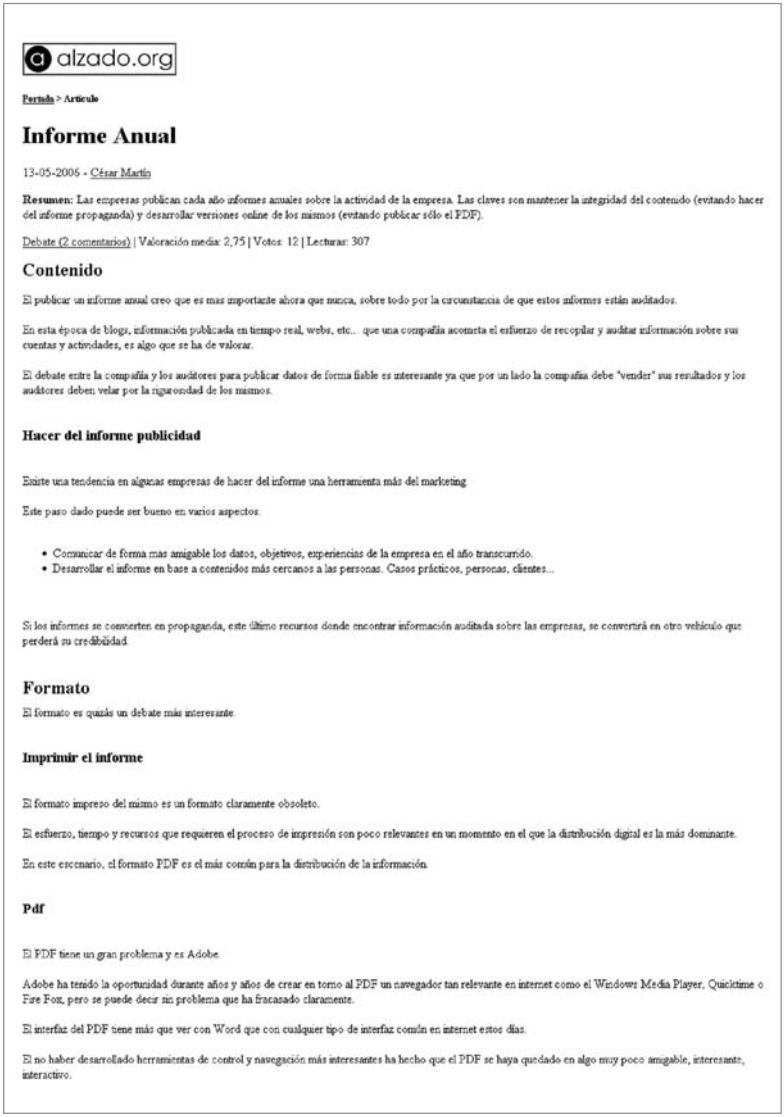
res utilizan, o bien una etiqueta propia dentro de la hoja de estilos, o bien la negrita. Ambas son consideradas como estéticas, pero que no aportan importancia al contenido. Todo destacado dentro de los textos debe estar bajo dichas etiquetas.

Una estrategia adecuada es utilizar para destacados y <emphasis> para citas o referencias.

Presentamos a continuación un magnífico ejemplo de organización en la web de Alzado.org.

Primero la página como se vería con los correspondientes estilos y después, la misma tal y como la lee un buscador sin estilo alguno, destacando la magnífica estructura de la misma.

Web tal y como se ve en el navegador.



alzado.org

Portada > Artículo

Informe Anual

13-05-2006 - César Martín

Resumen: Las empresas publican cada año informes anuales sobre la actividad de la empresa. Las claves son mantener la integridad del contenido (evitando hacer del informe propaganda) y desarrollar versiones online de los mismos (evitando publicar sólo el PDF).

Debate (2 comentarios) | Valoración media: 2,75 | Votos: 12 | Lecturas: 307

Contenido

El publicar un informe anual creo que es más importante ahora que nunca, sobre todo por la circunstancia de que estos informes están auditados.

En esta época de blogs, información publicada en tiempo real, webs, etc... que una compañía acometa el esfuerzo de recopilar y auditar información sobre sus cuentas y actividades, es algo que se ha de valorar.

El debate entre la compañía y los auditores para publicar datos de forma fiable es interesante ya que por un lado la compañía debe "vender" sus resultados y los auditores deben velar por la rigurosidad de los mismos.

Hacer del informe publicidad

Existe una tendencia en algunas empresas de hacer del informe una herramienta más del marketing.

Este paso dado puede ser bueno en varios aspectos.

- Comunicar de forma mas amigable los datos, objetivos, experiencias de la empresa en el año transcurrido.
- Desarrollar el informe en base a contenidos más cercanos a las personas. Casos prácticos, personas, clientes...

Si los informes se convierten en propaganda, este último recursos donde encontrar información auditada sobre las empresas, se convertirá en otro vehículo que perderá su credibilidad.

Formato

El formato es quizás un debate más interesante.

Imprimir el informe

El formato impreso del mismo es un formato claramente obsoleto.

El esfuerzo, tiempo y recursos que requieren el proceso de impresión son poco relevantes en un momento en el que la distribución digital es la más dominante.

En este escenario, el formato PDF es el más común para la distribución de la información.

Pdf

El PDF tiene un gran problema y es Adobe.

Adobe ha tenido la oportunidad durante años y años de crear en torno al PDF un navegador tan relevante en internet como el Windows Media Player, Quicktime o Fire Fox, pero se puede decir sin problema que ha fracasado claramente.

El interfaz del PDF tiene más que ver con Word que con cualquier tipo de interfaz común en internet estos días.

El no haber desarrollado herramientas de control y navegación más interesantes ha hecho que el PDF se haya quedado en algo muy poco amigable, interesante, interactivo.

La misma página tal y como se ve sin estilos: organización jerárquica del contenido y uso adecuado de las etiquetas H.

d) Equilibrio código/contenido. Maquetación con capas y css

Recordemos. A los buscadores les gusta el contenido. Texto organizado adecuadamente sobre un tema concreto. No buscan imágenes, ni diseños, ni películas, ni flash… buscan contenido de calidad adecuadamente organizado.

Y en una página web hay dos tipos de información:

- Visible por el usuario. Aquella por la cual nos visita. Aquella que puede leer y analiza.
- Invisible para el usuario. Código fuente utilizado para presentar el anterior. Puede ser en forma de tablas o de capas controlables desde la hoja de estilo. El usuario no necesita tener idea alguna de cómo funciona.

Y el contenido visible para el usuario puede presentarse con mucho contenido invisible o con poco. Con tablas y más tablas anidadas que generan infinidad de líneas de código o con capas marcadas desde el archivo css que genera el mínimo código posible.

Este concepto se resume en la capacidad de separar contenido de presentación, haciendo los diseños más fáciles de cambiar, más eficaces en su presentación, de menor carga de datos – al estar muchas de sus pautas en el archivo css – y esta eficacia en la ejecución de un web tiene efectos colaterales positivos para los buscadores.

Cuanta mayor sea la proporción de contenido visible sobre contenido invisible mucho mejor.

El mismo ejemplo de Alzado.org vale para este punto.

e) Alts en imágenes

Las imágenes no son leídas por los buscadores. Entienden que hay una imagen, pero no pueden saber si es un espacio negro o un cuadro de Miró.

La única manera de que interpreten qué hay en una determinada imagen es incluir la etiqueta <Alt> que presenta el texto alternativo.

Usar esta etiqueta es importante porque:

- Permite indicar que representa una imagen incluyendo en dicha etiqueta nuestras palabras clave.
- Permite considerar los posibles usuarios discapacitados que encontrarán contenido en la misma.
- Permite que si una imagen es enlazada dicho alt actúe como texto de enlace apoyando dicho link y su correspondiente fuerza como enlace.

La etiqueta Alt convierte a una imagen en contenido y el contenido es información para usuarios y buscadores.

f) Links internos

La popularidad de enlaces es tanto interior como exterior. Son claves los enlaces externos, pero también los internos tienen peso e importancia.

Por enlaces internos entendemos todos aquellos que enlazan páginas dentro del propio web. Si una página llama a otra con un determinado texto de enlace, está apoyando a la misma con dicho texto.

Para conocer que páginas enlazan hacia un site o página se puede consultar con el comando link:www.nombredominio.com. Como vemos al hacerlo, aparecen tanto links externos como internos. En muchos casos, la mayoría son internos y cuando una página está enlazada desde toda la web o parte de ella considera que la misma es más relevante que el resto.

Recordemos que el comando link:www.nombrededominio.com ofrece una muestra de los enlaces y que el mismo comando es también válido en Yahoo con link:http//www.nombrededominio.com donde normalmente encontrará muchísimos más resultados dado que Yahoo presenta todos los que tiene en su base de datos.

De hecho, esta es una estrategia interna que siguen determinados portales para apoyar aquellas secciones que les interesa en un periodo dado e indica la importancia de la adecuada selección de los menús que contienen enlaces constantes dentro de todo un site.

Así, tomemos como ejemplo la palabra "CENTRO COMERCIAL" que aparece en el menú superior de elcorreodigital.es y otros soportes del grupo Vocento. Ello hace que la página de la Trastienda Digital aparezca dentro de la primera página para esa búsqueda pero sólo para la URL de la subtienda de el Correo digital www.latrastiendadigital.com/elcorreodigital/, a pesar de que dicha palabra no aparece en la web a la que se enlaza.

Este ejemplo deberá destacarse posteriormente al hablar de contenido duplicado.

g) Relevancia de contenido

Por último y aunque parezca evidente, el contenido debe ser relevante en conjunto. Toda la estructura comentada anteriormente, junto con su contexto, que demuestre que no sólo las palabras, sino aquello que las rodea está relacionado con la temática que se está tratando y, por tanto, es relevante y de calidad.

¿Cómo Google es capaz de identificar las palabras y su contexto?

Su algoritmo ya ha afinado hasta el punto de ser capaz de identificar patrones donde determina qué otras palabras tienen en su entorno y por tanto las ubica en contexto.

Y para demostrarlo el mejor ejemplo es la herramienta de sugerencia de palabras clave de Google para las campañas en AdWords. Cuando se le hace una sugerencia Google no sólo presenta palabras en plural o asociadas a lugares o adjetivos sino también sinónimos u otras variaciones contextuales de la misma palabra como recomendación.

Por ejemplo, ante la palabras "Billetes" la herramienta de Google sugiere "avión", "numismática", "vuelos"… todas ellas asociadas a los contextos donde suele encontrar la inicial.

De esta manera Google lucha contra aquellas páginas orientadas a hacer spam de buscadores donde potencian una sola palabra pero olvidan situarla en el contexto adecuado.

No basta la palabra sino también todo aquello que la rodea.

Y todo esto en el fondo se limita a generar contenido de calidad e interés para el usuario presentado de la manera más sencilla y organizada posible para el buscador.

2.3.3. Eligiendo palabras objetivo

a) Herramientas de sugerencia de palabras clave

¿Cómo nos busca el usuario? ¿Cuáles son los términos que utiliza?

Descubrir las pautas de comportamiento son claves para orientar nuestro web site y construir una posterior estrategia de captación de enlaces.

Y debemos de pensar en dos formas de buscar:

- Aquella que es la estructural o principal. Cuáles son las palabras que definen de manera global nuestro negocio y que suelen ser las de mayor volumen. Esto es, si somos un centro comercial, serán las palabra "Ofertas" "Centro Comercial", "Compras", etc.
- La larga cola de palabras. Si como centro comercial tenemos 200.000 productos, cada uno de ellos se convierte en una puerta de entrada. Y cada producto o categoría debe de quedar enfocado hacia ese nicho. Esa larga cola acaba suponiendo, en la mayoría de las ocasiones, mucho más volumen que las palabras clave estructurales.

Para averiguar cómo nos busca el usuario tenemos una principal herramienta para el cálculo de volumen y dos para estimación de coste:

- Herramienta de sugerencia de palabras clave de Yahoo Search Marketing. A pesar de representar un porcentaje pequeño de las búsquedas es lo suficientemente representativo estadísticamente hablando. Esta herramienta tiene la gran ventaja de mostrar cantidades permitiendo asignar importancia a cada palabra. Es pública y se encuentra en esta URL: http://inventory.es.overture.com/d/searchinventory/suggestion/
- Herramienta de ofertas de Yahoo. Permite ver cuál es el precio y las empresas que están dispuestas a pagar por cada click. http://uv.bidtool.overture.com/d/search/tools/bidtool/?mkt=es&lang=es_ES
- La herramienta de estimación de tráfico de Google permite mayores opciones de segmentación pero sólo facilita el número de clicks estimados y el coste que tiene en la actualidad en el sistema sin contar el historial del mismo, por lo que vale como análisis económico, pero no para volumen u orientación. https://adwords.google.com/select/TrafficEstimatorSandbox.

Antes de iniciar una labor de optimización o puesta en marcha de campaña debemos elegir palabras objetivo y construir un Excel que nos permita hacer un seguimiento de la evolución posterior.

◆ 2.4. El concepto de popularidad

Dos ejemplos ilustran muy bien el concepto de popularidad.

Primero, el colegio. ¿Recuerdas cuán popular eras en el colegio?. De una clase de 40 sólo 3 ó 4 lo eran y tu posición en el ranking de si eras más o menos "cool" dependía de tus relaciones con los demás. Si andabas con los "cool" y encima tenías amigos de clases superiores subían tus puntos. Si andabas con los torpes y tus colegas eran del curso anterior estabas listo.

Segundo, el mundo científico. Un científico tiene como meta final publicar. Publicar es el mayor reconocimiento que puede obtener. Cuanto más publique, más importancia tendrá en su especialidad. Y además de la cantidad, está la calidad. ¿Dónde le publican? Es mayor el éxito de un ensayo en "The Lancet" que en "Gaceta universitaria". Y por último ¿Cuántos otros científicos hablan de tu trabajo? El concepto de citación es clave en la investigación que muchas veces se basa en las investigaciones ya reconocidas por otros, ya aprobadas por la "comunidad científica".

Esta misma situación marca tu popularidad en la red. No importa si tienes muchos enlaces. Importa si éstos son buenos o si tratan del mismo tema que tú. Importa si la comunidad te ha reconocido. Importa cuán arriba estás de esa pirámide social cuyo puesto debes ganarte con calidad y credibilidad.

En formato de definición podría decirse que la popularidad de enlaces o link popularity es la medida de cantidad y calidad de otros web sites que enlazan hacia el tuyo en la red en un momento dado.

Los 3 buscadores actuales basan buena parte de su algoritmo en este concepto. Y para ello miden el número de enlaces que otros sitios hacen hacia tu página. Cada enlace es un voto. Pero un voto no democrático dado que dicho enlace trae mensajes consigo. Un voto de alguien importante vale más. Y un voto de alguien que sabe mucho de un tema vale más principalmente para ese tema.

Así nace también el concepto de <Authority site> que define por qué un web sabe más que otro sobre una temática concreta y que ayuda a entender por qué suelen ser webs especializados los que aparecen para muchas de las búsquedas.

2.4.1. La importancia de los enlaces entrantes

Así, cada link es un voto y por ello debemos ganarnos votos dentro de la comunidad de Internet y definir una estrategia de captación de enlaces entrantes.

Estos enlaces entrantes definen no sólo el número de votos, sino también las palabras con la que nos votan o nos consideran relevantes (lo cual es válido también de manera interna en nuestro site).

"Un enlace, un voto"

La teoría de la citación científica llevada a Internet.
Cada voto cuenta, pero no todos los votos son iguales ni tienen una autoridad adecuada para cada tema. Un enlace de El Mundo puede tener menos valor que el de una web especializada para una determinada búsqueda.
Democracia, pero sólo en parte.

La estrategia de captación debe de enfocarse tanto en cantidad como en calidad y temática:

- Cuantos más enlaces mejor (siempre que no provengan de malos vecinos aunque sea inevitable la ausencia de control sobre este punto).
- Calidad. Intentar que sean otros sites de calidad y Page Rank elevado quienes nos enlacen.
- Temática. Intentar que sean authority sites sobre nuestra temática o páginas sobre nuestra temática los que nos enlacen

Y esos enlaces entrantes deben tratar de ser conseguidos con las palabras clave en los mismos para conseguir así su máxima eficacia.

2.4.2. Presencia en directorios

Los directorios organizados por hombres como Yahoo o Dmoz se consideran web sites de alta calidad y enlaces relevantes. Al haber tenido una revisión manual, dicho enlace implica que ha pasado unos controles de calidad antes de ser aprobados dadas las filosofías que utilizan ambos. Esos dos directorios por sí solos, permiten garantizar un cierto peso en la red y deben ser definidos como los primeros objetivos de cualquier estrategia de captación de links.

En segundo lugar, deben buscarse directorios temáticos sobre nuestra materia.

Para encontrar los mismos recomendamos utilizar Buscopio www.buscopio.net con una adecuada clasificación de buscadores por regiones y temática.

En general será muy difícil lograr que los directorios nos enlacen con una palabra clave diferente al título de nuestro site. A cambio, suele ser sencillo lograr que, web sites de calidad, sean incluidos. De hecho, lo normal es que si un web lleva tiempo en la red y es reconocido en su sector, ya esté incluido aunque éste no lo haya solicitado.

"Comandos útiles de Google"

Google tiene una serie de comandos que nos ayudan a conocer información muy relevante sobre nuestro site y su situación.

- Comando link. Link:www.dominio.com. Presenta toda una muestra de todas las páginas que enlazan a nuestra web.
- Comando allinurl. Allinurl:palabra. Presenta todos los webs que tienen dicha palabra en su URL.
- Comando allintitle. Allintitle: Nokia 6600. Presenta todas las páginas que incluyen dichas palabras en su título.
- Comando allintext. Allintext: palabra o frase clave. Presenta todas las páginas que incluyen dicha palabra en su body.
- Comando site. Site.www.dominio.com Indica cuántas páginas existen en la base de datos de Google. Si además se añade +palabra hace la búsqueda de la misma sólo en ese dominio.
- Comando filetype. Filetype:pdf nokia. Presenta todos los archivos de dicho tipo con la palabra clave determinada.

2.4.3. Técnicas para conseguir ser enlazado

Servicios y contenido

En esos dos puntos suele resumirse las razones por las que otro web tenderá a enlazarnos.

Tener algo de interés – contenido – u ofrecer un servicio gratuito.

El arte de ser enlazado ha generado incluso su propia palabra "Link baiting" o "Cebando enlaces" que busca encontrar ganchos que permitan lograr esos enlaces. Es lograr algo lo suficientemente interesante como para conseguir ser enlazado.

En este punto, es importante hacer un guiño a los usuarios de Blogs. Nunca ha sido tan fácil publicar en Internet y, por tanto, enlazar hacia otros sites, que en este momento gracias al extensivo uso de los Blogs. Poner un enlace no significa esfuerzo alguno y es precisamente una de las bases del estilo de comunicación en Blogs.

Y además, los Blogs están – por su propia construcción en capas y clasificación temática y cronológica y su fuerte cantidad de enlaces entre sí – muy bien preparados para ser adecuadamente indexados y ser además, generadores de enlaces de calidad.

Entre las ideas para promover el ser enlazado:

- Liberar contenido a cambio de enlaces. Una estrategia muy habitual por parte de los Blogs y profesionales y fácil de implementar en el mundo editorial. Liberar parte del contenido que se publica diariamente a cambio del reconocimiento vía enlaces puede ser una manera sencilla y eficaz de conseguir los mismos.
- Construir herramientas de utilidad que puedan ser incluidas en otros sites.
- Crear concursos o competiciones.
- Hacer tests que impliquen participación.
- Humor.
- Gratuidad.

- Retransmisión en directo on line de algún acontecimiento.
- Utilización de herramientas de comunidad para colgar determinados materiales interesantes. Por ejemplo Flirck – fotos – o YouTube – vídeos – pueden conseguir que se haga referencia a dicho material e indirectamente a la fuente.

2.4.4. La importancia de las palabras de enlace

Nos encontramos ante el punto más importante en las estrategias de posicionamiento actual. Conseguir enlaces con las palabras adecuadas puede disparar nuestra posición para las mismas. Es tan estratégico que es posible aparecer en las primeras páginas, sin siquiera mencionar dicha palabra en el contenido.

Y el ejemplo de su poder está en los Google Bombings que antes explicábamos.

Las palabras de enlace o "anchor text" son las palabras usadas para enlazar hacia una determinada página. Es la base de la citación. Cuenta a los buscadores por qué una determinada página es relevante.

Lógicamente, no se tiene control de cuáles son las palabras con las que otros websmasters nos enlazan y es muy habitual que sea el propio nombre del web site, el autor o simplemente frases estándar como "click aquí" o "ver más". Ese enlace estará aportando como voto pero, no indica exactamente al buscador para qué concepto es relevante – aunque sí que lo hace el contexto y contenido general de la página que enlaza –.

Si se pudiese elegir, una adecuada estrategia de selección de palabras de enlace sería:

- Natural. No conseguir enlaces siempre a la misma página ni con la misma palabra. Los enlaces deben aparecer de forma natural. Algunos con palabras clave otros con genéricas, otros con el nombre del dominio…
- Variada. No enlazar siempre con la misma palabra. Una vez definidas las diferentes palabras clave, usar una variedad de las mismas.
- Pensada para el usuario. Recordemos que estos textos debe leerlos también el usuario. Y él debe ser la primera preocupación, teniendo sentido y aportando valor en la comprensión de la página y su destino.

• En lo posible, sin reciprocidad. Cuando un web site enlaza hacia otro y en el mismo existe un enlace de vuelta hacia el original, se considera que existen enlaces recíprocos o intercambio de enlaces. Esta acción no está considerada negativamente o como spam pero hará que la importancia de dicho enlace sea algo inferior. Es una estrategia muy común entre redes de portales o web sites que puede llegar a ser penalizada cuando se trata de granjas de enlaces – web sites creados exclusivamente para generar enlaces hacia sites que se desean promocionar y que se cruzan entre sí para generar popularidad sin aportar valor alguno al usuario – .

◆ 2.5. Normas de calidad de contenido

2.5.1. Normas de calidad y contenido de Google y Yahoo

Presentes desde el principio en las páginas de Google son tan básicas que pecan de ingenuas. Pero su lectura entre líneas permite descubrir dónde está la esencia de todo éxito en posicionamiento.

Directrices sobre el diseño y los contenidos

• Cree un sitio con una jerarquía y vínculos de texto claros. Se debe poder acceder a todas las páginas desde, al menos, un vínculo de texto estático.
• Ofrezca a los usuarios un mapa del sitio con vínculos que apunten a las secciones importantes de su sitio. Si el mapa del sitio contiene más de 100 vínculos, puede dividirlo en varias páginas.
• Cree un sitio útil con mucha información y redacte páginas que describan su contenido con claridad y exactitud.
• Determine las palabras que los usuarios escribirían para encontrar sus páginas y asegúrese de que su sitio realmente las incluya.
• Intente utilizar texto en lugar de imágenes para mostrar nombres, contenido o vínculos importantes. El rastreador de Google no reconoce texto incorporado dentro de las imágenes.
• Asegúrese de que las etiquetas <title> y <alt> sean descriptivas y precisas.
• Compruebe que no haya vínculos rotos o código HTML incorrecto.
• Si decide utilizar páginas dinámicas, (por ejemplo, una URL que conten-

ga un carácter "?"), tenga en cuenta que no todos los motores de búsqueda rastrean páginas dinámicas, además de páginas estáticas. Se recomienda que los parámetros sean cortos y escasos.

• No ponga un número excesivo de enlaces en una página determinada (no debería superar los 100).

Directrices técnicas

• Utilice un navegador de sólo texto como Lynx para examinar su sitio, ya que la mayoría de las arañas de los motores de búsqueda ven su sitio de la misma manera que lo haría Lynx. Si las funciones avanzadas como JavaScript, cookies, ID de sesión, marcos, DHTML o Flash impiden la visualización de todo su sitio en un navegador de texto, es posible que las arañas de los motores de búsqueda tengan problemas para rastrearlo.

• Permita que los robots de búsqueda rastreen sus sitios sin ID de sesión o argumentos que hacen un seguimiento de su ruta por el sitio. Estas técnicas son de gran utilidad para el seguimiento del comportamiento de un usuario individual, pero el patrón de acceso de los robots es completamente distinto. El uso de estas técnicas puede llevar a una indexación incompleta de su sitio, ya que es posible que los robots no puedan eliminar las URL que tengan un aspecto distinto, pero que en realidad apuntan a la misma página.

• Asegúrese de que su servidor web admite la cabecera HTTP "If-Modified-Since" (Si se modificó desde). Esta función permite que su servidor web avise a Google si el contenido de su sitio ha cambiado desde la última vez que lo rastreamos. La compatibilidad con esta función le ahorra ancho de banda e inversiones generales.

• Utilice el archivo robots.txt en su servidor web. Este archivo comunica a los rastreadores qué directorios pueden o no rastrearse. Asegúrese de que el archivo esté actualizado para su sitio de modo que no bloquee accidentalmente el rastreador Googlebot. Si desea obtener más información acerca de cómo configurar los robots que visitan su sitio, vaya a http://www.robotstxt.org/wc/faq.html.

• Si su compañía adquiere un sistema de administración de contenido, asegúrese de que éste pueda exportar el contenido de modo que las arañas de los motores de búsqueda puedan rastrear su sitio.

• No use "&id=" como parámetro en sus URLs, puesto que no incluimos este tipo de páginas en nuestro índice.

Y, con la misma intención, están las de Yahoo!

Páginas web que Yahoo! quiere en su índice

- Contenido original y único con un valor demostrable.
- Páginas diseñadas principalmente para personas, con las consideraciones sobre motores de búsqueda en un segundo plano.
- Enlaces pensados para ayudar a la gente a encontrar – cuando aplique – contenidos relacionados e interesantes.
- Metadata (incluidos título y descripción) que describen con precisión el contenido de una página web.
- Buen diseño web general.

Por desgracia, no todas las páginas web contienen información útil para el usuario. Algunas son creadas expresamente para engañar el motor de búsqueda y ofrecer resultados inapropiados, redundantes o de ínfima calidad; esto es lo que se suele denominar "spam." Éste es el tipo de páginas que Yahoo! no quiere en su índice.

Lo que Yahoo! no quiere incluir

Ejemplo de algunos – que no todos – tipos de páginas que Yahoo! no quiere incluir:

- Páginas que dañan la exactitud, relevancia o diversidad de los resultados de búsqueda.
- Páginas que sólo sirven para redirigir al usuario a otra página.
- Páginas que tienen prácticamente el mismo contenido que otras páginas
- Sitios con varios e innecesarios hostnames virtuales.
- Páginas generadas abundante y automáticamente y que tienen escaso valor
- Páginas que usan métodos artificiales de incremento de relevancia en motores de búsqueda.
- Uso de contenido no visible para el usuario.
- Páginas que ofrecen al motor de búsqueda contenidos diferentes a los que el usuario puede ver.
- Excesivos enlaces cruzados creados para incrementar la popularidad aparente de un sitio.
- Páginas pensadas sólo para motores de búsqueda.
- Uso indebido de marcas de la competencia.

- Múltiples sitios que ofrecen el mismo contenido.
- Páginas que producen un gran número de pop-ups que interfieren la navegación.
- Páginas engañosas, fraudulentas o que proporcionan una pobre experiencia al usuario.

◆ 2.6. Errores más comunes

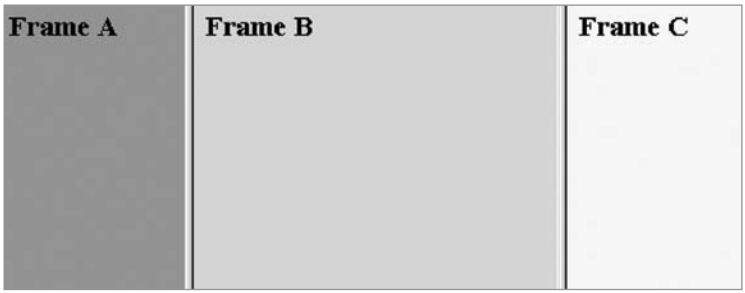
A la hora de diseñar y programar un web site hay errores básicos y estratégicos, que, de cometerse, harán prácticamente imposible una adecuada indexación. Como en los casos anteriores, estos errores complican la experiencia no sólo a los robots, sino también a los propios usuarios y en muy pocas ocasiones se justifica el uso de estas técnicas.

Ya comienzan a ser tan obvios y conocidos que empiezan a escasear en la red, pero todavía sorprende cómo profesionales muy conocedores de su trabajo insisten en su uso.

2.6.1. Uso de Frames

Un frameset es una página que permite presentar un conjunto de páginas al llamar a cada una de ellas de manera independiente. Su uso se justifica en raras ocasiones, principalmente cuando se presentan listados muy extensos que necesitan continuar presentes durante todo el proceso de navegación.

Una página web típica con frames tiene este aspecto:



Erróneamente, muchos programadores utilizan los frames para conseguir controlar las posiciones del contenido al incluir páginas de ancho o alto variable en los laterales o la parte superior e inferior.

Las posibles combinaciones de frames son casi infinitas dado que un frame set puede llamar a otro dentro del mismo – convirtiéndose en un frame anidado – y complicando enormemente su integración.

Para ver ejemplos de uso de frames extremos recomendamos visitar www.queweb.org que presenta casos de diseños y usos web erróneos.

2.6.2. Redirects

Un redirect hace que una página redirija a otra automáticamente sin intervención alguna del usuario. Un ejemplo claro de redirect sería el uso de un adserver. Un adserver presenta un anuncio con un determinado enlace dentro del servidor del que, al hacer click, primero visita la URL del adserver que después dirige al usuario a la página del cliente. Esta acción se realiza para poder controlar el número de impresiones y, principalmente, el número de clicks para controlar la eficacia de la campaña.

El redirect, además, no transmite popularidad a través de sus enlaces.

Esta técnica no tiene penalización usada de esta manera, pero estos redirects pueden ser usados de manera fraudulenta al crear una página exclusivamente para buscadores con contenido sobre optimizado que, cuando es visitada por el usuario, es redirigido de inmediato a la página de destino.

2.6.3. Enlaces Javascript y Flash

Los enlaces javascript no son seguidos por los buscadores al considerar que existe una acción de usuario. Los enlaces deben ser enlaces planos de texto o imagen con alt, pero no basados en acciones javascript. Usar estos enlaces no implica diferencia en programación y evita que el robot pueda seguir el mismo y aportar el valor que dicho enlace pueda tener para la página a la que dirige.

Similar problemática aparece en el uso de flash. La respuesta no es siempre sencilla, pues depende de varios factores. Google, por ejemplo, es capaz de

indexar textos y seguir enlaces que se encuentren en una película Flash. Al hablar de texto contenido nos referimos a texto que fue introducido directamente en modo de edición, no información recibida de fuentes externas.

Algunas soluciones, como colocar textos de fuentes externas en forma de comentarios en el código HTML, suelen funcionar, mientras que redireccionar spiders a páginas especiales es considerado abuso y puede repercutir negativamente en el posicionamiento de una web en los diferentes motores de búsqueda. Para ellos, no existe nada mejor que contenido relevante y accesible. En ese sentido, aplicaciones híbridas (Flash/HTML) suelen tener los mejores resultados.

En general, recomendamos no usar Flash dentro de menús y elementos de navegación y sacar el máximo partido a esta herramienta en demostraciones interactivas donde el diario Elmundo.es presenta magníficos ejemplos.

2.6.4. Cloacking o páginas exclusivas para buscadores

Existe una técnica que resulta muy tentadora por sus resultados inmediatos y que permite compaginar el uso de páginas basadas en Flash y otros elementos muy gráficos con un buen posicionamiento.

El cloacking es la técnica que se basa en crear páginas exclusivas para los robots de los buscadores a los cuales se les identifica por su IP o agente y se le presenta una página diferente a la que ven los usuarios. Por supuesto, esa página está preparada al 100% para las palabras clave y nunca será vista por los usuarios. Porque los usuarios, al hacer click sobre el enlace del buscador serán llevados a las páginas normales – dado que no es identificado como uno de los robots –.

Pero sus riesgos son muy elevados.

A principios del 2006, BMW.de fue eliminado del directorio de Google durante un mes por utilizar esta técnica. Esta decisión fue discutida dentro de los foros de Internet dado que Google estaba utilizando tiempos de penalización diferentes al tratarse de una empresa como BMW y, por otro lado, lleva a la pregunta de ¿puede Google permitirse eliminar de sus resultados páginas tan relevantes como esa?

2.6.5. Textos ocultos

Este truco consiste en incluir dentro de capas invisibles o con textos con el mismo color de fondo, palabras clave de manera repetitiva y plagada de enlaces para lograr que esa página sea considerada relevante.

Son contenidos ocultos y por tanto penalizados. Y es una técnica muy fácil de descubrir.

Recordemos que cualquiera de estas técnicas son susceptibles de ser denunciadas por cualquier usuario a través de un spam report que, en caso de prosperar, nos eliminará de la base de datos durante un importante periodo de tiempo.

2.6.6. Sobreoptimización y Contenido duplicado

Conseguir una mayor densidad de página, especialmente cuando se utilizan bases de datos, puede resultar relativamente sencillo.

Siempre un ejemplo será mucho más eficaz.

Una empresa de alquiler de coches tiene 3 delegaciones en Alicante, Málaga y Mallorca. Puede optimizar para dichas palabras, pero al no tener gran cantidad de contenido no podrá:

- Tener una importante densidad de página.
- Atacar palabras clave secundarias.

Al tratarse de un alquiler de coches decide crear páginas para cada semana del año y así, en vez de tener 3 páginas:

- Alquiler de coches Alicante.
- Alquiler de coches Málaga.
- Alquiler de coches Mallorca.

Pasa a tener 156:

- Alquiler de coches Alicante del 1 al 7 de enero.
- Alquiler de coches Alicante del 8 al 15 de enero…

Y a cada una de esas páginas, cuyo contenido es en un 99% similar al original, puede adaptarse a otras palabras objetivo.

Esta técnica puede ser un ejemplo de sobre optimización y contenido duplicado. ¿Cuál es el límite de optimización y duplicación de contenidos? ¿Es lícita una acción de este tipo? En general son acciones que resultan penalizadas y descubrir nuestra base de datos no debe implicar multiplicarla por todas las variaciones posibles.

Hay una excepción: las noticias de agencias.

Nos encontramos ante un caso de contenido duplicado que no es susceptible de ser penalizado. Decenas de medios en todo el mundo se nutren de las mismas fuentes: las agencias de noticias. Un contenido que se corta y pega y que en todos los soportes es exactamente el mismo. Un mismo contenido por el que cada medio existente compite.

Por otra parte, hay prácticas como el personalizar partes del site duplicando los contenidos en distintos subdominios o en distintos directorios del mismo dominio. El inconveniente que presentan es que devalúan la potencia del dominio y la posibilidad de obtener primeros puestos por sus palabras clave, al dispersar los enlaces.

3. Posicionamiento publicitario y pago por click

◆ 3.1. Publicidad basada en resultados

Google es el medio de comunicación que más ha facturado en EEUU durante el año 2005. Más que cualquier periódico o cadena de TV. Todo ello sin tener un precio fijo o tarifa. Y sin límites de espacio disponible. Basado en anuncios sin imágenes y con menos de 70 caracteres.

La publicidad basada en el Pago por Click o Pay Per Click (PPC) es el mercado de mayor crecimiento en el sector y una de las grandes revoluciones jamás ocurridas en la comunicación comercial.

El anunciante, con un presupuesto que empieza en 5 euros mensuales, elige sus palabras, redacta sus anuncios y analiza sus resultados sin intervención de cualquier otro intermediario o comercial y con posibilidad de actuar mundialmente.

Y con sus nulos costes de producción aplana por completo el terreno de juego permitiendo que la tienda de la esquina pueda pelear con la gran superficie de tú a tú.

Y en este medio se cambia el concepto de público objetivo. En vez de basarnos en perfiles sociodemográficos pasamos a basarnos en interés en un determinado momento. Un momento donde el usuario hace, con su búsqueda, una declaración de intenciones y necesidades convirtiendo un click posterior en el mejor impacto publicitario que pueda existir.

Es, además, un soporte donde la publicidad es presentada como información y sin carácter intrusivo, consiguiendo que tanto usuario como

anunciantes queden satisfechos de su experiencia. El anunciante al recibir una visita con unos criterios y anuncios que ha definido personalmente y el usuario al encontrar relevante el servicio que le ofrece el anunciante.

En EEUU el 5% de toda la inversión publicitaria está dirigida a este soporte que hace 6 años no existía y las previsiones es que pueda alcanzar el 10% de toda la tarta publicitaria antes del 2010 gracias a que haya añadido decenas de miles de anunciantes que antes no se planteaban inversiones en marketing.

Y esta tendencia tendrá su reflejo en España que acabará asumiéndola.

Este sorprendente mercado está gobernado por Google con su plataforma AdWords, seguido por Yahoo! – que adquirió Overture y que ahora es Yahoo Search Marketing – y al que se unirá MSN con su propia plataforma antes de acabar el presente año. El resto de jugadores tiene porcentajes mínimos que aquí no merecen ser comentados.

◆ 3.2. Breve revisión de los sistemas de apuestas de AdWords y Yahoo Search marketing

Los sistemas de pago por click ofrecen unas herramientas de administración a los clientes que permiten gestionar y administrar sus campañas y presupuestos. En general, tanto Google como Yahoo! se componen de:

Configuración de campañas
Permite definir

- Ámbitos geográficos.
- Idiomas de actuación de las campañas.
- Redes donde actuará la campaña (sólo búsqueda, sólo contenido – la conocida red AdSense en el caso de Google que introduce anuncios relevantes acorde con el contenido de la página – o ambos).
- Presupuesto diario de la campaña. Cuanto está el anunciante dispuesto a gastar cada día.

Configuración de grupos de anuncios
Una vez creada una campaña se le adjudican grupos de anuncios. Estos grupos de anuncios, identificados por un determinado nombre, se componen de:

- Palabras clave en las que se postrará el anuncio. Eligiendo las mismas con diferentes posibilidades o variaciones – búsqueda exacta, búsqueda amplia, negativa…– con herramientas que ayudan a seleccionar las mismas indicando una estimación de coste y volumen de clicks.
- Creación de los anuncios que se mostrarán. Pudiendo crear varios para poder hacer test de la eficacia de los mismos.
- Definición de precios. En el caso de Overture se puede conocer exactamente la apuesta que está realizando la competencia y apostar por una posición fija, mientras que la posición de Google es una mezcla entre el coste por click dispuesto a pagar y el porcentaje de ratio de click.

Informes
Ambos sistemas facilitan todo tipo de informes que permiten analizar la rentabilidad de la inversión: número de impresiones, ratio de click, coste medio de visita, ratio de conversión en el determinado objetivo – una venta por ejemplo – ratio de click por palabra clave, por creatividad, etc.

Es importante destacar entre las redes de búsqueda y las redes de contenido de ambos sistemas. Por búsqueda, entendemos aquellos resultados donde nuestros anuncios se han mostrado cuando el usuario ha tecleado en una caja de búsqueda nuestras palabras clave objetivo, teniendo una participación activa en el proceso.

Mientras las redes de contenido son aquéllas donde nuestro anuncio ha sido mostrado después de que el contenido de la página haya sido analizado y el algoritmo del buscador haya decidido que nuestro anuncio es relevante para el mismo. En el caso del contenido el usuario tiene una actitud más pasiva donde está navegando en búsqueda de información u entretenimiento y nuestros anuncios no son vistos con toda la percepción informativa de la búsqueda.

Habitualmente tanto los costes de los pagos por click, como los ratios de click y de conversión son muy inferiores en las redes de contenido.

4. El futuro de los buscadores

¿Dónde va el mercado de búsqueda? ¿Cómo se comportarán los usuarios? ¿Cuáles son los jugadores que aparecerán en el futuro inmediato?

Los buscadores ganarán en potencia y en ubicuidad permitiendo ser usados desde cualquier soporte y cruzando datos geográficos o basados en experiencias de usuario. Pero tratar de adivinar el futuro en una industria que tiene menos de una década de vida en la que aparecen nuevas tecnologías cada día es, cuanto menos, una temeridad algo presuntuosa.

Sí es evidente que aunque veamos muchas novedades con aspecto de tener gran interés y posibilidades en el futuro – lanzamientos de los que Google es el principal exponente – cuando llegan los datos sobre su uso sorprenden por su baja penetración. Así, esta es la tabla de los 20 principales dominios y servicios de Google y lo que representa cada uno de ellos en penetración de mercado. Muchos de los sectores donde se espera una transferencia de aplicaciones de escritorio o completas novedades tienen todavía participaciones en el mercado muy bajas y son mercados a la vez muy segmentados.

Rankings of Top 20 Google Domains
Week Ending 5/13/06

Rank	Name	Market Share
1	Google	79.98%
2	Google Image Search	9.54%
3	Google Mail	5.51%
4	Google News	1.49%
5	Google Maps	0.82%
6	Froogle	0.46%
7	Google Video Search	0.45%
8	Google Groups	0.43%
9	Google Scholar	0.27%
10	Google Book Search	0.25%
11	Google Earth	0.22%
12	Google Desktop Search	0.18%
13	Google Directory	0.10%
14	Google Answers	0.09%
15	Google AdWords	0.07%
16	Google - Local	0.05%
17	Google Finance	0.03%
18	Google Calendar	0.01%
19	Google Talk	0.01%
20	Google Labs	0.01%

Apuntemos unas líneas maestras hacia donde creemos que el mercado se dirige.

◆ 4.1. Especialización

Surgirán buscadores verticales especializados en tipos de búsqueda o tipos de publicación.

Como ejemplos inmediatos está el énfasis en las búsquedas locales cuya importancia será creciente y que se disparará en el momento en que la posición del usuario sea un criterio determinante. Google será el proveedor de acceso a Internet Inalámbrico en San Francisco (EEUU) ofreciendo el mismo de manera gratuita y financiado por publicidad que tendrá un acentuado carácter local, dado que se presentará publicidad acorde con la posición física del usuario – imagínese la eficacia de un anuncio a la hora de comer del restaurante que tiene en la esquina –.

O los buscadores de libros – Amazon y Google –, o los identificadores de imagen, o los que buscan sólo contenidos con licencias creative commons, o metabuscadores como Dogpile.com que agregan a los principales motores bajo un mismo interface o servicios como Yahoo Answers, donde son los propios usuarios los que dan respuestas…

Igualmente, tendrá una creciente importancia la búsqueda móvil que en su mejor momento WAP resultó un cierto fiasco y que con la tecnología, pantallas y actitudes actuales tiene grandes posibilidades.

◆ 4.2. Buscadores sociales

Como en el caso anterior, su base de datos se alimenta de las suscripciones a los archivos RSS o las solicitudes de agregación que llegan desde la página que se quiere añadir.

Son buscadores sociales o agregadores sociales de contenido, especializados en noticias o en temáticas concretas donde las más importantes son las presentadas, son las más votadas por los usuarios.

Digg fue el primer ejemplo que en España imitó meneame.net y que ahora ha seguido fresqui. Menéame ha hecho público su código permitiendo que cualquiera pueda montar su buscador especializado en la temática que desee.

◆ 4.3. La importancia creciente de los Blogs: Technorati

Los blogs son una de las revoluciones silenciosas de Internet que al ser una forma de publicación tan rápida, eficaz y sencilla y tener una difusión basada en la suscripción RSS, están permitiendo un cierto tipo de periodismo social y amateur que en muchos casos sorprende por su frescura y calidad.

Technorati es el gran ganador en este sentido. Se trata de un buscador especializado en Blogs que indexa en la actualidad más de 40 millones de Blogs y los actualiza por los mensajes o que les envía el propio sistema de publicación del usuario – RSS y ping – lo cual le evita todos los costes de ancho de banda, robots y almacenamiento de páginas innecesarias. Y además consigue información con un nivel de actualización que no tiene la base de los grandes buscadores.

Sus búsquedas se basan en el contenido, pero también en las etiquetas con las que cada usuario categoriza la página o post de blog en cuestión.

Si se quiere saber de qué habla la blogosfera –cada vez más grande e influyente–, Technorati es el medio.

Las etiquetas de Technorati permiten crear nubes de las mismas donde el tamaño del texto indica la importancia de la misma y al ser información enviada cada vez que se publica, permite, como en esta imagen, presentar el contenido de última hora – lo cual no consigue ningún buscador tradicional –.

◆ 4.4. Etiquetas o tags. Búsquedas basadas en la experiencia del usuario

El conocimiento social puede ser muy superior a la experiencia conseguida a través de máquinas. La participación de los usuarios es el eje de una nueva revolución en Internet que se ha dado en llamar Web 2.0 y que, en general, permite que los usuarios sean capaces, a través de la comunidad que crean, de definir cuáles son los contenidos destacados, las ideas más importantes y dar opiniones sobre cualquiera de los elementos que pueda aparecer en Internet: un resultado, una imagen, un vídeo, una noticia…

Sobre esta idea se asientan multitud de proyectos que están teniendo éxitos sorprendentes. Y esta idea se combina con el hecho de abrir los contenidos a otros desarrolladores para que creen sus propias aplicaciones logrando que, aquellos huecos o mejoras que no fue capaz de hacer la empresa, los hagan los propios usuarios. Pongamos algunos ejemplos.

Flirck
Un organizador de fotos donde todo usuario puede subir sus imágenes, compartirlas, valorarlas, asignarle etiquetas y colgar sus álbumes en cualquier otra página. Así, otros usuarios han programado aplicaciones que buscan sólo en determinadas etiquetas creándose buscadores especializados de fotos de toda categoría imaginable (adquirida por Yahoo).

YouTube
El Flirck de los vídeos. Una página que en menos de 1 año está posicionada como la referencia de subida y búsqueda de imágenes, que está permitiendo a particulares un canal de difusión de enorme potencia a coste cero. Con todo tipo de herramientas sociales – votaciones, opiniones, visualizaciones, etiquetado, … – ha machacado a Google Video. Y como ejemplo de su poder en nuestro país, no hay más que decir que ha sido la base de difusión de El Koala.

Del.icio.us
Una web donde puedes añadir tus páginas favoritas, insertarle etiquetas y comentarios, y compartir dichas listas con otros usarios. Además, permite hacer búsquedas con criterios como las veces que dicha página ha sido añadida o el volumen de etiquetas que existen con una determinada palabra, indicando de esa manera las tendencias o importancia que un determinado tema tiene en la red.

Gennio

Un proyecto nacional que se basa en la valoración que los usuarios dan a los resultados de búsqueda y el etiquetado posterior que éstos hacen de dichos resultados. Un buscador social que para que su funcionamiento sea eficaz debe alcanzar una importante masa crítica y en el que se crean super usuarios capaces de tener más valor con sus opiniones y valoraciones dependiendo de las votaciones que ese usuario alcance.

Gennio permite a cada usuario modificar título, resumen y palabras clave con la que identificar una página, creando una inteligencia colectiva de todo ello.

"La importancia de la participación"

Internet está volviéndose a pasos agigantados un medio social donde cada usuario tendrá múltiples relaciones y participación en forma de blogs, etiquetas, comentarios, favoritos, imágenes, valoraciones, opiniones… cuando estos sistemas alcanzan masa crítica suficiente son los propios usuarios quienes definen el contenido de la red y los resultados de búsqueda.

Esta tendencia se hará cada vez más creciente y debe ser uno de los puntos que todo profesional de la red siga y trate de mantener e integrar en su estrategia.

5. El mundo editorial y los buscadores

◆ 5.1. La relación entre editores y buscadores

Dos entornos condenados a entenderse.

Las editoras de contenido ofrecen el mismo para ser indexado alimentando la calidad de los resultados de búsqueda. Contenido que los buscadores indexan para presentar sus resultados y dar satisfacción a sus usuarios. Los editores de contenido así lo permiten, pues de esta manera reciben tráfico cualificado.

Una simbiosis aparentemente perfecta.

¿Sólo aparentemente?

Así parece ser. Al igual que los equipos de fútbol cobran por sus retransmisiones – contenido – que les permite a las televisiones traer audiencia y con ella vender publicidad, los editores empiezan a hacerse la misma pregunta.

¿Qué utilidad tiene un buscador si no tiene el contenido de mejor calidad, editado profesionalmente por equipos bien formados y pagados? ¿Por qué los buscadores están llevándose una buena tarta del mercado publicitario simplemente presentando los contenidos de los demás? ¿Por qué una empresa editora tiene que soportar unilateralmente todos los gastos de generación de contenidos para que otras se beneficien de ello?

¿Es esto realmente una simbiosis donde ambas partes ganan o hay una que hace el esfuerzo y otra que con mucho menos, obtiene proporcionalmente más beneficio?

Esta discusión ya está activa en la red. De hecho, ya ha habido los primeros casos de webs especializados que han incluido instrucciones en el archivo "robots.txt" para que los buscadores no entren en sus páginas ni indexen los resultados. Estas webs consideran que:

- El tráfico que le traen los buscadores no es de calidad.
- Al tratarse de webs de referencia, sus visitantes ya saben que deben acudir a las mismas para mantenerse al día.
- Su crecimiento orgánico es menor, pero de muchísima mayor calidad. Los nuevos usuarios que llegan lo hacen por recomendaciones, links de otros webs o comentarios de usuarios.
- Implica un ahorro importante de ancho de banda, dado que los robots prácticamente tumbaban los servidores, obligando a dimensionar sistemas sólo para la visita de un solo usuario: la araña.

Algunas asociaciones de editores ya han sugerido que los buscadores deberían compartir o pagar por indexar sus resultados y puede que sea el primer paso en ese camino.

Pero mientras tanto el status quo es el siguiente: los editores ponen los contenidos y los buscadores les mandan el tráfico. Y cuanto más tráfico tengamos más posibilidades tendremos de convertirlo en ingresos.

◆ 5.2. El ejemplo de 20minutos.es

Aunque todo el mundo editorial – incluyendo Elmundo.es o Elpais.es – han hecho grandes mejoras en trabajo de optimización, uno de los ejemplos de mayor calidad es 20minutos.es que en un plazo muy breve de tiempo ha conseguido resultados espectaculares.

Haciendo un análisis de su estructura ha conseguido organizar un website magníficamente optimizado con una gran estrategia de captación de enlaces y logrando que estas sean herramientas de utilidad para el usuario, además de para el buscador.

Expliquemos con detalle algunos de dichos puntos.

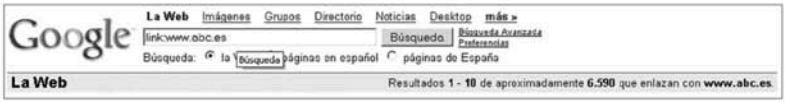
Maquetación

Una magnífica maquetación con CSS que separa contenido y presentación con gran calidad. Así es como el usuario ve la noticia

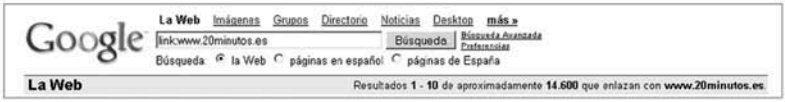
Y así es como la ve el buscador con su etiqueta H destacando un código limpio y jerárquico.

Captación de links entrantes

En un plazo muy inferior a otros soportes (como el ejemplo de ABC) 20 minutos ha logrado duplicar el número de enlaces entrantes.



ABC cuenta con 6.500 links entrantes.



Mientras 20 minutos tiene 14.600.

Las estrategias han sido:

- Potenciar directorios y no subdominios. Esto es: www.20minutos.es/murcia en vez de Murcia.20minutos.es
- Hacer un guiño a la comunidad de la blogosfera con los premios 20 minutos. Estos premios le han permitido contar con multitud de enlaces entrantes y los incentivos de los propios blogs para ser votados enlazando hacia el premio.

Un sencillo premio que se convierte en un acto de relaciones públicas y a la vez capta cientos o miles de enlaces entrantes con muchas palabras clave y de diferentes dominios y direcciones IP.

Creación de archivo o mapas

Desde nuestro punto de vista uno de los mapas más eficaces que hemos visto tanto por ser muy útil para los buscadores como por ser una gran herramienta para los usuarios.

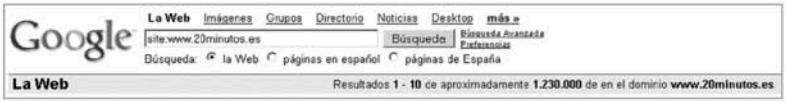
La sección archivo está estructurada cronológicamente por meses y años con acceso a cada uno de los días.

Y después cada día presenta una lista de todas las noticias del mismo clasificada jerárquicamente e incluyendo el titular como enlace lo que potencia dichas palabras clave.

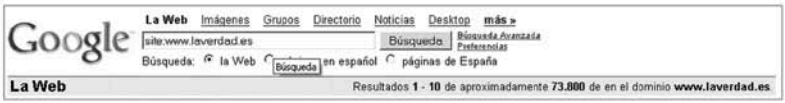
Número de páginas indexadas

Y este mapa o archivo conjuntamente con la estrategia de Blogs les ha llevado a tener más de un millón de páginas relevantes indexadas a pesar de su corto periodo de tiempo en la red.

Así, 20 minutos cuenta con 1.200.000 páginas indexadas que son, al fin y al cabo, puertas de entrada a su página.



Mientras que laverdad.es cuenta con 73.000 a pesar de llevar mucho más tiempo online.



Búsquedas convertidas en páginas estáticas

Al hacer una búsqueda, ésta se convierte en una página estática con las posibilidades de indexación que eso permite.

Nótese que el término que se incluye en la caja de búsqueda aparece en el resultado convertido en un directorio de la URL. De esta manera, se podría indexar las búsquedas más populares una vez convertidas en páginas estáticas y cacheadas y generar un número infinito de las mismas (tal y como hace Kelkoo en el ejemplo antes mencionado).

"En el momento de generar contenido de búsquedas: prudencia"

Resulta tentador pensar en las infinitas posibilidades que permite generar contenido acorde con las búsquedas incluidas por los usuarios –como Kelkoo- o las pensadas por los propios editores.

No obstante, sugerimos que esta técnica sea usada prudentemente, pues si se abusa puede caer en la generación de contenido duplicado y ser penalizado por ello.

Marketing rentable: Buscadores

anetcom

C/ Luis Vives 6, 4º, 12ª
46003 Valencia
Tel. 96 392 39 16
Fax 96 392 40 83
informacion@anetcom.es
www.anetcom.es

GENERALITAT VALENCIANA
CONSELLERIA D'EMPRESA, UNIVERSITAT I CIÈNCIA