

Marketing de buscadores. El reino de la relevancia

POR JOSÉ MANUEL REDONDO

Un poco de historia

Hablar hoy día de buscadores es hablar inevitablemente del gigante de Mountain View, pero no siempre fue así. La historia de los buscadores de Internet comienza hace casi 20 años con el nacimiento de Archie. Archie fue creado por Alan Emtage, un estudiante de la Universidad McGill en Montreal, y recibe su nombre como abreviatura de *archives* (archivos). Básicamente se trataba de una gran base de datos de ficheros en los que se podía buscar cualquier documento utilizando palabras contenidas exclusivamente en el nombre del propio archivo. Recordemos que por aquella época Internet tan sólo extendía sus tentáculos entre algunas privilegiadas universidades alrededor del mundo. Además, Sir Tim Berners Lee todavía no había puesto en marcha la *World Wide Web* (WWW), por lo que los ficheros se compartían alojándolos en sitios mediante FTP (*File Transfer Protocol* o Protocolo de Transferencia de Ficheros), lo que hacía de Archie un sistema eficaz sólo para trabajar en ámbitos reducidos.

En 1991 Mark McCahill, otro estudiante universitario de Minnessota, crea Gopher (llamado con este curioso nombre en honor a la mascota de la Universidad de Minnessota). A diferencia de Archie, Gopher permitía indexar documentos de texto plano. Se trataba de un gran avance, ya que las búsquedas extendían su potencial a cualquier parte del documento y no sólo al nombre del fichero. Precisamente gracias a naturaleza textual de Gopher la mayoría de los sitios así creados pudieron convertirse fácilmente en sitios web con el nacimiento de la WWW.

De la mano de Gopher nacieron Veronica (*Very Easy Rodent-Oriented Net-wide Index to Computerized Archives*) y Jughead (*Jonzy's Universal Gopher Hierarchy Excavation And Display*), dos programas que buscaban información dentro de los índices de los sistemas Gopher.

En 1993 un estudiante del MIT, Matthew Gray, introduce el concepto de robot al crear el *World Wide Web Wanderer*. Su objetivo inicial era medir el tamaño de Internet a través de la creación de un proceso automatizado para contar los servidores que la componían. A través de sus viajes por Internet, el *Wanderer* (viajero) creó la primera base de datos de direcciones

de Internet (URL), llamada *Wandex*.

El inicio de los robots

Merece la pena hacer un pequeño inciso sobre este asunto de los robots. Siguiendo la definición de Robotstxt.org, «Un robot es un programa que automáticamente recorre la estructura hipertextual de la web recuperando un documento y recuperando, a su vez, todos los documentos que están referenciados en éste». Es decir, los robots recorren Internet recogiendo información del contenido que se encuentran en una página web (es decir el contenido de una URL única, por ejemplo *www.elmundo.es/economia*) y después siguen cada enlace que encuentran en dicha página hacia otras páginas web, para recorrer el contenido de éstas y así sucesivamente.

Es importante entender la diferencia entre un sitio web (*website*) y una página web (*webpage*). Un sitio web estará alojado normalmente bajo un dominio único (*www.loquesea.com*) y estará compuesto por múltiples páginas web que se alojarán en el mismo servidor, cada una de ellas con una URL única (*www.loquesea.com/quienessomos*). Es un error muy extendido el referirse a una página web cuando en realidad se quiere describir un sitio web. Por proporcionar un dato que ilustre la importancia del matiz: la última estimación de sitios web en el mundo supera los 186 millones, mientras que el dato que Google proporcionaba de manera oficiosa en su *blog* corporativo mencionaba un número superior al trillón americano (nuestro billón de toda la vida) de URL únicas o páginas web indexadas.

Retomando el hilo de la breve historia de los buscadores, es más o menos a partir de ese momento (1993) cuando los inversores de la soleada California empiezan a intuir el potencial de los buscadores de Internet y emerge un sinfín de proyectos. En ese mismo año, seis estudiantes de la Universidad de Stanford comienzan *Excite*, aportando a su propuesta sofisticación estadística en el tratamiento de la relación entre las palabras halladas en el contenido. *Excite* se convirtió en uno de los buscadores más populares en los primeros tiempos de Internet y hoy día forma parte del proyecto *Askjeeves*.

En 1994 David Filo y Jerry Jang comienzan a agregar sus sitios favoritos de Internet en un mismo lugar, creando un directorio de sitios web al que denominaron Yahoo. Yahoo enriquecía su propuesta aportando una breve descripción de cada uno de los sitios que recogía. Dos años más tarde aparecería Inktomi, que luego sería adquirido por Yahoo. Ese mismo año, Webcrawler ofrecía la primera opción que indexaba todo el contenido de cada página y no sólo ciertos elementos o partes del mismo.

En 1995 nace Altavista, que llegaría a convertirse en el buscador líder de Internet durante mucho tiempo. Altavista aportaba al mercado de las búsquedas la capacidad de usar lenguaje natural (escribir lo que se busca como si se lo preguntásemos a otra persona y no a una máquina), así como otras técnicas avanzadas.

Por primera vez un buscador ambicionaba no sólo ofrecer búsquedas de páginas web mediante su contenido, sino que ofrecía la posibilidad de hacer búsquedas de otro tipo de contenidos como imágenes, música o vídeos.

El nacimiento de un gigante

En el año 1995 Larry Page, por entonces un estudiante de *Computer Science* en la Universidad de Stanford, buscaba un tema para su tesis doctoral. Siendo hijo de un profesor de esa misma disciplina, Page era más que consciente de la importancia de su proyecto para su futuro profesional. Entre los distintos temas que estaba considerando para la tesis, la incipiente WWW despertaba su curiosidad.

Para Page era obvia la importancia de seguir los enlaces (*links*) de un lugar a otro; lo que parecía no ser tan obvio era la relación inversa, es decir, el poder conocer qué *sites* le enlazaban a uno. En la cabeza de Page estaba la referencia al sistema de publicaciones académicas tradicionalmente sustentado sobre un modelo de citación. Los autores de una publicación generalmente citaban a otros autores como referencias para sustentar argumentos de una determinada hipótesis. El número de citas que un autor recibía determinaba de alguna manera su relevancia. Así, Page pensaba que los enlaces, en definitiva, no diferían mucho del concepto de cita académica; por lo tanto, una página web que recibiera muchos enlaces apuntándola debería ser más importante que otra que no recibiera ninguno. De esta manera comenzó su proyecto, al que denominó *Backrub*.

Page pronto empezó a darse cuenta de lo hercúleo de la tarea. No sólo era una cuestión de medir cuántos enlaces recibía una determinada página, también era de suma importancia saber la importancia de cada una de esas páginas, medida igualmente por el número de enlaces que éstas recibían. Y aquí es donde hace su aparición Brin.

Sergei Brin llegó a EEUU a los seis años de edad procedente de Rusia. Era reconocido como un auténtico genio matemático desde muy joven y no se puede negar que llevaba las matemáticas en sus genes: sus padres eran uno ingeniero en la NASA y otro profesor de matemáticas en la Universidad de Maryland.

Juntos, Page y Brin idearon un sistema que permitía dar mayor importancia a aquellos sites que recibían gran cantidad de enlaces y castigar a aquellos que, por el contrario, no los recibían. A este sistema decidieron llamarle *PageRank*, en honor del propio Page.

Pronto empezaron a entender la importancia de lo que habían concebido. Su sistema introducía por primera vez un elemento -los enlaces- que aportaba a la simple presencia de palabras dentro de un contenido una nueva forma de categorizar los sitios web en función de su importancia. De esta forma empiezan a ser capaces de ofrecer respuestas mucho más relevantes a una misma pregunta lanzada a sus entonces competidores Excite o Altavista.

Es en ese año, 1997, cuando nace el que hoy es el mayor buscador de Internet y el mayor negocio que ha dado el mundo digital hasta la fecha.

El marketing de buscadores

El primero en utilizar el concepto de *marketing* de buscadores (*Search Engine Marketing*) fue Danny Sullivan, en un artículo publicado en *SearchEngineWatch* el 5 de noviembre de 2001. Tomando a Sullivan como referencia, una definición amplia del *marketing* de buscadores podría ser la siguiente: «El *Search Engine Marketing* (SEM) o *marketing* de buscadores es una disciplina del marketing cuyo objetivo es el de ayudar a un anunciante a alcanzar el mejor posicionamiento posible en las paginas de resultados (SERP, *Search Engine Result Pages*) de los motores de búsqueda en un conjunto de palabras clave (*keywords*) relevantes».

Siguiendo la definición de la Organización Profesional de *Marketing* de Buscadores (*Search Engine Marketing Professional Organization*, SEMPO), el SEM se dividiría a su vez en diferentes disciplinas: el posicionamiento de pago o enlaces patrocinados, la optimización orgánica de buscadores (*Search Engine Optimization*, SEO), la publicidad contextual, la optimización de activos digitales y el pago por inclusión. Por su importancia, en este artículo nos vamos a concentrar en las dos primeras disciplinas.

El posicionamiento de pago o enlaces patrocinados

EL primer concepto dentro de esta disciplina del *marketing* comienza en 1996, con un proyecto denominado *Open Text*. Sin embargo, no sería hasta dos años más tarde cuando el buscador GoTo.com lanzaría con éxito un programa, SEM.

Según cuenta John Batelle en su libro *The Search*, Bill Gross -uno de los fundadores de Idealab, la incubadora de negocios donde nació GoTo.com-, demostró el 'gen emprendedor' desde su más tierna infancia. Según parece, Bill Gross se dedicaba a comprar caramelos en una tienda fuera de su barrio, donde el precio era sensiblemente inferior al de las tiendas de su entorno, para revenderlos a sus amigos y vecinos obteniendo un margen interesante con la diferencia entre el precio de compra y el de venta. Un modelo que en inglés recibe el nombre de *arbitrage*.

Gross empezó entonces a comprar tráfico a bajo coste mediante acuerdos con terceros, como el buscador Inktomi, con la esperanza de revender ese tráfico obteniendo una ganancia por la diferencia, como ya había hecho antes con las chucherías. Gross acertaba al pensar que tenía todo el sentido ofrecer a un anunciante la posibilidad de aparecer en una posición preferente en los resultados de búsqueda cuando un usuario buscase palabras genéricas de algunas categorías que podían atraer el interés de los clientes potenciales de un anunciante y, de paso, el dinero de este último.

Para hacer su oferta aún más atractiva, Gross ofreció un modelo donde el anunciante no necesitaba asumir ningún riesgo; tan solo tenía que pagar cuando un usuario hacía clic en su anuncio. Había nacido el *pay-per-click* (pago por clic), un modelo completamente transgresor con respecto a los que utilizaban en aquel momento competidores como Yahoo o Microsoft, basados en el CPM (pago por cada mil impactos de mi publicidad). El anunciante que estuviera dispuesto a pagar más por un clic ganaría su derecho a aparecer en primeras posiciones.

Posteriormente, en 2001 GoTo se transformaría en Overture y finalmente acabaría siendo vendido a Yahoo en 2003 por 1,63 billones de dólares.

En octubre de ese mismo año, si entrabas a realizar una búsqueda en Google podías leer el siguiente mensaje: *'Have a credit card and 5 minutes? Get your ad on Google today'*. (¿Tiene una tarjeta de crédito y 5 minutos? Consiga su anuncio en Google hoy). De esta manera comenzaba a promocionar su programa *Google Adwords*. *Adwords* tomaba como referencia el modelo creado por GoTo, pero sólo ofrecía como fórmula de contratación el CPM y reducía la capacidad creativa a las dimensiones de una serie limitada de caracteres en un anuncio de texto.

En el año 2001 Google ya sumaba unos ingresos cercanos a los 85 millones de dólares; sin embargo Overture (el antiguo GoTo) le llevaba claramente la delantera, con más de 288 millones. En 2002 Google introduce un cambio importante en su modelo publicitario *Adwords*, al crear un mercado abierto en el que el anunciante podía, mediante un sistema basado en pujas, ganar su derecho a la primera posición de los resultados de búsqueda. El modelo toma tanto esta característica como el pago por clic del exitoso modelo de Overture.

Sistema de cómputo por clic

Sin embargo, *Adwords* incorpora una importante novedad. ¿Qué ocurre si un anunciante, gracias al valor de su puja, aparece en primera posición para un término para el que no es el más realmente relevante para la búsqueda? Probablemente lo que ocurrirá será que se reducirán las posibilidades de que su anuncio consiga un clic. Y si el modelo se basa en pago por clic, ¿tiene realmente sentido que el anunciante que más pague sea el que aparezca en primera posición? Parece evidente que no es así. Por eso Google incorpora el CTR (*Clickthrough Rate*, o ratio de clic, que se expresa como un porcentaje de los clics que obtiene un anuncio por cada 100 veces que aparece) como ingrediente en la fórmula, de manera que la probabilidad de que un usuario haga clic -medida en términos del CTR que obtienen sus anuncios-, ayudará a éste a aparecer en las primeras posiciones. Google consigue así dos objetivos: para el anunciante un mejor retorno de su inversión y para Google un mayor volumen de clics o, lo que es lo mismo, mayores ingresos.

En palabras del propio John Batelle, «la decisión de Google de incorporar el CTR como un factor del ranking de anunciantes forzó la creación de una economía de relevancia y beneficios en un modelo de pago por clic».

Overture acabó denunciando a Google y a FindWhat por infracción de los derechos derivados de su patente 6.269.361 relativa a las tecnologías desarrolladas por éste en su modelo de pago por resultados y la plataforma de pujas que lo sustentaba. En agosto de 2004, Google y Yahoo (que por entonces ya había adquirido Overture) entierran el hacha de guerra, alcanzando un acuerdo por el que Google hacía a Yahoo un único pago en forma de un paquete de más de 2,7 millones de acciones de la compañía a cambio de la licenciamiento de la tecnología.

Sin embargo, Google no ha dejado de sofisticar su plataforma publicitaria desde entonces, manteniendo en su cabeza siempre un mismo objetivo: conseguir el máximo retorno del número de clics recibidos de los enlaces patrocinados en sus páginas de resultados.

El Índice de Calidad (QS)

De esta sofisticación surge el concepto del Índice de Calidad (*Quality Score*, o QS). En un algún punto, Google se da cuenta de que el CTR, aunque es una buena referencia, no es el único indicador que incide en la probabilidad de que un usuario haga clic en un anuncio y posteriormente cumpla el objetivo fijado por el anunciante para la campaña (generar un *lead*, una venta...); al fin y al cabo no deja de ser un indicador que sólo se puede medir a posteriori. Pero ¿habría alguna manera de saber de antemano si un anuncio tendrá éxito para una determinada búsqueda?

Es entonces cuando nace el Índice de Calidad (QS), un índice que comprendía una serie de elementos de una puja para una determinada búsqueda. El QS sigue tomando en consideración el CTR, pero va mucho más allá, incorporando elementos que pretenden ayudar a predecir su probabilidad de éxito. De esta manera, Google introduce en la fórmula del *Quality Score* factores como la relación entre el anuncio y la palabra clave por la que se puja, el contenido de la página de destino y otros relativos a la página de destino, como el tiempo de carga, el uso de estándares *w3c*,...

Hasta aquí el modelo funciona de manera que me atrevería a decir brillante. Sin embargo, no puedo dejar de hablar de un 'lado oscuro' del Quality Score: cuando el modelo se basaba en el CTR, era evidente para un anunciante cómo podía afectar a su competitividad en Google Adwords la respuesta de los usuarios a su propuesta; de este modo el anunciante podía variar elementos de su propuesta para mejorar el CTR -y por tanto su competitividad y su posición en el ranking-. Sin embargo, la fórmula secreta del Quality Score se cuece en las cocinas de Google, lejos de ojos indiscretos, y ya no resulta tan evidente cuáles son los ingredientes, ni la cantidad y orden en que se combinan. Pero el QS es un elemento clave para fijar el precio de entrada de un anunciante para una determinada puja, es decir, el CPC (coste por clic) mínimo por el que Google permite a un anunciante pujar para una determinada palabra clave. En un modelo de puja teóricamente transparente convive un factor -el QS- muy poco transparente y que, además, determina de manera directa los precios de cada puja y, por lo tanto, lo que Google ingresa por cada clic.

Otros elementos del Adwords

El modelo de *Adwords* tiene otros elementos además del QS y del coste por clic. Pasaré a describirlos brevemente:

- La estructura de la campaña de *Adwords*: una acción en *Google Adwords* se estructura en torno a los siguientes elementos: campañas, grupos (*ad group*) y anuncios; todos ellos son elementos importantes para el rendimiento de la campaña.

A nivel de campaña se controlan los presupuestos. Se establece un presupuesto máximo diario, es decir, lo que un anunciante está dispuesto a gastar diariamente para esa campaña. A nivel de grupo se recogen los conjuntos de palabras clave y sus concordancias (amplia, frase o exacta). Bajo cada grupo se reúnen todas las palabras clave que están relacionadas con un mismo concepto de campaña.

Los anuncios se establecen a nivel de grupo, es decir, cada grupo de palabras tendrá asociados una serie de anuncios que aparecerán cuando el usuario busque alguna de esas palabras. Los anuncios están limitados en cuanto a su contenido por la política editorial de Google y en cuanto a su extensión por el número máximo de caracteres por cada una de las líneas (titular, cuerpo y URL visible).

Optimización de motores de búsqueda o *Search Engine Optimization*

Probablemente cualquier usuario de Google será perfectamente capaz de reconocer y diferenciar los resultados de pago de los naturales u orgánicos. Si tiene alguna duda, no tiene más que hacer una búsqueda en Google y podrá comprobar que los resultados de pago aparecen normalmente en la columna de la derecha y hasta en los tres primeros resultados de la parte superior de la columna izquierda. Ambos están claramente identificados con el mensaje 'enlaces patrocinados', y los de la parte superior además tienen un tenue fondo color tierra. El resto de los resultados de la columna izquierda son orgánicos. Se denominan resultados naturales u orgánicos a aquellos que son resultado del proceso de indexación del universo web por parte de los sistemas de Google.

El proceso que sigue Google para ello es más o menos el siguiente: un robot o araña recorre las páginas que encuentra a su paso, siguiendo cada enlace de esa página a otras y así sucesivamente. Las arañas de Google reciben el nombre de *Googlebot*. En realidad, lo que hace *Googlebot* es solicitar al servidor donde está guardada una determinada página web su contenido; entonces lo escanea en busca de los enlaces (*hyperlinks* o *links*) que contenga y solicita de nuevo al servidor correspondiente la página enlazada. Google asigna un número a cada una de esas páginas, con el objeto de poder referirse a ellas posteriormente. Pero todas estas páginas todavía no son 'buscables', es necesario un proceso de indexación.

El proceso de indexar el contenido de las páginas recogidas por el robot consiste en hacer

una relación de cada página que contiene una determinada palabra clave. Por simplificar, usaré un ejemplo con una única palabra clave, 'cámara'. Supongamos que 'cámara' aparece en las páginas 12, 34, 56, 64 y 99. Si un usuario teclea la palabra 'cámara' en el cajetín de búsqueda de Google, lo primero que ocurre es una solicitud a la base de datos para saber en cuáles de sus páginas aparece la palabra 'cámara'. A continuación tratará de establecer el *ranking* de relevancia de cada una de esas páginas para dicha búsqueda y, en función de éste, decidirá el orden en que aparecerán en los resultados de búsqueda.

PageRank y otros criterios

Para establecer cuál es más importante o más relevante para la búsqueda, Google emplea varios factores en la fórmula. El *PageRank* es uno de ellos. Antes he descrito el concepto y el origen del *PageRank* pero merece la pena adentrarse un poco más en este concepto.

El *PageRank* valora dos cosas: el número de enlaces que recibe nuestra página y la importancia de cada uno de dichos enlaces. Cualquier página de Internet tiene un *PageRank* que se representa, al menos a nivel público, como un valor numérico entre 0 y 10. Conocer el *PageRank* de cualquier página es muy sencillo, tan sólo necesitamos instalar la barra de Google en nuestro navegador favorito y nos proporciona el valor del *PageRank* de cualquier página que visitemos.

El valor real del *PageRank* suele ser del orden de miles de unidades. Sin embargo, Google coge su valor logarítmico para hacerlo visible a los usuarios. La base de este logaritmo es desconocida. Pero para ilustrarlo podemos seguir el ejemplo que utiliza el *blog* de Dirson. Por ejemplo, para una base 7 tendríamos los siguientes valores (ver tabla 1).

TABLA 1. CORRESPONDENCIA ENTRE EL PAGERANK DE LA BARRA DE HERRAMIENTAS DE GOOGLE Y EL REAL

PageRank de la Barra	PageRank Real
0	0 - 3
1	4 - 19
2	20 - 110
3	111 - 907
4	908 - 6.351
5	6.352 - 44.458
6	44.459 - 311.209
7	311.210 - 2.178.466
8	2.178.467 - 15.249.262
9	15.249.263 - 106.765.603
10	> 106.765.608

Como se puede observar, la dificultad para pasar de un nivel al siguiente es creciente, es decir, pasar de un nivel 1 a un nivel 2 es relativamente sencillo, pero pasar de un nivel 6 a un nivel 7 requiere de un grandísimo esfuerzo.

Además del *PageRank* para calcular el *ranking* de los resultados, Google incorpora otro tipo de criterios que sólo Google conoce a ciencia cierta. Sin embargo, algunos son relativamente conocidos, como por ejemplo la ubicación de la palabra en el contexto del contenido. Para Google no es lo mismo si la palabra aparece una sola vez o varias veces en el texto, si aparece en el cuerpo o en el titular, si aparece en negrita o en cursiva, si además forma parte del dominio, si forma parte de la etiqueta *Title* del HTML y un largo etcétera de variables.

¿Y por qué es tan importante un buen posicionamiento orgánico para un anunciante? En primer lugar hay un argumento racional incontestable: el coste. La optimización SEO suele ser más rentable que la SEM, aunque ambas tienen un excelente ROI. Sin embargo, diversos estudios y la propia lógica demuestran que la discusión no debe estar en la elección entre las dos opciones, sino que es en la combinación de ambas donde se encuentra el mayor retorno. Detrás de esta afirmación hay una lógica aplastante: si en la página de búsqueda el usuario tiene dos opciones para hacer clic en nuestro anuncio, las posibilidades de que lo haga se multiplican.

Algunas cifras sobre el marketing de buscadores

Un informe publicado recientemente por eMarketer (*Search Marketing Trends: back to basics*) muestra una tendencia incipiente al crecimiento de la inversión de los anunciantes en SEO. En el mercado de EEUU la inversión en SEO llegará a alcanzar un total del 16,5 por ciento sobre el total de la inversión en *marketing* de buscadores desde el 12,9 por ciento que está representando durante 2009. Sin embargo, la previsión para el posicionamiento de pago muestra una cierta ralentización en su ratio de crecimiento respecto de años anteriores, pasando de un 15,3 por ciento en 2009 hasta alcanzar un 11,9 por ciento en 2013. A pesar de esta ralentización, el SEM seguirá representando el mayor volumen manteniendo la cuota del 63 por ciento del total de la inversión en *marketing* de buscadores.

En nuestro país, la evolución del marketing de buscadores, según los datos publicados por el estudio de inversión de IAB/PWC para el año 2008, muestra un destacado crecimiento de la inversión, con un 36,7 por ciento de crecimiento respecto a 2007. En 2009 la inversión del primer semestre ha mostrado, incluso en momento de clara recesión económica y con un descenso generalizado de la inversión en medios, un crecimiento del 7,9 por ciento respecto al mismo periodo del año anterior.

Si bien es cierto que existen algunas nubes en el horizonte -sobre todo aquellas provenientes de las cálidas tierras de los medios sociales- que podrían afectar a la salud del *marketing* de buscadores, todos los indicadores muestran una excelente salud del SEM y una envidiable esperanza de vida.