

Combinación de logs internos y externos en la predicción de estacionalidad de búsquedas para el rediseño de webs

Por Jorge Serrano-Cobos

Resumen: Se estudia la efectividad del análisis diacrónico de los search logs como herramienta para mejorar la arquitectura de información de sitios web. Se describe una metodología para extraer la estacionalidad de búsquedas a través del estudio combinado de search logs internos (extraídos del buscador interno) y externos (obtenidos de buscadores como Google). Los resultados apuntan a que una mejor arquitectura de información produce mejores resultados tanto de cara a los buscadores externos, como al buscador interno.

Palabras clave: Análisis de search logs, Search analytics, Minería web, Diseño de interacción, Diseño web centrado en el usuario, Arquitectura de información, Métrica web, Analítica web, Conducta de búsqueda del usuario.

Title: Combination of internal and external search log analysis for forecasting of seasonality of queries to enhance websites redesign

Abstract: Effectiveness of longitudinal search analytics is studied, in the sense that search log analysis can be used as a tool to enhance website information architecture. A methodology of analysis is described, obtaining the seasonality of queries through the combination of studies on internal search logs (from internal search engines) and external search logs (from Google). Results show that a better information architecture produces better information retrieval results, both for external or internet search engines and for internal search engines.

Keywords: Search log analysis, Search analytics, Web usage mining, Query mining, Interaction design, User centred design, Information architecture, Web analytics, User search behaviour.

Serrano-Cobos, Jorge. "Combinación de logs internos y externos en la predicción de estacionalidad de búsquedas para el rediseño de webs". *El profesional de la información*, 2009, enero-febrero, v. 18, n. 1, pp. 11-19.

DOI: 10.3145/epi.2009.ene.02



Jorge Serrano-Cobos, diplomado en biblioteconomía y documentación por la Universidad de Zaragoza, y licenciado en documentación por la Universidad de Granada, es jefe del Departamento de Contenidos de MASmedios. Es especialista en gestión de contenidos, arquitectura de información, diseño de web e intranets, marketing digital, optimización y posicionamiento en buscadores y usabilidad.

1. Introducción y objetivos

En la disciplina del diseño web centrado en el usuario existen diversas formas de conocer al usuario para diseñar o rediseñar tanto la arquitectura de información como los flujos de interacción en la realización de tareas (Allerton Institute; **Cheuk, Dervin**).

Una de las menos conocidas y sin embargo más útiles es el análisis de los *search logs* (**Angiolillo**), es decir, los archivos del servidor en los que se guardan las consultas de los usuarios en un buscador. Estos archivos guardan un auténtico tesoro (**Rosenfeld**, 2002), pues muestran las *queries* en su lenguaje natural, aportando información sobre quién (o al menos desde dónde), cuándo, y cuánto se busca algo. Nosotros usamos un método en el que combinamos el estudio de los *search logs* internos y los externos (obtenidos de buscadores tipo *Google*).

El presente trabajo intentará demostrar la utilidad de este tipo de examen para conocer la estacionalidad de las búsquedas, es decir, las épocas del año en que se buscan más ciertos conceptos. Con la información obtenida se puede rediseñar o implementar desde cero la arquitectura de información de sitios web, y por tanto mejorar la recuperación de información (**Ellis**).

“Los archivos de *search logs* guardan un auténtico tesoro pues muestran las consultas en su lenguaje natural, aportando información sobre quién, cuándo, y cuánto se busca algo”

Artículo recibido el 10-10-08

Aceptación definitiva: 30-12-08

2. Análisis de search logs internos y externos

Hay dos momentos en los que se puede hacer uso del análisis de search logs: al diseñar la estructura de un sitio web por primera vez, y cuando necesita ser rediseñada o reestructurada en todo o en parte.

En cada ocasión podemos utilizar diferentes herramientas que nos servirán para entender mejor a los usuarios, tanto a aquellos que buscan algo que potencialmente pueden encontrar en nuestro sitio web (**Bradford**), como a aquellos que una vez han llegado a él, consultan algo en el buscador interno del sitio.

Prediseño	Post diseño
TRENDS INSIGHTS ADWORDS	SEARCH LOG INTERNO GOOGLE ANALYTICS GOOGLE WEBMASTER ESTUDIOS CUALITATIVOS BUSINESS INTELLIGENCE
Externo	Interno

Figura 1. Herramientas para la analítica de las búsquedas en los webs

Existen distintos instrumentos de análisis de los *search logs* externos, entre los que destaca *Google Adwords* y *Google Insights* (una evolución de *Google Trends*) (figura 1).

Adwords permite descubrir *queries* relacionadas con una búsqueda inicial o “semilla”, que suele ser un concepto genérico, del cual desglosar palabras clave más específicas que concuerden con aquellos contenidos que vayamos a ofrecer en nuestro sitio web. Asimismo, *Adwords* posibilita ver la estacionalidad de cada palabra clave (figura 2, ver Tendencias).

Google Insights, como versión mejorada de *Google Trends*, facilita comparar la evolución de estas palabras clave para dirimir cuáles de las previamente obtenidas son más interesantes tanto para incluirlas en el catálogo de contenidos del sitio web, como para inferir el momento del año en que conviene más destacarlas en la arquitectura de información (pues cuanto más prominentes o más fácilmente visibles en el sitio web, más probabilidades tendrán de ser encontradas) (figura 3).

Al rediseñar un sitio web hay que tener en cuenta no sólo las herramientas que ayudan a conocer al usuario externo al portal, sino también al interno. En este sentido, *Google Webmasters* informa de las búsquedas

Palabras clave	Competencia del anunciante	▼ Volumen de búsquedas aproximado: noviembre	Prom. volumen de búsquedas aprox.	Tendencias de volumen de búsquedas (nov 2007 - oct 2008)	Tipo de concordancia: ?
Palabras clave relacionadas con los términos introducidos - ordenar por relevancia ?					
agua	<div style="width: 20px; height: 10px; background-color: #ccc;"></div>	2.240.000	1.830.000		Añadir ▼
aguas	<div style="width: 20px; height: 10px; background-color: #ccc;"></div>	550.000	368.000		Añadir ▼
el agua	<div style="width: 20px; height: 10px; background-color: #ccc;"></div>	368.000	246.000		Añadir ▼
agua potable	<div style="width: 20px; height: 10px; background-color: #ccc;"></div>	33.100	27.100		Añadir ▼
ciclo agua	<div style="width: 20px; height: 10px; background-color: #ccc;"></div>	27.100	22.200		Añadir ▼
contaminacion agua	<div style="width: 20px; height: 10px; background-color: #ccc;"></div>	27.100	18.100		Añadir ▼
contaminacion del agua	<div style="width: 20px; height: 10px; background-color: #ccc;"></div>	22.200	12.100		Añadir ▼
consumo agua	<div style="width: 20px; height: 10px; background-color: #ccc;"></div>	18.100	14.800		Añadir ▼
presion agua	<div style="width: 20px; height: 10px; background-color: #ccc;"></div>	18.100	14.800		Añadir ▼
agua beber	<div style="width: 20px; height: 10px; background-color: #ccc;"></div>	14.800	14.800		Añadir ▼
agua mineral	<div style="width: 20px; height: 10px; background-color: #ccc;"></div>	14.800	12.100		Añadir ▼
fuentes agua	<div style="width: 20px; height: 10px; background-color: #ccc;"></div>	14.800	14.800		Añadir ▼
tratamiento agua	<div style="width: 20px; height: 10px; background-color: #ccc;"></div>	14.800	12.100		Añadir ▼
bailame el agua	<div style="width: 20px; height: 10px; background-color: #ccc;"></div>	12.100	8.100		Añadir ▼
filtros agua	<div style="width: 20px; height: 10px; background-color: #ccc;"></div>	12.100	9.900		Añadir ▼
consumo de agua	<div style="width: 20px; height: 10px; background-color: #ccc;"></div>	9.900	9.900		Añadir ▼
dureza agua	<div style="width: 20px; height: 10px; background-color: #ccc;"></div>	9.900	8.100		Añadir ▼
filtro agua	<div style="width: 20px; height: 10px; background-color: #ccc;"></div>	9.900	8.100		Añadir ▼
fuelle agua	<div style="width: 20px; height: 10px; background-color: #ccc;"></div>	9.900	8.100		Añadir ▼
instalacion agua	<div style="width: 20px; height: 10px; background-color: #ccc;"></div>	9.900	6.600		Añadir ▼
propiedades agua	<div style="width: 20px; height: 10px; background-color: #ccc;"></div>	9.900	6.600		Añadir ▼
agua sanitaria	<div style="width: 20px; height: 10px; background-color: #ccc;"></div>	8.100	6.600		Añadir ▼
calidad agua	<div style="width: 20px; height: 10px; background-color: #ccc;"></div>	8.100	8.100		Añadir ▼

Figura 2. Sugerencia de palabras clave con Google Adwords



Figura 3. Estacionalidad de búsquedas con Google Insights

en Google en las que nuestro sitio web es encontrado, y cuáles de ellas nos aportan más porcentaje de visitas. De esa forma podemos medir el éxito del sitio en función de cuánto se encuentra a través de Google y a través de qué queries.

Por tanto podremos estudiar en detalle cuáles son las consultas en las que aparece más nuestra web entre los primeros resultados, y cuáles son las que han llevado más a los usuarios a visitar nuestro sitio (figura 4).

Si no hay correlación entre ambas cosas, ello indicaría que los snippets o extractos de la web que aparecen en las respuestas de Google a una búsqueda dada no son suficientemente pertinentes para el usuario y no

responden al concepto de *information scent* (aroma de la información) y relevancia subjetiva, por lo que no muestran visualmente una respuesta probablemente ajustada a la pregunta que el usuario está haciendo.

La conclusión lógica será que deberemos transformar los textos de la página web que sirve como respuesta concreta, así como el título y los metadatos (al menos, el metadato *description*) que suelen ser utilizados para informar al usuario del contexto en el que se ha encontrado esa palabra clave.

Asimismo deberemos tener en cuenta no sólo lo que buscan los usuarios antes de entrar en el portal sino –en el caso de sitios web ya diseñados y con una historia detrás– lo que buscan cuando han entrado.

Diversos estudios muestran ciertas diferencias entre lo que busca el usuario en un buscador, y lo que busca dentro de un sitio web (Serrano-Cobos, 2006). La bibliografía al respecto señala que los usuarios buscan más o menos veces dependiendo de cómo se haya diseñado la arquitectura de información y la interacción (Koch, Ardö, Golub, 2004) con el riesgo añadido de que el buscador interno no responda a sus necesidades de información.



Figura 4. Análisis de queries externas en Google Webmaster

El análisis de los logs internos sólo permite realizar un estudio limitado de quién, cuándo, desde dónde, y cuántas veces, buscó una palabra clave (*query*). No permite vincular ese conocimiento con la navegación de los usuarios, pero aporta una información enormemente valiosa, al descubrir el lenguaje natural usado por los usuarios.

En muchas ocasiones no habrá correlación entre la frecuencia de búsqueda de los términos más buscados en el buscador interno y los términos más buscados en los externos (como *Google*). Esto se debe principalmente a que las búsquedas realizadas en *Google* que llevan a nuestro sitio son más genéricas o están encaminadas a encontrar el portal por su propio nombre de dominio, y en cambio las búsquedas realizadas en el buscador interno son más específicas, puesto que vienen acompañadas y son modificadas por una mayor comprensión del contexto en el que se realiza la consulta, al haber hecho *browsing* por los contenidos del portal antes de realizar esa consulta (Kuhlthau, 1999; Wildemuth, 2004; Zhang, Anghelescu, Yuan, 2005).

Otras herramientas menos utilizadas pero más potentes son los sistemas de *Business intelligence* o *Data mining*, que gracias al enorme volumen de datos que pueden manejar permiten inferir predicciones en cuanto a estacionalidad de búsquedas y de comportamiento de los usuarios frente a determinadas estructuras de contenidos, e incluso modificarlas dinámicamente. Normalmente utilizan lo que se denomina un cubo de datos, formado por las coordenadas X, Y, Z, donde:

X = páginas (nivel de arquitectura de información, páginas concretas, agrupaciones de contenidos, relaciones no obvias entre contenidos visitados).

Y = personas (edad, género, gustos..., dependiendo de los niveles de privacidad que se empleen en el análisis).

Z = tiempo (cuándo se hicieron las transacciones o movimientos).

Finalmente, como “hermano menor” de estos sistemas, encontraríamos los programas de analítica web. *Google Analytics* (como otros productos de esta índole) permite trabajar con ambos logs al mismo tiempo, tanto con el *search log* como con el log de navegación. Esto es debido a que se basa en etiquetas o *tags* insertas en el html, con lo que se recuperan las palabras clave usadas en las búsquedas internas. Las diferencias entre trabajar sólo con logs o con *tags* han sido estudiadas. Siempre se encontrará mayor riqueza de información en sistemas de análisis que ayuden a combinar ambos tipos de transacción, porque si sólo nos preguntamos por las palabras más o menos buscadas, no podremos saber qué ocurrió antes o después de la búsqueda, por

qué se buscó, y lo más importante, si la búsqueda fue exitosa o infructuosa.

Una de las grandes virtudes de *Google Analytics* es que permite visualizar grandes cantidades de datos mediante un cuadro de mando en el que se consultan de un vistazo los indicadores de éxito / fracaso más comúnmente exigidos a estas herramientas, y entender, en la medida de lo posible, ese indicador que con números sólo, sin contextualizar y por sí mismo, no nos diría mucho. Por ejemplo, en la figura 5, el indicador en verde de Site Usage +13,30%, significa que está aumentando el número de visitas en el período estudiado en comparación con el período anterior, los mismos meses del año pasado, por lo que podemos deducir una mejora en ese aspecto, pero sólo en ese.

Por tanto, para interpretar qué respuestas queremos, necesitamos pensar detenidamente qué preguntas haremos, así como entender cuándo un dato se juzga positivo o negativo.

Partiendo del indicador web más tradicional, el número de páginas vistas, debemos cuestionarnos si significa lo que nosotros creemos. Si al mes nos visitan miles de usuarios que ven millones de páginas, pero no adquieren nada ni clican en ningún banner de publicidad o –en el caso de los opacs– no piden prestado online ningún libro, o no ven satisfecha una demanda de información, lo que les obliga a seguir llamando al teléfono de atención pública, ese indicador puede serlo de un fracaso, no de un éxito. Poner en contexto el indicador es clave en el análisis.

Lo mismo ocurrirá al estudiar las conductas de búsqueda de los usuarios.

Así, en un opac de un portal bibliotecario se podría medir cuántos usuarios se apuntan a un boletín de novedades, o con la nueva funcionalidad *Site Search* de *Google Analytics*, responder a algunas de estas preguntas (Avinash):

- ¿Cuánto usan mi buscador interno los usuarios?
- ¿Qué palabras clave buscan?
- ¿Dónde comienzan a buscar?
- ¿Cómo buscan los diferentes grupos de usuarios dentro de mi sitio web?
- ¿Qué beneficios obtienen quienes buscan en mi sitio web?
- ¿Están satisfechos con lo que encuentran?

“En el análisis es clave poner el indicador en contexto”



Figura 5. Evolución comparativa anual de visitas

3. Browsing versus Searching

Al diseñar la arquitectura de información del sitio web se deben tener en cuenta que existen dos modos fundamentales de visita:

- a través de la navegación interna entre enlaces (organizados de alguna forma), lo que se denomina *browsing* u “hojeo”.

- mediante la búsqueda, que puede ser interna (cuando el usuario utiliza ese buscador, y elige una respuesta para navegar hasta ella), y externa (cuando un usuario, al utilizar un buscador como *Google* por ejemplo, encuentra una página –no necesariamente la principal o *home*, sino una página interna, que da respuesta directa a su consulta– y accede a ella sin pasar por la principal).

Como se puede observar, en realidad todas son formas distintas de hacer *browsing* u hojeo por enlaces, por lo que nuestro análisis debe contemplar todos estos modos de forma conjunta y coordinada, con un enfoque sistémico. Así podríamos saber, por ejemplo:

- Si en nuestro sitio web se comprueba lo que ya postulaba **Nielsen** en 1997, que más de la mitad de los usuarios prefieren buscar (*search dominant*) que mirar en un directorio o similar (*browsing, link dominant*).

- Si debemos diseñar nuestro sitio web más para los que buscan o para los que navegan (si el público potencial de nuestro sitio web es más *search-dominant* que *browsing-dominant*).

- Si debemos cambiar la arquitectura de información dependiendo de las palabras clave que usan, o si lo que se busca existe ya y es fácil de encontrar.

“Cuanto mayor sea el porcentaje de salidas del portal, menos satisfechos estarán nuestros usuarios con los resultados de la búsqueda”

- Si lo que el buscador interno ofrece es lo que el usuario está demandando y está bien posicionado o no dentro de las respuestas que da el buscador interno (*internal search engine optimization*, o *SEO* pensando en nuestro buscador interno, no en que esté bien posicionado en *Google*, puesto que normalmente lo que se busca dentro de un sitio web, no tiene por qué ser lo que se busca fuera de ese sitio a través de un buscador).

- En qué momento la navegación por enlaces (organizados por categorías, facetes, alfabéticamente, etc.) no está dando a los usuarios la información que ellos buscan, en el lenguaje (natural o no) en el que quieren verlo, siguiendo los principios de la *Information foraging theory* –o teoría de la búsqueda y recolección de información– (**Pirolli, Card**, 1999): esto es, cuál es el porcentaje de usuarios que dejó nuestro sitio web

después de buscar algo en el buscador interno. Porque cuanto mayor sea el porcentaje de salidas del portal, menos satisfechos estarán nuestros usuarios con los resultados de la búsqueda (**Kaushik**).

4. Estudio de caso

4.1. Aplicación del método de combinación

Para ilustrar la metodología de análisis propuesta se toma el caso del rediseño del sitio web de la *Entidad Pública de Saneamiento de Aguas Residuales de la Comunidad Valenciana* llevado a efecto durante el año 2007, y con el que se está desarrollando un proyecto de mejora continua (figura 6).

1. Se observaron las palabras clave específicas relevantes que atañen a las temáticas del portal que más se buscan en *Google*, mediante *Google Adwords*.
2. Se comprobó su estacionalidad externa en *Google Trends / Insights* y en el mismo *Google Adwords*.
3. Se rediseñó el portal con estas palabras en mente –además de otros múltiples requerimientos– a

la hora de generar una estructura de contenidos o arquitectura de información que permitiera a los usuarios que llegaban de los buscadores encontrar rápidamente la información, y a los usuarios que visitaban el portal descubrir intuitivamente los contenidos de su interés, teniendo como objetivo que utilizaran poco el buscador (lo que, en portales dinámicos con gran variedad y actualización de información no permite controlar sin ciertos costes el orden de relevancia de las búsquedas, salvo programación ad hoc y/o mediante el uso de *best bets* –mejores apuestas–).

4. Pasado un tiempo se analizaron tanto las palabras clave por las que el portal era más encontrado y visitado desde *Google* (mediante *Google Webmaster* y *Google Analytics*) como las palabras clave que se buscaban en el buscador interno (mediante *Site Search* de *Google Analytics*).
5. En ambos casos se buscó confirmar la estacionalidad externa e interna para prever en qué momentos del año mostrar contenidos de forma prominente en el portal, así como sacar a la luz búsquedas constantes a lo



Figura 6. Sitio web de Epsar de la Comunidad Valenciana <http://www.epsar.gva.es/>



Figura 7. Estacionalidad externa de la búsqueda “adjudicaciones”

largo del año, para que tuvieran una prominencia fija, clara y visible.

6. Se evaluaron diversos indicadores de éxito, como la mejora en el posicionamiento del portal en buscadores y el ratio de salida del portal tras la búsqueda interna (*bounce rate* o tasa de fracaso).

4.2. Resultados

La cantidad de análisis y resultados obtenidos fue múltiple, de los que destacamos algunas conclusiones interesantes que pueden servir como modelo de extracción de inferencias:

Muchas de las palabras por las que es encontrado muy a menudo el portal eran indicadoras de usuarios que buscaban específicamente el portal: “epsar”, “edar”, “entidad de saneamiento”.

Sin embargo, una vez encontrado el portal, las palabras que más se buscaban internamente eran mucho más específicas (“modelo 301”), encontrándose argot o sinónimos significativos (“modelo 203” o “md 203”).

Además, se confirmó que ciertas palabras clave o *queries* tenían similares estacionalidades tanto en los buscadores externos (con *Google Insights*) (figura 7) como en el buscador interno (mediante *Site Search* de *Google Analytics*) (figura 8).

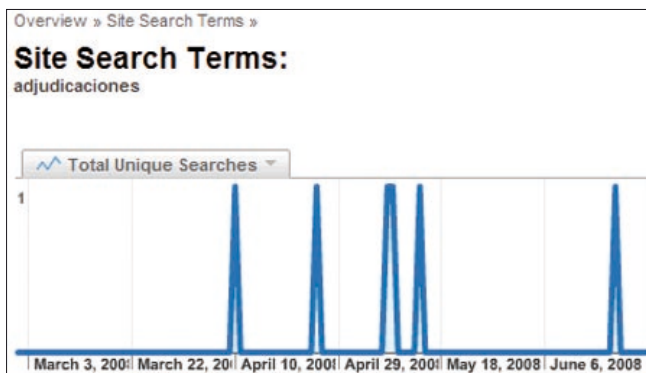


Figura 8. Estacionalidad interna de la búsqueda “adjudicaciones”

En este ejemplo podemos observar cómo coinciden las estacionalidades de la *query* “adjudicaciones”. El estudio de su frecuencia (volumen de búsquedas internas y externas) debería llevarnos a determinar si merece la pena llevar a cabo o no esfuerzos suplementarios.

Esta comprensión de los modelos y tendencias de búsqueda llevó a preparar los contenidos de forma que la estructura respondiera mejor a esas preguntas, y el sitio web las contestara lo antes posible en la jerarquía de contenidos (tanto visual como temáticamente).

Así, como resultado de este esfuerzo, el análisis posterior al rediseño dejó huella de un aumento de búsquedas externas que llevaban a contenidos que habían sido mejorados debido al análisis anterior, hasta incluso de casi el 1.000%, dato que corrobora estudios anteriores (*Serrano-Cobos*, oct. 2006) (figura 9).

modelo 010	
September 1, 2008 - September 30, 2008	41
September 1, 2007 - September 30, 2007	18
% Change	127.78%
entidad de saneamiento	
September 1, 2008 - September 30, 2008	33
September 1, 2007 - September 30, 2007	26
% Change	26.92%
decreto 32/2006	
September 1, 2008 - September 30, 2008	32
September 1, 2007 - September 30, 2007	3
% Change	966.67%

Figura 9. Comparativa de frecuencia de búsquedas en periodos distintos

Y por último, y quizá sea el mejor dato, se comprobó que las visitas que llevaban aparejadas búsquedas en su navegación, ofrecían un menor ratio de abandono del portal, en torno al 0,15% solamente (figura 10), lo que suponía una diferencia importante con respecto a la media del portal, y aportando visitas de los usuarios que habían empleado más tiempo al profundizar en los documentos que los buscadores daban como respuesta.

“Las palabras clave más utilizadas en buscadores externos que llevan al portal suelen tener diferencias con las usadas en el buscador interno”

Este dato puede hacernos sospechar que las búsquedas estaban ofreciendo información valiosa, puesto que de otro modo, el usuario rápidamente habría cambiado su modo de búsqueda, o habría salido del portal.

5. Conclusiones e investigación futura

Se ha estudiado la evolución de las palabras clave o *queries* en buscadores externos como *Google* y en el buscador interno, lo que ha llevado a mejorar de forma escalonada el portal, analizando la estacionalidad de las *queries* y la influencia que el *browsing* tiene en el uso del buscador interno.

El análisis de los resultados nos lleva a confirmar que:

- Una mejor arquitectura de información produce una mejor experiencia de *browsing*.
- Un mejor *browsing* produce mejores resultados tanto de cara a los buscadores externos, como al buscador interno.
- Dependiendo del portal, las palabras clave más buscadas en buscadores externos que llevan al portal suelen tener diferencias con las buscadas en el portal interno.

- El estudio del lenguaje natural de las *queries* y su agrupación (*clustering*) permiten responder mejor a posibles variaciones del lenguaje, dando valor al control del vocabulario mediante el análisis de *search logs*.

Los estudios futuros nos perfilan un escenario en el que sistemas más complejos de minería de datos nos permitan inferir conclusiones más precisas y predicciones más ajustadas a través de la combinación de una mayor riqueza de datos a priori inconexos, tales como la procedencia del usuario (**Kralisch, Berendt, 2004**) y la reserva de un ítem, o los enlaces que ha seguido en

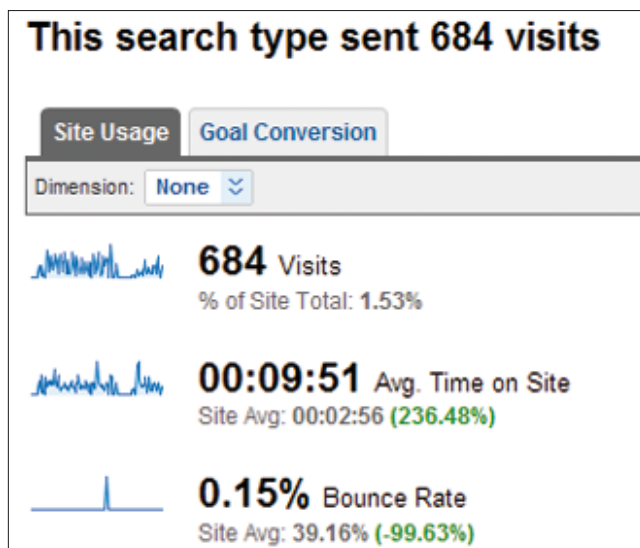


Figura 10. Evolución del ratio de fracaso o bounce rate en visitas con búsquedas

internet un usuario hasta salir de nuestro portal. Cuanto mayor sea el número, período histórico, volumen y variedad de estos datos, mejores inferencias a largo plazo se podrán hacer.

6. Direcciones de interés

- Google Webmaster
<http://www.google.com/webmasters/>
- Google Trends
<http://www.google.es/trends>
- Google Adwords Keyword Tool
<https://adwords.google.com/select/KeywordToolExternal>
- Google Insights
<http://www.google.com/insights/search/>
- Google Analytics
<http://www.google.com/analytics/es-ES/>

7. Bibliografía

1. Allen, B. L. "Cognitive research in information science: implications for design". *Annual review of information science and technology*, 1991a, 26, pp. 3-37.
2. Allen, B. L. "Topic knowledge and online catalog search formulation". *Library quarterly*, 1991b, v. 61, n. 2, pp. 188-213.
3. Allerton Institute. *How we do user-centered design and evaluation of digital libraries: a methodology*, 1995.
<http://edfu.lis.uiuc.edu/allerton/95/>
4. Angiolillo, J. *Search log analysis as a usability engineering tool*.
<http://home.earthlink.net/~searchworkshop/docs/chi-search-angiolillo3.doc>
5. Borgman, C. L. "All users of information retrieval systems are not created equal: an exploration into individual differences". *Information processing & management*, 1989, v. 25, n. 3, pp. 237-251.
<http://informationr.net/ir/10-2/paper217.html#bor89>
6. Bradford, C.; Marshall, I. *Analysing users WWW search behaviour*.
<http://www.cs.kent.ac.uk/people/staff/iwm2/personal/ieewebcoll.pdf>

7. **Cheuk, W.; Dervin, B.** "A qualitative sense-making study of the information seeking situation: professionals in three workplace contexts". *Electronic journal of communication*, 1999, v. 9, n. 2-4. <http://communication.sbs.ohio-state.edu/sense-making/art/artabscheukdervin99ejoc.html>
8. **Ellis, D.** et al. "Information searching. Part 5. User - intermediary interaction". *Journal of the American Society for Information*, 2002, pp. 883-893.
9. **Ellis, D.** "A behavioural approach to information retrieval design". *Journal of documentation*, 1989, n. 46, pp. 318-338
10. **Fidel, R.** "Moves in online searching". *Online review*, 1985, v. 9, n. 1, pp. 61-74.
11. **Jones, S.; Cunningham, S. J.; McNab, R.** *Usage analysis of a digital library*. Working paper 98/13, Department of Computer Science, University of Waikato. Hamilton, New Zealand, 1998. <http://www.cs.waikato.ac.nz/~stevej/Research/PAPERS/dl98logs.pdf>
12. **Kaushik, A.** Five questions to ask of your site search data, 2008. <http://www.google.com/support/conversionuniversity/bin/answer.py?answer=82020>
13. **Koch, T.; Ardö, A.; Golub, K.** *Log analysis of user behaviour in the Renardus web service*, 2004. <http://www.it.lth.se/knownlib/renardus-log/log-analysis.html>
14. **Kralisch, A.; Berendt, B.** "Cultural determinants of search behaviour on websites" En: *Proceedings of the IWIPS 2004 Conference on culture, trust, and design innovation*. Vancouver, Canada, 8-10 Jul 2004. http://www.wiwi.hu-berlin.de/~berendt/Papers/kralisch_berendt_CATAC2004_to_appear.PDF
15. **Kuhlthau, C. C.** The role of experience in the information search process of an early career information worker: perceptions of uncertainty, complexity, construction, and sources. *Journal of the American Society for Information Science*, 1999, v. 50, n. 5, pp. 399-412.
16. **Mangano, V.; Beaulieu, M.; Robertson, S.** "Evaluation of interfaces for IRS: Modelling end-user searching behaviour". *Proceedings of 20th British Computing Society Information Retrieval Special Group (BCS/IRSG), Colloquium on information retrieval*, Springer-Verlag, Grenoble, 1998, pp. 137-146.
17. **Marchionini, G.** *Information seeking in electronic environments*. Cambridge: Cambridge University Press, 1995.
18. **Moe, W.** "Buying, searching, or browsing: differentiating between online shoppers using in-store navigational clickstream". *Journal of consumer psychology*, 2003, n. 13, pp. 29-39.
19. **Navarro-Prieto, R.; Scaife, M.; Rogers, Y.** "Cognitive strategies in web searching". En: *Proceedings of the 5th Conference on human factors and the web*. Jul 1999.
20. **Nielsen, J.** *Designing web usability*. Indianapolis: New Riders, 1999.
21. **Nielsen, J.** "Search and you may find". En: *Alertbox*, 15-07-1997. <http://www.useit.com/alertbox/9707b.html>
22. **Pirolli, P.; Card, S.** (1999). "Information foraging". *Psychological review*, 1999, v. 106, n. 4, pp. 643-675. <http://www2.parc.com/istl/groups/uir/publications/items/UIR-1999-05-Pirolli-Report-InfoForaging.pdf>
23. **Rosenfeld, L.** "Search log analysis: the 'quiet' user study". *User studies and information architecture. Asist 2002 annual conference*, 18-10-2002. <http://www.louisrosenfeld.com>
24. **Serrano-Cobos, J.** "Casos prácticos de 'information seeking' en el diseño de sistemas de información web". En: *VIII Jornadas de gestión de la información*, nov. 2006. Madrid: Sedic, pp. 11-27. http://eprints.rclis.org/archive/00007933/01/Ponencia_1_Jorge_Cobos.pdf
25. **Serrano-Cobos, J.** "Interaction design based on user search behaviour through diachronic search logs analysis". En: *I International conference on multidisciplinary information sciences and technologies - InSciT 2006*, Mérida, España, 25-28 oct. 2006. <http://www.instac.es/inscit2006/papers/pdf/374.pdf>
26. **Spink, A.** *Modelling users' successive searches in digital environments*, 1998. <http://www.dlib.org/dlib/april98/04spink.html>
27. **Vakkari, P.; Pennanen, M.; Serola, S.** "Changes of search terms and tactics while writing a research proposal: a longitudinal case study". *Information processing & management*, 2003, v. 39, n. 3, pp. 445-463.
28. **Wildemuth, B. M.** "The effects of domain knowledge on search tactic formulation". *Journal of the American Society for Information Science & Technology*, 2004, v. 55, n. 3, pp. 246-258.
29. **Wilson, T. D.** "Human information behaviour". *Information science*, 2000, v. 3, n. 2, pp. 49-55.
30. **Zhang, X.; Anghelescu, H.; Yuan, X.** "Domain knowledge, search behaviour, and search effectiveness of engineering and science students: an exploratory study". *Information research*, 2005, Jan., v. 10 n. 2, paper 217. <http://informationr.net/ir/10-2/paper217.html>

Jorge Serrano-Cobos

jorgeserrano@gmail.com

<http://trucosdegoogle.blogspot.com>

<http://www.masmedios.com>

Suscripciones

Renovar (o comenzar) la suscripción a "El profesional de la información" es ágil y sencillo.

Usted puede gestionar online su suscripción conectándose a esta web:

<http://www.elprofesionaldelainformacion.com/suscripciones.html>

Si lo desea puede comunicar con nosotros dirigiéndose a:

suscripciones@elprofesionaldelainformacion.com

o al teléfono: +34-609 352 954