



Artículo Original / 091007-2016

Posicionamiento SEO del Centro de Investigación Flamenco Telethusa: Palabras claves

SEO Positioning of Telethusa Flamenco Research Center: Keywords

Pedro Corchero Murga, MA (1)

María del Rosario Fernández-Falero, PhD (2) Email: rferfal@unex.es

M^a Antonia Hurtado Guapo, PhD (2)

(1) Máster Universitario en Gestión de la Información Digital (2015). Universidad de Extremadura. Badajoz, España.

(2) Grupo de investigación Sapiencia. Universidad de Extremadura. Badajoz, España.

Recibido: 20 may 2016 Revisión editorial: 21 may 2016 Revisión por pares: 24 may 2016 Aceptado: 25 may 2016 Publicado online: 26 may 2016.

Resumen

El estudio, clasificación y organización de los metadatos extraídos de los sitios web es una herramienta de gran utilidad para optimizar su posicionamiento. El objetivo de este trabajo es posicionar el web site del Centro de Investigación Flamenco Telethusa. Para ello se han analizado los términos utilizados en los buscadores para encontrar el sitio web de la entidad. Los 10 términos más usados para la búsqueda según el sistema de estadísticas AWStats son los siguientes: flamenco, baile, danza, ballet, estructura, lesiones, telethusa, alegrías, bailar, pie. El uso de estos términos de búsqueda ha sido analizado y compilado por meses y años, desde diciembre de 2008 hasta diciembre de 2014. Los resultados de búsqueda para cada uno de estos términos han sido: flamenco, 9600; baile, 4196; danza, 3147; ballet, 1819; estructura, 1651; lesiones, 989; telethusa, 888; alegrías, 887; bailar, 803; y pie, 851. El número de búsquedas del sitio ha ido aumentando de forma progresiva desde 2009 hasta año 2012 donde alcanzó su máximo. Se produce una disminución durante los años 2013 y 2014. El mes de noviembre, para todos los años, es en el que se produce un mayor número de búsquedas, y el término flamenco el más buscado. Como conclusión proponemos que sea en este mes de noviembre cuando se publiquen contenidos nuevos para satisfacer la demanda de los usuarios.

Palabras Clave Minería web, sitio web, metadatos, AWStats, motores de búsqueda, términos de búsquedas.

Abstract

The study, classification and organization of metadata extracted from websites is a useful tool to optimize their ranking. The aim of this work is to optimize the Telethusa Flamenco Research Center (CIFT) Website. To do this, we have analyzed the keywords that search engines use to find the website of the organization. The most frequently used terms, according to AWStats statistics, are: flamenco, baile, danza, ballet, estructura, lesiones, telethusa, alegrías, bailar, pie. The use of these keywords has been analysed and compiled, both monthly and annual. From December 2008 to December 2014, the search results for each of these terms were: flamenco, 9600; baile, 4196; danza, 3147; ballet, 1819; estructura, 1651; lesiones, 989; Telethusa, 888; alegrías, 887; bailar, 803; and pie, 851. The number of website searches increased steadily from 2009 till 2012, when it reached its peak, followed by a decrease in the following two years, 2013 and 2014. In conclusion, we suggest to publish new content in November, since this is the month with the highest number of web searches.

Keywords Web minning, website, metadata, AWStats, search engine, search term.

Introducción

La minería Web permite recuperar información de un sitio web de manera automática (Fuentes Reyes, 2007), hecho de gran valor para las organizaciones y muy utilizado para el estudio del posicionamiento Web y el comercio electrónico.

La Minería Web de contenido es la que clasifica y organiza los metadatos extraídos, con el objeto de recuperar y facilitar el acceso a la información. Tales metadatos pueden ser estudiados estáticamente mediante instantáneas de la web en un periodo determinado. La minería de datos consta de cuatro fases: recolección automática de la información importante para procesarla posteriormente; procesamiento de datos, ordenándolos y clasificándolos automáticamente; descubrimiento de patrones mediante hallazgo de frecuencias, reglas de asociación que permiten establecer estrategias de difusión de la información en las organizaciones; y análisis de patrones -se interpretan y validan los patrones¹-.

La minería web es objeto de investigaciones para mejorar los sitios web, contribuyendo al desarrollo de un plan de marketing para empresas², a la mejora de la actividad comercial electrónica del sitio³, a la observación de los distintos tipos de comportamiento respecto al trabajo colaborativo⁴, o a la mejora de sitios dedicados a la teleformación -mediante la obtención de datos que permitan al profesorado evaluar el proceso de aprendizaje y a los estudiantes sentirse ayudados en la realización de sus tareas⁵-. El trabajo sobre el ciclo de vida de una campaña completa de posicionamiento⁶, describe trabajos sobre posicionamiento SEO y minería de consultas de usuarios con el fin de centrarse al máximo en las intenciones de búsqueda de los interesados⁷. Además, la relación entre las palabras claves y el SEO viene siendo estudiada práctica y teóricamente, estableciéndose un marco teórico que avala las investigaciones en este campo⁸.

El posicionamiento web consiste en el estudio de una página para optimizarla de cara a los resultados en buscadores⁹. Puede ser de dos tipos: optimización de motores de búsqueda, Search Engine Optimization (SEO), o marketing de motores de búsqueda, Search Engine Marketing (SEM). El posicionamiento natural u orgánico, también llamado SEO, consiste en una serie de técnicas que permiten que el sitio aparezca en las primeras posiciones en los buscadores. El posicionamiento patrocinado o inorgánico, SEM, consiste en el pago por enlaces patrocinados por los buscadores¹⁰.

Las Palabras claves son un factor de relevancia para el SEO: deben aparecer en el título, que debe contener las palabras claves más importantes en las páginas distribuidas inteligentemente (es decir, cuando es preciso y no añadiéndolas por añadir); junto con palabras relacionadas y sinónimos refuerza la relevancia del sitio web en el dominio: aunque Google ya no le da importancia a una coincidencia total entre las palabras claves introducidas por el usuario y las que se encuentran en el dominio, sigue siendo un factor a considerar en otros componentes de la página web como en los enlaces internos, URL, atributos al en imágenes, encabezados <h1>, bloques de texto y ritmo de actualización del sitio, ya que los contenidos recientes son más relevantes, como sucede con los acontecimientos estacionales o noticias entre otros¹¹.

Material y Método

Muestra

Para la elaboración de este estudio se extraen todos los términos utilizados en los navegadores: MS Internet Explorer, Firefox, Opera, Safari, Google Chrome, Mozilla y Netscape para encontrar el sitio web del Centro de Investigación Flamenco Telethusa (CIFT). De los términos recuperados se seleccionan y analizan los 10 términos de mayor popularidad.

Método y Diseño de la investigación

El objeto de estudio es el sitio web del Centro de Investigación Flamenco Telethusa (CIFT), organización privada sin ánimo de lucro, cuyo objetivo más importante es la investigación del baile flamenco basándose en las Ciencias de la Actividad Física y la Salud, dirigido con un fin de asesoramiento científico, divulgativo y editorial. También colabora con investigaciones relacionadas con otros aspectos del flamenco así como de la danza en general¹².

El centro edita una revista científica donde se publican artículos multidisciplinares sobre evolución histórica, técnicas y teorías sobre flamenco¹³, cuyo objetivo es el estudio del flamenco desde distintas áreas del saber. Para ello, el CIFT, además de contar con sus propios miembros, entre ellos doctores universitarios y licenciados, se asocia con investigadores externos que forman parte de grupos de investigación

en distintas universidades españolas. El objetivo principal es científico y el centro se encarga de divulgarlo para así poder englobar a un mayor número de personas relacionadas con el flamenco. Por último, profesionales del mundo del flamenco y de las manifestaciones artísticas en general ayudan al proyecto de manera desinteresada.

El Centro de Investigación Flamenco Telethusa también ofrece a la comunidad científica un repositorio de documentos rescatados de fuentes antiguas. Está dedicado a documentos descatalogados y cuyos derechos de autor hayan vencido. Los propios artistas a su vez pueden subir sus colecciones en el repositorio, siempre que obedezcan los requisitos legales de la página 12.

El material de trabajo procede de las estadísticas de la página web del Centro de Investigación Flamenco Telethusa (CIFT) (<http://www.flamencoinvestigacion.com/>). Concretamente el estudio de los términos de búsqueda del CIFT utilizados en los motores de búsqueda, google, altavista, yahoo y otros.

Las estadísticas recogidas están generadas con AWStats, una herramienta poderosa y gratuita que genera datos por streaming, FTP o servidor de correo electrónico de forma gráfica. AWStats es un software gratuito distribuido bajo la licencia GNU General Public License. Entre sus características se encuentra la obtención de palabras clave y frases clave usadas para encontrar tu sitio procedente de los 115 motores de búsqueda más famosos detectados como Google, Altavista y otros 14.

La primera etapa del trabajo consiste en la determinación de las variables de estudio. Las estadísticas de AWStats proporcionan gran cantidad de información. Por ello se recuperan los datos de las Top 25 palabras claves de cada mes, desde Diciembre de 2008 hasta Diciembre de 2014, desde el PDF generado por AWStats, para ser insertados en el editor de texto Coda. Se eliminan las palabras vacías, y los términos en japonés y chino se traducen al español.

En la segunda etapa se reorganizan los datos en tablas desde el programa Pages, procediendo a su estudio. Para ello se analizan las 10 palabras más populares de cada mes, de las Top 25 recuperadas.

Los softwares empleados en la realización del trabajo han sido Numbers (el programa de hojas de cálculo de Apple) y Coda (un editor de texto)¹⁵.

Finalmente las palabras claves presentan la misma ortografía con la que aparecen en el documento fuente (las estadísticas generadas con AWStats).

Resultados

Cómo se muestra en el gráfico 1, las 10 palabras clave más buscadas en la Red para acceder al Centro de Investigación Flamenco Telethusa desde diciembre de 2008 hasta diciembre de 2014 han sido: flamenco (9600), baile (4196), danza (3147), ballet (1819), estructura (1651), lesiones (989), telethusa (888), alegrías (887), bailar (803) y pie (851).

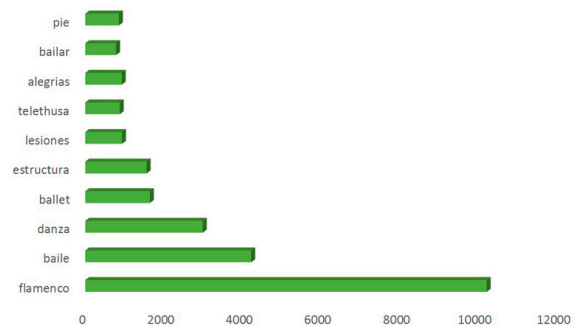


Fig. 1 Palabras más buscadas en la red para acceder al Centro de Investigación Flamenco Telethusa desde diciembre de 2008 hasta diciembre de 2014.

Tabla 1. Número total de búsquedas por término y mes

	flamenco	baile	danza	ballet	estructura	lesiones	telehusa	alegrias	bailar	pie
enero	891	349	196	92	135	54	112	65	114	52
febrero	929	358	248	178	123	56	111	44	45	103
marzo	575	373	278	183	102	88	51	39	72	114
abril	615	398	322	174	130	114	68	67	40	56
mayo	884	365	338	185	152	116	91	102	19	94
junio	832	346	244	115	172	46	98	77	43	45
julio	650	256	178	101	158	36	35	58	70	38
agosto	708	261	221	139	148	73	31	57	74	70
septiembre	779	339	236	156	134	70	42	92	108	52
octubre	968	392	321	206	119	101	52	90	116	87
noviembre	1032	464	347	184	159	138	114	117	67	96
diciembre	737	295	217	106	119	97	83	79	35	44
TOTAL	9600	4196	3146	1819	1651	989	888	887	803	851

A continuación, se expone la búsqueda de términos por año.

Tabla 2. Resumen del número total de búsquedas de términos por año

	2008	2009	2010	2011	2012	2013	2014
flamenco	11	600	1277	2327	3362	1968	663
baile	2	231	499	832	1573	803	282
danza	0	163	586	746	822	502	170
ballet	1	53	183	438	603	252	115
estructura	0	0	0	180	691	538	153
lesiones	3	152	228	294	169	28	63
telehusa	2	107	169	296	205	100	0
alegrias	0	0	0	47	398	322	103
bailar	0	40	101	275	187	128	55
pie	0	0	0	188	461	150	58
TOTAL	19	1346	3043	5623	8471	4791	1662

Las siguientes gráficas muestran el análisis de cada una de las 10 palabras claves mediante la suma del tráfico de los años estudiados.

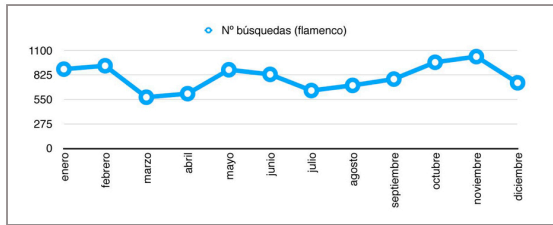


Figura 1. Tráfico del término flamenco por meses desde 2008 a 2014

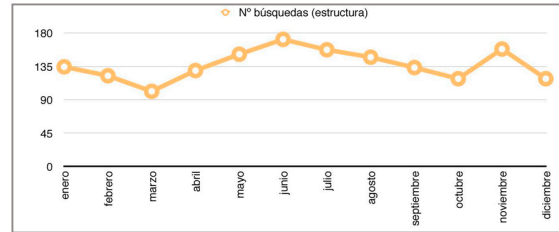


Figura 5. Tráfico del término estructura por meses desde 2008 a 2014

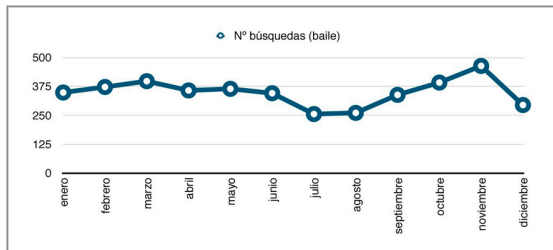


Figura 2. Tráfico del término baile por meses desde 2008 a 2014

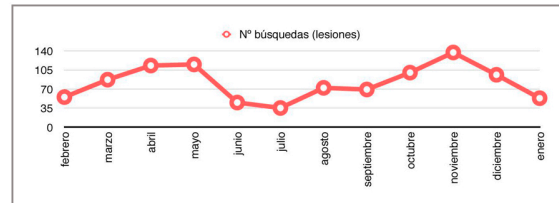


Figura 6. Tráfico del término lesiones por meses desde 2008 a 2014

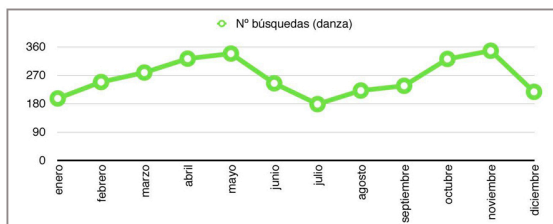


Figura 3. Tráfico del término danza por meses desde 2008 a 2014

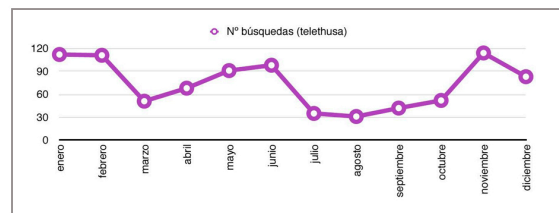


Figura 7. Tráfico del término telethusa por meses desde 2008 a 2014

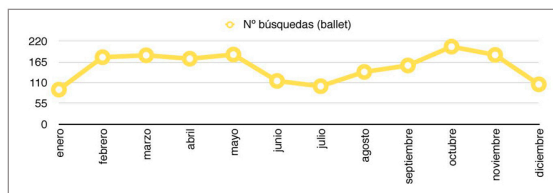


Figura 4. Tráfico del término ballet por meses desde 2008 a 2014

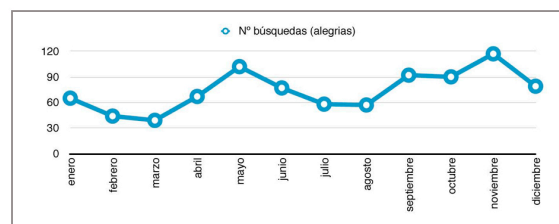


Figura 8. Tráfico del término alegrías por meses desde 2008 a 2014

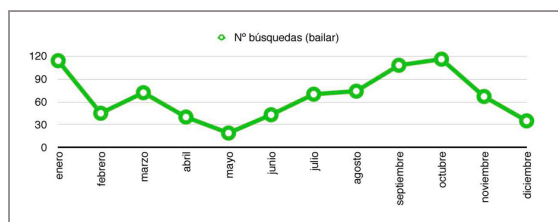


Figura 9. Tráfico del término bailar por meses desde 2008 a 2014

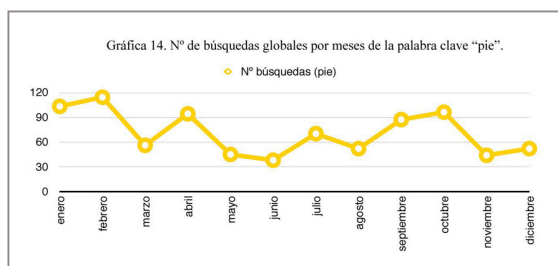


Figura 10. Tráfico del término pie por meses desde 2008 a 2014

Análisis de los resultados

Tras el estudio de los resultados obtenidos se constata que en los años 2008 y 2009 el CIFT contaba con baja popularidad y aún no tenía notoriedad en el buscador. El nivel más bajo de búsquedas se produce en estos dos años. A lo largo del período estudiado, como se observa en la gráfica, el número de búsquedas ha ido aumentando y con ello las visitas, llegando a ser el año 2012 el de mayor ratio de búsquedas.

Analizando los resultados globales por términos, expuestos en las 10 figuras, el número de búsquedas de la palabra flamenco sufre un descenso a partir del mes de febrero, alcanzando su mínimo en marzo-abril. En el mes siguiente de nuevo aumenta, para decrecer desde junio hasta julio, momento en el que asciende el número de búsquedas hasta alcanzar su máximo en noviembre.

El descriptor baile a diferencia del descriptor flamenco, mantiene su tráfico de búsqueda constante hasta junio, momento en que empieza a decaer para alcanzar su mínimo en julio, incrementándose posteriormente hasta llegar a noviembre alcanzando su pico máximo.

La evolución del tráfico danza no es constante, ya que empieza a incrementarse desde enero hasta mayo, para luego descender bruscamente y llegar a su mínimo en julio. Posteriormente aumenta y consigue su pico máximo en noviembre.

El término ballet se mantiene constante desde febrero hasta mayo, disminuyendo hasta llegar a julio, que es el mes de menor actividad, para luego crecer hasta alcanzar su punto más alto en octubre.

El término estructura no tiene cambios tan bruscos, aunque sus niveles mínimos se producen en marzo y en octubre, aumentando sucesivamente hasta conseguir su tráfico máximo en junio y noviembre.

El descriptor lesiones sufre oscilaciones variables; el número menor de búsquedas se realiza en enero, y su mayor pico se encuentra en noviembre.

En cuanto al término de búsqueda telethusa, es notable su variación, teniendo un importante descenso en el mes de febrero y junio, llegando a su mínimo en agosto, para luego ir aumentando hasta conseguir su mayor tráfico en noviembre.

El descriptor alegrías suele tener una progresión global ascendente a lo largo de todo el año, con un mínimo de búsquedas en marzo y un máximo en noviembre.

El término bailar tiene dos niveles: en la primera mitad del año desciende hasta llegar a su mínimo en mayo; posteriormente se produce un aumento hasta llegar a su máximo en octubre, un mes antes que otros descriptores, que tienen su pico máximo en noviembre.

Finalmente, la palabra clave pie contiene muchos ascensos y descensos, aunque su punto más bajo se sitúa en junio, el rango donde los descriptores tienen menos tráfico. Al igual que en el caso de bailar, su punto más elevado corresponde a octubre.

Conclusiones

Esta investigación supone un instrumento de trabajo para el webmaster del site estudiado, con el cual puede mejorar el posicionamiento SEO. Con este análisis queda demostrado que es necesario utilizar los descriptores adecuados para aumentar el posicionamiento del sitio web. Para ello es

recomendable incluir, de manera estratégica, los diez términos analizados, en las distintas zonas del site (título, encabezados, URL y otros).

Las palabras totales extraídas de las estadísticas para localizar el CIFT son: flamenco, injurias, lesiones, tesis, pdf, biomecánica, cruz, Telethusa, revista, jb, calvo, Cádiz, baile, Vargas, dance, doctoral, danzas, dolencia, ballet, danza, investigación, flamencos, dolor, bailarines, investigación, Alfonso, bailar, zapateado, zapateados, balonmano, tipos, física, espalda, Macías, preparación, actividad, centro, actuales, bailes, estudio, descriptivo, biomecánico, física, preparación, condición, espalda, clásica, posiciones, importancia, zapatos, España, braceo, tacón, cervicalgia, bailaoras, usadas, bailarinas, clásica, lesiones, www.flamencoinvestigacion.es/-, flamenca, técnica, ejercicios, clásico, lozano, lesión, investigación, cambre, giro, adelgazar, andehor, brujo, aeróbica, lumbar, giros, amor, manuel, sebastian, deporte, alto, cerrado, pase, deors, técnica, vuelta, física, abierto, técnica, pie, hallux, valgus, coreografía, partes, alegrías, estructura, pies, bailarina, estructuras, elementos, rodilla, apoyo, puntos, corporal, estiramientos, artículos, cifosis, leve, anaizquierdoangulo, genoma, génesis, バレエ (ballet), フラメンコ (flamenco), 奪います (quitar), フラマンシア (flamenco shea), ジュース (zumo), http, abductus, dances, patología, letra.

Las palabras claves que habría que usar, según los resultados obtenidos, para aumentar el posicionamiento del sitio serán: flamenco, baile, danza, ballet, estructura, lesiones, telethusa, alegrías, bailar y pie, que son las más populares por orden de tráfico descendente.

En lo que respecta a las búsquedas realizadas, el mes de noviembre, para todos los años, es en el que se produce un mayor número de búsquedas, y el término flamenco el más buscado al contrario que el término bailar, el menos buscado.

Como corolario al estudio, podemos concretar que el establecimiento de los meses de mayor uso de búsqueda permitirá al CIFT elaborar un cronograma de contenidos para publicarlos cuando lo demanden los usuarios del centro de investigación. Finalmente, los resultados indican que el mes de noviembre es el más adecuado para informar sobre flamenco (específicamente alegrías), baile, danza y lesiones; el mes de octubre para ballet y bailar; el mes de junio para estructura y el mes de marzo para pie.

En las búsquedas globales de cada descriptor

clasificadas por año se puede comprobar que el número de búsquedas del sitio ha ido aumentando de forma progresiva desde 2009 hasta un máximo en el año 2012. Se produce una disminución durante los años 2013 y 2014. El CIFT debe plantearse un proceso de dinamización de su web para recuperar los usuarios perdidos.

Se podría mejorar el posicionamiento del sitio web siguiendo alguna de las siguientes pautas:

- Actualmente el sitio web estudiado, y según la web de W3C, no obtiene ninguna validación, teniendo 24 errores y 6 advertencias. Es recomendable validar el sitio para que cumpla con los estándares W3C. Esto ayudará en el posicionamiento web del mismo en la red.
- En el momento del estudio dicha web no se encontraba optimizada para móviles, este hecho influye negativamente en el posicionamiento, ya que en la actualidad el buscador Google obliga a que los sitios web tengan optimizado el sitio para dispositivos móviles. Sería conveniente adaptar el diseño de la página para que cumpliera con estos requisitos.

Referencias documentales

1. Fuentes SC, Ruiz M. 2007. Minería Web: un recurso insoslayable para el profesional de la información. *Acimed* 16(4): 1-10.
2. Araya S, Silva M, Weber R. 2004. A methodology for web usage mining and its application to target group identification. *Fuzzy Set Syst* 148(1): 139-152.
3. Büchner AG, Mulvenna MD. 1998. Discovering internet marketing intelligence through online analytical web usage mining. *Sigmod Rec* 27(4): 54-61.
4. Chau, M, Zeng, D et al. 2003. Design and evaluation of a multi-agent collaborative Web mining system. *Decis Support Syst* 35(1): 167-183.
5. Zorrilla ME, Menasalvas E, Marín D, et al. 2005. Web usage mining project for improving web-based learning sites. En *Computer Aided Systems Theory-EUROCAST 2005*. Heidelberg: Springer. Pp 205-210.
6. Codina L, Marcos MC. 2005. Posicionamiento web: conceptos y herramientas. *El profesional de la información*. 14(2): 84-99.
7. Gonzalo C, Godina L, Rovira C. 2015. Recuperación de información centrada en el usuario y SEO: categorización y determinación de las intenciones de búsqueda en la web. *Index Comunicación* 5(3): 19-27.
8. Codina L, Iglesias M, Pedraza R, García L. 2016. Visibilidad y posicionamiento web de informaciones periódicas: el framework SEO-RCP. Barcelona : Universitat

- Pompeu Fabra. Departament de Comunicació, 2016. 21 p. (Active Audiences and Journalism). <http://hdl.handle.net/10230/26098> . Consultada 18 abr 2016.
9. Codina L. 2004. Posicionamiento web: conceptos y ciclo de vida. *Hipertext.net* (2): 1-10.
 10. García CL, Calmet RF. Posicionamiento en buscadores aplicando un model SEO. 2011. *Rev Investig Sistemas Informática* 8(2): 75-112.
 11. De Teresa T. ¿Pero qué diablos es SEO?. <http://deteresa.com/que-es-seo/> . Consultada 19 may 2013.
 12. Centro de Investigación Flamenco Telethusa. http://www.flamencoinvestigacion.com/index_revista.htm Consultada 12 may 2016.
 13. Fernández-Falero MR, Vargas-Macías, A, Gómez Lozano, S. 2015. Revista interdisciplinaria de acceso abierto: calidad. En: 5ª Conferencia internacional sobre revistas de ciencias sociales y humanidades. Universidad de Murcia. Pp 1-3.
 14. AWStats. AWStats official web site: Free log file analyzer for advanced statistics. <http://www.awstats.org/> . Consultada 10 mayo 2016.
 15. Corchero P. 2015. Análisis de los términos de búsqueda para mejorar el posicionamiento SEO: Centro de Investigación Flamenco Telethusa [Tesis Master]. Badajoz: Universidad de Extremadura. <http://dehesa.unex.es/handle/10662/3238?show=full> . Consulta 1 abril 2016.
-
-