

FUNCIONALIDADES DE CLASIFICACIÓN EN SISTEMAS DE GESTIÓN DE CONTENIDOS

Evaluación de prestaciones

JESÚS TRAMULLAS
Univesidad de Zaragoza
tramullas@unizar.es

PIEDAD GARRIDO
Universidad de Zaragoza
piedad@unizar.es

ANA I. SÁNCHEZ-CASABÓN
Univiverdad de Zaragoza
asanchez@unizar.es

Resumen Los sistemas de gestión de contenidos (*content management systems*, CMS) son paquetes de software que permiten crear, publicar y archivar contenidos digitales. En estas plataformas desempeñan un papel fundamental los mecanismos de clasificación de los contenidos informativos. Este trabajo evalúa la implementación de funcionalidades para esquemas de clasificación en estas herramientas. Para ello, se ha definido un método de análisis, con el que se han estudiado diez paquetes, con la finalidad de determinar el soporte a diferentes tipos de esquemas de clasificación, en relación con los contenidos que gestionan. La evaluación se ha llevado a cabo mediante análisis heurístico, tablas de presencia/ausencia de prestaciones, e indicadores numéricos. Los resultados han sido recogidos en tablas comparativas. Las conclusiones obtenidas permiten afirmar que las soluciones aportadas son incompletas y parciales, y no aprovechan todo el potencial subyacente a los esquemas de clasificación.

Palabras-clave sistemas de gestión de contenidos. Esquemas de clasificación. Evaluación.

Abstract The content management systems (*content management systems*, CMS) are software packages that allow to create, publish, and archive digital content. On these platforms play a key role classification schemes. This work evaluates the implementation of features for classification schemes in these tools. To do this, we have defined a test method, and have been studied ten software packages, with the aim of determining the different types of classification schemes supported, in relation to the content that they manage. The evaluation was carried out by heuristic analysis, presence / absence of functionalities, and numerical indicators. The results have been collected in comparative tables. The conclusions obtained are such that the solutions provided are incomplete and partial, and do not exploit the full potential underlying classification schemes.

Keywords Content management systems. Classification schemes. Evaluation.

Planteamiento y objetivos

Los sistemas de gestión de contenidos (*content management systems*, CMS) son el eje fundamental sobre el cual giran las sedes web. Estas aplicaciones permiten desarrollar espacios de información que dan soporte a las más variadas necesidades informativas que puedan tener los usuarios. Se trata de herramientas que ofrecen las prestaciones necesarias para crear, editar, evaluar, publicar y archivar contenidos digitales, en el marco de un flujo de trabajo, y para

administrar el ciclo de vida de la información digital (Tramullas, 2005; Pérez-Montoro, 2005). Las principales tareas que llevan a cabo los usuarios de los CMS son la publicación, administración y consulta de contenidos. Para cualquiera de estos perfiles de usuario, las funciones de clasificación y etiquetado de los contenidos se convierten en un componente nuclear de los servicios, en tanto en cuanto tienen tres usos principales:

1. El establecimiento de estructuras de organización de contenidos y de arquitecturas de la información en los productos y servicios de información digital.
2. La creación de sistemas de navegación, en cuanto se utilizan para crear estructuras de exploración y trayectos.
3. La recuperación de información, en cuanto ofrecen puntos de acceso de alto valor añadido.

En estas plataformas, y dadas las características que ofrecen para la información digital, así como los posibles contextos de uso, desempeñan un papel fundamental los mecanismos disponibles para la clasificación de los contenidos informativos. La creciente importancia de las herramientas de clasificación se aprecia en el auge que están teniendo los esquemas de clasificación en el contexto de la gestión de la información digital, en especial en los últimos años, en los cuales el desarrollo de servicios del web semántico se fundamenta en la utilización de ontologías. Sin embargo, estas herramientas deben complementarse en los CMS con la existencia de las funcionalidades técnicas necesarias para conseguir su correcta implantación y uso. La posibilidad de implementar un tesoro o una clasificación facetada en un sistema de gestión de contenidos, no quiere decir que automáticamente se encuentren disponibles todos los posibles usos deseables del esquema. Este trabajo tiene como objetivo evaluar las funcionalidades con las que los CMS implementan las soluciones técnicas necesarias para utilizar correctamente los esquemas de clasificación.

1 Uso de esquemas de clasificación

En la última década, la utilización de esquemas de clasificación en internet se ha popularizado de manera notable, en particular desde la explosión de los servicios del web 2.0. Los usuarios se han acostumbrado a etiquetar contenidos de todo tipo, principalmente con el objetivo de poder recuperar y acceder a información de su interés. Mai (2004) ya argumentó que la clasificación bibliográfica podía aportar un gran avance a la organización de los contenidos del web, siempre y cuando la investigación en la misma se ajustase a las características del medio digital, y se centrase en clasificaciones basadas en las necesidades y usos de la información. Sin embargo, sólo en entornos controlados se ha procedido a sistematizar el uso de esquemas de clasificación, como en bibliotecas digitales, repositorios, o entornos corporativos.

Gran parte de la actividad de etiquetado que los usuarios llevan a cabo puede identificarse como una aproximación intuitiva a la clasificación facetada. El uso de los principios de las clasificaciones facetadas ha sido analizado por Zins (2002), quien ha estudiado diferentes modelos de organización de información en portales web. Este autor ha ideado Dal Porto, S., & Marchitelli, A. (2006) identificado hasta un total de ocho modelos de clasificaciones facetadas, correspondientes a temas, objetos, aplicaciones, usuarios, localizaciones, referencias, medios y lenguajes. Se pueden encontrar modelos que combinan varios de ellos. Los cinco primeros señalados hacen referencia al contenido informativo, siendo ésta la aproximación más utilizada en los servicios y productos de información digital. Las limitaciones inherentes al modelo de etiquetado plano, y los problemas que conlleva han sido puestos de manifiesto por Noruzi (2006), que ha señalado la necesidad de

perfeccionar el uso de los mismos, integrando principios procedentes de los esquemas de clasificación clásicos.

Garshol (2004), en su revisión de esquemas de clasificación y metadatos de aplicación en el web semántico, señala las diferencias entre vocabulario controlado, taxonomía (o clasificación jerárquica), tesauro, clasificación facetada, ontología y *topic map*. La experiencia de los usuarios de información digital permitiría afirmar que, si bien los dos primeros son los dos más utilizados por su sencillez, como puede ser en comunidades virtuales, foros o blogs, en los entornos de servicios de información más especializados se ha optado por las clasificaciones facetadas y los tesauros, mientras que las ontologías y los *topic maps* continúan usándose sobre todo en entornos y productos muy avanzados y en investigación académica, en especial en aquellos desarrollos de servicios del web semántico.

2 La evaluación de sistemas de gestión de contenidos

No son abundantes las evaluaciones sistemáticas y rigurosas sobre sistemas de gestión de contenidos. La mayoría de la información disponible proviene de informes técnicos parciales, elaborados por profesionales independientes o consultoras que trabajan en el sector. En otras ocasiones, se limitan a estudios de caso, de lo que resulta sumamente dificultoso para extraer datos o aspectos que puedan usarse para estudios comparativos bajo los mismos parámetros. Robertson (2002) estableció un esquema básico de evaluación de sistemas de gestión de contenidos, dentro del cual situaba como una fase fundamental la estructuración previa de requerimientos, atendiendo a las funcionalidades nucleares consideradas como necesarias. Dentro de las mismas ya destacaba un apartado dedicado a los metadatos, identificados como críticos para gestionar el repositorio de información. Dentro de este grupo señalaba el uso de palabras clave, taxonomías o clasificaciones temáticas y mapas de tópicos.

Africa (2008) ha desarrollado un marco de evaluación para CMS libres, *Web Site Quality Evaluation Method*, que sigue el estándar ISO 9126. Tras definir el objetivo y el alcance de la evaluación, establece las características de calidad, las subcaracterísticas y los atributos, así como las métricas de calidad para cada uno de ellos. Finalmente, obtiene un indicador numérico que indica la preferencia global para cada CMS.

Quinn y Gardner-Madras (2010) han llevado a cabo un estudio comparativo general sobre cinco CMS libres, en el que han evaluado funcionalidades básicas sobre instalación y configuración, administración, interfaz gráfica, accesibilidad y optimización para búsquedas, flexibilidad estructural, gestión de usuarios y comunidad, seguridad, extensibilidad y fortaleza de la comunidad. Sólo en un único apartado, concretamente en el de accesibilidad, dedican un indicador simple sobre el uso de metadatos o categorías, sin profundizar más en la cuestión. El Centro de Apoyo Tecnológico a Emprendedores de la Fundación Parque Científico y Tecnológico de Albacete (2012) ha redactado un trabajo con un enfoque similar, en el que amplía a ocho el número de herramientas estudiadas. Sin embargo, en el mismo no se contemplan específicamente los aspectos de clasificación y etiquetado de contenidos.

Aproximación diferente es la adoptada por la consultora Smile (Bécue, 2012a), que, además de un estudio comparativo general sobre varias herramientas (Bécue, 2012b), plantea una evaluación basada en una lista de comprobación de funcionalidades. En las que conciernen a las prestaciones sobre metadatos, se señala la capacidad de capturar metadatos automáticamente, relacionados con el proceso de gestión del contenido, y la de asignar, de manera directa o automática, metadatos descriptivos de contenido informativo, atendiendo al estándar Dublin Core, y a la posibilidad de incorporar metadatos tomados de diferentes taxonomías usadas en el CMS. Señala además la importancia de que éstas, taxonomías o clasificaciones, puedan ser jerárquicas, con la finalidad de permitir una estructuración automática de los contenidos.

3 Metodología

La evaluación de software, a pesar de existir normas ISO (ISO 9126-1) que sirven de guía para determinar su calidad, resulta una tarea compleja y de elevada dificultad. Jadhav y Sonar (2009) han analizado un amplio repertorio bibliográfico sobre la cuestión, concluyendo que, aunque existen una gran variedad de propuestas, métodos y técnicas, una amplia mayoría ha usado un proceso analítico jerárquico, en el marco del cual se suelen asignar valores numéricos para obtener un indicador útil para la toma de decisiones.

El proceso de evaluación que se ha llevado a cabo en este trabajo ya ha sido probado, en una primera iteración simple, por los autores (Tramullas & Garrido, 2006). La evaluación se basa en el estudio analítico de funcionalidades requeridas. El proceso se estructura en las siguientes fases:

1. Definición del objetivo de la evaluación;
2. Establecimiento de las funcionalidades necesarias, de su jerarquía y de su valor;
3. Selección de los CMS a evaluar;
4. Prueba empírica de las funcionalidades disponibles;
5. Síntesis y validación de resultados.

En la primera fase, se ha establecido que el objetivo es evaluar el soporte a la utilización de esquemas de clasificación en sistemas de gestión de contenidos web. La segunda fase ha determinado que, en un primer nivel, los sistemas deben dar soporte a la implementación de una lista de palabras clave, de una clasificación jerárquica, de una clasificación facetada, de un tesauro y de una ontología, y a su utilización en entornos de creación de contenidos. En un segundo nivel se contemplan la posibilidad de usar varias de ellas de forma simultánea en diferentes tipos de contenidos, la generación de arquitecturas de información para estructurar el servicio, y la capacidad para intercambiar esquemas de clasificación con otras plataformas. Cada criterio ha recibido un valor de cumplimiento de 10 puntos. Dal Porto y Marchitelli (2006) usaron un enfoque similar, pero de menor complejidad, y sólo contra un único CMS, *Movable Type*.

La selección de CMS se ha circunscrito a aquéllos que se distribuyen con licencias de software libre. Esto excluye a las plataformas comerciales. Sin embargo, la creciente utilización de software de este tipo, tanto en servidores corporativos, institucionales y personales, y la existencia de grandes comunidades de usuarios, tanto empresariales como individuales, que utilizan y ofrecen soporte para estos productos, hacen más interesante esta clase de herramientas. A ello se une la facilidad con la que es posible implementar una instalación de los mismos para proceder a su análisis y estudio. El análisis se ha llevado a cabo sobre una instalación básica o nuclear de cada paquete de software. La arquitectura de las aplicaciones hace posible que se puedan aumentar y mejorar sus prestaciones con la incorporación de módulos o *plugins* complementarios. Dada la numerosa casuística que puede crearse si se toma en consideración este enfoque, el estudio se ha limitado a las versiones básicas.

CMS	URL de referencia	Versión	Licencia
WordPress	www.wordpress.org	3.5.1	GNU GPL 3
Plone	www.plone.org	4.3	GNU GPL 2
Drupal	www.drupal.org	7.22	GNU GPL 3
Joomla	www.joomla.org	3.1.1	GNU GPL 2
e107	www.e107.org	1.0.3	GNU GPL 2
Typo3	www.typo3.org	6.1	GNU GPL 2
Xoops	xoops.org	2.5.5	GNU GPL 2
phpwcms	www.phpwcms.de	1.6	GNU GPL 2
SPIP	www.spip.net	3.0.8	GNU GPL 3
ez Publish	www.ez.no	5	GNU GPL 2

Tabla 1. Paquetes de software analizados

Para la prueba empírica, se han evaluado diez sistemas que reúnen las características indicadas, y que se recogen en la tabla 1. Todos ellos tienen una asentada presencia en Internet, y numerosos grupos de desarrolladores y usuarios, merced a los cuales existen gran cantidad de sedes web que han implementado las soluciones indicadas.

[Inicio](#) » [Administración](#) » [Estructura](#) » [Taxonomía](#)

Puede reorganizar los términos en *Artículos* usando sus iconos de arrastrar y soltar, y agrupar términos bajo un término padre deslizándolos debajo y a la derecha del padre.

[+ Agregar término](#)

[Mostrar pesos de la fila](#)

NOMBRE	OPERACIONES
+ Boletines electrónicos	editar
+ Congresos y jornadas	editar
+ Convocatorias	editar
+ Cursos y seminarios	editar
+ Divulgación	editar
+ Enlaces a recursos	editar
+ Exposiciones y ferias	editar
+ Premios	editar
+ Publicaciones	editar
+ Subvenciones y ayudas	editar

[Guardar](#)

[Restablecer a alfabético](#)

Fig. 1. Listado de términos en taxonomía en Drupal.

Una vez instalados con la configuración básica, se crearon varias entradas de diferentes tipos de contenido, a crear los diferentes esquemas de clasificación indicados en párrafos superiores, y a asignar términos a los contenidos. Casi todas las herramientas hacen necesario introducir los elementos de la clasificación, términos o categorías con anterioridad a las entradas, ya que se seleccionan desde las pantallas de creación y edición de contenido. Los resultados obtenidos se recogen en la tabla 2.

CMS	Palabras clave	Clasificación Jerárquica	Clasificación Facetada	Tesauro	Ontología
WordPress	SI	SI	NO	NO	NO
Plone	NO	SI	NO	NO	NO
Drupal	SI	SI	SI	NO	NO
Joomla	SI	SI	NO	NO	NO
e107	SI	NO	NO	NO	NO
Typo3	SI	SI	NO	NO	NO
Xoops	NO	NO	NO	NO	NO
phpwcms	SI	SI	NO	NO	NO
SPIP	SI	NO	SI	NO	NO
ez Publish	SI	SI	NO	NO	NO

Tabla 2. Funcionalidades de clasificación básicas.

Los resultados han sido recogidos en tablas comparativas, y evaluados mediante tablas de presencia/ausencia de prestaciones y análisis heurístico. En todos los paquetes estudiados es necesaria la existencia de un perfil de administrador que gestione los esquemas de clasificación, tanto en lo concerniente a su creación y mantenimiento, como a la configuración de reglas y permisos de uso por parte de los usuarios. Algunos de los criterios usados pueden inducir a confusión, Por ejemplo, *Plone* usa como esquema de clasificación jerárquico la noción de “carpeta”; una aproximación similar es también empleada por *e107*, por *ezPublish* y por *Typo3*. Se trata de una aproximación al problema que tiende a replicar, en la interfaz de usuario, ideas clásicas de organización jerárquica de información en ordenadores. El caso de *Xoops* es particular, dado que una instalación básica no permite añadir contenidos, para lo cual es necesario instalar módulos complementarios; en consecuencia, no ofrece funcionalidades básicas para ningún esquema de clasificación.



Fig. 2. Gestión de categorías para artículos en Joomla.

Sólo *Drupal* es capaz de ofrecer funcionalidades que hacen posible crear y usar esquemas facetados, sobre cualquier tipo de contenido. Como se deduce del contenido de la tabla 2, ninguno de los paquetes analizados ofrece soporte para tesauros en sus instalaciones básicas, y ninguno de

ellos está preparado para implementar directamente ontologías y, por extensión, servicios basados en tecnologías del web semántico.

CMS	Búsqueda específica por términos de clasificación	Uso combinado esquemas de clasificación	Generación de arquitecturas de información	Import./export. de esquemas de clasificación
WordPress	NO	SI	SI	NO
Plone	NO	NO	SI (parcial)	NO
Drupal	NO	SI	SI	NO
Joomla	NO	SI	SI (parcial)	NO
e107	NO	NO	NO	NO
Typo3	NO	NO	SI (parcial)	NO
Xoops	NO	NO	NO	NO
phpwcms	NO	NO	SI	NO
SPIP	NO	SI	SI (parcial)	NO
ez Publish	NO	SI (parcial)	SI (parcial)	NO

Tabla 3. Funcionalidades avanzadas para esquemas de calificación.

Las funcionalidades consideradas por el análisis como avanzadas también muestran notables carencias. Sólo *WordPress*, *Joomla* y *Drupal* son capaces de combinar esquemas de clasificación, aunque en el caso de los dos primeros son etiquetas simples y una clasificación jerárquica, mientras que el tercero, *Drupal*, permite combinar simultáneamente varios esquemas, y discriminar por tipo de contenido si resulta necesario. La generación de arquitecturas de información, que parece una funcionalidad bastante común, en realidad se beneficia de aquellos paquetes que usan la aproximación de la “carpeta” como esquema jerárquico (marcados como SI parcial), indicados en el párrafo anterior. Resulta destacable que, pese a emplear esquemas de clasificación, ninguno de los paquetes analizados ofrezca directamente la posibilidad de usar los esquemas de clasificación en sus interfaces de búsqueda.

CMS	Funcionalidades básicas	Funcionalidades avanzadas	Indicador numérico final
WordPress	20	20	40
Plone	10	5	15
Drupal	30	20	50
Joomla	20	15	35
e107	10	0	10
Typo3	20	5	25
Xoops	0	0	0
phpwcms	20	10	30
SPIP	20	15	35
ez Publish	20	10	30

Tabla 4. Indicadores cuantitativos.

La tabla 4 muestra los resultados obtenidos de asignar los valores de cumplimiento de los requisitos, sobre un total de 90 puntos posibles. *Drupal* es el paquete que ofrece más prestaciones y flexibilidad. Tras éste, viene un segundo grupo formado por *WordPress*, *Joomla* y *SPiP*, que muestran capacidades similares, siendo más amigable la interfaz del primero. Teniendo en cuenta que el valor máximo posible era de 90, los paquetes con valores por debajo de 30 muestran notables carencias en sus versiones básicas de instalación.



Fig. 3. Gestión de secciones y etiquetas en SPiP.

Conclusiones

La evaluación desarrollada permite extraer conclusiones de gran interés para analizar el estado de la integración de las funcionalidades para clasificación de contenidos en los sistemas de gestión de contenidos:

- En primer lugar, resulta obligatorio destacar que los sistemas de gestión de contenidos web existentes dan, por el momento, más importancia a los procesos de publicación de contenidos y de presentación de usuario final, que a las funcionalidades de etiquetado y clasificación y de recuperación de información fundamentadas en esquemas de clasificación.
- Las funcionalidades disponibles para la implementación de herramientas de clasificación son muy heterogéneas, oscilando desde simples esquemas jerárquicos, hasta la posibilidad de implantar esquemas facetados.

- Las tareas de gestión de los esquemas de clasificación se entienden como parte de la necesaria administración general y configuración del sistema que se trate.
- Muchos sistemas tienden a identificar los esquemas de clasificación jerárquicos con la jerarquía de carpetas y ficheros de los sistemas operativos.
- Los esquemas de clasificación no se han integrado con las interfaces de búsqueda de información.
- No existe la posibilidad de utilizar ontologías, o usar otras aproximaciones al web semántico.

Como conclusión general, puede afirmarse que la actual generación de sistemas de gestión de contenidos no ofrecen todas las prestaciones deseables para las herramientas de clasificación, y presentan además un panorama muy heterogéneo. Gran parte de los servicios de información más avanzados construidos sobre estos paquetes, que muestran un uso más desarrollado de esquemas de clasificación en su arquitectura de la información, han hecho uso de la instalación y configuración de módulos complementarios que aumentan las funcionalidades y prestaciones para los esquemas de clasificación. Esto supone la necesidad de expertos y desarrollo especializado para implementar sistemas de gestión de contenidos que aprovechen todo el potencial que ofrecen los esquemas de clasificación.

Referencias

- AFRICA, A.D.M. (2008). Quantitative Evaluation of Open Source Content Management Systems. *IEEE Multidisciplinary Engineering Education Magazine*, 3(2), 45-51.
- BÉCUE, G. (2012b). *200 Questions pour choisir un CMS*. Smile Open Source Solutions. Recuperado de: <<http://www.smile.fr/Livres-blancs/Gestion-de-contenu-et-GED/Choisir-un-CMS>>.
- BECUÉ, G. (2012b). *Gestion de contenus. Le meilleur des solutions open source*. Smile Open Source Solutions. Recuperado de: <<http://www.smile.fr/Livres-blancs/Gestion-de-contenu-et-GED/Les-CMS-open-source>>.
- CENTRO DE APOYO TECNOLÓGICO A EMPRENDEDORES, FUND. PARQUE CIENTÍFICO Y TECNOLÓGICO DE ALBACETE. (2012). *Estudio de los sistemas de gestión de contenidos web. Análisis de las mejores soluciones del mercado (2012)*. Junta de Comunidades de Castilla-La Mancha. Recuperable en: <<http://www.bilib.es/noticias/noticia/articulo/estudio-de-los-sistema-de-gestion-de-contenidos-w/>>.
- DAL PORTO, S., & MARCHITELLI, A. (2006). Functionality and flexibility of traditional classification schemes applied to a Content Management System (CMS): facets, DDC, JITA. *Knowledge Organization*, 33(1), 35-44.
- GARSHOL, L.M. (2004). Metadata? Thesauri? Taxonomies? Topic Maps! *Journal of Information Science*, 30(4), 378-391.
- JADHAV, A.S. & SONAR, R.M. (2009). Evaluating and selecting software packages: A review. *Information and Software Technology*, 51(x), 55-563.

- MAI, J-E. (2004). Classification on the Web: Challenges and Inquiries. *Knowledge Organization*, 31(2), 92-97.
- NORUZI, A. (2006). Folksonomies: (Un)Controlled Vocabulary? *Knowledge Organization*, 33(4), 199-203
- PÉREZ-MONTORO, M. (2005). Sistemas de gestión de contenidos en la gestión del conocimiento. *bid textos universitarios de biblioteconomía i documentació*, 14. Recuperado de: <http://www2.ub.es/bid/consulta_articulos.php?fichero=14monto2.htm>.
- QUINN, L. & GARDNER-MADRAS, H. (2010). *Comparing Open Source Content Management Systems: WordPress, Joomla, Drupal and Plone*. Idealware. Recuperado de: <http://www.idealware.org/reports/2010-os-cms>
- VITARI, C. RAVARINI, A. & RODHAIN, F. (2007). An Analysis Framework for the Evaluation of Content Management Systems. *Communications of the Association for Information Systems* 18(1). Recuperado de: <http://works.bepress.com/aurelio_ravarini/5>.
- ROBERTSON, J. (2002). *How to evaluate a content management system*. Step Two Designs. Recuperado de: <http://www.steptwo.com.au/papers/kmc_evaluate/index.html>.
- TRAMULLAS, J. (2005). Open Souce Tools for Content Management. *Hipertext.net*, 3, Recuperado de: <<http://www.upf.edu/hipertextnet/en/numero-3/software-libre.html>>.
- TRAMULLAS, J. & GARRIDO, P. (2006). Weblogs content classification tools: performance evaluation. *I Internacional Conference on Multidisciplinary Information Sciences & Technologies: Current Research on Information Sciences and Technologies, Multidisciplinary Approaches to Global Information Systems*, vol. I. (p. 532-536). Mérida.
- ZINS, C. (2002). Models for Classifying Internet Resources. *Knowledge Organization*, 29(1), 20-28.