

Arquitectura y prestaciones de los sistemas de gestión de contenidos

Dr. Jesús Tramullas

Dep. Ciencias de la Documentación, Univ. de Zaragoza

<http://tramullas.com>

0. Resumen

Los sistemas de gestión de contenidos (*Content Management Systems*, o CMS) se han convertido en los últimos años en la herramienta estándar para la creación y edición de información digital en servicios y unidades de información. Estos sistemas, que en ocasiones se ven como meras herramientas de publicación digital, muestran, en realidad, una arquitectura interna modular compleja, y características propias que los diferencian de otros sistemas de publicación, como puedan ser los weblogs o blogs. De la misma forma, las funcionalidades y prestaciones que ofrecen hacen posible extender el proceso de gestión de la información digital más allá de los cánones clásicos, permitiendo crear verdaderos sistemas integrales de gestión de información. Los sistemas para repositorios digitales y para bibliotecas digitales, los esquemas de metadatos y ontologías, los integradores de información, son expresiones avanzadas de la integración entre la gestión de contenidos y la gestión de información.

1. De la gestión de contenidos a la gestión integral de la información

El procedimiento tradicional de gestión de la información ha centrado su actividad en las tareas de recopilación, descripción y tratamiento y difusión de la información, siendo especialmente en estas fases finales en las que se llevaba a cabo el trabajo de valor añadido que era aportado por los especialistas en información y documentación. Sin embargo, con el acelerado desarrollo de la Sociedad de la Información y del Conocimiento, los cauces tradicionales de producción y publicación de información se han visto ampliados de manera exponencial, ya que no sólo ha aumentado la información producida y disponible en sí misma, como ya habían previsto los especialistas que ocurriría, sino que este incremento ha llevado parejo el aumento de los medios y técnicas de producción de la misma, así como la multiplicidad de medios de acceso a la misma.

Los sistemas de gestión y publicación digitales han estado disponibles desde la década de 1970, y han ido evolucionando hacia sistemas cada vez más desarrollados y potentes. Sin embargo, en el momento actual la publicación digital se entiende como un producto

de alto valor añadido, que puede ser preparado para y distribuido por diferentes canales. Esto supone que el esquema clásico de publicación también ha sufrido significativas variaciones, entre las cuales debe destacarse la posibilidad de hacer múltiples versiones del mismo documento, lo cual tiene, evidentemente, interesantes implicaciones para la reutilización de contenidos y la personalización de servicios para el usuario final. El proceso de publicación clásico consistía en la obtención de un contenido original, su preparación (maquetación) y corrección (ortográfica y de estilo), y su impresión final, con lo que terminaba el mismo. Actualmente el proceso de publicación digital se configura como más complejo, ya que atiende a factores más variados para la obtención de contenido original, la intervención de diferentes profesionales en el proceso, la adecuación del producto a diferentes entornos, etc. En este proceso complejo y dinámico es donde más se ha ido incorporando valor añadido, en cuanto los usuarios demandan cada vez mayores prestaciones y servicios. El mero acceso a la información, que podría ser el objetivo hace unos pocos años, ya no parece suficiente.

Los dos procesos de gestión, de información y de publicación, señalados en los párrafos anteriores, ha desembocado en una disciplina a la que se ha dado en llamar gestión de contenidos. La gestión de contenidos, entonces, se configura como un proceso informativo en entorno digital, cuyo resultado es un producto de información digital, que ofrece, junto al propio contenido, elementos de valor añadido. Puede argumentarse que esta definición no añade nada nuevo al proceso de publicación, sino que es meramente una cuestión de escala y entorno. Sin embargo, el proceso de gestión de contenidos tiene uno de sus fundamentos en la información de soporte al propio proceso. Como se verá más adelante, el proceso de publicación digital inherente a la gestión de contenidos incluye el tratamiento informativo-documental de esos contenidos, mediante el uso de diferentes tipos y esquemas de metadatos. En realidad, todo el proceso de gestión de contenidos se soporta en la gestión de flujos de trabajo, que permite definir los procesos, las actividades y tareas, los responsables de las mismas, las fuentes de información y los puntos de control, y en el proceso informativo-documental inherente al contenido informativo que se publica, y al de todas las tareas incluidas en la gestión de flujos de trabajo. Los productos de información digital resultantes llevan incluidas, además, prestaciones y características de localización y acceso a la información que anteriormente era productos diferenciados, propios de la gestión de información. Ahora es común encontrar publicaciones y documentos digitales dinámicos, que incorporan en su misma interfaz información externa relacionada, esquemas de clasificación, herramientas de valoración y recomendación de contenidos, sistemas de alerta, etc. De esta forma, gran parte de los servicios de valor añadido, procedentes de la actividad informativo-documental, que antes eran servicios independientes, se han visto ahora integrados directamente en los propios documentos a los que antes referenciaban, rompiendo los límites considerados tradicionalmente como válidos. Las técnicas de la gestión de contenidos, y los sistemas y herramientas que la hacen posible, han facilitado que los principios y métodos de la gestión de información

se extiendan a toda la actividad informativa que se desarrolla en internet y el world wide web, de la misma forma que están permitiendo que la actividad de los especialistas en información y documentación se extienda más allá de los límites que tenían fijados tradicionalmente.

2. La gestión de contenidos

En la segunda mitad de la década de los 90 se podía diferenciar entre productos para gestión documental, para recuperación de información, etc., Desde el año 2000 se ha producido una convergencia entre todas las plataformas, de forma que en la actualidad se pueden encontrar soluciones que pretenden ser globales y ofrecer soporte a todo el proceso de gestión de información en una organización. En este contexto deben situarse la gestión de contenidos. Los sistemas de gestión de contenidos están orientados a gestionar objetos que actúan como componentes de documentos virtuales, en el contexto de lo que llama segmentación. Como punto de partida, baste señalar que los sistemas de gestión de contenidos deben ser capaces de gestionar toda la información no estructurada de una organización. Los objetos informativos, los contenidos, son tratados mediante un conjunto de procesos estructurados con la finalidad de producir publicaciones digitales basadas en la metáfora del documento. Este objetivo de la producción de publicaciones y documentos digitales por integración de componentes se puede encontrar en las principales referencias sobre gestión de contenidos (BOIKO, 2001; BROWNING y LOWNDES, 2001).

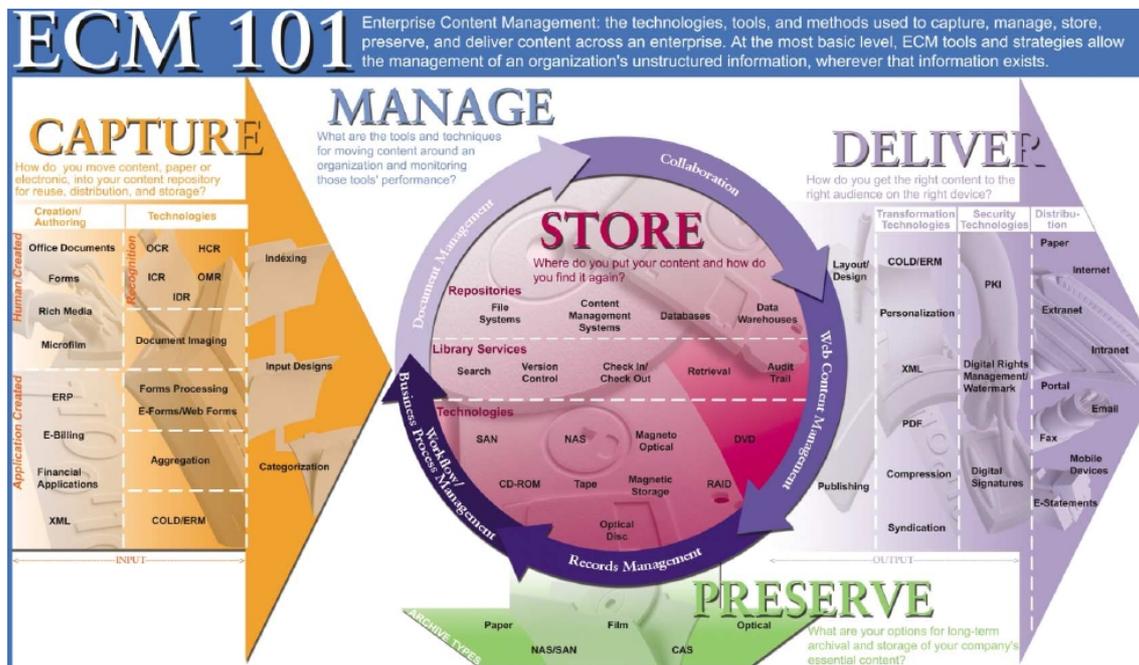


Fig. 1. El proceso de gestión de contenidos (según AIIM, 2005)

El proceso de publicación digital más sencillo define un proceso en el cual se diseña un marco de integración para un conjunto de objetos, cuyo resultado final es la publicación del documento digital. Este proceso se complica si se sitúa en el contexto de una unidad o sistema de información que debe crear información digital más compleja, y que, además, debe servir a comunidades de usuarios especializados, o a diferentes comunidades de usuarios. Los productos de información digital se vuelven más complejos, y los procesos técnicos que se desarrollan sobre ellos se hacen igualmente más complejos (ROBERTSON, 2003). El trabajo en colaboración obliga a desarrollar flujos de trabajo en los que es necesario asignar diferentes roles a las personas que intervienen en el mismo. Los objetos que componen los documentos pueden proceder de diferentes fuentes y tener diferentes formatos, con lo que es necesario controlar repositorios de información. Además, los usuarios demandan servicios de apoyo a los productos de información digital como la búsqueda de información o el acceso a documentos generados anteriormente, lo que hace necesario disponer de herramientas de archivo digital. Los cada vez más complicados procesos de gestión de contenidos demandan la formalización de sistemas de metadatos que resultan necesarios tanto para los procesos de gestión como para los de publicación o los de recuperación de información.

Esta complejidad ha producido el desarrollo de dos especializaciones principales de la gestión de contenidos: *Web Content Management*, WCM, y *Enterprise Content Management*, ECM. La primera especialización (MILLER y DAVID, 2002) se centra en el entorno web, y su objetivo principal y sus métodos se orientan a la producción de documentos e información digitales para el web, especialmente para portales y webs corporativos. La segunda (GINGELL, 2003) encuentra su fundamento en la idea de gestión total de la información en las organizaciones mediante la integración de toda la información necesaria para alcanzar los objetivos de la organización proveniente no sólo de sistemas de publicación, sino también de sistemas ERP, gestión de documentos, datawarehouses, etc. La expresión más generalizada del ECM son las intranets y los portales internos de las organizaciones. En todos los casos, la gestión de contenidos tiene como factor ineludible el etiquetado de la información en formatos XML, o en lenguajes específicos para cada contexto, derivados del mismo. A esto es necesario añadir la cada vez mayor necesidad de gestionar lo que se ha dado en llamar *digital assets*, documentos multimedia en formato digital, que es necesario almacenar, describir, integrar y gestionar, con especial referencia a la gestión de los derechos inherentes a los mismos.

3. Arquitectura y componentes de los sistemas de gestión de contenidos

La bibliografía disponible sobre sistemas de gestión de contenidos coinciden en señalar que un sistema de este tipo debe ofrecer, como mínimo, una aplicación nuclear CMS

que soporte los procesos de publicación, de flujo de trabajo y de repositorios de información; un repositorio de información; herramientas de integración de información externa; y modelos y *templates* para productos finales. la arquitectura técnica de un sistema de gestión de contenidos debe mostrar varios subsistemas (BOIKO, 2001) que interaccionan entre ellos:

- Subsistema de colección: encargado de la creación y/o adquisición de información. Debe dar soporte a los procesos de creación de contenidos, soporte a flujos de trabajo, sindicación e integración de fuentes externas. Además, debe ofrecer soporte a procesos de conversión entre formatos diversos, y a la agregación de contenidos de fuentes diversas en estructuras específicas.
- Subsistema de gestión: encargado de la gestión y control de los repositorios de información, de los grupos de usuarios, y de los procesos de soporte para los otros subsistemas. Se encarga de definir y controlar los flujos de trabajo que son utilizados por los otros subsistemas, y de la definición de parámetros para el funcionamiento del sistema.
- Subsistema de publicación: encargado de la producción final de publicaciones o productos de información digital, de manera automática o casi automática. Utiliza un modelo basado en *templates* y deberá ofrecer posibilidades de personalización para usuarios y la posibilidad de producir para diferentes tipos de plataformas y/o clientes.

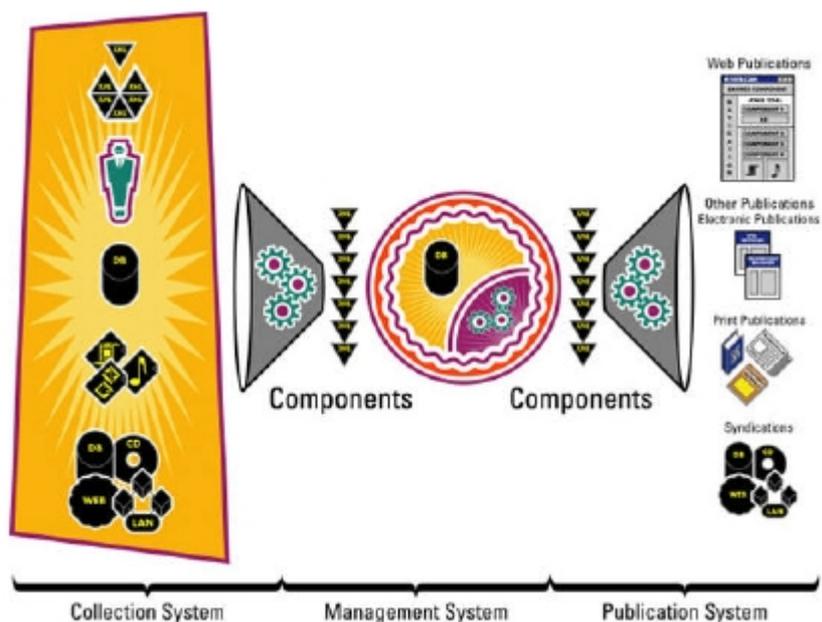


Fig. 2. Los tres subsistemas de los CMS (según Boiko, 2002)

Otros autores han indicado la existencia de diferentes componentes en los sistemas de gestión de contenidos. Nakano (2002) señala cuatro subsistemas, correspondientes a

creación/edición de contenidos, repositorio, flujo de trabajo y gestión de operaciones. Browning y Lowndes (2001) señalan autoría/creación, flujo de trabajo, almacenamiento/repositorio y publicación. En cualquier caso, la dinámica de los productos y aplicaciones que se engloban bajo la etiqueta de gestión de contenidos ha hecho posible la aparición de herramientas que siguen diferentes enfoques y que, en consecuencia, ofrecen diferentes prestaciones. Dada la importancia que la elección e implantación de una herramienta de este tipo tiene para la organización que se trate, se han realizado detallados estudios que evalúan las prestaciones y características de los productos disponibles. De estos *toolkits* de evaluación para CMS pueden deducirse las prestaciones básicas que deben ofrecer. Por ejemplo, véanse los de Step Two (2004), CMS-Watch (2005), o la Guía CMS-Spain (CMS-SPAIN, 2005).

Siguiendo a Jennings (2004), las principales funcionalidades que debe ofrecer un sistema de gestión de contenidos deben responder a:

- Creación y autoría de contenidos
- Soporte de *templates*
- Captura de documentos
- Repositorio de contenidos
- Flujos de trabajo
- Desarrollo de sedes web
- Indización y búsqueda
- Categorización
- Control de versiones
- Personalización
- Publicación en web
- Distribución de contenidos
- Publicación multicanal
- Colaboración
- Sindicación de contenidos
- Agregación de contenidos
- Seguridad de recursos digitales
- Gestión del ciclo de vida del contenido
- Gestión de documentos (*records management*)

- Integración de aplicaciones

La selección, implantación y puesta en marcha de una herramienta para gestión de contenidos es fruto de un estudio y de un análisis detallado de la organización que lo instala, de los objetivos de la misma, de los procesos de trabajo y recursos de información que utiliza, y de los usuarios que van a usarlo. En consecuencia, una exhaustiva planificación, resultado del estudio previo, y una rigurosa metodología de puesta en marcha son ineludibles para tener éxito con cualquier iniciativa de gestión de contenidos. Si, se revisa, por ejemplo, la monografía de Rockley (2003), obra de referencia para ECM, puede apreciarse que gran parte está dedicada a los aspectos relacionados con el análisis, planificación y metodología de selección, integración, implantación y puesta en marcha. Nakano centra gran parte de su texto (2002) en la organización y dinamización del trabajo colaborativo en comunidades estructuradas alrededor del web.

Los especialistas en gestión de la información y la documentación están encontrando en los sistemas de gestión de los contenidos en las organizaciones nuevos nichos en los que desarrollar su labor y practicar sus habilidades. Las tareas de tratamiento especializado de información y de contenido, el diseño de clasificaciones y taxonomías para etiquetar contenidos, y el desarrollo de esquemas y sistemas de metadatos pueden ser abordadas mediante la aplicación de los conocimientos básicos adquiridos durante los procesos formativos. En los próximos años se producirá un completo cambio de paradigma, desde una posición secundaria, pasiva y receptora, hacia una posición central, activa y dinámica, desde la cual ponen en valor los recursos de información de las organizaciones.

Bibliografía

BLUEBILL ADV. The Classification & Evaluation of Content Management Systems. En: The Gilbane Report, vol. 11, n. 2, 2003, (disponible en http://www.gilbane.com/gilbane_report.pl/86/The_Classification_Evaluation_of_Content_Management_Systems.html; consultado 22-1-2005)

BOIKO, B. Content Management Bible. Wiley, 2001.

BROWNING, P. y LOWNDES, M. JISC TechWatch Report: Content Management Systems. TechWatch Report TSW 01-02, The Joint Information Systems Committee, 2001.

CMS-SPAIN. Guía CMS-Spain 2004. (disponible en <http://www.ecm-spain.com/interior.asp?IdItem=1851>; consultado 24-11-2004)

CMS-WATCH. *The CMS Report: In-Depth Analysis of Web Content Management Solutions*. 2005 (disponible en <http://www.cmswatch.com/TheCMSReport/>; consultado 3-2-2005)

GINGELL, D. *A 15 Minutes Guide to Enterprise Content Management*. Documentum Inc. 2003.

GUPTA, V.K., GOVINDARAJAN, S., JOHNSON, T. *Overview of content management approaches and strategies*. En: *Electronic Markets*, vol. 11, n. 4, 2001, p. 281-287.

DOYLE, B. *Open Source Content Management Redux*. En: *The Gilbane Report*, vol. 11, n. 3 2003 (disponible en http://www.gilbane.com/gilbane_report.pl/87/Open_Source_Content_Management_Redux.html; consultado 2-2-2005)

ERP SOFTWARE, *Content Management Tutorial*, 2003 (disponible en <http://erptoday.com/CMS/Content-Management-Tutorial.aspx>; consultado 15-01-2005)

GILBANE, F. *Whats content management?* En: *The Gilbane Report*, vol. 8, n. 8, 2000 (disponible en http://www.gilbane.com/gilbane_report.pl/6/What_is_Content_Management; consultado 24-1-2005)

GINGELL, D. *A 15 Minute Guide to Enterprise Content Management*. EMC Documentum, 2003.

GROFF, T.R., JONES, T.P. *FileNet. A Consultant's Guide to Enterprise Content Management*. Burlington: Elsevier Butterworth-Heinemann, 2004.

HAN, Y. *Digital content management: the search for a content management system*. En: *Library Hi Tech*, vol. 22, n. 4, 2004, p. 355-365.

HOLST, S. *Open Source Content Management: A Parallel Universe?* En: *The Gilbane Report*, vol. 9, n. 4, 2001 (disponible en http://www.gilbane.com/gilbane_report.pl/1/Open_Source_Content_Management_A_Parallel_Universe.html; consultado 22-1-2005)

JENNINGS, T. *Defining the Document and Content Management Ecosystem*. Butler Direct Limited, 2002.

MCKEEVER, S. (2003), "Understanding web content management systems: evolution, lifecycle and market." En: *Industrial Management & Data Systems*, vol. 103, n. 9, p. 686-692.

MILLER, B. y DAVID, D. *Directions in Web Content Management*. Burntsand, Inc., 2002.

NAKANO, R. *Web Content management. A Collaborative Approach*. Prentice Hall, 2002.

ROBERTSON, J. So, what is a content management system? En: KM Column, June 2003 (disponible en http://www.steptwo.com.au/papers/kmc_what/index.html; consultado 1-2-2005)

ROCKLEY, A. *Managing Enterprise Content. A Unified Content Strategy*. New Riders, 2003.

SERENA SOFTWARE. *The 10 Critical Elements in Choosing a Content Management Platform*. 2004.

STEP TWO. *Content Management Requirements Toolkit, Step Two*, Sydney, 2004.

WHITE, M. *Selecting a content management system*. En: VINE, vol. 32, n. 2, 2002, p. 34-39.