

Gestió de continguts

Documents de lectura

PID_00154839

Material docent de la UOC



Universitat Oberta
de Catalunya

www.uoc.edu

Primera edició: febrer 2010
© Fundació per a la Universitat Oberta de Catalunya
Av. Tibidabo, 39-43, 08035 Barcelona
Disseny: Manel Andreu
Dipòsit legal: B-9.555-2010

Cap part d'aquesta publicació, incloent-hi el disseny general i de la coberta, no pot ser copiada, reproduïda, emmagatzemada o tramesa de cap manera ni per cap mitjà, tant si és elèctric, com químic, mecànic, òptic, de gravació, de fotocòpia, o per altres mètodes, sense l'autorització prèvia per escrit dels titulars del copyright.

Continguts

Lectures bàsiques

1.

History, Digitized (and abridged)

By Katie Hafner

Published: March 10, 2007

Disponible en línia a:

<http://www.nytimes.com/2007/03/10/business/yourmoney/11archive.html>

2.

La memoria del periodismo

Anàlisi: quaderns de comunicació i cultura, Any: 2006 Núm.: 33

Alicia Tapia, Nereida López, Elena Medina, Pedro Gómez Fernández

Disponible en línia a:

<http://www.raco.cat/index.php/Analisi/article/viewArticle/51744/0>

3.

García Avilés, José Alberto and Masip Masip, Pere and Micó Sanz, Josep Lluís.

La redefinición del perfil y funciones del documentalista en las redacciones digitales de medios españoles, 2007 . In IX Jornadas de Gestión de la Información, Madrid (Spain), 22-23 November 2007. [Conference Paper]

4.

Cómo aprovechar su gestor de contenidos para tener una web bien posicionada en buscadores.

24/11/2006 - Autor: Fernando Maciá

Disponible en línia a:

<http://www.humanlevel.com/recursos.asp?IdNoticia=551>

5.

Herramientas de software libre para la gestión de contenidos

Autor: Jesús Tramullas

Situación en la jerarquía: Inicio -> Números publicados -> Núm. 3, mayo 2005 -> Software Libre

Citación: Jesús Tramullas. Herramientas de software libre para la gestión de contenidos [en línia]. "Hipertext.net", núm. 3, 2005.

<<http://www.hipertext.net>> [Consulta: 22/09/2009]. ISSN 1695-5498

Disponible en línia a:

<http://www.hipertext.net/web/pag258.htm>

6.

XML y la gestión de contenidos

Autor: Ricardo Eíto Brun Situació en la jerarquia: Inicio -> Números publicados -> Núm. 3, mayo 2005 -> XML

Citació: Ricardo Eíto Brun. XML y la gestión de contenidos [on line].

"Hipertext.net", núm. 3, 2005. <<http://www.hipertext.net>>

[Consulta: 22/09/2009]. ISSN 1695-5498

Disponible en línia a:

<http://www.hipertext.net/web/pag256.htm>

Lectures complementàries

1.

The classification & evaluation of content management systems

The Gilbane Report Vol. 11, No. 2 March, 2003 www.gilbane.com

Disponible en línia a:

<http://gilbane.com/artpdf/GR11.2.pdf>

2.

Diseño y creación de la base de datos documental del "Grupo Godó"

Autors: Maita Corbera, Ernest Abadal Falgueras, Andreu Sulé Duesa, Miquel Llevat, Mònica Sancho, Carles Salmurri

Localització: El profesional de la información, ISSN 1386-6710, Vol. 11, Núm. 3, 2002, pàg. 195-204

Disponible en línia a:

<http://dialnet.unirioja.es/servlet/articulo?codigo=243271>

3.

Sistemas de gestión de contenidos en la gestión del conocimiento

Mario Pérez-Montoro Gutiérrez

Estudis de Ciències de la Informació i de la Comunicació

Internet Interdisciplinary Institute (IN3)

Universitat Oberta de Catalunya

Textos universitarios de biblioteconomía i documentació

número 14 juny de 2005

Disponible en línia a:

http://www2.ub.es/bid/consulta_articulos.php?fichero=14monto2.htm

History, Digitized (and Abridged)

By KATIE HAFNER

Correction Appended

#SALINAS, Calif.

THE National Steinbeck Center, at the top of Main Street in this farming community, exhibits an array of artifacts from John Steinbeck's life and works: family memorabilia, a passport from the 1960s and movie stills from "The Grapes of Wrath." Downstairs, in a climate-controlled vault, is the original manuscript of "The Pearl," his novella published in 1947. There is also an exuberant letter that Steinbeck wrote to a distant relative when he was a teenager, as well as rare footage of him on 16-millimeter film, introducing a 1961 movie, "Flight."

Steinbeck aficionados wishing to examine the manuscript of "The Pearl," which he wrote in pencil in small, precise handwriting on a yellow legal pad, have to travel here — after making an appointment with a part-time archivist, who is in on Mondays, Wednesdays and Fridays.

The center takes great care to preserve these relics of Steinbeck, a Nobel laureate, yet it has no plans to take the collection a step further, to adapt to a digital age. As a result, the manuscript of "The Pearl" is no more likely to be digitized than is the camper with the canine-motif curtains that Steinbeck immortalized in his book "Travels With Charley," and that is parked in perpetuity in the center's main exhibition hall.

These Steinbeck artifacts are not the only important pieces of history that are at risk of disappearing or being ignored in the digital age. As more museums and archives become digital domains, and as electronic resources become the main tool for gathering information, items left behind in nondigital form, scholars and archivists say, are in danger of disappearing from the collective cultural memory, potentially leaving our historical fabric riddled with holes.

"There's an illusion being created that all the world's knowledge is on the Web, but we haven't begun to glimpse what is out there in local archives and libraries," said Edward L. Ayers, a historian and dean of the college and graduate school of arts and sciences at the University of Virginia. "Material that is not digitized risks being neglected as it would

not have been in the past, virtually lost to the great majority of potential users."

To be sure, digitization efforts over the last 10 years have been ambitious and far-reaching. For many institutions, putting collections online, for both preservation and accessibility, is a priority. Yet for every letter from Abraham Lincoln to William Seward that can be found online, millions of documents bearing fine-grained witness to the Civil War will never be digitized. And for every CD re-release of Bessie Smith singing "Gimme a Pigfoot," the work of hundreds of lesser-known musicians from the early 20th century are unlikely to be converted to digital form. Money, technology and copyright complications are huge impediments.

It is not for a lack of trying.

At the Library of Congress, for example, despite continuing and ambitious digitization efforts, perhaps only 10 percent of the 132 million objects held will be digitized in the foreseeable future. For one thing, costs are prohibitive. Scanning alone on smaller items ranges from \$6 to \$9 for a 35-millimeter slide, to \$7 to \$11 a page for presidential papers, to \$12 to \$25 for poster-size pieces. (The cost of scanning an object can be a relatively minor part of the entire expense of digitizing and making an item accessible online.)

Similarly, at the National Archives, the repository for some nine billion documents, only a small fraction are likely to be digitized and put online. And at thousands of smaller, local collections around the country, the bulk of the material is languishing on yesterday's media: paper, LPs, magnetic tape and film.

Strapped for money, archivists around the country are looking to private partners for help. Google has donated \$3 million to help start an effort led by the Library of Congress that will digitize and share materials around the globe, and has also provided technical resources for digitizing various printed materials at the library. Google, on its own, is digitizing books at the Library of Congress, which has its hands full with other items. And a number of other companies and foundations, including Reuters, I.B.M. and the Andrew W. Mellon Foundation, have financed digitization projects around the world.

Even with outside help, experts say, entire swaths of political and cultural history are in danger of being forgotten by new generations of amateur researchers and serious scholars.

Consider the Library of Congress archive of one million photo prints from The New York World-Telegram & Sun; only 5,407 have been digitized. Of the 1.2 million images from U.S. News and World Report, the library has digitized only 366. Its collection of five million images from Look magazine, spanning the period from 1937 to 1971, creates what Jeremy E. Adamson, director of collections and services at the library, calls "a fascinating portrait of America through photo stories on social and political subjects, personalities, food, fashion and sports." Yet only 313 of those images have been digitized.

"It's a crying shame," Mr. Adamson said, "as today's public is acutely visually literate and comfortable with pictures as a means to understand the past and experience for themselves the direct look and feel of history."

The reason for not digitizing these collections? "Not enough money," Mr. Adamson said.

THE decision to put off digitizing a significant collection is seldom easy, archivists at the Library of Congress say. Plans to digitize The National Intelligencer, a newspaper published in Washington during much of the 19th century and filled with Colonial script not easily recognized by digitizing equipment, eventually had to be put on hold because of the high expense.

"If researchers conclude that the only valuable records they need are those that are online they will be missing major parts of the story," said James J. Hastings, director of access programs at the National Archives. "And in some cases they will miss the story altogether."

Maritime buffs, for example, hoping to use the Internet to piece together the story of the Silenus, one of the finest ships ever built in North America, will find a spotty narrative. The papers of its captain, Joseph King, who lived a brief but adventurous life, from 1782 to 1806, can be found courtesy of the Mellon Foundation, in a digitized archive from the Mystic Seaport's collection. Researchers will see how much Captain King paid for "1 potte lijn oli" in 1803, when the ship was in the Netherlands.

What they will not see is that two years after Captain King's death, at the Cape of Good Hope, the ship itself was advertised for sale on May 4, 1808, in Calcutta. This clue remains paperbound, on the front page of The Asiatic Mirror, an English-language newspaper published in Calcutta during that era, whose only known remaining copies now reside in large bound volumes in a remote storage room outside

Washington. The relative obscurity of the newspaper, and its odd size, make it impractical to digitize.

A Google search will pick up the next chapter of the story at the Princeton University's special collection, which includes the papers of James and Dolley Madison. It reveals that in 1817, President Madison signed over the ship's papers to William Gallup.

"The story of what happened to the good ship Silenus between 1806 and 1817 will never be complete," said Mr. Adamson of the Library of Congress, "but what happened in 1808 in Calcutta is the kind of little crumb that can be picked up and become a significant research item."

The ultimate fate of information relating to potentially valuable but obscure people, places, events or things like the Silenus highlights one of the paradoxes of the digital era. While the Internet boom has made information more accessible and widespread than ever, that very ubiquity also threatens records and artifacts that do not easily lend themselves to digitization — because of cost, but also because Web surfers and more devoted data hounds simply find it easier to go online than to travel far and wide to see tangible artifacts.

"This is the great problem right now, and it's a scary thing," said the documentary filmmaker Ken Burns. "The dots are only connected by a few of us who are willing to go to the places to make those connections."

In its digitization efforts, the Library of Congress is focusing mainly on special collections, hewing to a philosophy that it should be digitizing objects that cannot be seen elsewhere. There are the obvious things, like the papers of Washington, Jefferson and Lincoln. And then there are the Farm Security Administration's collection of photographs from the Depression, and a set of mounted photographs of the America's Cup yacht race since the 1890s.

Elizabeth S. Dulabahn, a senior manager at the Library of Congress who oversees part of the library's digitizing effort, said the library was examining closely the behavior of those who use its Web site.

"We're trying to do a better job of understanding the kinds of information that people are looking for on the Web, and the kinds of searches that bring users to the library's site," she said. She cited Women's History Month and the centennial of the first Wright Brothers flight as "examples of events of interest to a broad constituency."

The Library of Congress and other archives are creating indexes that refer to the contents of a physical collection, in the hope that they will entice researchers away from their computers.

But the reality remains that a new generation of researchers prefers to seek information online, a trend made all too clear to Mr. Hastings of the National Archives last year, after Google, in an experiment of sorts, digitized 101 of the National Archives' films — including World War II newsreels and NASA footage — and put them up on its site, at video.google.com/nara.html.

"Before that happened, we had 200 requests total for the whole year in our research room," Mr. Hastings said. "The first month the films were available on Google, there were about 200,000 hits on them — a thousandfold increase."

In some cases, strange bedfellows have conspired to help solve the problem.

Over the years, the New Orleans Public Library has steadily been digitizing its photographs, but its documents have gone largely untouched. The collection, which rivals the holdings of many university special collections, contains millions of historical documents, going back to 1769 and the Spanish colonial era.

The records survived Hurricane Katrina unscathed, but are still at risk for damage and loss, said Irene Wainwright, an archivist at the library.

"I can't tell you how many people have suggested to us, 'Oh, you just need to " Ms..digitize all that stuff down in the basement and you'll be all right,' Wainwright said. "They have no idea how much effort that requires."

Enter the Genealogical Society of Utah, an organization financed by the Mormon Church, for whom the search for ancestors is a core mission. The society has embarked on a three-year, \$200,000 project to digitize all of the library's genealogically relevant records from 1805 to 1880 (www.familysearch.org).

"The records we gather document the lives of people," said Wayne J. Metcalfe, vice president of the society. "Births, christenings, land records and other documents that provide information about individuals who have lived on the earth."

To that end, genealogy experts affiliated with the Church of Jesus Christ of Latter-day Saints are fanning out, digital cameras in hand,

making copies of genealogically relevant records in 200 cities around the world, including New Orleans. Over the next five years, the church expects to have hundreds of millions of digital images available.

Mr. Metcalfe said economies of scale helped his organization bring down the cost of capturing each image to roughly 20 cents — far less than what a commercial company might charge.

Similarly, I.B.M.'s digitization efforts — dating to the mid-1990s, when the company converted a healthy chunk of the Vatican Library's archives — are done in a way to benefit the company as well as the institution looking to digitize its holdings.

"We look for projects that will highlight I.B.M.'s most innovative technologies or help us develop those technologies with very specific partners who have a problem to solve," said Paula Baker, vice president for global community initiatives at I.B.M. The company looks for projects that require the newest technology.

Such is the case with its most recent multiyear, multimillion-dollar project: a virtual version of the vast Forbidden City in Beijing, which I.B.M. is building in partnership with China's Ministry of Culture. When it is finished, early next year, the site will include interactive, three-dimensional images of ancient thrones, artwork and military implements.

Ms. Baker added that each time I.B.M. embarks on a new venture, requests start coming in from other institutions in need. "When we do these projects everyone else comes out of the woodwork," she said. "But we have to be very selective."

Donald J. Waters, program officer for scholarly communication at the Mellon Foundation, said his foundation had also become increasingly selective over the years.

By way of example, Dr. Waters pointed to the papers of Matthew Parker, the archbishop of Canterbury in the 16th century who collected ancient manuscripts to prove the early existence of an independent English-speaking church that was responsible not to the pope but to the king of England. For centuries, those papers have been locked up at Corpus Christi College at Cambridge University. Mellon is financing a project to put them online.

"It takes a special skill to select stand-alone collections that have a durable appeal in the marketplace of scholars, which is the marketplace that Mellon cares most about," Dr. Waters said. "As interesting and as

important as standout collections in individual libraries and archives might be, the mere fact of digitizing them does not mean that once they are online they will attract and sustain an audience."

The Parker collection, Dr. Waters said, meets all these criteria — it is a core collection for a variety of fields: linguistics, ecclesiastical and religious history, English history, art history, medieval studies. He added, however, that the materials have a long history of restricted access, largely to protect the materials because they are so important.

"Digitization would allow much broader access to the contents," he said, "which is sufficient for much research, without exposing the physical manuscripts to added handling."

WHILE copyright is not a concern for those digitizing documents that are hundreds of years old, copyright restrictions play a significant role when it comes to modern material. Even if the Steinbeck Center in Salinas were to find the money to digitize, say, the manuscript of "The Pearl," its copyright would limit its distribution.

"At this point, online materials are best for authors no longer under copyright," said Susan Shillinglaw, a professor of English at San Jose State University and scholar in residence at the Steinbeck Center.

When Leonard Bernstein's family donated the composer's papers to the Library of Congress in 1993, it was with the goal of digitizing portions of the collection and making them broadly accessible. Although more than a thousand items from the collection have been digitized and placed on the library's Web site, there is still an enormous quantity of material that, because of sheer volume and copyright concerns, is still accessible only to researchers who travel to the library.

For instance, the collection includes a seven-page letter that Jacqueline Kennedy wrote by hand to Bernstein at 4 a.m. on June 8, 1968, the day after the funeral for Robert F. Kennedy, thanking him for conducting a Gustav Mahler work. The letter is an extraordinary window into her grief: "Your music was everything in my heart, of peace and pain and such drowning beauty," she wrote. But the library would need permission from the estate of Mrs. Onassis to digitize it.

When it comes to sound recordings, copyright law can introduce additional complications. Recordings made before 1972 are protected under state rather than federal laws, and under a provision of the 1976 Copyright Act, may be entitled to protection under state law until 2067.

Also, an additional copyright restriction often applies to the underlying musical composition.

A study published in 2005 by the Library of Congress and the Council on Library and Information Resources found that some 84 percent of historical sound recordings spanning jazz, blues, gospel, country and classical music in the United States, and made from 1890 to 1964, have become virtually inaccessible.

"Copyright is a very blunt instrument," said Tim Brooks, the author of "Lost Sounds: Blacks and the Birth of the Recording Industry, 1890 to 1919" (University of Illinois, 2004). "Once you have copyright, you have total control; there's very little room in the copyright law even for preservation, much less reissuing material."

Generally, rights owners like Sony BMG have reissued on CD only a small portion of the recordings they control.

For example, John Philip Sousa's own band made scores of recordings for Victor Records in the early 20th century. BMG bought Victor in 1986, and few if any of those recordings have since been reissued on CD. "There is probably an odd track out somewhere," Mr. Brooks said, "but they've certainly never done any kind of retrospective of him that I'm aware of." And of the hundreds of recordings made in the same period by Noble Sissle, an African-American tenor who recorded for several labels now owned by Sony BMG, few if any have made it onto CD.

.

THE result, Mr. Brooks said, is a series of gaps in the popular understanding of the nation's musical heritage. "It's as if before Bessie Smith, there was nothing," he said. "It has the effect of narrowing our own understanding of our own history."

Another factor that determines what is digitized is how straightforward it is to copy the material.

In some cases, said Theresa Salazar, curator of Western Americana at the Bancroft Library at the University of California, Berkeley, the two go hand in hand. "Agencies and organizations providing funding often want large volume for their money," Ms. Salazar said.

For example, she pointed out, objects like books can be handled in a straightforward way. It is easy to capture these materials because they are printed, and many of these titles are more or less the same size.

No one knows this better than Google, whose digitization efforts focus mainly on books.

In its quest to scan every one of the tens of millions of books ever published, Google has already digitized one million volumes. Google refuses to say how much it has spent on the venture so far, but outside experts estimate the figure at at least \$5 million. The company has also been scanning and indexing academic journals to make them searchable, and is working with the Patent Office to digitize thousands of patents dating back to 1790.

David Eun, Google's vice president for content partnerships, said that rather than dwell on what is being left behind, he preferred to take a more optimistic view.

"We're talking about a huge, huge universe of content," Mr. Eun said. "If you look at the glass as half-empty it becomes too overwhelming."

Correction: March 18, 2007

An article last Sunday about the digitization of historical documents referred incorrectly to a 1968 letter from Jacqueline Kennedy to Leonard Bernstein, part of a Bernstein collection that was donated to the Library of Congress. The letter, which thanked Bernstein for conducting a Gustav Mahler work at the funeral of Robert F. Kennedy, was referring to the Adagietto movement from Mahler's Fifth Symphony; Mrs. Kennedy did not thank him in the letter for conducting "Mahler's Requiem." (Mahler never composed a requiem.)

La memoria del periodismo

Alicia Tapia

Departamento de Periodismo
Universidad Francisco de Vitoria, Madrid
atapia@telemadrid.es

Nereida López

Facultad de Comunicación y Humanidades
Universidad Europea, Madrid
nereida.lopezvidales@uem.es

Elena Medina

Pedro Gómez

Departamento de Comunicación Audiovisual
Universidad Francisco de Vitoria, Madrid
esgaya@esgayafilms.es
pegoma@inicia.es

Resumen

La implantación de las nuevas tecnologías en las televisiones ha supuesto una mayor utilización de la información documental audiovisual gracias a un notable incremento en la accesibilidad de los contenidos conservados.

Pero la implantación de los vídeoservidores y la digitalización del archivo han propiciado paralelamente un incremento del trabajo y la responsabilidad del profesional o documentalista. También está cambiando la lógica documental como consecuencia de la aparición de los «contenidos documentales» y los «ficheros» que desplazan a los antiguos documentos y registros.

Los cambios son importantes, aunque parece que esta revolución tecnológica no acaba de solucionar el problema acerca del legado audiovisual que heredarán los tiempos venideros...

Palabras clave: documentación informativa, documentación audiovisual, archivos digitales, documentalista, memoria histórica.

Abstract. *«Remastering» part of our history... Journalistic Reporting*

The implantation of new technologies in television has supposed the greater use of audiovisual documentary information thanks to a remarkable increase in the accessibility of archived content. But the implantation of video servers and the digitalization of archives have caused, at the same time, an increase in the work and responsibility of professionals or documentalists. Documentary logic is also changing as a result of the appearance of

«documentary contents» and «files» that move old documents and registers. These changes are important, although it seems that this technological revolution does not ultimately solve the problem of the audio-visual legacy that will be inherited in future times...

Key word: News media documentation, Audiovisual documentation, Digital archives, Historical memory, Documentalist, Information manager.

Sumario

- | | |
|---|---|
| 1. Introducción | 5. El documentalista en la televisión digital |
| 2. La digitalización del proceso de producción | 6. El derecho a la memoria audiovisual |
| 3. Neorutinas: para desplazarse... a ningún sitio | 7. A modo de conclusión... |
| 4. El impacto de la digitalización: la catalogación no lineal | 8. Bibliografía y hemerografía |

1. Introducción

Cualquier periodista conoce la importancia de la documentación como parte indisociable del proceso de investigación que conduce a una adecuada exposición de la información ante la audiencia. La documentación sugiere un amplio concepto de búsqueda, selección, orden, comprobación, relación y exposición de multitud de datos referentes a una misma información. Actualmente todos estos procesos que componen la documentación audiovisual se encuentran en una fase de transición.

La tecnología digital permitirá resolver, mejor que la analógica, algunos de los problemas de los archivos de televisión (preservación, documentación y acceso). Por ello, los servicios de documentación de las emisoras de televisión están adaptando sus sistemas de gestión documental al nuevo entorno, pero todavía no existe una metodología suficientemente establecida ni una perspectiva válida para evaluar su impacto real en la gestión de los archivos de las televisiones¹.

2. La digitalización del proceso de producción

Hace siete años Telecinco emprendía un proceso de digitalización que acaba de culminar con éxito. La redacción digital de los servicios informativos fue la primera de Europa y abrió las puertas a un proceso de cambio imparable que, sobre todo en la televisión, ha ido a un ritmo trepidante en los últimos tiempos. En el año 2000 nació el archivo digital (Virage/Informix/StorageTek/Sun) para agilizar el almacenamiento de los contenidos. Esta primera digitalización impor-

1. RODRÍGUEZ BRAVO, B. (2004). «El documento audiovisual en las emisoras de televisión, conservación y tratamiento». *Biblios*, núm. 20, año 5. Edición Grupo Nexo, Lima (Perú). P. 29-39. En: <<http://www.bibliosperu.com/index.shtml>>.

tante permitió introducir unos 250.000 documentos dentro de un robot, en cintas de datos. Desde 2001 se fueron incorporando otras redacciones de programas no informativos, las salas de edición y postproducción y, en general, todos los estudios hasta que el pasado mes de septiembre la cadena privada anuncia la finalización del proceso².

Telemadrid comenzaba en 1999 la digitalización de su redacción de informativos de la mano de Sony. Desde entonces, todo el material visual se guarda dos veces, en alta y en baja resolución, mediante formato MPEG-2. Los periodistas reciben por red local en *streaming* todo el material en sus propios terminales (en baja resolución) y desde allí pueden trabajar con él y montar sus propias piezas, generando un EDL por pieza (lista de decisión de edición)³.

Un servidor de vídeo permite, en primer lugar, almacenar. Debe disponer de una gran capacidad para ello pues el vídeo profesional ocupa mucho espacio. El almacenamiento de datos se realiza sobre discos duros que permiten el acceso múltiple, instantáneo y aleatorio a los datos digitales que contiene. Se puede apoyar igualmente sobre un depósito o *Juke-box* de CD-ROM, o de cintas magnéticas de vídeo, o de datos de gran capacidad, gestionadas por uno o varios robots⁴.

En marzo de 2005 Telemadrid llevó a cabo una ampliación del servidor de vídeo, tanto de baja resolución como de alta, triplicando la capacidad del mismo. Actualmente el servidor de materiales (Daily Server) cuenta con 1.090 horas. En el *daily server*, al que se accede a través de más de 300 terminales distribuidos por todo el edificio de la cadena autonómica, se encuentran los VTR, las emisiones que llegan por satélite —como los servicios de noticias de las agencias (Reuters, APTN, enviados especiales)—, las retransmisiones por cable del Congreso de los Diputados, los partidos de fútbol, las transmisiones por microondas de las unidades móviles, etc. Todos los documentalistas, tanto de informativos como de programas no informativos, y de Deportes, trabajan con el material digitalizado a través de un programa llamado *Clipedit* que permite una selección de los materiales al tiempo que una edición sencilla de vídeo.

Con respecto a la digitalización del archivo, existe un proyecto en curso para automatizar la videoteca trasladando a ésta las imágenes con un sistema robotizado en formato, no ya de vídeo, sino de datos. Hoy en día, se utiliza una base de datos desarrollada por Sony en 1999 con un motor Oracle. La BZA 7000 es una base de datos multimedia asociada a materiales de vídeo grabados en cinta y con posibilidad de transmisión de «metadatos» al servidor de vídeo y viceversa.

2. Producción profesional, núm. 65 (2005). «Telecinco concluye su digitalización». Editorial Bolina, Madrid. P. 4-5.
3. TOLEDO PÉREZ, V. (2004). «Telemadrid apuesta por un futuro digital». *PC Actual*. Computación Actual a. c., Valencia, 10 de marzo.
4. VALLE GASTAMINZA, F. DEL (2004). «Digitalización del patrimonio audiovisual y análisis documental: la utilización de los metadatos en el tratamiento de los documentos de televisión». En: NUÑO MORAL, María Victoria (coord.). *Sistemas de tratamiento y gestión de información*. Cáceres: Universidad de Extremadura. P. 43-55.

Las televisiones que han dado el paso hacia la redacción digital han transformado sus instalaciones: de la disposición de múltiples cabinas para montar las piezas, a una situación en la que cada terminal de ordenador permite al periodista editar la pieza y prepararla para su emisión inmediata sin moverse del asiento. Actualmente Tele 5, Antena 3, Sogecable, Telemadrid, TV3 y las televisiones autonómicas de Castilla La Mancha y Murcia han digitalizado sus redacciones. También se encuentran en marcha Canal Sur y la televisión de Aragón.

3. Neorutinas: para desplazarse... a ningún sitio

Con los archivos digitalizados a través de librerías robotizadas, el redactor, el documentalista, el realizador y el productor pueden volcar al servidor de vídeo, mediante *retrieve*, el material encontrado en la base de datos. El *retrieve* se puede realizar a través del minutado, conocido en algunos ámbitos como *shots*, o a través de los fotogramas destacados en el registro, también llamado *storyboard*, basado en cambios de plano y contenidos. El redactor, una vez realizada la consulta, enlazará el pasado con el presente para elaborar la noticia o reportaje. En este momento registra esa consulta (EDL) en el sistema de alta resolución desde donde se emitirá directamente. En muchos casos, el periodista incluirá la pieza en la escaleta del programa informativo desde su propio terminal añadiendo así su trabajo en la lista de piezas en el orden cronológico con el que se van a emitir.

El material conservado en cintas de datos podrá ser visionado en baja resolución antes de efectuar su volcado definitivo al servidor de vídeo. De esta forma, se puede elegir el *time code* de las imágenes requeridas. El material que por antigüedad o capacidad de espacio de la librería digital se encuentre en cintas de vídeo se volcará al servidor pero sin visionado previo.

A pesar de estos avances tecnológicos, se corre el riesgo de que con los servidores de vídeo y la posibilidad de proteger materiales durante bastante tiempo, se utilice a menudo una misma información del archivo en lugar de buscar variantes de imagen y/o sonido. No es lo mismo elaborar una secuencia de imágenes históricas cuando se puede elegir de entre las existentes en una filmación de veinte minutos que cuando uno se encuentra con un fragmento de seis previamente «volcado en el servidor».

El rasgo común de la mayoría de cadenas de televisión es que se hallan en un universo híbrido y deben navegar entre lo analógico y lo digital. Los archivos no pueden conservar exclusivamente las nuevas producciones efectuadas en el entorno digital, sino que deberán seguir haciendo accesible para sus usuarios todo el fondo documental existente en registro analógico, formato frágil, y que sólo puede leerse con equipos ya desfasados y casi inexistentes⁵.

En la mayoría de las televisiones se encuentran los formatos de conservación Betacam SP, Betacam digital, el sistema digital SX (imagen comprimida) y cin-

5. HIDALGO, P. (2005). «La documentación audiovisual en las televisiones. La problemática actual y el reto de la digitalización». *Documentación de las ciencias de la información*, Vol. 28. Universidad Complutense, Madrid. P. 159-171.

tas de datos. Y aunque actualmente TVE y Telecinco se encuentran digitalizando el material viejo, tendrán que pasar veinte años para disponer de todos los materiales digitalizados. No es necesario digitalizar todo el archivo; se debe realizar una selección del mismo. Aún así, siempre quedarán en las videotecas cintas analógicas, al igual que en todos los medios escritos hay alguna fotografía en papel y en blanco y negro.

Actualmente todos los sistemas de automatización y digitalización de las televisiones integran módulos de ingesta de materiales (satélite, grabaciones programadas, VTR, etc.), gestión de contenidos (archivo, búsqueda y administración de contenidos, compactados a cinta, control de librerías robotizadas), producción gestionada de noticias (escaleta con edición de textos, *play out* multiformato, tituladora y *teleprompter* sincronizados), automatización de continuidad (edición en escaleta, creación de escaletas en remoto, gestión de publicidad, generador de logos, etc.), copia legal (grabación legal de emisión hasta seis meses 24x7, completamente automática) y almacenamiento en *petasite* (librería de cintas de datos, con una capacidad de 20.000 horas en adelante)⁶.

Debemos destacar a Activa 3 y su sistema de producción y archivo digital, DigitonSuite, inicialmente creado y desarrollado para dar respuesta a las necesidades de Televisió de Catalunya y Catalunya Ràdio. La versatilidad del sistema ha permitido que se hayan adaptado versiones para otras televisiones y centros de archivo, como la nueva televisión autonómica de las Islas Baleares, IB3. La configuración de Digiton contará con una completa gestión de *assets* que permitirá la catalogación y búsqueda de los contenidos, además del archivo en calidad broadcast mediante una librería de datos robotizada de gran capacidad⁷.

El hecho de que la televisión autonómica balear funcione con el mismo sistema de producción y almacenaje que Televisió de Catalunya abre la posibilidad de que las dos televisiones puedan compartir sus archivos en el futuro. Al tratarse de un sistema escalable y totalmente adaptable, está siendo utilizado para la producción y el archivo de TV3, así como para la digitalización y la gestión de archivos audiovisuales de gran volumen provenientes del Parlamento de Cataluña, la Filmoteca o el Fórum Universal de las Culturas, Barcelona 2004.

Siguiendo la línea tecnológica documental española, cabe mencionar Vsnnews, otro de los sistemas digitales de informativos con más implantación en el mercado. RTVA (TV Andorra), TV de Vic, Universidad de Burgos, televisiones locales como Tv Eivissa i Formentera, TV Sanlúcar de Andalucía e Info-TV en Valencia han optado por la implantación de este sistema. RTVA (Andorra) ha ampliado el servidor de archivo hasta 10.000 horas en WM9⁸.

6. El Departamento de Documentación de Programas no Informativos de Sogecable comenzó con una librería digital de 20.000 horas, que fue ampliada posteriormente a 40.000.

7. POZO, N. DEL (2005). «Los servicios informativos del presente y del futuro». *Producción Profesional*, núm. 63. Madrid: Editorial Bolina. P. 38-44.

8. POZO, N. DEL (2005). «Los servicios informativos del presente y del futuro». *Producción Profesional*, núm. 63. Madrid: Editorial Bolina. P. 38-44.

4. El impacto de la digitalización: la catalogación no lineal

Los problemas relativos a los centros de documentación de los archivos televisivos siempre han estado orientados a solucionar los problemas de espacio, de conservación, búsqueda, clasificación y recuperación rápida. Se necesitaba un elemento básico de registro de los documentos audiovisuales y una tecnología integradora de todo el proceso documental.

Son tres los beneficios de la digitalización de los archivos: una mayor capacidad de almacenamiento en cintas de datos, el acceso rápido para un mayor número de personas simultáneamente a los ficheros y la garantía de conservación del documento original que ya no corre el riesgo de ser modificado o borrado por algún usuario. El problema de estos sistemas con arquitecturas centralizadas son los llamados «cuellos de botella». A pesar de la operatividad demostrada, la inyección de los materiales al vídeo servidor en las horas de mayor petición informativa provoca que se saturen o colapsen los puertos de entrada e inyección de los materiales.

Inmersos en la era óptica donde términos como archivado, búsqueda y recuperación informativa están siendo desbancados por otros como *storage*, *records management*, *information research* o *information retrieval*, y donde se ha superado el modelo de base de datos basado en un concepto estático y bibliográfico, resulta, sin embargo, evidente que todavía no se han desarrollado soportes y modelos nuevos de gestión⁹.

El problema de las bases de datos audiovisuales, analógicas y digitales, es gestionar y analizar toda la cantidad de información que contienen. Las bases de datos correspondientes a la era digital oscurecen los registros tradicionales. Antes, cada registro representaba un tema y éste a su vez podía estar formado por distintas secuencias y planos. Con las bases de datos digitales, como el sistema de la empresa alemana Blue Order, desaparece el concepto de registro; la esencia es la imagen que deja de ser textual para ser enteramente visual. A la hora de recuperar la información requerida a través de la búsqueda en el minutado o en los shots, las bases de datos digitales ofrecen en pantalla la información audiovisual: *story board*¹⁰ o *keyframes* de imágenes si se refiere al registro o minutado completo del documento, y *strata* si la relación es con una secuencia o fragmento de la misma. El usuario deja de leer textualmente las imágenes y en su lugar se abre un abanico de fotogramas que representan el documento audiovisual. También puede apreciarlas en directo desde su terminal.

Estos nuevos sistemas cambian drásticamente el concepto documental, la esencia de la documentación audiovisual tradicional, antes documentación

9. SAGREDO, F. (2004). «Desarrollo y gestión de soportes avanzados de información documental». En: NUÑO MORAL, M^a V. *Sistemas de tratamiento y gestión de información*. Cáceres: Universidad de Extremadura. P. 87-92.

10. Es aquel dibujo o gráfico que se emplea para representar una idea o vender un producto. Se trata de una especie de guión gráfico que se usa en publicidad, como paso previo para la creación de un anuncio. En: LÓPEZ YEPES, J. (editor) (2004). *Diccionario enciclopédico de ciencias de la documentación*. Madrid: Síntesis. P. 470.

textual para televisión. A la hora de seleccionar la información, de analizar las imágenes, desaparece el concepto global de un tema como registro para llegar a la idea de contenidos audiovisuales casi independientes. Programas como el Virage/Informix, utilizado en Telecinco, no generan un solo registro por documento, sino que aprovechan la segmentación y jerarquización propia del documento para obtener un registro matriz del que dependen otros «miniregistros» con sus correspondientes «metadatos». Debe tenerse en cuenta que el TCR actúa como señal física que permite ir automáticamente al punto del documento que nos interesa: «El resultado es que los niveles de descripción vienen determinados por el contenido; la unidad clásica se disgrega y sólo es perceptible si es necesaria. La recuperación es mucho más precisa, el índice de relevancia aumenta y el acceso físico a la documentación solicitada es mucho más rápido»¹¹.

Algo más complejo es el ejercicio del documentalista: para obtener el máximo rendimiento de estos sistemas es preciso una descripción de las imágenes minuciosa. En el momento actual no existen registros globales sino *contenidos documentales* a los que hay que titular, establecer descriptores, describir los planos, etc., catalogación que con las bases tradicionales se efectuaba en un solo registro o documento; hoy, un registro se ha podido convertir en cinco contenidos o documentos.

Conviene reflexionar detenidamente sobre la idea de *strata* en estos sistemas documentales. El concepto es amplio y nos aleja de las bases de datos textuales o audiovisuales referenciales y de la lógica documental tradicional. De un mismo material se pueden establecer diferentes niveles de descripción y, por tanto, diferentes contenidos. Un contenido es la declaración de un político, y otro lo constituyen las reacciones gestuales del resto de los políticos durante la intervención del mismo. De esta forma, se realizan dos niveles de catalogación, convirtiendo a cada nivel en contenidos diferentes de un mismo material. Se trata de una *catalogación no lineal*. El documentalista, a la hora de analizar un documento, podrá determinar los niveles de catalogación que desee realizar dándole a cada uno la categoría de «contenido» para atender a su recuperación futura.

Los sistemas tradicionales no permitían diferentes niveles de catalogación. La descripción de las imágenes debía realizarse siguiendo el código de tiempo de la cinta. Con los nuevos sistemas el documentalista, dentro de un código de tiempo, puede establecer contenidos con diferentes marcas e incluso puede insertar pequeñas secuencias y eliminar otras no pertinentes dentro del contenido analizado.

Los nuevos sistemas permiten transcribir las declaraciones, es decir, cuentan con un sistema de reconocimiento del habla que transforma la señal de audio en texto. Esto facilita el trabajo del documentalista, que deja de transcribir

11. VALLE GASTAMINZA, F. (2004). «Digitalización del patrimonio audiovisual y análisis documental: la utilización de los metadatos en el tratamiento de los documentos de televisión». En: NUÑO MORAL, María Victoria (coord.). *Sistemas de tratamiento y gestión de información*. Cáceres: Universidad de Extremadura. P. 43-55.

las declaraciones (este ejercicio se efectúa de forma automática). Sin embargo, habrá que esperar a que su completo desarrollo permita la utilización efectiva y masiva de estos sistemas; por el momento, contamos con las posibilidades tecnológicas.

Es preciso tener en cuenta que el principal problema de acceso a las imágenes es de lenguaje. Los sistemas tradicionales se basan en la producción de textos más o menos controlados por parte de un documentalista que, a partir de su propia experiencia, conocimientos, memoria visual, intuición, etc., extrae del contenido visual lo que considera que representa la imagen¹². Los nuevos sistemas digitales establecen lenguajes documentales controlados a través de descriptores asociados a contenidos o capas de catalogación, o asignados a documentos completos o registros, como se indizaba tradicionalmente.

El sistema de la compañía norteamericana Harris se basa en el Paradigma de Laswell¹³ para establecer cinco bloques de descriptores y asimila los campos de metadatos designados por el Dublín Core¹⁴.

Los tesauros permiten la consulta y navegación por los descriptores hasta llegar a la información audiovisual requerida. Tanto los tesauros como los campos de metadatos¹⁵, organizados homogéneamente, son necesarios para la aparición de estos archivos en Internet. «Se considera conveniente la dedicación interna y externa de recursos humanos, económicos y materiales para el mantenimiento periódico del tesoro, lo que incluirá un sistema controlado de incorporación de descriptores, de revisión de relaciones, etc., absolutamente necesario en un campo tan imprevisible como la información de actualidad.¹⁶»

En contradicción con lo que establece la tecnología, el trabajo del documentalista se ha visto ampliado. A la facilidad que aportan los sistemas mencionados hay que restarle la cantidad de trabajo y tareas acumuladas que llevan consigo. La administración de un tesoro supone contemplar un grupo de trabajo asignado al *keyword manager* o sistema de administración de descriptores. La posibilidad de catalogación no lineal y, por tanto, indización y análisis de varios contenidos por material ingestado amplía las tareas documentales y el trabajo de los documentalistas. Como señala Leonor García, es importante subrayar que tras la revolución tecnológica se tiende a guardar más cantidad de imágenes debido a la facilidad con la que se introducen en archivo los fiche-

12. CHAIN NAVARRO, C. (2004). *Técnicas y métodos de recuperación de información*. Murcia: Diego Marín. P. 128.
13. CALDERA, J.; ZAPICO, F. (2004). «La fórmula de comunicación de Laswell como método para implementar bases de datos documentales en los medios audiovisuales». *Investigación bibliotecológica*, Vol. 18, núm. 37. México: Centro Universitario de Investigaciones Bibliotecológicas de la Universidad Nacional Autónoma. P. 110-131.
14. Dublín Core metadata initiative. En: <<http://purl.org/dc/>>.
15. JONG, A. de (2003). *Los metadatos en el entorno de la producción audiovisual. Una introducción*. Traducido al castellano por Jesús Andérez. Federación Internacional de Archivos de Televisión. Hilversum (Holanda): Netherlands Institute for Sound and Vision.
16. VALLE GASTAMINZA, F. DEL; GARCÍA JIMÉNEZ, A. (2003). «Tesauro e información audiovisual. Estudio de caso». *Documentación de las ciencias de la información*, Vol. 26. Madrid: Universidad Complutense. P. 165-180.

ros digitales: «ello supone un problema porque la bolsa de “pendientes de documentar” llega un momento en que no se puede asumir.¹⁷»

En resumen, gracias a la tecnología se pueden conservar físicamente más contenidos; otra cosa es su ordenación suficiente para permitir un uso óptimo de todos los materiales conservados. La asignatura pendiente es una selección efectiva de los materiales que asegure mediante unos criterios racionales la conservación cotidiana de los mismos y su uso en condiciones óptimas.

5. El documentalista en la televisión digital

En este protocolo metodológico es fundamental la conjunción de elementos u objetos audiovisuales: fotos, declaraciones, entrevistas, imágenes relacionadas, archivos de todo tipo, vídeos profesionales y domésticos, y un largo etcétera complican a la vez que enriquecen las labores de documentación.

La realidad tecnológica actual revela además otras necesidades, entre ellas, reescribir las tareas del documentalista sin perder de vista que el concepto de trabajo en grupo supone ahora la base sobre la que se desarrollan muchas de las nuevas herramientas documentales.

Las televisiones han transformado sus redacciones en sistemas de gestión y emisión de noticias, que además se encuentran dotadas de un gran y eficiente sistema de archivo y documentación. En el sistema Estructure, un sistema español instaurado actualmente en el canal temático «Telecorazón» (MediaPro), en la Agencia EFE, en el Congreso de los Diputados, en las universidades Carlos III y Rey Juan Carlos de Madrid, y en el Canal Real Madrid TV, la gestión de todo este material se lleva a cabo por parte de los documentalistas y responsables de archivo, que son quienes toman nota del material ingestado o inyectado en el servidor de vídeo y de qué material puede ser archivado y, por lo tanto, borrado del vídeo servidor y trasladado de nuevo a cintas o a la librería digital (con cintas de datos) para su almacenamiento y conservación¹⁸. Ésta es la misma forma de organización del trabajo en la cadena autonómica madrileña. Los documentalistas, distribuidos por secciones y por programas, se encargan de seleccionar el material ingestado, conservar o protegerlo en el servidor, compilar los contenidos seleccionados en cinta de vídeo (Betacam SX) o trasladarlos al *juke-box* o robot de cintas de vídeo y borrar definitivamente el material desechado.

La digitalización de los procesos de producción en las televisiones implica que el documentalista adquiera nuevos conocimientos y desarrolle nuevas aptitudes: «Además de la formación general en documentación, los documenta-

17. Entrevista realizada a Leonor García, jefe de Documentación de Sogecable, durante la realización del proyecto de investigación *Incidencia de las nuevas tecnologías en técnicas y procedimientos de trabajo*. Investigadora principal, Alicia Tapia López; investigadores Nereida López Vidales, Elena Medina de la Viña y Pedro Gómez Martínez. Proyecto ICOM/03. Financiada por la Universidad Francisco de Vitoria, Madrid. 2004.

18. POZO, N. del (2005). «Los servicios informativos del presente y del futuro». *Producción Profesional*, núm. 63. Madrid: Editorial Bolina. P. 38-44.

listas de televisión necesitan unos conocimientos y habilidades específicos, que generalmente adquieren de forma autodidacta y una vez que comienzan a trabajar en este medio de comunicación. Un buen conocimiento de la información de actualidad, de la producción del medio y la empresa en la que trabajan, el correcto manejo de los equipos técnicos necesarios para la lectura de los documentos audiovisuales con los que deben trabajar y la puesta al día sobre las innovaciones tecnológicas, son algunas de las exigencias habituales en los profesionales que desarrollan este trabajo.¹⁹» Las nuevas tecnologías no simplifican el trabajo; con la tecnología, el trabajo se complica y se trabaja cada vez más²⁰.

El acercamiento a la revolución tecnológica ha ampliado las funciones del documentalista y, con ello, la formación necesaria para ejercerla en el nuevo siglo apenas comenzado. «Debe prepararse para trabajar en equipo, ya que han aparecido nuevas formas de organización del trabajo, debe estar convenientemente informado, como cualquier periodista; y necesita familiarizarse con la informática y las nuevas tecnologías. Estos esfuerzos, obligatorios, se compensan con un nivel de trabajo bien distinto, más cómodo y donde asume más responsabilidad.²¹»

6. El derecho a la memoria audiovisual

A medida que los hombres han ido desarrollando conceptos audiovisuales se ha generado una cultura audiovisual que hace evidente que la historia del siglo XX no se puede estudiar seriamente si no se concede la misma importancia a las fuentes audiovisuales que a las escritas²². La historia de la Humanidad está indisolublemente unida a la revolución tecnológica, y la comunicación es sin duda alguna uno de los aspectos que más relación ha tenido con dicha evolución de la ciencia y de la técnica²³.

La imprenta, revolución tecnológica del siglo XV, llevó a cabo una labor de vulgarización de la ciencia y de la cultura, y, a su vez, un desarrollo y evolución de las mismas y de la sociedad. Hace ya 22 años que la UNESCO soli-

19. HIDALGO, P. (2005). «La documentación audiovisual en las televisiones. La problemática actual y el reto de la digitalización». *Documentación de las ciencias de la información*, Vol. 28. Madrid: Universidad Complutense. P. 159-171.
20. Entrevista realizada a Andoni Ortúzar, director general de EITB, durante la realización del proyecto de investigación: *Incidencia de las nuevas tecnologías en técnicas y procedimientos de trabajo*. Investigadora principal, Alicia Tapia López; investigadores Nereida López Vidales, Elena Medina de la Viña y Pedro Gómez Martínez. Proyecto ICOM/03. Financiada por la Universidad Francisco de Vitoria, Madrid. 2004.
21. TAPIA LÓPEZ, A. (2003). «La memoria de la televisión: un viaje al archivo digital». En: LÓPEZ VIDALES, N.; PEÑAFIEL SAIZ, C. (coord.). *Odisea 21. La evolución del sector audiovisual*. Madrid: Editorial Fragua. P. 275-277.
22. RODRÍGUEZ BRAVO, B. (2004). «El documento audiovisual en las emisoras de televisión, conservación y tratamiento». *BIBLIOS*, Año 5, núm. 20. Lima (Perú): Edición Grupo Nexo. P. 29-39.
23. CALDERA SERRANO, J. (2003). «La documentación audiovisual en las empresas televisivas». *BIBLIOS*, Año 4, núm. 15. Lima (Perú): Edición Grupo Nexo. P. 3-18.

citó a los gobiernos del mundo que se preocuparan de cuidar las imágenes de sus países para preservar la identidad como parte de la herencia cultural de las naciones. Asimismo, el 8 de noviembre de 2001 el Consejo de Europa instó a los países miembros para hacer efectivo el depósito legal de todo su material de imágenes en movimiento y de depósito voluntario del material accesorio.

Para Lorenzo Vilches es necesario que junto a la creación de un Consejo de lo Audiovisual español se provea de una ley de depósito legal de los archivos y de un organismo ejecutivo para la conservación, la investigación, la producción de creación, la formación y la consulta del patrimonio: «El desarrollo coordinado de las fuentes audiovisuales y de los programas de investigación constituye un medio esencial para garantizar uno de los derechos más fundamentales, el de la información»²⁴.

Según una información de Pretospace, proyecto de ámbito europeo que tiene como objetivo desarrollar soluciones técnicas y sistemas integrados para una completa preservación digital de todos los tipos de colecciones audiovisuales, al ritmo actual de trabajo de preservación y con el material de audio y vídeo con más de 20 años, comenzará a quedar degradado en un porcentaje del cinco por ciento cada año que pasa. Desde ahora hasta 2045, el 40% del material que existe simplemente va a desaparecer. El Sony Recording Media Restoration Centre, inaugurado en julio de 2004, tiene como objetivo la preservación de archivos, convertirse en el «guardián de la herencia audiovisual». La fábrica está dedicada a la limpieza, restauración y digitalización de un amplio espectro de material audiovisual e integra un laboratorio químico exclusivo que hace investigación en todos los formatos de grabación existentes y con conocimientos esmerados en la formulación de cinta. De acuerdo con los datos presentados por Henri Chite, director de la división de Marketing de Sony, en Europa cerca del 85% de los archivos de vídeo son en formato analógico. El 15% restante está distribuido por los formatos D2, D3, D5, Betacam Digital y CD-ROM. En cuanto a los formatos de película, los tres primeros lugares van hacia el *16 mm* positivo (23%), el *16 mm* reversible —todos los tipos— (21%) y el *35 mm* positivo (15%)²⁵.

Está en manos del Estado y de las universidades intervenir para garantizar y conservar este patrimonio²⁶. Es necesario crear un organismo nacional que conserve estos fondos y una normativa que afecte a todos los organismos pro-

24. VILCHES, L. (2003). «Salvemos el patrimonio audiovisual». *Telos*, núm. 54. Madrid: Fundación Telefónica.

25. NOGUEIRA, S. (2005). «Recuperando la historia. La fábrica de preservación de Sony en Dax». *Producción Profesional*, núm. 63. Madrid: Editorial Bolina. P. 54-60.

26. El artículo 46 de la Constitución española hace referencia a los documentos sonoros y de imagen, y destaca que forman parte del Patrimonio Documental los documentos de cualquier época generados, conservados o reunidos en el ejercicio de su función por cualquier organismo o entidad de carácter público, por las personas jurídicas en cuyo capital participe mayoritariamente el Estado u otras entidades públicas y por las personas privadas, físicas o jurídicas, gestoras de servicios públicos, en lo relacionado con la gestión de dichos servicios.

ductores y emisores. España es el único país de la Unión Europea que carece de un Consejo Audiovisual de ámbito estatal²⁷.

Es importante que las televisiones comiencen a hacer copias a cintas de datos de sus materiales más antiguos y deteriorados. De hecho, TVE está llevando a cabo un proyecto de transformación y digitalización de los fondos audiovisuales. El proyecto completo tiene un desarrollo previsto entre julio de 2003 y junio de 2011.

El 13 de diciembre de 2005 se celebró la comisión de control parlamentario del ente público donde se destacó que el estándar elegido como soporte de los fondos audiovisuales recuperados y transformados fuese a base de cintas digitales de Betacam IMX²⁸. Al mismo tiempo y en paralelo se obtendrán otras copias del circuito en formato de cinta SX, que es la utilizada para préstamos a programas. El proceso final de digitalización comenzará en 2007, momento en el que uno de los archivos históricos más valiosos en poder de los medios tendrá que estar preparado: «el archivo histórico audiovisual de RTVE constituye sin duda uno de los activos más valiosos que posee la cadena pública y que su conservación y adecuada administración debe formar parte de los principales objetivos de sus responsables.²⁹» En esta comisión se aprobó una proposición no de ley, propuesta por el PP, que impide la enajenación o venta parcial o total de los fondos documentales de RTVE y se insta a la Dirección General de RTVE a que siga cumpliendo los plazos del proyecto de digitalización.

Actualmente casi todos los nuevos sistemas digitales de archivo incorporan un sistema de copia jurídica³⁰. El canal temático AXN acaba de adoptar, tanto para sus emisiones de España como en Portugal, el sistema de copia jurídica «Natural Black Box», desarrollado por Sgo. El sistema Black Box es capaz de almacenar 24 horas de emisión durante seis meses, aunque es totalmente configurable de manera que puede grabarse un número de horas indeterminado. Al superar los seis meses, el sistema permite que el material obsoleto se elimine automáticamente dejando siempre espacio para la nueva emisión, de modo que nunca sea necesaria la intervención de un operador para mantener este sistema de copias judiciales. El software está diseñado para que se pueda acceder directamente desde cualquier ordenador y en cualquier momento al

27. ARNANZ, C. M.; GARCÍA CASTILLEJO, Á.; GONZÁLEZ, B. (2004). *¿Queréis un buen consejo? El sector ante el consejo audiovisual*. Madrid: Gabinete de Estudios de la ATV.

28. Palabras de la diputada del PSOE Remedios Martel Gómez.

29. Congreso de los diputados (2005). *Comisión de control parlamentario de RTVE*, VIII Legislatura, núm. 450, 13 de diciembre de 2005. P. 1-20. El diputado del PP Ramón Moreno Bustos mostró de este modo su preocupación por el tratamiento que recibirá este archivo en los próximos meses.

30. En España no existe ninguna ley sobre conservación del material audiovisual por su importancia histórica. Sólo se deben conservar las emisiones durante seis meses y a efectos legales. Es lo que se conoce como copia jurídica. Durante este tiempo el telespectador puede acudir al juzgado para denunciar a una cadena. Aún así todas las televisiones españolas conservan sus emisiones como memoria dinámica de la cadena.

material almacenado indicando el día, hora, minuto y segundo al que se quiere acceder³¹.

Por su parte, Natural Tools presentó en el IBC de Amsterdam de 2005 su nueva tecnología Natural News, Broadcasting Blackbox y los últimos desarrollos en la edición de noticias pensada para integrarse con las más avanzadas soluciones de *asset management*. Broadcasting Black Box es un sistema automático de grabación que permite mantener on line la copia legal de emisión durante los seis meses que establece la ley o incluso por periodos mayores de tiempo³².

7. A modo de conclusión...

Gracias a la tecnología estamos viviendo una época de madurez audiovisual. A los responsables de las televisiones, políticos y académicos nos empieza a preocupar nuestra propia memoria y el legado que dejamos al futuro. Pero esta inquietud tiene una relación directa con los activos, *los assets*, el valor económico de la imagen. No debemos olvidar que cuando nacieron los archivos audiovisuales a principios del siglo XX y de la mano de las primeras producciones cinematográficas, no interesaba su valor histórico ni artístico. Sólo destacaba su valor económico; eran activos de las productoras, imágenes que podían ser reutilizadas o vendidas. Una vez recorrido un siglo seguimos cuestionando o defendiendo la importancia de los archivos como memoria del periodismo. Es cierto que la tecnología nos está conduciendo a una preservación más definitiva de los fondos audiovisuales, pero la conservación tecnológica de nuestros días no se rige por su esencia fundamental: la importancia de nuestra memoria como parte de la historia.

8. Bibliografía y hemerografía

- ARNANZ, C. M.; GARCÍA CASTILLEJO, Á.; GONZÁLEZ, B. (2004). *¿Queréis un buen consejo? El sector ante el consejo audiovisual*. Madrid: Gabinete de Estudios de la ATV.
- CALDERA, J.; ZAPICO, F. (2004). «La fórmula de comunicación de Laswell como método para implementar bases de datos documentales en los medios audiovisuales». *Investigación bibliotecológica*, Vol. 18, núm. 37. México: Centro Universitario de Investigaciones Bibliotecológicas de la Universidad Nacional Autónoma, p. 110-131.
- CALDERA SERRANO, J. (2003). «La documentación audiovisual en las empresas televisivas». *BIBLIOS*, Año 4, núm. 15. Lima (Perú): Edición Grupo Nexa, p. 3-18.
- CASTILLO, A.; POZO, N. DEL (2005). «Amsterdam mira al mundo de los contenidos». *Producción Profesional*, núm. 65. Madrid: Editorial Bolina, p. 34-42.

31. *Producción Profesional*, núm. 56 (2004). «Copia jurídica en AXN con Natural Black Box». Madrid: Editorial Bolina. P. 8.

32. CASTILLO, A.; POZO, N. DEL (2005). «Amsterdam mira al mundo de los contenidos». *Producción Profesional*, núm. 65. Madrid: Editorial Bolina. P. 34-42.

- CHAIN NAVARRO, C. (2004). *Técnicas y métodos de recuperación de información*. Murcia: Diego Marín.
- Congreso de los Diputados (2005). *Comisión de control parlamentario de RTVE*, VIII Legislatura, núm. 450, 13 de diciembre de 2005. P. 1-20.
- HIDALGO, P. (2005). «La documentación audiovisual en las televisiones. La problemática actual y el reto de la digitalización». *Documentación de las ciencias de la información*, Vol. 28. Madrid: Universidad Complutense, p. 159-171.
- Incidencia de las nuevas tecnologías en técnicas y procedimientos de trabajo*. Proyecto de investigación. Investigadora principal, Alicia Tapia López; investigadores Nereida López Vidales, Elena Medina de la Viña y Pedro Gómez Martínez. Proyecto ICOM/03. Financiado por la Universidad Francisco de Vitoria, Madrid. 2004.
- JONG, A. de (2003). *Los metadatos en el entorno de la producción audiovisual. Una introducción*. Traducido al castellano por Jesús Andérez. Federación Internacional de Archivos de Televisión. Hilversum (Holanda): Netherlands Institute for Sound and Vision.
- LÓPEZ YEPES, J. (editor) (2004). *Diccionario enciclopédico de ciencias de la documentación*. Madrid: Síntesis.
- NOGUEIRA, S. (2005). «Recuperando la historia. La fábrica de preservación de Sony en Dax». *Producción Profesional*, núm. 63. Madrid: Editorial Bolina, p. 54-60.
- POZO, N. del (2005). «Los servicios informativos del presente y del futuro». *Producción Profesional*, núm. 63. Madrid: Editorial Bolina, p. 38-44.
- Producción Profesional*, núm. 65. (2005). «Telecinco concluye su digitalización». Madrid: Editorial Bolina, p. 4-5.
- Producción Profesional*, núm. 56 (2004): «Copia jurídica en AXN con Natural Black Box». Madrid: Editorial Bolina, p. 8.
- RODRÍGUEZ BRAVO, B. (2004). «El documento audiovisual en las emisoras de televisión, conservación y tratamiento». *Biblios*, núm. 20, Año 5. Lima (Perú): Edición Grupo Nexa, p. 29-39.
- SAGREDO, F. (2004). «Desarrollo y gestión de soportes avanzados de información documental». En: NUÑO MORAL, M^a V. *Sistemas de tratamiento y gestión de información*. Cáceres: Universidad de Extremadura, p. 87-92.
- TAPIA LÓPEZ, A. (2003): «La memoria de la televisión: un viaje al archivo digital». En: LÓPEZ VIDALES, N.; PEÑAFIEL SAIZ, C. (coord.). *Odisea 21. La evolución del sector audiovisual*. Madrid: Editorial Fragua, p. 275-277.
- TOLEDO PÉREZ, V. (2004). «Telemadrid apuesta por un futuro digital». *PC Actual*. Computación Actual, a. c., Valencia, 10 de marzo.
- VALLE GASTAMINZA, F. del (2004). «Digitalización del patrimonio audiovisual y análisis documental: la utilización de los metadatos en el tratamiento de los documentos de televisión». En: NUÑO MORAL, María Victoria (coord.). *Sistemas de tratamiento y gestión de información*. Cáceres: Universidad de Extremadura, p. 43-55.
- VALLE GASTAMINZA, F. DEL; GARCÍA JIMÉNEZ, A. (2003). «Tesauro e información audiovisual. Estudio de caso». *Documentación de las ciencias de la información*, Vol. 26. Madrid: Universidad Complutense, p. 165-180.
- VILCHES, L. (2003). «Salvemos el patrimonio audiovisual». *Telos*, núm. 54. Madrid: Fundación Telefónica.

Webgrafía

<http://www.bibliosperu.com/index.shtml>

<http://www.campusred.net/telos>

<http://cuib.unam.mx/publicaciones.htm>

<http://purl.org/dc/>

www.unesco.org/webworld/portalarchives

Alicia Tapia López es doctora en Ciencias de la Información por la Universidad Complutense. Desde 1997 es profesora de Documentación Informativa en la Universidad Francisco de Vitoria, de Madrid. Mantiene dos líneas de investigación, de las que se han publicado diferentes artículos: una línea sobre la enseñanza de la documentación informativa y otra relacionada con su profesión, la documentación audiovisual.

Nereida López Vidales es doctora en Ciencias Políticas y en Sociología por la Universidad del País Vasco. Ha participado en investigaciones científicas financiadas por la UPV-EHU, Gobierno Vasco e IES de Burdeos, UFVI de Madrid, de las que se han publicado ya siete libros. Es autora de varios artículos científicos nacionales e internacionales sobre la influencia de las nuevas tecnologías en los medios actuales.

Elena Medina de la Viña es doctora en Geografía e Historia por la Universidad de Oviedo. Desde 1999 se dedica a la docencia universitaria en las áreas de Producción Audiovisual y de Historia del Cine Español en la Universidad Francisco de Vitoria y Universidad Europea de Madrid respectivamente. Su actividad artística se ha centrado en la realización de varios programas de videocreación y de exposiciones fotográficas. Paralelamente desarrolla su actividad profesional en el campo de la producción audiovisual desde su productora Esgaya Films.

Pedro J. Gómez es doctor en Ciencias de la Información y Guionista en diversas series de televisión. Desde 1999 es profesor de Narrativa e Historia General de la Imagen en la Universidad Francisco de Vitoria de Madrid. En julio de 2004 estrena la obra de teatro *La curva de la felicidad*. También ha publicado diversos artículos sobre cine y televisión.

IX

JORNADAS DE GESTIÓN DE LA INFORMACIÓN

Informar y difundir:
servicios documentales y comunicación

ACTAS DE LAS JORNADAS

Madrid, 22 y 23 de noviembre de 2007
Biblioteca Nacional de España



LA REDEFINICIÓN DEL PERFIL Y FUNCIONES DEL DOCUMENTALISTA
EN LAS REDACCIONES DIGITALES DE MEDIOS ESPAÑOLES

REDEFINING THE PROFILE AND FUNCTIONS OF NEWS LIBRARIANS
IN SPANISH DIGITAL NEWSROOMS

García Avilés, José Alberto. Profesor. Responsable del Área de Periodismo. Universidad Miguel Hernández, jose.garciaa@umh.es; **Masip Masip, Pere.** Profesor. Facultat de Comunicació. Universitat Ramon Llull, peremm@blanquerna.url.edu; **Micó Sanz, Josep Lluís.** Profesor. Facultat de Comunicació. Universitat Ramon Llull, joseplluisms@blanquerna.url.edu

Resumen: La introducción de las tecnologías digitales en el periodismo, no sólo ha supuesto la aparición de un nuevo tipo de periodista, el ciberperiodista, sino que también ha contribuido a la modificación de las prácticas profesionales de los periodistas “tradicionales” de prensa, radio y televisión. En particular, la digitalización de las redacciones, la irrupción de los sistemas de edición no lineal y de internet llevan aparejada una transformación de los procesos productivos, la aparición de prácticas profesionales, así como el desarrollo de nuevas habilidades. Las tareas asociadas a la labor del periodista se han visto modificadas desde diversos puntos de vista: surgen nuevas tareas, se reconfiguran algunas ya existentes y el periodista asume funciones tradicionalmente asumidas por otros profesionales. Asimismo, el periodista ha alcanzado una mayor autonomía en los procesos de documentación, con lo que disminuye su dependencia de los archivos y los documentalistas. En la presente comunicación se ofrece una visión de cómo los periodistas han asumido las tareas de búsqueda y recuperación de la información, tradicionalmente atribuidas a los documentalistas. Se analizan qué retos han debido abordar, su percepción de la labor del documentalista y cómo se ha visto alterada su relación con los documentalistas y el archivo. Finalmente, a la luz de los resultados obtenidos, se ofrecen algunas indicaciones sobre la dirección que deberían tomar los nuevos centros de documentación periodística en la era de la convergencia mediática.

Palabras clave: Periodismo, documentación, centros de documentación, medios de comunicación, digitalización, perfiles profesionales

Abstract: The introduction of digital technologies in journalism has not only brought about a new type of journalist, the so called ciberjournalist, but it has also contributed to modify journalists’ professional practices in traditional media (print, radio and television). More specifically, newsroom digitization, the implementation of non lineal editing systems and widespread internet usage have led to a transformation of productive processes, the rise of professional practices as well as the development of new skills. Many tasks associated with the job of the journalist have been modified from different viewpoints: new tasks are introduced, some of the existing ones are redefined and journalists also take on new roles traditionally carried out by other professionals. These trends have allowed journalists to increase autonomy in the process of documentation, so that their dependence on information professionals and news archives has decreased.

This communication provides an overview of how journalists have adopted the roles of information search and retrieval, traditionally carried out by documentalists. It also analyzes the new challenges faced by journalists, their perception of the job of documentalists and changes in their relationship both with documentalists and the documentation centre. Finally, the results are discussed and some proposals about the design of the new documentation centres in the age of media convergence are made.

Keywords: journalism, documentation, news archives, media, digitization, professional profiles

1. La digitalización de las redacciones y los nuevos perfiles profesionales

La introducción de cualquier tecnología contribuye a la transformación de las prácticas periodísticas, al establecimiento de otras (Weispfenning, 1993) y al desarrollo de nuevas habilidades (Marjoribanks, 2000; Bromley, 1997; Cottle, 2000). Así, gracias a la introducción de la tecnología digital y los sistemas de edición no lineal, los periodistas de televisión han debido asumir tareas que hasta el momento eran desempeñadas por trabajadores especializados como operadores de cámara, montadores o técnicos de sonido. Igualmente, de la mano fundamentalmente de internet (Garrison 2000, 2001; Nicholas et al. 2000), aunque no exclusivamente (Garrison, 2000), los periodistas han adoptado mayor protagonismo en los procesos vinculados a la búsqueda y recuperación de la información, llegando a modificar sus rutinas y su relación con los documentalistas (McKercher 1995; Ward, et al. 1988).

Teóricos como Bohlin (2000) han estudiado detalladamente las transformaciones laborales y las oportunidades profesionales derivadas de la Sociedad de la Información en el ámbito de la industria de los medios. Y algunas administraciones, como la Consejería de Ocupación y Desarrollo de la Junta de Andalucía, también han analizado este asunto. De hecho, el interés primordial por la cuestión se remonta en España al mismo momento en que las redacciones empezaron a automatizarse (Díaz Arias, 1990).

La digitalización de los medios todavía ha llamado más poderosamente la atención de los investigadores sobre el estudio de las competencias de sus profesionales. Así, en opinión de Cebrián Herreros, el periodista televisivo ha dejado de ser el profesional procedente exclusivamente del campo audiovisual especializado en la construcción de un relato concreto, puesto que el contexto actual ha desencadenado la integración de todos ellos en nuevos equipos de diseño y producción (Cebrián, 2001). La profundidad de los cambios revelados en las sucesivas investigaciones incluso ha impulsado a algunos autores a bautizar con denominaciones muy variadas al profesional que combina las actividades tradicionales del periodista con nuevas labores, como la documentación, la grabación de audio y vídeo o la edición. Entre otras, encontramos las etiquetas “informador audiovisual” (Díaz Arias, 1990), “ciberperiodista” (Parra Valcarce y Álvarez, 2004), “teleperiodista” o “teleinformador” (Micó, 2003; 2006).

Las transformaciones afectan tanto a las empresas tradicionales como a los medios y soportes de reciente creación. Un informe elaborado por el Grupo de Periodistas Digitales y el Sindicato de Periodistas de Cataluña con la colaboración del Colegio de Periodistas de Cataluña, señala: “Entre las tareas que habitualmente realiza un periodista digital (están) [...] la producción de contenidos originales, la dirección de proyectos, la edición de contenidos elaborados por terceras personas, el retoque de fotografías, la maquetación web, el

mantenimiento de directorios o bases de datos, la moderación de espacios virtuales y el diseño gráfico avanzado [...] Aun cuando la tarea más realizada es la generación de nuevos contenidos –como hacen el resto de periodistas, sean o no digitales– sí que se detecta que el profesional que trabaja en internet asume nuevas labores más técnicas con tal de adaptar los contenidos periodísticos a nuevos soportes (web, multimedia)” (GPD/SPC, 2003).

En numerosas empresas de comunicación se tiende a la convergencia y a la creación de un marco común de trabajo, la redacción multimedia, concebida como una factoría de la información. Esta redacción se concibe como el ámbito en el que se centralizan todos los mensajes, se realizan asignaciones y se canaliza el flujo de información para editar las versiones impresas, audiovisuales y online de unos contenidos cada vez más personalizados, en función de los destinatarios y del soporte de difusión (García Avilés, 2006).

Sin embargo, la polivalencia periodística y la aparición de nuevos perfiles también han supuesto la desaparición o la mutación de algunas “viejas” profesiones. Rintala y Suolanen (2005), a partir de una investigación realizada en Finlandia entre 2001 y 2002, sintetizan estas transformaciones del siguiente modo:

Transferencia de tareas: las funciones que antes llevaba a cabo un profesional determinado, ahora las asume otro. Por ejemplo, así sucede cuando el periodista radiofónico se hace cargo de la edición sonora.

- Fusión de roles: ciertas actividades que en el pasado eran competencia de dos o más profesionales, ahora son ejecutadas por uno solo. Verbigracia: los perfiles del secretario de redacción de televisión y el secretario de redacción en línea acaban conformando una única figura: el secretario de redacción o managing editor en inglés.
- Incremento de labores: los nuevos medios generan nuevas tareas para el periodista, como ocurre con la necesidad de intervenir en todas las fases del proceso productivo (desde la documentación hasta la gestión de los comentarios de los lectores) o con la obligación de elaborar una misma información en diferentes formatos.

Hasta ahora hemos aludido en varias ocasiones al periodista polivalente. Para entender mejor cuál es su alcance y relevancia, en este punto deberíamos aclarar que se trata del profesional de la información capacitado para desarrollar diferentes tareas con los medios tecnológicos que, en mucho casos, anteriormente empleaban otras figuras para producir contenidos destinados a diferentes medios. Ésta sería una posible clasificación de la polivalencia periodística:

- Polivalencia tecnológica. El profesional de la información utiliza instrumentos (software y hardware) que le permiten producir y gestionar contenidos en diferentes soportes. Por ejemplo, el periodista domina al mismo tiempo programas d’escritura, de retoque fotográfico, sistemas de edición no lineal de vídeo, software para la gestión de redes, bases de datos, etc.
- Polivalencia mediática. El profesional de la información diseña y produce contenidos en los siguientes formatos: escrito, audio, gráfico, vídeo e interactivo. P. ej., el periodista, tras cubrir un acontecimiento, prepara un texto (para la prensa y/o el ciberdiario), una intervención radiofónica y edita un vídeo para la televisión.
- Polivalencia temática. El profesional de la información se ocupa de elaborar informaciones para secciones diferentes (sociedad, economía, política, cultura, etc.). Por ejemplo, el mismo periodista cubre acontecimientos muy variados, desde un partido de fútbol hasta la inauguración de una exposición (GRID, 2007).

Las formas que adopta la polivalencia no se excluyen entre sí, más bien todo lo contrario, puesto que tienden a superponerse. Un mismo periodista puede estar capacitado para confeccionar informaciones para diferentes medios y, al mismo tiempo, generar contenidos sobre política, cultura o sociedad (tal y como sucede en la Agència Catalana de Notícies o en el *Diari de Barcelona*). En otros ámbitos, los profesionales dominan medios y lenguajes diversos, pero se especializan en un área temática. Ésta es la situación actual de los periodistas de la CCRTV.

Las modificaciones en los perfiles profesionales empiezan a tener consecuencias contractuales y jurídicas (GRID, 2007). Los convenios laborales pensados para un cierto tipo de modelo productivo, fundado en una determinada división del trabajo, han quedado desfasados, a menos que nos refiramos a grandes corporaciones públicas como TVE o RNE, donde todavía se mantiene la filosofía imperante hace un par de décadas. Hoy por hoy, ya no resulta extraño encontrar organismos –como el *Skills Council for the Audio Visual Industries* del Reino Unido o *Barcelona Activa-Porta 22* en Cataluña– que se dedican a identificar los cambios experimentados en los perfiles profesionales con tal de mejorar la formación y ayudar a optimizar los recursos dentro de la industria de la comunicación.

2. Objetivos y metodología

La digitalización de las redacciones y la irrupción de internet han incidido irremediabilmente en el advenimiento del nuevo periodista polivalente, que ha asumido algunas de las funciones tradicionalmente desarrolladas por los documentalistas. Internet se ha convertido en una herramienta inevitable para los periodistas, el vínculo con la red es cada vez más intenso y estrecho y ha supuesto un cambio radical en la manera en que los periodistas se relacionan con los centros de documentación y las fuentes. Sin embargo, no es menos cierto, que la red no ha cumplido las esperanzas depositadas inicialmente en ella, según las cuales internet cambiaría la naturaleza misma de la profesión y conduciría hacia un mejor periodismo (Pavlik, 2001).

Gracias al proceso de digitalización, la documentación periodística también ha evolucionado notablemente, tanto por la variedad de las fuentes como por la amplitud geográfica de la procedencia de las imágenes (Caldera y Zapico, 2003). Ha adaptado sus servicios al entorno multimedia (Martín Muñoz y López Pavillard, 1999) y ha creado un proceso más versátil y transparente (Caldera y Nuño, 2002). La creciente competitividad entre los medios, especialmente en televisión, ha obligado a las empresas a ofrecer informaciones de mayor calidad, aportando más datos y a contrastar mejor las fuentes (López de Quintana, 1995).

Observamos como la digitalización de los archivos ha transformado las funciones profesionales que tradicionalmente desempeñaban tanto los documentalistas como los redactores. Cabe plantearse, por ejemplo, qué nuevas funciones está desarrollando el documentalista en redacciones diseñadas para que el periodista acceda directamente desde su terminal a las distintas bases de datos del centro de documentación y que le relegan, por tanto, de una de sus tareas fundamentales que era la recuperación de la información. De la misma manera, cabe examinar en qué medida afecta a los redactores la mayor accesibilidad a la información. Y es que, en principio, el trabajo del redactor puede ser un trabajo de mayor calidad. Éste asiste a todo el proceso de producción, y el control que ejerce sobre cada pieza en teoría debe permitir buenos resultados. El periodista ha ganado en autonomía y control

sobre el producto final, pero se ha visto obligado a desarrollar nuevas habilidades relacionadas con el uso de las tecnologías, las técnicas de búsqueda de información, de gestión de la información, etc. mucho más avanzadas que las que poseía. Unas habilidades por las cuales con frecuencia no han recibido la formación preceptiva.

El objetivo genérico de esta comunicación, pues, es analizar cómo los periodistas, a partir de la digitalización y la irrupción de internet, han asumido tareas tradicionalmente atribuidas a los documentalistas, como son la búsqueda y la recuperación de la información. De manera más concreta, se arroja un poco de luz sobre las características del uso de internet como fuente de información, se explica cómo se ha visto alterada la relación de los periodistas con los documentalistas y el archivo, y se cuenta cómo los periodistas perciben la labor de los documentalistas.

Los resultados provienen fundamentalmente del trabajo de campo realizado en diversas redacciones de medios españoles entre 2001 y 2006, con el objetivo de analizar cómo la digitalización ha modificado las rutinas periodísticas y ha definido nuevos perfiles profesionales. En concreto, se han mantenido entrevistas con más de 70 profesionales de distintos medios de comunicación (Antena 3, Tele 5, TVE, Telemadrid, TV3, Castilla la Mancha Televisión, Catalunya Ràdio, *La Vanguardia*, Ser, RNE, EFE...) y de diversos perfiles. Igualmente, se llevó a cabo observación de campo en las redacciones de *La Vanguardia*, Catalunya Ràdio y Televisió de Catalunya, durante los años 2004 y 2005, y en el *Diari de Girona* entre 2000 y 2005.

3. Informatización y centros de documentación informativa

Garrison (2001) sitúa en la década de los 50 los primeros casos de aplicación de los ordenadores en la tarea periodística, en concreto para poder realizar el seguimiento de las campañas presidenciales. Sin embargo, no fue hasta la década de los 70 que los terminales empezaron a proliferar en las redacciones de los periódicos norteamericanos y canadienses. Durante los años 80 la informática invade los medios de comunicación, no sólo las redacciones, sino también sus respectivos centros de documentación. Estos, conscientes de las nuevas oportunidades que los ordenadores les ofrecen, inauguran un proceso lento y paulatino de incorporación y desarrollo de recursos de información electrónica.

Nora Paul distingue tres etapas en los servicios de documentación periodística. Durante la primera etapa, la basada en el papel, la función principal de los documentalistas era la de conservación. La segunda etapa supuso el inicio de la digitalización de los documentos y la creación de bases de datos para su posterior comercialización. En esta fase, debido a los elevados costes asociados a la consulta de los nuevos recursos de información electrónica y a la complejidad de los lenguajes de interrogación, los documentalistas ejercen un marcado rol de intermediarios entre los periodistas y la información. Los centros de documentación desarrollan entonces un papel importante como proveedores de información, fuentes e ideas. El panorama cambiaría radicalmente en la tercera etapa identificada por Paul, en la que los periodistas pueden consultar desde sus propios ordenadores los fondos digitalizados del medio, así como las bases de datos comerciales suscritas. Igualmente, las interfaces son cada vez más intuitivas y mejoran en usabilidad, los lenguajes de interrogación se simplifican, y, desde mediados de los noventa, irrumpe internet en las redacciones. Durante esta fase, progresivamente, los periodistas ganan en autonomía y los documentalistas observan como tareas que les eran delegadas pasar a ser asumidas por los propios redactores. Los resultados

obtenidos por Garrison (2000) muestran como el protagonismo de los periodistas en la búsqueda de información aumenta significativamente, pasando de un 23,5% en 1995 a un 68,8% el 1999.

Ante esta nueva situación, la mencionada Nora Paul estima que los documentalistas deben asumir un nuevo rol. Los periodistas ya no precisan de intermediarios, ahora necesitan consultores, consejeros, formadores. Con respecto a la labor formativa, diversos autores señalan que los centros de documentación deben convertirse en auténticas pasarelas de acceso a la información digital, desarrollando intranets que faciliten el acceso al archivo y recursos de interés para los periodistas, como cronologías, efemérides, selección de webs, etc. (Nicholas, et al., 1998). Las intranets ofrecen beneficios en términos de mejora de acceso a la información. Permiten que los periodistas tengan acceso a un abanico más amplio de recursos, mediante un formato único, de fácil consulta y desde su propia mesa. Además, posibilitan que los periodistas ahorren tiempo, y también los documentalistas, que no deben responder preguntas triviales y de fácil respuesta, como la manera correcta de escribir un apellido o un topónimo (Drennan, 1999). María Rubio Lacoba (2004) destaca que los documentalistas deben desarrollar una función gestora que pueda satisfacer las nuevas necesidades de los usuarios de los centros de documentación. No se trataría de una tarea suministradora, sino que debería anticiparse a las necesidades de los periodistas y acompañarles para satisfacer necesidades informativas más complejas, que el periodista por sí mismo no puede lograr.

En España la situación ha sido diametralmente distinta a la vivida en los Estados Unidos. La disponibilidad de recursos digitales en los centros de documentación españoles a mediados de los noventa era prácticamente nula. En 1994, Fuentes y Conesa subrayaban la ausencia casi total de bases de datos en los principales medios de comunicación del país. Las autoras atribuían el escaso desarrollo de los centros de documentación a la falta de tradición, la escasez de inversiones, la baja especialización profesional, el poco desarrollo de la industria de la información y la poca concentración empresarial.

Transcurridos 10 años, el panorama ha cambiado sustancialmente, aunque no sin dificultades. Así, cuando los periodistas españoles, y los de la mayor parte de Europa, empezaban a sustituir las máquinas de escribir por los primeros terminales, los norteamericanos hacía más de una década que los usaban. Esta familiaridad con los ordenadores y la información electrónica permitió, a nuestro entender, que años más tarde, con la irrupción de internet, ésta se viese como un paso adelante más en la relación entre periodismo y tecnología. Una impresión que difícilmente podrían tener los periodistas españoles, que acababan de tener acceso a las redacciones digitales. Sin prácticamente tradición en el uso de la información electrónica, la introducción de internet como fuente de información entre los periodistas se produciría de forma lenta, gracias a la llegada de otras tecnologías –en Tele 5 las primeras conexiones a internet aparecen ligadas a la renovación del parque informático debido a la implantación de los sistemas de edición no lineal (Micó, 2003)-, y sin políticas activas de formación de los profesionales. La llegada de internet a las redacciones, pues, estuvo más marcada por los frenos que por los impulsos.

4. Periodistas y documentalistas: nuevas funciones, nuevas relaciones

A pesar de que el nivel de los centros de documentación de los medios españoles ha mejorado ostensiblemente en la última década, su situación aún dista mucho de la de los

centros homólogos de otros países, como los Estados Unidos. Con todo, con frecuencia los investigadores españoles se han dejado imbuir por la realidad norteamericana y han adoptado discursos, a nuestro juicio, demasiado triunfalistas con relación a la situación de los centros de los medios españoles y el futuro de la profesión. Unos discursos más próximos a una manifestación de deseos que a un análisis reposado de la realidad.

El trabajo de campo realizado en los últimos años nos permite identificar un panorama con luces y sombras. Mientras que, por un lado, los archivos audiovisuales se han visto reforzados gracias a la digitalización, por el otro, los archivos de documentación impresa, y particularmente los documentalistas que en ellos ejercen sus labores, se hallan en una situación ciertamente paradójica.

Como decimos, la digitalización de los archivos permite una mejor conservación de los fondos, y garantiza la versatilidad y la facilidad de uso de la información. El periodista audiovisual asiste a todo el proceso de producción. El nuevo sistema permite realizar el proceso completo para la edición, emisión y archivo de noticias, incluyendo la ingesta de imágenes, la redacción de noticias, la edición de vídeo y la emisión de los programas. El proceso afecta a los diferentes colectivos del medio, no sólo a los redactores, también a los documentalistas, operadores de vídeo, ayudantes y técnicos, realizadores, productores, etc. Se incrementa la capacidad de generar noticias al disponer de nuevas herramientas, para ganar también en la inmediatez de la emisión y en la mejor documentación de los contenidos (García Avilés y León, 2002).

En televisión, el archivo ya no puede separarse de los restantes procesos que integran el ciclo de producción de noticias (Hidalgo, 2005; Risoto 2004). La decisión acerca del material que pasa a engrosar el archivo de la redacción, los procedimientos de acceso y de uso posterior de ese material, constituyen las cuestiones esenciales en la gestión de contenidos. Como subraya el director técnico de Telecinco, José Luis Romero, los nuevos sistemas ofrecen una enorme versatilidad:

“Para acceder a imágenes del archivo, el periodista ya no depende de nadie, sólo de sí mismo. Desde su propia terminal accede a la información del archivo, a ficheros de baja velocidad; selecciona, marca y formula su petición. La información le llega a su servidor y puede seguir trabajando. Es decir, que lo busca en una base de datos, lo relaciona y lo solicita, lo revisa y lo descarga”.

El periodista antes se quejaba de la falta de tiempo para acudir al archivo. Ahora, puede utilizar el material de archivo más a menudo en sus informaciones. La digitalización de los archivos y los recientes sistemas de integración de vídeo por ordenador permiten el acceso directo a las imágenes disponibles y no sólo a la habitual descripción escrita. El uso inteligente del archivo constituye por sí solo el origen de un buen número de informaciones y en muchos casos, enriquece notablemente las piezas de actualidad, al ligarlas a su contexto. Por ejemplo, al preparar una información sobre la retirada de un futbolista, mediante un solo clic en su ordenador puede acceder y ver todas las imágenes de ese futbolista que están disponibles en el archivo e incorporar las que mejor se ajusten a la historia que desea contar. Ello indudablemente permite, sobre el papel, mejorar su trabajo.

Así pues, la documentación en una redacción multimedia se ha convertido en un auténtico valor añadido que sitúa favorablemente a un medio con respecto a la competencia y a los medios tradicionales (Rubio, 2005). El documentalista puede considerarse un coproductor de informaciones en los medios, por su trabajo de organización, archivo y apoyo de documentos al periodista. Su actividad se concreta en organizar, archivar y apoyar al

periodista en su trabajo. Un subdirector de informativos de Telemadrid destaca la función de los documentalistas como gestores de información:

“Los documentalistas procuran que el material disponible en el servidor central sea lo que interesa, y que eliminen lo que no interesa. El documentalista se convierte en el gestor del material útil, tanto en anticipo como en retraso. Además de sacarle un gran partido al material, permite que el sistema genere satisfacción entre los periodistas”.

Junto a esta visión ciertamente esperanzadora de los centros de documentación audiovisuales - en el caso de Telecinco, incluso el equipo de documentalistas se ha duplicado respecto a la situación anterior, con el sistema analógico,- el aumento de la autonomía de los periodistas, la facilidad de uso de las bases de datos e internet revelan una realidad menos boyante, especialmente en los archivos de los medios escritos, que en algunos casos se ha materializado en verbalizar la posibilidad de externalizar el servicio.

Efectivamente, la llegada de internet y la digitalización de los archivos, tanto audiovisuales como textuales, ha ido aparejada al descenso del número de consultas de los centros de documentación. En el caso de los periodistas que precisan información textual para completar sus noticias, para preparar entrevistas o para cualquier otra actividad de la práctica periodística, la visita al centro de documentación se ha visto substituida por la consulta de la base de datos del propio medio o, muy especialmente, de internet.

Los periodistas acuden a la red fundamentalmente para buscar información de *background*, que no siempre aparece reflejada en el texto de las noticias, y muy especialmente, para localizar determinadas informaciones u obtener datos concretos. Desde este punto de vista, se ha convertido en un recurso de primera magnitud, con frecuencia el recurso exclusivo. En estas circunstancias, la consulta del centro de documentación es minoritaria. Algunos reporteros reconocen que únicamente acuden a los documentalistas y les delegan las búsquedas cuando el tiempo les apremia, ya que la ayuda de los documentalistas les permite trabajar en otros temas. La escasa fiabilidad de numerosas fuentes en internet, obligaría al periodista a extremar las precauciones y a contrastar, aún más, la información. Aunque debido a la premura en el cierre de las ediciones, algunos reconocen no hacerlo.

Las hemerotecas digitales accesibles por internet se han convertido en fuentes de información recurrentes, sustituyendo la consulta de la base de datos del propio medio. Esta circunstancia ha sido observada, por ejemplo, tanto en Catalunya Ràdio como en Televisió de Catalunya, donde, a pesar de que los respectivos centros de documentación elaboran conjuntamente una excelente base de datos, de amplia cobertura temática y temporal, y además fácil de usar, los periodistas prefieren la búsqueda de información en internet, en concreto en hemerotecas digitales, como las de la BBC o *El Mundo*. A la pregunta sobre las causas de este comportamiento, los reporteros respondían con un lacónico “no tengo la costumbre [de consultar la base de datos del centro de documentación]”. La responsable de documentación de Televisió de Catalunya reconoce que la inversión realizada para la digitalización del archivo no ha obtenido los resultados esperados, en cuanto al número de consultas directas de los periodistas, y “[esto] deberá hacer replantear algunas cosas”

El procedimiento más habitual por localizar información en la red es, obviamente, el uso de buscadores. Google es, sin duda, el motor de búsqueda preferido por los periodistas catalanes, 38,41%, seguido a mucha distancia de Yahoo o Altavista (Masip, 2005). La rapidez es la característica que más se valora, en concreto un 31,46% de los periodistas encuestados consideran que esta es la propiedad principal que debe tener un buscador. A

continuación se sitúa la exactitud de los resultados obtenidos. No deja de ser significativo que se valore más un aspecto técnico que la calidad de los resultados obtenidos. Al igual que en estudios efectuados en otras latitudes (Ewart, 2002; Garrison, 2000), el uso de las opciones de busca avanzadas es minoritario, un 40% no las usan nunca o casi nunca, y sólo un 13,15% las usan siempre. Las opciones de búsqueda avanzadas más usadas son los operadores booleanos (41,31%), la búsqueda por frase exacta (31,92%) y la limitación por campos (27,23%)

A menudo, los periodistas justifican este deficiente uso de la red en la medida que, afirman, se necesita demasiado tiempo de aprendizaje para llegar a dominar las posibilidades que ofrece internet y obtener resultados satisfactorios en la búsqueda de información, un tiempo del que los periodistas a menudo no disponen. La falta de formación en búsqueda de información es precisamente uno de los principales inconvenientes que tienen que afrontar los periodistas. La formación nunca ha sido una prioridad para las empresas. Como hemos apuntado anteriormente, la introducción de internet se produjo habitualmente vinculada a actualizaciones tecnológicas. Según datos de la profesora Bella Palomo (2002), tan solo un 31,2% de los periódicos españoles ofrecieron cursos de formación a sus trabajadores. La cifra es ciertamente modesta, aunque se aprecian diferencias significativas entre lo que la autora denomina “prensa central de Barcelona”, que en el año 2001 había organizado algún tipo de actividad formativa en el 100% de los casos, y su homóloga de Madrid, con tan solo un 25% de casos.

Así las cosas, al tradicional menosprecio con el que los grupos de comunicación españoles han tratado los centros de documentación, habría que sumar que, más recientemente, en ciertos círculos hemos detectado la sensación de que estas áreas tienen un carácter supuestamente superfluo, debido a la irrupción de internet, que ofrece gran cantidad de información y de forma gratuita. Un problema que se agudiza en el momento en que algunos periodistas asumen labores del documentalista y consideran que pueden prescindir de sus servicios.

Un ejemplo que ilustra esta circunstancia lo encontramos en *La Vanguardia*. Los periodistas del rotativo del Grupo Godó valoran muy positivamente el centro de documentación, y se muestran orgullosos de su hemeroteca centenaria y de su fondo fotográfico histórico. A pesar de ello, el centro de documentación presenta un problema de percepción, de visibilidad entre los redactores. A diferencia de lo descrito en los medios audiovisuales de la CCRTV, la base de datos de *La Vanguardia* es ampliamente consultada por los periodistas. Aurora, así se denomina la base de datos, incorpora todos los artículos publicados en el periódico desde 1994 y tanto las fotografías de los profesionales del medio como las recibidas por agencia. El acceso a la base de datos es directo desde el programa Hermes, lo que permite el trabajo simultáneo en la redacción del texto y la búsqueda documental, así como la integración automática de las fotografías. Como consecuencia de este acceso directo, las consultas de la base de datos son frecuentes, sin embargo, los periodistas no asocian la información obtenida a través de Aurora a la labor de catalogación previamente realizada por los documentalistas. La mayoría distingue entre la intranet (Aurora) y el centro de documentación, que son percibidas como dos cosas diferentes. Cuando el periodista menciona el centro de documentación, se refieren exclusivamente al espacio físico que éste ocupa y a las consultas directas realizadas a los documentalistas. No consideran que Aurora sea un producto suyo.

Los periodistas afirman que tras la puesta en funcionamiento de la intranet el número de peticiones de información al centro de documentación ha disminuido de forma considerable.

En algunos casos, también se detecta un desconocimiento del perfil profesional de los documentalistas y de la función de apoyo al periodista que pueden desempeñar. Así, un redactor de la sección de Economía afirmaba que usaba poco el centro de documentación porque no ofrecía los servicios que requería habitualmente y echaba en falta “un profesional capaz de realizar tareas de búsqueda documental, propias del periodismo de investigación”. Con estas palabras se pone en evidencia que todavía se mantiene una concepción de esta profesión heredera de la visión tradicional del documentalista como organizador y guardián de la documentación.

En los medios de dimensiones más reducidas la situación es todavía más delicada. Es el caso, por ejemplo, del *Diari de Girona*, el periódico con más difusión de la ciudad, con una media de 8.000 lectores por día. En este periódico del grupo Prensa Ibérica una única persona se encarga de mantener y gestionar el archivo. Los redactores, que, por cierto, también maquetan sus informaciones y disparan las fotografías que necesitan, buscan la documentación desde sus ordenadores, sin la intermediación del documentalista, mayoritariamente en Internet. Sin embargo, ni para llevar a cabo esta tarea ni para desempeñar el resto de labores adicionales que conforman actualmente su perfil profesional han recibido ninguna formación. Únicamente reciben un complemento económico por las fotografías realizadas y publicadas.

La búsqueda y recuperación de información audiovisual en internet también ha incidido en las redacciones televisivas, aunque de forma diferente a lo observado con la información textual. El intercambio de vídeo y audio por la red -aunque aún no tenga calidad *broadcast*- sirve como una valiosa fuente de información que complementa la tarea redaccional. Los periodistas acceden a internet cuando están trabajando sobre *breaking news*, para ampliar información, contactar con expertos o implicados y también para incluir material gráfico de la web en sus piezas. Algunos periodistas señalan que internet en ocasiones plantea problemas para verificar las fuentes y que también han de obtener los correspondientes derechos antes de emitir ciertas imágenes.

Anteriormente, hemos destacado que los periodistas, gracias a la tecnología, han tomado el control directo sobre el proceso de producción. Su responsabilidad sobre el producto final se ha incrementado y ello les ha aportado una mayor autonomía: ya no dependen necesariamente de otros profesionales y servicios, como los documentalistas y los centros de documentación. Esta circunstancia, en teoría, debería contribuir a mejorar la calidad del producto final y, en consecuencia, ofrecer a los lectores y telespectadores unos mejores periódicos e informativos (Dahlgren, 1996; Pavlik, 2001). La realidad, sin embargo, nos muestra un panorama menos alentador. Efectivamente, los periodistas han ganado en autonomía, pero cabe subrayar que ello no equivale necesariamente a una mejora de las informaciones.

Las innovaciones tecnológicas han propiciado que se acumulen las tareas encomendadas a los periodistas, que se han visto obligados a desarrollar o perfeccionar nuevas habilidades. Según la opinión de los propios profesionales, la calidad de los productos finales se ha resentido por esta razón. Además de asistir a una rueda de prensa o cubrir un acontecimiento noticioso, ahora tienen que realizar fotografías -lo hemos visto en el *Diari de Girona*- tomar imágenes de vídeo -por ejemplo en l'Agència Catalana de Notícies- y, al llegar a la redacción, montar las piezas -en la mayoría de las televisiones-, buscar documentación, etc. La combinación de una formación incompleta para nuevas responsabilidades y el mantenimiento de la misma presión -o mayor- en los tiempos de

entrega de las noticias tiene un resultado claro: algunas de las tareas encomendadas se efectúan de forma deficiente. Diversos responsables de televisiones y periódicos reconocen que la calidad de los productos es inferior en el presente. No necesariamente desde una perspectiva periodística, pero sí desde una óptica técnica. Conviene preguntarse, por ejemplo, en qué medida los documentalistas en el nuevo sistema se convierten en parte del engranaje de una maquinaria, en la que su trabajo está limitado, y donde disminuye –o incluso desaparece– su capacidad creativa. Algunos profesionales argumentan que las redacciones digitalizadas refuerzan el individualismo y reducen las oportunidades de intercambio de información entre periodistas y documentalistas.

Sirva como ejemplo para ilustrar esta situación la de un redactor de un importante periódico que para completar una noticia sobre el auge de las compañías aéreas de bajo coste, deseaba realizar un cuadro comparativo de los principales aeropuertos regionales europeos que acogen este tipo de aerolíneas. Tras buscar en Google –y no encontrar– la información necesaria para elaborar el cuadro, y sin plantearse recurrir a los documentalistas del medio, decidió sustituir el cuadro previsto por una fotografía de un avión en el aeropuerto de Girona. A la pregunta sobre el motivo de no acudir a los documentalistas del medio que le ayudasen a obtener la información que precisaba, el redactor sencillamente respondió que “los documentalistas tampoco habrían sabido sabrían encontrarla”

Esta anécdota sirve para mostrar, la incidencia sobre la calidad editorial de una formación deficiente, y también uno de los problemas que los documentalistas padecen todavía en los medios españoles: el desconocimiento de su labor. El problema, sin embargo, también permite identificar algunas de las posibles soluciones.

De acuerdo con Martín y López (1999), la documentación en una redacción multimedia desarrolla dos funciones esenciales: la integración de información de distinta naturaleza en una misma base de datos -plataforma multimedia- y la integración de dichas bases de datos en otras mayores, a escala internacional, para el intercambio de información, lo que obliga a una estandarización en los procesos documentales más allá de las normas particulares de cada centro de documentación. Ambos autores apuntan la necesidad de:

“Redefinir el papel del documentalista en el entorno multimedia, una vez que los usuarios han ganado autonomía en la recuperación y la introducción de información se ha universalizado dentro de la empresa. Crear un plan de formación específico para estos “nuevos documentalistas”, que incida en la polivalencia de tareas dentro de las funciones propias de un documentalista”.

En el mismo sentido se manifiesta Rubio Lacoba (2005), que considera que el nuevo entorno exige una adaptación de las funciones del servicio documental, que debe forjar una nueva asociación informativa con la redacción. De acuerdo con la directora de documentación de Informativos Telecinco, las nuevas tecnologías facilitan el trabajo del documentalista, “ahorrando rutinas mecánicas y esfuerzos inútiles y permitiendo que sus esfuerzos se centren en lo fundamental: la elaboración de documentación específica, cronologías, recopilaciones temáticas, investigación sobre las fuentes documentales más adecuadas para resolver las diferentes tipologías de consultas, etc. Lo verdaderamente importante es el valor añadido, el trabajo intelectual, el tratamiento conceptual que aporta el profesional de la documentación a la información”. Sin embargo, pese a las ventajas que proporciona la tecnología digital, el trabajo del documentalista también se ha visto ampliado, tal y como sugieren Tapia et al. (2006).

5. Conclusiones

A partir de las entrevistas realizadas y las experiencias vividas en las redacciones, consideramos que para que los esfuerzos de los servicios documentales para adaptarse a la nueva realidad den sus frutos, es necesario superar el problema de imagen y visibilidad que todavía arrastran los centros de documentación. Lo hemos constatado en el caso de *La Vanguardia*, donde los periodistas no tienen la percepción de que la base de datos Aurora sea parte del servicio de documentación, que lo reducen a un espacio más o menos identificado a un extremo de la redacción. Pero hay más ejemplos.

Ante las noticias del día y *breaking news* hemos visto como los centros de documentación tienen pocas oportunidades de competir con la agilidad, la inmediatez y la aparente sencillez de uso de internet. Sin embargo, en las noticias de largo recorrido - efemérides, conmemoraciones, reportajes, etc.- los documentalistas pueden resultar mucho más útiles a los periodistas.

El tiempo que los documentalistas dedicaban antes a localizar todo tipo de peticiones formuladas por los periodistas, ahora lo pueden invertir en otro tipo de búsqueda de información más compleja, que lleva mucho más tiempo, como el material para elaborar informaciones en profundidad. En ocasiones, los documentalistas elaboran productos con garantías de emisión o publicación. Se trata de ir preparando aquellas imágenes y textos que pueden ser de utilidad a los periodistas, incluso antes de que el propio periodista los solicite, y a las propias empresas. Entre esas tareas, se puede señalar:

- Elaborar biografías o programas con motivo de algún aniversario o de un acontecimiento que sucederá en breve.
- Reelaborar nuevos productos, como programas de revisión histórica y montajes de archivo.
- Reponer fragmentos por su interés inmediato dentro de una determinada emisión.
- Elaborar dossiers electrónicos sobre acontecimientos previsibles y que sirvan de base para la labor de los redactores encargados de cubrir los acontecimientos cuando se produzcan.
- Comercializar y difundir externamente el fondo a petición de particulares, organismos públicos o productores audiovisuales y publicitarios.

En este sentido, el documentalista debe reivindicarse como potencial creador de contenidos. Existen ya algunas experiencias en esta dirección. *La Vanguardia* permite que los documentalistas dispongan de un espacio donde publicar piezas elaboradas por ellos mismos a partir de la explotación del propio archivo. El personal de documentación de *El Periódico de Catalunya* ha reivindicado intensa y constantemente su derecho a firmar aquellas informaciones que contengan una proporción elevada de material de archivo, como sucede, por ejemplo, con la infografía, cuyos responsables aparecen explícitamente como autores de una parte de la información. Según su punto de vista, el trabajo de búsqueda de datos es equiparable al que efectúan los redactores al confeccionar sus textos. Por lo tanto, también ellos deberían obtener un cierto reconocimiento. En el presente, su petición ha sido atendida, aunque sólo en parte.

La relación entre periodistas y documentalistas es una de las que más se ha visto alterado a raíz de la llegada de internet y la digitalización de las redacciones. La red y los archivos informatizados han permitido que los periodistas puedan abordar sus rutinas de

búsqueda de información sin ningún contacto con los documentalistas y los centros de documentación. Esta nueva realidad ha provocado una pérdida de visibilidad del archivo, que con frecuencia no ha sabido comunicar el rol que juega, o podría jugar, en el complejo engranaje de la producción de la noticia.

En la mayor parte de los medios analizados, tanto escritos como audiovisuales, se detecta un importante descenso del uso de los respectivos centros de documentación, especialmente de los archivos de prensa, cuya consulta se ha visto sustituida por la de internet. Esta pérdida de usuarios se observa también en las televisiones, si bien se centra en los fondos de documentación escrita. Los sistemas redaccionales integrados, por el contrario, ofrecen nuevas oportunidades a los centros de documentación audiovisual.

El centro de documentación en una redacción multimedia comparte rasgos comunes con los recursos documentales tradicionales, aun cuando también posee elementos significativamente distintos que lo convierten en un servicio especial por la peculiaridad del usuario al que atienden, por el número de medios dependientes del servicio, o por su mismo planteamiento físico –con un menor volumen de fuentes bibliográficas tradicionales y un incremento del empleo de Internet como fuente.

El futuro de los centros de documentación informativa pasa por una adaptación a la nueva realidad, que ya se está produciendo, pero también por un “lavado de imagen” que permita que los periodistas vean en los documentalistas unos profesionales aliados, especialistas en el uso de la tecnología y con los que comparten un lenguaje común. Labor difícil si atendemos al tradicional ostracismo en que ha vivido la documentación informativa en las empresas de comunicación y al individualismo que cada vez más se instala en las redacciones.

Referencias bibliográficas

1. Bohlin, E. [et al.] *Information and Communication Technologies and the Information Society*. Sevilla . IPTS, 2000
2. Bromley, M. “The End of Journalism? Changes in Workplace Practices in the Press and Broadcasting in the 1990’s”. En: Bromley, Michael; O’malley, Tom, eds. *A Journalism Reader*. Londres: Routledge, 1997 pp. 330-350.
3. Caldera, J. y Nuño Moral, M.V. “Etapas del tratamiento documental de imagen en movimiento para televisión” En: *Revista General de Información y Documentación*, 2002, v. 12, n. 2, pp. 375-392.
4. Caldera, J. y F. Zapico. *Procedencia de información audiovisual en las empresas televisiva*. En: *Hipertext.net*, n. 1, Consultado en: 26/06/2007.
<http://www.hipertext.net>.
5. Cebrián Herreros, M. “La edición en el sector audiovisual”. En: *ZER*, 2001, n. 11.
6. Cottle, S. “From BBC newsroom to BBC newscenter: on changing technology and journalistic practice”. En: *Convergence: the Journal of Research into New Media* 1999, v. 5, n. 3, pp. 22-43.
7. Dahlgren, P. “Media Logic in Cyberspace: Repositioning Journalism and its Politics”. En: *Javnost: The Public*, 1996, v. 3, n 3, pp. 59-72.

8. Díaz Arias, R. "La redacción automatizada, un sistema de redacción". En: *Mensaje y Medios*, 1999, n. 15, p. 57.
9. Drennan, J.. "The Introduction of Intranets into the Newspapers Industry". En: *ASLIB Proceedings*. 1999, v. 51, n. 8, pp. 269-274.
Ewart, J. *Searching the Net: A Study of Internet search techniques used by some Australian journalists*. 2002. Inédito
10. Fuentes, M. E.; Conesa, A. *La documentació periodística: Catalunya, Espanya i altres experiències europees*. Barcelona: Centre d'Investigació de la Comunicació.
11. Generalitat de Catalunya, 1994
12. García Avilés, J. A. *El periodismo audiovisual ante la convergencia digital*. Elche: U. Miguel Hernández, 2006
13. García Avilés, J. A.; León, B. "Journalistic Practice in Digital Television Newsrooms. The Case of Spain's Tele 5 and Antena 3". En: *Journalism*. 2002, v. 3, n. 3, pp. 355-371.
14. Garrison, B. "Diffusion of online information technologies in newspapers newsrooms". En: *Journalism*, 2001, v. 2, n. 2, pp. 221-239.
15. Garrison, B. "Online Information Use in Newsrooms. A longitudinal diffusion study". En: *Newspaper Division, Association for Education in Journalism and Mass Communication*. Phoenix, 2000. Consultado en 12/02/01.
<http://www.miami.edu/com/car/phoenix1.pdf>
16. GPD/SPC *Informe sobre la situació laboral i professional dels periodistes digitals a Catalunya*, Barcelona. 2002 Consultado en: 11/12/06.
<http://www.periodistesdigitals.org/cat/informes.htm>
17. GRID *Nous perfils professionals de l'actual panorama informatiu audiovisual i multimèdia de Catalunya*. Vic: Eumo, 2007
18. Hansen, K. A.; Ward, J.; Mcleod, R. "Role of the Newspaper Library in the production of News". En: *Journalism Quarterly*, 1987, v. 64, n. 4, pp. 714-720.
19. Hidalgo, P. "La documentación audiovisual en las televisiones. La problemática actual y el reto de la digitalización" En: *Documentación de las ciencias de la información*, 2005, v. 28, pp. 159-171.
20. López de Quintana, E. "Información multimedia en el entorno de la imagen. El centro de documentación de Antena 3 Televisión" En: *Cuadernos de Documentación Multimedia*, 1995, n. 4, pp. 69-90. Consultado en 15/07/07.
<http://www.ucm.es/info/multidoc/multidoc/revista/cuadern4/antena3.htm>
21. Marjoribanks, T. *News Corporation, Technology and the Workplace*. Cambridge: Cambridge University Press, 2000
22. Martín, J.; López Pavillard, J. "RTVE: Reorganización de la Documentación en un entorno multimedia" En: *Cuadernos de Documentación Multimedia*, 1999, n. 6-7. Consultado en 26/06/07. <http://www.ucm.es/info/multidoc/multidoc/revista/cuad6-7/rne.htm>
23. Masip, P. *Presencia i ús d'internet a les redaccions. Periodistes, rutines professionals i tecnologia*. Barcelona: Unversitat Ramon Llull, 2005. Tesis doctoral

24. McKecher, C. "Computer and Reporters. Newsroom Practices at Two Canadian Daily Newspapers". En: *Canadian Journal of Communication*. 1995, v. 20, n. 2.
25. Micó, J. L. *Teleperiodisme digital*. Trípodos: Barcelona, 2006
26. Micó, J. L. *La edición digital no lineal en los programas informativos de televisión*. València. Universitat Politècnica de València, 2003. Tesis doctoral.
27. Nicholas, D., et al. *The Media and the Internet*. Londres: ASLIB, 1998.
28. Palomo, B. *El uso redaccional de internet en la prensa diaria española*. Málaga: U.de Málaga, 2002. Tesis doctoral
29. Parra Valcárcel, D.; Álvarez Marcos, J. *Ciberperiodismo*. Madrid: Síntesis, 2004
30. Paul, N. "Media Libraries and New Media". En: *Netmedia97*. Londres: City University, 1997. pp. 1-4.
31. Pavlik, J. *Journalism and the new media*. Nueva York: Columbia University Press, 2001
32. Rintala, N.; Suolanen, S. "The Implications of Digitalization for Job Descriptions, Competencies and the Quality of Working Life". En: *Nordicom Review*, 2005, n. 2. Consultado en: 13/12/06.
http://www.nordicon.gu.se/common/publ_pdf/222_rintala_suolanen.pdf
33. Risoto, J. "Gestión de la documentación audiovisual en Televisión Valenciana" En: *Hipertext.net*, 2004, n. 2. Consultado en: 26/08/07. <http://www.hipertext.net>
34. Rubio Lacoba, M. "Nuevos tiempos para la documentación informativa en el periodismo digital: viejas y nuevas funciones del servicio de documentación digital" En: *Comunicación y Sociedad*, 2005, v. 18, n. 1, pp. 153-170.
35. Rubio Lacoba, M. "Seis reflexiones sobre la documentación en la Sociedad del Conocimiento". En: *Información Pública*, 2004, v. 2, n. 2, pp. 107-122.
36. Tapia, A., N. López, E. Medina y P. Gómez "La memoria del periodismo" En: *Anàlisi*, 2006, n. 33, pp. 119-133
37. Ward, J.; Hansen, K. A.; Mcleod, R. "Effects of Electronic Library on News Reporting Tools Protocols". En: *Journalism Quarterly*, 1988, v. 65, n. 4, pp. 845-852.
38. Weispfenning, J. "The Routinization of News Production", En Greenberg, B. S.; Gantz, W., eds. *Desert Storm and the Mass Media*. Chesskill, N.J.: Hampton Press, 1993 .

Cómo aprovechar su gestor de contenidos para tener una web bien posicionada en buscadores.

24/11/2006 - Autor: [Fernando Maciá](#)

© Fernando Maciá, 2006.

Director de Human Level Communications, consultora de posicionamiento en buscadores, gestión de campañas de enlaces patrocinados y desarrollo de webs inmobiliarias. Autor del libro Posicionamiento en buscadores, publicado por Anaya Multimedia. Profesor de Marketing Digital en Fundesem Business School.

Tener en cuenta los aspectos de optimización web para el posicionamiento en buscadores a la hora de implementar un gestor de contenidos puede evitar que nuestra inversión en contenidos quede devaluada por una pobre presencia en los buscadores.

Los gestores de contenido

La generación, publicación y archivo de un enorme número de páginas en los grandes portales y sitios Web plantea múltiples retos que los sistemas de gestión de contenidos web (o CMS, de *Content Management System*) han tratado de resolver en los últimos años:

- **Facilitar la generación y edición de contenidos** para la Web por parte de personal sin formación específica en programación.
- **Asegurar una apariencia homogénea** de todos los contenidos y su presentación de acuerdo con un diseño corporativo y con una línea editorial predefinida.
- **Mantener una coherencia en la estructura** de los sitios Web, permitiendo la incorporación de nuevos contenidos en las secciones adecuadas tras la aprobación y el control previo de las personas indicadas.
- **Mantener una navegación consistente** que permita a los usuarios llegar a cada uno de los contenidos que están publicados en cada momento.
- **Evitar la existencia de contenido duplicado** (distintas direcciones URL que muestran un mismo contenido), contenido huérfano (archivos que quedan en el servidor de forma innecesaria ya que son páginas a las que ya no apunta ningún enlace o archivos de imagen o multimedia que se mostraban en páginas suprimidas) o bien enlaces rotos, que apuntan a páginas que no existen en el servidor.

Los gestores de contenidos o CMS son herramientas de software que permiten descentralizar las labores de mantenimiento del contenido de un portal, de forma que personal no técnico de los distintos departamentos de una empresa puede añadir, editar y gestionar su propio contenido en una Web corporativa.

CMS y posicionamiento en buscadores: ¿una simbiosis imposible?

Sin embargo, y a pesar de sus evidentes ventajas, el enfoque tradicional de este tipo de herramientas se ha centrado en facilitar al máximo la gestión de los contenidos simplificando los procesos de producción, aprobación y publicación de los mismos, más que en generar páginas web debidamente optimizadas para ser competitivas en los buscadores.

Los factores que hacen que un portal web sea amigable para los buscadores se analizaron en distintos artículos de esta sección, así como en la Guía Básica de Posicionamiento en Buscadores.

Entre los problemas que, desde el punto de vista de la optimización para el posicionamiento en buscadores, aparecen de forma recurrente en los portales soportados por sistemas gestores de contenido, destacan los siguientes:

- **URLs dinámicas:** los buscadores limitan en ocasiones el número de variables dinámicas

presentes en las URLs que indexan. Las páginas generadas por muchos gestores de contenidos incluyen frecuentemente un elevado número de variables dinámicas en su dirección URL.

- **Títulos únicos:** el título de una página es uno de los factores más importantes a la hora de posicionar bien en buscadores un contenido. Sin embargo, muchos sistemas gestores de contenidos no permiten a los usuarios asignar un título único relevante a cada página.
- **Falta de soporte para metaetiquetas:** muchos CMS no disponen de campos específicos para que el usuario pueda especificar el contenidos de las metaetiquetas Keywords y Description. Aunque no son tan importantes como el título para lograr una buena posición en los buscadores, estas etiquetas siguen jugando un papel relevante a la hora de que el usuario prefiera hacer click sobre nuestro sitio en una página de resultados de búsqueda.
- **Ausencia de palabras clave en las URL:** las URLs dinámicas generadas por muchos sistemas de gestión de contenido suelen ser poco amigables tanto para el usuario como para los buscadores, y no incluyen términos de búsqueda que contribuyen a un mejor posicionamiento.
- **Imposibilidad de una optimización posterior:** el proceso de producción de contenidos que impone el uso de un sistema CMS dificulta enormemente la optimización a posteriori de los contenidos generados y, en el mejor de los casos, añade una carga de trabajo extra que podría evitarse si los aspectos SEO se hubieran tenido en cuenta en la implementación del gestor.

Resulta, pues, paradójico que precisamente sean aquellas empresas que más recursos invierten en el mantenimiento y generación de nuevos contenidos para sus portales las que, al mismo tiempo, menos se benefician del volumen que dicho tráfico podría movilizar debido a una pobre implementación de sus sistemas de gestión de contenidos, desde el punto de vista del posicionamiento en buscadores. Un fallo que, en muchos casos, no se debe a carencias de la herramienta en sí, sino a un desconocimiento por parte de los técnicos que la implementan de la importancia que tiene que los contenidos generados por ella puedan ser competitivos en los buscadores.

Hacer del gestor de contenido la mejor herramienta SEO

Pero, de la misma forma que un gestor de contenidos mal implementado puede mermar el rendimiento de la inversión en generación de contenidos de un portal, uno que haya tenido en cuenta los aspectos básicos de la optimización web para el posicionamiento en buscadores puede ser el aliado más eficaz para generar contenidos que conseguirán escalar los puestos de las búsquedas más competidas. Veamos cómo.

- **Emplear código válido según W3C:** los gestores de contenido parten de plantillas prediseñadas, que los usuarios no pueden alterar, para generar las nuevas páginas. Si validamos el código de estas plantillas en origen, nos aseguramos de que las páginas generadas a partir de ellas contendrán también código válido. El empleo de código válido asegura que la página se mostrará correctamente en los distintos navegadores y que los buscadores serán capaces de rastrearla sin problemas.
- **Crear un mapa del sitio:** casi todos los gestores de contenidos permiten crear y mantener actualizado un mapa del sitio. Los buscadores limitan el número de enlaces que pueden seguir a 100 por página y han de ser enlaces normales de texto HTML. Si ajustamos nuestro gestor de contenido para que genere y mantenga un mapa jerárquizado del sitio web con estas premisas, facilitamos a los buscadores que puedan rastrear todas y cada una de las páginas de nuestro sitio web.
- **Generar páginas con URLs amigables:** algunos gestores de contenido generan archivos estáticos cuyas URL no contienen variables dinámicas, mientras que otros emplean direcciones URL con múltiples parámetros dinámicos, como por ejemplo: <http://www.midominio.org/portal/site/Equipo/menuitem.6dab26af2ec93c76a68e76b1805101ca/?vgnnextoid=88c8c55c1edb7010VgnVCM100000081510acRCRD>. Entre éstos últimos, algunos permiten crear URL alias más amigables (conteniendo palabras clave o eliminando los parámetros dinámicos) que después el sistema sustituye internamente por la URL dinámica que precisa. En igualdad de condiciones, elijamos un CMS que genere direcciones URL ricas en palabras clave o con un reducido número de parámetros dinámicos. Direcciones URL como la del ejemplo son muy poco usables así como poco amigables para los buscadores.
- **Limitar el nivel de subdirectorios:** los buscadores dan mayor importancia a una página

cuanto más cerca está de la página home del portal. Es por esto que debemos limitar el número de subdirectorios que muestra la URL: muchos gestores de contenidos permiten organizar jerárquicamente el contenido independientemente de la localización física de los archivos en el servidor, presentando URLs mucho más sencillas que la estructura real del directorio correspondiente. Por ejemplo, la URL http://www.midominio.com/Viajes/Destinos/Ciudades/MasInfo/DatosGenerales/Congresos/Congresos_negocios.htm hace que este contenido esté situado seis niveles de profundidad por debajo de la página home.

- **Conectar el control de validación de enlaces** del CMS: la mayoría de gestores controla la publicación de enlaces rotos que apuntan a contenidos controlados por el propio gestor, pero pocos validan que un enlace que apunta a una web externa no es un enlace roto. Si el control existe, asegurémonos de conectarlo para impedir que un usuario pueda introducir un enlace hacia una página web inexistente.
- **Dejar el control del archivo robots.txt al webmaster:** algunos gestores de contenidos permiten editar el contenido del archivo robots.txt por parte del autor de una página. En general, es mejor que sólo el webmaster controle el contenido de este archivo para impedir que, por desconocimiento, un usuario pueda bloquear el rastreo de los robots de una importante parte de la Web.
- **Evitar URLs duplicadas:** los buscadores son tremendamente selectivos a la hora de castigar el contenido duplicado de una Web, de modo que debemos asegurarnos de que cada página sólo existe bajo una única URL. En cualquier caso, si queremos que los usuarios puedan llegar a un mismo contenido desde URLs homólogas, es mejor programar redireccionamientos permanentes 301, que no son castigados por los buscadores.
- **Evitar variables de sesión en la URL:** si nuestro portal realiza transacciones electrónicas o algún otro proceso que requiera mantener el estado del cliente, es preferible emplear una cookie de sesión que una variable de sesión que aparece en la URL. La razón es que si los buscadores detectan esto, se abstendrán de rastrear la página para evitar indexar como páginas distintas la misma, pero con distinta variable de sesión. Por ejemplo: <http://www.midominio.com/index.php?module=Health&action=DefaultIndex&PHPSESSID=98ab41f9acd8c74df4b6e063985bda8d>. En esta URL podemos ver que existe un parámetro de identificación de la sesión (PHPSESSID) que, si es detectado por los buscadores, impedirá que la página sea rastreada, ya que sucesivas visitas del buscador archivarían como páginas distintas ésta misma, en la que únicamente cambiaría el valor del identificativo de sesión. En cualquier caso, se debe restringir el empleo de identificadores de sesión exclusivamente a las áreas del portal en los que sea absolutamente necesario.
- **Reducir el código basura:** simplificar el código HTML que emplean las plantillas y optar por usar Hojas de Estilo en Cascada (CSS) en lugar de tablas para maquetar el contenido. El empleo de Hojas de Estilo en Cascada facilita la actualización del diseño de una Web, reduce considerablemente el peso de los archivos (la maquetación viaja una sola vez del servidor al usuario, después está ya disponible en la memoria temporal o caché del navegador para las siguientes páginas visitadas) y da mayor preponderancia al contenido significativo de la página respecto al peso total del código de la misma.
- Seleccionar la opción de **texto para la navegación** del sitio: evitando siempre que sea posible el uso de menús Javascript o Flash, ya que sus enlaces no pueden ser seguidos por los buscadores. En muchos casos, podremos lograr con el uso de CSS efectos parecidos a los de los menús Javascript o Flash. Si el gestor de contenidos permite crear un rastro de migas de pan (*breadcrumb*), activémoslo. Mejora la usabilidad del sitio, ayuda a situarse al usuario en el conjunto de la estructura de la web y es un inmejorable atajo para que los buscadores rastreen todo el contenido.
- **No olvidar que existen los encabezamientos:** el uso de estilos nos hace olvidar la existencia de etiquetas HTML de jerarquía (H1, H2, H3, etc.) cuyo aspecto visual final podemos modificar también con estilos, pero que ayudan a los buscadores a entender mejor la estructura lógica de la página y señalizan qué aspectos son más importantes. Es por ello importante animar a los editores de contenido a usar encabezamientos en lugar de simplemente definir el texto más grande o más pequeño con font size y, si es posible, limitar a un único encabezamiento de máximo nivel (H1) por página.
- **Exigir la creación de un título único,** y la adición de metaetiquetas relevantes: programar el CMS de tal forma que la cumplimentación de título y metaetiquetas sea un requisito para la publicación del contenido y, si es posible, activar un control para

comprobar la unicidad del título.

- **Requerir cumplimentar la etiqueta alt** al añadir una imagen al contenido. Esta etiqueta permite a los buscadores indexar mejor las imágenes, colabora en la relevancia de los términos clave de la página y mejora la accesibilidad a los contenidos por parte de las personas con problemas de visión.
- **Implementar controles para evitar la publicación de contenido duplicado.**
- **Animar a usar textos descriptivos en enlaces:** en lugar de "Haga click aquí", emplear "Más información sobre nuestro servicio de asistencia al cliente 24H".

Lo mejor de ambos mundos

Es evidente, pues, que si al implementar un gestor de contenidos para generar y mantener el contenido de un gran portal se tienen en cuenta los aspectos relacionados con la optimización web y el posicionamiento en buscadores de dicho portal podemos animar o, de alguna forma, imponer una cierta disciplina a través del propio sistema CMS que implique a los editores de contenido a crear páginas fácilmente rastreables e indexables, que puedan competir adecuadamente en los buscadores.

Las grandes empresas cuentan con la materia prima favorita de los buscadores: contenido abundante, original, dinámico y de rápida actualización. Aprovechemos toda la capacidad de los gestores de contenidos para extraer el máximo rendimiento de la inversión que efectúan en su presencia en la Red.

Este artículo puede ser reproducido siempre que se respete su integridad y se acompañe de la siguiente línea de crédito en la misma página en que aparezca el mismo y conservando los respectivos enlaces:

© Fernando Maciá, 2006.

Director de Human Level Communications, consultora de posicionamiento en buscadores, gestión de campañas de enlaces patrocinados y desarrollo de webs inmobiliarias. Autor del libro Posicionamiento en buscadores, publicado por Anaya Multimedia. Profesor de Marketing Digital en Fundesem Business School.

Herramientas de software libre para la gestión de contenidos

Autor: Jesús Tramullas
Situación en la jerarquía: [Inicio](#) -> [Números publicados](#) -> [Núm. 3, mayo 2005](#) -> Software Libre
URL de esta página: <http://www.hipertext.net/web/pag258.htm>
Citación: Jesús Tramullas. *Herramientas de software libre para la gestión de contenidos* [on line]. "Hipertext.net", núm. 3, 2005. <<http://www.hipertext.net>> [Consulta: 28/01/2010]. ISSN 1695-5498

■ Sumario

- ▼ 1. [Gestión de información y gestión de contenidos](#)
- ▼ 2. [Actividades y procesos de gestión de contenidos](#)
- ▼ 3. [Soluciones de software libre para gestión de contenidos](#)
- ▼ 4. [Una propuesta de tipología](#)
- ▼ 5. [Perspectivas de desarrollo](#)
- ▼ 6. [Referencias](#)
- ▼ 7. [Enlaces](#)

▲ 1. Gestión de información y gestión de contenidos

Los servicios de información y documentación accesibles a través de internet, más concretamente mediante servidores web, están aumentando de una forma exponencial. La lógica evolución del web desde hace más de 10 años ha ido produciendo la sustitución de páginas y documentos estáticos por documentos generados dinámicamente, merced a la interacción del usuario con la lógica de procesos y flujos de trabajo definida por los creadores del servicio y a la disponibilidad de cada vez mayores repositorios de información. Evidentemente, se ha ido pasado progresivamente de un concepto de publicación de páginas web, bastante simple en su origen, a esquemas más complejos y diferenciados, fundamentados en procedimientos y técnicas basados en la gestión de información. La cada vez mayor complejidad de los servicios y de los sistemas que los soportan, ha hecho necesaria la formulación de un corpus teórico y práctico en el que se combinen las técnicas clásicas de gestión de información en las organizaciones con las características propias del medioambiente digital.

Esta evolución, que se ha acelerado durante la primera mitad de la década de 2000, ha tenido un impacto no sólo en los métodos y técnicas de gestión de información, sino también en la propia tecnología para gestión de información y, en consecuencia, en el mercado de productos y servicios (GILBANE, 2000). Si bien 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. Las herramientas para este trabajo han recibido la denominación de sistemas de gestión de contenidos (o *Content Management Systems*, CMS), y se han integrado con los sistemas de gestión documental y con los de recuperación de información. A ello hay que unir que, en la concepción actual de la gestión de información, el control de los procesos es un elemento nuclear, por lo que se acompañan de sistemas de *workflow*, o de flujos de trabajo. Con todo ello se puede delinear un paisaje en el cual las herramientas de gestión documental han ido incorporando a sus prestaciones las capacidades necesarias para gestionar los procesos que crean, almacenan, tratan y presentan información, en entorno digital.

Sin embargo, no por ello cabe afirmar que existe una igualdad directa entre los sistemas de gestión de contenidos y los sistemas de gestión documental. Se pueden encontrar en el mercado sistemas de gestión de contenidos que no ofrecen las prestaciones documentales que serían deseables, y viceversa. De la misma forma, no es posible igualar mediante una ecuación gestión de contenidos y gestión documental o *records management*. Íntimamente relacionadas sí, pero iguales no. Sólo el estudio y la evaluación de las características y prestaciones presentes en las herramientas disponibles pueden determinar la adecuación de una solución en un contexto o problema dado. A ello cabe añadir las diferencias existentes entre la gestión de contenidos para web y la gestión de contenidos para empresas, y que pueden encontrarse en los informes o documentos especializados que publican las propias compañías del sector. Si a esto se suma la complejidad creciente de muchos portales, tanto internos como externos, de las organizaciones, que son soportados por sistemas de gestión de contenidos, y a los procesos de publicación digital necesarios para su producción, los sistemas de gestión de contenidos se configuran como aplicaciones de varias escalas, que pueden llegar a alcanzar una alta complejidad.

En la parte humana, las herramientas para gestión de contenidos sólo se pueden entender en un entorno de trabajo en colaboración y distribuido, ya que es en este tipo de ambientes donde se puede aprovechar todo su potencial. De la lectura de los párrafos anteriores se deduce que, frente a enfoques tradicionales de tratamiento documental, pasivos, las herramientas actuales han llevado más allá

los límites del tratamiento documental, ya que se han extendido hasta el proceso de creación, por su parte inicial, y al proceso de nueva publicación y de personalización, por su parte final. Esto configura ahora un esquema circular e iterativo para el tratamiento documental, a lo que se debe unir, entonces, la necesaria reflexión sobre el concepto y características de los documentos digitales. El ciclo de vida de los documentos digitales muestra significativos cambios sobre el ciclo de vida tradicional, al igual que el concepto de documento. Evidentemente, las actividades informativo-documentales deben reformularse para hacer frente al nuevo medioambiente digital.

▲2. Actividades y procesos de gestión de contenidos

La gestión de contenidos debe definirse desde la perspectiva de su objetivo y actividades. Frente a la gestión de documentos, que tiene como objeto de trabajo documentos, como su propio nombre indica, la gestión de contenidos está orientada a gestionar objetos que actúan como componentes de documentos virtuales, en el contexto de lo que llama segmentación. Los objetos 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). Como en otros métodos y técnicas directamente relacionados con el desarrollo de la Sociedad de la Información y del Conocimiento, las formulaciones teóricas han seguido a la necesidad y a su resolución tecnológica, o han discurrido en paralelo. En este sentido, la gestión de contenidos se ha beneficiado de las técnicas y procesos de producción y edición de publicaciones digitales en un primer momento. Sin embargo, rápidamente pudo apreciarse que el proceso de publicación digital no cubría por completo todos los procesos de gestión de información necesarios para una adecuada gestión de contenidos.

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. Evidentemente, este proceso, que puede reflejar el esquema más simple de funcionamiento de un weblog, 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 -en el momento actual- 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. Otro enfoque especializado es el que considera las herramientas para desarrollo de bibliotecas digitales como sistemas de gestión de contenidos, como por ejemplo el reciente trabajo de Han (2004), que compara Fedora, Dspace y Greenstone.

Los trabajos sobre 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, flujo de trabajo y repositorios de información; un repositorio de información; herramientas de integración de información externa; y modelos y *templates* para productos finales. Un CMS se compone de varios subsistemas (BOIKO, 2001) que interaccionan entre ellos:

- Colección: subsistema 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.
- Gestión: subsistema 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.
- Publicación: subsistema 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.

Nakano (2002) señala cuatro subsistemas, en lugar de tres, 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), Gilbane Report (BLUEBILL, 2003), CMS-Watch (2005), o la Guía CMS-Spain (CMS-SPAIN, 2004).

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. ASILOMAR (2003) ha publicado una interesante encuesta sobre los problemas que plantean los CMS a sus usuarios. Como ha señalado Veen (2004), el éxito o el fracaso de un CMS en una organización no es tanto una cuestión tecnológica como una cuestión de personas y de procesos.

▲3. Soluciones de software libre para gestión de contenidos

El software libre es un fenómeno que está llamado a revolucionar modelos de negocio de la industria del software. No es este el momento de analizar con detalle el mundo del software libre, ni las plataformas tecnológicas, ni la tipología de licencias existente, ya que para todo ello son muy abundantes los recursos de información disponibles. Para los objetivos de este texto es suficiente con establecer los criterios básicos que deben cumplir las herramientas reseñadas:

- ofrecer el código fuente de la aplicación
- distribuirse bajo alguna de las licencias consideradas de referencia (véanse las licencias disponibles en *Open Source Initiative* , <http://www.opensource.org>)
- poder ser modificadas, copiadas y distribuidas libremente, respetando los términos establecidos en la licencia respectiva.

Una revisión somera de las herramientas para gestión de contenidos que son distribuidas bajo alguna de las licencias consideradas para software libre muestra la existencia de un gran número que ofrecen variadas prestaciones y orientaciones. Brevemente, su arquitectura técnica se fundamenta en el tercer servidor web, intérprete de lenguaje de programación y gestor de base de datos. A este esquema responde el conocido acrónimo LAMP (Linux, Apache, MySQL, PHP), o su versión Windows, WAMP. Precisamente han sido PHP (<http://www.php.net>) y MySQL (<http://www.mysql.com>) las herramientas más extendidas entre los sistemas libres para gestión de contenidos, ya que la mayoría de las soluciones se basan en ellos. Ejemplos de todo esto se pueden encontrar en *OpenSourceCMS* (<http://www.opensourcecms.com/>), y en *CMS Matrix* (<http://www.cmsmatrix.org/>), que ofrece una matriz de comparación muy útil y exhaustiva para comparar los requerimientos y prestaciones de las diferentes herramientas. Muchas de estas herramientas son dignos contendientes de soluciones propietarias de alto coste, y están siendo utilizadas en numerosas intranets y portales de todo tipo de organizaciones, lo cual ha permitido desarrollar un mercado y un modelo de negocio, que por el momento parece sostenible a corto y medio plazo. En la última edición celebrada de *LinuxExpo*, Mambo (<http://www.mamboserver.com>), un CMS bajo licencia GPL, ha ganado el premio *Best Open Source Solution*, ante productos de grandes empresas. El interés de este tipo de soluciones, tanto en los clientes como en los proveedores de servicios de valor añadido, puede verse reflejado en Holst (2001), Dolye (2003) y Robertson (2004).

▲4. Una propuesta de tipología

Si bien la gestión de contenidos es una disciplina en auge, y que cada vez ofrece mayor nivel teórico y metodológico, junto a herramientas en constante desarrollo, también es cierto que bajo este "metatérmino" se pueden encontrar herramientas y plataformas que muestran diferentes orientaciones, diferentes prestaciones, y enfocadas a diferentes objetivos y grupos de usuarios. Para los propósitos de este trabajo, se ha optado por crear una tipología esquemática.

- Plataformas para desarrollo de gestión de contenidos: se trata de soluciones que ofrecen la plataforma necesaria para desarrollar e implementar aplicaciones que den solución a necesidades específicas. Ofrecen un entorno y unas herramientas de desarrollo. En consecuencia, su interés radica

en la posibilidad de construir soluciones adaptadas a cada caso. Demandan un alto nivel de conocimiento de lenguajes de programación de alto nivel, como por ejemplo Java o Python, mientras que otros usan PHP. Pueden servir para construir soluciones de gestión de contenidos, para soluciones de comercio electrónico, o para construir, a su vez, herramientas que respondan a las características de los tipos que se explican en los apartados posteriores. En muchos de ellos diferentes desarrolladores y usuarios han ido creando módulos, de forma que es posible integrar desarrollo propio con herramientas disponibles. Las características de cada plataforma varían, ya que se pueden encontrar desde sólo la plataforma, como *Zope*, a entornos que ofrecen las prestaciones básicas necesarias para gestión de contenidos, con interfaces y flujos de trabajo especialmente diseñados para ello, como *OpenCMS*. Sin pretender ser exclusivo, las plataformas más extendidas son:

- Zope, <http://www.zope.org/>
- Typo3, <http://www.typo3.org/>
- Midgard Project, <http://www.midgard.project.org/>
- OpenCMS, <http://www.opencms.org/>
- Apache Lenya, <http://lenya.apache.org/>

Fig 1. Zope

ZOPE
Community

You are here: Home

You are not logged in [Log In](#) [Join](#)

Download

- Download Zope
- Zope Products
- Zope CMP

Resources

- Project Status
- Mailing Lists
- User Groups
- Licenses

Documentation

- The Zope Book
- API reference
- ZPT reference
- Developer Guide
- CMP Documentation
- Zope Articles
- Zope How-To's
- Books In Print

About Zope

- What is Zope?
- Zope Powered Sites
- About Zope.org
- Member FAQ

Business

- Zope Corporation
- Solution Providers
- Zope Training

Developers

- Zope Projects
- Zope 3
- Zope CVS

Welcome to Zope.org

The Web Site for the Zope Community

Zope is an open source application server for building content management systems, intranets, portals, and custom applications. The Zope community consists of hundreds of companies and thousands of developers all over the world, working on building the platform and Zope applications. Zope is written in Python, a highly-productive, object-oriented scripting language. [more](#)

Zope for...

- New Users**
[Introduction To Zope](#), [About Python](#)
- Product Developers**
[Developer's Guide](#), [Unit Testing](#)
- Web Masters**
[Using Apache](#), [Security](#), [ZPT](#), [GITM](#)
- Zope Developers**
[Content Management Framework \(CMP\)](#), [Products](#), [Using CVS/SVN](#), [Bus Days](#)

Zope 4 Intranets Training
February 21 - 23, 2005

Zope 3 Training for Developers
April 18 - 20, 2005

News [RSS](#)

- [Open Source Realize Forum](#)
Mar 19 - 19 February 22
- [LTOnlineTest v0.2.1 Released](#)
February 21
- [QIQ/Zope on G/NET Taiwan 4 2-Day Seminar at 222](#)
February 21
- [New development version of LocalImageSystem](#)
February 21
- [Infrae release: Site 1.2 final](#)
February 21
- [February Zope Bus Day](#)
February 21
- [Zope3 Live in Copenhagen](#)
February 19

Products [RSS](#)

- [CORPUS - Zope Weblog Product \(1.11\)](#)
February 21
- [\(0.9.9220\)](#)
February 21
- [GMailFolder \(0.5\)](#)
February 21
- [SMTP MailHost \(2.0.0\)](#)
February 19
- [SMTP MailHost \(2.0.0\)](#)
February 19

- Portales: la creación y mantenimiento de portales, como soporte y herramienta básica de los servicios de información web, es una de las funciones principales que desempeñan los sistemas de gestión de contenidos para portales. Su funcionalidad, administración y mecanismos de control están especialmente orientados a ofrecer a sus usuarios un portal con diferentes tipos de contenidos y de servicios, desde la publicación de noticias, al repositorio de documentos, pasando por foros, encuestas, sindicación de contenidos, creación de perfiles y de grupos de usuarios, personalización de la información y de su presentación, etc. Siguen una arquitectura modular, ya que se componen de módulos, encargados de diferentes funciones, que son administrados desde una interfaz centralizada. Adoptan el esquema de espacios en su interfaz, para lo cual usan los ya citados *templates*, en los cuales se distribuyen los diferentes módulos. Técnicamente, la mayoría de ellos requieren LAMP/WAMP, aunque existen algunas herramientas que usan Python (Plone sobre Zope) o Perl (como WebGUI). Los más extendidos son:

- PHP Nuke, <http://phpnuke.org/>
- Drupal, <http://www.drupal.org/>
- Mambo, <http://www.mamboserver.com/>
- Plone (requiere Zope), <http://www.plone.org/>

Fig. 2. Mambo



- Aula virtual: entornos que ofrecen las prestaciones necesarias para crear contenidos para aprendizaje en línea, y ciertos mecanismos de interacción, como foros, chats, evaluación interactiva, etc. En realidad, la mayoría de ellos siguen esquemas de aula clásica, traspasados al entorno web, más que ser entornos de aprendizaje virtual proactivo. Este tipo de plataformas han ajustado las prestaciones de gestión de contenidos a las características del material docente en formato

digital. Es altamente recomendable que la plataforma que se utilice sea capaz de empaquetar sus contenidos en formatos SCORM y/o IMS. Las plataformas más conocidas son:

- Claroline, <http://www.claroline.net/>
- Moodle, <http://www.moodle.org/>

- Bibliotecas digitales: las bibliotecas digitales, el paradigma para los servicios de información de finales del siglo XX y comienzos del siglo XXI, organizadas alrededor del terceto que forman los usuarios, las colecciones, y los servicios de valor añadido, se configuran como un espacio altamente especializado para la gestión de contenidos. El desarrollo de colecciones digitales, la organización y creación de los mecanismos de acceso o la gestión de metadatos se benefician de la combinación de los procesos de publicación digital, y de los principios de gestión de información. Además, tienden a configurar espacios de colaboración, mediante el acceso y la gestión de colecciones distribuidas. El estado actual de las principales herramientas es diverso, así como sus enfoques, ya que van desde repositorios federados de documentos (como Fedora), hasta herramientas de usuario final (como Greenstone). Este tipo de herramientas es más exigente, en sus requerimientos, que los otros tipos indicados en este texto.

- Fedora, <http://www.fedora.info/>
- Dspace, <http://www.dspace.org/>
- Greenstone, <http://www.greenstone.org/>

Fig. 3. Greenstone

greenstone digital library software PREFERENCES

[download](#) [examples](#) [docs](#) [faq](#) [support](#)

About Greenstone

Greenstone is a suite of software for building and distributing digital library collections. It provides a new way of organizing information and publishing it on the internet or on CD-ROM. Greenstone is produced by the **New Zealand Digital Library Project** at the **University of Waikato**, and developed and distributed in cooperation with **UNESCO** and the **Human Info NGO**. It is **open-source**, **multilingual** software, issued under the terms of the GNU General Public License.

NEW There is now a Greenstone Spanish Users Group, more information [here](#).

NEW [Download](#) all the language interfaces available for the Greenstone digital library software (version 2.52 only): the four "core" languages English, French, Spanish, Russian, and interfaces for Arabic, Armenian, Catalan, Chinese, Croatian, Czech, Dutch, Farsi, Finnish, Galician, Georgian, German, Greek, Hebrew, Hindi, Indonesian, Italian, Japanese, Kannada, Kazakh, Latvian, Maori, Portuguese (Brazil), Portuguese (Portugal), Serbian, Thai, Turkish, Ukrainian, Vietnamese.

NEW If you are running the Local Library server on Windows, you can download a small package to upgrade to Greenstone 2.52. [Upgrade from Greenstone 2.51](#), or [upgrade from Greenstone 2.50](#). If you are upgrading a UNESCO 2.50 CD-ROM, you should delete the macros\home.dm file and rename the macros\home.dm.bak file to macros\home.dm after the upgrade has finished.

NEW [Download](#) Greenstone v2.52, which includes many minor bug fixes as well as several major additions, including: a built-in OAI server, export to METS functionality, a new "mirror" pane in the Librarian Interface for downloading web resources, a new applet version of the GUI, an improved metadata set editor (called GEMS), partial query matching for MGPP, an OggVorbis plugin, and Lucene building support.

- Publicaciones digitales: son plataformas especialmente diseñadas teniendo en cuenta las necesidades de las publicaciones digitales, tales como periódicos, revistas, etc. En consecuencia, sus prestaciones están orientadas al control de los procesos de edición, creación y publicación de contenidos, más que al desarrollo de aplicaciones o portales. De especial interés van a resultar en los próximos años todos los sistemas dedicados a la publicación de revistas científicas que den soporte al proceso de envío, revisión y publicación. Aunque no es exactamente un sistema de publicación de revistas se ha incluido en el listado *ePrints* por su interés para el ámbito académico y su posibilidad de ser adaptado para publicaciones.
- Cofax, <http://www.cofax.org/>
- Open Journal Systems, <http://www.pkp.ubc.ca/ojs/>

- ePrints, <http://www.eprints.org/>
- Entornos para colaboración: corresponderían a la clásica definición de *groupware*, herramientas para trabajo en grupo. Su objetivo es dar a grupos de usuarios especializados las prestaciones necesarias para llevar a cabo trabajos y proyectos en común. En este tipo de entornos, los flujos de trabajo, los usuarios, los puntos de control y los documentos entregables (*deliverables*) son los contenidos clave a gestionar. Además, se necesitan herramientas de control de tiempos y actividades junto a posibilidades de comunicación síncrona y asíncrona. Se trata, entonces, de una especialización evidente de la gestión de contenidos que se relaciona con la gestión de proyectos. Un entorno colaborativo muy conocido y extendido es *wiki*, sobre todo destinado a la elaboración de documentos para comunidades de interés. Existen gran número de plataformas *wiki* que han ido aumentando progresivamente sus prestaciones, de manera que unos autores las sitúan entre las herramientas para portales, mientras otros las colocan en este apartado:
 - eGroupware, <http://www.egroupware.org/>
 - phpCollab, <http://www.php-collab.org/>
 - Wiki (web de referencia), <http://www.wiki.org/>
- Blogs o bitácoras: los blogs son el fenómeno, ya consolidado, que ha dado un verdadero potencial democrático a la web, a pesar de los inconvenientes que pueda conllevar. Los blogs muestran un modelo de gestión de contenidos bastante simplificado, ya que suelen ser monousuario, y con un sencillo flujo de trabajo, lo que ha facilitado su expansión entre amplios grupos de usuarios sin conocimientos técnicos profundos. A pesar de que los grandes servidores de blogs dan al usuario la solución ya implementada, existen herramientas para blogs, dentro del modelo LAMP/WAMP, que puede ser instalados y operados bajo licencia libre. Debe indicarse, en cambio, una tendencia a aumentar las prestaciones de los blogs, intentando dotarlos de parecidas características a las disponibles en las herramientas para portales.

- WordPress, <http://wordpress.org/>

Fig. 4. WordPress



▲5. Perspectivas de desarrollo

La gestión de contenidos, y las herramientas de software libre que facilitan su implementación, son una realidad que ni puede ni debe pasar desapercibida, ya que ofrecen soluciones para los servicios que debe ofrecer, en el entorno digital, una unidad de información. Existen soluciones de tipo general y soluciones especializadas. Además, es innegable la existencia de un creciente mercado de

productos y soporte para estas herramientas. Sin embargo, y a modo de conclusión, es necesario delinear varios factores críticos que deben ser tenidos en consideración:

- **Requerimientos:** si bien en el momento actual unos conocimientos técnicos de nivel medio son suficientes para poner en marcha un servicio con estas herramientas, en un futuro cercano aumentará el requerimiento técnico, sobre todo debido al creciente volumen de información, a la complejidad de los documentos digitales y a la exigencia de niveles de seguridad.
- **Documentación:** el soporte profesional para estas soluciones es un coste oculto que debe ser valorado. La documentación de administración y de usuario final todavía no alcanza, en muchos casos, el nivel de exigencia que sería deseable.
- **Objetivos y planificación:** a pesar de lo que pueda parecer, todavía se siguen montando servicios de información en Internet en los que los objetivos y una correcta planificación son difuminados por el espejismo tecnológico. Sólo un correcto análisis y planificación podrán ayudar a determinar la elección de la herramienta más adecuada: "la forma sigue a la función".
- **Web semántico:** a pesar del desarrollo del XML y de sus lenguajes derivados, y de los esquemas de descripción de recursos de información y de metadatos, son todavía escasas las soluciones para gestión de contenidos que trabajan directamente con ellos. No existe apenas soporte para la utilización de RDF o de ontologías. En un entorno digital enfrentado al crecimiento exponencial de la información, este es un problema al que hacer frente con celeridad. Las herramientas que lo hagan tendrán una ventaja competitiva más que notable.

▲6. Referencias

ASILOMAR INSTITUTE FOR INFORMATION ARCHITECTURE (2003), <i>The Problems with CMS</i> (disponible en http://aifia.org/pg/the_problems_with_cms.php ; consultado 12-11-2004).
BLUEBILL ADV. (2003), "The Classification & Evaluation of Content Management Systems."En: <i>The Gilbane Report</i> , vol. 11, n. 2 (disponible en http://www.gilbane.com/gilbane_report.pl/86/The_Classification_Evaluation_of_Content_Management_Systems.html ; consultado 22-1-2005).
BOIKO, B. (2001), <i>Content Management Bible</i> . Wiley.

BROWNING, P. y LOWNDES, M. (2001), <i>JISC TechWatch Report: Content Management Systems</i> . TechWatch Report TSW 01-02, The Joint Information Systems Committee.
CMS-SPAIN (2004), <i>Guía CMS-Spain 2004</i> . (disponible en http://www.ecm-spain.com/interior.asp?IdItem=1851 ; consultado 24-11-2004).
CMS-WATCH (2005), <i>The CMS Report: In-Depth Analysis of Web Content Management Solutions</i> . (disponible en http://www.cmswatch.com/TheCMSReport/ ; consultado 3-2-2005).
GINGELL, D. (2003), <i>A 15 Minutes Guide to Enterprise Content Management</i> . Documentum Inc.
GUPTA, V.K., GOVINDARAJAN, S., JOHNSON, T. (2001), "Overview of content management approaches and strategies." En: <i>Electronic Markets</i> , vol. 11, n. 4, p. 281-287.
DOYLE, B. (2003), "Open Source Content Management Redux. " En: <i>The Gilbane Report</i> , vol. 11, n. 3(disponible en http://www.gilbane.com/gilbane_report.pl/87/Open_Source_Content_Management_Redux.html ; consultado 2-2-2005).
ERP Software (2003), <i>Content Management Tutorial</i> , (disponible en http://erptoday.com/CMS/Content-Management-Tutorial.aspx ; consultado 15-01-2005).
GILBANE, F. (2000),"Whats content management? " En: <i>The Gilbane Report</i> , vol. 8, n. 8(disponible en http://www.gilbane.com/gilbane_report.pl/6/What_is_Content_Management ; consultado 24-1-2005).
HAN, Y. (2004), "Digital content management: the search for a content management system." En: <i>LibraryHi Tech</i> , vol. 22, n. 4, p. 355-365.
HOLST. S., (2001), "Open Source Content Management: A Parallel Universe? "En: <i>The Gilbane Report</i> , vol. 9, n. 4 (disponible en http://www.gilbane.com/gilbane_report.pl/1/Open_Source_Content_Management_A_Parallel_Universe.html ; consultado 22-1-2005).
MCKEEVER, S. (2003), "Understanding web content management systems: evolution, lifecycle and market." En: <i>Industrial Management & Data Systems</i> , vol. 103, n. 9, p. 686-692.
MILLER, B. y DAVID, D. (2002), <i>Directions in Web Content Management</i> . Burntsand, Inc.
NAKANO, R. (2002). <i>Web Content management. A Collaborative Approach</i> . Prentice Hall.
ROBERTSON, J. (2003), "So, what is a content management system?". En: <i>KM Column</i> , June 2003 (disponible en http://www.steptwo.com.au/papers/kmc_what/index.html ; consultado 1-2-2005).

ROBERTSON, J. (2004), "Open-source content management systems." En: <i>KM Column</i> , January 2004, (disponible en http://www.steptwo.com.au/papers/kmc_opensource/index.html ; consultado 12-2-2005).
ROCKLEY, A. (2003), <i>Managing Enterprise Content. A Unified Content Strategy</i> . New Riders.
STEP TWO (2204), <i>Content Management Requirements Toolkit</i> . Step Two, Sydney.
VEEN, J. (2004), <i>Why Content Management Fails</i> . Adaptive Path. (Disponible en http://www.adaptivepath.com/publications/essays/archives/000315.php ; consultado 16-10-2004).
WHITE, M. (2002), "Selecting a content management system". En: <i>VINE</i> ,vol. 32, n. 2, p. 34-39.

▲7. Enlaces

AIIM Enterprise Content Management Association, http://www.aiim.org/
EDOC Magazine, http://www.edocmagazine.com/
CMS Matrix, http://cmsmatrix.org/
CMS Professionals, http://www.cmprofessionals.org/
CMS Review, http://www.cmsreview.com/
CMS Watch, http://www.cmswatch.com/
CMS Wiki, http://www.cmswiki.com/tiki-index.php
CSM Wire, http://www.csmwire.com/
CSM Spain, http://www.cms-spain.com/
OSCOM, http://www.oscom.org/

XML y la gestión de contenidos

Autor: Ricardo Eíto Brun

Situación en la jerarquía: [Inicio](#) -> [Números publicados](#) -> [Núm. 3, mayo 2005](#) -> XML

URL de esta página: <http://www.hipertext.net/web/pag256.htm>

Citación: Ricardo Eíto Brun. *XML y la gestión de contenidos* [on line]. "Hipertext.net", núm. 3, 2005. <<http://www.hipertext.net>> [Consulta: 28/01/2010]. ISSN 1695-5498

■ Sumario

- ▼ 1. Una breve introducción al lenguaje XML
- ▼ 2. Evolución continua del lenguaje
- ▼ 3. El reconocimiento de la normalización
- ▼ 4. Definición de gestión de contenidos
 - ▼ 4.1. Origen de la gestión de contenidos
 - ▼ 4.2. Descripción funcional de un sistema de gestión de contenidos
 - ▼ 4.2.1. Mantenimiento de la estructura lógica del sitio
 - ▼ 4.2.2. Edición de contenidos mediante plantillas
 - ▼ 4.2.3. Mantenimiento automático de la estructura de navegación
 - ▼ 4.2.4. Aprobar, revisar y validar contenidos
 - ▼ 4.2.5. Periodos de vigencia de las páginas
 - ▼ 4.2.6. Control de cambios y revisiones
 - ▼ 4.2.7. Páginas conectadas: compartir contenidos entre páginas
 - ▼ 4.3. Descripción técnica de un sistema de gestión de contenidos
 - ▼ 4.3.1. Páginas dinámicas
 - ▼ 4.3.2. Bases de datos y repositorios
 - ▼ 4.3.3. Metadatos
 - ▼ 4.3.4. Gestión de flujos de trabajo
 - ▼ 4.3.5. Integración de contenidos y aplicaciones
- ▼ 5. La función de XML en la gestión de contenidos
 - ▼ 5.1. XML como base para el almacenamiento de contenidos
 - ▼ 5.2. XML como modelo para la representación de metadatos
 - ▼ 5.3. XML como medio de intercambio e integración de contenidos
- ▼ 6. Conclusiones
- ▼ 7. Bibliografía
- ▼ 8. Notas

▲1. Una breve introducción al lenguaje XML

El lenguaje *XML (eXtensible Markup Language)* comenzó a desarrollarse en septiembre de 1996 auspiciado por el W3C con un claro propósito: diseñar un lenguaje de marcas optimizado para Internet. XML debía combinar la simplicidad de HTML con la capacidad expresiva de su predecesor, SGML [1].

En su definición participaron empresas como Microsoft, IBM, Sun Microsystems, Novell y Hewlett-Packard. La versión 1.0 fue ratificada por el W3C en la conferencia sobre SGML/XML celebrada en Washington en diciembre de 1997. En pocos años, XML se ha convertido en el lenguaje con mayor impacto en el desarrollo de aplicaciones de publicación de contenidos para Internet e Intranet.

La edición de documentos XML persigue los siguientes objetivos:

- Distinguir el contenido y la estructura de los documentos de su presentación en papel o en pantalla.
- Hacer explícita su estructura y sus contenidos informativos.
- Crear documentos que puedan intercambiarse y procesarse con facilidad en sistemas informáticos heterogéneos.

Para alcanzar estos objetivos XML propone un formato en el que se intercalan marcas en el texto de los documentos con el objeto de distinguir las partes o elementos estructurales del mismo. Las principales características del lenguaje son:

- La posibilidad de marcado descriptivo, con un conjunto de marcas abierto. En HTML y XML se intercalan marcas en los documentos. La principal diferencia entre uno y otro está en la función de estas marcas. En XML las marcas tienen la función de diferenciar los contenidos informativos de los documentos, frente al uso que se hace en HTML, donde las marcas sirven para indicar cómo se deben visualizar los contenidos. Por otra parte, mientras que HTML nos indica qué marcas podemos utilizar cuando creamos un documento, XML no especifica un conjunto válido de marcas, sino que nos ofrece las reglas que nos permiten crear nuevos vocabularios o conjuntos de marcas aplicables para la codificación de distintos tipos de documentos.
- La distinción entre la estructura y la presentación de los documentos. En XML se establece una clara diferencia entre la estructura de un documento y su presentación. Las marcas de un documento XML no indican nada sobre cómo debe presentarse el documento. Para indicar cómo se debe presentar

un documento en pantalla o en papel, será necesario crear una hoja de estilo aparte, y asociarla posteriormente al documento.

▲2. Evolución continua del lenguaje

Desde su formulación inicial en 1997 el lenguaje XML ha dado lugar a un elevado número de iniciativas relacionadas con el intercambio y la codificación de contenidos y metadatos. XML se ha convertido en una opción unánimemente aceptada para facilitar su gestión y recuperación. La lista de iniciativas es amplia, y si bien no todas ellas han logrado el mismo nivel de aceptación, contamos con numerosos ejemplos de la aplicación exitosa de este lenguaje en los ámbitos académicos, empresariales e institucionales.

Durante el año 2004 hemos asistido a la publicación de nuevas propuestas e iniciativas relacionadas con el uso del lenguaje y hemos visto cómo algunas de sus especificaciones han continuado evolucionando hacia la madurez. Como ejemplos podemos citar la continua publicación de nuevas versiones de aplicaciones informáticas destinadas a trabajar con XML, el diseño de nuevos vocabularios o la adopción del lenguaje en distintos ámbitos de trabajo. Más concretamente, debe ponerse de relieve la publicación definitiva de la primera versión del vocabulario UBL (Universal Business Language) para el intercambio de documentos comerciales, y la elección del gobierno de Dinamarca de este vocabulario para sus proyectos de Administración electrónica; el desarrollo del vocabulario DITA (Darwin Information Typing Architecture) por OASIS para la codificación y reutilización de documentación técnica; o el lanzamiento de nuevas versiones de programas informáticos como DB XML de Sleepycat Software, Astoria 4.3 o la adquisición de Advent Publishing por parte de Arbortext.

Por otro lado, las líneas de trabajo que recibieron una mayor atención a lo largo del 2004 y que han logrado unos resultados más productivos son las siguientes:

Un interés creciente en las especificaciones orientadas a la integración de aplicaciones informáticas mediante el intercambio de datos a través de los llamados "servicios web" [2]. Este interés se ha materializado en la publicación de nuevas especificaciones que

facilitan la orquestación y coordinación de los interfaces entre aplicaciones (BPEL4WS, BPML, etc.) o el intercambio de datos XML en formato binario.

Un mayor respaldo por parte de los organismos de normalización internacional a las especificaciones destinadas a lograr la estandarización del comercio electrónico entre empresas. Prueba de ello es la publicación como norma ISO de las especificaciones ebXML [3], y la reciente actualización del estándar UDDI (Universal Description and Discovery). Estos dos estándares han vuelto a despertar un interés que se vio ralentizado en años anteriores por la reducción de las inversiones en tecnologías de la información y la desconfianza en los mercados relacionados con Internet.

▲3. El reconocimiento de la normalización

El año 2003 finalizó con una noticia curiosa: el 31 de diciembre, Tim Berners-Lee [4], "inventor" del World Wide Web, director del W3C y principal promotor del concepto de "web semántica", recibía el nombramiento de caballero " *Knight Comander* " de la Orden del Imperio Británico por parte de la reina Isabel de Inglaterra. En las declaraciones que hizo Tim Berners-Lee tras la comunicación de este nombramiento, señaló como merecedores del mismo a todas las personas que han contribuido al desarrollo descentralizado y abierto del Web [5], sustentado en tecnologías libres, y accesibles para toda la comunidad. Este nombramiento tiene carácter anecdótico, pero también puede interpretarse como una evidente señal de la trascendencia que el Web ha alcanzado en nuestros días no sólo como herramienta aplicable con fines académicos y comerciales, sino también como medio de comunicación en el que está llamada a sustentarse la sociedad futura.

El 2004 concluyó con la celebración del décimo aniversario de la fundación del W3C (World Wide Web Consortium) con un simposio celebrado el 1 de diciembre en Boston al que asistieron un amplio número de directivos de las principales empresas de tecnologías de la información [6]. Destacó, nuevamente, la presencia entre los patrocinadores del evento de empresas tecnológicas rivales, lo que demuestra el consenso alcanzado unánimemente en torno a los estándares publicados por el W3C. Entre las comunicaciones presentadas en este evento se destacó el impacto del Web y de sus estándares y tecnologías, entre ellas el lenguaje XML y su función en

la creación, gestión y difusión de información y contenidos. A modo de ejemplo, en la presentación de Denis Lacroix (Amadeus e-Travel) y Teri Richman (National Association of Convenience Stores) titulada *Impact on Science and Industry* se citaron los servicios web y la web semántica como las dos tecnologías clave de la futura Internet, ya que facilitan la interacción de aplicaciones informáticas a partir de una representación estandarizada de la información que hace posible su comprensión, recuperación y tratamiento. En los dos casos se trata de tecnologías basadas en el uso de XML. En otro apartado de su presentación titulado significativamente " *Cómo el W3C ha ayudado a nuestra industria* ", Teri Richman incluía una mención a los lenguajes HTML y XML como medios que han hecho posible el intercambio de datos a escala global..

Estas contribuciones son tan sólo una muestra del protagonismo alcanzado por el lenguaje XML desde su creación hace ya 8 años y del reconocimiento de su importancia en los procesos de creación, gestión, y sobre todo intercambio de contenidos y conocimiento.

Pero cuando nos referimos a la *gestión de contenidos* vemos que no siempre se asocia este término a la utilización del lenguaje XML. ¿Qué relación podemos establecer entre esta práctica y el formato? ¿A qué nos referimos cuando hablamos de gestión de contenidos?

▲4. Definición de gestión de contenidos

La elaboración de una definición de gestión de contenidos encontraríamos las mismas dificultades que encontramos al hablar de gestión documental o de gestión de conocimiento. Esta dificultad se debe, principalmente, al uso que de este término se viene haciendo en diferentes contextos. Por una parte, el mercado de aplicaciones informáticas ha acuñado su propia definición. Existe un conjunto de aplicaciones informáticas que presentan unas características comunes bajo este nombre. Esto ha hecho que actualmente, y por extensión, el término gestión de contenidos se haya equiparado al conjunto de funciones que forman parte integral de este tipo de programas. Por otra parte, la gestión de contenidos también puede equipararse a un planteamiento más amplio enfocado a la gestión global de los recursos de información de una institución o empresa mediante tecnologías web (Internet e Intranet). Este enfoque da a la tecnología un papel facilitador, y el mayor peso recae en los aspectos relacionados con la identificación de recursos de información internos

y externos, su valoración, gestión y tratamiento eficiente. Se trata de una propuesta que puede verse como una evolución de la gestión documental tradicional y, si nos remontamos más atrás en el tiempo, en iniciativas como el Information Resources Management (IRM) propuesto por Horton en la década de los 80.

Una tercera definición de gestión de contenidos procede del área de la gestión de publicaciones y documentos electrónicos. En este ámbito, se usa el término para referirse a la aplicación de una serie de técnicas y herramientas para la codificación, almacenamiento y distribución de publicaciones en formato digital. Es en este ámbito donde la utilización de los lenguajes de marcas " inicialmente SGML y posteriormente XML " ha sido una constante debida a su carácter abierto, la independencia de proveedores y de plataformas hardware/software específica, y la posibilidad de reutilizar el mismo contenido en múltiples productos y publicaciones [7]. En los tres casos se pueden encontrar puntos comunes. El más importante de ellos, la necesidad de utilizar tecnologías de la información y sistemas informáticos para el almacenamiento y distribución de información de naturaleza textual.

▲4.1. Origen de la gestión de contenidos

La gestión de contenidos tiene un origen pragmático. Surge como respuesta a las dificultades que suponía la gestión y el mantenimiento de sitios web y publicaciones en línea basadas en colecciones de documentos HTML. Los primeros sitios web, caracterizados por sus contenidos estáticos y compuestos en su mayor parte de páginas HTML y archivos de imagen, fueron dando paso a soluciones más sofisticadas en las que se adoptaron nuevas tecnologías que hacían posible la generación dinámica de contenidos. Inicialmente se usó para esto los programas o pasarelas CGI y posteriormente las páginas dinámicas escritas en lenguajes de script de servidor como ASP (Active Server Pages), JSP (Java Server Pages) o el popular PHP, que permitían la extracción de información de bases de datos relacionales.

La arquitectura de los sitios web se fue haciendo más compleja con la adopción de estas tecnologías. En paralelo se tenía que dar respuesta a la necesidad de mantener vivas las publicaciones y facilitar el acceso a información actualizada en tiempo real por equipos de trabajo multidisciplinares compuestos por autores, editores,

diseñadores gráficos y personal técnico. Este problema llevó a algunos proveedores de sistemas informáticos a diseñar aplicaciones que facilitasen la gestión de todos los componentes o recursos que forman parte de un sitio web. Entre estos componentes se incluían los contenidos propiamente dichos (pudiendo tratarse de materiales en distintos formatos: texto, imagen, video, audio, etc.) así como páginas dinámicas, hojas de estilo o archivos de script que implementan los aspectos visuales y dinámicos del sitio.

Las características de estas aplicaciones informáticas han conducido a una definición del término "gestión de contenidos" muy acoplada a su alcance funcional. Por ejemplo, en un glosario publicado en el sitio web de la empresa Interwoven "uno de los principales proveedores de este tipo de aplicaciones", se define gestión de contenidos como:

"la capacidad de gestionar y controlar el proceso de recoger, etiquetar, aprobar y distribuir contenido a un número variable de destinatarios (..)"

En este mismo documento se señala que:

"las funciones de la gestión de contenidos incluyen la agregación a partir de fuentes externas, el uso de plantillas para la entrada de datos y la visualización, control de versiones, workflow inteligente, etiquetado, clasificación, distribución y sindicación".

La gestión de contenidos ha tenido un importante éxito comercial. Han sido numerosas las organizaciones que han adoptado una aplicación informática de este tipo, y el mercado se ha abierto paso con relativa facilidad. Esto se debe a que en todo momento los proveedores han tratado de dar respuestas concretas a los problemas presentes en la práctica diaria de las organizaciones. Para identificar estos problemas, basta con plantearnos preguntas como las siguientes: ¿estamos seguros de que usuarios acceden realmente a los últimos contenidos o textos que se han actualizado? ¿Dependemos de la disponibilidad de una persona "normalmente el administrador del sitio web" para poner a disposición de los usuarios cierta información? ¿Nos exige un trabajo y tiempo excesivo publicar

una nueva información siguiendo las pautas y guías de estilo adoptadas por la organización? ¿En qué medida es posible hacer un seguimiento del estado en el que se encuentra la validación y aprobación de un artículo o de un texto determinado? ¿Puede un usuario publicar y editar contenidos con facilidad e independencia, sin necesidad de conocer los detalles del lenguaje HTML, hojas de estilo, o la distribución física del sitio web en carpetas, subcarpetas, etc.?

Todas estas preguntas resumen la problemática a la que se enfrenta cualquier organización que haya optado por usar las tecnologías web como canal de comunicación interno o externo, en cuanto su sitio adquiere un tamaño y una complejidad media. Se trata de unas dificultades que podríamos resumir en los siguientes puntos:

- Dificultades relacionadas con la necesidad de contar con conocimientos técnicos para poder publicar información y contribuir a los contenidos del sitio. Esto se traduce en una dependencia del personal técnico para mantener el sitio actualizado.
- Dificultad de mantener manualmente la integridad de la información, controlar versiones y cambios realizados en los contenidos y asegurar que sólo los contenidos aprobados llegan a estar disponibles para los usuarios, según una política de accesos y seguridad establecida.
- Incapacidad de programar la publicación de contenidos a ciertas horas y durante periodos de tiempo específicos.
- Dificultad para coordinar el trabajo de un equipo multidisciplinar, formado por perfiles heterogéneos (autores, editores, diseñadores, programadores, etc.).

El no disponer de una aplicación y de unos procedimientos de trabajo que faciliten la gestión de un sitio web desemboca irremisiblemente en consecuencias negativas, entre las que cabe destacar las siguientes:

- El administrador del sitio web y el personal técnico acaban viéndose desbordados ante el número creciente de peticiones y tareas relacionadas con la publicación de contenidos y la administración del sitio.

- Los contenidos no están disponibles cuando se espera que estén, se ocasionan retrasos que afectan a la validez e integridad de la información.
- Resulta muy difícil mantener actualizadas las herramientas de navegación a medida que se añaden y eliminan páginas al sitio, al tratarse de un proceso que requiere la actualización manual de múltiples páginas.
- Resulta muy difícil, si no imposible, reutilizar contenidos (por ejemplo, publicar el mismo contenido para los usuarios de un sitio web externo y para los de una Intranet con distintas plantillas conllevaría necesariamente una duplicidad de tareas y contenidos).

Los sistemas de gestión de contenidos comerciales "y las alternativas código abierto u *open source* " pretenden ofrecer respuesta a estos problemas, y han definido la disciplina de la gestión de contenidos tal y como la conocemos a día de hoy. De esta forma, podemos afirmar que una organización gestiona correctamente sus contenidos si da una respuesta de forma ordenada a los problemas anteriormente citados, y si implementa las herramientas y procedimientos necesarios para dar respuesta a los problemas anteriores.

▲4.2. Descripción funcional de un sistema de gestión de contenidos

Los sistemas de gestión de contenidos ofrecen las siguientes ventajas a las organizaciones que los implementan.

- Democratizar la creación y edición de contenidos, al no ser necesario conocer lenguajes como HTML, CSS, o similares, ni aplicar costosas conversiones para publicar información.
- Definir y aplicar ciclos de validación y aprobación sobre los contenidos.
- Reducir el tiempo que transcurre desde la creación del contenido a su publicación y distribución.
- Descargar al personal técnico y al administrador del sitio web de tareas administrativas, rutinarias, que no aportan valor añadido (copia manual de archivos entre entornos de edición, pruebas, y publicación, conversiones de formatos, etc.).

La aplicación de la gestión de contenidos integrará las herramientas necesarias para que las personas encargadas de mantener y gestionar un sitio web (normalmente departamentos de marketing, comunicación, etc.) puedan actualizar fácilmente los contenidos del sitio sin necesidad de conocer los detalles de la codificación HTML o la ubicación física de las páginas en el servidor web. Las funciones que suelen incorporar estas aplicaciones incluyen:

- Mantenimiento de la estructura física y lógica del sitio
- Creación de nuevos contenidos y edición de contenidos existentes mediante plantillas, normalmente a través de un navegador web.
- Mantenimiento automático de la navegación del sitio y de los hiperenlaces entre páginas.
- Aprobación, revisión y validación de los contenidos hasta que estos se hacen públicos en el sitio web.
- Periodos de vigencia de los contenidos.
- Control de cambios y revisiones.
- Compartir contenido entre distintas páginas (páginas conectadas).

▲4.2.1. Mantenimiento de la estructura lógica del sitio

En un sitio web podemos diferenciar dos estructuras. Por una parte, la estructura física de las páginas HTML, PHP, etc., que lo forman. Esta estructura está determinada por la forma en la que se almacenan las páginas físicamente en el servidor web (carpetas, subcarpetas, tablas en bases de datos, ficheros, etc.). Por otra parte contamos con una estructura lógica que establece la forma en la que se presentarán las páginas a los visitantes del sitio web (agrupadas en secciones, sub-secciones, etc.). Si bien es habitual que exista una correspondencia entre la estructura física y la estructura lógica del sitio, ésta correspondencia no tiene por qué producirse en el cien por cien de los casos.

En un sistema de gestión de contenidos, la estructura lógica del sitio se gestiona mediante una jerarquía de "agrupaciones" o "secciones". En cada sección se reúnen las páginas que contienen la información relevante para esa sección. La aplicación de gestión de contenidos se encargará de mantener la estructura física de forma transparente a los usuarios y a los administradores, que pueden "despreocuparse" de los detalles referentes al almacenamiento físico de los contenidos.

En un sistema de gestión de contenidos, la estructura de secciones y subsecciones de un sitio se suele guardar en una base de datos, junto con el resto de la información (textos, contenidos, etc.) De esta forma en la base de datos utilizada por el gestor de contenidos se guardan las secciones y subsecciones que forman parte de nuestro sitio, y la relación entre páginas y secciones (esto es, qué páginas pertenecen a cada sección o grupo).

▲4.2.2. Edición de contenidos mediante plantillas

Para evitar a los usuarios la complejidad del lenguaje HTML, la creación de nuevas páginas y la modificación de páginas existentes se realizan mediante el uso de plantillas. Las plantillas son formularios HTML donde los autores puedan escribir textos y añadir imágenes, hiperenlaces, y cualquier otro componente característico de las páginas web.

En una plantilla se distinguen zonas editables y zonas no editables. Los autores únicamente pueden modificar el contenido de las zonas editables, con lo que se garantiza la integridad de los elementos fijos de la página (cabeceras, menús, imágenes, etc.)

Cuando un usuario crea contenidos a través de una plantilla, lo que hace en realidad es añadir un texto a un repositorio de contenidos, donde el texto se guarda de forma independiente a las partes fijas del texto. Puede usarse como repositorio de la información una base de datos relacional o un sistema de ficheros (las distintas aplicaciones pueden optar por un sistema u otro de almacenamiento). El sistema también guarda información relativa a la plantilla con la que se debe visualizar ese contenido.

Una herramienta de gestión de contenidos debe permitirnos crear plantillas para los distintos tipos de páginas que vayamos a utilizar en nuestro sitio. Cada plantilla puede tener tantos campos como sean necesarios. Al tratarse de páginas de naturaleza textual, normalmente se tratará de campos de longitud variable. La edición puede llevarse a cabo desde un navegador estándar al que se incorporan utilidades visuales que permiten al usuario marcar y

formatear textos para que aparezcan en negrita, cursiva, mediante listados, etc.

El concepto de plantilla permite garantizar la consistencia en la presentación de contenidos, ya que todas las páginas basadas en la misma plantilla comparten una presentación idéntica. Por otra parte, las plantillas facilitan el mantenimiento de los aspectos visuales del sitio, ya que si se necesita realizar una modificación en la presentación de una colección de páginas, únicamente será necesario hacer el cambio en la plantilla, y el cambio se hará efectivo inmediatamente en todas las páginas que se basan en ella (aunque se hayan creado con anterioridad).

▲4.2.3. Mantenimiento automático de la estructura de navegación

Uno de los problemas más frecuentes en la gestión de un sitio web es el mantenimiento de los hiperenlaces que determinan la navegación del sitio. Por ejemplo, si se añade una nueva página y se quiere incluir un hiperenlace en un menú existente en varias páginas desde el que se pueda acceder a la nueva página, sería necesario editar individualmente todas y cada una de las páginas que contienen el menú para incluir el nuevo enlace manualmente. De la misma forma, si se elimina del sitio una página para la que existen enlaces desde otras páginas, sería necesario revisar estas páginas y eliminar manualmente los hiperenlaces que tienen como destino la página que hemos borrado.

Si bien esto no resulta problemático en sitios web con pocas páginas, el problema se acentúa en sitios con un número elevado de páginas y que se actualizan frecuentemente.

Con una aplicación de gestión de contenidos, en el momento en el que se definen las plantillas para el sitio se puede indicar qué hiperenlaces o menús se deben incluir en cada una de ellas. Los hiperenlaces se generan mediante consultas al repositorio de páginas que gestiona el gestor de contenidos, por lo que las opciones disponibles en menús, índices, etc., siempre se generan dinámicamente y reflejan el estado real del sitio en cada momento.

▲4.2.4. Aprobar, revisar y validar contenidos

Otra función que suelen ofrecer las aplicaciones de gestión de contenidos son los ciclos de aprobación y validación de contenidos. Las páginas pueden tener un ciclo de aprobación asociado, de forma que siempre que un autor crea una nueva página o modifica una existente se solicitará la aprobación de los cambios por parte de un usuario autorizado. Este usuario tendrá la función de *editor*, si usamos la terminología tradicional en los medios impresos, y su función será revisar, aprobar o desestimar los cambios.

En el caso de páginas y contenidos sujetos a ciclos de aprobación, los cambios no son visibles en el sitio web hasta que el editor los aprueba. De esta forma, se asegura que toda la información que se publica sigue un proceso de control de calidad que impide la publicación de información errónea, desactualizada o malintencionada.

Los usuarios autorizados a aprobar páginas deben disponer de alguna herramienta que les permita acceder a las páginas pendientes de aprobación, ver su contenido, aceptarlas o rechazarlas. También es aconsejable que la aplicación de gestión de contenidos incorpore un sistema de envío de notificaciones (normalmente el correo electrónico) para comunicar a los editores la necesidad de revisar los cambios.

▲4.2.5. Periodos de vigencia de las páginas

Otro problema que deben afrontar los encargados de mantener el contenido del sitio web se deriva del hecho de que ciertos contenidos tienen un periodo de validez o vigencia, pasado el cual es necesario eliminarlos del sitio. Pongamos como ejemplo una noticia que debe estar disponible únicamente durante dos o tres días, o una publicación semanal que todos los lunes a primera hora debe ser actualizada sustituyendo los artículos disponibles por otros nuevos.

Sin una aplicación de gestión de contenidos, el mantenimiento de estos sitios resulta sumamente problemático. Por ejemplo, en casos

en los que los contenidos deben reemplazarse en bloque, a un mismo tiempo, (por ejemplo, los lunes a las 7 de la mañana), sería necesario contar con la presencia física de una persona que se encargue de hacer el reemplazo de archivos en el servidor web, en el momento preciso.

El problema se acentúa con los sitios web que contienen información financiera, datos sobre precios de productos, etc., en los cuales es un requerimiento crítico asegurar que la información está disponible en el instante correcto "ni antes ni después.

Con una aplicación de gestión de contenidos podremos ser capaces de indicar, para cada página individual, grupo de páginas, o incluso para toda una sección, cual va a ser su periodo de vigencia, de forma que estos contenidos sólo están disponibles en el sitio web público durante el periodo que se indique, incluso si la página ha sido creada y aprobada con anterioridad.

En este apartado tenemos que tener en cuenta que la visibilidad no sólo debe aplicarse a la página en sí misma, sino también a todas las referencias e hiperenlaces que apunten a ella. Estos sólo deben permanecer activos durante el periodo de vigencia de la página: todos los enlaces que la tengan como destino, deben dejar de ser visibles en el sitio web en cuanto la página alcance el fin de su periodo de validez. Para lograr esto, ¿en qué se basa la aplicación de gestión de contenidos? La fecha de inicio y de fin del periodo de validez de cada página formará parte de las propiedades o metadatos administrativos de la página. Esta información se guardará en el repositorio que gestiona la aplicación, de forma que una simple consulta al mismo bastará para saber qué páginas, hiperenlaces y referencias deben estar visibles en cada momento y han sido aprobadas.

▲4.2.6. Control de cambios y revisiones

Otra funcionalidad ofrecida por las aplicaciones de gestión de contenidos es el control de cambios y versiones de las páginas. El control de cambios permite mantener un histórico del contenido que tenía una página. Así, cuando se modifica una página, en lugar de

sobrescribir su contenido y perderlo, el sistema creará una nueva versión y mantendrá el estado anterior de la página y su información.

El control de cambios nos permitirá recuperar versiones anteriores de las páginas. En un entorno no gestionado con un sistema de gestión de contenidos, nos veríamos obligados a perder la información anterior al cambio si sobrescribimos la página o a mantener un costoso sistema manual de copia de archivos que resultaría complicado y susceptible de tener errores.

▲4.2.7. Páginas conectadas: compartir contenidos entre páginas

En determinadas ocasiones, se impone la necesidad de compartir el mismo contenido en distintas páginas del mismo sitio web o de sitios web diferentes (por ejemplo, entre un sitio internet y un sitio Intranet, o entre los usuarios registrados y los no registrados en un sitio público). Sin un sistema de gestión de contenidos, la necesidad de compartir contenidos entre distintas páginas nos obligaría a duplicar el mismo contenido en distintos archivos. Esto acarrea el problema del mantenimiento y exige un mayor control, ya que si ese contenido tiene que modificarse, es necesario hacer los cambios en más de una página. También es necesario saber con exactitud qué páginas comparten contenidos, para que la actualización del sitio o sitios sea posible y consistente.

Con una aplicación de gestión de contenidos podremos compartir contenidos entre páginas fácilmente, mediante una funcionalidad que se suele llamar *páginas conectadas*. Son páginas independientes, normalmente basadas en distintas plantillas, que comparten el mismo contenido. De esta forma, cuando se modifica el contenido de una de las páginas conectadas, se actualizará automáticamente el contenido de todas las demás. Esto es posible porque el contenido de los campos se almacena de forma independiente a su presentación en el repositorio de la aplicación de gestión de contenidos, lo que nos permite extraerlo y visualizarlo de una u otra forma, aplicándole distintas plantillas.

▲4.3. Descripción técnica de un sistema de gestión de contenidos

Las funciones descritas en los apartados anteriores precisan de una infraestructura técnica y de la utilización de distintas tecnologías. La gestión de contenidos hace uso de las siguientes tecnologías: páginas dinámicas (PHP, JSP, ASP, etc.); bases de datos y repositorios; metadatos; gestión de flujos de trabajo y mensajería; e integración entre aplicaciones.

▲4.3.1. Páginas dinámicas

Las páginas dinámicas "también denominadas "scripts de servidor" constituyen la base de los sitios web actuales. Se trata de archivos similares a las páginas HTML, en los que se intercalan "junto a las etiquetas y marcas propias del lenguaje HTML" órdenes o comandos escritos en un lenguaje de programación. Cuando el servidor web recibe una petición de una página dinámica, en lugar de enviarla directamente al navegador cliente del que proviene la solicitud "tal y como sucede en el caso de las páginas HTML" se interpretan o ejecutan las órdenes o comandos para a continuación enviar al navegador web el resultado obtenido por su ejecución, que será código HTML.

Esta tecnología ha permitido dotar a los sitios web de un gran dinamismo y llevar a Internet e Intranet aplicaciones informáticas de la más diversa índole (comercio electrónico, gestión administrativa, etc.) Su principal aplicación es la consulta y extracción de información desde bases de datos. En el caso de las aplicaciones para la gestión de contenidos, hemos señalado cómo se establece una clara diferencia entre el contenido propiamente dicho y las plantillas en las que éstos se visualizan y editan. La tecnología de las páginas dinámicas hace posible esta diferenciación, ya que a partir de un mismo contenido pueden ensamblarse distintas presentaciones mediante esta tecnología.

▲4.3.2. Bases de datos y repositorios

Una aplicación de gestión de contenidos necesita almacenar numerosa información. A los textos y contenidos propiamente dichos, tenemos que añadir los metadatos y propiedades para cada una de sus páginas o secciones. Dentro de éstos incluiríamos los metadatos administrativos (relativos a la vigencia de la página, su creador,

estado de aprobación, etc.) y los necesarios para facilitar su posterior recuperación (palabras clave, códigos de clasificación, etc.). A parte de esto, la aplicación de gestión de contenidos debe conocer la estructura de los sitios que gestiona, las secciones dentro de cada sitio, las páginas que se incluyen en cada sección, los ciclos de aprobación asociados a cada página, los estados posibles, y la información relativa a qué usuarios están autorizados y son responsables de validar cada tipo de contenido.

Por lo tanto, una aplicación de gestión de contenidos precisa de un repositorio de información y contenido con el que interactúan los distintos usuarios del sistema (autores, editores, programadores, diseñadores y consumidores y usuarios de la información) a través de las distintas herramientas e interfaces de usuario que ofrezca la aplicación. Este repositorio puede adoptar distintas formas: una base de datos relacional, un sistema de ficheros, un repositorio de datos XML, o la combinación de ellos.

Un aspecto que debe considerarse en lo relativo a la gestión del repositorio de información, es la necesidad de almacenar también todas las imágenes, archivos JavaScript, hojas de estilo CSS, etc., que forman parte de la estructura de un sitio web. Se utiliza el término "recursos" para referirnos en general a todos estos elementos. Al considerar el almacenamiento y la gestión del repositorio de elementos que forman el sitio web, debe tenerse en cuenta este requerimiento, ya que todos los elementos deben estar perfectamente identificados, controlados y administrados desde un punto de acceso común.

▲4.3.3. Metadatos

Por metadatos se entiende la información que se asocia a los distintos objetos que forman parte del repositorio del sistema gestor de contenidos, y que facilitan la gestión administrativa y la recuperación. En el apartado anterior hemos citado distintos ejemplos de metadatos aplicables a nivel de página, sección, o a cualquier otro tipo de recurso.

La utilización de los metadatos depende de las tecnologías anteriores, en el sentido de que debemos ser capaces de almacenarlos en un

entorno que haga posible su recuperación, y de utilizarlos y acceder a ellos mediante páginas dinámicas.

Entre las funciones que cumplen los metadatos destaca el permitir la personalización del sitio: la comparación de los metadatos asociados a una página con las preferencias establecidas para los usuarios registrados hace posibles las estrategias de personalización y distribución selectiva de la información, adecuando los contenidos que se muestran en el sitio web o que se distribuyen mediante notificaciones por correo electrónico a las necesidades e intereses particulares de cada usuario.

▲4.3.4. Gestión de flujos de trabajo

La gestión de flujos de trabajo "o si preferimos usar el término inglés, workflow" permite organizar secuencias de actividades en torno a una o más páginas. Este elemento tiene una parte funcional (al tratarse de una de las características deseables en toda aplicación de gestión de contenidos), pero también un componente técnico.

Para disponer de esta función, la aplicación necesita disponer de un "motor de workflow", capaz de distribuir y asignar tareas de forma dinámica a partir del tipo de acción que se tenga que completar, de las características de los contenidos, y de la carga de trabajo asignada en cada momento a los distintos actores que interactúan con el sistema.

Pongamos un ejemplo: en un periódico digital, es probable que cada área o sección cuente con editores especializados. De esta forma, cuando un autor finaliza la redacción de una noticia, la revisión podría ser asignada al editor más apropiado dependiendo de la temática del artículo; además de esto, la aplicación podría considerar las asignaciones actuales de cada editor, y dirigir la solicitud a aquel que cuente con un menor número de solicitudes pendientes de revisar.

Obviamente, el uso que se hace en una aplicación de gestión de contenidos de la automatización de flujos de trabajo no exige la complejidad que se precisa en otros entornos. Sin embargo, sí se

trata de una pieza necesaria para asegurar que todas las informaciones han sido debidamente revisadas y validadas antes de hacerse públicas.

El subsistema de gestión de flujos de trabajo puede incluir algunas características más avanzadas, como por ejemplo:

- La definición de distintos niveles de revisión y autorización dependiendo del tipo de contenido.
- La integración con un sistema de correo electrónico para enviar notificaciones sobre las revisiones pendientes y el resultado de cada revisión.
- La definición del orden de ejecución de las tareas que se deben completar para un contenido específico, la función de las personas que deben realizarla, la asignación a personas específicas y los estados se alcanzan a medida que éstas se completan.
- La definición de los plazos disponibles para completar cada una de las tareas o revisiones, y las acciones que deben ejecutarse en caso de alcanzarse dicho plazo sin haberse concluido la tarea en cuestión.
- La aplicación de distintos niveles en la asignación de tareas, distinguiendo entre flujos secuenciales (en los que todos los usuarios deben completar una tras otra las tareas que les han sido asignadas) y paralelos (en los que una tarea se asigna simultáneamente a varios usuarios, de forma que ésta se considera finalizada cuando uno de ellos la realiza o cuando la realizan parte de ellos).
- La provisión de información sobre la definición de los ciclos de aprobación y el estado de aprobación de cada elemento gestionado por el sistema.

▲4.3.5. Integración de contenidos y aplicaciones

Otras tecnologías que forma parte integrante de las aplicaciones de gestión de contenidos son aquellas que facilitan la integración de datos procedentes de otras aplicaciones o de fuentes externas.

A medida que la Web ha ido creciendo, resulta sumamente complejo mantener y generar internamente contenidos atractivos para los

usuarios. Esto ha hecho que las organizaciones recurran a proveedores de información externos que pueden facilitar contenidos en forma de artículos, resúmenes, referencias a sitios externos, etc., en un formato que permita su integración en el sitio Web. Esta actividad ha llegado a generar un negocio en sí mismo: la sindicación de contenidos. Con este término se suele hacer referencia a la actividad que consiste en agregar noticias o titulares procedentes de distintos sitios web y categorizarlos mediante unos códigos de clasificación, descriptores o metadatos que hagan posible su redistribución a una serie de empresas consumidoras.

Los procesos de agregación añaden un importante valor para las empresas usuarias de estos servicios, al realizar un filtrado profesional en un volumen de datos y contenidos creciente, y al distribuir únicamente aquellos contenidos relevantes para la empresa cliente.

La sindicación de contenidos requiere del uso de dos componentes: a) formatos normalizados para la transferencia de información a través de la red, y b) un mecanismo que permita a la empresa cliente descargar los contenidos agregados que son de su interés de forma desatendida, en unos intervalos de tiempo preestablecidos.

El primer componente está estrechamente relacionado con el uso del lenguaje XML y de distintos vocabularios que se han diseñado para este fin, siendo los más populares RSS (Resource Site Summary) con sus distintas versiones, y ATOM. En los dos casos se trata de vocabularios XML que definen la estructura que deben cumplir los documentos con los que se transmiten titulares de noticias o contenidos para que éstos puedan ser procesados fácilmente por la empresa consumidora, e integrados en los sitios web y en las publicaciones oportunas.

El segundo componente, el mecanismo de descarga de contenidos agregados, está relacionado con la tecnología de los servicios web, a los que nos referiremos posteriormente, y que nos permiten acceder a funciones disponibles en servidores web remotos mediante una URL (Uniform Resource Locator) que nos devolverá un documento XML con los datos resultado de la ejecución de dicha función por parte del equipo remoto.

La forma de invocar la ejecución de un servicio web será siempre a través de una URL, siendo posible adjuntar a la solicitud una serie de parámetros que el servidor web remoto utilizará en la ejecución del servicio. Así, en nuestro ejemplo de integración de contenidos, la llamada al servicio web del proveedor remoto podría recibir como parámetros un código de clasificación relativo a la temática de los artículos que queremos recuperar, y la fecha a partir de la cual queremos filtrar para obtener sólo los artículos añadidos a partir de un momento dado.

▲5. La función de XML en la gestión de contenidos

Una vez hemos definido las funciones que caracterizan una aplicación de gestión de contenidos, es necesario dar respuesta a dos cuestiones: ¿qué función cumple el lenguaje XML en este tipo de sistemas? y ¿cuáles son las ventajas que podemos obtener de su aplicación?

En los apartados anteriores hemos señalado algunas características de las aplicaciones de gestión de contenidos que nos recuerdan los principales aspectos del lenguaje XML. Así, hemos señalado entre las ventajas de la gestión de contenidos la clara diferenciación entre los contenidos propiamente dichos y las plantillas que se utilizan para su edición y visualización, de forma que es posible reutilizar un mismo contenido en distintos contextos. Como recordaremos, esta es una de las principales ventajas del formato XML: la distinción entre contenidos y presentación.

Otros de los aspectos donde se manifiesta el potencial del lenguaje XML aplicado a la gestión de contenidos es la gestión del repositorio de metadatos asociados a las distintas páginas. Son numerosos los sistemas de metadatos propuestos en los últimos años para la recuperación de información que han optado por XML como formato de codificación. En este grupo tendríamos Dublin Core, MARCXML, MODS, etc. Pero donde resulta más evidente la aplicación de XML es en los procesos de intercambio, transferencia e integración de información. De hecho, si bien sería posible obviar la utilización de XML en el almacenamiento de contenidos y en la gestión de metadatos, la integración y sindicación de contenidos resultaría prácticamente inabordable sin la utilización de este formato.

En los siguientes apartados se incluye una breve descripción de cada una de estas áreas de aplicación.

▲5.1. XML como base para el almacenamiento de contenidos

Los contenidos gestionados y publicados a través del sitio web pueden almacenarse de distinta forma, aunque para garantizar su posterior reutilización y recuperación debemos considerar las ventajas que ofrece XML como lenguaje de almacenamiento frente a otras alternativas como HTML.

Debemos citar la importancia que en los últimos años están alcanzando las llamadas bases de datos XML nativas. Con este término nos referimos a bases de datos que almacenan y gestionan una colección de documentos XML sin realizar ningún tipo de transformación previa. En este modelo, el documento XML es la principal unidad de almacenamiento de la información. Contrasta este tipo de bases de datos con las bases de datos relacionales, en las que la información se guarda en una colección de tablas formadas por campos y registros, con una estructura sumamente rígida que debe ser predefinida con anterioridad, y que en muchas ocasiones no da cabida a la flexibilidad que exige el almacenamiento de documentos de naturaleza textual.

Entre los principales exponentes de las bases de datos XML nativas podemos destacar sistemas comerciales como Tamino, de la empresa alemana Software AG, Textil, o el sistema código abierto DBXML, que se puede obtener gratuitamente.

No obstante, como señalamos inicialmente, una aplicación de gestión de contenidos no tiene que estar vinculada, forzosamente, a una base de datos XML nativa. Aunque se opte por almacenar los textos en formato XML, cabe la opción de guardarlos en un sistema de bases de datos relacional utilizando los tipos de datos que permiten guardar y gestionar campos de texto de longitud variable.

Para facilitar la edición de contenidos en XML, los principales proveedores de herramientas de edición han publicado utilidades que permiten la interacción entre éstas y el repositorio de datos de la aplicación de gestión de contenidos. Encontramos ejemplos de esta integración en las propuestas de las empresas Altova, Blast Radius, XYEnterprise o Stylus.

▲5.2. XML como modelo para la representación de metadatos

Algunos de los sistemas de metadatos que se han publicado en los últimos años han optado por el lenguaje XML como principal mecanismo de representación y codificación de los mismos. Las ventajas de usar XML frente a otras alternativas se encuentran en su orientación hacia Internet, la facilidad de su intercambio y posterior procesamiento utilizando una única sintaxis común, y la opción de combinar e intercalar los metadatos dentro del texto completo de los documentos. Así, sería posible codificar un texto usando un conjunto de etiquetas o marcas relevantes para ese tipo de información, y agregarle metadatos descriptivos o administrativos como parte de la cabecera de dicho documento, codificados en cualquier otro vocabulario siempre que éste también esté basado en la sintaxis XML.

Si se opta por usar XML para la codificación de los metadatos debe considerarse la necesidad de disponer de un sistema de indexación y recuperación que permita discriminar documentos a partir del contenido de elementos o atributos específicos. Es decir, el sistema de búsqueda no sólo debe permitir la búsqueda en texto completo (esto es, poder recuperar el documento si contiene una combinación de palabras en particular), sino también si dichas palabras aparecen dentro de un elemento específico o en alguno de sus elementos descendientes. Sólo de esta forma se pueden obtener las ventajas de la organización jerárquica de la información que caracteriza al lenguaje XML.

En este sentido, las bases de datos XML nativas incorporan este tipo de herramientas y utilidades de búsqueda. Las aplicaciones de gestión de contenidos que utilizan otro sistema de repositorio (normalmente una base de datos relacional), presentan aquí algunas

limitaciones, ya que esta tecnología no suele incorporar herramientas de indexación adaptadas a las características del formato.

▲5.3. XML como medio de intercambio e integración de contenidos

Señalábamos con anterioridad la función primordial de XML como facilitador del intercambio y agregación de contenidos. Decíamos que es en este área donde mayor y más evidente se hace la capacidad y el potencial del formato en las actividades relacionadas con la gestión de contenidos.

Pues bien, debemos recordar que XML no se trata, únicamente, de un formato para codificar textos y documentos, sino de una familia de especificaciones que establece la forma en la que se pueden procesar y presentar dichos textos. Especificaciones como XSLT, DOM o XPath hacen posible procesar documentos XML basados en distintos vocabularios mediante distintos lenguajes de programación (Visual Basic, Java, etc.), utilizando un modelo común, estándar y claramente documentado.

La posibilidad de obtener a través de la red documentos XML y procesarlos con facilidad para cualquier fin (por ejemplo, para integrarlos en un repositorio o base de datos, o para visualizarlos como parte de nuestro sitio web), nos ofrecen una flexibilidad extrema y abre las puertas a cualquier tipo de integración.

▲6. Conclusiones

En los anteriores apartados se han descrito las principales características de la gestión de contenidos, y hemos visto como este término está estrechamente vinculado a un tipo de aplicaciones informáticas que ha surgido en los últimos años como respuesta a la problemática derivada de la gestión de sitios web.

En relación al lenguaje XML, se han señalado el potencial que ofrece la aplicación de este lenguaje en los procesos de gestión de contenidos, sea como formato para el almacenamiento de los

contenidos y metadatos, o como herramienta para la integración y el intercambio de contenidos procedentes de aplicaciones o de organizaciones externas.

Estas son las principales áreas de aplicación del formato en la gestión de contenidos, y podríamos encontrar distintos escenarios de su utilización; dependiendo del escenario, la importancia dada a una o a otra será mayor o menor.

De esta forma, si pensamos en aplicaciones orientadas al comercio electrónico entre empresas, donde el intercambio de documentos comerciales estructurados resulta clave, el mayor protagonismo recaerá en el uso de XML como formato de intercambio. Sin embargo, si pensamos en aplicaciones orientadas a la publicación de información, se dará una mayor relevancia a la aplicación del lenguaje como medio de almacenamiento de la información y como formato para la creación y edición de los documentos.

De hecho, una de las diferencias que suele hacerse al tratar de la aplicación de XML en la gestión de documentos y contenidos, es entre la gestión de documentos persistentes frente a los documentos transaccionales.

Con el primer término se hace referencia a los documentos que se crean en formato XML, usando herramientas de edición, y que se almacenan en ese formato, sin hacer ningún tipo de transformación; se trata de documentos dirigidos a personas (editores, lectores, autores, etc.), que se almacenan de forma permanente. Por documentos transaccionales se entiende cualquier documento que se genera con el único fin de intercambiar una información a través de Internet, en formato XML, pero cuyo ciclo de vida finaliza en cuanto el documento es recibido y procesado por la organización receptora. En el caso de los documentos transaccionales se trata de documentos volátiles, que tienen como única finalidad facilitar el intercambio de información y que normalmente se crean a partir de la información disponible en cualquier otro sistema o base de datos. Los documentos transaccionales son los documentos que hacen posible las transacciones propias del comercio electrónico entre empresas, la función que ha correspondido tradicionalmente al modelo EDI

(Electronic Data Interchange) y a su familia de estándares X12, EDIFACT, etc.

Cabría la opción de ampliar el alcance de la gestión de contenidos para referirnos también a este tipo de información. Como hemos visto, la amalgama de tecnologías que hacen posible la gestión de contenidos, y la posibilidad de optar por distintas alternativas para su almacenamiento y gestión, hacen difícil dar una única definición, y siempre quedan aspectos abiertos que nos permiten ampliar dicho alcance para incluir o excluir cuestiones concretas.

En cualquier caso, se hace manifiesto el potencial del lenguaje en estos distintos escenarios, como demuestra la adopción que del mismo han realizado los principales proveedores de sistemas informáticos que, a la larga, son quienes han dotado a esta disciplina de un significado.

▲7. Bibliografía

Boiko, Bob (2001) <i>The Content Management Bible</i> . New York: Wiley. ISBN 076454862X
Bussler, Christoph (2003). <i>B2B Integration: concepts and architecture</i> . Berlin [etc.]: Springer. ISBN 3-540-43487-9.
Donovan, Truly (1997) <i>Industrial-Strength SGML: an Introduction to Enterprise Publishing</i> . New Jersey: Prentice Hall PTR. ISBN 0-13-216243-1.
English, Bill. (2003) <i>Microsoft Content Management Server 2002: A Complete Guide</i> . Boston [etc.]: Addison-Wesley. ISBN 0321194446.
Hammersley, Ben (2003). <i>Content Syndication with RSS</i> . Beijing [etc.]: O'Reilly. ISBN 0596003838.
Marcos, Mari Carmen; Baiget, Tomàs (2003). "Integración y personalización de contenidos: Factiva". En: <i>El Profesional de la Información</i> , vol. 12, no. 1, p. 26-34.
Meloni, Julie (2004). <i>Plone Content Management Essentials</i> . Indiana: SAMS. ISBN 1-8000-382-3419.
Nakano, Russell. (2002) <i>Web Content Management: a Collaborative Approach</i> . Boston [etc.]: Addison-Wesley. ISBN 0-201-

65782-1.
Rockley, Ann. (2002) Managing Enterprise Content: A Unified Content Strategy. New York: New Ridder Press. ISBN 0735713065.
XML Data Management (2003). Chaudhri, Akmal B.; Rashid, Awais I; Zicari, Roberto (eds.) Boston [etc.] : Addison-Wesley. ISBN 0-201-84452-4.

▲8. Notas

[1] SGML (Standard Generalized Markup Language) es el antecesor de los lenguajes de marcas. Su origen se encuentra en el trabajo desarrollado por IBM en la década de los 60 bajo la dirección de Charles Goldfarb para la codificación de documentos legales. XML hereda sus principales características de este lenguaje. [[volver](#)]

[2] Se pueden definir como una tecnología basada en estándares públicos que facilita la interacción entre aplicaciones informáticas distribuidas en Internet. Los servicios web permiten que una aplicación cliente invoque la ejecución de una operación en una aplicación remota (que actuaría como aplicación servidora. Los servicios web se invocan mediante el envío de mensajes codificados en XML, usando la sintaxis que se define en la especificación SOAP (Simple Object Access Protocol), también un estándar del W3C. [[volver](#)]

[3] Corresponde a la familia de normas ISO 15000 Electronic Business eXtensible Markup Language, que pretenden acercar las ventajas del comercio electrónico Internet entre empresas a empresas de cualquier sector y tamaño. Estas normas establecen un marco para la definición de procesos, documentos comerciales, registros de empresas, etc. En el caso de UDDI, se establece un modelo de directorio de empresas y organizaciones web basado en el intercambio de información XML. [[volver](#)]

[4] Tim Berners-Lee es el inventor del World Wide Web, proyecto que desarrolló como parte de su trabajo en el CERN en Ginebra, en torno a 1989. El precedente de este invento fue un sistema informático llamado Enquire. En 1990 completó el primer servidor y cliente web,

capaces de comunicarse mediante el protocolo httpd. El lenguaje HTML también se debe a su trabajo. [[volver](#)]

[5] Puede obtenerse más información en <http://www.cabinet-office.gov.uk/> [[volver](#)]

[6] El programa y el detalle de las comunicaciones se puede consultar en <http://www.w3.org/2004/09/W3C10.html> [[volver](#)]

[7] Esta práctica suele conocerse con el término anglosajón "single sourcing", y se refiere a los procesos que permiten obtener distintas publicaciones o productos de información a partir de un único repositorio de contenidos mediante procedimientos automatizados. [[volver](#)]

THE GILBANE REPORT™

Content, Computing, and Commerce – Technology & Trends

THE CLASSIFICATION & EVALUATION OF CONTENT MANAGEMENT SYSTEMS

The continued dramatic growth in content management systems (CMSs) and technologies – there are hundreds of CMSs, including dozens of open source tools – has defied the usual rules of business software markets. The number of new product launches by old and new companies somehow still manages to keep ahead of the ongoing consolidation. This is very healthy. However, even a full-time market analyst paid to be a content management expert is not going to be able to keep up with all the products and features, especially since managing content involves technologies that go well beyond a CMS (See last month’s report on integration, infrastructure, and content services technology.) IT strategists and project managers who have other responsibilities don’t stand a chance of keeping up-to-date.

Fortunately, there is an industry effort gathering steam to provide an open and free list of CMS products and features. This public domain classification will be based on an XML schema (CMSML) so that anyone can use the information. Although the project is in its early stages, there is already one public domain tool for mining this CMS data, which we will make available at gilbane.com in early June. This is a huge and complex task and promises to be ongoing. Even given a predictable perpetual state of flux, this project should make information collection about CMSs and their features a lot easier for everyone. This month Bob Doyle, editor of the CMSML project, describes the project and how to participate, and shares the draft feature list.

CONTENTS

The Classification & Evaluation of Content Management Systems ..2	
Industry News	14
Back Issues.....	31
Subscription Form & Calendar	32

The Gilbane Report

▲1

March, 2003

Vol. 11, No. 2
March, 2003

www.gilbane.com

Published by:
Bluebill Advisors, Inc.
763 Massachusetts Ave.
Cambridge, MA 02139
USA
(617) 497.9443
Fax (617) 497.5256
www.bluebilladvisors.com

Editor:
Frank Gilbane
frank@gilbane.com
(617) 497.9443

Editors Emeriti:
Tim Bray
tbray@textuality.com
(604) 708.9592
David Weinberger
self@evident.com
(617) 738.8323

Associate Editors:
Bill Trippe
bill@gilbane.com
(617) 497.9443

Recent Contributors:
Sebastian Holst
sebastian@gilbane.com
Girish Altekhar
girish@deepcpclear.com
Bob Boiko
bob@metatorial.com
JoAnn Hackos
joann.hackos@comtech-serv.com
Rita Warren
rita@ziacontent.com

Associate Analyst:
Jared Spataro
jared@gilbane.com

Production Assistant:
Sarah G. Dionne
sarah@gilbane.com
(617) 497.9443

Subscriptions:
sales@gilbane.com
(617) 497.9443

Customer Service:
customersevice@gilbane.com

Consulting Inquiries:
consulting@gilbane.com
(617) 497.9443

Gilbane Report White Papers:
Jared Spataro
jared@gilbane.com
(617) 497.9443

Speaking Engagements:
speakersbureau@gilbane.com
(617) 497.9443

THE CLASSIFICATION & EVALUATION OF CONTENT MANAGEMENT SYSTEMS

Content Management Systems (CMSs) are a most unusual category of computer software. Most software categories as well known as content management are dominated by a few products that have captured market share and users' mind share. But the CMS space is different, very different, especially the most familiar Web Content Management System segment.

There are several reasons for this difference.

First, since a CMS has been notoriously hard to define (we address this below), there are many very different products that think they are providing a "complete" CMS. This includes fast-growing personal web publishing weblogs, group wikis, and organizational news portals, which have received a lot of press lately.

As the Gilbane Report suggested three years ago, "content management" has become an all-encompassing buzzword for document management, information management, and knowledge management. "It is largely because web publishing is so different from other types of publishing that content management has come to encompass such a wide variety of functions." (Volume 8, Number 8, October 2000).

Second, the huge potential market has attracted many independent software developers. Although we have seen a shakeout among the major early players, market entrants continue to appear, with several hundred companies listed on the DMOZ Open Directory Project under Content Management.

Third is the very low cost of market entry. With online marketing and delivery over the web, many companies dream of being the next *Vignette*. They don't realize that software is a small fraction of a content management solution.

Perhaps most important is that almost everyone who is managing a web site is fighting, some desperately, to get control over their site management. As a result, tens, perhaps hundreds, of thousands of organizations and individuals have produced something like their own CMS. A significant fraction of these have decided to market their CMS to others, selling them or renting them as Application Service Providers (ASP), and thus created a software marketplace like no other.

Since websites vary from individual user weblogs (personal web publishing), through organizational news-oriented portals and e-commerce sites, to massive enterprise portals that are the public front end for a complex of intranet corporate portal sites (Microsoft has 8000 of these), we expect a wide range in the power, complexity, and cost of the right CMS.

Any attempt to evaluate a CMS must begin with understanding what it does. Even more important though, is to understand what it is the website needs done. Hundreds of millions of dollars have been spent in recent years on early

versions of a CMS that did not have the functionality needed. In-house developers have added their own functions, and smart CMS developers have adapted many of these solutions back into their general offerings. So we are seeing some convergence on the core functionality of a CMS, whether for an enterprise or a small website.

The number of features associated with a CMS continues to grow. For example, most of the major CMS vendors have now made a significant commitment to the tools of, what some would call, "knowledge management". Some of these tools are new, but most are venerable techniques and terminology from print publishing, library science, linguistics, semiotics, and even philosophy.

The time has come for a serious effort to identify these core features and functions. If we are lucky, we might even get some agreement from vendors on what to call them.

A number of industry analysts, consultants, vendors, and users have joined us in a collaborative effort to provide a public domain taxonomy and descriptive tool to aid in the understanding of what a CMS is. "We", in this article, refers to this collaborative group.

A CMS IS WHAT A CMS DOES

Regular readers know we prefer definitions that reflect reality and are practical, so a definition of a CMS should not be based on what some think it *should* be. Rather, we prefer to attempt a definition of a CMS by describing all the things people actually are *doing* with one. And we can use the words they use to describe what they are doing.

Nearly three years ago we asked "What Is Content Management?" in the single most popular issue of the Gilbane Report. Now we want to continue our attempt at an answer by starting an industry-wide collaborative online effort to seek agreement on the core features and functions of a CMS. We will look for agreement from analysts, from industry experts in their books and articles, but especially from vendors and ultimately from the users.

INVESTIGATING CMS WITH CMS TOOLS

We will present our research in controlled vocabularies, faceted classification systems, topical searches, *etc.*, as a demonstration that uses Content Management tools to study Content Management. We will do this online, we will do it collaboratively, and we will do it with content management tools and techniques themselves. How?

First, we are building an online database of features and functions, plus a list of evaluatory questions, that vendors can access themselves with edit permission to enter quantitative values and qualitative answers.

The database is designed to be very flexible and extensible, allowing feature names and functions to be changed, definitions to be revised, and room for every vendor to enter their own qualified description when the generic feature or function does not fit their product. The inherent ability for CMS version control

will allow the work to evolve with the moving target both vendors and users are aiming at.

Second, we think only the widest possible collaboration is likely to get the necessary support from the vendors. If the resulting research is publicly available, not just on our website but on any collaborating CMS-related website that provides advice on CMSs, we expect the vendors will keep their data accurate and up to date. They will know that they enter data once and it comes out in many places on the web.

Distributing the knowledge base will take advantage of syndication technology, similar to news feeds, but designed to allow complete repository replication if desired, so multiple institutions can have full copies of the data. We will publish syndicated "knowledge feeds," which allow collaborating websites to provide straight lists of CMSs, sophisticated searches for CMSs with a particular set of features and functions, a faceted classification directory of CMSs, and a unique feature comparator that lines up the features and functions of two CMSs side by side to facilitate user decisions.

These knowledge feeds are a special web service of the CMS industry, by the industry, and for the industry.

We will encourage the major vendor-neutral CMS consultants, and even smaller vendor-specific consultants, to publish these interactive knowledge feeds on their websites to educate their clients. The question that usually comes right after "What is a CMS?" is "Which CMS is the right one?" Knowledge retrieval through a faceted classification scheme will allow users to put their most important CMS facets (characteristics or properties) at the top of their custom-built hierarchies. This should help bring the "right one" to the top of their list.

SYNDICATING CMS KNOWLEDGE SERVICES IN A WIDE VARIETY OF FORMATS

We will export our CMS data in unstructured HTML pages, as well as structured XML, RDF, OPML, XFML, and possibly XTM Topic Maps. We will do this to demonstrate the kinds of publishing possible with a CMS today, and to feed structures that will be maximally useful to those who want to post-process them into a more useful display format for their own needs. If web services are to succeed, we should demonstrate their success as an industry analytic tool.

The **XML** will be our own special markup language for describing Content Management Systems called CMSML. CMSML is a joint development of the OSCOM softwareML project, the University of Washington iSchool CMS Evaluation Lab, and CMS Review. It is described at www.cmsml.org. The mailing list is softwareML@oscom.org.

The **RDF** structure will enable inference engines to crawl the available CMS and draw conclusions about them. People may develop semantic web inferencing programs to help choose the right CMS from the many available. This is a major commitment to permanent URIs that describe CMS Resources.

OPML is Dave Winer's Outline Processor Markup Language. An organization can prepare a specific set of facets ordered by their priorities, and then create an OPML document that captures their choices, for further study and critical annotation in OPML readers and browsers. See www.opml.org.

XFML is Peter van Dijck's eXtensible Faceted Metadata Language, and it will transport the entire knowledge base in a file that can be browsed in tools like those from Travis Wilson's www.FacetMap.com. See www.xfml.org.

Although in many respects Topic Maps are isomorphic functionally to RDF Resources and both can implement the model of a semantic web, there are probably advantages to the CMS industry in a set of published Topic Maps that establish the core concepts and terms as well as the functions and features of content management systems. Others can then build their own Topic Maps that associate with these CMS Topics.

THE DRAFT FEATURE AND FUNCTION LIST

To develop our initial list we have looked at the best-known books and reports on the CMS business. Here is a bibliography.

- *Content Critical*, by Gerry McGovern
- *Content Management Bible*, by Bob Boiko
- *Content Management for Dynamic Web Delivery*, by JoAnn Hackos
- *Content Management Systems - Tools of the Trade*, by Dave Addey, James Ellis, Phil Suh, David Thiemecke
- *Managing Enterprise Content*, by Anne Rockley
- *Web Content Management*, by Russell Nakano

- *CMS Report*, by Tony Byrne, CMS Watch
- *CM Requirements Toolkit*, by James Robertson, Step Two Designs
- *The Gilbane Report*

Our list is divided into five major areas.

1. A **Product Overview** where we collect general information about the product (Description, Status, Marketing, Installation, and Support).

Then we look at four main phases of Content Management:

2. **Content Creation** (Acquisition, Aggregation, Authoring)
3. **Content Management Proper** (Workflow, Editing, Approvals, Staging, Repository, etc.)
4. **Content Delivery** (Live Server, Publishing, Syndication)
5. **Content Lifecycle Enhancements** (Security, Business rules, Integration, Metadata management, Transformations, Associations, Analysis, Search and Locate)

The current proposed feature and function list is attached in **Appendix A**.

NOTES ON THE COMPARATOR TOOL

How Some Features Become Facets

When a feature or function can be described only with a text entry, it has limited use in a faceted classification scheme. However, when the possible values of a feature can be enumerated from a list, this feature becomes a facet that can be used to sort the CMS in different ways.

One such facet is the computer operating system. Another is the programming language. When deciding on a CMS, arranging systems in a hierarchy that puts operating system and programming language in the top layers aggregate systems that will suit different IT strategies.

Figure 1 shows an example of choosing a platform (OS), a web server, and a programming language in the **Faceted CMS Directory**. The list at the right now shows the CMS Frameworks (our fourth facet choice) available.

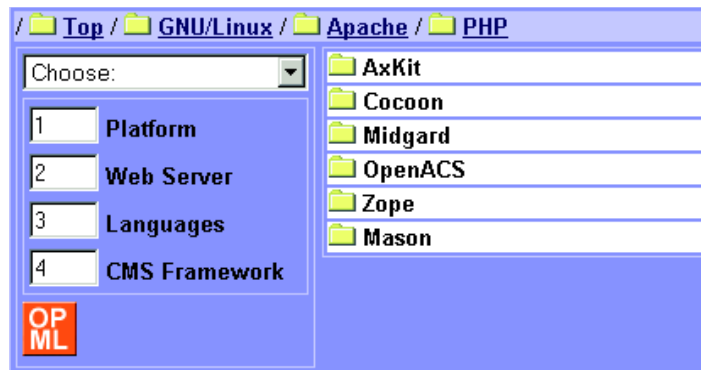


Figure 1.

Figure 2 is an example of choosing a platform (OS), a web server, and a programming language. The list at the right now shows the CMS products available.

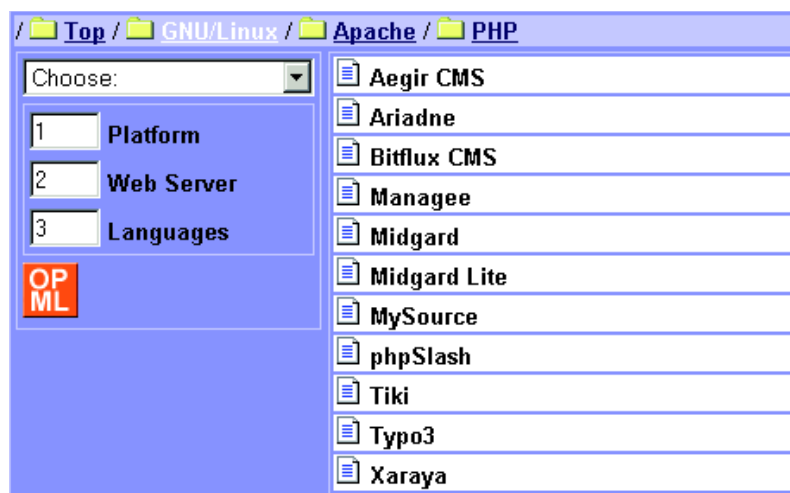


Figure 2.

Figure 3 is the result of choosing two CMS for a side-by-side comparison in the Feature Comparator



Midgard 		Tiki 
▼ CMS		
▼ Overview		
▼ Description		
Product Name	Midgard	Tiki
Organization Name	The Midgard Project	Tiki
Organization Website	http://www.midgard-project.org/	http://tikiwiki.sourceforge.net
Product web page	http://www.midgard-project.org/midgard/	http://tikiwiki.sourceforge.net
Description	Midgard is an Open Source Content Management Framework based on the popular Linux, Apache, MySQL and PHP (LAMP) platform. Midgard's key benefits include high performance, robustness, flexible templating system, internationalization and the MidCOM component model. Midgard is available both as compilable Apache and PHP modules, and as a pure-PHP implementation, Midgard Lite.	Tiki is full featured content management system suited to many types of online communities. Features include news, topics, wiki, polls, trackers, image galleries, forums, blogs, webmail, and much more. Using PHP, MySQL and Smarty.
▼ Technology		
License	LGPL	LGPL
Type	Framework	General CMS, News Portal, Blog, Wiki
Platform	GNU/Linux, Unix	GNU/Linux, Windows, Macintosh OS X, Unix
Web Server	Apache	Apache, IIS
Application Framework	PHP	PHP

Figure 3.

Cautionary Notes on CMS Evaluation

Some clever programmer may take a knowledge service feed and try to generate a rating for various CMSs based on our data. We encourage serious users to regard these data as a starting point for their own CMS evaluation and selection process. Work with an established CMS consultant or your own in-house IT staff who knows how to use the data to create the right fit for your CMS needs. Tools such as these cannot automatically choose “the right” CMS.

Vendor marketing teams will naturally try to use any classification scheme to make their products look better. Our approach to problem is to support vendor-qualified descriptions for each feature. This will encourage a clear statement of what the vendor means in the particular case. Also, any particular feature may be open to (mis)interpretation, either deliberate or accidental. Different vendors may genuinely mean different things by the same term. These are both further reasons to work with a qualified internal or external consultant.

Note that we are making no effort whatever to create new terminology. We firmly believe that “meaning is use,” and we are trying to discover that use at the same time we try to discover what CMSs have in common to help understand and define them. We think we can help by publishing a glossary of terms – a controlled vocabulary for the CMS industry – that is used to describe the core features. We want to get everyone at least familiar with industry jargon. This will increase the chances that they will converge on shared concepts. Formally we describe this as building a thesaurus of terms. As in ordinary language, a thesaurus can have rich semantic relationships, with many terms substituting for one

another, and single words having different meanings in different contexts. We have the technology to identify preferred terms and common variants.

HOW TO ACCESS AND USE CMS KNOWLEDGE SERVICES

For Users of the Services

Visit one of the collaborating CMS-related sites who are supporting the development version of the knowledgebase and tools:

- www.cmsInfo.org
- www.CMSReview.com
- www.Content-Wire.com
- www.Gilbane.com (available in early June 2003)
- www.OSCOM.org

For CMS-related sites that would like to add CMS Knowledge Services for their site visitors

- Go to www.CMSReview.com/CMSFeeds.html
- Copy the code snippet for the desired CMS Knowledge Service and paste it into one of your web pages. The service is now available on your site. Further help is available at <http://www.cmsreview.com/HowTo/ForVendors.html>

For CMS vendors that would like to edit their feature data

- Go to www.CMSReview.com
- Login without a user name and password. You will be invited to join as an individual or organization and you will be sent a password. You will be given edit privileges by the administrator.
- Click on the Edit links next to your product name, log in, and edit your data. Further help is available at <http://www.cmsreview.com/HowTo/ForVendors.html>

For all interested who would like to criticize the feature list

- Join the CMSML mailing list at <http://www.oscom.org/mailman/listinfo/softwareml>
- Send your thoughts to me (editor@cmsreview.com).

PLEASE JOIN US!

The knowledge services tools described here are in development. We need the participation of the CMS vendor community to fill in their data. We need them and we need CMS users to help us refine the most meaningful features and functions to be included in the master list.

A progress report on this project will be delivered at the *Gilbane Conference on Content Management @ Seybold* in San Francisco, September 9-10, 2003.

Bob Doyle bobdoyle@skybuilders.com, with Gregor Rothfuss gregor@apache.org, and Frank Gilbane frank@gilbane.com

APPENDIX A: PROPOSED CMS TAXONOMY

This is a work in progress. Please comment, suggest, critique, or congratulate.
editor@cmsreview.com.

1. Product Overview

- Description
 - Product Name
 - Company Name
 - Company/Organization website
 - Product web page
 - Company's description
 - Our Description
- Technology
 - License - Open-source, Proprietary, which
 - Type - General CMS, Framework, Front end (UI), News Portal, Blog, Wiki
 - Platform - Windows, Linux, Mac, *etc.*
 - Web Server - IIS, Apache, *etc.*
 - Application Framework - Perl, Python, .NET, J2EE, PHP, Cold Fusion, *etc.*
 - CMS Framework - AxKit, Cocoon, Midgard, Zope, *etc.*
 - Languages - Perl, VB, Java, PHP, Python, *etc.*
 - Databases - Oracle, SQL Server, MySQL, PostgreSQL, any ODBC, *etc.*
 - API - public to allow extensibility.
- Status
 - Release - 2.0, *etc.*
 - Year introduced
 - Number of Installs, Downloads
 - Developer Community (website?, mail list?)
- Marketing
 - Price
 - License (per CPU, per user, *etc.*)
 - Market Position (Revenues, Competitors)
 - Sales Methods (Sales Force, Online)
 - Support Contracts, Consultants
 - Online Demos, Sandbox, Prototype, Trial
- Installation
 - Online How To
 - Hours/Days for Typical Install
 - Documentation online/printed
 - Download site/CD-ROMs
 - Code Commented
- Support
 - Online Help
 - Tutorials
 - Training Classes
 - Cost

-
- Commercial Contracts
 - Help Desks
 - Independent Consultants

2. Content Creation (Acquisition, Aggregation, Authoring)

- Acquisition
 - Native support for filetypes
 - Multiple file transfers (FTP, site import)
 - Conversion tools (e.g., Word to XML “chunks”)
 - Rights management
 - Mandatory tagging (force structure and semantics)
 - Supported RDF ontologies (e.g., Dublin Core)
- Aggregation
 - Incoming syndicated feeds
 - Metadata management (read incoming metadata)
 - Integrated Web services (e.g., currency conversion)
 - UDDI tools
- Authoring (Editing Tools, Templating, Tagging)
 - Content Element Editors (Naive and Power Users)
 - WYSIWYG Through-The-Web
 - Text-only Forms
 - XML Editor
 - Spell checker
 - Content objects use templates
 - Asset repository (images, sounds, Flash, video, etc.)
 - Template Editor
 - WYSIWYG Through-The-Web
 - Template Gallery
 - XML Editor
 - Tag Editor (semantics and style)
 - Drop-down menus of all tags
 - Metadata Thesaurus
 - Taxonomies/Ontologies online
 - Help online
 - Context-sensitive help
 - Documentation
 - Examples

3. Content Management Proper (Workflow, Editing, Approvals, Staging, Repository, etc.)

- Workflow
 - Access Permission Levels (Privilege granularity)
 - Number of levels
 - Per User, Per Folder, Per Role, Per Item
 - Flexible assignments to workflow
 - Creator automatic owner
 - User subscription to workflow
 - Check In/Check Out
 - Open page on web (Edit this page)
 - Automatic file lock on open

-
- Conflict Resolution (who has it?)
 - Instant Messaging (email, phones)
 - Merge Tools, Diffs
 - Workflow Messaging
 - Email notifications (links to work)
 - Status (stage in workflow)
 - Comments at each stage
 - Audit trail (workflow log)
 - Arbitrary Roles (Writers, Editors, Graphic Artists, Rights Managers, Publishers, etc.)
 - Versioning
 - Scheduling, Expiration
 - All elements, templates date/time stamped
 - Archive with rollback (per file or site?)
 - Personalization
 - Relationship Management (History)
 - Actions tracking
 - Session/Click/Behavior analysis
 - Individual visitor ID (cookies)
 - Localization
 - Multilingual server
 - Respond to browser language requests
 - Gist translation option
 - Workflow
 - Automatic notifications
 - Quality checkers
 - UI multilingual
 - Reporting
 - Chrono workflow and by worker
 - WebTrends-style for whole site
 - Specific monitors
 - Performance (page delivery times)
 - Storage
 - Format (text, HTML, XML)
 - Database only
 - Files
 - Files and database
 - Backup
 - Onsite and offsite
 - Files and database
 - To nonvolatile media
 - Disaster recovery plan
 - Security
 - Firewall rules
 - Encrypted sessions
 - Staging Server for QA
 - Testing methodology
 - Replicates publishing environment

4. Content Delivery (Live Server, Publishing, Syndication)

- Publishing (Delivery)
 - Separate Delivery from Creation/Staging/Testing

-
- Use different server platform?
 - Replication
 - Synchronization of mirror sites
 - Multi-Publishing to different clients
 - HTML, XHTML, XML
 - PDF
 - PDAs
 - Cell phones
 - Handicap accessibility
 - Syndication
 - RDF Syndicated News Feeds
 - Web services
 - Information feeds

5. Lifecycle Enhancements (Apply to all three stages above)

- Security
 - Audit trails
 - Users
 - System
 - Network
- Business rules
 - Records policies
 - Privacy policies
- Integration
 - Single authentication
 - Enterprise portal
 - Legacy database reuse
 - Data warehousing
- Metadata management
 - Digital rights management
 - Digital rules management
- Transformations
- Associations
 - Hierarchy, taxonomy
 - Index
 - Cross reference
- Analysis
 - Analytic Tools
 - Pattern recognition
- Search and Locate

INDUSTRY NEWS

Current news, old news (to January 1999), and commentary is available at www.gilbane.com. Free RSS 2.0 news feeds are available at www.gilbane.com/syndication.html.

PERCUSSION INTRODUCES RHYTHMYX 5 ENTERPRISE CONTENT MANAGEMENT

3/31/2003

Percussion Software introduced version 5 of its Rhythmyx Content Management System (CMS), a major release that extends Rhythmyx into an ECM solution. It also features an architecture for system scaling - De-Coupled Delivery - which supports new delivery platforms as they are adopted. New capabilities in release 5 include: A new Document Management Module; new Rhythmyx Content Connectors for Oracle9i Applications Server and BEA WebLogic portals, the MediaBin Digital Asset Management (DAM) system, and Kofax' Ascent Capture; a new user interface for managing all ECM content; enhanced globalization support for management of multi-lingual content delivered to multiple audiences; and a new Web Services Developer Kit. Rhythmyx 5.0 is sold as a comprehensive ECM solution, or as a basic solution for Web Content Management (WCM) implementations, with a smooth upgrade path to the full ECM product. Both configurations are built on the same ECM foundation code base. Rhythmyx 5 supports Windows NT 4.0, Windows 2000 and Solaris platforms and is available for licensing immediately. www.percussion.com

MERANT INTRODUCES LATEST VERSION OF ENTERPRISE CONTENT MANAGEMENT SOLUTION

3/31/2003

Merant announced Merant Collage 4.0. Merant Collage 4.0 is easier to use, delivers advanced deployment capabilities, and offers enhanced support for Oracle9i Database, Microsoft SQL Server 2000, Linux servers and Macintosh clients. Merant Collage 4.0 features enhancements to deployment capabilities, enabling organizations to deploy code and content to databases and portals in addition to staging and production servers. Merant Collage also allows companies to deploy assets residing outside of the Collage repository, associate custom commands with deployment at multiple points of time, and execute custom commands on deployment servers. Merant Collage 4.0 supports Microsoft SQL Server 2000, Linux servers and Macintosh clients, as well as the comprehensive, integrated platform provided by Oracle's end-to-end technology stack, including the Oracle9i Database and Oracle9i Application Server. Merant Collage also supports multiple other operating systems, relational database platforms, application servers, desktop browsers and industry protocols. Merant Collage 4.0 will be available in April 2003. www.merant.com

REDDOT LAUNCHES CONTENT MANAGEMENT SERVER 5.0

3/31/2003

RedDot Solutions announced the launch of RedDot Content Management Server (CMS) 5.0. Among the new features are improved database support including native Oracle connections, and a graphical workflow designer. The RedDot translation editor is also now available as a server option. RedDot now offers full support for Oracle (in addition to MS-SQL Server and MSDE) as the internal database platform for a RedDot deployment. There is a new graphical browser-based interface for the design and modification of workflow processes within RedDot.

The RedDot Translation Editor supports a global enterprise with translation tools to assist with the deployment of multilingual sites. Translators are notified via workflow e-mail when content has been created or modified and that content needs to be translated into the language that s/he is responsible for. The RedDot Translation Editor gives the translator a split-screen text editor with one side displaying the original content and the other a text editor ready for the translator to input the content in the language needed. For a detailed list of CMS 5.0 features and functionality, see www.reddot.com/product/product_v5.htm

LEGATO INTEGRATES VERITY SOFTWARE INTO CONTENT MANAGEMENT SUITE

3/31/2003

Verity Inc. announced its Verity K2 Developer (K2D) software has been embedded into ApplicationXtender, an integrated suite of enterprise content management solutions from LEGATO Systems Inc. Verity K2D complements ApplicationXtender's foreign language support in the enhanced search capability by supporting the languages offered by ApplicationXtender -- Japanese, Spanish and Brazilian Portuguese. www.verity.com

ENDECA & CLEARFOREST ANNOUNCE STRATEGIC PARTNERSHIP

3/31/2003

Endeca and ClearForest announced a partnership that will provide customers with the ability to search, explore and navigate unstructured enterprise data. ClearForest will offer unstructured data auto tagging to customers deploying Endeca ProFind enterprise search application solutions. Endeca and ClearForest have integrated their technology products to provide a more powerful information retrieval system for large corporate and government users. Through its automated creation of metadata, ClearTags transforms unstructured content and makes it as accessible as the structured information typically found in databases. In combination with Endeca ProFind, an advanced search and Guided Navigation application, users will benefit from search and navigation of unstructured data. www.endeca.com, www.clearforest.com

PLUMTREE ANNOUNCES INTEGRATION WITH FILENET ENTERPRISE CONTENT MANAGEMENT

3/27/2003

Enterprise Web technology leader Plumtree Software announced an integration suite that delivers access to FileNet's P8 enterprise content management (ECM) architecture through the Plumtree Corporate Portal. The integration suite, developed and supported by Soft Cell, consists of portlets for incorporating FileNet's content and business process management, and a certified Crawler Web Service for indexing and meta-tagging FileNet content in the Plumtree Corporate Portal Document Directory. The Crawler Web Service for FileNet Content Services by Soft Cell periodically scans the FileNet repository for the latest document versions and new content, and creates links in the portal's document directory to text documents, images, spreadsheets or other forms of content. The Crawler and portlets are available from Soft Cell and Plumtree. www.softcell.be, www.plumtree.com

ARTESIA RELEASES TEAMS ADAPTER FOR VIGNETTE CONTENT MANAGEMENT

3/26/2003

Artesia Technologies announced the immediate availability of a newly developed integration adapter linking the Artesia TEAMS solution with Vignette's family of content management and portal solutions. With the release of this product, enterprises can integrate Artesia's digital asset management solutions with Vignette's solutions. The jointly developed integration package is currently available. Pricing information may be obtained through Vignette and Artesia.

www.artesia.com

HUMMINGBIRD ACQUIRES KEY AUTOMATION/DISPRO

3/26/2003

Hummingbird Ltd. announced the acquisition of Key Automation Nederland B.V. and its affiliate Dispro B.V. (herein referred to as "Key Automation/Dispro") privately held distributors and integrators of document management solutions in the Netherlands and Benelux. Key Automation/Dispro are long standing Hummingbird partners in the deployment and integration of Hummingbird products. While their expertise spans multiple industries, Key Automation/Dispro is mostly recognized in the government sector for document, records and correspondence management solutions. Key Automation/Dispro provide product integration, deployment, implementation and educational services that deliver single source point of support to manage the entire life cycle associated with enterprise information management solutions.

www.hummingbird.com

MEDIABIN & SMARTPATH PARTNER

3/25/2003

SmartPath Inc. announced a strategic partnership with MediaBin Inc. The integration of SmartPath's MRM (Marketing Resource Management) solution with MediaBin's DAM (Digital Asset Management) solution provides comprehensive functionality to support production and project management initiatives for all marketing activities. This integration will offer large marketing organizations the capabilities they need to speed marketing campaigns and product launches while increasing efficiencies and reducing costs. www.smartpathinc.com, mediabin.com

WEBWARE INTRODUCES ACTIVEMEDIA ACCESS

3/25/2003

WebWare Corporation announced the introduction of WebWare ActiveMedia Access, an enterprise-ready digital asset management system for under \$50,000. ActiveMedia Access is engineered for the enterprise but designed for the needs of the workgroup. It enables enterprise workgroups, creative agencies, video post-production facilities, and similar environments to create a digital library for rich media: key graphics, layouts, illustrations, slide presentations, and video. ActiveMedia Access is built on the web services-based ActiveMedia 4 content management engine. ActiveMedia Access lets project managers assign sophisticated business rules to a company's digital assets at the workgroup level without the need for database administrators or outside professional services. It also allows authorized users to access, approve, track, audit, and transform product and marketing content instantly and securely over the Internet. ActiveMedia Access runs on Microsoft Windows 2000 and SQL Server, or Sun Solaris. It is available as both an installed software solution or a fully hosted service. www.webwarecorp.com

ARTESIA INTEGRATES DIAMONSOFT'S FONT SENSE

3/25/2003

DiamondSoft Inc. and Artesia Technologies announced the integration of Diamondsoft's Font Sense technology into the Artesia TEAMS Creative Client. Font Sense enhances the Artesia TEAMS digital asset management solution by identifying and recording critical information about fonts used in digital content, thus preventing workflow interruptions that may occur due to missing fonts or incorrect font usage. When ingesting files, the TEAMS Creative Client can now automatically extract and store the Font Sense identification provided by Font Reserve Server as metadata in TEAMS. If a user needs to reprint an asset -- perhaps years after its creation -- the correct fonts are always available via Font Reserve Server, eliminating wasted time spent searching for the correct font. Additionally, when assets must be sent to a third-party collaborator or partner, fonts can be collected and automatically attached to the content upon export from Artesia's TEAMS. www.artesia.com, www.fontreserve.com, www.diamondsoft.com

W3C ISSUES XPOINTER RECOMMENDATION

3/25/2003

The World Wide Web Consortium released XPointer as a three-part W3C Recommendation. The "XPointer Framework" is an extensible system for identifying regions in XML documents which provides for multiple addressing schemes. The "element()" scheme allows basic addressing of XML elements in terms of a document's tree structure. The "xmlns()" scheme is used to interpret namespace prefixes in scheme names and pointers. www.w3.org/TR/2003/REC-xptr-framework-20030325/

XEROX LICENSES DRALASOFT WORKFLOW FOR DOCUSHARE

3/24/2003

Dralasoft announced that Xerox Corporation has licensed Dralasoftware Workflow for inclusion into DocuShare, Xerox's Web-based document and content management software. By embedding Dralasoftware Workflow engine into DocuShare, Xerox can help its DocuShare customers route documents more efficiently across their enterprises through workflows built with modeling tools that automate and map customer business processes. As a document repository, DocuShare keeps critical information safe, yet easy to access. Users can store any file type, including office documents, images, media files and graphics. As a content management system, it updates and manages the content of a Web site. As a collaborative project management system, it enables users to organize, manage and share all of a project's information. Its Java-based DocuShare platform will enable third parties to develop custom workflow applications through Dralasoftware Workflow. DocuShare runs on Windows, Linux and Solaris systems. Dralasoftware products provide the capability to integrate workflow technology into Java-based systems using Java, XML, SOAP, LDAP, HTTP and JDBC. www.dralasoft.com, <http://docushare.xerox.com>

FAST UNVEILS DATA SEARCH 3.2

3/24/2003

Fast Search & Transfer (FAST) unveiled FAST Data Search 3.2, an advanced version of its real-time search solution that is designed for organizations seeking either departmental or enterprise wide solutions. Offered as FAST Data Search Professional, FAST Data Search Enterprise, and FAST Data Search SST, the FAST Data Search product line now includes Live Analytics, which enables "on-the-fly" data and information analysis with a dynamic drill-down capability, where end users can tap the specific information they need based on the actual data in the result set, rather than pre-defined binning of the data. Also now included is the Business Manager's Con-

trol Panel, so business managers can tune the search attributes of the data source. Additional enhancements include improvements to the combination of structured and unstructured data capabilities, automatic document categorization, dynamic result clustering, and similarity based results analysis. Advanced linguistics support for Japanese was added, and automatic language detection was increased to support 54 languages. FAST Data Search also now includes enhanced security features and supports HP-UX and Red Hat Linux Advanced Server, in addition to IBM AIX, Red Hat Linux, Solaris, and Windows 2000/XP. www.fastsearch.com

IXOS & CONVERA PARTNER

3/24/2003

Convera and IXOS Software Inc. announced a new partnership designed to assist enterprise customers in managing large volumes of business documents and email communications. IXOS will sell Converas RetrievalWare search and categorization software as an optional component of the IXOS-eCONserver product. The combined solution will add Converas search capabilities to enable customers to retrieve email messages, attachments and other business documents from the IXOS-eCONserver archive for a variety of purposes, including records management, discovery and general knowledge management. The solution, available immediately, will be sold to the commercial and federal market by IXOS North American sales teams in the United States, Canada and Mexico. www.ixos.com, www.convera.com

CONVERA LAUNCHES TAXONOMY DEVELOPER PROGRAM WITH 10

PARTNERS

3/24/2003

Convera announced that it has certified ten partners as part of its new Taxonomy Developer Certification Program. Under the program, Convera trains and certifies IT and knowledge management professionals in the development and deployment of custom taxonomies and classifications using Converas new RetrievalWare Categorization and Dynamic Classification product. Participants who completed the five-day certification program include Access Innovations, Advanced Technology Systems (ATS), Bleuphish, Creative Technology Inc., IBM Italia, KAPS Group LLC, Mitre, MZM Inc., SRA International Inc., and Veridian Engineering. Certified Taxonomy Developers will serve as consultants and service providers to assist Convera customers in understanding their information management needs, organizing their information in a way that best serves their objectives, and developing well-tuned taxonomies for use with Convera software. Certified Taxonomy Developers also will have the opportunity to market cartridges that they have developed within the Convera customer base as pre-built taxonomies. www.convera.com

BACKWEB ENHANCES PROACTIVEPORTAL SERVER

3/24/2003

BackWeb Technologies announced Version 3 of its ProactivePortal Server offering. The new product release enables mobile professionals to not only access their portal offline, but also update portal data while they are disconnected from the network. The BackWeb ProactivePortal Server integrates with portal frameworks to enable offline interaction with critical content and applications with essentially the same personalized user experience that occurs online. ProactivePortal Server v3 delivers a two-way synchronization capability that enables end users to fill and submit portal-based forms while disconnected from the network. Once the user reconnects online, the forms are synchronized up to the BackWeb ProactivePortal Server and applied to through the portal to the underlying applications. This ensures all updated and altered information is automatically stored and synchronized. BackWeb ProactivePortal Server v3 is available immediately and can be purchased through SAP, IBM and BackWeb. www.backweb.com

EMETA UNVEILS ERIGHTS SUITE

3/24/2003

eMeta Corporation announced the availability of the eRights Suite. The eRights Suite provides the infrastructure, applications and insight necessary to secure and sell digital assets and services. The new Suite enables enterprises to create, control and commercialize digital offerings that are synchronized to specific business needs and models, assets and audiences. The eRights Suite is composed of RightAccess 3.0, RightCommerce 3.0, and RightServices 1.0. RightAccess provides the infrastructure necessary to protect and manage access to digital assets and services. Offering advanced authentication and authorization, delegated administration, and metadirectory features. RightCommerce contains packaging, pricing, customer care, billing and payment applications. RightServices provides a solution for the protection and commercialization of those assets, and is optimized for non-traditional online assets such as software services and software code. RightCommerce may be integrated with Content Management, CRM, and DRM systems. www.emeta.com

ADVENT ANNOUNCES 3B2 V8

3/20/2003

Advent Publishing Systems announced a new and substantially enhanced version of its 3B2 software, Version 8, at XML Europe on May 5th 2003. Version 8 of 3B2 implements a range of XML-related international standards including XSL-FO, XPath, XSLT, and SVG. Additional features in Version 8 of Advent 3B2 software include: page building on the fly wherein content can determine page layout, more consistent 3B2 programming commands, a new de-bugging tool that streamlines the development and debugging of templates and formats, a new format profiler for producing job statistics on the processing speed of a job, font quick tables to make fonts more accessible, new and improved CALS libraries for use with SGML and XML, entity libraries to support ISO standard entity sets, and an improved and enhanced user interface. Version 8 will be available for customer shipment beginning Monday 5th May 2003.

www.3b2.com/v8

CAMBRIDGE DOCS & ALTOVA ANNOUNCE xDOC XML CONVERTER PLUG-IN FOR XMLSPY

3/19/2003

Altova software and CambridgeDocs announced the xDoc Plug-in, a tool for migrating unstructured content from Microsoft Word, HTML, and PDF documents, directly into Altova's XMLSPY. The combined solution helps businesses transform volumes of pre-existing unstructured documents into DocBook XML, LegalXML, NewsML, HR-XML, SCORM, and other customer-specific schemas. The resulting XML content can be stored in content management systems, databases and XML repositories. XMLSPY users would be able to select "Open From Microsoft Word", "Open From HTML", "Open From PDF" or from any other supported format, and have the documents automatically converted into XML, along with an appropriate XSLT stylesheet required to view the XML as a well-formatted document. The xDoc Plug-in is a Java 2 platform engine and works on Windows 2000 and Windows XP. The application can be downloaded for immediate evaluation and purchase. www.cambridgedocs.com/products_xmlspyplugin.htm, www.altova.com/download

VERIDOCs RELEASES XMLDOCS 1 WITH "TEMPLATE DESIGNER"

3/19/2003

Veridocs announced the availability XMLdocs Version 1, a hosted XML document authoring and management system. For the release, Veridocs has added a flexible new "Template Designer" feature not available on the products Beta Version. Template Designer makes it simple for organizations to create templates from schemas (or use an existing template from XMLdocs as a starting point for a new template) for everything from press releases and product descriptions to knowledge-based articles and help documents. Using these templates, non-technical users write XML documents via the XMLdocs Editor's word-processor-like interface. Documents can then be exported for publication on a Web site or elsewhere through a Web Service. XMLdocs is priced to be affordable for everyone interested in authoring content in XML. Individual users can register for the product and create a personal workspace for free. Group fees start at \$99 per month. my.xmldocs.com/register

MONDOSOFT & SITECORE TO PROVIDE ENHANCED CONTENT MANAGEMENT

3/18/2003

Mondosoft announced an integration between MondoSearch site search and SiteCore's content management system. The integration enables MondoSearch enterprise search to index and retrieve content managed by SiteCore content management system to provide information retrieval and immediate access to business content. Both solutions use Microsoft .NET built technology. Web site owners will also have the advantage of visitor search activity reporting to determine who is using the search function, what information they are looking for and, what they are finding and not finding on the site. Content can be simply adjusted and refined using both SiteCore and MondoSearch. www.mondosoft.com

XYENTERPRISE ANNOUNCES WEB SERVICES DEVELOPMENT KIT FOR XML PUBLISHING

3/18/2003

XyEnterprise announced a new Web Services Development Kit for its XML Professional Publisher (XPP) software line. This new Web Services interface makes available the XML publishing capabilities of XPP to Web interfaces, portals, and application integrations, providing rapid deployment of automated XML publishing solutions in environments as diverse as corporate IT services to consumer "publish on demand" applications. The new Web Services offering exposes the composition, transformation and rendering capabilities of the XPP software through a SOAP/XML interface. This API provides access for customers and developers, enabling them to build customized end-user interfaces in support of distributed XML publishing services. Web Services also supports the deployment of XPP as an embedded batch processing engine able to transform XML content to PDF, print, and Web output on demand. www.xyenterprise.com

SDLX TRANSLATION SUITE 2003 LAUNCHED

3/18/2003

SDL Desktop Products has announced the launch of a new version of its SDLX Computer Aided Translation system with over 20 new enhancements. SDLX Translation Suite 2003 is a complete and secure translation environment for corporations, translation companies and freelance translators. www.sdlx.com

SEMAGIX SHIPS FREEDOM 3.0

3/17/2003

Semagix announced that it is shipping version 3.0 of its Freedom platform for enhancing meta-data and integrating enterprise information from structured and unstructured content sources. Semagix Freedom is a semantics-based Enterprise Information Integration (EII) software platform that is used to aggregate and relate large volumes of heterogeneous information held in disparate locations. The enhanced features of Freedom 3.0 simplify and reduce the cycle for System Integrators to implement the Freedom platform. Semagix has enhanced several usability, scalability and performance features. Semagix's Freedom 3.0 is available immediately with prices starting at \$75,000 per processor. www.semagix.com

INTRASPECT ANNOUNCES NEW COLLABORATION SOLUTIONS

3/17/2003

Intraspect Software, Inc. announced a new set of collaboration solutions. Comprised of the Solution Suite for High Technology, Solution Suite for Financial Services and Solution for Project Delivery/PMO, these new offerings incorporate best practices from over 400 solutions implementations with 200 customers. These solutions can be implemented faster when compared to generic collaboration products. Intraspect Solutions are a combination of its collaboration software platform -- Intraspect 5, application-specific user interfaces and data models, business process models and workflows, connectors to enterprise applications and packaged consulting services designed to meet the needs of Fortune 1000 customers. www.intraspect.com

VERITY PROVIDES STELLENT CERTIFIED CATEGORIZATION FOR CONTENT MANAGEMENT

3/17/2003

Verity, Inc. announced Stellent Inc. has certified specially-tailored Verity categorization technology for use with its content management system. The integration of the Verity Profiler product with the Stellent Content Management system allows enterprise users and administrators to populate Stellent Content Server metadata quickly, accurately and automatically when new or modified content is added into a Stellent data repository. It complies with Stellent's API for categorization, offering Stellent customers a new choice to best meet their categorization needs. www.stellent.com, www.verity.com

ALTOVA & DATAPOWER TEAM

3/17/2003

DataPower Technology, Inc. and Altova, Inc. announced the availability of XMLSPY 5 integrated with the DataPower XS40 XML Security Gateway. The unified solution addresses enterprise the need for centralized XML Web Services security without forcing application developers to alter pre-existing design and deployment practices in any way. The XMLSPY 5 XML development environment provides visual design, development and debugging for all types of XML documents including XML schema, XSLT, SOAP and WSDL. DataPower's XS40 is an XML security gateway providing XML filtering, schema validation, XML Encryption, XML Signature, WS-Security and others. Security policies within the DataPower XS40 are XML-based so developers can use existing XML Schema, WSDL and other XML application files to assign filtering rules, access control and overall policy management. XMLSPY 5 and XS40 as integrated solution is available immediately. Pricing for the solution starts at \$64,995. The XMLSPY 5 Professional Edi-

tion includes an XSLT debugger and is available for stand-alone purchase for \$399 from the Altova online shop. www.altova.com, www.datapower.com

STELLENT ACQUIRES ACTIVE IQ ASSETS

3/17/2003

Stellent, Inc. announced it has acquired selected assets of Active IQ Technologies for \$650,000 in cash, or approximately one times the acquired annualized recurring revenue. The transaction will have no revenue effect on Stellent's March 31, 2003 quarter and is expected to be accretive for Stellent's fiscal year 2004. There is no ongoing business relationship between the company and Active IQ. Under terms of the agreement, Stellent has acquired technology, customer agreements and other assets related to the hosted application business that Active IQ built based upon the Stellent Content Management system. The company has assumed Active IQ's current subscription revenue agreements with customers and will continue to market the hosted applications for vertical markets. Stellent will also offer general hosting capabilities to other content management prospects and customers. Stellent has hired Active IQ sales, marketing and development staff, and Active IQ president and chief executive officer Jack Johnson is now Stellent's vice president of hosted solutions. www.stellent.com

INSTRANET ANNOUNCES CHANNELS IN-LINE 3.2

3/13/2003

InStranet, Inc. announced the release of version 3.2 of its Channels In-Line (CIL) application featuring built-in analytics. With a strong emphasis on feedback and measurement, these analytics enable managers to make informed sales and marketing decisions based on market pressures, competition and channel needs. In addition to publishing and syndicating information, companies can measure its effectiveness in real-time. CIL 3.2 ships with the following business indicators as part of the new analytics module: Field Audit, Awareness Ratio, and Document Effectiveness Metric. CIL is a web-based application that ships with InStranet's Content Application Foundation. It includes the following content and document management modules: publishing, workflow, file system agent, syndication, user management, security, metadata tracking and organization, multi-dimensional segmentation engine and application content exchange. CIL is targeted at companies with multiple distribution channels, large sales organizations, and multiple customer touch points within the following industries: retail, financial, manufacturing, CPG, healthcare and technology. www.InStranet.com

VIGNETTE ANNOUNCES V7 CERTIFICATION FOR IBM

3/13/2003

Vignette Corp. announced certification of Vignette V7 for IBM software and hardware solutions. Vignette V7's support for IBM infrastructure provides current and prospective customers with the ability to standardize on a unified solution for managing content and portals across the enterprise. The Vignette and IBM solution addresses the challenges of creating and managing portals and Web sites to deliver the right enterprise information and business processes to each audience. When combined, IBM's application infrastructure and hardware and Vignette's content management and portal applications enable organizations to deploy enterprise Web applications that protect and leverage investments in enterprise applications, databases, application servers, security and operating systems. Vignette's support for the IBM platform now includes the AIX operating system, WebSphere Application Server, DB2 Universal Database, IBM Directory Server, WebSphere Portal Server, IBM MQSeries, IBM Content Manager 8, Lotus Notes and Lotus Domino. www.vignette.com

HiSOFTWARE INTEGRATES WITH MICROSOFT CONTENT MANAGEMENT SERVER

3/13/2003

HiSoftware, Inc. announced integration between its AccVerify, AccMonitor Server and Microsoft Content Management Server (CMS) 2002, enhancing the ability to ensure that websites meet Section 508 recommendations and organizational Web content quality standards. The combined solutions allow for integrated access for verification of accessibility and content quality throughout the different layers of CMS managed Web sites. AccVerify, AccMonitor and Content Management Server integration can be used together "out-of-the-box" with no additional development resources required, so that content quality and accessibility verification is seamlessly integrated into the development, deployment and delivery of the CMS-driven Web sites. These solutions also allow organizations to address accessibility and content quality verification issues related to XML-based Web services. www.hisoftware.com

KMTECHNOLOGIES & SWINGTIDE PARTNER

3/12/2003

KMtechnologies Inc. announced a new partnership agreement with Swingtide. FSwingtide has licensed KMtechnologies' work2gether platform to deliver the Swingtide QoB Assistant. Leveraging the work2gether collaborative features, the QoB Assistant is a productivity tool for architects and developers to understand, define and manage the interoperability of XML used in service-oriented architectures. The QoB Assistant is a software application packaged with an online knowledgebase of XML standards that's accessible from your network. The QoB Assistant automates planning in several key ways. First, the QoB Assistant lets you make specific choices about XML and make them persistent across your local copy of the knowledge-base as design guidelines. Second, several people can work and collaborate on a design simultaneously. The QoB Assistant comes complete with discussion thread capability, email notification when team members update a portion of the design, and even support for distance collaboration. Additionally, the QoB Assistant enables you to add custom content that can be leveraged by the rest of the team. www.kmtechnologies.com www.swingtide.com

FATWIRE INTRODUCES FREE CONTENT MANAGEMENT PORTLETS FOR WEBSPHERE

3/11/2003

FatWire Software introduced a new dCM Integration kit with six free FatWire dCM portlets for IBM's WebSphere Portal. The portlets allow business users to easily create and modify both structured and unstructured content and publish it to the WebSphere Portal to provide users with fresh, relevant information. The new dCM Integration Kit provides J2EE integration between FatWire's dynamic content management software, UpdateEngine, and IBM's WebSphere Portal v4.1 to allow business users to easily create, edit and publish content to WebSphere Portals. FatWire's six new dCM content-rich display portlets include The User Display Portlet, The Launch Pad Portlet, The News Portlet, The Job Listing Portlet, The Document Portlet, and The Advertisement Portlet. The dCM Integration Kit and six portlets are available from the IBM Portlet Catalog (Code: 1WP1000CK). www.ibm.com, www.fatwire.com

INTERWOVEN OPENS ITALIAN OFFICE

3/11/2003

Interwoven, Inc. announced the official opening of its office in Milan, Italy. The new Interwoven Italian office is located at Via Monte di Pietà, 21, 20121 Milano, Italia; Tel: +39 02 86 33 7 237. www.interwoven.com

SCENE7 INFINITE IMAGING PLATFORM INTEGRATES WITH IBM DB2 CONTENT MANAGER

3/11/2003

Scene7, Inc. announced the integration of its Infinite Imaging Platform with IBM DB2 Content Manager Version 8.2. IBM and Scene7 jointly market Scene7's platform with IBM software and hardware platforms. Scene7 clients can computer generate any digital image they need -- in photo quality -- for uses including Websites, catalogs, partner extranets and marketing collateral. Scene7 integration includes the ability to publish dynamically generated imagery, including related image attributes and server requests, to IBM DB2 Content Manager v8.2 as a secure repository. Within Content Manager's standard e-client, users can directly view all published images and related metadata. Using Scene7's enhanced e-client, these users can search, browse, edit and transform image assets directly from IBM Content Manager. In support of this integrated solution, Scene7's platform operates on all major IBM environments, including DB2, WebSphere Application Server, Linux, AIX, Portal Server, and eServer hardware (xSeries, pSeries). www.scene7.com www.ibm.com/software/data/cm

DOCUMENT SCIENCES INTRODUCES XPRESSION

3/10/2003

Document Sciences Corporation announced the release of xPression, a content processing services architecture. xPression extends the capabilities of enterprise Content Management, CRM, Portal and Business Application systems through a J2EE/XML component-based architecture. xPression uses XML to connect an organization's existing workflow components. xPression works with Microsoft Word to design documents based on business rules. xPression can communicate with any solution, customize content at a high level of granularity, generate output through multiple channels and handle both transactional and high-volume batch processes. It runs on an IBM WebSphere, and is compatible with Windows 2000, Solaris, AIX, and HP/UX platforms. Support for BEA WebLogic, and z/OS will be added later this year. Xpression is Unicode compatible. It integrates with Oracle, DB2, and SQL Server. Xpression's Upgrade Utility allows companies currently using Document Sciences' Autograph DLS product to migrate their applications to the xPression environment if application server operability is desired. Supported output formats include: PDF, HTML, Postscript, AFP, Metacode, and PCL. www.docscience.com

MICROSOFT RELEASES "OFFICE SYSTEM" BETA 2

3/10/2003

Microsoft Corp. began distributing a half-million copies of the beta 2 version of the new Microsoft Office System to customers and partners worldwide. Seven Office System products are provided in the beta evaluation kit, including beta 2 versions of the new Microsoft Office Suite (Word, Excel, Outlook, PowerPoint and Access), the two new Office System additions, Microsoft Office InfoPath, Microsoft Office OneNote, as well as Microsoft Office FrontPage, Microsoft Office Publisher, Microsoft Windows SharePoint Services and Microsoft Office SharePoint Portal Server "2.0.". The products were developed with four key design goals in mind: information in-

telligence, process management, effective teaming, and personal impact. The release of the Microsoft Office System to beta 2 testing enables Microsoft business partners to build information worker solutions that incorporate collaboration and portal capabilities and improved desktop tools. Business partners also can begin testing the new information and research services they will be able to offer as Web services. Microsoft Office System products are scheduled for release in mid-2003. Pricing and availability are unavailable. www.microsoft.com/office/preview/

FACTIVA SERVICES TO BE INTEGRATED INTO MICROSOFT OFFICE 2003

3/10/2003

Factiva and Microsoft, Inc. announced a relationship that will deliver news and business information from Factiva into the Microsoft Office 2003 suite. Factiva's first solution for Office 2003, Factiva News Search, will allow information workers to conduct research on Factiva's 8,000 sources directly from a report or presentation they're creating. For example, while drafting a competitive brief in Word 2003, competitive intelligence professionals can use Factiva News Search to quickly look up industry trends from newspapers, journals, and newswires from around the world and insert that research into their document. Factiva News Search will be integrated into the Research Task Pane of all Microsoft Office 2003 applications, including Word 2003, Excel 2003, Access 2003, Outlook 2003, and PowerPoint 2003. Factiva News Search is currently available as part of the Microsoft Office 2003 Beta 2 trial. Factiva's investment in a platform based on XML and Web services enabled its integration with the Microsoft Office 2003 system. While evaluating beta test feedback, Factiva and Microsoft will also be researching customer solutions that can be built on top of the Office 2003 System. www.factiva.com

ZYLAB ADDS XML SUPPORT

3/10/2003

ZyLAB announced that it has enhanced its ZylIMAGE data storage and records management & archival product suite by incorporating XML as the data format to store scanned and electronic documents. A key driver to incorporating XML data formats is to help ZyLAB customers better adhere to corporate governance legislation such as Sarbanes Oxley, which is designed to ensure organizations have clear policies, proper records management systems, and adequate audit trails for its enterprise-wide information. www.zylab.com

TRADOS ANNOUNCES TM SERVER

3/10/2003

TRADOS Incorporated announced that it would release, in late April, TRADOS TM Server, their next-generation Translation Memory technology. TRADOS TM Server is a central translation memory database server that can be accessed simultaneously by hundreds of translators and enables corporations to maintain extremely large translation memory databases. TRADOS TM Server processes tasks on translation memories with tens of thousands or millions of translation units equally fast, whether there are 2, 20 or 200 users accessing the database simultaneously. Microsoft SQL Server and Oracle database back-ends are supported on both Windows and Unix. TRADOS TM Server can be integrated into all common document and content management systems, including Documentum, Tridion, Astoria, Vignette and Interwoven. It can also be used to power custom workflow applications, translation portals, etc. TRADOS TM Server will be available for commercial deployment late April 2003. www.trados.com

IXIASOFT RELEASES XMP-ENABLED TEXTML SERVER 2.3

3/10/2003

IXIASOFT announced the general availability of TEXTML Server 2.3, the latest version of its native XML database. IXIASOFT unveiled features that enhance the performance and flexibility of TEXTML Server, including support for Adobe's XMP (Extensible Metadata Platform) framework, improved support for Namespaces, and a new WebDAV Server. Introduced by Adobe, the XMP framework enhances workflows so that content can be applied between print, Web, eBooks and other media. Adobe XMP is available as an open-source license and uses a W3C based RDF syntax, it provides an XML-based metadata framework for embedding metadata within application files. With its support of XMP, TEXTML Server 2.3 enables the extraction and indexing of metadata embedded inside images, PDFs or any other binary format. Since the metadata is embedded in a file, the information remains with the file throughout its life, facilitating the exchange of metadata between applications. XMP was designed to be extensible and provides the means for customizing metadata based on specific requirements. IXIASOFT has also added improved support for Namespaces, enabling the use of Namespace validation as part of the index definition process. www.ixiasoft.com

ATOMICA RELEASES GURUNET 4.0

3/6/2003

Atomica Corporate announced the immediate availability of GuruNet 4.0, an advanced version of its application for Windows PC's, and the successor to Atomica Slingshot. With a single click on any word or phrase on the screen, GuruNet delivers real-time and reference information, including dictionary and encyclopedia terms, thesaurus, acronyms, technical terms, company descriptions, stock quotes, biographies, news, sports, weather, maps, translation and more. GuruNet's technology displays an unobtrusive pop-up window on demand, without interrupting users' work. It works within any Windows application: e-mail, Office, enterprise apps and web browsers. Answers are activated by an alt-click on any word on the screen; GuruNet provides an instant definition, overview or snapshot description. GuruNet displays pre-compiled and organized reference data from over 50 premium content sources. Among its improvements are: user interface innovations; richer, more dynamic content; expanded customization; instant audible word pronunciation; and an advanced power query capability. In addition to single end-user licenses, the company offers educational, corporate and large-organization group discounts. www.atomica.com

IMANAGE ADDS ECM & HUBBARD ONE AS PARTNERS

3/5/2003

iManage, Inc. announced two partners in the legal and corporate markets: eCM, a provider of outsourced electronic content management and cataloging systems, and Hubbard One, a provider of software to legal departments and law firms. iManage partners resell, deploy and build products and add-ons that complement the iManage WorkSite suite, which delivers document management, collaboration, workflow and knowledge management accessible through an integrated portal in a single integrated Internet solution. eCM provides complete outsourced content management and cataloging solutions for corporate and corporate legal clients. The firm converts content to structured formats to prepare it for electronic distribution applications including e-commerce, electronic publishing and online information services. Hubbard One develops intranets, extranets and Web sites for corporate legal departments and law firms increase. Hubbard One will work with customers to integrate the iManage WorkSite suite with their own software suites: LegalConnect for corporate legal departments and FirmConnect for law firms. www.electrictcm.com, www.hubbardone.com, www.imanage.com

SOFTWARE AG ADDS DIGITAL SIGNATURE SUPPORT TO TAMINO

3/5/2003

Software AG, Inc. announced the availability of a new software component for its Tamino XML Server 4.1 that allows the native XML server to validate digital signatures in support of secure electronic communications. Tamino XML Security Extensions makes the completion of legally binding contracts and other sensitive transactions over the Internet more secure and convenient. Digital signature technology allows the author of an electronic document to digitally "sign" the document and check later to see if the document has been modified without permission. Tamino XML Security Extensions automatically checks whether an electronic document has been given a digital signature, and verifies the validity of the signature during the database storage transaction for the signed document. A Developer Community download for Tamino XML Security Extensions is available. <http://developer.softwareag.com/tamino/download.htm>

COREL ANNOUNCES WORDPERFECT OFFICE 11

3/5/2003

Corel Corporation unveiled the new version of WordPerfect Office 11, scheduled for North American availability in spring 2003. WordPerfect Office 11 includes new Document Map and Collaborative Review functions, Reveal Codes, and RealTime Preview. A new Classic Mode features the WordPerfect 5.1 user environment, complete with a blue screen interface and WordPerfect 5.1 keystrokes. An enhanced legal toolbar adds features from WordPerfect Law Office Editions, including the Pleading Wizard, Concordance macro tool and Publish to EDGAR. Improved XML integration gives users the ability to publish slide shows, spreadsheets and WordPerfect files to XML. File sharing capabilities include Publish to PDF features in WordPerfect and Presentations, suite-wide Publish to HTML capabilities, and support for ODMA (Open Document Management API). New email Document Routing capabilities and Microsoft Outlook Address Book connectivity lets users collaborate via Microsoft Outlook. Set for availability in late April, WordPerfect Office 11 will be offered in Standard and Professional editions, with the Professional edition as a license-only offering. WordPerfect Office 11 Standard edition has a suggested retail price (SRP) of \$299.99US. Existing users of WordPerfect Office will be eligible to purchase WordPerfect Office 11 Standard at the upgrade SRP of \$149.99US. A special edition to serve the education market will be available at an SRP of \$99.99 US. www.corel.com

PROGRESSIVE INFORMATION TECHNOLOGIES & ROCKLEY GROUP FORM PARTNERSHIP

3/4/2003

Progressive Information Technologies announced a new partnership with The Rockley Group. Progressive and The Rockley Group will work together in an educational forum to be able to offer organizations the best content management solutions possible for their publishing needs. Vasont 8.0, Progressive's content management system for cross-media publishing, solves the challenge of repurposing content across all media - print, Web, CD-ROM and wireless channels - while ensuring enterprise-wide content integrity. www.vasont.com, www.rockley.com

CROWNPEAK ADDS EMAILLABS' EMAIL SOLUTION TO ADVANTAGE CMS

3/4/2003

CrownPeak Technology unveiled AdvantageMail, a complete email campaign solution that provides customers with interface components for rich media email creation, publishing, and management, all incorporated into CrownPeak's Advantage CMS. The AdvantageMail system draws

on a partnership with EmailLabs, a provider of advanced email marketing technology solutions. Like CrownPeak, EmailLabs follows the Software-as-a-Service (SaaS) model. CrownPeak's AdvantageMail system enables customers to integrate Website workflow with both email and site content. Advantage CMS customers can easily create, design, and publish an email. Simultaneously, the system can automatically update the site, create a customized landing page or even automatically create an updated archive of the newsletter listing. CrownPeak customers can manage their email campaigns from the Advantage CMS interface. Users can run their rich-media email through workflow like any other document. When ready to publish, the system automatically feeds the HTML, text and AOL-friendly versions directly into an EmailLabs system for distribution and tracking. www.crownpeak.com, www.EmailLabs.com

DATAMIRROR ANNOUNCES "GO LIVE FOR \$55K"

3/4/2003

DataMirror announced a special data integration solution package aimed at small and medium sized businesses - "Go Live for \$55K." Comprised of DataMirror LiveIntegration software, implementation, training, maintenance and support, this complete data integration solution empowers companies to integrate their incompatible systems and databases in order to achieve real-time, enterprise-wide data flows. This offering for \$55,000 (USD) is only available until April 30th, 2003. DataMirror's family of LiveIntegration software - including Transformation Server, DB/XML Transform, Constellar Hub, LiveConnector and iDeliver-lets companies bridge the gap between disparate systems, applications and relational databases. Prices may vary based on local currency and services. Special terms and conditions apply. www.datamirror.com/integratenow

REALOBJECTS RELEASES JAVA BEAN SDK

3/4/2003

RealObjects released their new JavaBean SDK. The SDK enables Developers to embellish their applications with the same rich-text WYSIWYG XHTML/XML editing functionality as the new edit-on Pro. This Bean is easily integratable into AWT- or Swing-based Java applications, and is compatible with existing server platforms. It empowers web developers to expand their applications to build distributed object-oriented business applications using documented APIs, combine the SDK with components from different vendors, and apply reusable code. One edit-on JavaBean SDK license is required per developer. A developer may use the license on multiple machines as long as they are not used concurrently nor used by another developer. There are no run-time royalty fees. The licensee is granted a non-exclusive and non-transferable license to use the JavaBean SDK for the purpose of creating and redistributing the Application Program(s). A single Developer License is 7,500 Euros, 5 Developer Licenses are 25,000 Euros. Volume pricing available upon request. www.realobjects.de

WINDFIRE ANNOUNCES VIRTUAL REPOSITORY MANAGEMENT TECHNOLOGY

3/4/2003

WindFire Technology announced the availability of its virtual repository management (VRM) technology which allows IT organizations to view, access and manage both structured and unstructured enterprise information as though they were in a single repository. As a core component of WindFire's Xtorm Data ServerT, the VRM technology allows data access and integration across disparate repositories including Oracle, IBM DB2, Microsoft SQL Server, legacy (mainframe) platforms, proprietary platforms and file systems. Designed to be highly scalable and secure, the Xtorm Data Server can be used to tie information together regardless of the data type, location, or application source. The Xtorm Data Server is a central component of WindFire's en-

terprise information management (EIM) offering which includes enterprise content management, business process management and virtual repository management applications. Developed on a J2EE, XML and Unicode foundation, the integrated suite is designed to be embedded into existing business applications, thus it is targeted to system integrators, ISVs and in-house IT developers. www.windfiretechnology.com

IUPLOAD & WEBPARTZ MERGE

3/4/2003

iUpload, a provider of content management services is merging with WebPartz, a provider of hosted e-business component applications. The new company will transact business under the iUpload brand, and expand its service to manage content from staff, application and visitor-generated. WebPartz and iUpload form an ASP delivering content management services across all categories. iUpload adds tools from WebPartz to manage content from applications and web site visitors. These functions include online surveys, e-mail marketing, newsletter publishing, site wide search and wireless applications. The new company provides these functions as "partz," allowing marketing departments to elevate a brochure-ware web presence to an e-Business site by blending partz into a solution that best meets their needs. The content management function will serve as the interface to the other partz, so instead of four or five separate applications, the iUpload interface is the hub for using the other partz. www.webpartz.com, www.iupload.com

BEA ANNOUNCES WEBLOGIC PLATFORM 8.1

3/3/2003

BEA Systems, Inc. announced BEA WebLogic Platform 8.1. BEA WebLogic Platform 8.1 is based on a unified architecture that includes new versions of BEA's application server and Java virtual machine, BEA's enterprise portal and application development framework, and BEA's integration solution. Another component of BEA WebLogic Platform 8.1, BEA WebLogic Portal 8.1 is designed to simplify the production and management of custom-fit portals. BEA's new portal product supports JSR 168, for portlet-level development, and JSR 170, for content management system integration. BEA WebLogic Portal 8.1 also will deliver a new portal design and development environment, based on BEA WebLogic Workshop. In addition, BEA WebLogic Portal 8.1 adds new content management capabilities, search, Web integration, and native support for wireless communications. BEA's application server will implement both Web services security and an infrastructure for delivering SOAP messages. A beta version of BEA WebLogic Platform 8.1 is available. General availability of BEA WebLogic Server 8.1 and BEA WebLogic JRockit 8.1 is scheduled for the spring. General availability of the BEA WebLogic Platform 8.1, which includes BEA WebLogic Workshop 8.1, BEA WebLogic Integration 8.1, BEA WebLogic Server 8.1 and BEA WebLogic Portal 8.1, is scheduled for the summer. BEA WebLogic software is available for download. www.bea.com

COAST ADDS CONTENT QC TO MICROSOFT CONTENT MANAGEMENT SERVER

3/3/2003

COAST Software Inc. announced the integration between its enterprise-based COAST Web Quality Central solution and Microsoft Content Management Server 2002. COAST Web Quality Central is a complementary solution that integrates with Microsoft Content Management Server 2002 to automate content testing and verification. Automated scans verify content based on user-defined rules within the workflow process. Pass/fail reports are then returned to the author or quality assurance team and before a document can be passed on to the next re-

cient or published to the Web site, any errors or issues are addressed. COAST Web Quality Central can also be used when organizations first deploy a CMS system and they have mountains of content that must be migrated into the content management repository. Using the solution from COAST Software, they can automatically scan through that content, identify any old, inappropriate or updated information and then take necessary action. www.coast.com

FATWIRE & GRAVITYROCK PARTNER FOR CONTENT MANAGEMENT IN EUROPE

3/3/2003

FatWire Software announced a strategic partnership with GravityRock, where they will be FatWire's authorized distributor and support center for FatWire's complete software line throughout Europe. GravityRock will distribute and resell FatWire's UpdateEngine enterprise content management suite and FatWire Spark portal Content Management (pCM) software. GravityRock has already implemented UpdateEngine solutions both at customer sites as well as in their own technology center, PortalRock.net. Portalrock.net offers rapid prototyping as a tool for decision makers to evaluate Portal and Content Management solutions prior to making final choices and investments. Additionally, GravityRock will work with portal server providers to sell FatWire Spark portal Content Management (pCM) software. www.gravityrock.com, www.fatwire.com

QL2 LAUNCHES WEBQL VERSION 2

3/3/2003

QL2 Software (formerly Caesius Software) announced the release of WebQL version 2 for information extraction from the Web and other unstructured data sources. WebQL is used for business intelligence, competitive intelligence, knowledge management, and business activity monitoring applications. WebQL version 2 includes: notation based on ANSI standard SQL including full support for joins, grouping, sorting, and set operations; an IDE includes syntax highlighting editor, graphical data flow monitor, and real-time delivery of results; support for Windows, Linux, and Unix; support for "all modern" technologies for parsing and extracting unstructured, semi-structured and structured data; and simultaneous and uniform input and output of all common file formats HTML, XML, PDF, DOC, CSV, TAB, images, databases, proprietary email formats, etc. WebQL version 2 is available for immediate shipment. Prices range from \$7,900 per year for a single processor license to \$100,000 per year for an unlimited usage single location license. Low cost licenses are available to educational institutions and charitable organizations. No-charge evaluation licenses are available to software developers. www.QL2.com

RECENT ISSUES

Issues from 1993 thru 1998 are \$15 *if* in print. More recent issues are available in PDF for various prices and *may* be available in print form for \$30. See www.gilbane.com or call for more information.

Volume 11 — 2003

Number 1 **Information Integration, Objects, Content Services & Infrastructures**, *Frank Gilbane*

Volume 10 — 2002

Number 10 **Corporate Portals – Success Kills the Market**, *Frank Gilbane*

Number 9 **XML 2002 – More Than Just Another Show**, *Sebastian Holst*

Number 8 **The Role of XML in Content Management**, *Lauren Wood*

Number 7 **Searching for Value in Search Technology**, *Sebastian Holst*

Number 6 **SVG – The Future of Web Rendering?**, *Bill Trippe*

Number 5 **A Framework for Understanding the Information Management Market**,
Jared Spataro, Bryan Crow

Number 4 **The Top 10 Trends in Content Management**, *Frank Gilbane*

Number 3 **In Search of Search Solutions**, *Sebastian Holst*

Number 2 **The Many [Inter]Faces of Content Management Systems**, *Rita Warren*

Number 1 **What is an Information Model, and Why Do You Need One?**, *JoAnn Hackos*

Volume 9 — 2001

Number 10 **An Alternative Model for Personal Information Management**,
Girish Altekar

Number 9 **Who Should Own Your Content Management System?**, *Bob Boiko*

Number 8 **Understanding Web Services**, *Sebastian Holst*

Number 7 **Editorial Interfaces & Enterprise-enabled Content**, *Bill Trippe & David R. Guenette*

Number 6 **Why Content & XML Integration Technologies are Fundamental**,
Frank Gilbane

Number 5 **The Application Server Cometh, II**,
Bill Trippe

Number 4 **Open Source Content Management Systems: A Parallel Universe?**,
Sebastian Holst

Number 3 **Privilege Management & Rights Management for Corporate Portals**, *David R. Guenette, Larry Gussin, and Bill Trippe*

Number 2 **Choosing an Architecture for Wireless Content Delivery**,
Girish Altekar, Regan Coleman

Number 1 **XHTML: What You Should do About it, and When**,
Sebastian Holst, David R. Guenette

Volume 8 — 2000

Number 10 **XML: The State of the Union** *Bill Trippe, David R. Guenette*

Number 9 **E-books: Technology for Enterprise Content Applications?**
Bill Trippe, David R. Guenette

Number 8 **What is Content Management?** *Frank Gilbane*

SUBSCRIPTION FORM

You can also order on our *secure* website www.gilbane.com.

Please start my electronic subscription to the Gilbane Report for **only \$99**. (10 issues/year). Subscription includes access to HTML and PDF versions at www.gilbane.com. (Call for print subscriptions, site license prices, and back issues.)

I am eligible for an affiliate discount* _____ Affiliate organization _____ Tracking # _____

My check for \$_____ is enclosed please bill me
Please charge my credit card MasterCard Visa American Express

Name as on card: _____ Number _____
Signature _____ Expiration date _____

Name _____ Title _____
Company _____ Department _____
Address _____
City _____ State/Province _____ Zip/Postal Code _____
Country _____ Tel. _____ Fax _____ E-mail _____

Checks from outside the U.S. should be made payable in U.S. dollars.

Funds may be transferred directly to our bank, please call for details.

Mail this form to: Bluebill Advisors, Inc. 763 Massachusetts Ave., Cambridge, MA 02139, USA.

You can also place your order at www.gilbane.com or by phone (+617.497.9443), or fax (+617.497.5256).

CALENDAR

XML Europe 2003. May 5-8th, 2003, *Hilton London Metropole, London*. The largest, longest-running XML event in Europe. XML Europe is a European forum for the XML community, spanning the worlds of electronic business, publishing, Internet, e-government, software and open standards development. XML Europe's remit is traditionally broad: mixing the newest technology with the latest thinking in the business and legal issues of information management. **Gilbane Report subscribers receive \$100 off the conference registration rate!** <http://www.xml europe.com/>

Seybold-Gilbane Content Management Intensive, Wednesday, June 11, *RAI Centre, Amsterdam*. Our one-day intensive update helps you get started confidently on new content management strategy, and helps you checkpoint your current implementation against what other businesses are doing and what can now be accomplished. See www.gilbane.com/events.html for more information on this and other events.

The Gilbane Conference on Content Management @ Seybold. September 9-10, exhibits September 9-11, *Moscone Center, San Francisco*. Our annual conference is *the* place to go to learn what is happening in content management from the industry thought leaders. We cover technologies, best practices, and market trends. The event is designed for beginners, experienced implementors, and IT and content strategists from all vertical industries. Co-located with our partners at Seybold Seminars. See www.gilbane.com/events.html for more information on this and other events.

© 2003 Bluebill Advisors, Inc. - all rights reserved. No material in this publication may be reproduced without written permission. To request reprints or permission to distribute call +617 497.9443. The Gilbane Report is a registered trademark of Bluebill Advisors, Inc. Product, technology, and service names are trademarks or service names of their respective owners.

The Gilbane Report is published 10 times a year (monthly, except for August and December). The Gilbane Report is an independent publication offering objective analysis of technology and business issues. The Report does not provide advertising, product reviews, testing, or vendor recommendations. We do discuss product technology that is appropriate to the topic under analysis, and welcome product information from vendors. Letters to the editor are encouraged. Send to: editor@gilbane.com. Visit our web site at www.gilbane.com

ISSN 1067-8719

Sistemas de información

Diseño y creación de la base de datos documental del Grupo Godó

Por Carles Salmurri; Ernest Abadal; Mònica Sancho; Miquel Llevat; Andreu Sulé y Maita Corbera

Resumen: Se describe y analiza el proyecto de diseño y creación de la base de datos corporativa del Grupo Godó. En primer lugar, se exponen las características de la situación de partida que originó la necesidad del proyecto. A continuación, se explican con detalle las características de la base de datos: su estructura, colecciones que integra, tratamiento documental, flujos de información e interfaces de consulta. Finalmente, se presenta una evaluación sobre el funcionamiento de la nueva base de datos hasta el momento presente y las iniciativas de futuro con las que se pretende mejorar su implantación y funcionamiento.

Palabras clave: Medios de comunicación, Centros de documentación, Bases de datos, Informatización, Grupo Godó.

Title: Design and development of the Grupo Godó's document database

Abstract: The article describes the process by which the corporate database of the mass media group, Grupo Godó, was designed and created. First the author reviews the prior situation that gave rise to the need for the project. Thereafter the characteristics of the database are described in detail: structure, contents, documentary analysis, information flow, and search interfaces. The article concludes by evaluating the functionality of the new database up to the present and outlining future initiatives that are expected to improve both the system's functionality and implementation.

Keywords: Databases, Automation, Mass media, Grupo Godó, Document centres.

QUIZÁ SEA UN MEDIO DE COMUNICACIÓN uno de los casos que concentran mayor número de complejidades en lo que se refiere a la organización y gestión de la información. Como es bien sabido, sus centros de documentación manejan todo tipo de

información (texto, fotografía, sonido y vídeo), en grandes cantidades y procesan un alto número de consultas que se han de resolver en un tiempo de respuesta lo más breve posible. Estas tres características condicionan especialmente las operaciones docu-

mentales que se deben realizar. Así pues, se ha de disponer de precisos criterios de selección que permitan jerarquizar el interés de los numerosos y diversos tipos de documentos.

Por otro lado, las operaciones de análisis han de permitir el acceso a la información no sólo desde el punto de vista descriptivo, sino especialmente por el contenido, con lo que será muy relevante contar con lenguajes documentales específicos. Finalmente, para la recuperación de la información será necesario disponer de sistemas informáticos que permitan almacenar y gestionar con rapidez los enormes volúmenes de datos que pueden reunirse en un medio de comunicación y de programas que faciliten el acceso al contenido de la información.

En este texto vamos a describir y analizar la experiencia de diseño y creación de la base de datos corporativa del Grupo Godó, un im-



Figura 1. Estructura del sistema de información documental. Situación inicial

portante grupo de comunicación español. De hecho, en una lectura rápida y sin profundizar, no se aprecia que exista ninguna singularidad en este proyecto, ya que se hace referencia a operaciones y procesos comúnmente utilizados en cualquier centro de información y documentación. La particularidad reside en la complejidad que comporta trabajar en un entorno en el que se dan cita a la vez las tres características a las que antes aludíamos: variedad de documentos y formatos, volumen de información y rapidez en la respuesta.

«El propósito general era crear un sistema documental único para gestionar la información del grupo»

El *Grupo Godó (GG)* es un grupo multimedia español constituido en 1998 que se articula en torno al diario *La Vanguardia* y que tiene presencia en el sector de la prensa deportiva (*Mundo deportivo*), en revistas (*Interiores, Historia y vida, etc.*), radio (*RadioCat, RAC1*) y televisión (*CityTv*). El proyecto al cual vamos a hacer referencia se inició en septiembre de



Figura 2. Estructura del sistema de información documental. Situación actual

1999 y finalizó en abril de 2001. Surgió de la necesidad, ya detectada a partir de la constitución del GG, de disponer de un sistema de información único y compartido por todas las empresas que lo constituyen, centralizando la información y optimizando los recursos. En él participaron diversos equipos de trabajo, y constó de las siguientes fases:

—Análisis de la situación (septiembre – diciembre 1999). Se realizó un estudio con profundidad

del sistema de información documental existente a fin de detectar las necesidades.

—Diseño (abril – septiembre 2000). Una vez conocido lo que se requería, se pasó a describir con detalle las características y estructura de la futura base de datos, sus colecciones, interfaces, etc., y a detallar los requerimientos técnicos para poder analizar y valorar las distintas soluciones tecnológicas. También se realizó un estudio sobre cómo tenían que ser los flujos de información, elaborándose un tesoro para facilitar el acceso al contenido de los documentos. El resultado final fue un manual de procedimiento que recogía todas las especificaciones técnicas del proyecto.

—Implementación (octubre 2000 - febrero 2001). Una vez seleccionados el sistema y el programa informáticos, se procedió a configurar las especificaciones del manual de procedimiento sobre el programa escogido.

2. Análisis

En este apartado se va a describir brevemente la situación de partida, con el fin de mostrar las limitaciones existentes, y a definir las

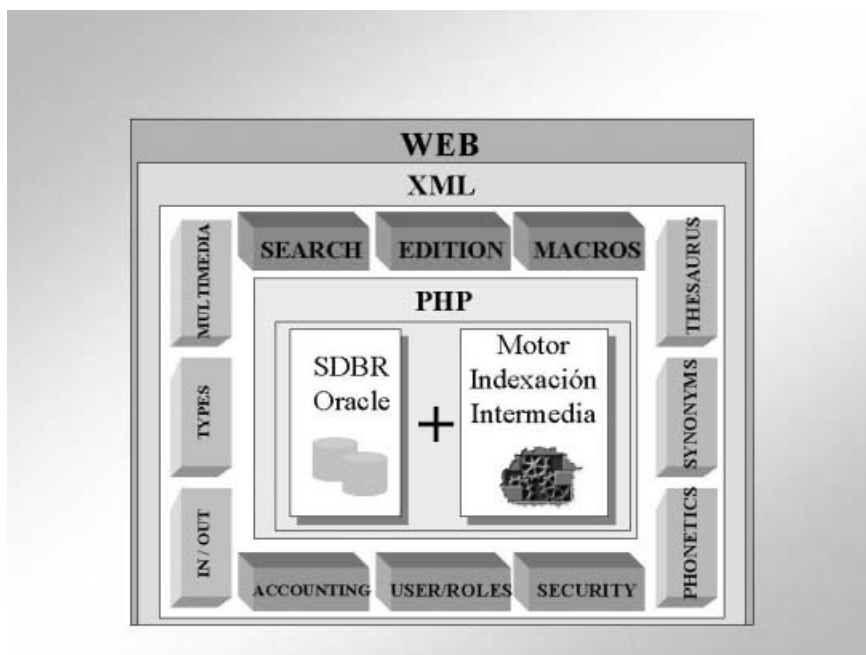


Figura 3. Diseño de la base de datos

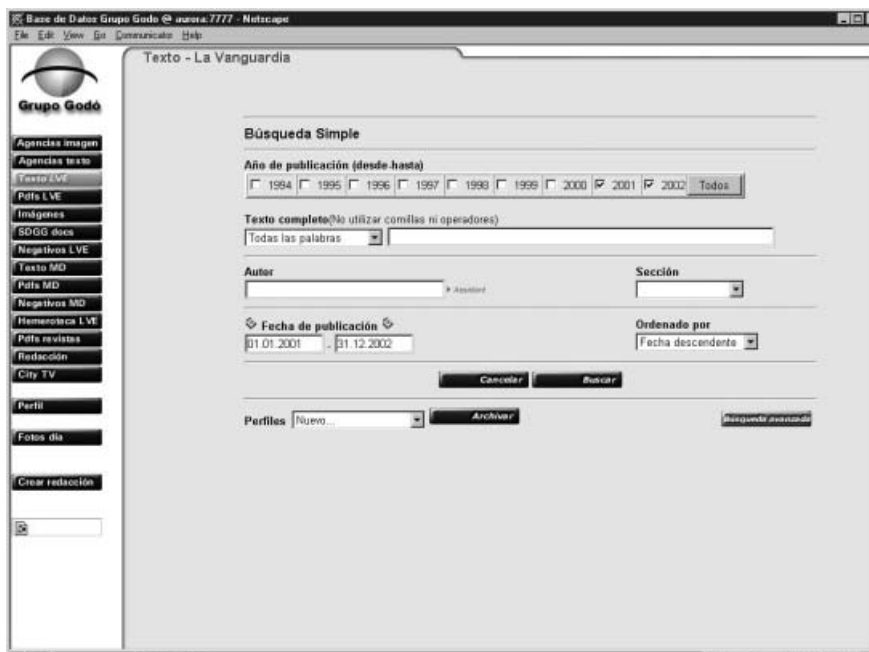


Figura 4. Página de consulta básica

líneas generales del proyecto que intenta subsanarlas.

En 1999 el sistema de información de *La Vanguardia* estaba gestionado por el programa *Topic (Verity)* y contenía unos 500.000 documentos estructurados en 6 colecciones: textos y fotos de producción propia de *La Vanguardia* y de *Mundo deportivo*, y de agencias, así como la documentación elaborada por el *Servicio de Documentación de La Vanguardia (Sdlv)* que sólo era accesible a usuarios internos. Por otro lado, se encontraba la base de datos de *La Vanguardia digital*, que contenía las informaciones publicadas en el periódico durante los últimos 6 meses y que era (sigue siendo) accesible para usuarios externos.

<http://www.verity.com>

Con el proceso de análisis se pudieron detectar las necesidades y limitaciones respecto de las entradas, tratamiento, sistema de recuperación y distribución de la información recibida y producida por las distintas empresas del GG. Fue una larga y ardua fase que supuso la realización de numerosas entrevistas así como el estudio de diversos y complejos procesos de trabajo. Es por ello que resulta difícil re-

sumir en poco espacio el resultado de este análisis. De todas formas intentaremos destacar los aspectos más sobresalientes.

A continuación vamos a exponer los objetivos generales y específicos que se establecieron para el proyecto, cada uno de los cuales intentaba subsanar alguna deficiencia, limitación o necesidad detectada. De esta forma se podrá tener una visión global y resumida de cuál era la situación de partida y cómo se dibujaba la estación de

llegada. Las figuras 1 y 2 intentan sintetizar y mostrar de forma gráfica estos dos escenarios.

El propósito general, como se ha avanzado anteriormente, era crear un sistema documental único para gestionar la información del grupo. De entre los objetivos específicos destacaremos los siguientes:

—Centralizar la información e integrar documentos. El nuevo sistema contiene todo tipo de recursos (textos, imágenes, vídeo, sonido) y de todos los medios de comunicación de la empresa, estableciendo criterios de selección de los documentos que han de formar parte de la base de datos. Está formado por varias bases de datos coordinadas e interrelacionadas que incluyen las informaciones textuales, gráficas, sonoras, etc., producidas por los diferentes medios de comunicación.

—Optimizar e integrar los recursos. La información se encontraba almacenada por duplicado en diversos sistemas. Un ejemplo son las noticias de agencia que se almacenaban en dos distintos: *Wire Center* y base de datos. Por otro lado, coexistían bases de datos similares para las diferentes empresas

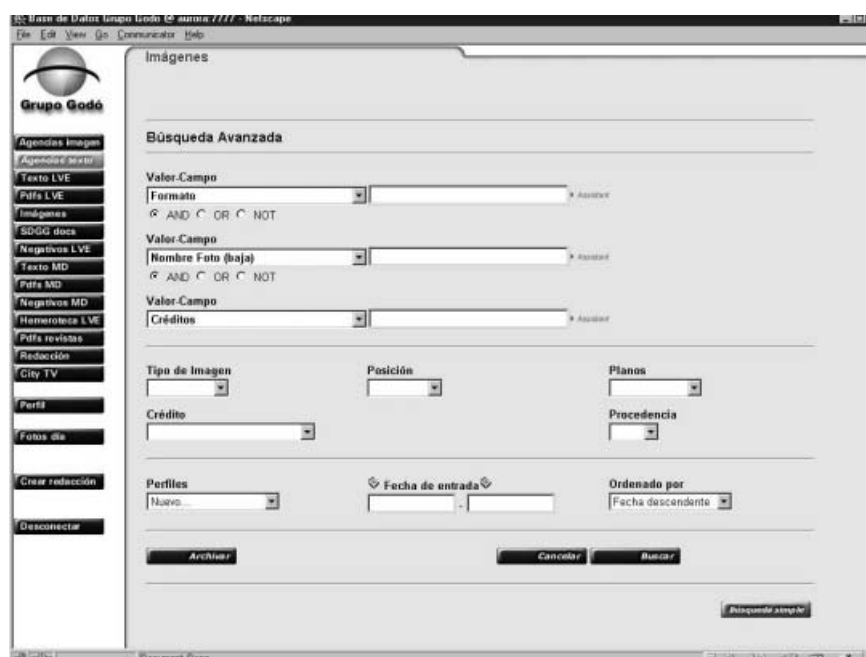


Figura 5. Página de consulta avanzada

del grupo; es el caso de *La vanguardia* impresa y su homónimo en internet, que habían creado herramientas separadas con unos contenidos muy similares. Curiosamente, además, tenían licencias independientes del mismo sistema de recuperación (*Topic*). El nuevo modelo se define como corporativo y tiene como prioridad elaborar productos y servicios compartidos.

«Uno de los aspectos del proyecto consiste en estudiar con detalle los derechos de autor de los textos y las fotografías ya publicadas para establecer un sistema de pago y distribución»

—Mejorar y simplificar los flujos. La transferencia de los documentos textuales desde la redacción (utilizaban el sistema *Hermes*) a la base de datos no era totalmente satisfactoria ya que no se producía un ajuste exacto a su estructura de campos, lo que suponía un trabajo de revisión considerable¹. Además no estaba sistematizado el proceso de transmisión de las fotografías, uno de los activos

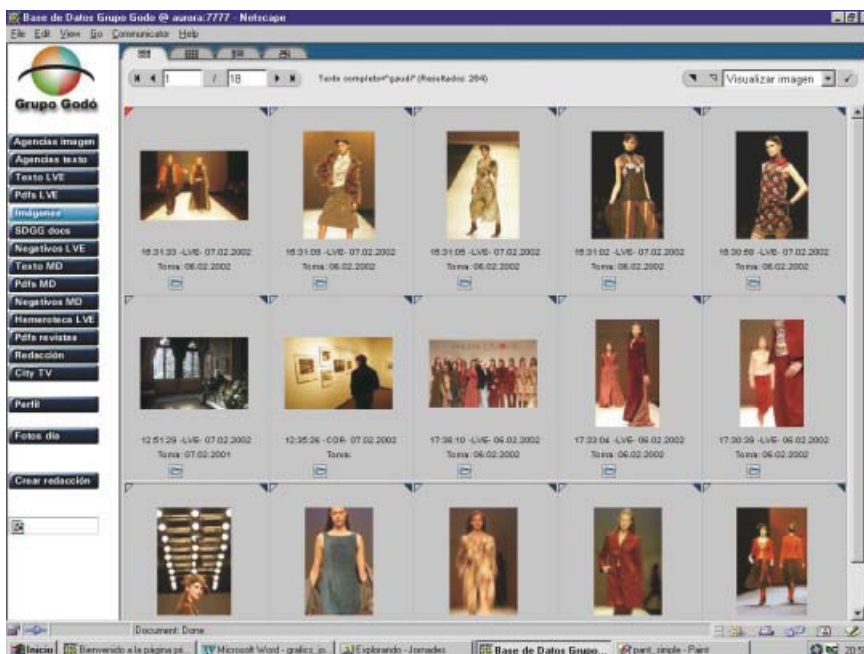


Figura 6. Página de resultados

documentales más importantes en un medio de comunicación. El nuevo sistema establece un circuito único de entrada y salida para los documentos.

—Mejorar y simplificar el proceso de análisis documental. Anteriormente no se podía realizar adecuadamente el análisis de contenido (asignación de descriptores y clasificación) de los textos. Por otro lado, en lo que se refiere a las imágenes (fotografías) se utilizaba una ficha documental muy básica

que permite el programa *Photoshop*, un sistema a todas luces insuficiente para permitir una buena recuperación de la información gráfica.

—Accesibilidad inmediata a la información tanto desde la red local como del exterior. La utilización de la base de datos no era posible por parte de los corresponsales y enviados especiales; el nuevo sistema facilita este acceso.

—Comercialización. La única vía de consulta posible de los datos retrospectivos para los usuarios externos (clientes) era a través del buscador disponible en *La vanguardia digital*, que permitía la recuperación de las informaciones publicadas en los últimos 6 meses en su edición impresa. Uno de los aspectos del proyecto consiste en estudiar con detalle los derechos de autor de los textos y las fotografías ya publicadas para establecer un sistema de pago y distribución.

3. Diseño del sistema de información

En este apartado se describen sus características generales: contenidos, estructura, interfaces de consulta y el sistema de recuperación de la información que utilizan.



Figura 7. Página de visualización de documento



energise your thinking



O V I D

En Ovid revolucionamos la información y transformamos contenidos, herramientas y servicios en una experiencia única para nuestros clientes.

Para recibir más información:

C/ Orense, 85 28020 Madrid

Madrid Tel: 91 567 84 83 • Barcelona Tel: 93 810 65 63

E-mail: spain@ovid.com

www.ovid.com

© 2002 Ovid Technologies, Inc. All Rights Reserved

1. Contenidos. Como se ha dicho, las empresas del GG en su conjunto producen una gran variedad de publicaciones de distintas características y trabajan con formatos también variados. La selección de los contenidos a incluir en la base de datos corporativa partió de los siguientes criterios:

- control de la producción propia,
- seleccionar qué contenidos se incluirían en la base de datos y en qué forma (a texto completo, referencias, etc.), y
- definir distintos niveles de descripción.

Cada una de las empresas participa en la selección de contenidos y en dicha definición en función de los recursos humanos existentes en cada caso. Ambas cuestiones constituyen la primera parte del diccionario de datos. En la definición de contenidos se establecen las bases para el control de los documentos sujetos a limitaciones de algún tipo (derechos de autor).

2. Estructura. Realizada la selección, nos encontramos ante tipologías de documentos muy diferentes que hacen necesario desarrollar diversos diccionarios de datos para cada una de las colecciones en las que se organizan los documentos en la base, conservando unos campos comunes que permitan la consulta global de contenidos.

La estructura de la base de datos permite representar todo tipo de documentos:

- Registros + texto.
- Registros + imagen fija (en alta resolución).
- Registros sonoros + sonido.
- Registros imagen + imagen en movimiento.

Cada registro consta de campos de control, descripción docu-

mental, contenido y gestión de derechos. Para los campos de descripción documental y de contenido se están desarrollando herramientas (listas de validación y tesoro) que permitan unificar criterios entre las distintas empresas del grupo y facilitar la búsqueda y recuperación de información a través de un único lenguaje de descripción (principio básico de uniformidad en el sistema de trabajo).

Para cada uno de los campos establecidos se especifican diversos parámetros entre los que podemos destacar la descripción del contenido, período cronológico, responsabilidad (automática o manual) e instrucciones para la entrada de datos.

3. Interfaz.

a. Estructura.

Sigue el mismo modelo en todas sus colecciones con la intención de facilitar su utilización, aunque las diferentes pantallas de visualización y funcionalidades disponibles se adaptan al tipo de documento ya sea éste imagen, texto, etc., y al perfil del usuario que accede al sistema.

«La realización de este proyecto puso de manifiesto la falta de compatibilidad entre los diferentes sistemas informáticos de la organización»

La página de inicio se divide en dos zonas: menú y trabajo. La primera, situada en el margen izquierdo de la pantalla, permanece fija durante toda la sesión y da acceso a las diferentes colecciones y a opciones generales a nivel de usuario, como la gestión de los perfiles de búsqueda, actualización automática de consultas y desconexión del sistema entre otras. Es en la zona de trabajo donde se interactúa en el momento de consultar las

Colección	Registros
Imágenes	144.344
Textos <i>La Vanguardia</i>	508.000
Textos <i>Mundo Deportivo</i>	191.472
Revistas	4.100
Negativos <i>La Vanguardia</i>	43.373
Negativos <i>Mundo D</i>	8.323
CityTV	5.721
Datos cronológicos	18.283
Total	923.525

Tabla 1. Contenidos de la base de datos

colecciones, crear nuevos documentos y ejecutar cualquiera de las opciones disponibles sobre los documentos.

b. Consulta.

Cada vez que se accede a una colección clicando sobre su nombre en la zona de menú, se abre una pantalla de consulta que permite su interrogación. Algunas bases de datos disponen de una segunda opción de búsqueda, la avanzada, que permite elaborar estrategias mucho más concretas.

—Simple: las pantallas ofrecen la posibilidad de interrogar la colección por cualquiera de sus campos. Para cada colección se ha establecido también una serie de recuadros de búsqueda para los campos considerados de especial relevancia (por ejemplo, en el caso de *Imágenes-fotos* se han escogido fotógrafo, créditos y fecha de toma).

—Avanzada: aquí se permite llevar a cabo búsquedas de forma más concreta que la anterior. La principal diferencia entre una y otra es que en ésta es posible realizar consultas específicas sobre un mayor número de campos y que la cantidad de elementos limitadores también es mayor.

c. Visualización.

Una vez elaborada y ejecutada la estrategia de búsqueda, el sistema muestra los documentos correspondientes a los criterios definidos. Pueden ser visualizados en varios formatos accesibles por medio de pestañas situadas en la parte superior de la pantalla:

—Abreviado de varios documentos (galerías).

—Visualización e impresión completa de un único documento (*print preview*).

—Edición de un documento (editar).

d. Funcionalidades.

Ya se ha mencionado que uno de los principios básicos que debía cumplir el nuevo sistema era el de la centralización de la información: un único circuito de entrada y salida para los documentos. A nivel de la interfaz esto implica incorporar características específicas para gestionar los flujos de trabajo además de los propios documentos.

«En referencia al proceso de implantación, hay que destacar la importancia de la formación de los usuarios»

Entre las funciones sobre los documentos encontramos de edición, eliminación, visualización, selección e impresión, así como algunas otras específicas como la edición de imágenes. Entre las prestaciones relacionadas con los flujos de trabajo se encuentra la exportación de objetos dentro o fuera del sistema y su control según los derechos de propiedad asignados a éstos. Conseguir la integración de los flujos de la información quizá sea uno de los aspectos especialmente interesantes del proyecto, ya que facilita la circulación de la información, permite ahorrar de

tiempo y redonda en la calidad del producto final.

El acceso a unas u otras funciones dependerá de la colección consultada y de los privilegios de cada usuario sobre la misma. Esto mejora la productividad de las distintas organizaciones del grupo ya que integra la gestión de los documentos en la producción en aspectos como la circulación de la información, ahorro de tiempo, mejora de la calidad del producto final, etc.

e. Perfiles de usuario.

Para acceder a la base de datos es necesario identificarse mediante un nombre y una contraseña. De esta forma el sistema asigna individualmente unos privilegios que permitirán realizar ciertas funciones y no otras. La complejidad de la gestión de estos perfiles en el proyecto viene dada por la participación de diferentes empresas y por el uso de la información que puede hacer cada una de ellas por un lado, así como los distintos niveles de usuarios a definir según sus funciones (redacción, compaginación, edición gráfica, dirección, documentación, etc.), por otro. Hay que tener en cuenta que, aunque la normalización y unificación de los procesos de trabajo sea deseable, cada una de las empresas del grupo se estructura según objetivos y prioridades distintas que se reflejarán tanto en la definición de los perfiles como de las funciones a realizar.

4. Sistema de recuperación de la información. *Digital Collections*, el producto que se ha adoptado, es una aplicación para gestionar información de prensa que está

construida sobre *Oracle* y dispone de la mayoría de las prestaciones necesarias para la creación y explotación de la base de datos documental corporativa del GG. Se trata de un programa muy adecuado para la creación del sistema de información global propuesto, ya que resuelve de forma integrada los circuitos y flujos de información.

<http://www.digicol.com/>

Su principal característica es la de facilitar el control de todos los documentos generados y recibidos dentro de un mismo entorno (flujos de gestión documental) así como el tratamiento de las imágenes de forma integrada. Posibilita, por tanto, una muy buena integración del producto con los procesos de producción de los medios de comunicación de la empresa. No podemos olvidar que uno de los principales requerimientos del proyecto consiste en facilitar la entrada y salida de información de todo tipo al sistema global (textos, noticias de agencia, fotografías, etc.). En este sentido, hay que apuntar que el programa resuelve de forma muy sencilla, rápida y transparente las conexiones (pasarelas) entre la base de datos y las diversas entradas de información (textos *Hermes*, noticias de agencia, fotografías, etc.).

La compatibilidad absoluta con *Oracle* permite, por otro lado, una fácil interrelación con otros sistemas de gestión de datos del grupo, como la gestión económica o el control de la publicidad, que funcionan también sobre este conocido *sgbd* relacional. Las otras opciones que también se analizaron (*Excalibur* y *Verity*), aunque resolvían mejor el acceso al contenido de los documentos, no permitían un grado de integración de aplicaciones tan óptimo.

Resumiendo, las opciones analizadas eran básicamente dos: utilizar un paquete integrado —con la

Acción	Cantidad
Consultas	32.000
Descargas de documentos	3.700
Exportaciones	2.500
Visualizaciones	41.000
Total	79.200

Tabla 2. Estadísticas de uso (mensuales)

ventaja de no tener que desarrollar pasarelas ya que integra muy bien los procesos y facilita la circulación de la información— u optar por un motor de búsqueda asociado a un sgbd, lo cual no resuelve la integración de todos los procesos en un único sistema de forma tan rápida y homogénea como el paquete integrado.

Más allá de la elección de uno u otro, la realización de este proyecto puso de manifiesto la falta de compatibilidad entre los diferentes sistemas informáticos de la organización y supuso una oportunidad para su reorganización. Esta situación queda ejemplificada en los gráficos de las figuras 8 y 9.

4. Conclusiones

Aunque es pronto para extraer datos concluyentes sobre el impacto del proyecto, vamos a presentar un balance sobre el grado de consecución de las metas planteadas, sobre el funcionamiento de la base de datos y también sobre la implementación y el mantenimiento del sistema.

Desde un primer momento, los principales logros se consiguieron en los dos primeros objetivos plan-

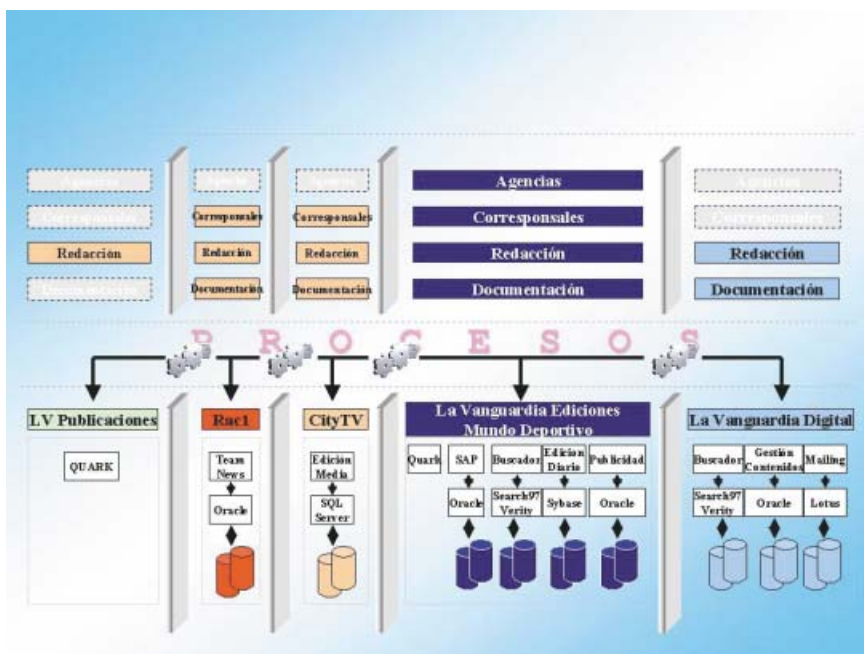


Figura 8. Situación inicial entre sistemas

teados: la centralización de la información e integración de los contenidos independientemente de su formato, así como la optimización de recursos entre las distintas empresas del GG. También se facilitó rápidamente el acceso interno y externo a los contenidos.

La integración de la base de datos corporativa con los distintos sistemas de producción (la “mejora y simplificación de flujos” aludida en los objetivos) ha sido más costosa por las implicaciones que su-

pone al nivel de organización del trabajo tanto de las personas como de los procesos y sistemas involucrados. Sin embargo la valoración es positiva por cuanto se ha creado un marco de reflexión y análisis que ha permitido mejorar en muchos aspectos los flujos de trabajo existentes antes de la implantación de la base de datos corporativa.

Un objetivo en el que se sigue trabajando actualmente es la uniformidad de los sistemas de trabajo, referida tanto a la adecuación de las herramientas de descripción a las necesidades como a la mejora de la coordinación y de los mecanismos de comunicación entre las diferentes empresas del grupo.

Respecto a la estructura y funcionamiento de la base de datos, la participación de los usuarios en la evaluación de la misma ha permitido introducir mejoras significativas en la usabilidad de la interfaz de búsqueda y ha exigido en algunos casos modificaciones en las pantallas de consulta y visualización para adecuarlas a perfiles de usuario concretos. Actualmente se trabaja en una ayuda en línea que resuelva las dudas más frecuentes y facilite el aprendizaje y resolución de errores durante la interac-

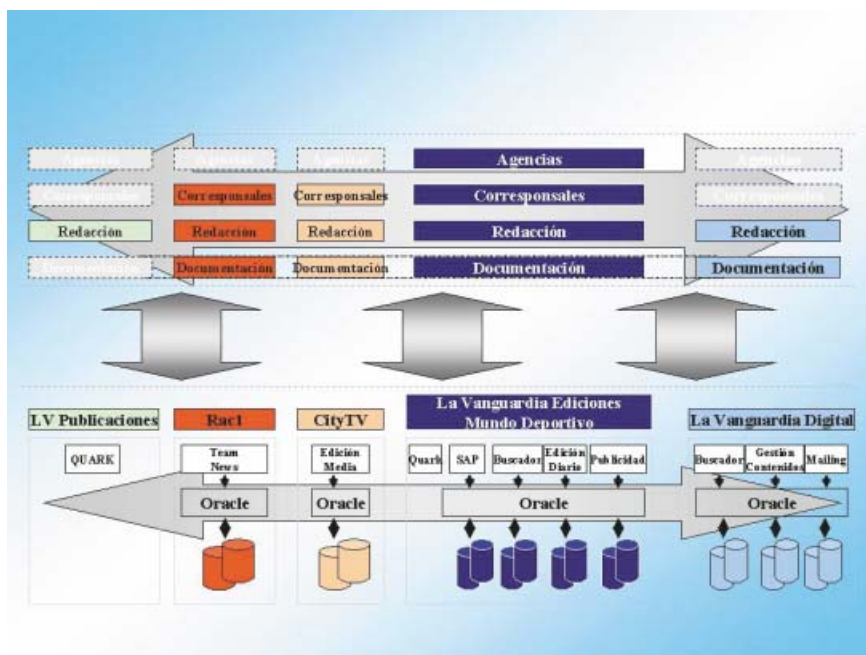
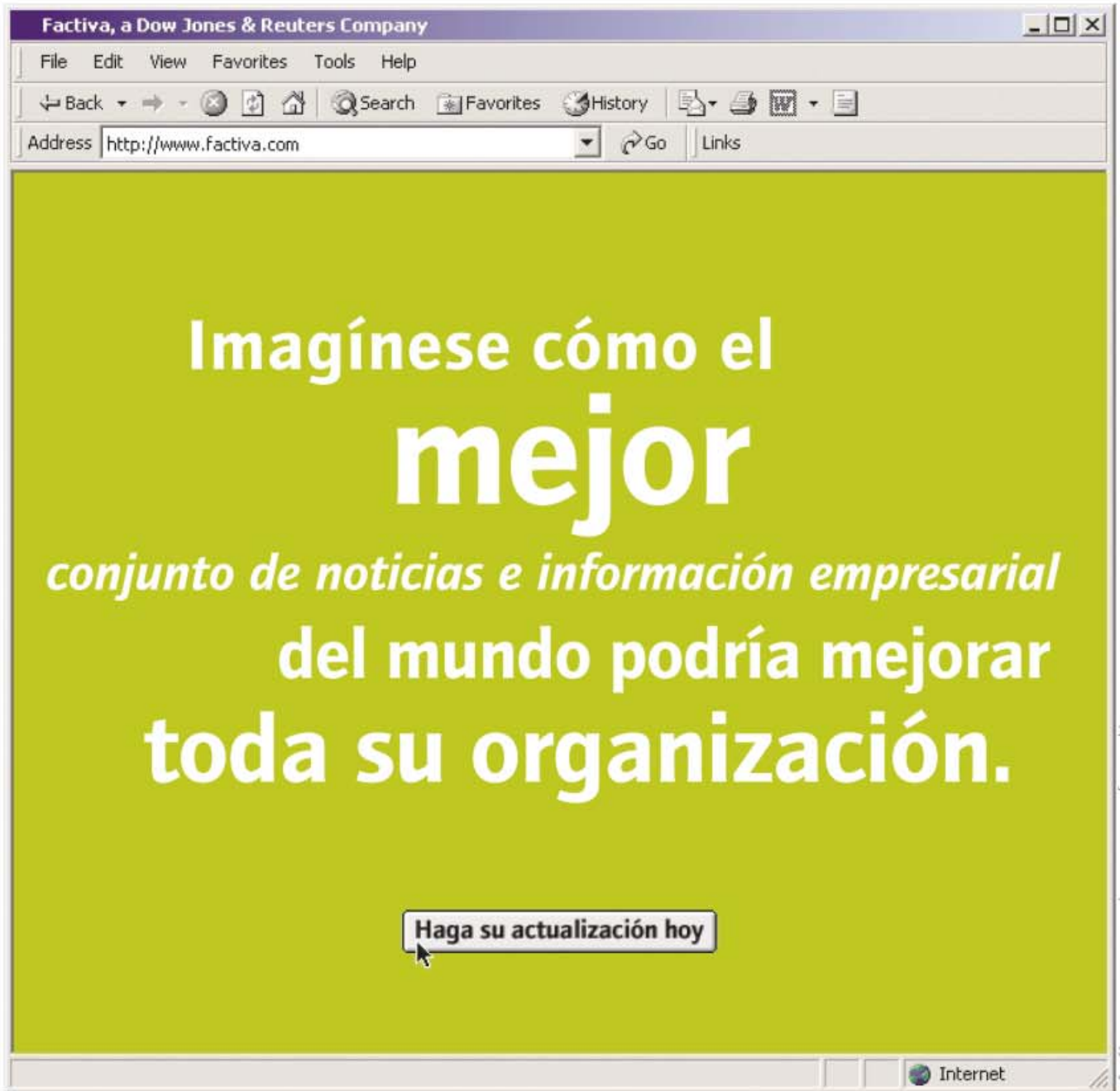


Figura 9. Situación actual entre sistemas



Derechos de autor 2002 Dow Jones Reuters Business Interactive (operando bajo el nombre de Factiva). Todos derechos reservados. F-792-S-1.02

Cambie a Factiva y mejore la *competitividad* de su organización aumentando la *productividad* de su personal. Gracias a la utilización de una nueva y potente tecnología, Factiva suministra a cada uno de sus colaboradores una selección personalizada de los recursos más valiosos en su campo de actividades. Resultado: *un proceso innovador que permite tomar decisiones más acertadas*. Factiva combina la potencia y la seriedad de la información suministrada por Dow Jones y Reuters a nivel mundial para ofrecer la gama más completa de fuentes de información que incluyen artículos enteros tomados de publicaciones de primera categoría del mundo de las finanzas y la industria.

Cambie a Factiva: es una manera totalmente nueva de conseguir el éxito.

Visítenos en www.factiva.com.

Aproveche las ventajas de actualizar en '02

FACTIVA
Reuters España, S.A.
Paseo de la Castellana, 36 – 38
28046 MADRID
Tel.: +34 91 585 21 00 – Fax: +34 91 435 96 54


Dow Jones & Reuters

ción con la aplicación. La incorporación de una base de datos de incidencias y soluciones durante la implantación del sistema para recoger de forma centralizada las sugerencias, problemas o errores por parte de usuarios y técnicos, ha sido de gran efectividad para su resolución, y su funcionamiento se ha mantenido como método de trabajo.

La participación de los usuarios también ha sido importante en la definición de los perfiles iniciales ya que el impacto de la nueva herramienta en los procesos de trabajo ha obligado a adaptar algunos de ellos o crear otros nuevos.

En referencia al proceso de implantación, hay que destacar la importancia de la formación de los usuarios. En este caso, además de prepararlos para su uso, se incidió especialmente en la explicación de los motivos del cambio y en los objetivos propuestos. La diversidad de usuarios y los distintos niveles existentes entre ellos plantean actualmente la necesidad de hacer hincapié en aspectos concretos de la formación como pueden ser las estrategias de búsqueda avanzada, la utilización de programas asociados, conocimientos en informática, etc.

Finalmente, no se puede olvidar que, aunque el proyecto y la implantación en sí ya han concluido, la base de datos del GG se utiliza para elaborar la producción en muchos casos de un día para otro y en ella intervienen múltiples departamentos, procesos, necesida-

des y sistemas de producción. Todo ello exige que se actúe con mucha meticulosidad en el proceso de mantenimiento del sistema para asegurar su calidad.

En lo que respecta a las líneas de trabajo que están abiertas, a corto plazo hay que señalar el estudio de actuaciones que faciliten la incorporación de nuevos productos y aplicaciones a partir de las necesidades tanto informativas como de gestión que puedan ir surgiendo y, por otra parte, el análisis de los aspectos relacionados con la comercialización de contenidos, ámbito en el que aún existen vacíos legales.

Nota

1. Se tenían que revisar manualmente los textos procedentes de *Hermes* para comprobar que la distribución de la información en los campos (autor, titular, etc.) fuera correcta.

Bibliografía

Codina, Lluís. "Metodología de análisis de sistemas de información y diseño de bases de datos documentales: aspectos lógicos y funcionales". En: *Anuari Socadi: 1998*. Barcelona: Socadi, 1998, pp. 195-209.

Cooper, Michael D. "Design considerations in instrumenting and monitoring web-based information retrieval systems". En: *Journal of the Asis*, 1998, v. 49, n. 10, pp. 903-919.

Fuentes, Ma. Eulàlia (ed). *Manual de documentación periodística*. Madrid: Síntesis, 1995.

Hearst, Marti A. "User interfaces and visualization". En: **Baeza-Yates, Ricardo; Ribeiro-Neto, B.** *Modern information retrieval*. New York: ACM; Harlow: Addison-Wesley, 1999, pp. 257-323.

Leloup, Catherine. *Motores de búsqueda e indexación*. Barcelona: Gestión 2000, 1998. Isbn 84-8088-257-3.

Muñoz, Jesús E. "Bancos de imágenes: evaluación y análisis de los mecanismos de recuperación de imágenes". En: *El profesional de*

la información, 2001, marzo, v. 10, n. 3, pp. 4-18.

Periago, Lucía. "Evolución de los sistemas de almacenamiento y distribución de fotografías en el diario regional 'La verdad'". En: *El profesional de la información*, 2001, octubre, v. 10, n. 10, pp. 12-21.

Raya, Fidel. *Database design for information retrieval: a conceptual approach*. New York [etc.]: John Wiley & Sons, 1987. Isbn 0-471-82786-X.

Rodríguez Yunta, Luis. "Evaluación e indicadores de calidad en bases de datos". En: *Revista española de documentación científica*, 1998, v. 21, n. 1, pp. 9-23.

Sabin-Kildiss, Luisa; Cool, Colleen; Xie, Hong (Iris). "Assessing the functionality of web-based versions of traditional search engines". En: *Online*, 2001, March-April, pp. 18-26.

Willitts, John. *Database design and construction: an open learning course for students and information managers*. London: Library Association, 1992. Isbn 1-85604-027-5.

Carles Salmurri, jefe del Departamento de Documentación de La Vanguardia Ediciones.
csalmurri@vanguardia.es

Ernest Abadal, profesor de tecnologías en la Facultat de Biblioteconomia i Documentació, Universitat de Barcelona.
abadal@fbd.ub.es

Mònica Sancho, archivo de La Vanguardia Ediciones.
msancho@vanguardia.es

Miquel Llevat, informático de La Vanguardia Ediciones. Project manager del proyecto.
mllevat@vanguardia.es

Andreu Sulé, profesor de tecnologías en la Facultat de Biblioteconomia i Documentació, Universitat de Barcelona.
sule@fbd.ub.es

Maita Corbera, archivo de El mundo deportivo.
archivo@elmundodeportivo.es

Sistemas de gestión de contenidos en la gestión del conocimiento

Opciones



[\[Versió catalana\]](#)

MARIO PÉREZ-MONTORO GUTIÉRREZ
Estudis de Ciències de la Informació i de la Comunicació
Internet Interdisciplinary Institute (IN3)
Universitat Oberta de Catalunya
perezmontoro@uoc.edu

Resumen [\[Abstract\]](#) [\[Resum\]](#)

Desde la segunda mitad de los noventa, la gestión del conocimiento ha ido gestándose y consolidándose con fuerza como una nueva disciplina en el ámbito de las organizaciones. El objetivo principal de este trabajo es evaluar si los sistemas de gestión de contenidos (CMS) pueden convertirse en el estándar tecnológico más adecuado para la consolidación de esta disciplina. Para cubrir ese objetivo se intenta caracterizar y analizar, en un primer movimiento, las principales funciones (de edición y explotación) de esos nuevos sistemas. En segundo lugar, se intenta abordar la disciplina de la gestión del conocimiento examinando primero cuáles son los diferentes tipos de conocimiento que podemos encontrar en el contexto de una organización e identificando y analizando, después, las operaciones críticas para la gestión adecuada de esa tipología cognitiva. Y, por último, se evalúa en qué medida las soluciones tecnológicas aportadas por los CMS pueden ayudar en la implementación apropiada de esas operaciones críticas.

1 Introducción

En un sentido intuitivo, podemos entender las organizaciones como comunidades o conjuntos de individuos cuyos miembros se estructuran y se articulan para cubrir unos objetivos determinados. A la luz de esta descripción, podemos considerar un hospital, una ONG, un centro educativo, un ministerio, un centro de investigación, una institución política o incluso, tensando la definición, una unidad familiar como ejemplos claros de organización.

Sin embargo, no tendríamos demasiados problemas en convenir que las empresas suelen coincidir con el paradigma de organización que todos tenemos en nuestras mentes. Las empresas están constituidas por un conjunto de individuos estructurados y articulados para cubrir unos fines comunes. En la mayoría de los casos, los principales objetivos que se persiguen son la obtención de beneficios y la supervivencia a largo plazo de la propia empresa.

Dentro del contexto de las organizaciones, se suele manejar el concepto de *activo*. En un sentido lato, podemos considerar como activo de una organización todo aquello que permite o ayuda a una organización en el cumplimiento de sus objetivos. O en un sentido más técnico: todo aquello que puede convertirse en valor para la organización. Los locales

que utiliza la organización, las personas que la forman, el capital del que dispone o la maquinaria y los ordenadores que emplea, son algunos ejemplos de activos de esa organización.

Históricamente, las organizaciones en general —y las empresas en particular—, sólo han considerado como activos elementos materiales o tangibles. En este sentido, en los estados contables tradicionales acostumbraban a aparecer descritos como activos ítems claramente identificados y sobre los que se tuviera una información y una valoración fiable. En muy pocos casos se reconocía como activos otros elementos que no fueran tangibles o materiales, que no tuvieran, por decirlo en un sentido extremo, una dimensión o naturaleza espacio-temporal.¹

Sin embargo, desde la mitad de la década de los noventa esta tendencia está cambiando. Las empresas se han dado cuenta de que para poder realizar unos estados contables más fiables y realistas han de incluir, además de elementos tangibles, activos de naturaleza intangible. Se trata de elementos que, aun siendo no tangibles, son capaces de producir igualmente valor para la organización, son capaces de contribuir a que la organización alcance sus objetivos.²

Entre esos activos intangibles destaca uno en especial por el papel crítico que juega en el contexto de las organizaciones: el conocimiento. En la última década, el conocimiento se ha perfilado como un recurso clave para las empresas, a veces más importante que el propio capital o el trabajo. Se ha considerado, en definitiva, como uno de los activos que pueden convertirse en la principal fuente de ventajas competitivas que permita la supervivencia de una compañía en el actual escenario de globalización de mercados y de feroz competencia.

Con el objetivo de intentar convertir el conocimiento en ese valor añadido que puede garantizar el correcto funcionamiento y la supervivencia de las empresas en el nuevo escenario económico, emerge con fuerza en la última década una nueva disciplina: la gestión del conocimiento.

Una de las características más importantes de esta nueva disciplina es su importante dimensión tecnológica. La gestión del conocimiento no sería lo que actualmente es si no hubiera aparecido de la mano del desarrollo de una serie de tecnologías de la información y de la comunicación. De esta manera, existe un vasto conjunto de soluciones tecnológicas que participan en el éxito pragmático de esta disciplina y que, en cierto sentido, compiten por ocupar un lugar privilegiado dentro de la misma. En los últimos años se está desarrollando y consolidando una nueva herramienta, los sistemas de gestión de contenidos (*Content Management Systems*, CMS) que, por sus características funcionales propias, puede llegar a convertirse en el estándar tecnológico para la implementación de los programas de gestión del conocimiento.

El objetivo principal de este trabajo es evaluar cómo estos nuevos sistemas de gestión de contenidos pueden convertirse en el componente tecnológico más adecuados para la consolidación de la disciplina de la gestión del conocimiento. Para cubrir ese objetivo vamos a desarrollar los siguientes puntos. En primer lugar, en el apartado 2, vamos a caracterizar y analizar las principales funciones de esos nuevos sistemas. A continuación, en el apartado 3, vamos a intentar definir la disciplina de la gestión del conocimiento examinando primero cuáles son los diferentes tipos de conocimiento que podemos encontrar en el contexto de una organización e identificando y analizando, después, las operaciones críticas para la gestión adecuada de esa tipología cognitiva. Y, por último, en el apartado 4, se evalúa en qué medida las soluciones tecnológicas aportadas por los CMS pueden ayudar en la implementación apropiada de esas operaciones críticas.

2 Sistemas de gestión de contenidos

Aunque existan muchos productos en el mercado y se hayan escrito verdaderos ríos de tinta sobre los Sistemas de Gestión de Contenido (CMS),³ no es una tarea sencilla caracterizar este tipo de sistemas. El problema no radica en la complejidad del objeto analizado, sino, más bien, en la naturaleza del mismo. Y es que, como señalan algunos autores (Browning y Lowndes, 2001, por ejemplo), los CMS son más un nuevo concepto que una nueva tecnología.

Pero, ¿qué son los realmente los CMS? Para contestar a esta pregunta vamos a introducir brevemente una descripción del contexto donde nacen, haciendo especial mención de las necesidades que generan su desarrollo y, en un segundo movimiento, pasaremos a presentar una caracterización de las principales funcionalidades de este tipo de sistemas.

2.1 El origen de los sistemas de gestión de contenidos

En términos generales, podemos decir que los sistemas de gestión de contenidos aparecen para dar respuesta a unas necesidades que surgen como fruto de una evolución tecnológica y del aprovechamiento de la misma.

Aunque ya existían desarrollos plenamente funcionales en la segunda mitad de la década de los noventa, es sobretodo a partir del año 2000 cuando los CMS comienzan a consolidarse dentro del panorama tecnológico. Sirva esto para ofrecer una fecha aproximada a partir de la opinión de diversos autores (Tramullas, 2005; Wilkoff, Walter y Dalton, 2001; o Cuerda y Minguillón, 2005)

Justo en ese momento se estaba produciendo un cambio importante en la explotación del entorno de Internet por parte de sus usuarios, y especialmente por parte de las empresas. A lo largo de la década de los noventa, se había identificado por parte de las corporaciones la posibilidad de explotar el fenómeno de Internet en su propio beneficio. En este sentido, por un lado, se comenzó a apostar por la red como fuente de negocio identificando nuevos canales de comercialización y dando lugar a lo que se hoy se conoce como comercio electrónico y a todo lo que rodea a esta estrategia empresarial. Pero también, por otro lado, se identificó la red como el recurso tecnológico idóneo para mejorar y hacer más eficiente el funcionamiento interno de las estructuras organizativas y de los procesos involucrados en este tipo de organizaciones.

En esta nueva situación, para afrontar esos nuevos retos con garantías, las páginas webs estáticas y poco articuladas comenzaban a presentarse como insuficientes y se requería, cada vez más, la utilización de otros tipos de páginas web más dinámicas y capaces de admitir cambios continuos —páginas web más *escalables*, podríamos decir— en función de las necesidades del entorno.

Pero el reto no era sólo tecnológico, sino que también debía responder a las restricciones de tipo económico: era necesario encontrar una herramienta que facilitase la realización de todo eso, pero además, que permitiese hacerlo con un bajo coste. O dicho en otros términos, que permitiese que una persona con unos conocimientos informáticos mínimos fuera capaz de utilizarla y de desarrollar, de una manera rápida y sencilla, las soluciones requeridas. Los sistemas de gestión de contenidos nacen como respuesta a esta doble necesidad tecnológica y económica.⁴

Pero, ¿cómo estos sistemas ayudan cubrir estos objetivos? En el contexto de una organización, encontramos habitualmente infinidad de documentos (o recursos digitales)

que contienen datos de diferentes tipos. Así, por ejemplo, podemos identificar documentos con datos textuales, otros con datos numéricos, otros con imágenes y/o sonidos o incluso documentos que incluyen simultáneamente todos esos distintos tipos de datos. Además, esos documentos, con su variedad de datos, pueden aparecer en formatos bien diferenciados. De esta manera, existen documentos generados por un procesador de texto, hojas de cálculo, páginas web, emails o ficheros multimedia, por citar sólo unos pocos ejemplos. Si buscamos una gestión adecuada de estos recursos digitales que aparecen bajo una gran diversidad de formatos, seguramente tendremos que apostar por la segmentación de esa gestión y por la utilización de diversos sistemas de gestión, con los problemas de reducida eficacia y limitada exhaustividad que conlleva este tipo de estrategias.

Estrictamente hablando, como algunos autores señalan (Browning y Lowndes, 2001), los sistemas de gestión de contenidos son herramientas tecnológicas creadas para cubrir el objetivo prioritario de incrementar y automatizar los procesos que sostienen de una manera eficaz y eficiente la comunicación por Internet. En este sentido, los CMS son conjuntos articulados de aplicaciones informáticas (aunque desde el punto de vista del usuario se tenga la sensación de que se encuentra delante de un único programa) que suelen integrar, en la medida de lo posible, los documentos con formatos distintos en formato XML y crear directamente nuevos documentos en este mismo formato. Esos documentos o recursos digitales resultantes reciben el nombre genérico de *contenidos*.⁵ Los CMS mantienen separados el contenido de su apariencia o presentación final, lo que permite obtener unos importantes beneficios en su explotación (como veremos en el siguiente apartado), y suministran las herramientas necesarias para realizar una gestión eficiente de esos contenidos que pueda potenciar la dimensión comunicativa (en un sentido amplio) de la web de la que forman parte. Las herramientas que suministran y que permiten cubrir el ciclo de vida completo de las páginas de una web se caracterizan por permitir implementar esa gestión de una manera muy cómoda y rápida que evita tener que encargar continuamente esas tareas a un webmaster o especialista, con el ahorro de tiempo y dinero y con la flexibilidad que esto supone.

2.2 Funciones de los sistemas de gestión de contenidos

Una vez que hemos apuntado los orígenes de los sistemas de gestión de contenidos y descrito brevemente su estructura, pasemos ahora a abordar las principales funciones que éstos presentan.⁶

FUNCIONES DE EDICIÓN DE CONTENIDOS	DE	FUNCIONES DE EXPLOTACIÓN DE LA WEB	DE
Creación de contenidos		Entornos colaborativos	
Gestión de contenidos		Perfiles de usuarios	
Publicación de contenidos	de	Sindicación de contenidos	de
Presentación de contenidos	de	Articulación de funcionalidades	de
		Integración de aplicaciones	de

Tabla 1. Funciones de un sistema de gestión de contenidos

Para introducir esas características funcionales vamos a agruparlas, como mostramos en la tabla número 1, en dos grandes grupos. En primer lugar abordaremos las funciones relacionadas con la edición de los contenidos. En segundo lugar, introduciremos las funciones relacionadas con la explotación de la web. Bajo esta segunda categoría introduciremos una serie de soluciones importantes que ofrece un CMS pero que no atañen exclusivamente al proceso de edición de contenidos.

2.2.1 Edición de contenidos

Bajo la categoría “edición de contenidos” vamos a agrupar todas las funcionalidades que ofrecen los sistemas de gestión de contenidos y que están relacionadas con el ciclo de vida básico de los contenidos. Para describir las funciones involucradas en la edición, vamos, a su vez, a utilizar la clasificación propuesta por Roberson (2002a, 2003c, 2003d) y que también otros autores recogen (Browning y Lowndes, 2001; Tramullas, 2005; o Cuerda y Minguillón 2005, por ejemplo). En esta línea, vamos a distinguir las funciones involucradas en la creación, gestión, publicación y presentación de esos contenidos.

Comencemos con funciones relacionadas con la creación de los contenidos. Los CMS ofrecen todo el utillaje informático necesario para desarrollar con comodidad la autoría y creación de los contenidos. En este sentido, acostumbran a suministrar un editor de texto WYSIWYG (*what you see is what you get*: ‘lo que ves es lo que obtendrás’, literalmente). Este editor, que funciona de una manera similar a los editores de texto tradicionales, permite que un contenido pueda ser creado por una persona que no posea conocimientos técnicos sobre la creación de páginas web, sobre lenguajes HTML o XML. Con este editor, la persona es consciente en todo momento del aspecto final del contenido y no tiene que estar pendiente, como en el caso de otros editores de páginas web, de los complicados códigos que generan ese aspecto.

Esta funcionalidad se refuerza, además, con el hecho de que los CMS tratan el contenido separadamente de la apariencia final que pueda adquirir el mismo. Esto permite que la persona que lo está creando pueda despreocuparse de los temas relacionados con esa apariencia formal y pueda centrarse, casi de forma exclusiva, en la calidad semántica del mismo.

Las consecuencias de estas dos soluciones (editor de texto y separación contenido-presentación), son claras y directas. Por un lado, la persona que posee una información relevante que es susceptible de ser publicada en la web puede, sin grandes problemas, ser también la que cree materialmente el contenido y puede, además, ir variando ese contenido a voluntad tantas veces como desee y de una manera flexible. Y todo ello sin tener que complicar, encarecer y ralentizar el proceso de creación e introducción de cambios encargando esas tareas a un especialista en creación y diseño de páginas web. La reducción de los costes procedimentales y económicos es clara. Y, por otro, y como consecuencia de todo ello, dota a la web de una mayor capacidad de crecimiento y adaptación a las necesidades, presentes y futuras, para la cual fue creada.

Pero las funcionalidades relacionadas con la creación no acaban aquí. Los sistemas de gestión de contenidos suministran también soluciones de *workflow* (soluciones para ciclos de trabajo) que permiten que los contenidos puedan ser creados por distintas personas, descentralizando la autoría de los mismos y ofreciendo las ventajas que esto supone. Estos sistemas, además, ofrecen la posibilidad de enriquecer, en el mismo proceso de autoría, los contenidos creados con estándares de control (metadatos), asegurando de esta manera una ulterior explotación documental mucho más efectiva.

Pasemos ahora a la gestión de los contenidos. La responsabilidad principal de esta tarea recae sobre un sistema de gestión de bases de datos que se encuentra incorporado en los sistemas de gestión de contenidos.⁷ La base de datos creada permite la posibilidad de dar de alta o baja un contenido en el depósito y la explotación (búsqueda y recuperación) de todos los contenidos que se han ido creando y añadiendo a ese depósito. Pero también, como bien señalan Cuerda y Minguillón (2005), los contenidos no son los únicos elementos integrados en esa base de datos. En ésta se almacenan principalmente también los datos relativos a los documentos (versiones hechas, autores, fecha de publicación, cambio y caducidad, etc.) y los datos y preferencias de los usuarios.

La inclusión de estos dos tipos de datos aporta beneficios muy importantes. Por un lado, permiten un control de versiones y un control de autoría en cada uno de los contenidos. El control de versiones y autoría habilita la posibilidad de que dos o más autores puedan crear y modificar simultáneamente contenidos sin confusiones. Pero también permite introducir perfiles entre los autores otorgándoles, de forma planificada, ciertos privilegios de autoría y transformación.

Este triple control de versiones, autoría y perfiles introduce un alto nivel de seguridad en la web. En cualquier momento se puede pedir al sistema un informe histórico de autoría y versiones, reduciéndose, de esta manera, los posibles problemas legales asociados a ese proceso de autoría.

Además, el sistema que gestiona la base de datos está diseñado también para realizar otras operaciones. Entre éstas, se podrían destacar, por ejemplo, la posibilidad de integrar tesauros, taxonomías y ontologías para mejorar la gestión, la importación y explotación de bases de datos previamente ya existentes y la incorporación al depósito de documentos en distintos formatos.

Abordemos ahora las funciones relacionadas con la publicación de los contenidos. Los sistemas de gestión de contenidos suministran también soluciones tecnológicas que permiten realizar la publicación de las páginas web (de los contenidos) de una forma muy cómoda. El éxito de las funciones de publicación se sostiene, de nuevo, sobre un tratamiento que ya hemos apuntado: mantener separados el contenido de las páginas y su apariencia final. Así, por ejemplo, esta separación permite que se controle la visualización de esas páginas mediante el uso de plantillas que previamente hayan sido creadas y sin que los eventuales cambios de visualización afecten al contenido de las mismas. También, ofrecen la posibilidad de publicar de forma programada los contenidos (introduciendo una fecha de inicio y otra final de publicación), de diseñar interfaces de forma personalizada y de publicar en los diferentes formatos existentes (PDF, HTML, etc.).

Describamos, por último, las funciones relacionadas con la presentación de los contenidos. De nuevo estas funciones se fundamentan sobre la separación entre aspecto final de las páginas y el contenido de las mismas y lo que permite tratar la publicación de los contenidos como un proceso *ad hoc* al de su creación. Entre estas funciones habría que destacar, por ejemplo, la de permitir un acceso personalizado a la web a través del diseño de los menús de navegación, manteniendo los enlaces y favoreciendo, de esta manera, la usabilidad de la misma. Aunque no suelen presentar problemas para que sus páginas puedan ser visualizadas por los principales navegadores (Mozilla, Explorer, etc.), ofrecen la posibilidad de transformar todos los contenidos en documentos con formato HTML, lo que garantiza la perfecta compatibilidad con los navegadores y plataformas y también permiten la posibilidad de restringir el tipo de usuarios de la web exigiendo un tipo de tecnología concreta (Java, Javascript, Flash, etc.) para la visualización de sus páginas.

2.2.2 Explotación de la web

Introduzcamos, ahora, las funciones relacionadas con la explotación de la web. Bajo esta segunda categoría vamos a presentar una serie de soluciones importantes que puede ofrecer un CMS pero que no atañen exclusivamente al proceso de edición de contenidos.

La primera de estas funciones hace referencia a la posibilidad de crear entornos colaborativos utilizando las soluciones tecnológicas que ofrece un CMS. Aprovechando las herramientas de *workflow* suministradas para la creación de los contenidos, se pueden solucionar, simplificar y mejorar aquellos procesos de la organización donde se puedan automatizar circuitos y flujos documentales. En la misma línea, estos sistemas permiten crear espacios virtuales de encuentro (foros) que habilitan la comunicación sincrónica y diacrónica entre los usuarios de los mismos

La segunda de las funciones se centra en la creación y el control de perfiles de usuarios gestionando de forma segura los privilegios de los mismos. Con los CMS es posible introducir y controlar los perfiles de los usuarios de la web otorgándoles, de forma planificada, ciertos privilegios de acceso, autoría y transformación. Este control también permite obtener informes sobre el comportamiento de explotación de los usuarios (qué miran, qué recuperan, qué imprimen, etc.) que habilitan la posibilidad de realizar estudios de usabilidad de la web e introducir mejoras a partir de los mismos. Estos sistemas permiten también que los usuarios autorizados puedan dar de alta (respecto a la autoría) a nuevos usuarios y éstos a unos terceros.

La tercera de las funciones se centra en la posibilidad de crear un sistema de sindicación de contenidos. Aprovechando la gestión de usuarios y perfiles, el sistema permite la creación automática, en función de esos perfiles, de nuevos contenidos a partir de los ya existentes en el depósito. Esta función se basa, como en el caso de algunas de las anteriores, en la gestión independiente del contenido y la presentación de las páginas web.

Otra de las características se encuentra relacionada con la introducción de nuevas funcionalidades. Los sistemas de gestión de contenidos, como bien señalan algunos autores como Cuerda y Minguillón (2005), permiten la integración a voluntad de nuevas funcionalidades en la web para cubrir el crecimiento y las nuevas necesidades que ésta deba cumplir. Son capaces de introducir esas nuevas funcionalidades de una manera cómoda y sin que esta introducción produzca problemas estructurales internos. Además, con las herramientas para la gestión de la presentación de las páginas, es posible mantener con cierta comodidad la consistencia visual de toda la web incluso después de introducir nuevas funcionalidades en la misma.

La última de las funciones que queremos señalar es la de la integración de otras aplicaciones. Los CMS, utilizando un recurso del tipo API (Applications Programming Interface), son capaces de integrar dentro del mismo otras soluciones informáticas (aplicaciones de marketing o de comercio electrónico, por ejemplo) externas al sistema. En la misma línea, los últimos productos de sistemas de gestión de contenidos permiten extender esta funcionalidad a las plataformas móviles (PDA o telefonía móvil, por ejemplo) y ofrecen la posibilidad, también, de realizar periódicamente imágenes (espejos) del sistema como medida de seguridad.

3 Gestión del conocimiento

Una vez analizada la funcionalidad de los sistemas de gestión de contenidos ya podemos pasar a abordar la gestión del conocimiento.

Como ya apuntábamos, la gestión del conocimiento debe ser entendida como una disciplina que tiene como principal objetivo diseñar sistemas que permitan que el

conocimiento pueda convertirse en valor para una organización. Esto significa que, con la implementación de esos sistemas, el conocimiento debería llegar a contribuir de una manera clara a la consecución de los objetivos que persigue la propia organización.

El diseño de este tipo sistemas, sin embargo, no puede realizarse adecuadamente sin tener en cuenta antes una serie de restricciones impuestas por el propio objeto sobre el que deben aplicarse. En este sentido, vamos comenzar examinando primero cuáles son los diferentes tipos de conocimiento que podemos encontrar en el contexto de una organización. Este examen nos permitirá entender las propiedades de la entidad gestionada y nos ayudará también a perfilar su gestión. Una vez concluido este examen, analizaremos las operaciones críticas que tendrían que implementarse para que el conocimiento alcanzara plenamente ese papel central en el entorno organizacional.

3.1 Tipologías de conocimiento

Dentro del contexto de una organización existen diferentes tipos de conocimientos que presentan características propias y que, por tanto, reclaman una gestión personalizada —a la carta— que tenga en cuenta esa naturaleza singular.

Existen muchas propuestas de clasificación del conocimiento involucrado en las organizaciones. Nosotros vamos a introducir una clasificación alternativa (resumida en tabla 2), basada en una doble e intuitiva distinción, que nos permitirá, en un segundo movimiento, caracterizar con mayor finura las operaciones implicadas en los procesos en los que se gestiona el conocimiento. La primera distinción hará referencia al formato del conocimiento, la segunda, en cambio, se centrará en las propiedades intrínsecas (o proposicionales) del mismo.

SEGÚN SU FORMATO		SEGÚN PROPIEDADES INTRÍNSECAS		SUS
Conocimiento capital	como humano	Conocimiento explícito	Conocimiento tácito	
Conocimiento información	como	Conocimiento explicitable	Conocimiento tácito no explicitable	

Tabla 2. Tipología del conocimiento en una organización

Comencemos primero con la distinción basada en el formato. Habitualmente el conocimiento en las organizaciones puede encontrarse en dos formatos distintos. Por un lado, nos encontramos con el conocimiento como capital humano. Se trata del conocimiento residente, como estado mental, en la cabeza de los miembros de la organización. Está compuesto por la suma de todo el conocimiento (tácito y explícito) que éstos poseen en sus mentes. Por otro, podemos identificar el conocimiento como información. Éste se encuentra formado por aquellos conocimientos que se hallan representados o materializados en forma de documentos (en cualquiera de sus formatos).

Pasemos ahora a la segunda distinción, que se centra en las propiedades intrínsecas o proposicionales del conocimiento. Utilizando estas propiedades, podemos discriminar entre el conocimiento explícito y el conocimiento tácito.

El conocimiento explícito se caracteriza por ser directamente codificable en un sistema de representación como el lenguaje natural, por ejemplo. En este sentido, es fácilmente transmisible o comunicable y, por lo tanto, sí que es accesible a otros individuos de una forma directa. Para señalar que alguien, A, posee un conocimiento de este tipo solemos utilizar la expresión “A sabe que P” (donde *P* acostumbra a ser un enunciado). De esta manera, saber que una molécula de agua está formada por dos átomos de hidrógeno y uno de oxígeno o saber que cuando a la fotocopiadora se le enciende la luz roja hay que cambiarle el recambio de tinta, son dos ejemplos de este tipo de conocimiento.

El conocimiento tácito, en cambio, se corresponde con ese conocimiento basado en la experiencia personal y en muchos casos se identifica con las habilidades del sujeto. Su principal característica es que es difícilmente transmisible o comunicable —verbalizable por así decirlo— y, por tanto, que no es accesible a otros individuos de una forma directa. Para señalar que alguien, A, posee un conocimiento de este tipo solemos utilizar la expresión “A sabe P” (donde *P* acostumbra a ser un verbo). De esta manera, “saber nadar”, “saber ir en bicicleta”, “saber conducir un automóvil”, “saber hablar en público” o “saber articular y dirigir un grupo de personas”, son varios ejemplos de este tipo de conocimiento. Es posible discriminar, a su vez, dos tipos de conocimiento tácito: el conocimiento tácito que, aunque con dificultades, puede ser explicitado (verbalizado), y el conocimiento tácito no explicitable. “Saber multiplicar” puede ser un ejemplo del primer tipo; y “saber hablar una lengua”, una muestra del segundo.

Si multiplicamos cartesianamente de forma adecuada⁸ los dos conjuntos de categorías que subyacen tras esta doble distinción (conocimiento como capital humano y conocimiento como información, por un lado; y conocimiento explícito, tácito explicitable y tácito no explicitable, por otro), obtenemos como resultado una clasificación de todo el conocimiento potencialmente existente en el contexto de una organización. Las cuatro categorías que la definen serían las siguientes: conocimiento explícito como capital humano, conocimiento tácito explicitable como capital humano, conocimiento tácito no explicitable como capital humano y conocimiento como información. Esta clasificación nos permitirá caracterizar de una manera más adecuada, en el siguiente apartado, las operaciones críticas de los programas de gestión del conocimiento.

3.2 Las operaciones críticas en la gestión del conocimiento

Una vez que tenemos claro cuáles son los diferentes tipos de conocimientos (y sus características propias) existentes en el entorno de una organización, ya podemos pasar a identificar cuáles son las operaciones críticas que deben ser realizadas para poder gestionar el conocimiento en esos contextos.

Como ya adelantábamos, la gestión del conocimiento persigue que el conocimiento pueda convertirse en valor para una organización, llegando a contribuir de una manera clara a la consecución de los objetivos que persigue la propia organización. En términos generales, esto se consigue cuando se logra que todo el conocimiento que reside en la organización pueda ser utilizado por quien lo necesite para actuar de manera adecuada en cada momento en ese contexto.

Sin embargo, tras ese propósito se encuentran un conjunto de operaciones críticas de cuyos correctos desarrollos depende en gran medida el éxito o el fracaso de un programa de gestión del conocimiento. Este conjunto estaría formado por las siguientes operaciones: la creación del conocimiento, su captura, su estructuración y procesamiento, su diseminación, su adquisición y la aplicación del conocimiento. Es importante señalar que, para el buen funcionamiento de estos programas, estas operaciones deben estar

articuladas entre sí formando una estructura de cadena circular, donde cada una de ellas se aplica sobre el resultado obtenido de la aplicación de la anterior.

Comencemos abordando la primera de las operaciones: la creación del conocimiento. Esta operación puede definirse como el proceso mediante el cual se genera nuevo conocimiento en la mente de un miembro de la organización. El resultado acostumbra a ser conocimiento en forma de capital humano, aunque puede ser tácito o explícito. Ese nuevo conocimiento suele generarse a partir de la praxis cotidiana del individuo dentro de la organización. Entre todo ese conocimiento creado destaca uno por el importante papel que juega en el contexto de una organización: las buenas prácticas (*best practices*). Las buenas prácticas son soluciones a problemas (conocimiento, en definitiva) que ha encontrado un individuo y que al ser imitado por el resto de la comunidad ahorra esfuerzos y ayuda a la organización.⁹

Pasemos ahora a la segunda de las operaciones: la captura del conocimiento. La captura puede entenderse como el conjunto de operaciones encaminadas a la identificación y extracción del conocimiento residente en la mente de una persona (conocimiento como capital humano) para ponerlo al alcance del resto de la comunidad que lo necesite. En el caso del conocimiento explícito, esta captura se realiza mediante la codificación o representación del mismo en forma de documentos. El resultado de esa captura es la transformación del conocimiento como capital humano en conocimiento como información. En el caso del tácito explicitable, se produce primero la verbalización del mismo y luego se somete, como en el caso del explícito, a su codificación o representación en forma de documentos. Sin embargo, en el caso del tácito no explicitable, la captura se vehicula a través de estrategias de socialización entre el individuo que lo posee y el resto de la comunidad. En este último caso el conocimiento continúa en su formato de conocimiento como capital humano.

La estructuración y procesamiento es la siguiente maniobra crítica y se concentra sobre el conocimiento (explícito o tácito explicitable) representado en forma documental. Esos documentos se someten a un tratamiento que incluye una serie de operaciones. Por un lado, el documento se somete a una doble revisión. En esa revisión se decide por parte de unos especialistas (un consejo editor) si el conocimiento representado es pertinente (si realmente puede ayudar a la consecución de los objetivos que persigue la propia organización) y si no encierra ningún riesgo (si no incluye algún tipo de información sensible que puede perjudicar en algún sentido a esa organización). Si el documento pasa la doble revisión puede considerarse como aprobado. Por otro lado, una vez aprobado el documento se somete a un examen formal y semántico en el que se realiza un análisis documental del mismo. En último lugar y después de ese examen, el documento se incorpora en un sistema documental (se publica en ese sistema) que ha sido diseñado para conseguir que los documentos como éste, en función de los intereses de la organización, puedan ser accesibles y compartidos por todos sus miembros que lo necesiten.

Pasemos ahora a describir la siguiente de las operaciones: la diseminación (o compartimiento). La diseminación puede entenderse como el conjunto de operaciones encaminadas a que el conocimiento involucrado en la organización llegue mediante su transferencia y su difusión a los miembros de la comunidad que lo necesite. En el caso del conocimiento explícito y del tácito explicitable, esta diseminación se suele realizar mediante la transferencia y la difusión de los documentos en los que se encuentra representado. En el caso del tácito no explicitable, en cambio, esa diseminación se realiza de nuevo a través de estrategias de socialización entre el individuo que lo posee y el resto de la comunidad interesada en el mismo.

La siguiente de las operaciones críticas es la adquisición. Esta operación puede entenderse como la aprehensión, por parte de los miembros de la comunidad, del conocimiento que circula por la misma como fruto de la diseminación. Esta aprehensión

cognoscitiva siempre se realiza a partir del *background* (conjunto de creencias y conocimiento que rigen la conducta de un individuo) previo de la persona (del receptor) que recibe ese conocimiento. En el caso del conocimiento explícito y del tácito explicitable representado en forma documental, esta operación se realiza a partir de la interacción entre el receptor y el documento donde se encuentra representado el conocimiento. El resultado de esa interacción es la transformación del conocimiento como información en conocimiento como capital humano. Sin embargo, en el caso del tácito no explicitable, la adquisición se realiza, de nuevo, a través de estrategias de socialización entre el individuo que lo posee y el potencial receptor del mismo. En este caso el conocimiento continúa en su formato de conocimiento como capital humano.

La última de las operaciones de esta cadena es la aplicación. En esta operación, la persona que ha adquirido el conocimiento lo aplica y lo reutiliza en su praxis diaria dentro de la organización. En muchos casos, el receptor reutiliza ese conocimiento en nuevos contextos diferentes al que lo originó creándose una reinterpretación del mismo que ofrece como fruto la creación de nuevo conocimiento. Con este nuevo conocimiento creado se incrementa el depósito de conocimiento involucrado en la organización y se cierra el círculo comenzando de nuevo todo el proceso de gestión al activarse el primer eslabón de la cadena de operaciones.

El resultado de aplicar la cadena circular de operaciones críticas articuladas en la gestión del conocimiento es altamente rentable y beneficioso. Cada vez que un conocimiento completa todos los procesos involucrados en esa cadena de gestión se produce un aumento de la cantidad de conocimiento útil que circula por la organización y se mejora el acceso al mismo. Esto revierte directa y positivamente en el funcionamiento interno de la propia organización mejorándose las estrategias diseñadas a la hora de cubrir los objetivos de la misma. Y, simultáneamente, de forma recursiva, gracias a la estructura circular del proceso, ese conocimiento vuelve a alimentar la misma cadena y a generar, a su vez, nuevo conocimiento.

4 Sistemas de gestión de contenidos en la gestión del conocimiento

Hasta este punto hemos caracterizado, en el primer apartado, las principales capacidades operativas que ofrecen los sistemas de gestión de contenidos. En un segundo apartado, hemos analizado las operaciones críticas que rigen los procesos de gestión del conocimiento. Nos queda, por último, examinar en qué medida esos sistemas pueden ayudar en la implementación adecuada de esas operaciones críticas. El fruto de este análisis nos permitirá concluir que todo apunta a que los CMS pueden convertirse, en un futuro no muy lejano, en uno de los estándares tecnológicos para la implementación de programas de gestión del conocimiento.¹⁰

Como estrategia expositiva para realizar ese análisis vamos a seguir linealmente la cadena de las operaciones que conforman la gestión del conocimiento. Este seguimiento nos permitirá evaluar la adecuación de las soluciones tecnológicas a esas operaciones críticas.

Comencemos abordando la primera de las operaciones: la creación del conocimiento. Para el desarrollo adecuado de esta operación, los CMS, con sus funcionalidades, pueden jugar un papel muy importante.

En este sentido si, con ayuda del sistema, hemos diseñado una base de datos donde se encuentran almacenados documentos que contienen representado conocimiento, los miembros de la organización pueden tener acceso y recuperar esos documentos del

dipósito y, utilizando sus contenidos, generar conocimiento adecuado (buenas prácticas, por ejemplo) para cubrir de una manera efectiva sus objetivos.

La búsqueda y recuperación de esos documentos que permiten generar conocimiento se puede hacer mucho más adecuada y eficiente si hacemos intervenir otras funciones de los sistemas de gestión de contenidos. Así, por ejemplo, podemos enriquecer todo ese proceso de búsqueda y recuperación integrando algún recurso como un tesoro, una taxonomía o una ontología que favorezca la explotación del depósito por parte del usuario. Pero, además, aprovechando que en el depósito se integran también los datos y preferencias de los usuarios, se pueden diseñar unos perfiles de usuario que, de una manera sistemática y ergonómica, ofrezcan a ese usuario aquellos documentos que pueden ser susceptibles de ayudarle a generar conocimiento. Esta misma gestión de usuarios y perfiles puede permitir que, mediante un sistema de sindicación de contenidos, el usuario obtenga nuevos documentos que contienen conocimiento y que han sido creados de forma automática a partir de documentos que representan conocimiento y que, previamente, se encontraban en el depósito. Y todo ello, con un completo control de seguridad, con la posibilidad de evitar, mediante el control del perfil de usuarios, que el usuario tenga acceso a otros documentos que puedan no estar relacionados con sus objetivos o que contengan información que si se diseminara por la organización supondría un riesgo potencial.

El papel de los CMS en la creación de conocimiento se puede completar mediante la posibilidad de desarrollar entornos colaborativos utilizando las soluciones tecnológicas que éstos ofrecen. Como señalan algunos autores (von Krogh, Ichijo y Nonaka, 2000, por ejemplo), uno de los cuellos de botella en la creación del conocimiento viene determinado por el hecho de que los miembros de una organización llegan a un punto en el que individualmente y de forma autónoma ya no pueden crear más conocimiento. Llegado ese momento, una buena estrategia para superar esta situación es trabajar en grupo para crear conocimiento. Aprovechando las herramientas de los CMS que permiten crear espacios virtuales de encuentro (foros), se puede lograr que los miembros de la organización que forman parte de esos grupos puedan comunicarse (de forma sincrónica y diacrónica). Mediante esta comunicación se suele producir un denso intercambio de ideas (de buenas prácticas, en muchos casos) que, de una manera natural, pueden provocar la generación de conocimiento entre los protagonistas de estos actos comunicativos.

Pasemos ahora a la segunda de las operaciones: la captura del conocimiento. Como ocurría en el caso de la creación, en esta segunda operación los sistemas de gestión de contenidos se perfilan también como una solución global muy adecuada para el correcto desarrollo de la misma.

En el caso de la captura del conocimiento explícito y del tácito explicitable, los CMS tienen un papel central. Permiten que la persona que posee este tipo de conocimiento sea la misma que lo captura y lo representa en un documento. Y que, además, sea la misma que pueda ir variándolo o revisándolo en función de la evolución del mismo. Y que realice esta captura, representación y revisión de una forma rápida, cómoda y, lo que también es importante, en línea. Para ello puede utilizar el editor de texto que incorporan, centrándose, casi de forma exclusiva, en la calidad del conocimiento y desentendiéndose del aspecto final y visualización del documento donde está representando ese conocimiento. Esta solución permite evitar que se complique, encarezca y ralentice el proceso de captura encargando esas tareas a una segunda persona (ajena a la persona que posee el conocimiento) con habilidades de captura y sólidas nociones sobre creación y diseño de páginas web.

La captura de estos tipos de conocimiento también se beneficia de otra de las funcionalidades de los sistemas de gestión de contenidos. En este caso, se puede utilizar también las soluciones de *workflow* que éstos suministran para permitir que los documentos que representan conocimiento puedan ser creados por distintas personas.

Diferentes miembros de la organización que poseen un mismo conocimiento pueden trabajar colaborativamente para la representación del mismo utilizando esta funcionalidad, descentralizando la captura y ofreciendo las ventajas que esto supone. Y, siempre, todos estos procesos con un alto control de seguridad y bajo un control legal mediante una inspección de autoría a partir de los informes históricos.

En el caso del conocimiento tácito no explicitable, la captura se fundamenta sobre otras utilidades suministradas por los CMS. Como ya vimos en el caso de la creación, podemos también ayudar a la captura de este tipo de conocimiento mediante el desarrollo de entornos colaborativos. Creando espacios virtuales de encuentro, permitimos que se produzcan, mediante los episodios de comunicación, las estrategias de socialización, entre el individuo que lo posee y el resto de la comunidad necesarias para que se produzca la captura directa de este tipo especial de conocimientos.

Concentremos ahora nuestra atención en la operación de la estructuración y procesamiento. Como ya indicamos anteriormente, esta operación crítica se concentra sobre el conocimiento (explícito o tácito explicitable) representado en forma documental.

Mediante la utilización de la solución de *workflow* que suministran los sistemas de gestión de contenidos, podemos simplificar y racionalizar en gran medida las operaciones involucradas en la doble revisión (riesgo y pertinencia) a la que los miembros del consejo editor someten los documentos que representan conocimiento y son candidatos a formar parte del depósito. Una vez que el documento pasa la doble revisión y puede considerarse como aprobado, es posible en el mismo proceso enriquecer los documentos evaluados con estándares de control (metadatos) fruto del análisis formal y semántico del mismo, asegurando de esta manera una ulterior explotación documental mucho más efectiva.

Por último, el CMS también suministra un sistema que nos permite diseñar una base de datos que acoja los documentos surgidos del proceso que acabamos de describir. Como ya adelantamos cuando afrontamos la operación de la creación, esta base de datos permite dar de alta o baja los documentos de este tipo y en función de los intereses de la organización, que puedan ser accesibles y compartidos por todos sus miembros que lo necesiten. En esta base de datos también se almacenan datos relativos a los documentos (versiones hechas, autores, fecha de publicación, cambio y caducidad, etc.) lo que permite un control de versiones.

Abordemos ahora la siguiente de las operaciones: la diseminación del conocimiento. Como hicimos en el caso de la captura, hemos de presentar la potencial contribución de los CMS a la operación de la diseminación especificando en cada momento el tipo concreto de conocimiento involucrado en la misma.

Como ya indicamos, en el caso del conocimiento explícito y del tácito explicitable, la diseminación se suele realizar mediante la transferencia y la difusión de los documentos en los que se encuentra representado. En este sentido, los Sistemas de Gestión del contenido son una herramienta altamente eficaz para implementar esta operación. Suministrando un sistema de gestión de bases de datos que permite diseñar una base de datos (que ya describimos anteriormente) donde se encuentran almacenados los documentos en los que se encuentra representado el conocimiento, los CMS facilitan la diseminación o compartimiento del conocimiento a través de la explotación (búsqueda y recuperación) de todos esos documentos transportadores de conocimiento y que conforman el depósito.

Al igual que pasaba en la operación de la creación de conocimiento, la diseminación desarrollada a través de la búsqueda y recuperación de esos documentos puede hacerse mucho más adecuada y eficiente. Así, por ejemplo, se puede enriquecer el proceso mediante el uso de recursos documentales como los tesauros, las taxonomías o las

ontologías. También, aprovechando la utilidad de gestión de perfiles de usuarios, se puede programar el sistema para que, de una manera sistemática y ergonómica, ofrezca a ese usuario aquellos documentos que pueden serle útiles y que, mediante la sindicación de contenidos, el usuario obtenga nuevos documentos que contienen conocimiento creados de forma automática a partir de documentos que previamente se encontraban en el depósito.

En el caso del conocimiento tácito no explicitable (y, en parte también, en el del que sí lo es), los procesos de socialización entre el individuo que lo posee y el resto de la comunidad son la estrategia más adecuada para la diseminación. Con los sistemas de gestión de contenidos se pueden diseñar entornos colaborativos virtuales donde se permita que se disemine el conocimiento a partir de los intercambios comunicativos y los intereses particulares de los usuarios de esos espacios.

La siguiente de las operaciones críticas es la de la adquisición del conocimiento. Como en los casos anteriores, los CMS, con sus funcionalidades, pueden jugar un papel central en el desarrollo adecuado de esta operación. Pero, como también ocurría anteriormente, la contribución varía según sea el tipo de conocimiento que se vea involucrado.

En el caso del conocimiento explícito y del tácito explicitable representado en forma documental, la adquisición es indirecta. Ésta se suele realizar mediante la interacción entre el miembro de la comunidad y el documento donde se encuentra representado el conocimiento y que ha obtenido tras un proceso de búsqueda y recuperación en el depósito creado a partir de las funcionalidades del CMS.

Para mejorar esa adquisición mediada por la interacción receptor-documento se pueden implementar diversas estrategias. Así, por ejemplo, por un lado, se puede reducir el ruido y el silencio en la recuperación de los documentos que contienen conocimiento crítico para un miembro de la organización aprovechando la utilidad de gestión de perfiles de usuarios y la sindicación de contenidos. Esta reducción permite obtener como resultado que el conjunto de documentos con los que tiene que interactuar el individuo sea mucho más adecuado y útil. Y, por otro lado, también se puede mejorar la interacción gestionando y personalizando la visualización de esos documentos utilizando las herramientas relacionadas con la publicación de contenidos.

Como ya ocurría en las operaciones anteriores, en el caso del tácito no explicitable (y, en parte también, en el del que sí lo es), la adquisición se realiza, de nuevo, a través de estrategias de socialización entre el individuo que lo posee y el potencial receptor del mismo. Mediante el uso de los sistemas de gestión de contenidos se pueden planificar entornos colaborativos virtuales para que se produzca esa socialización (en este caso virtual) que habilite la adquisición.

Es importante señalar que la adquisición, mediada o no por documentos, se ve altamente favorecida por una de las capacidades que caracterizan a los últimos productos de sistemas de gestión de contenidos: su funcionalidad en las plataformas móviles (PDA o telefonía móvil, por ejemplo). Esta funcionalidad permite que la adquisición de conocimiento, ya sea a través de documentos o por medio de la socialización, pueda realizarse sin depender del espacio físico donde se encuentren los individuos implicados en esa adquisición.

Acabemos este análisis abordando la última de las operaciones de esta cadena: la aplicación del conocimiento. En esta operación, el papel de los sistemas de gestión de contenidos es un poco más indirecta que en los casos anteriores.

Directamente, no intervienen en la aplicación y reutilización del conocimiento adquirido en la praxis cotidiana que desarrollan los miembros de una organización. Sin embargo, sí

intervienen, y de una manera muy directa, en la difusión de los resultados de esa aplicación. Sobre todo, cuando el receptor utiliza ese conocimiento en nuevos contextos diferentes al que lo originó y, a partir de una reinterpretación del mismo, se termina obteniendo como fruto la creación de nuevo conocimiento.

En estos casos, los CMS ofrecen la posibilidad de integrar este nuevo conocimiento en la base de datos donde se encuentran almacenados los documentos que contienen representado conocimiento, activándose de nuevo toda la cadena de operaciones críticas que se pueden apoyar sobre este recurso tecnológico. Pero también, mediante la posibilidad de desarrollar entornos colaborativos, estos sistemas permiten que el conocimiento creado a partir de la aplicación a nuevos contextos se difunda al resto de la comunidad con el objetivo de que todo ese conocimiento pueda ser utilizado por quien lo necesite para actuar de manera adecuada en cada momento. Este tipo de recurso tiene especial efectividad cuando se aplica sobre las comunidades de práctica, es decir, sobre comunidades o grupos de personas que dentro de la organización comparten unos intereses y un bagaje de experiencias sobre una serie de temas comunes.

Es importante terminar apuntando una de las virtudes más interesantes de los Sistemas de Gestión de Contenido para ayudar en la implantación y el correcto funcionamiento de los programas de gestión del conocimiento: su capacidad de integrar aplicaciones externas al sistema. Sirviéndonos de esa capacidad de los CMS basada en recursos del tipo API (*Applications Programming Interface*), podemos enriquecer todo el sistema de gestión del conocimiento integrando de una forma articulada y ergonómica dentro de una misma solución tecnológica otros recursos informáticos que suelen utilizarse en este tipo de programas y de poder utilizar en línea todos estos recursos aprovechando los estándares de Internet.¹¹

En este sentido, por ejemplo, dentro del CMS se pueden incluir otros recursos como los programas de *data warehouse*, *data mining* o *text mining* que pueden ayudar en la creación, captura, estructuración, diseminación y adquisición del conocimiento a partir de la gestión de documentos que contienen representado conocimiento. Pero también podemos enriquecer el sistema y completar estos recursos con la inclusión de un índice temático digital que, aprovechando la taxonomía derivada de un mapa de conocimiento y permitiendo una navegación interna cómoda, clasifique la totalidad de esos documentos involucrados.

Es posible, además, integrar un sistema de soporte a la toma de decisiones que facilite que los miembros puedan mejorar sus contribuciones a los objetivos organizacionales a partir de una toma de decisiones más adecuada, más ágil y menos tutelada. Por último, no es en absoluto superfluo integrar directamente, además, algún tipo de motor de búsqueda y de metabuscador. El motor de búsqueda posibilita que los miembros puedan navegar con comodidad por el interior de la propia web (intranet) de la organización y que tengan la oportunidad de buscar sin intermediarios nueva información externa en Internet. El metabuscador, en cambio, al combinar todo el poder de recuperación de información de distintos motores de búsqueda facilita y refuerza en gran medida una búsqueda¹² exhaustiva de nuevos e interesantes conocimientos de calidad para la organización en ese mismo contexto de Internet.

Bibliografía

Asilomar Institute For Information Architecture (2003). "The problems with CMS". *The information architecture institute*. <http://aifia.org/pg/the_problems_with_cms.php>. [Consulta: 1/3/2005].

Bluebill Adv. (2003), "The classification and evaluation of content management systems". *The Gilbane report*, vol. 11, n. 2 <
http://www.gilbane.com/gilbane_report.pl/86/The_Classification_Evaluation_of_Content_Management_Systems.html>. [Consulta: 22/1/2005].

Boiko, B. (2001). *Content management bible*. New Jersey: Wiley.

Boisot, Max H. (1998). *Knowledge assets*. Oxford: Oxford University Press.

Browning, Paul y Lowndes, Mike (2001). "JISC techwatch report: content management systems". *Techwatch report TSW 01-02, The joint information systems committee*. September 2001.

Cuerda, Xavier; Minguillón, Julià (2005). "Introducción a los sistemas de gestión de contenidos (CMS) de código abierto". *Mosaic*, nº 36. <
<http://www.uoc.edu/mosaic/articulos/cms1204.html>>. [Consulta: 1/3/2005].

Davenport, T.; Prusak, L. (1998). *Working knowledge*. Boston: Harvard Business School Press.

Doyle, B. (2003). "Open source content management redux". *The Gilbane report*, vol. 11, n. 3. <
http://www.gilbane.com/gilbane_report.pl/87/Open_Source_Content_Management_Redux.html>. [Consulta: 1/3/2005].

ERP Software (2003). *Content management tutorial*. <<http://erptoday.com/CMS/Content-Management-Tutorial.aspx>>. [Consulta: 1/3/2005].

Fahey, Liam; Prusak, Laurance (1998). "The eleven deadliest sins of knowledge management". *California management review*, vol. 40, no. 3, p. 265–276.

Fraser, S. (2002). *Real World ASP.NET: building a content management system*. Berkley: Apress.

Gilbane, Frank (2000). "What is content management?". *The Gilbane report*, vol. 8, no. 8. <
http://www.gilbane.com/gilbane_report.pl/6/What_is_Content_Management>. [Consulta: 1/3/2005].

Gingell, D. (2003). *A 15 minutes guide to enterprise content management*. Pleasanton: Documentum Inc.

Gorey, R.M; Dobat, D.R. (1996). "Managing in the knowledge era". *The systems thinker*, vol. 7, no. 8, p. 1–5.

Gupta, V.K.; Govindarajan, S.; Johnson, T. (2001). "Overview of content management: approaches and strategies." *Electronic markets*, vol. 11, no. 4, p. 281–287.

Han, Y. (2004). "Digital content management: the search for a content management system". *Library Hi Tech*, vol. 22, no. 4, p. 355–365.

Jennings, Tim (2002). *Defining the document and content management ecosystem*. London: Butler Direct.

McKay, Andy (2004). *The definitive guide to plone*. Berkley: Apress.

McKeever, S. (2003). "Understanding web content management systems: evolution, lifecycle and market". *Industrial management & data systems*, vol. 103, no. 9, p. 686–692.

Medición del capital intelectual: modelo Intellect (1998). Madrid: Instituto Universitario Euroforum Escorial.

Nakano, R. (2002). *Web content management: a collaborative approach*. New Jersey: Prentice Hall.

Nonaka, Ikujiro; Takeuchi, Hirotaka (1995). *The knowledge creating company*. Oxford: Oxford University Press.

Pérez-Montoro Gutiérrez, Mario (2000). *El fenómeno de la información. Una aproximación conceptual al flujo informativo*. Madrid: Trotta.

Robertson, James (2002a). "How to evaluate a content management system". *Step two designs*. <http://www.steptwo.com.au/papers/kmc_evaluate/index.html>. [Consulta: 1/3/2005].

Robertson, James (2002b). "What are the goals of a content management system?". *Step two designs*. <http://www.steptwo.com.au/papers/kmc_goals/index.html>. [Consulta: 1/3/2005].

Robertson, James (2003a). "Is it document management or content management?". *Step two designs*, <http://www.steptwo.com.au/papers/cmb_dmorcm/index.html>. [Consulta: 1/3/2005].

Robertson, James (2003b). "Metrics for knowledge management and content management". *Step two designs*, <http://www.steptwo.com.au/papers/kmc_metrics/index.html>. [Consulta: 1/3/2005].

Robertson, James (2003c). "So, what is a content management system?". *Step two designs*. <http://www.steptwo.com.au/papers/kmc_what/index.html>. [Consulta: 1/3/2005].

Robertson, James (2003d). "Why a small website needs a content management?". *Step two designs*. <http://www.steptwo.com.au/papers/cmb_needcms/index.html>. [Consulta: 1/3/2005].

Robertson, James (2004). "Open-source content management systems". *Step two designs*. <http://www.steptwo.com.au/papers/kmc_opensource/index.html>. [Consulta: 1/3/2005].

Rockley, A. (2003). *Managing enterprise content: a unified content strategy*. Worcester: New Riders.

Senge, Peter M. (1990). *The fifth discipline: the age and practice of the learning organization*. London: Century Business.

Suh, P. [et al.] (2003). *Content management systems*. Birmingham: Glasshaus.

Tramullas Saz, Jesús (2005). "Herramientas libres para gestión de contenidos". [Inédito].

Veen, Jeffrey (2004). "Why content management fails?". *Adaptive Path*. <<http://www.adaptivepath.com/publications/essays/archives/000315.php>>. [Consulta: 1/3/2005].

Von Krogh, Georg; Ichijo, Kazui; Nonaka, Ikujiro (2000). *Enabling knowledge creation*. Oxford: Oxford University Press.

Wilkoff, N.; Walker, J.; Root, N.; Dalton, J. (2001). "What's next for content management?". *The techRankings techInsight*. <<http://www.forrester.com/ER/Research/TechInsight/Excerpt/0,4109,13920,00.html>>. [Consulta: 1/3/2005].

Enlaces de interés

AIIM Enterprise Content Management Association. <<http://www.aiim.org/>>

CMS Professionals. <<http://www.cmprofessionals.org/>>

CMS Review. <<http://www.cmsreview.com/>>

CMS Spain. <<http://www.cms-spain.com/>>

EZ-Publish. <<http://www.ez.no/>>

Geeklog. <<http://www.geeklog.net/>>

Mambo. <<http://www.mamboserver.com/>>

Moodle. <<http://moodle.org/>>

OpenCMS. <<http://www.opencms.org/opencms/en/>>

PHPNuke. <<http://phpnuke.org/>>

Plone. <<http://plone.org/>>

PostNuke. <<http://www.postnuke.com/>>

Tiki. <<http://tikiwiki.org/tiki-index.php>>

Xoops. <<http://www.xoops.org/>>

Fecha de recepción: 10/04/05. Fecha de aceptación: 10/04/05.

Notas

¹ Las normas contables internacionales FASB (diciembre, 1984) que regulan lo que debe ser considerado como un activo en los estados contables de una empresa pueden ser consideradas como un claro indicador de esta tendencia.

² Existe, incluso, una fórmula muy intuitiva para calcular el valor de los activos intangibles de una organización que cotiza en bolsa: valor de los activos intangibles = valor de mercado de la organización (valor en bolsa) – valor de activos tangibles. Para introducirse en el tema de los activos intangibles y en el de su medición contable, se pueden consultar *Medición del capital intelectual: modelo Intellect* (1998).

³ A partir de este punto vamos a utilizar con sinónimas las expresiones *sistemas de gestión de contenido* y *CMS*. Hemos decidido utilizar el acrónimo inglés *CMS* (*Content Management Systems*) en detrimento del español *SGC* ya que hemos constatado que las siglas de la versión inglesa aparecen como estándar terminológico en toda la literatura consultada sobre el tema, incluyendo la bibliografía en lengua española.

⁴ Estos orígenes, tan estrechamente relacionados con la web, han provocado que algunos autores (como Wilkoff, Walter y Dalton, 2001, entre otros) se refieran a este tipo de sistemas como *Web Management Systems* en lugar de *Content Management Systems*.

⁵ Como señala McKay (2004), un contenido es una unidad de datos con alguna información asociada. En este sentido, estrictamente hablando, este tipo de recursos contendría la parte semántica de un contenido con un mínimo grado de presentación.

⁶ En este apartado vamos a abordar las funciones que cubriría un producto CMS de tipo estándar. Es importante señalar que existe una gran variedad de productos CMS y que, por tanto, en este apartado puede quedar sin cubrir alguna funcionalidad específica de alguno de esos productos.

⁷ MySQL suele ser el sistema de gestión de bases de datos que incorporan los CMS de código abierto.

⁸ En este caso no tiene sentido multiplicar la categoría “conocimiento como información” de la primera distinción por las categorías de la segunda porque, como ya hemos indicado anteriormente, este tipo de conocimiento es, por definición, conocimiento (explícito o tácito explícitable) que ha sido explicitado y representado en forma de documento.

⁹ Unido a este concepto aparece también el de las malas prácticas (*worst practices*): soluciones erróneas (pseudoconocimiento) que no deben ser imitadas y en las que no se debe invertir más esfuerzos en el contexto de la organización. Éstas deben ser consideradas tan importantes o más que las buenas prácticas.

¹⁰ Estrictamente hablando, la principal aportación de los CMS a los programas de Gestión del Conocimiento se centra en su contribución a la construcción de la Intranet. La Intranet puede ser considerada como uno de los recursos tecnológicos más interesantes para fomentar el flujo y el intercambio global de conocimiento dentro de una organización. En términos tecnológicos, una Intranet es una red informática que utiliza los protocolos de comunicación propios de Internet y que se encuentra desarrollada, en la mayoría de las ocasiones, para uso interno y exclusivo en el seno de una organización.

¹¹ Habitualmente, estas intranets integradoras de la tecnología necesaria para la implantación de un programa de Gestión del Conocimiento se conocen también bajo el nombre de *portales del conocimiento* (*Knowledge Portals*).

¹² Algunos tipos de búsqueda, sobre todo aquellos que presentan ciertos patrones repetitivos en la estrategia desarrollada, pueden automatizarse mediante la utilización de *agentes inteligentes* diseñados para implementar esas acciones

