

# Sistemas de documentación de prensa ¿cuáles son y cómo evaluarlos?

di Ernest Abadal, Javier Guallar, Lluís Codina

## 1. Introducción

El incremento de la cantidad de información de prensa accesible en la red es constante y parece no tener fin. Nunca como ahora había sido posible poder consultar tanta información de actualidad, en todos los idiomas y, además y muy especialmente, de todas las épocas. Esta abundancia de contenidos y la diversidad de demandas, no obstante, generan algunas dificultades para la localización de informaciones concretas y precisas. ¿Dónde se pueden encontrar las noticias sobre temas concretos que pueden interesar a un usuario específico en un momento determinado? Está claro que los buscadores generalistas no son la solución aunque los usuarios poco experimentados piensan que así es. Este es el caso, por ejemplo, de Lysbeth Salander, la atípica y heterodoxa investigadora de las novelas de la serie Millennium. Acostumbrada a entrar en los ordenadores ajenos para satisfacer sus necesidades de información, se encuentra, justo al principio de la saga, con una pregunta que no se puede resolver hurgando en ficheros personales. En este momento piensa en la solución fácil y, tal y como harían muchas personas, escribe “Magda + asesinato” en el cuadro de búsqueda de Google confiando en que los primeros resultados le van proporcionar información relevante sobre unos misteriosos asesinatos de mujeres acontecidos en Suecia en un momento muy anterior a la trama (entre 1949 y 1966). Lo realmente sorprendente es que la segunda referencia facilitada por Google en esta obra de ficción es un reportaje de la televisión sueca que ofrece una pista muy importante para avanzar en la investigación.

Estamos de acuerdo con Stieg Larsson en que la información de prensa tiene un valor muy relevante aunque discrepamos totalmente en que acceder a ella sea tan simple como nos indica en su novela. Quien ingenuamente intente reproducir la búsqueda anterior, ni por asomo va a encontrar el valioso reportaje.

ERNEST ABADAL, Universitat de Barcelona, Facultat de Biblioteconomia y Documentación, Melcior de Palau, 140, 08014 Barcelona, e-mail [abadal@ub.edu](mailto:abadal@ub.edu).

JAVIER GUALLAR, Universitat de Barcelona, Facultat de Biblioteconomia y Documentación, Melcior de Palau, 140, 08014 Barcelona, e-mail [jguallar@ub.edu](mailto:jguallar@ub.edu).

LLUÍS CODINA, Universitat Pompeu Fabra, Barcelona, Departamento de Comunicación, C/ Roc Boronat, 138, 08018 Barcelona, e-mail [lluís.codina@upf.edu](mailto:lluís.codina@upf.edu).



Actualmente, la consulta de noticias retrospectivas se realiza a través de productos especializados que son poco conocidos no sólo por el gran público sino a veces también por los profesionales del sector. En realidad, los perfiles de usuario interesados por estos contenidos son variados: investigadores sociales, periodistas de investigación, historiadores, bibliotecarios, cargos públicos y privados y, en menor medida, también el gran público.

A partir de esta preocupación, y desde la perspectiva del documentalista especializado, hemos realizado diversos estudios que han tenido un doble objetivo. En primer lugar, categorizar y describir las fuentes de información para la consulta y recuperación de información de prensa. En este apartado hemos distinguido entre hemerotecas históricas, hemerotecas del propio medio y bases de datos de prensa. En segundo lugar, hemos establecido un sistema de indicadores que sirven para la evaluación de estos sistemas documentales y, además, los hemos aplicado a conjuntos específicos como pueden ser las hemerotecas digitales españolas (de los diarios de mayor difusión), las catalanas, etc. y en períodos de tiempo distintos, lo que ha permitido disponer de un análisis evolutivo.

En este artículo vamos a presentar someramente los resultados de ambas líneas de investigación.

## 2. Sistemas de documentación de prensa: tipología y características

Su objetivo fundamental es facilitar la consulta a la información de actualidad, es decir, poder encontrar una noticia concreta o un grupo de noticias sobre un tema determinado publicadas en uno o diversos medios, ya sean noticias del día o información retrospectiva.

Hemos distinguido tres tipos de servicios: las hemerotecas históricas, con fondos antiguos digitalizados procedentes de varios medios; las hemerotecas digitales de un medio, que permiten la consulta a los fondos impresos y digitales de un diario, y las bases de datos de prensa, que indexan los contenidos de múltiples medios.

El estudio de las fuentes de información sobre prensa es un ámbito con gran diversidad de trabajos y enfoques, desde los estudios globales o que analizan tipologías de productos a aquellos estudios que tratan de productos concretos. Se trata de un sector que, al estar en continua evolución, justifica realizar nuevas revisiones y actualizaciones. De manera somera y sin pretensiones de exhaustividad, se puede citar, entre los estudios que ofrecen una visión amplia de recursos y fuentes de información sobre medios y prensa digital, a Fuentes, Jiménez y González<sup>1</sup>, Pareja Pérez<sup>2</sup>, Abadal y Guallar<sup>3</sup> y Guallar, Abadal y Codina<sup>4</sup>. Más concretamente en lo que se refiere a hemerotecas digitales de prensa, Jiménez, González y Fuentes<sup>5</sup> han examinado las caracte-

1 Maria Eulàlia Fuentes; Àngels Jiménez; Alfons González, *Servicios y sistemas de acceso a la información periodística en Internet*, «Anuari de Biblioteconomia, Documentació i Informació: Bibliodoc 2000», Barcelona: COBDC, 2001, p. 85-102.

2 Víctor Manuel Pareja Pérez, *Periodismo digital. Medios de comunicación en Internet*. In: *La información especializada en internet: directorio de recursos de interés académico y profesional*, coords. Ángeles Maldonado, Luis Rodríguez Yunta, Madrid: CSIC, 2006, p. 237-276.

3 Ernest Abadal; Javier Guallar, *Prensa digital y bibliotecas*, Trea: Gijón, 2010.

4 Javier Guallar; Ernest Abadal; Lluís Codina, *Sistemas de acceso a la información de prensa digital: tipología y evolución*, «Investigación», vol. 27 (2013), n. 61, p. 29-52. <<http://eprints.rclis.org/21044/>>.

5 Àngels Jiménez, Alfons González; Maria Eulàlia Fuentes, *Las hemerotecas digitales de la prensa en internet*, «El profesional de la información», vol. 9 (2000), n. 5, p. 15-22. <<http://www.elprofesionaldelainformacion.com/contenidos/2000/mayo/2.pdf>>.

## OSSERVATORIO

rísticas de las hemerotecas de diarios internacionales, y Guallar, Abadal y Codina, en diferentes estudios que se detallan más adelante, han establecido unos indicadores de evaluación de hemerotecas y los han aplicado en varios periódicos españoles.

Las hemerotecas digitales han tenido un notable reflejo en la bibliografía italiana. Brunella Longo<sup>6</sup> hace un repaso a la presencia de los archivos de prensa italiana en Internet. Se trata de un estudio no exhaustivo que presenta la panorámica en dos grandes ámbitos: bases de datos internacionales y archivos de prensa de los propios medios. En primer lugar, se analizan los periódicos italianos cuyos contenidos están comercializados por las principales bases de datos (Factiva, LexisNexis) o distribuidores internacionales (Dialog). A continuación, se describen las prestaciones de los archivos de prensa de los principales periódicos digitales, tanto estatales como regionales y se destaca el cambio de las hemerotecas digitales hacia un modelo de pago o, al menos, de identificación para la consulta de la información retrospectiva de la prensa italiana.

Stefano Bonanni et al<sup>7</sup> describen la situación del proceso de digitalización de la prensa impresa italiana. Su estudio incluye 43 iniciativas de carácter diverso financiadas fundamentalmente por la administración pública (entes territoriales, universidades, bibliotecas, etc.) y que incluyen unos 2.500 títulos de prensa. Los principales parámetros analizados son el grado de digitalización, el tipo de prensa digitalizada, los aspectos técnicos, los aspectos legales, la facilidad de acceso, etc. Los autores ponen de manifiesto la necesidad de incrementar el grado de cooperación y de estandarización de estas iniciativas, así como también una relación más directa entre los objetivos de digitalización y de preservación.

A partir de nuestras líneas de investigación en este ámbito proponemos una tipología y caracterización de esta clase de sistemas de documentación de prensa que presentamos a continuación.

### 2.1. Hemerotecas históricas

Son servicios en los que se ha llevado a cabo una digitalización de las páginas de diarios y revistas históricos y se han puesto a disposición del público en Internet. Una buena parte de los impulsores de estas iniciativas son bibliotecas (nacionales, municipales, etc.) que cuentan con el respaldo económico de las organizaciones que las sustentan, ya sean ministerios, ayuntamientos, u otras administraciones. Los formatos más habituales de presentación de los documentos son dos: la imagen gráfica de las páginas escaneadas (por ejemplo, en jpg) y el formato pdf, en ocasiones con opción de búsqueda de palabras dentro de los documentos.

Uno de los primeros proyectos de estas características fue la Internet Library of Early Journals (<http://www.bodley.ox.ac.uk/ilej/>), impulsada por diversas universidades del Reino Unido<sup>8</sup>. La International Coalition on Newspapers (ICON) ofrece datos de múltiples iniciativas de digitalización de prensa correspondientes a una cincuentena de países<sup>9</sup>. De cada una de ellas se indica el nombre, una breve descripción y la URL de acceso. En el caso de Italia constan L'Emeroteca Digitale, Gazzette Bolognesi [1645-1796], La Stampa y l'Unità.

6 Brunella Longo, *Gli archivi dei giornali online*, «Biblioteche oggi», vol. 24 (2006), n. 1, p. 9-21. <<http://www.bibliotecheoggi.it/2006/20060100901.pdf>>.

7 Stefano Bonanni; Elisa Fiocchi; Enrico Martellini, *Iniziative italiane di digitalizzazione di periodici a stampa*, «Biblioteche oggi», v. 27 (2009), n. 10, p. 7-14. <<http://www.bibliotecheoggi.it/content/20091000701.pdf>>.

8 Javier Díaz-Noci, *Biblioteca digital de periodistas vascos: proyecto de investigación*. Barcelona, 2007. Inedito.

9 International Coalition on Newspapers, *Newspapers digitization*: <<http://icon.crl.edu/digitization.htm>>.

Son fuentes muy útiles en investigaciones y búsquedas de informaciones de contenido histórico. La cobertura geográfica puede ir desde la ciudad a la provincia, la región, el país e incluso tener ámbito internacional. Se da el caso de diversos títulos que pueden formar parte de más de una hemeroteca histórica. Algunos ejemplos destacados son los siguientes:

– Chronicling America

<<http://www.loc.gov/chroniclingamerica/>>

Estados Unidos es el país que cuenta con mayor número de proyectos de hemerotecas históricas en marcha. Uno de los más amplios y activos es el National Digital Newspaper Program gestionado en los EUA por el National Endowment for Humanities y que cuenta con la asistencia de la Library of Congress (<http://www.neh.gov/grants/guidelines/ndnp.html>). La mayoría de las digitalizaciones realizadas proceden de microfilm. El período incluido va de 1836 a 1922. Chronicling America es el portal que permite consultar todos los periódicos del programa de forma conjunta.

– Biblioteca Virtual de Prensa Histórica

<<http://prensahistorica.mcu.es/prensahistorica/es/consulta/busqueda.cmd>>

Es la hemeroteca histórica más importante de ámbito español. Impulsada por el Ministerio de Cultura español, ofrece el contenido de unos 2.140 títulos de los siglos XVIII-XXI. Según los datos de su sitio web, en enero de 2014 contenía 6,5 millones de páginas digitalizadas. El formato de visualización era en principio jpg, ampliado posteriormente a pdf, con opciones de búsqueda por palabras dentro de las páginas de los diarios.

– L'Emeroteca Digitale della Biblioteca Nazionale Braidense

<<http://emeroteca.braidense.it/>>

Hemeroteca que tiene la finalidad de divulgar en formato digital los periódicos de interés para la historia y la cultura italiana. Integra diferentes proyectos de digitalización, como son: EPOCA (Emeroteca Politica e Culturale Antifascista), EVA (Emeroteca Virtuale Aperta), GEA (Giornali & Altro) y TRIPEDI (Trieste Periodici Digitali), conteniendo unas 960 cabeceras y un millón y medio de páginas.

The screenshot shows the website interface for L'Emeroteca Digitale. At the top, there is a navigation menu with links for Home, Presentazione, Catalogo, Ricerche, and Guida. Below the menu, there is a search bar with the text "Cerca testate per parola chiave" and a search button. A dropdown menu for "Parola chiave:" is set to "Titolo". Below the search bar, there is a row of circular buttons for each letter of the alphabet (A-Z) and a "Tutti" button. The main content area displays a table of newspaper titles with columns for "Titolo", "Posseduto", "Luogo", "Spoglie", "Istt.", and "Proa.".

Titolo (A-Z, 960 testate, pag. 1/39)	Posseduto	Luogo	Spoglie	Istt.	Proa.
Il 22 Marzo : primo giorno dell'indipendenza lombarda, giornale ufficiale	1848-1848		SP	IT:BND	EVA
Il 26 Maggio : foglio settimanale	1883-1884	Varese		IT:BND	GEA
Il 26 maggio a Caravaggio : supplemento al giornale Il Risveglio (periodico di Cremona) del 26 maggio 1894	1894-1894			IT:BND	GEA
L'Abbandonato : periodico settimanale	1881-1881	Lodi		IT:BND	GEA
Accademia di Brera : atti	1896-1948	Milano	SP	IT:BND	EVA
L'Adda : foglio che tratta degli interessi di Cassano d'Adda	1904-1904	Cassano d'Adda		IT:BND	GEA
L'Adda : giornale settimanale dei partiti popolari valtellinesi	1908-1911	Tirano		IT:BND	GEA
L'Adunata dei refrattari : pubblicazione settimanale	1942-1948	New York		IT:INSMILI	EPOCA
L'Affarista alla berlina	1880-1882	Mantova		IT:BND	GEA
L'Affondatore : giornale letterario umoristico teatrale	1879-1884	Milano		IT:BND	GEA

Los cambios producidos en los últimos años van en la dirección de ir completando los procesos de digitalización. A medio plazo se conseguirá disponer de la (casi) totalidad de la prensa histórica en formato digital, lo que constituirá un producto del máximo interés para historiadores e investigadores de prensa en general.

## OSSERVATORIO

### 2.2. Hemerotecas digitales del propio medio

La hemeroteca digital es uno de los servicios específicos del entorno digital que los diarios ofrecen a sus usuarios y que los diferencian más claramente de las ediciones impresas.

Resaltaremos aquí fundamentalmente su papel como fuentes de información, es decir, desde la posición del usuario (especialmente el profesional) que necesita localizar una o varias informaciones publicadas por un medio de comunicación en concreto. Desde este punto de vista, son fuentes de información muy importantes y, en muchos casos, insustituibles. En este sentido, algunas cuestiones a tener en cuenta de estos servicios son su alcance temporal, el coste, y las prestaciones de consulta.

Si desde los inicios de la prensa en Internet hasta el año 2002 aproximadamente, la mayoría de diarios ofrecían una hemeroteca en línea meramente “testimonial”, de los últimos días o semanas, desde entonces la cantidad de fondos accesibles ha crecido de forma notable y es claramente visible en los últimos años la tendencia por parte de las cabeceras de prensa más importantes del mundo a ofrecer la totalidad de su archivo en el sitio web del diario. Con todo, hay una gran diferencia en estos momentos entre los diarios que ofrecen grandes hemerotecas en línea, con la totalidad de sus fondos, y aquellos que siguen ofreciendo hemerotecas con pocos contenidos.

En cuanto al coste, la tendencia a ofrecer los contenidos de archivo gratuitos o de pago no está del todo clara, aunque la actual coyuntura de crisis haya agudizado desde 2009 la necesidad de cobro, y el equilibrio entre ambas opciones parezca decantarse cada vez más hacia los sistemas de pago, mediante diferentes fórmulas.

La consulta de las hemerotecas ha presentado tradicionalmente dos formas de acceso: la consulta de ejemplares anteriores íntegros por navegación cronológica (browsing), y el acceso mediante búsqueda (searching) por palabra clave utilizando un buscador. Ambos sistemas coexisten en estos momentos, si bien la preponderancia del acceso por buscador es cada vez más clara. La calidad de las prestaciones de consulta ha ido mejorando a lo largo de los últimos años, aunque incluso en el caso de algunos diarios importantes la funcionalidad y la calidad del sistema de búsqueda y recuperación puede llegar a ser bastante deficientes.

También se da la paradoja de que, en ocasiones, la tendencia de muchos sitios web hacia la simplificación de sus sistemas de búsqueda con el objeto de ponerles las cosas fáciles al usuario medio (por ejemplo, haciendo desaparecer las opciones de búsqueda avanzada o algunos operadores booleanos), haya supuesto desde el punto de vista del uso profesional, un empobrecimiento y retroceso considerable incluso en medios considerados de referencia<sup>10</sup>.

– The New York Times

<<http://www.nytimes.com/ref/membercenter/nytarchive.html>>.

El diario norteamericano tiene un archivo en línea desde 1851. El servicio de archivo está dividido en dos bloques, 1851-1980; y 1981-actualidad, y cuenta con opciones de búsqueda avanzada y un completo sistema de ayuda. Es una de las hemerotecas de diarios de referencia a nivel internacional

– La Vanguardia

<<http://www.lavanguardia.com/hemeroteca/index.html>>

Ofrece la totalidad del fondo desde la fundación del diario en 1881. Combinan los contenidos gratuitos (menos los últimos días, en formato pdf) con los contenidos de pago para suscriptores. Además, destaca por la variedad de servicios y contenidos de la sección de Hemeroteca, como efemérides, reportajes históricos, perfil en redes sociales, y otros.

**10** Javier Guallar; Ernest Abadal; Lluís Codina, *Hemerotecas de prensa digital: evolución y tendencias* «El profesional de la información», vol. 21 (2012), n. 6, p. 595-605. <<http://eprints.rclis.org/18199/>>.

- La Repubblica

<<http://ricerca.repubblica.it/>>

Tiene buenas prestaciones en su sistema de consulta y en la página de resultados, con opciones, por ejemplo, de búsqueda avanzada, acotación por colecciones, gestión de resultados, agrupación de resultados por categorías, etc. La cobertura temporal del archivo tanto por navegación como en el buscador, es de treinta años, hasta el año 1984, alcanzando por tanto buena parte de su fondo aunque no su totalidad (fue fundado en 1976).

**la Repubblica.it** Archivio

Home Pubblico EconomiaFinanza Sport Spettacoli Cultura Motori Viaggi Moda Ca

tiscali: **17.95 €/MESE PER 2 ANNI!**

RICERCA **RICERCA AVANZATA**

ARCHIVIO

2008	2009	2010	2011	2012	2013	2014
Gennaio						
Febbraio						
Marzo	Marzo	Marzo	Marzo	Marzo	Marzo	Marzo Aprile
Aprile	Aprile	Aprile	Aprile	Aprile	Aprile	Maggio
Maggio	Maggio	Maggio	Maggio	Maggio	Maggio	Giugno Luglio
Giugno	Giugno	Giugno	Giugno	Giugno	Giugno	Agosto
Luglio	Luglio	Luglio	Luglio	Luglio	Luglio	Settembre
Agosto	Agosto	Agosto	Agosto	Agosto	Agosto	Ottobre
Settembre	Settembre	Settembre	Settembre	Settembre	Settembre	Novembre
Ottobre	Ottobre	Ottobre	Ottobre	Ottobre	Ottobre	Dicembre
Novembre	Novembre	Novembre	Novembre	Novembre	Novembre	
Dicembre	Dicembre	Dicembre	Dicembre	Dicembre	Dicembre	

PERSONE

- matteo renzi (227)
- enrico letta (155)
- silvio Berlusconi (130)
- berlusconi (85)
- giorgio napolitano (69)
- beppe grillo (66)
- davis (57)
- verona (55)
- garcia (53)

ENTI E SOCIETÀ

- facebook (219)
- senato (210)
- twitter (204)
- sebastien (162)
- di freud (158)
- buyers club (149)
- governo (130)
- google (109)
- ue (93)

LUOGHI

- roma (2065)
- italia (1704)
- milano (1041)
- napoli (706)
- torino (473)
- firenze (462)
- bologna (394)
- new york (378)
- europa (368)

En esta tipología se dispone prácticamente ya de todos los fondos en Internet. La gran mayoría de cabeceras ya ha completado la fase de digitalización de sus colecciones retrospectivas y los ofrecen a suscriptores y usuarios en general como complemento a la información diaria.

### 2.3. Bases de datos de prensa

Las bases de datos de prensa se caracterizan por realizar el análisis y la indización de las noticias aparecidas en centenares e incluso miles de medios y por disponer de sistemas avanzados de consulta, que suelen ser muy superiores a los que ofrecen los propios medios (aunque las diferencias se hayan ido reduciendo con los años).

Están dirigidas a un usuario profesional, no sólo por sus contenidos sino también especialmente debido a su coste y además pueden ofrecer también servicios de seguimiento de prensa.

Además, llevan a cabo sus servicios mediante acuerdos contractuales con los medios, con lo cual sus páginas de resultados proporcionan el acceso directo a los documentos, y no a una lista de enlaces, como sucede con otras fuentes de información de prensa. En suma, se pueden considerar como los mejores sistemas de búsqueda existentes para localizar contenidos periodísticos.

Tienen una importante tradición e implantación en el ámbito anglosajón y han tenido una introducción más reciente en otras áreas. Las más importantes internacionalmente son Lexis Nexis y Factiva, diferenciadas claramente del resto; en español es muy relevante el caso de la base de datos de prensa MyNews, presente en la

## OSSERVATORIO

mayoría de las empresas de comunicación de nivel medio-alto de España. Google News, sistema muy conocido y utilizado, no se encuentra dentro de esta categoría porque carece de sistema de búsqueda retrospectiva.

- Lexis Nexis

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

Está considerada la mayor base de datos comercial del mundo. Además de información de prensa, tiene información económica, fiscal y legislativa, con más de 3.000 millones de documentos de unas 42.000 fuentes de información. El apartado de noticias de actualidad incluye unas 9.700 fuentes, de las cuales 2.100 son diarios. Por ejemplo, en español el número de fuentes es superior a 400.

- Factiva

<<http://global.factiva.com>>

Creada en 1999 por la agencia de noticias Reuters, que anteriormente comercializaba Reuters Business Briefing, y la empresa Dow Jones. Además de información periódica, contiene información económica y financiera. Da acceso al contenido de más de 10.000 fuentes, de las cuales unas 230 son en español.

- Mynews

<[www.mynews.es](http://www.mynews.es)>

Mynews es la base de datos con contenidos de prensa española más importante. Tiene también servicio de seguimiento de prensa. Creada en 1996 (desde 1998 en formato web), ofrece el texto completo y los pdf de unas 200 fuentes de prensa impresa y también de más de 800 fuentes exclusivamente digitales, con una cobertura temporal que puede llegar para algunos diarios hasta 1996. Los servicios de hemeroteca tienen las prestaciones documentales más altas entre los productos españoles, con tres opciones de búsqueda: simple, avanzada y profesional. Es probablemente la herramienta en documentación de prensa española más utilizada en el sector periodístico por profesionales de la información.

The screenshot shows the MY NEWS HEMEROTECA website interface. At the top, there are language options (Español, Català, English) and a 'Usuario' button. The date '21 Feb 2014' is displayed. Below this is a navigation bar with 'Buscador', 'Qué es', 'Cómo funciona', and a phone number '(+34) 933 568 410'. The main content area features a search bar with the text 'Encuentra todas las noticias publicadas' and 'Más de 136.448.155 artículos'. Below the search bar is a search input field with the placeholder 'Busca noticias, personas, documentos...', a dropdown menu set to 'Hoy', and a 'Buscar' button. There are also links for 'Búsqueda avanzada' and 'Búsqueda profesional'. A statistics table is shown below the search area:

Usuarios	Consultas realizadas	Fuentes	Noticias desde 1996
655	414.568	1.072	

Below the statistics, there are two sections: 'Las más vistas de las últimas 24h' and 'Términos más buscados 24h'. The 'Las más vistas' section lists several news items, including 'EL ACTOR BILBAÍNO SE EXPONE Y SE DESHIBE AL MÁXIMO EN EL SHOW MUSICAL 'EL INTERPRETE', QUE LLEGA MAÑANA AL GAYARRRE (21 HORAS)'. The 'Términos más buscados' section lists terms like 'nueva', 'nuevo', 'metalla', 'azul', 'azora', 'amadeus', 'tc-street', 'alemania', 'otro', 'aplicación', 'metalla', 't-ventu', 'nou', 'estad', 'sí', 'quiero', 'el', 'compromiso', 'de', 'la', 'familia', 'nuevo', 'estado', 'telefonía', 'agr', 'gucci', 'telefonía', 'alpha', 'deady', 'gpeuro', 'permilloc', 'enrique', 'tomás', 'surgenia', 'united', 'collegio', 'benjin', 'movil', 'app', 'vodafone', 'onda', 'joven', 'fm'.

La base de datos de prensa es un producto dirigido también a profesionales del sector de la comunicación (periodistas y documentalistas de empresas de medios), y a gabinetes de prensa y servicios de información de empresas y organismos de diversos ámbitos,

incluyendo partidos políticos y administraciones públicas, por ejemplo. Aparece con pocos cambios en los últimos años, salvo mejoras continuadas en las interfaces de usuario, cada vez más amigables. Presenta pocas expectativas de variación a corto y medio plazo, y representa una incógnita a largo plazo en un horizonte de plena digitalización de la noticia desde su misma creación y en el futuro ecosistema de la web móvil.

### 3. ¿Cómo evaluar los sistemas de documentación de prensa?

La segunda parte de nuestros estudios se ha orientado a establecer indicadores para la evaluación de los servicios de consulta de prensa y también para aplicarlos a conjuntos diversos de sistemas de documentación de prensa (hemerotecas digitales y bases de datos de prensa).

#### 3.1. Sobre los indicadores

Para el establecimiento de los indicadores se ha partido de los estudios previos de evaluación realizados para hemerotecas digitales así como también para recursos web y bases de datos en general. Por otro lado, la propuesta se ha contrastado con la realidad mediante su aplicación a diversas muestras de hemerotecas digitales (españolas y catalanas, fundamentalmente) lo que ha permitido su ajuste. De esta forma, nuestra propuesta combina la vertiente teórica con la práctica.

A partir de estos antecedentes, hemos establecido un conjunto de 27 indicadores articulados en cuatro grandes apartados: aspectos generales, contenidos, sistema de consulta y presentación de resultados. En un trabajo anterior publicamos su descripción completa<sup>11</sup> aunque en otros textos<sup>12</sup> había tenido formulaciones distintas, con menor número de indicadores.

Apartado	Indicador
A. Aspectos generales	A.1. Denominación A.2. Ubicación del acceso a la hemeroteca A.3. Información sobre la hemeroteca A.4. Coste A.5. Productos documentales
B. Contenidos	B.1. Colecciones disponibles B.2. Formatos disponibles B.3. Cobertura temporal
C. Sistema de consulta	C.1. Tipos de consulta por interrogación (palabras clave) C.2. Lenguaje de consulta C.3. Visualización de los índices C.4. Ajuste de relevancia C.5. Acotación temporal C.6. Acotación en colecciones C.7. Acotación en formatos C.8. Acotación por campos del registro C.9. Reutilización de búsquedas C.10. Consulta de tesauros u otros vocabularios controlados C.11. Acceso por navegación cronológica ( <i>browsing</i> ) C.12. Ayudas

11 Javier Guallar; Ernest Abadal; Lluís Codina, *Sistema de análisis de hemerotecas de prensa digital*, «Trípodos», vol. 21 (2012), n. 6, p. 37-64. <<http://eprints.rclis.org/19929/>>.

12 Javier Guallar; Ernest Abadal, *Evaluación de hemerotecas de prensa digital: indicadores y ejemplos de buenas prácticas*, «El profesional de la información», vol. 18 (2009), n. 3, p. 255-269. <<http://eprints.rclis.org/13048/>>. Javier Guallar; Ernest Abadal, *The digital press archives of the leading Spanish online newspapers*, «Information Research», v. 15 (2010), n. 1, paper 424. <<http://informationr.net/ir/15-1/paper424.html>>.

## OSSERVATORIO

Apartado	Indicador
D. Presentación de resultados	D.1. Gestión de las listas de resultados D.2. Campos visualizados del registro del documento D.3. Agrupación de resultados por categorías D.4. Identificación de los términos de búsqueda en el documento D.5. Formatos de visualización de un mismo documento. D.6. Gestión de los documentos recuperados D.7. Presentación de documentos relacionados

Tabla 1. Indicadores para la evaluación de hemerotecas

Todos los indicadores se pueden ponderar para la valoración de los servicios, ya que su presencia, ausencia y grado de desarrollo se pueden considerar valores objetivos en la evaluación de los sistemas, con la única excepción del indicador 'coste', que solamente se ha incluido en la relación a título informativo.

En un método de evaluación hay que considerar si se establece una asignación numérica que permita realizar análisis cuantitativos, y por ejemplo, en el caso que nos ocupa, rankings de diarios. En esta propuesta se utiliza una escala basada en los tres valores característicos de los análisis de usabilidad (Nielsen):

- El indicador es deficiente: 1
- El indicador es correcto pero admite mejoras: 2
- El indicador es excelente: 3

Además, hay un cuarto valor para el caso de que el indicador sea inexistente o su implementación sea tan deficiente que no pueda tomarse en consideración. En ambos casos, el valor asignado es o (cero). Asimismo, algunos indicadores no son escalares, y no se les ha asignado por ello una puntuación, incluyéndose por tanto a título informativo. En nuestro caso son solamente dos: los indicadores de 'Denominación' y 'Coste'.

En este método no se incluyen ponderaciones relativas entre indicadores o entre bloques o secciones de indicadores. El motivo es que toda ponderación es inevitablemente discrecional cuando no subjetiva, ya que responde a valores, aunque sean muy legítimos o justificados. En todo caso, entendemos que mientras un método de evaluación puede ser de aplicación general, o abierta, toda ponderación está vinculada a unos objetivos concretos. Por ejemplo, ¿el equipo investigador está comparando la utilidad de diversas hemerotecas a fin de dar soporte a investigaciones históricas de tipo académico? En tal caso, la ponderación debería ayudar a destacar las hemerotecas con acceso abierto a colecciones más amplias y de mayor rango cronológico, etc. Por este motivo, entendemos que el sistema que proponemos aquí puede ser objeto de ponderación cuando se aplique con unos objetivos concretos.

### 3.2. Sobre la evaluación de hemerotecas digitales

La tabla de indicadores se ha aplicado a los diarios españoles de mayor difusión, en dos etapas, y también a la prensa catalana. Otros investigadores también la han utilizado para valorar hemerotecas de prensa del País Vasco<sup>13</sup> e internacional<sup>14</sup>.

**13** Idoia Camacho; Teresa Agirreazaldegi; Milagros Ronco; Carmen Peñafiel, *Evolución o revolución?: Adaptación de los servicios de documentación a la era digital*. In: *II Congreso Internacional de la Asociación Española de Investigadores de la Comunicación (AE-IC)*, Málaga, 3 - 5 febrero 2010. <<http://www.aeic2010malaga.org/upload/ok/243.pdf>>.

**14** Federico Nanni, *L'archiviazione delle pagine dei quotidiani online*. «Diacronie. Studi di Storia Contemporanea: Spazi, percorsi e memorie», n. 15, 29/10/2013. <[http://www.studistorici.com/2013/10/29/nanni\\_numero\\_15/](http://www.studistorici.com/2013/10/29/nanni_numero_15/)>.

Nuestro primer estudio de evaluación de hemerotecas<sup>15</sup> se realizó sobre una muestra de 12 diarios de Catalunya en base a 20 indicadores, y mostraba muchas diferencias entre los diarios mejor y peor situados en el ranking.

En un estudio posterior<sup>16</sup> se seleccionó una muestra con los 12 medios de mayor difusión en España y se les aplicaron los indicadores durante los meses junio-agosto de 2007. Los resultados obtenidos indican que las hemerotecas de los diarios no alcanzan el nivel de los productos comerciales especializados (las bases de datos de prensa). Se establecieron cinco niveles de desarrollo de los archivos de prensa en función del cumplimiento de los indicadores analizados: medio-alto (El País, El Mundo, La Vanguardia); medio (Abc, El Periódico); medio-bajo (La Verdad, As, El Mundo Deportivo, Sport); bajo (20 minutos, Libertad digital, Periodista digital) e inexistente (Marca).

Como continuación del anterior análisis<sup>17</sup> se muestran los resultados de un análisis realizado en 2010 sobre la misma muestra de diarios del estudio anterior, y se relacionan los resultados obtenidos en ambos estudios. Se señalan los aspectos mejores y peores de las hemerotecas de prensa digital, se establece un ranking de diarios en el que ocupan las primeras posiciones El mundo, Abc, El país y La vanguardia, y se destacan 5 grandes tendencias observadas: ampliación del fondo temporal online, hemerotecas gratuitas versus hemerotecas de pago, simplificación (y empobrecimiento) de la búsqueda, filtros en la presentación de resultados y productos documentales en la sección Hemeroteca.

Por su parte, Idoia Camacho et al.<sup>18</sup> utilizan este método para evaluar las hemerotecas de la prensa digital del País Vasco, recogiendo la mayoría de nuestros indicadores y obtienen que casi todas las hemerotecas de los diarios digitales vascos tienen una valoración global media-baja, exceptuando la de El Diario Vasco.

Asimismo, Acosta-Valdés et al.<sup>19</sup> recogen nuestros indicadores correspondientes a la página de resultados en su método de análisis de la organización de información de los diarios, que aplican sobre una pequeña muestra de diarios de varios países.

Por último, Federico Nanni<sup>20</sup> también utiliza nuestro sistema de indicadores de evaluación y lo aplica a una muestra de 30 hemerotecas de prensa de los principales medios internacionales.

#### 4. Estudios futuros

Se considera importante el mantenimiento y la actualización en el futuro tanto del método de análisis de hemerotecas como de la clasificación de fuentes de información de prensa digital, ya que al ser la propia Web un entorno cambiante, lo es tam-

15 Ernest Abadal; Javier Guallar, *Les hemeroteques digitals de la premsa catalana: anàlisi dels diaris catalans de més difusió*. In: Jornades Catalanes de Documentació (11es: Barcelona, 22-23 maig 2008). Barcelona: COBDC, 2008, p. 149-162. <<http://eprints.rclis.org/11696/>>.

16 Javier Guallar; Ernest Abadal, *The digital press archives of the leading Spanish online newspapers*, cit.

17 Javier Guallar; Ernest Abadal; Lluís Codina, *Hemerotecas de prensa digital: evolución y tendencias*, cit.

18 Idoia Camacho; Teresa Agirreazaldegui; Milagros Ronco; Carmen Peñafiel, *Evolución o revolución?* cit.

19 Misleiny Acosta-Valdés; Zulia Ramírez-Céspedes; Liliam Marrero-Santana, *Organización de la información en sitios web periodísticos*. «El profesional de la información», vol. 20 (2011), n. 1, p. 54-60. <<http://eprints.rclis.org/15756/>>.

20 Federico Nanni, *L'archiviazione delle pagine dei quotidiani online*, cit.

## OSSERVATORIO

bién el sector de los cibermedios, y con ellos, sus hemerotecas. La ventaja para estas nuevas exploraciones es que pueden partir del conjunto de indicadores que aquí se han discutido como un punto de partida sólido, aunque sea para poder centrarse únicamente en los aspectos más novedosos.

Una segunda posibilidad importante son los estudios longitudinales, de modo que la continuación del método de trabajo que aquí se ha presentado se puede efectuar en esta misma muestra de diarios en el futuro para ver su evolución.

Por último, también se puede adaptar este método de evaluación de hemerotecas de prensa a otras colecciones de fondos documentales de medios de comunicación. Es el caso de los archivos digitales de televisiones en Internet o videotecas, y de los archivos digitales de radios en Internet o audiotecas. En el primer caso, se cuenta ya con el estudio de Laura Anton y Javier Guallar<sup>21</sup> en el que se estudian los archivos en Internet de una serie de cadenas de televisión españolas, a partir de un conjunto de indicadores similares a los aquí presentados junto a otros específicos para contenidos audiovisuales.

## Agradecimientos

Este trabajo ha contado con el apoyo de la AGAUR, Grupo de Investigación Consolidado “Cultura y Contenidos Digitales” 2009-SGR-177, y de los proyectos del Ministerio español de Economía y Competitividad CSO2012-39518-Co4-01 y CSO2012-39518-Co4-02.

Ai fini della citazione, utilizzare esclusivamente il testo in lingua italiana, che presenta il DOI, la paginazione, l'abstract e gli altri dati ufficiali.

When citing, please always refer to the Italian translation only, complete with DOI, page numbers, abstract and other data.

[Ernest Abadal, Javier Guallar, Lluís Codina, *Sistemi di documentazione della stampa periodica: quali sono e come valutarli?*

Aib studi, vol. 54 n. 1 (gennaio/aprile 2014), p. 75-86, DOI 10.2426/aibstudi-9486.]

**21** Laura Anton; Javier Guallar, *Análisis de los archivos audiovisuales en internet de las televisiones autonómicas españolas*, «Revista española de documentación científica», v. 37 (2014), n. 1.