



HEMEROTEQUES. ESTAT ACTUAL I PERSPECTIVES

PROYECTO DE DIGITALIZACIÓN DE PRENSA EN LA BIBLIOTECA NACIONAL



PROYECTO DE
DIGITALIZACIÓN
DE PRENSA EN
LA BIBLIOTECA
NACIONAL

Dolores Rodríguez, Jefe de Sección de Sala de Publicaciones Periódicas. Biblioteca Nacional

FICHA TÉCNICA:

BIBLIOTECA NACIONAL. MADRID

- Sala de Publicaciones Periódicas
- Microformas y Prensa Digitalizada

Dirección: Recoletos, 20-22. 28071 Madrid

Tel. / 91-5807793/94 Fax: 91-8168009

Correo electrónico: infoseriadas@bne.es

Horario de atención al público:

lunes a viernes, 9-21 h. sábados, 9-14 h.

Accesos: Autobuses: 1, 5, 9, 19, 21, 27, 37, 45, 51,

53, 74, 150 - Metro: Colón - Cercanías: Recoletos

Personal: En la Sala: 7 bibliotecarios y 14 auxiliares

.En el Servicio de Publicaciones Seriadas:

27 bibliotecarios y 13 auxiliares.

Superficie de depósitos:

20.600 metros lineales ocupados (en la sede de Recoletos).

62.140 metros lineales ocupados (en la sede de Alcalá).

Sala de lectura: 52 puestos para consulta en papel,

10 puestos para CD-ROM, 25 puestos para

microforma y 6 puestos de consulta para OPAC e

Internet. Fondos: 111.000 títulos

INTRODUCCIÓN

La Biblioteca Nacional, en su función conservadora del patrimonio bibliográfico español, ha venido invirtiendo enormes esfuerzos en la preservación de sus colecciones. Comprobado el avanzado estado de deterioro en que encontraba gran cantidad de volúmenes de prensa, se pensó en tomar una serie de medidas para detener el daño que la manipulación y el uso indebido producen en un material tan frágil como es el papel de periódico, ya que, en aquel momento, no se podía llevar a cabo la restauración del soporte. Dichas medidas consistieron en prohibir la realización de fotocopias directas de los ejemplares, práctica muy extendida, obligando a realizar todas las reproducciones en el laboratorio, por personal cualificado y mediante microfilm, y se continuó con la encuadernación sistemática, no permitiendo la consulta de la colección procedente de depósito legal hasta que los ejemplares no estuviesen convenientemente encuadernados.

Paralelamente, la política de conservación de la Biblioteca se centró en abordar la preservación de los contenidos mediante su reproducción en otros soportes. En 1991 comenzó el programa de microfilmación de prensa, que en un primer momento se centró en la prensa anterior a 1945, pero posteriormente, tras la integración de la Hemeroteca Nacional en la Biblioteca, hubo que dar prioridad a las cuestiones relativas al ahorro de espacio en los depósitos y a la preservación de los títulos de prensa actual, que sufrían un gran deterioro debido a su continuo uso en las salas de lectura. Hasta ahora se ha microfilmado la práctica totalidad de las cabeceras españolas custodiadas por la Biblioteca hasta 1998, y también se han microfilmado colecciones de revistas.

La microfilmación se realiza en rollos de 35 mm, sobre película de seguridad en sales de plata, y conforme con todas las normas ISO pertinentes

Las cifras totales correspondientes al fondo microfilmado entre 1991 y el ejercicio 2002/2003 son las siguientes:

Títulos: 1.631

Volúmenes: 58.776

Fotogramas: 23.680.232

Rollos de microfilm: 41.711

Microfichas: 7.213

En 1999 se inicia el Programa de Digitalización. Dado que la microfilmación abarca, por regla general, hasta 1998, se decidió que la digitalización comenzara a partir de 1999 para no duplicar esfuerzos, salvo los casos de adquisición digital de colecciones retrospectivas, en los que se ha adquirido la colección



1,
Sala microformas.

2,
La portada de
La Renaixensa
de 1896





PROYECTO DE DIGITALIZACIÓN DE PRENSA EN LA BIBLIOTECA NACIONAL

completa debido a la facilidad de la consulta y a la posibilidad de completar nuestra colección en alguna ocasión. Las listas de títulos microfilmados y digitalizados están accesibles en el web de la BN (www.bne.es).

A continuación examinaremos más detalladamente el Proyecto de Digitalización.

PROYECTO DE DIGITALIZACIÓN

En 1997 la Biblioteca realizó las primeras tentativas en materia de digitalización de prensa, microfilmando y digitalizando a partir del microfilm 187 títulos del siglo XIX y mediados del XX, pero no se generó ninguna herramienta útil para visionar las imágenes, por lo que la consulta era demasiado complicada y estas primeras pruebas se desestimaron. En 1998 se había dado una solución a la consulta de los fondos retrospectivos de prensa mediante el Programa de Microfilmación, que se encontraba ya en fase muy avanzada, pero con motivo de las obras de remodelación del edificio de la Biblioteca hubo de cerrar la "Sala de Consulta de Prensa Diaria" y buscar una solución para la consulta de la prensa actual.

Comenzaban entonces a aparecer los primeros proyectos de digitalización de prensa, como el de ABC, que digitalizó su colección retrospectiva y la editó en cd-rom. En ese momento en la Biblioteca se pensó que no debíamos permanecer al margen de las nuevas tecnologías en reproducción de documentos y en la posibilidad de llevar a cabo un Proyecto de Digitalización como solución a la consulta de la prensa actual.

Objetivo

Aunque se ha discutido mucho sobre la primacía o no de la digitalización frente a la microfilmación, en principio la Biblioteca Nacional no se planteó la digitalización como alternativa a la microfilmación, sino como un proceso paralelo y complementario.

El principal objetivo de este proyecto no era la conservación a largo plazo, sino facilitar la consulta de los fondos de prensa actual, proporcionando un rápido acceso y evitando las esperas a las que obliga la encuademación y/o microfilmación de esos títulos.

Para llevar a cabo este objetivo, en 1999 se comenzó por seleccionar, para su digitalización, 108 cabeceras de la prensa que ingresaba en la Biblioteca por depósito legal. Esta selección se hizo conforme a tres parámetros:

- Exhaustividad, intentando abarcar el mayor número de títulos existentes en ese momento
- Representatividad, seleccionando títulos que representaran la prensa de todas las Comunidades Autónomas, incluyendo al menos un título de cada una de ellas.
- Uso, priorizando aquellos títulos que estadísticamente se revelaron como los más consultados en la Biblioteca.

Estándares

Se tenía muy claro que las prescripciones técnicas debían responder a estándares reconocidos en el mundo bibliotecario, tanto para facilitar el posible intercambio de imágenes con otras instituciones, como para conseguir la mayor permanencia de éstas en el tiempo, por ello se siguieron las recomendaciones del programa "Memoria del Mundo" de la UNESCO, que señala las siguientes características para la digitalización de periódicos: resolución mínima de 300 p.p.p., 1 bit por pixel, el formato de fichero de las imágenes es TIFF, con comprensión CCITT Grupo IV. También se buscó asesoramiento de expertos en la materia, a fin de perfilar los criterios que debían regir el futuro proyecto.

Almacenamiento

Las imágenes resultantes del proceso de digitalización se almacenan en cd-rom. De todas ellas se conservan dos copias, una de ellas como máster de seguridad desde el que se pueda migrar la información a cualquier otro soporte que garantice su pervivencia. Este máster se guarda en otro depósito de conservación, sito en Alcalá de Henares.

Aunque algunas voces hablan de la necesidad periódica de migrar a otro soporte para conservar la información, por ahora no se ha observado ningún deterioro en nuestra colección como para plantearnos esta necesidad, aunque no se descarta la migración a soportes con mayor capacidad, DVD, para responder tanto a la conservación como al ahorro de espacio.

Coste y resultados

Desde 1999 la Biblioteca ha invertido en el Proyecto de Digitalización unos **105.000.000** de pesetas, además de los costes de hora/hombre dedicadas a tareas necesarias para llevar a cabo el proyecto y de las inversiones en adquisición de copias digitalizadas por otros centros. Esta inversión se ha materializado en más de 16.000.000 de imágenes, almacenadas en unos 20.000 cd-rom. Para los próximos años (2004-2005) hay prevista una inversión en este proyecto de 450.000 €.

3,
www.bne.es

4,
portada
Las Provincias



ARQUITECTURA DEL SISTEMA ACTUAL

Actualmente nos hallamos implementando una nueva versión de la primera aplicación denominada "Visor Hemerográfico". Se trata de un sistema basado en arquitectura cliente/servidor y su diseño se ha basado en mejorar tres aspectos que nos resultan básicos:

- La capacidad para gestionar archivos de distintos formatos (tiff y pdf) y para disponer de acceso tanto en línea como en monopuesto, lo que nos permite controlar las **359** cabeceras (unos 24.000 discos) que actualmente tenemos digitalizadas e incrementar la colección con imágenes procedentes de distintas instituciones sin necesidad de utilizar aplicaciones diferentes para cada una.
- La optimización del rendimiento en la consulta. El usuario se evitará realizar peticiones, manipular los discos y esperar a ser atendido en momentos de mucha demanda, lo que redundará en mayor rapidez y eficacia.
- El ahorro de espacio para la ubicación de los discos en la sala y de personal para servir las peticiones.

Esta nueva aplicación está compuesta de los siguientes elementos:

1. **Servidor:** Con el fin de no tener que manipular los discos se ha comenzado a grabar imágenes en un servidor conectado mediante red local a los terminales de consulta. La gestión de este servidor la llevará a cabo el propio personal de la Sala de Consulta, que por tanto podrá decidir cuando crea conveniente cualquier cambio de las imágenes albergadas en el servidor.
2. **Gestión de imágenes:** Las imágenes digitales están asociadas a una base de datos de tipo relacional. Esta base de datos contiene el directorio de todas las imágenes e informaciones asociadas a ellas e incluye un programa de gestión para todos los títulos y discos que permite su actualización. Esta base de datos cuenta con los siguientes campos:

- Título de la publicación
- Referencia de signatura
- Número secuencial de disco de cada título
- Etiqueta de volumen de discos
- Primera fecha del disco
- Última fecha del disco

· Incidencias: falta de fechas y/o páginas, mal estado de páginas, errores de fechas o numeración.

3. **Subsistema de consulta:** Encontramos en él tres utilidades que cubren distintas necesidades en la consulta de imágenes digitales:

· **Acceso a la información:** existen dos formas de recuperación de la información: cuando se conoce el título y fecha que se desea consultar se accede al índice de títulos, que a su vez dispone de un calendario con todas las fechas disponibles; esta opción incluye la consulta de varias fuentes a la vez para cotejo y de búsquedas por lugar de edición, muy útil cuando se busca prensa de alguna provincia concreta sin conocer los títulos. Otra opción es realizar búsquedas a texto completo mediante lectura OCR de cada imagen, esta funcionalidad permite múltiples combinaciones de búsqueda sobre una o varias fuentes simultáneamente con posibilidad de acotar por periodos y genera un índice de resultados desde el que se puede acceder al texto de la noticia o a la imagen digital. Cuando solicitamos un documento que aún no ha sido grabado en el servidor, el sistema lo indica con un mensaje donde aparece el título de la cabecera y el número del cd-rom que se debe insertar para consultar el documento.

· **Visualización de imágenes:** las pantallas de visualización se caracterizan por su claridad, ya que desde el principio se ha primado la facilidad de uso y se han ido introduciendo numerosas adaptaciones para paliar las posibles dificultades en su manejo, así en lugar de iconos, muchas veces inexplicables, cada botón indica su uso. Además se han simplificado los pasos a seguir permitiendo la consulta de los índices desde cualquier pantalla de la aplicación, reduciendo al mínimo los golpes de ratón necesarios para llegar a la página seleccionada y con el apoyo de pantallas de ayuda.

Las posibilidades de navegación dentro de la búsqueda realizada son completas, avanzar, retroceder en páginas, fechas, títulos, etc. , ir a título, fecha o página, etc.

Las posibilidades de mostrar los documentos son varias: podemos ver el ejemplar completo en miniaturas, podemos ver el documento página a página, ampliar o reducir la totalidad o partes seleccionadas de texto, y también permite la presentación en dos imágenes respetando el aspecto del original.

· **Obtención de documentos:** Una vez que hemos localizado y visualizado el documento tenemos varias posibilidades de obtención de copias: podemos imprimir las imágenes directamente, toda la página o una selección ampliada de ella. En las copias resultantes aparece por defecto la información de la fuente, fecha y página, a no ser que decidamos su exclusión con la intención de generar un facsímil. También podemos exportar imágenes al portapapeles de Windows o a un fichero de imágenes para generar nuestro propio dossier de prensa.



PROYECTO DE DIGITALIZACIÓN DE PRENSA EN LA BIBLIOTECA NACIONAL

5,
El objetivo de la digitalización de los fondos de la hemeroteca posteriores a 1999 no era la conservación a largo plazo, sino facilitar la consulta de prensa actual, proporcionando un rápido acceso y evitando las esperas a las que obliga la encuadernación y/o microfilmación de





PROYECTO DE DIGITALIZACIÓN DE PRENSA EN LA BIBLIOTECA NACIONAL

CONCLUSIONES

Después de trece años de vida (1991-2003) podemos analizar con cierta perspectiva los resultados obtenidos con la política de conservación de la prensa. Los dos proyectos (microfilmación y digitalización) han contribuido a que la Biblioteca Nacional, como cabecera del sistema bibliotecario español, pueda facilitar a cuantas instituciones lo deseen copia de sus colecciones reproducidas, previo pago de los costes, con el objetivo de aunar esfuerzos en materia de preservación y localización de las colecciones de prensa y evitando siempre duplicar trabajo y costes. Centrándonos en el Proyecto de Digitalización (1999-2003) y en los beneficios que ha aportado a los servicios de la Biblioteca, se pueden señalar principalmente dos:

Beneficios en materia de conservación: Aunque, como señalaba al principio, el principal objetivo del proyecto no era la conservación, hemos podido comprobar el enorme impacto que sobre la colección de prensa ha tenido la digitalización. Actualmente la colección original, que no se emplea para el escaneado, se ordena en cajas recubiertas interiormente por papel de ph neutro y se guardan en el depósito de conservación en posición horizontal. Este tratamiento conserva los ejemplares exactamente en su estado original sin sufrir ningún tipo de manipulación (guillotinas, encuadernaciones defectuosas, etc.) y protegidos de agentes contaminantes que degradan el soporte (polvo, humedad, microorganismos, etc.). Además al retirarse de la circulación no se ven sometidos al uso, que es el factor de deterioro más importante. Otra ventaja importante es que se dispone de una colección digital de reserva, o máster, que en caso de pérdida o rotura de un cd-rom, puede emplearse para realizar la copia, sin necesidad de utilizar el original.

Beneficios en el acceso a la información: Debido a la falta de espacio en los depósitos del edificio de Recoletos, la prensa se trasladó al segundo depósito en Alcalá de Henares, lo que suponía que los usuarios tenían que solicitarla y esperar un día para poder consultarla. Además la encuadernación retrasaba en gran medida la disponibilidad de los números más recientes, y aunque los títulos importantes pudieran consultarse en otros centros, nos encontrábamos con que era muy difícil acceder a las cabeceras regionales en un plazo de tiempo razonable, lo que creaba bastantes decepciones en muchos usuarios. También hay que tener en cuenta que la investigación en prensa ha experimentado un incremento considerable y a menudo no se podía servir una publicación por estar siendo consultada por otro usuario, a lo que había que añadir las solicitudes de reproducción de algún ejemplar (lo que es muy frecuente) que implicaba la retirada de éste para pasar al laboratorio. Todo esto hacía que la disponibilidad disminuyera cada vez más. Con la digitalización se ha acortado notablemente el plazo de tiempo en que la prensa está disponible, pero además, los discos no salen de la Sala de Consulta y las copias las realiza el mismo usuario en el momento de la consulta. Próximamente, con la nueva aplicación, varios usuarios podrán acceder al mismo tiempo a las mismas imágenes y con la aplicación de la técnica de OCR, que nos permite la recuperación por texto libre, podremos solucionar, en gran medida, una de las más importantes carencias con la que se encuentran los investigadores que trabajan con la prensa, la falta de índices temáticos.

Hay que decir, por último, que nuestros usuarios también han sabido apreciar el esfuerzo que la Biblioteca ha realizado para facilitar sus consultas, y así lo hemos podido constatar por medio de encuestas y con los datos comparativos del uso de la colección digital respecto a otros soportes, como se recoge en la siguiente tabla.

Nº DE CONSULTAS POR AÑO Y SOPORTE *

	1999	2000	2001	2002	2003
CD	2.423	8.074	14.328	21.021	26.403
Micros	28.973	30.870	30.085	32.340	24.616
Papel	69.668	41.637	63.896	46.315	38.788

* nº total de consultas por soporte (incluye prensa y revistas).

El porcentaje de revistas en CD y micro es de aproximadamente un 1%

6,
gráfica estadística
nº de consultas

Política actual: Está previsto continuar con el Proyecto de Digitalización de la prensa actual, y ya se ha aprobado el gasto para los dos próximos años.

Además, la Biblioteca suscribe las reproducciones digitales que los propios periódicos realizan cada vez en mayor número y les presta su apoyo, siempre que es posible. En este sentido sería deseable que cundiese el ejemplo de cabeceras como ABC o las de la prensa balear, que se editan en formato digital y cumplen con el depósito legal.

Terminada la microfilmación de prensa retrospectiva, las actuaciones en materia de conservación se centrarán en el tratamiento sistemático de la colección de revistas. En cada caso se realizará un estudio previo para evaluar la conveniencia de la microfilmación o la digitalización. Se comenzará con la digitalización de algún título de revista, lo que nos servirá como experiencia para proyectos futuros.

