

PROCESO DE DIGITALIZACIÓN

EN LA BIBLIOTECA NACIONAL DE ESPAÑA

BIBLIOTECA DIGITAL HISPÁNICA

Actualizado 25/02/2015



Contenido

PROCESO DE DIGITALIZACIÓN	1
EN LA BIBLIOTECA NACIONAL DE ESPAÑA.....	1
BIBLIOTECA DIGITAL HISPÁNICA.....	1
1. LA BIBLIOTECA DIGITAL HISPÁNICA (BDH).....	3
2. PASOS PREVIOS A LA DIGITALIZACIÓN	4
2.1 Criterios de selección	4
2.2 Extracción del catálogo automatizado	6
2.3 Préstamo de títulos.....	6
2.4 Revisión y selección de ejemplares	7
2.5 Catalogación y Clasificación	9
2.6 Planificación y lotes de trabajo	9
2.7 Préstamo de obras al digitalizador	10
3. DIGITALIZACIÓN	10
3.1 Criterios de creación de marcadores	13
4. FICHEROS MÁSTER Y DERIVADOS	14
4.1 Archivos de preservación.....	14
4.1.1. Aspectos técnicos	16
4.2 Archivos de difusión.....	16
4.2.1. Formato de los archivos de difusión	16
4.2.1.1. PDF de difusión.....	17
4.2.1.2. Generación de los PDF	17
4.2.1.3. JPEG de difusión	18
4.2.1.4. Marca de agua de los archivos JPEG de difusión	18
5. CONTROL DE CALIDAD	18
5.1 Control de calidad previo a la carga en Pandora	18
5.2 Control de calidad posterior a la carga en Pandora.....	19
6. METADATOS	20
6.1 Metadatos descriptivos.....	20
6.2 Metadatos de preservación (METS-PREMIS)	21
6.3 Almacenamiento de metadatos en el SGOD.	36
6.4. Exportación de metadatos. OAI-PMH.....	36
7. ENTORNO TECNOLÓGICO	37
7.1 Escáneres.....	38
7.1.1. Tipos de escáneres	39
8. VOLCADO EN SISTEMA DE PRESERVACIÓN DIGITAL	40
9. MOTOR DE BÚSQUEDA	40
glosario de términos y abreviaturas	42

1. LA BIBLIOTECA DIGITAL HISPÁNICA (BDH)

La [Biblioteca Digital Hispánica \(BDH\)](#) es un recurso en línea que permite la consulta gratuita de decenas de miles de documentos conservados en los fondos de la Biblioteca Nacional de España (BNE). Este portal fue creado en 2008 con el objetivo de contribuir a cumplir la misión de conservación, gestión y difusión del Patrimonio Bibliográfico español en cualquier soporte encomendada a la BNE. Igualmente es la forma de dar visibilidad a un gran proyecto de digitalización sistemática que entonces se iniciaba.

Originalmente se compuso de las obras digitalizadas que durante años anteriores había ido acumulando la BNE como resultado de proyectos de digitalización puntuales o de reproducciones llevadas a cabo por el Departamento de Preservación y Conservación en el desarrollo de su trabajo.

La aparición del portal en 2008 hizo patente la voluntad de la BNE no sólo de contar con un sitio único de consulta de los objetos digitales sino también de desarrollar desde ese momento una política de digitalización sistemática de sus fondos. Este proyecto recibió además un gran impulso gracias al patrocinio que durante cinco años, el período 2008-2012, se comprometió a ofrecer la empresa Telefónica.

La digitalización sistemática de colecciones se debe considerar a estas alturas, más que un proyecto, un proceso que afecta a casi todos los Departamentos y Áreas de la BNE. A día de hoy las tareas implicadas desde que se decide qué se va a digitalizar hasta que finalmente está accesible en el portal se han incorporado de modo casi pleno en todas las rutinas de los diferentes Servicios de la Biblioteca.

Este proceso, por tanto, hay que valorarlo no sólo por lo que mejora el servicio que la BNE ofrece a los usuarios, sino también por lo que ha supuesto de adaptación al cambio para la institución y para las personas que la componen.

Aunque el proyecto nace sobre la base de algunas colecciones ya existentes, lo cierto es que el camino recorrido desde su aparición ha sido casi por completo una novedad. Ha sido necesario fijar los procedimientos de trabajo, establecer los criterios de selección, los controles de calidad, diseñar e implantar un nuevo interfaz de búsqueda, etc. Además de las dificultades propias de esta tarea, hay que tener en cuenta la evolución tecnológica que no se detiene y que exige permanente revisión y actualización de criterios y procesos.

La BNE cuenta desde hace tiempo con otro gran proyecto de digitalización sistemática de prensa antigua coordinado desde el Departamento de Prensa y Revistas. Este proyecto, la Hemeroteca Digital, ha logrado la digitalización de millones de páginas de gran valor histórico y muy apreciadas por los usuarios.

La BDH ofrece desde enero de 2011 la posibilidad de consultar desde su interfaz las publicaciones digitalizadas en el marco de este proyecto, concluyendo así el propósito inicial de presentar en un único punto de consulta todos los documentos digitalizados de la BNE.

Como resultado de todo esto, la BDH se ha convertido a su vez en una pequeña (o no tanto) biblioteca que requiere organizar los fondos, describirlos, mantener un catálogo, facilitar el acceso y atender a los usuarios; tareas en las que de un modo y otro están implicados, como ya se ha señalado, todos los Servicios de la BNE.

Cabe señalar también que la BDH se ha desarrollado según un modelo económico de colaboración entre el sector público y el privado que ha sido igualmente novedoso por su importancia. Este modelo se puede considerar en todos los sentidos plenamente beneficioso para la institución que ha recibido una significativa financiación por parte de Telefónica, manteniendo, sin embargo, toda la autonomía necesaria para organizar el proyecto y para establecer los criterios de calidad que son exigibles a esta institución. Del mismo modo, la titularidad de las imágenes sigue siendo exclusivamente de la BNE.

Este conjunto de documentos representan la contribución de la BNE a Europea, a Biblioteca Digital Europea, que pretende ofrecer en un único interfaz de consulta, acceso a los recursos digitales de archivos, bibliotecas, museos y archivos audiovisuales de toda Europa. También está representado en Hispana, el recolector de objetos digitales del Ministerio de Cultura.

En este documento se detalla el flujo de trabajo necesario para llevar a cabo el proceso que permite poner una obra a disposición del usuario en la BDH, así como los criterios de selección, técnicos y de calidad que guían todo el proceso.

2. PASOS PREVIOS A LA DIGITALIZACIÓN

La creación de colecciones digitales se enmarca dentro de un proceso transversal, que a grandes rasgos, puede resumirse en las siguientes etapas:



2.1 Criterios de selección

Como ya se ha dicho más arriba, el proyecto de digitalización que comenzó en 2008 ha sido una novedad en la BNE. En el proceso ha sido necesario ir creando procedimientos y estableciendo criterios que, que en ocasiones, han evolucionado en función de las necesidades o de las posibilidades que iban apareciendo.

La selección de colecciones se planteó desde el comienzo teniendo presente un criterio temático que permitiría organizar los documentos de acuerdo a características comunes. Los diferentes tipos de documentos que se han ido incorporando han hecho que el propio concepto de colección haya evolucionado y en cierto modo se haya ampliado en función del aumento de documentos digitalizados y de la incorporación de nuevos materiales.

El proyecto de digitalización de la BNE se centra únicamente en los fondos de dominio público, es decir, aquellos que están libres de derechos de autor. Esto supone referirse a las obras conservadas en la BNE cuyo autor haya fallecido hace 70 u 80 años (dependiendo de su fecha de nacimiento) y, por supuesto, siempre que no se trate de ediciones nuevas también protegidas en el Texto Refundido de la Ley de Propiedad Intelectual (Ley 23/2006, de 7 de julio) actualmente vigente en España.

Hay que señalar sin embargo un proyecto piloto que ha llevado a cabo la BNe para ofrecer contenido sujeto a derechos de autor. Este proyecto, Enclave, en colaboración con los editores ha permitido incorporar 2812 obras que pueden ser encontradas en el portal de la BDH. En el propio portal el usuario puede consultar alrededor del 20% de la obra y, en caso de interesarle, puede navegar hacia la página del editor donde llevar a cabo la compra.

Al margen de su pertenencia al dominio público hay otra serie de principios generales que guían la selección de documentos:

- La relevancia del contenido. Muchas de las colecciones se han seleccionado mediante un trabajo exhaustivo por parte del Servicio de Información Bibliográfica, con el fin de crear relaciones de documentos especialmente pertinentes para un determinado tema: ocio, viajes, ciencia, independencia hispanoamericana. En otras ocasiones esta unidad proviene de tratar de ofrecer la producción completa de un autor.
- El interés del material. Desde las Salas de lectura de la BNE se han seleccionado colecciones de documentos cuyas características materiales son interesantes en sí mismas: manuscritos, incunables, dibujos de arquitectura, grabado alemán...
- Otro criterio general que se combina a la hora de seleccionar los documentos es el interés que puede tener para los usuarios. La aplicación de este criterio es posible gracias al conocimiento de la colección y de los usuarios que tienen los bibliotecarios de la BNE.
- El valor patrimonial es otro criterio que inspira la selección de documentos y que persigue que las obras aceptadas tradicionalmente como obras maestras figuren en la BDH.
- Aspectos relativos a la preservación del documento físico que se digitaliza. Dado que un documento digitalizado sufrirá un menor número de consultas, la digitalización se convierte así en una técnica privilegiada de preservación.
- El seguimiento de unos criterios u otros puede suponer en ocasiones tomar decisiones diferentes. Éste es el caso, por ejemplo, la elección de las ediciones que deben ser digitalizadas. En colecciones cuyo fin principal es tratar una materia se optó por seleccionar una única edición. Sin embargo, cuando se persigue recoger la producción íntegra de un autor, se digitalizan las sucesivas ediciones de una obra al entender que en este caso las variaciones pueden ser de interés para el especialista.

- En el caso del proyecto de digitalización de la prensa histórica, los criterios generales son los mismos aunque conviene añadir algunas precisiones. El objetivo es cubrir la evolución histórica de la prensa española, desde sus inicios hasta principios del siglo XX, respetando siempre las limitaciones que marca nuestra legislación en temas de propiedad intelectual. El criterio que ha guiado la composición de esta colección ha sido seleccionar periódicos y revistas representativos de su época, que reflejaran la riqueza temática de la edición hemerográfica hispana y de los que se conservaran colecciones completas. Quienes visiten la hemeroteca encontrarán prensa política, satírica, humorística, científica, religiosa, ilustrada, amena, deportiva, artística, literaria, etc.

2.2 Extracción del catálogo automatizado

Los títulos seleccionados se marcan en Unicorn (SIGB) en un campo local del formato MARC (899), con un código asignado a tal fin, para proceder después a la extracción de los mismos.

Cada vez que un registro se considere idóneo para la colección, deberá comprobarse antes si la obra ha sido ya digitalizada por la BNE.

Finalizada la fase de selección de títulos, el Área de Automatización y Organización de Procesos realiza la extracción de los registros marcados de acuerdo a los criterios transmitidos por el Área de Biblioteca Digital.

Tras la extracción, dichos registros se cargan en una base de datos de trabajo interno, comenzando de esta forma el flujo de tareas que implican la digitalización de una obra.

2.3 Préstamo de títulos

El personal de salas se encarga de realizar el préstamo de los títulos seleccionados, así como de supervisar la colocación de los fondos, la devolución y el traslado de los fondos a los diferentes implicados en el flujo de trabajo (servicios responsables del proceso técnico y/o digitalizadores).

El préstamo de fondos se lleva a cabo en dos procesos dentro del proyecto de digitalización:

- *Selección de fondos a digitalizar:* En este momento tiene lugar la correspondiente entrega de fondos al Departamento de Proceso Técnico (fondo Moderno) o al Departamento especializado en determinado tipo de material (Patrimonio Bibliográfico, Bellas Artes y Cartografía, Música y Audiovisuales), con el fin de que se revisen los documentos originales y se seleccione el ejemplar a digitalizar. Cada sala establece una media de fondos a sacar diariamente, en función del volumen de trabajo asumible acorde con los recursos humanos y materiales de que disponga. En el caso de fondo moderno, dicha media es de 40 títulos diarios. Condición sine qua non para este movimiento masivo de fondos, es la colocación del código de barras con el IDITEM del ejemplar.

- o *Recopilación de fondos que componen el/los lote(s) de digitalización diaria.*

2.4 Revisión y selección de ejemplares

En esta fase, el Departamento de Proceso Técnico (en el caso del fondo moderno) o el Departamento especializado (para los fondos de carácter singular, incluido el antiguo) proceden a la unificación del registro bibliográfico. Esta tarea consiste en la revisión de los títulos de interés, identificación de bibliográficos repetidos y gestión de altas y bajas de registros y/o firmas que esto origine. De todo esto se deduce fácilmente que la tarea de digitalización está suponiendo una labor de depuración del catálogo bibliográfico.

A continuación se revisa y selecciona el ejemplar más adecuado para ser digitalizado. En el caso del fondo moderno esta revisión manual la realiza directamente un técnico del Departamento de Preservación y Conservación. En el resto de casos, esta participación se hace patente siempre que el personal de los departamentos así lo requiera.

2.4.1. Pautas para determinar si un ejemplar es apto para digitalizar:

Si el libro ha sido microfilmado recientemente y el estado de conservación y características del original –grado de apertura- ha permitido la obtención de microformas adecuada siempre se deberá dar prioridad a la digitalización de la copia sobre el original, excepto en los casos en los que el documento original presente imágenes en color, fotografías o grabados.

En principio la selección de las obras a digitalizar ha tenido en cuenta que el material esté en las condiciones adecuadas para su reproducción. Si bien puede haber casos en los que haya pasado desapercibido un mal estado A continuación se reflejan los criterios de exclusión de la digitalización masiva y/o aspectos que el operario de digitalización deberá tener en cuenta para la manipulación.

- o No se deberán escanear masivamente o se deberá prestar atención durante los procesos de selección a los siguientes documentos:
 1. **Libros quebradizos.** Si las hojas presentan pérdidas, se parten fácilmente o se desprenden del libro no podrá ser escaneado automáticamente. Se deberá localizar, en caso de que exista, un ejemplar en mejores condiciones u optar, en su defecto, porque la digitalización se realice con el equipamiento técnico y personal del Departamento de Preservación y Conservación.
 2. **Libros quebradizos en la zona del lomo:** Encuadernaciones a gutapercha o a la americana, rústicas y encuadernaciones de lomo rígido (especialmente entre los siglos XV-XVIII). El exceso de refuerzo en el lomo y de encolado puede provocar la rotura de la encuadernación. Se rechazarán todos aquellos ejemplares que no puedan ser abiertos con comodidad en un ángulo de 135°.
 3. **Algunas encuadernaciones presentan encolados ácidos en el lomo.** Se examinará la costura del ejemplar y se rechazarán todos aquellos en los

que aparezcan hojas sueltas por rotura en el pliegue. Este problema suele ser más frecuente en las primeras y últimas hojas.

4. **Encuadernaciones rotas** en la zona del cajo –bisagra– de la tapa delantera o trasera.
5. **Ejemplares con ataques graves de microorganismos o de insectos**, en los que aparezcan fragmentos de papel, hojas sueltas o débiles.
6. Ejemplares con graves problemas físicos, como **hojas desgarradas, sueltas o con pérdidas**.
7. Ejemplares con mapas o grabados **desplegables**.
8. Ejemplares gravemente **deformados** por acción del agua o de la mala colocación. En impresos del XVII y XVIII pueden presentar deformaciones importantes en la caja de texto.
9. **Papel rígido**. A pesar de que no presenten problemas de acidez, no se podrán escanear los libros en los que el papel sea grueso y no se flexione con facilidad.
10. Cuando existan varios ejemplares de una misma obra y uno de ellos sea una **edición facticia**, se elegirá otra opción.
11. Cuando se trate de **folletos** –signatura VC– y existan varios ejemplares se elegirá siempre el que esté encuadernado.
12. Cuando existan varios ejemplares de una misma obra y uno de ellos presente los **cuadernillos cerrados in-tonso** se escogerá otro ejemplar. Si no es posible, se anotará la incidencia y se enviará al Departamento de Preservación y Conservación.
13. Cuando se trate de un **ejemplar único** debería digitalizarse con especial cuidado. Requerida marcado o identificación especial.
14. En el caso de que no existan otras opciones y deban elegirse ejemplares con contenedores de conservación hay que prestar especial atención, ya que pueden estar protegiendo ejemplares con encuadernaciones valiosas, ejemplares muy deteriorados o ejemplares únicos.

En estos casos podemos suponer que:

En caso en que el ejemplar tiene *punto rojo* (retirado de consulta):

- Si se trata de un ejemplar único y en buen estado se le ha puesto punto rojo para asegurar su conservación. En ese caso se aconseja digitalizar desde el microfilme existente, excepto en los casos en los que el documento original presente imágenes en color, fotografías o grabados.
- Si el ejemplar no es único se ha retirado de consulta por estar ácido o deteriorado. En ese caso se aconseja digitalizar otro

ejemplar en mejor estado o digitalizar desde el microfilme existente

En caso en que el ejemplar tenga *punto verde* –

- Si se trata de un ejemplar único debería digitalizarse con especial cuidado por parte de personal especializado.
- Si no es único sería mejor elegir otro ejemplar para digitalizar, porque en principio el punto verde indicaría el mejor ejemplar, que es el que tiene reproducción en microfilme.

2.5 Catalogación y Clasificación

Una vez se ha seleccionado el ejemplar más adecuado para digitalizar, el personal encargado de la catalogación y clasificación de los registros:

- Corrige y amplía la descripción bibliográfica

- Y asigna una colección temática a la obra según el esquema temático abreviado que se maneja en BDH y que deriva de la CDU (Clasificación Decimal Universal).

2.6 Planificación y lotes de trabajo

El material que cumple los criterios de selección, se somete a una revisión de su información bibliográfica, así como a una selección del ejemplar a digitalizar por parte del personal de la BNE (proceso técnico y/o salas especializadas). A continuación el personal de BDH realiza las correspondientes extracciones del catálogo y organiza los lotes de digitalización, en función del ritmo diario que puede asumirse (teniendo en cuenta turnos de digitalización, nº de equipos y tipo de documento).

En concreto la preparación de los lotes de digitalización implica:

- o Generación de los ficheros de metadatos descriptivos asociados en formato MARC 21. Antes de su envío a la empresa responsable de la carga en el SGOD de BDH, se procede a una revisión de los marcados en el 899, y la presencia/ausencia de 856. De esta forma se minimiza el riesgo de introducir en la cadena de digitalización obras que ya cuentan con una reproducción de calidad. Estos ficheros se obtienen a través de UNICORN, y se asocian a cada uno de los lotes de digitalización. Posteriormente, se transforman a formato MARC21XML, con las especificaciones requeridas para su carga en el sistema de gestión de objetos digitales (SGOD) de la BNE.
- o Elaboración de lotes diarios de trabajo para su digitalización, manteniendo en lotes independientes los títulos simples de los títulos multivolumen y, a su vez, organizados por colección. Para la preparación de los lotes se manejan dos parámetros básicos:
 - ✓ El límite máximo de páginas a digitalizar/día

- ✓ El número máximo de ejemplares diarios que pueden asumir las salas prestatarias.

Una vez elaborados los lotes, a todos los implicados en el proceso se les distribuye un listado con los datos mínimos que permiten identificar los títulos y ejemplares a digitalizar cada día. Dichos implicados son: salas prestatarias, departamento de preservación y conservación y digitalizadores. Esta información se envía mínimo con 1 semana de antelación. Los lotes se podrán ajustar según el ritmo real de digitalización y/o las incidencias que puedan producirse en el día a día.

2.7 Préstamo de obras al digitalizador

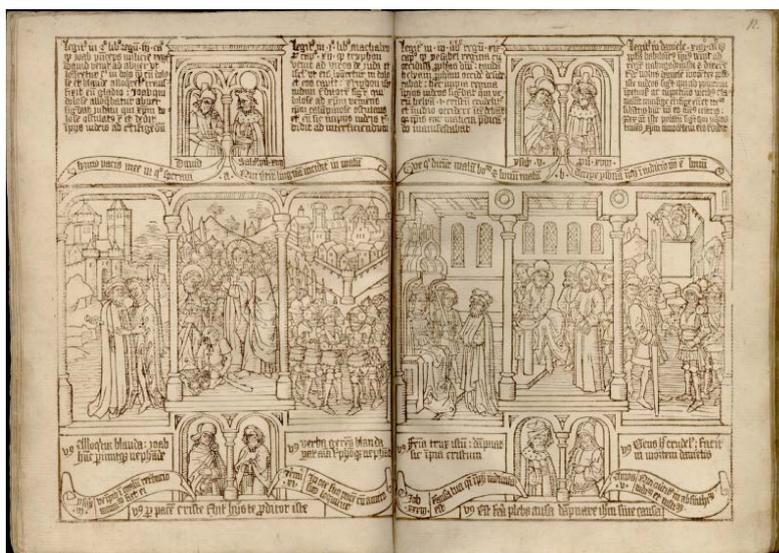
A partir del listado de lotes elaborado se realizan las tareas necesarias que permiten el préstamo al digitalizador a nivel lógico y a nivel físico comenzando así el proceso de escaneado propiamente dicho.

3. DIGITALIZACIÓN

Se detalla a continuación el proceso general por el que pasa una obra desde que se digitaliza hasta que se carga en el Sistema de Gestión de Objetos Digitales de la BNE (Pandora):

1. En primer lugar se procede a la digitalización de la obra, obteniendo imágenes en formato TIFF sin compresión a doble página en RGB 24 bits. La resolución óptica de impresión de la digitalización de las obras maestras se ha fijado en 300 dpi., en gris o a color, dependiendo del tipo de obra. Para ejemplares que pertenezcan a la colección de Bellas Artes y Cartografía se ha fijado esta resolución óptica en 400 dpi.

La manipulación de la obra durante el proceso de digitalización respetará las especificaciones de conservación establecidas por la BNE.



2. En la digitalización de cada ejemplar se utilizara una carta de color en la primera página sobre las 2 hojas del ejemplar.

3. En el proceso de digitalización de las obras se usará una cartulina de color negro debajo de los ejemplares para visualizar de forma correcta los bordes de las obras. En el caso de haber faltas de materia (por ejemplo, orificios por xilófagos) se interpondrán hojas de papel japonés para evitar ver las siguientes páginas.

4. Se lleva a cabo un control de calidad de las imágenes TIFF MÁSTER una a una, para detectar cualquier salto de página o desenfoque, que de producirse, se corrige de inmediato, digitalizando la imagen nuevamente. De manera simultánea las imágenes TIFF MÁSTER se almacenan en un servidor.

5. A continuación se realiza el enderezamiento de las imágenes con el programa LIMB Processing.

6. Para el fondo digital Manuscritos, Monografías e Impresos e Incunables, las imágenes se digitalizan a doble página.

7. Para el fondo digital Bellas Artes y Cartografía, las imágenes se digitalizan en una única página. Tras el proceso de escaneado, las imágenes TIFF MÁSTER son recortadas en dos partes, es decir, un fichero por cada página. El recorte se lleva a cabo con el programa LIMB Processing, obteniendo de esta forma un derivado de las imágenes originales. El fondo digital de Bellas Artes se digitalizará a página simple, recto y verso en el caso de que el verso pueda contener algún tipo de información relevante sobre el ejemplar (sello, nº de registro, etc...).



8. Cada imagen resultante de la digitalización (tanto TIFF MÁSTER como TIFF recortado), queda identificada con la signatura del ejemplar de la Biblioteca Nacional.

9. Se revisan de nuevo las imágenes, ya que tras el primer recorte automatizado puede quedar algún marco negro a su alrededor. Estos bordes se recortan manualmente.

10. En los casos en los que se vaya a obtener OCR y, sólo si se considera que se precisa un tratamiento previo para mejorar su calidad, se tratará la imagen con el fin de mejorar el texto sin perder información, desechando manchas y suciedad que han sido originadas por el paso del tiempo, el uso y/o la calidad del papel.



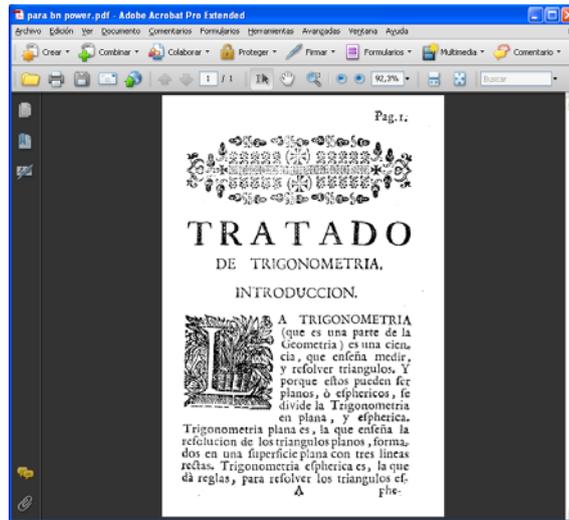
11. El siguiente proceso analiza la inclinación del texto para corregirlo en caso necesario. Se centra el texto manualmente y se unifica la medida del tamaño de todas y cada una de las imágenes, a excepción de mapas y láminas de color.

Esto se consigue utilizando una media sobre la medida del tamaño de las páginas, logrando así que sea lo más fidedigno al libro.



12. Sólo si se considera necesario para mejorar los resultados del OCR las imágenes se procesan automáticamente para convertirlas a blanco y negro. En el proceso se eliminan la suciedad o manchas que hayan podido quedar después del tratamiento de las imágenes TIFF. De manera general, casi la totalidad de fondos a digitalizar serán a color.

13. Se generan los archivos de difusión, en formato PDF o JPEG. En el caso de que se digitalicen monografías (fondo moderno) se realizará el proceso de OCR (Reconocimiento Óptico de Caracteres). También se crean los marcadores en los PDFs, y se inserta la marca de agua de la BNE a través de un archivo de imagen GIF.



14. Se procede después a la verificación de nombrado de imágenes, ejecutando un proceso que comprueba que todos los ficheros de cada ejemplar tienen exactamente el mismo nombre y que su número de secuencia comienza por 0000 y continúa sin saltos.

15. De cada archivo TIFF MÁSTER se genera la estructura de metadatos de preservación PREMIS.

16. Se relaciona posteriormente cada archivo de difusión PDF/JPEG con su registro/s MARC, generando la estructura METS/MARC/COMPLEX/SIMPLE correspondiente.

17. Antes de la carga en el SGOD, se realiza el control de calidad del 20% de los títulos de un lote y dentro de esa muestra se revisa un 20% de páginas incluyendo marcadores.

La tasa de fiabilidad de la imagen deberá ser del 99,25%. Y la tasa de fiabilidad de marcadores deberá ser del 100%. Si no se cumplen las tasas de fiabilidad, la imagen se digitaliza nuevamente.

18. Una vez superados los controles de calidad, se procede a ejecutar la carga de las obras digitalizadas en Pandora (SGOD de la BNE).

3.1 Criterios de creación de marcadores

Los criterios de creación de marcadores, dependiendo de si el libro tiene índice o no y del tamaño en número de páginas, son los siguientes:

- Si el libro NO tiene índice, se generan los marcadores correspondientes a las siguientes partes físicas/lógicas del libro:
 - ✓ Encuadernación
 - ✓ Portada (página del título y autor)
 - ✓ Índice general
 - ✓ Sección de ilustraciones (cuando aparecen todas juntas)

- ✓ Bibliografía
- ✓ Introducción prólogo
- ✓ Apéndice
- ✓ Fe de erratas
- ✓ División intelectual del contenido

Los marcadores se realizarán respetando el orden lógico del libro y se generarán en el caso de que el libro contase con esas partes.

- Si el libro SÍ tiene índice, los marcadores se establecerán en función del número de páginas:
 - ✓ Si el número de páginas es inferior a 500, se realizarán 25 marcadores.
 - ✓ Si el número de páginas es superior a 500, el número total de marcadores deberá ser el equivalente al 5% del número de páginas de libro.



Ejemplo de marcador indicando los preliminares de una obra

4. FICHEROS MÁSTER Y DERIVADOS

Tras la digitalización de una obra se generan dos tipos de archivos:

4.1 Archivos de preservación

Se considera archivo máster de preservación (TIFF máster) a aquél que se ha realizado con la máxima calidad posible para los fines establecidos en cada caso.

TIPO DE DOCUMENTO	OBJETIVO	RESOLUCION	PROFUNDIDAD DE COLOR	NOTAS
Texto impreso SIN ilustraciones, prensa, panfletos, páginas mecanografiadas	Imagen del Texto	300 ppp mínimo	Escala de grises 8 bits *	*Color (24 bits) cuando el color sea una característica importante del documento
	Texto con OCR	400 ppi	Escala de grises 8 bits *	
Música: partituras, escalas anotadas, manuscritos de música	Acceso al contenido	300 ppp mínimo	Escala de grises 8 bits*	*Color (24 bits) cuando el color sea una característica importante del documento
	Reconocimiento de sus características materiales	400 ppi	Escala de grises 8 bits*	
Manuscritos: escritos a mano, copias mecanografiadas	Acceso al contenido	300 ppp mínimo	Escala de grises 8 bits*	*Color (24 bits) cuando el color sea una característica importante del documento
	Reconocimiento de sus características materiales	400 ppp	Escala de grises 8 bits*	
Mapas: caracteres impresos color impreso hasta un tamaño 56 cm x 87 cm	Búsqueda	250 ppp mínimo *	24-bit color	*La resolución (ppp) depende del tamaño del mapa, sobre todo en los casos en los que las secciones del mapa tienen que unirse y el tamaño del archivo sobrepase los 500 MB
	Reproducción	400 ppp	24-bit color mínimo	
Fotografías: tono continuo, color	Acceso al contenido	300 ppp mínimo	Escala de grises 8 bits*	*Color (24 bits) cuando el color sea una característica importante del documento
	Reproducción	Máximo soportado	24-bit color mínimo	
Material gráfico	Acceso al contenido	300 ppi mínimo	Escala de grises 8 bits*	*Color (24 bits) cuando el color sea una característica importante del documento
	Reproducción	Máximo soportado	24-bit color	

Libros Especiales o Raros: Objetos de gran valor	Reconocimiento de sus características materiales	300 ppp mínimo	24-bit color	
	Investigación sobre sus características materiales	600 ppp mínimo	24-bit color mínimo	

Sin alteraciones posteriores a su realización representa la copia de seguridad y se utiliza para producir los ficheros derivados o de difusión.

Las recomendaciones descritas a continuación, se han definido teniendo en cuenta las directrices de distintas instituciones bibliotecarias que se enfrentan a una casuística analoga a la de la BNE. Tales recomendaciones deben considerarse como valores de referencia mínimos, y pueden variar en función de los avances tecnológicos y/o necesidades de la institución y/o sus usuarios.

4.1.1. Aspectos técnicos

El cuadro siguiente presenta las características técnicas adecuadas a cada tipo de documento relacionado en la primera columna.

La columna *Objetivo* refleja la intención de la reproducción. En cuanto a la *Resolución*, se trata de valores ópticos, no interpolados, en escala 1:1. Como norma general las imágenes deberán incluir escala de color si la digitalización es en color, y escala de grises si la digitalización se realiza en gris. El formato de imagen adoptado, en la actualidad, como óptimo para la preservación es el TIFF. Independientemente de esto, hay que considerar que la evolución técnica puede proporcionar, en un futuro, otros formatos con garantía de calidad de archivo a tener en cuenta.

4.2 Archivos de difusión

Los archivos de difusión, son los ficheros resultado de las imágenes TIFF máster, y que se someten a un procesado acorde con las especificaciones técnicas de digitalización de la Biblioteca Digital Hispánica.

Los archivos de difusión se cargan en la Biblioteca Digital Hispánica como dos tipos de objetos:

- Simples: un registro bibliográfico con un único archivo digital (PDF o JPEG).
- METS: un registro bibliográfico con varios archivos digitales (PDFs o JPEG).

4.2.1. Formato de los archivos de difusión

Los archivos de difusión van a tener dos tipos de formatos según el tipo de material: PDF y JPEG.

TIPO DE DOCUMENTO	ARCHIVO DE DIFUSIÓN
Texto impreso procedente de microforma	PDF con marcadores y OCR
Texto impreso procedente del original (incluidas partituras impresas)	PDF con marcadores y OCR
Incunables digitalizados directamente del soporte original	PDF con marcadores sin OCR
Incunables digitalizados de microformas	PDF con marcadores sin OCR
Material gráfico digitalizado directamente del soporte original (grabados, estampas, dibujos, fotografías, carteles)	JPEG a 150 ppp
Material gráfico procedente de negativo	JPEG a 150 ppp
Mapas y planos	JPEG a 150-200 ppp. En caso de que la toponimia y detalles del mapa o plano no se lean correctamente, se aumentará la calidad del JPEG.
Manuscritos digitalizados directamente del soporte original (incluida música manuscrita)	JPEG a 150-200 ppp

4.2.1.1. PDF de difusión

Las láminas, ilustraciones a color y portadas vistosas (así como cualquier otro motivo destacable que perdiera sus detalles al presentarse en blanco y negro), la imagen de difusión se generará a color o escala de grises, con el fin de representar fielmente el original físico digitalizado. De forma que los PDF presentados en Web podrán ser unos completamente en blanco y negro, otros con portadas a color y el resto en blanco y negro, o en blanco y negro y páginas internas en color o escala de grises.

4.2.1.2. Generación de los PDF

Se generará un archivo PDF a 125 ppp para su difusión.

Marca de agua de los PDF: Todas las páginas de los PDF deberán llevar la marca de agua de la BNE en la parte inferior centrada de cada página (salvo en manuscritos e incunables que irá a la derecha. En todo caso se evitará el tapar información y/o imágenes relevantes en la obra.

En el caso de que el color de la encuadernación de la portada o el color de fondo de las páginas sea oscuro, se utilizará la marca de agua inversa (color blanco del texto de la BNE).

Marcadores: Los ficheros PDF tendrán marcadores con la información de los capítulos/partes/secciones.

4.2.1.3. JPEG de difusión

Se generará un archivo JPEG a 150 ppp para su difusión, que garantice una visualización de calidad desde la Web.

Los manuscritos y libros antiguos deben mantener la encuadernación y las hojas de guarda o en blanco, pues conservan en la mayoría de los casos información de interés para la identificación de posibles procedencias o acerca de la forma de construir el libro.

4.2.1.4. Marca de agua de los archivos JPEG de difusión

La marca de agua se insertará en el ángulo inferior derecho. Esta nunca deberá tapar o superponerse sobre información del original.

© *Biblioteca Nacional de España*

Es necesario controlar el peso de las imágenes, tratando de que no resulten muy pesadas. En algunos casos, reducimos la calidad a unos 150 ppp aproximadamente, siempre que al aumentar la imagen no se pixele, para reducirlas un poco de peso.

5. CONTROL DE CALIDAD

De los ficheros obtenidos en el proceso de digitalización, se realiza un control de calidad previo y posterior a la carga en Pandora (SGOD), consistente en las siguientes tareas:

5.1 Control de calidad previo a la carga en Pandora

- Comprobación y revisión de firmas que se van a cargar.
- Revisión de xml: De los xml generados se procede a su revisión total de cada uno de los lotes que se va a cargar en Pandora. En la revisión de los xml se centra la atención en las siguientes etiquetas:
 - ✓ Nombre que en el caso de los simples debe de coincidir con el del correspondiente fichero de difusión.
 - ✓ Cabecera: Comprobamos que en la secuencia de las posiciones 21-24 de la cabecera no hay ningún espacio en blanco (ej.: debe de ser 4500 y no 45 0). Y que el tipo de material también está bien codificado.

- ✓ Etiqueta 008: Comprobamos que las posiciones correspondientes a los datos relativos a fecha de edición se expresan de la forma adecuada (no deben de haber espacios en blanco y los cuatro dígitos han de estar completos) y que el formato de fecha se ajusta a los esquemas del mapeo.
 - ✓ Etiqueta 080: Revisaremos que aparece esta etiqueta con su codificación de materia (CDU) o, en su defecto, que en la etiqueta 899 aparece la codificación simplificada de materia elaborada a tal efecto y basada también en la CDU. Estos códigos y marcas simplificados los podemos consultar en el documento de trabajo elaborado a tal fin.
 - ✓ Etiqueta 300: Se revisa esta etiqueta para comprobar si la obra es o no multivolumen y por tanto que la elección del tipo de carga en Pandora (simples, mets) es la correcta.
 - ✓ Etiqueta 899: Se revisa que esté marcada la signatura digitalizada que se va a cargar. En el caso de que exista otra signatura marcada en el catálogo, añadimos aquella que se va a cargar en BDH.
 - ✓ Etiqueta 856: En los casos que exista el enlace a BDH, se estudia si se sustituye o se mantiene el que hay. Otro aspecto importante es el relacionado con los manuscritos. En estos casos es conveniente verificar que si la obra aparece en el Catálogo Inventario General de Manuscritos de la BNE el registro cuenta con la 856 correspondiente que lo enlaza.
 - ✓ Etiquetas 9XX: Se eliminan todas las etiquetas 9XX, puesto que su contenido no es relevante como información para el usuario.
 - ✓ Si se trata de un XML METS, se comprueba el nombrado de archivos del fileSec y el structMap.
- Control de calidad de todos los ficheros PDF y JPG: Comprobación de la existencia de la marca de agua (en su caso), recorte, calidad de la imagen, OCR (en su caso), marcadores, etc.

5.2 Control de calidad posterior a la carga en Pandora

La carga de objetos digitalizados se realiza en el servidor de producción de Pandora. Una vez se ha realizado dicha carga, se comprueba su correcta carga en el sistema, prestando especial atención a los siguientes puntos: Visualización correcta de los metadatos:

- Visualización correcta de los metadatos
- Visualización correcta de los archivos.
- Liberación de las obras, una vez comprobada su carga correcta.

- En caso de obras con derechos de autor, se activa el acceso restringido.
- Actualización del listado de control de cargas, con identificadores y estado de las mismas.
- Actualización del recuento global en BDH.
- Actualización del estado de los lotes en la BBDD interna de BDH.

Una vez resueltas todas las posibles incidencias detectadas, se procede a la carga de los archivos de difusión y ficheros xml correspondientes, lo que supone la puesta a disposición de los documentos digitalizados para los usuarios de la [Biblioteca Digital Hispánica](#).

6. METADATOS

Los metadatos son el conjunto de informaciones relacionadas con los objetos digitales, cuyo objetivo es facilitar la descripción, búsqueda, uso y gestión de las colecciones digitales.

Los metadatos son las herramientas de las que disponemos para especificar la información contextual asociada a cada documento: su contenido, el historial de las transformaciones sufridas por cada objeto digital, las especificaciones de los equipos físicos necesarias para la construcción de los emuladores, los formatos de cada fichero, los programas que permitirán acceder a cada registro.

Los objetos digitales de la BDH se dotan de metadatos descriptivos (consultables en el interfaz público) y de preservación (a través de un perfil METS-PREMIS sólo accesible en los servidores de la BNE).

6.1 Metadatos descriptivos

De cada una de las obras digitalizadas se generan metadatos descriptivos en formato Marc XML. Los ficheros ".mrc" de las obras a digitalizar se obtienen desde Symphony (ISO 2709). Estos ficheros se desglosan en dos partes:

- .mrc que correspondan a objetos simples (documentos compuestos por una sola imagen).
- .mrc que correspondan a objetos complejos (documentos compuestos por varias imágenes). Se generará un fichero XML siguiendo la estructura METS XML por cada objeto complejo que vaya a cargarse en la Biblioteca Digital Hispánica.

Para adecuar el formato de metadatos descriptivos a las características específicas de carga en Pandora (SGOD), en el caso de los objetos simples, se deben nombrar archivos de metadatos y archivo de difusión de idéntica forma. Esto no hace falta en el caso de los objetos tipo METS; en estos casos la correcta carga y el orden de los archivos de difusión irá marcado por las asociaciones del fileSec.

6.2 Metadatos de preservación (METS-PREMIS)

De cada una de las obras digitalizadas dentro del proyecto de digitalización SISTEMÁTICA se obtiene un fichero de metadatos METS-PREMIS que se ingesta en los sistemas de preservación digital de la BNE.

Desde 2008 dentro de las múltiples recomendaciones internacionales que hay en materia de preservación digital, la Library of Congress empezó a dar algunas pautas sobre el uso combinado de dos modelos de metadatos ampliamente extendidos en el mundo bibliotecario: METS (*Metadata Encoding Transmission Standard*) y PREMIS (*Preservation Metadata: Implementation Strategies*). Siguiendo dos documentos fundamentales en esta materia (*Guidelines for using PREMIS with METS for Exchange*, 2008¹ y *A checklist for Documenting PREMIS-METS Decisions in a METS profile*, OCLC, 2010²), la Biblioteca Nacional de España ha diseñado un primer perfil de metadatos METS-PREMIS para materiales digitalizados dentro de su proyecto de digitalización sistemática de materiales impresos en formato papel. A este primer perfil le seguirán otros que se adapten a las características y necesidades del resto que figuran dentro de la casuística posible (sonoros, audiovisuales, digitales con soporte físico, nacidos digitales...).

A grandes rasgos este primer perfil tiene como objetivo recoger en un único fichero METS-PREMIS todos los metadatos descriptivos, administrativos y de preservación referentes a todos los archivos digitales que se generan de una obra y frente a los cuales se asume la preservación digital a largo plazo. En el caso de la digitalización sistemática, se ha tomado la decisión de preservar:

- Archivos máster tiff en crudo.
- Archivos máster tiff editados (si hubiera, y/o fuera pertinente/posible su conservación). La preservación de este material puede tener justificación por ser material que se haya utilizado como base para la generación de archivos de difusión y se desee preservar la labor de edición intermedia. Si bien, dependiendo de la disponibilidad de infraestructura tecnológica puede considerarse suficiente la preservación de los archivos máster en *crudo*.
- Archivos de difusión (en formato jpg y/o pdf dependiendo de la obra).

Toda la información se estructura en torno a las secciones de un archivo METS (<dmdSec>; <amdSec>; fileSec>; <structMap><behaviorSec>). En base a las pautas antes señaladas, ejemplos de otras instituciones y el diccionario de metadatos PREMIS³ se decidió que se utilizarían unidades semánticas PREMIS referidas a dos de las 5 entidades a las que puede referirse un archivo PREMIS: objeto y evento. Dichas unidades semánticas encontrarían ubicación fundamentalmente en la sección de metadatos administrativos de un fichero METS.

De manera muy sintética la estructura de este primer perfil METS-PREMIS sería la siguiente:

- Sección de metadatos estructurales para la descripción en marc21xml mediante un elemento mdWrap
- Una única sección de metadatos administrativos en la que se consignan:

¹ Disponible en: <http://www.loc.gov/standards/premis/guidelines-premismets.pdf>

² Disponible en: http://www.loc.gov/standards/premis/premis_mets_checklist.pdf

³ Disponible en: <http://loc.gov/standards/premis/v2/premis-2-2.pdf>

- Tantos <techMD> como imágenes se quieran describir. Dentro de este elemento se incluyen:
 - metadatos de preservación mediante unidades semánticas de la entidad OBJETO PREMIS
 - Dentro de la unidad semántica <premis:objectCharacteristicsExtension>se utiliza como extensión el modelo de metadatos MIX para la especificación de metadatos técnicos de las imágenes.
- Tantos <rightsMD> como tipo de imágenes que se quieran preservar. En el caso de la BNE serían: máster, máster editados, jpg, pdf. Se ha recurrido al modelo propuesto por la Universidad de Stanford⁴
- Tantos <digiprovMD> como eventos se realicen sobre las imágenes y de los que se guarda información acudiendo a unidades semánticas de la entidad evento de PREMIS
- Finalmente se cierra el fichero con las secciones:
 - <fileSec> en donde a modo de "inventario" se agrupan todos los tipos de imágenes (másters, másters editados, jpeg, pdf) descritos, su identificador y todos los metadatos que se han ido asignando a lo largo del fichero METS-PREMIS.
 - <StructMap> donde se informa de la estructura y ordenación de todos los ficheros descritos.

Esquemáticamente la estructura del fichero sería la siguiente:

```
<?XML version="1.0" encoding="UTF-8" ?>
- <mets xmlns:xsi="http://www.w3.org/2001/XMLSchema-instance"
  xmlns="http://www.loc.gov/METS/" xsi:schemaLocation="http://www.loc.gov/METS/
  http://www.loc.gov/standards/mets/mets.xsd http://www.loc.gov/standards/mix/
  http://www.loc.gov/standards/mix/mix.xsd http://www.w3.org/1999/xlink
  http://www.loc.gov/standards/mets/xlink.xsd" xmlns:premis="info:lc/xmlns/premis-v2"
  xmlns:xlink="http://www.w3.org/1999/xlink"
  xmlns:MARC="http://www.loc.gov/MARC21/slim" PROFILE="BNE:BNELibnovaProfile2011-
  01" xmlns:mix="http://www.loc.gov/mix/" TYPE="BNE_MonoTypeI" LABEL="¡¡La mar!!"
  OBJID="URN:BNE:1103015389">
- <metsHdr CREATEDATE="2013-01-10T15:01:58" LASTMODDATE="2012-03-08T18:37:46">
  - <agent ROLE="CREATOR" TYPE="OTHER" OTHERTYPE="SOFTWARE"> <name>Libnova
  METSWorks MetsWorks v2.06</name>
  </agent>
  - <agent ROLE="CREATOR" TYPE="ORGANIZATION">
  <name>Biblioteca Nacional de España</name> </agent>
</metsHdr>
- <dmdSec ID="dmd01-marc">
  - <mdWrap MIMETYPE="text/XML" MDTYPE="MARC" LABEL="Registro bibliográfico
  de la obra MARC21">
  - <xmlData>
  - <record> </mdWrap>
</dmdSec>
- <amdSec>
  - <techMD ID="techMD_MA_MP_000263_009_0000_001">
  - <mdWrap MDTYPE="PREMIS:OBJECT">
```

⁴ Disponible en: <http://cosimo.stanford.edu/sdr/metsrights.xsd>

```

- <xmlData>
  - <premis:object>
  - <premis:objectIdentifier>
    <premis:objectIdentifierType> </premis:objectIdentifierType>
    <premis:objectIdentifierValue> </premis:objectIdentifierValue>
    </premis:objectIdentifier> <premis:objectCategory>
  </premis:objectCategory>
  - <premis:preservationLevel>
    <premis:preservationLevelValue> </premis:preservationLevelValue>
    <premis:preservationLevelRole> </premis:preservationLevelRole>
    <premis:preservationLevelRationale> </premis:preservationLevelRationale>
    <premis:preservationLevelDateAssigned></premis:preservationLevelDateAssigned>
    </premis:preservationLevel> <premis:originalName>
  </premis:originalName>
  - <premis:objectCharacteristics>
  - <premis:format>
  - <premis:formatDesignation> <premis:formatName> </premis:formatName>
    <premis:formatVersion> </premis:formatVersion>
    </premis:formatDesignation>
  </premis:format>
  </premis:objectCharacteristics>
  - <premis:creatingApplication>
    <premis:creatingApplicationName> </premis:creatingApplicationName>
    <premis:creatingApplicationVersion> </premis:creatingApplicationVersion>
    <premis:dateCreatedByApplication></premis:dateCreatedByApplication>
    </premis:creatingApplication>
  - <premis:objectCharacteristicsExtension>
    - <mix:mix>
    - <mix:BasicDigitalObjectInformation>
    - <mix:Compression>
      <mix:compressionScheme></mix:compressionScheme>
    </mix:Compression>
    </mix:BasicDigitalObjectInformation>
    - <mix:BasicImageInformation>
    - <mix:BasicImageCharacteristics>
      <mix:imageWidth></mix:imageWidth>
      <mix:imageHeight></mix:imageHeight>
    - <mix:PhotometricInterpretation>
      <mix:colorSpace> </mix:colorSpace> </mix:PhotometricInterpretation>
    </mix:BasicImageCharacteristics> </mix:BasicImageInformation>
    - <mix:ImageAssessmentMetadata>
    - <mix:ImageColorEncoding>
    - <mix:BitsPerSample>
      <mix:bitsPerSampleValue></mix:bitsPerSampleValue>
      </mix:BitsPerSample>
      <mix:samplesPerPixel></mix:samplesPerPixel>
    </mix:ImageColorEncoding>
    </mix:ImageAssessmentMetadata>
    </mix:mix>
  </premis:objectCharacteristicsExtension>
</premis:object>

</xmlData>
</mdWrap>
</techMD>
-<rightsMD ID="rightsMD-MA-0001">
  - <mdWrap MDTYPE="OTHER" OTHERMDTYPE="METSRights">
  - <xmlData>
    - <rts:RightsDeclarationMD
      xmlns:rts="http://cosimo.stanford.edu/sdr/metsrights/"
      RIGHTSCATEGORY="COPYRIGHTED">

```

```

<rts:RightsDeclaration>(declaración sobre derechos de uso y
    acceso)</rts:RightsDeclaration>
- <rts:RightsHolder>
  <rts:RightsHolderName>Biblioteca Nacional de
    España</rts:RightsHolderName> <rts:RightsHolderComments>Para
    cualquier consulta más detallada acerca de los derechos asociados a las
    imágenes puede consultar nuestra información en
    http://www.bne.es/es/Servicios/ReproduccionDocumentos/index.html</rts:
    RightsHolderComments>
- <rts:RightsHolderContact>
  <rts:RightsHolderContactAddress>Pº de Recoletos 20-22. 28071.
    Madrid.</rts:RightsHolderContactAddress>
  <rts:RightsHolderContactPhone PHONETYPE="BUSINESS">(34) 91 580 78
    00</rts:RightsHolderContactPhone>
  <rts:RightsHolderContactEmail>info.repro@bne.es</rts:RightsHolderConta
    ctEmail>
  </rts:RightsHolderContact>
</rts:RightsHolder>
- <rts:Context CONTEXTCLASS="GENERAL PUBLIC"> <rts:Permissions
  DISCOVER="false" DISPLAY="false" COPY="false"
  DUPLICATE="false" MODIFY="false" DELETE="false" PRINT="false"
  OTHER="false" />
</rts:Context>
- <rts:Context CONTEXTCLASS="REPOSITORY MGR"> <rts:Permissions
  DISCOVER="true" DISPLAY="true" COPY="true"
  DUPLICATE="true" MODIFY="false" DELETE="false" PRINT="true"
  OTHER="false" />
</rts:Context>
</rts:RightsDeclarationMD>
  </xmlData>
</mdWrap>
</rightsMD>
-<digiprovMD ID="digiprovMD_MA_MP_000263_009_0000_001_Event001">
- <mdWrap MDTYPE="PREMIS: EVENT">
- <xmlData>
  - <premis:event>
  - <premis:eventIdentifier>
    <eventIdentifierType> </eventIdentifierType> <eventIdentifierValue>
    </eventIdentifierValue>
  </premis:eventIdentifier> <premis:eventType> </premis:eventType>
  <premis:eventDateTime></premis:eventDateTime> <premis:eventDetail>
  </premis:eventDetail>
  - <premis:eventOutcomeInformation> <premis:eventOutcome>
    </premis:eventOutcome>
  - <premis:eventOutcomeDetail>
    <premis:eventOutcomeDetailNote> </premis:eventOutcomeDetailNote>
    </premis:eventOutcomeDetail>
  </premis:eventOutcomeInformation>
  - <premis:linkingAgentIdentifier>
    <premis:linkingAgentIdentifierType>
    </premis:linkingAgentIdentifierType>
    <premis:linkingAgentIdentifierValue>
    </premis:linkingAgentIdentifierValue> <premis:linkingAgentRole>
    </premis:linkingAgentRole>
  </premis:linkingAgentIdentifier>
  </premis:event>
  </xmlData>
</mdWrap>
</digiprovMD>
</amdSec>
-<fileSec>
  - <fileGrp ID="MA- " USE=" Máster">

```

```
- <file ID=" " CREATED="2011-12-21T18:12:58"
  ADMID="techMD_MA_MP_000263_009_0000_001 rightsMD-MA-0001
  digiprovdMD_MA_MP_000263_009_0000_001_Event001
  digiprovdMD_MA_MP_000263_009_0000_001_Event002
  digiprovdMD_MA_MP_000263_009_0000_001_Event003
  digiprovdMD_MA_MP_000263_009_0000_001_Event004"
  MIMETYPE="image/TIFF" SEQ="1"
  CHECKSUM="922c6836adc6aa938b9e62559cca91fa" CHECKSUMTYPE="MD5"
  SIZE="11712023">
  <FLocat LOCTYPE="URL" xlink:href="\LOTE
    1819\TIF_MASTER\MP_000263_009__1103015389\MP_000263_009_0000.tif"
  />
</file>
</fileGrp>
```

</fileSec>

- <structMap TYPE="logical" LABEL="Organización lógica de ficheros de la obra por su tipo">

```
- <div>
- <div TYPE="Imágenes Máster" > <fptr FILEID="FPMA000001" /> </div>
- <div TYPE="Derivado - Imágenes editadas"> <fptr FILEID="FPED000001" />
  </div>
- <div TYPE="Derivado - Archivos PDF"> <fptr FILEID="FPPD000001" /> </div>
</div>
```

</structMap>

</mets>

A continuación se facilita un ejemplo simplificado e hipotético de este modelo en el que se describe las imágenes digitales de una obra que consta de 1 imágenes máster, 1 imágenes máster editada, 1 PDF, las cuales se someten a 4 eventos cada una (2 validaciones, 1 obtención de checksum o huella digital y un chequeo de antivirus).

```
<?XML version="1.0" encoding="UTF-8" ?>
- <mets xmlns:xsi="http://www.w3.org/2001/XMLSchema-instance"
  xmlns="http://www.loc.gov/METS/"
  xsi:schemaLocation="http://www.loc.gov/METS/
  http://www.loc.gov/standards/mets/mets.xsd
  http://www.loc.gov/standards/mix/
  http://www.loc.gov/standards/mix/mix.xsd http://www.w3.org/1999/xlink
  http://www.loc.gov/standards/mets/xlink.xsd"
  xmlns:premis="info:lc/xmlns/premis-v2"
  xmlns:xlink="http://www.w3.org/1999/xlink"
  xmlns:MARC="http://www.loc.gov/MARC21/slim"
  PROFILE="BNE:BNELibnovaProfile2011-01" xmlns:mix="http://www.loc.gov/mix/"
  TYPE="BNE_MonoType1" LABEL="¡¡La mar!!" OBJID="URN:BNE:1103015389">
- <metsHdr CREATEDATE="2013-01-10T15:01:58" LASTMODDATE="2012-03-
  08T18:37:46">
- <agent ROLE="CREATOR" TYPE="OTHER" OTHERTYPE="SOFTWARE"> <name>Libnova
  METSWorks MetsWorks v2.06</name>
  </agent>
- <agent ROLE="CREATOR" TYPE="ORGANIZATION">
- <name>Biblioteca Nacional de España</name>
  </agent> </metsHdr>
- <dmdSec ID="dmd01-marc">
- <mdWrap MIMETYPE="text/XML" MDTYPE="MARC" LABEL="Registro bibliográfico de la
  obra MARC21">
- <xmldata>
- <record>
  <leader>00983ncm a2200265 4500</leader>
  <controlfield tag="001">bipa0000061141</controlfield>
```

```

<controlfield tag="005">20110711</controlfield>
  <controlfield tag="007">qu</controlfield>
  <controlfield tag="008">001112q1877 espsga# nn spa</controlfield>
- <datafield tag="015" ind1="" ind2="">
- <subfield code="a">CCIM</subfield>
  </datafield>
- <datafield tag="016" ind1="7" ind2="">
  <subfield code="a">bipaBNE20010290124</subfield>
  <subfield code="z">PARBI00004021</subfield>
  <subfield code="2">SpMaBN</subfield>
  </datafield>
- <datafield tag="028" ind1="2" ind2="0">
- <subfield code="a">A. 91 V.</subfield>
  <subfield code="b">Andrés Vidal y Roger</subfield>
  </datafield>
- <datafield tag="040" ind1="" ind2="">
- <subfield code="a">SpMaBN</subfield>
- <subfield code="b">spa</subfield>
- <subfield code="c">SpMaBN</subfield>
- <subfield code="e">rdc</subfield>
  </datafield>
- <datafield tag="084" ind1="" ind2="">
  <subfield code="a">3.1</subfield>
  <subfield code="2">MU</subfield>
  </datafield>
- <datafield tag="100" ind1="1" ind2="0">
  <subfield code="a">Vehils, Joaquín María</subfield>
  <subfield code="d">1857-1934</subfield>
  </datafield>
- <datafield tag="240" ind1="1" ind2="3">
- <subfield code="a">¡La mar!</subfield>
  </datafield>
- <datafield tag="245" ind1="1" ind2="3">
- <subfield code="a">¡¡La mar!!</subfield>
- <subfield code="h">[Música notada] :</subfield>
  <subfield code="b">canción para voz de tiple, op. 4</subfield>
  <subfield code="c">letra de El cantor de Mayagüez ; música de J.M.
    Vehils</subfield>
  </datafield>
- <datafield tag="260" ind1="" ind2="">
- <subfield code="a">Barcelona</subfield>
  <subfield code="b">Andrés Vidal y Roger</subfield>
  <subfield code="c">[1877?]</subfield>
  </datafield>
- <datafield tag="300" ind1="" ind2="">
  <subfield code="a">1 partitura (8 p.)</subfield>
  <subfield code="c">35 cm</subfield>
  </datafield>
- <datafield tag="500" ind1="" ind2="">
  <subfield code="a">Dedicada a la Srta. Da. Clotilde Miquel y Sagás</subfield>
  </datafield>
- <datafield tag="594" ind1="" ind2="">
- <subfield code="a">Voz y piano</subfield>
  </datafield>
- <datafield tag="596" ind1="" ind2="">
  <subfield code="a">Fecha de publicación tomada de La música en el Boletín de la
    Propiedad Intelectual, 1997</subfield>
  </datafield>
- <datafield tag="650" ind1="" ind2="7">
  <subfield code="a">Canciones (Soprano) con piano</subfield>
  </datafield>
- <datafield tag="700" ind1="0" ind2="2">

```

```

<subfield code="a">El Cantor de Mayagüez</subfield>
</datafield>
- <datafield tag="899" ind1="" ind2="">
- <subfield code="j">MP/263/9</subfield>
- <subfield code="x">dig78dig</subfield>
  </datafield>
  </record>
  </xmlData>
  </mdWrap>
  </dmdSec>
- <amdSec>
- <techMD ID="techMD_MA_MP_000263_009_0000_001">
  - <mdWrap MDTYPE="PREMIS:OBJECT">
- <xmlData>
- <premis:object>
- <premis:objectIdentifier> <premis:objectIdentifierType>bdh</premis:objectIdentifierType>
  <premis:objectIdentifierValue>MP_000263_009_0000.tif</premis:objectIdentifier
  Value>
  </premis:objectIdentifier> <premis:objectCategory>file</premis:objectCategory>

- <premis:preservationLevel>
<premis:preservationLevelValue>full</premis:preservationLevelValue>
<premis:preservationLevelRole>requirement</premis:preservationLevelRole>
<premis:preservationLevelRationale>legislation</premis:preservationLevelRationale>
<premis:preservationLevelDateAssigned>2011-12-
21T18:12:58</premis:preservationLevelDateAssigned> </premis:preservationLevel>
<premis:originalName>MP_000263_009_0000.tif</premis:originalName>
- <premis:objectCharacteristics>
- <premis:format>
- <premis:formatDesignation> <premis:formatName>image/tif</premis:formatName>
<premis:formatVersion>Información no disponible</premis:formatVersion>
</premis:formatDesignation> </premis:format> </premis:objectCharacteristics>
- <premis:creatingApplication> <premis:creatingApplicationName>Información no
disponible</premis:creatingApplicationName>
<premis:creatingApplicationVersion>Información no
disponible</premis:creatingApplicationVersion>
<premis:dateCreatedByApplication>2011-12-
21T18:12:58</premis:dateCreatedByApplication> </premis:creatingApplication>
- <premis:objectCharacteristicsExtension>
- <mix:mix>
- <mix:BasicDigitalObjectInformation>
- <mix:Compression> <mix:compressionScheme>0</mix:compressionScheme>
</mix:Compression> </mix:BasicDigitalObjectInformation>
- <mix:BasicImageInformation>
- <mix:BasicImageCharacteristics> <mix:imageWidth>3330</mix:imageWidth>
<mix:imageHeight>3517</mix:imageHeight>
- <mix:PhotometricInterpretation> <mix:colorSpace>RGB</mix:colorSpace>
</mix:PhotometricInterpretation> </mix:BasicImageCharacteristics>
</mix:BasicImageInformation>
- <mix:ImageAssessmentMetadata>
- <mix:ImageColorEncoding>
  = <mix:BitsPerSample> <mix:bitsPerSampleValue>8</mix:bitsPerSampleValue>
</mix:BitsPerSample> <mix:samplesPerPixel>1</mix:samplesPerPixel>
</mix:ImageColorEncoding> </mix:ImageAssessmentMetadata> </mix:mix>
</premis:objectCharacteristicsExtension> </premis:object>
</xmlData> </mdWrap> </techMD>
- <techMD ID="techMD_ED_MP_000263_009_0000_001">
- <mdWrap MDTYPE="PREMIS:OBJECT">
- <xmlData>
- <premis:object>
- <premis:objectIdentifier> <premis:objectIdentifierType>bdh</premis:objectIdentifierType>
  <premis:objectIdentifierValue>MP_000263_009_0000.jpg</premis:objectIdentifie

```

```

rValue>
  </premis:objectIdentifier> <premis:objectCategory>file</premis:objectCategory>
- <premis:preservationLevel>
  <premis:preservationLevelValue>full</premis:preservationLevelValue>
  <premis:preservationLevelRole>requirement</premis:preservationLevelRole>
  <premis:preservationLevelRationale>legislation</premis:preservationLevelRationale>
  <premis:preservationLevelDateAssigned>2012-08-
    17T12:08:31</premis:preservationLevelDateAssigned> </premis:preservationLevel>
  <premis:originalName>MP_000263_009_0000.jpg</premis:originalName>
- <premis:objectCharacteristics>
- <premis:format>
- <premis:formatDesignation> <premis:formatName>image/jpg</premis:formatName>
  <premis:formatVersion>Informacion no disponible</premis:formatVersion>
  </premis:formatDesignation> </premis:format> </premis:objectCharacteristics>
- <premis:creatingApplication> <premis:creatingApplicationName>Información no
  disponible</premis:creatingApplicationName>
  <premis:creatingApplicationVersion>Información no
    disponible</premis:creatingApplicationVersion>
  <premis:dateCreatedByApplication>2012-08-
    17T12:08:31</premis:dateCreatedByApplication> </premis:creatingApplication>
- <premis:objectCharacteristicsExtension>
- <mix:mix>
- <mix:BasicDigitalObjectInformation>
- <mix:Compression> <mix:compressionScheme>0</mix:compressionScheme>
  </mix:Compression> </mix:BasicDigitalObjectInformatio>
- <mix:BasicImageInformation>
- <mix:BasicImageCharacteristics> <mix:imageWidth>1584</mix:imageWidth>
  <mix:imageHeight>2056</mix:imageHeight>
- <mix:PhotometricInterpretation> <mix:colorSpace>RGB</mix:colorSpace>
  </mix:PhotometricInterpretation> </mix:BasicImageCharacteristics>
  </mix:BasicImageInformation>
- <mix:ImageAssessmentMetadata>
- <mix:ImageColorEncoding>
- <mix:BitsPerSample> <mix:bitsPerSampleValue>8</mix:bitsPerSampleValue>
  </mix:BitsPerSample> <mix:samplesPerPixel>1</mix:samplesPerPixel>
  </mix:ImageColorEncoding> </mix:ImageAssessmentMetadata> </mix:mix>
  </premis:objectCharacteristicsExtension> </premis:object>
  </xmlData> </mdWrap> </techMD>
- <rightsMD ID="rightsMD-MA-0001">
- <mdWrap MDTYPE="OTHER" OTHERMDTYPE="METSRights">
- <xmlData>
- <rts:RightsDeclarationMD
  xmlns:rts="http://cosimo.stanford.edu/sdr/metsrights/"
  RIGHTSCATEGORY="COPYRIGHTED">
  <rts:RightsDeclaration>(declaración sobre derechos de uso y
    acceso)</rts:RightsDeclaration>
- <rts:RightsHolder>
  <rts:RightsHolderName>Biblioteca Nacional de España</rts:RightsHolderName>
  <rts:RightsHolderComments>Para cualquier consulta más detallada acerca de los
    derechos asociados a las imágenes puede consultar nuestra información en
    http://www.bne.es/es/Servicios/ReproduccionDocumentos/index.html</rts:
    RightsHolderComments>
- <rts:RightsHolderContact>
  <rts:RightsHolderContactAddress>Pº de Recoletos 20-22. 28071.
    Madrid.</rts:RightsHolderContactAddress> <rts:RightsHolderContactPhone
    PHONETYPE="BUSINESS">(34) 91 580 78
    00</rts:RightsHolderContactPhone>
  <rts:RightsHolderContactEmail>info.repro@bne.es</rts:RightsHolderContactEmail>
  </rts:RightsHolderContact> </rts:RightsHolder>
- <rts:Context CONTEXTCLASS="GENERAL PUBLIC"> <rts:Permissions DISCOVER="false"
  DISPLAY="false" COPY="false"

```

```

DUPLICATE="false" MODIFY="false" DELETE="false" PRINT="false" OTHER="false" />
</rts:Context>
  = <rts:Context CONTEXTCLASS="REPOSITORY MGR">
<rts:Permissions DISCOVER="true" DISPLAY="true" COPY="true" DUPLICATE="true"
  MODIFY="false" DELETE="false" PRINT="true" OTHER="false" /> </rts:Context>
</rts:RightsDeclarationMD> </xmlData>
</mdWrap> </rightsMD>
- <rightsMD ID="rightsMD-ED-0001">
- <mdWrap MDTYPE="OTHER" OTHERMDTYPE="METSRights">
- <xmlData>
- <rts:RightsDeclarationMD
  xmlns:rts="http://cosimo.stanford.edu/sdr/metsrights/"
  RIGHTSCATEGORY="PUBLIC DOMAIN">

  <rts:RightsDeclaration>(declaración sobre derechos de uso y
  acceso)</rts:RightsDeclaration>
- <rts:RightsHolder>
  <rts:RightsHolderName>Biblioteca Nacional de España</rts:RightsHolderName>
  <rts:RightsHolderComments>Para cualquier consulta más detallada acerca de los
  derechos asociados a las imágenes puede consultar nuestra información en
  http://www.bne.es/es/Servicios/ReproduccionDocumentos/index.html</rts:
  RightsHolderComments>
- <rts:RightsHolderContact>
  <rts:RightsHolderContactAddress>Pº de Recoletos 20-22. 28071.
  Madrid.</rts:RightsHolderContactAddress> <rts:RightsHolderContactPhone
  PHONETYPE="BUSINESS">(34) 91 580 78
  00</rts:RightsHolderContactPhone>
  <rts:RightsHolderContactEmail>info.repro@bne.es</rts:RightsHolderContactEmail>
  </rts:RightsHolderContact> </rts:RightsHolder>
- <rts:Context CONTEXTCLASS="GENERAL PUBLIC"> <rts:Permissions DISCOVER="false"
  DISPLAY="false" COPY="false"
  DUPLICATE="false" MODIFY="false" DELETE="false" PRINT="false" OTHER="false" />
  </rts:Context>
- <rts:Context CONTEXTCLASS="REPOSITORY MGR">
  <rts:Permissions DISCOVER="true" DISPLAY="true" COPY="true" DUPLICATE="true"
  MODIFY="false" DELETE="false" PRINT="true" OTHER="false" /> </rts:Context>
  </rts:RightsDeclarationMD>
  </xmlData>
  </mdWrap> </rightsMD>
- <rightsMD ID="rightsMD-PD-0001">
- <mdWrap MDTYPE="OTHER" OTHERMDTYPE="METSRights">
- <xmlData>
- <rts:RightsDeclarationMD
  xmlns:rts="http://cosimo.stanford.edu/sdr/metsrights/"
  RIGHTSCATEGORY="PUBLIC DOMAIN">
  <rts:RightsDeclaration>(declaración sobre derechos de uso y
  acceso)</rts:RightsDeclaration>
- <rts:RightsHolder>
  <rts:RightsHolderName>Biblioteca Nacional de España</rts:RightsHolderName>
  <rts:RightsHolderComments>Para cualquier consulta más detallada acerca de los
  derechos asociados a las imágenes puede consultar nuestra información en
  http://www.bne.es/es/Servicios/ReproduccionDocumentos/index.html</rts:
  RightsHolderComments>
- <rts:RightsHolderContact>
  <rts:RightsHolderContactAddress>Pº de Recoletos 20-22. 28071.
  Madrid.</rts:RightsHolderContactAddress> <rts:RightsHolderContactPhone
  PHONETYPE="BUSINESS">(34) 91 580 78
  00</rts:RightsHolderContactPhone>
  <rts:RightsHolderContactEmail>info.repro@bne.es</rts:RightsHolderContactEmail>
  </rts:RightsHolderContact> </rts:RightsHolder>
- <rts:Context CONTEXTCLASS="GENERAL PUBLIC">
  <rts:Permissions DISCOVER="true" DISPLAY="true" COPY="true" DUPLICATE="false"

```

```

MODIFY="false" DELETE="false" PRINT="true" OTHER="false" />
  </rts:Context>
- <rts:Context CONTEXTCLASS="REPOSITORY MGR">
  <rts:Permissions DISCOVER="true" DISPLAY="true" COPY="true" DUPLICATE="true"
    MODIFY="false" DELETE="false" PRINT="true" OTHER="false" /> </rts:Context>
  </rts:RightsDeclarationMD> </xmlData>
  </mdWrap> </rightsMD>
- <digiprovMD ID="digiprovMD_MA_MP_000263_009_0000_001_Event001">
- <mdWrap MDTYPE="PREMIS:EVENT">
- <xmlData>

- <premis:event>
- <premis:eventIdentifier> <eventIdentifierType>TYPE_ID</eventIdentifierType>
  <eventIdentifierValue>VALIDATION_000000000001_Event001</eventIdentifierValue>
  </premis:eventIdentifier> <premis:eventType>VALIDATION</premis:eventType>
  <premis:eventDateTime>2013-01-10T15:01:09</premis:eventDateTime>
  <premis:eventDetail>JHOVE validation</premis:eventDetail>
- <premis:eventOutcomeInformation>
  <premis:eventOutcome>successful</premis:eventOutcome>
- <premis:eventOutcomeDetail> <premis:eventOutcomeDetailNote>Well-formed and
  valid</premis:eventOutcomeDetailNote> </premis:eventOutcomeDetail>
  </premis:eventOutcomeInformation>
- <premis:linkingAgentIdentifier>
  <premis:linkingAgentIdentifierType>AgentID</premis:linkingAgentIdentifierType>
  <premis:linkingAgentIdentifierValue>jhove_1.5</premis:linkingAgentIdentifierValue>
  <premis:linkingAgentRole>software component</premis:linkingAgentRole>
  </premis:linkingAgentIdentifier> </premis:event>
  </xmlData>
  </mdWrap> </digiprovMD>
- <digiprovMD ID="digiprovMD_MA_MP_000263_009_0000_001_Event002">
- <mdWrap MDTYPE="PREMIS:EVENT">
- <xmlData>
- <premis:event>
- <premis:eventIdentifier> <eventIdentifierType>TYPE_ID</eventIdentifierType>
  <eventIdentifierValue>VALIDATION_000000000001_Event002</eventIdentifierValue>
  </premis:eventIdentifier> <premis:eventType>VALIDATION</premis:eventType>
  <premis:eventDateTime>2013-01-10T15:01:09</premis:eventDateTime>
  <premis:eventDetail>Validation LIVIT - Libnova Image Validation Indata
  Tool</premis:eventDetail>
- <premis:eventOutcomeInformation>
  <premis:eventOutcome>successful</premis:eventOutcome>
- <premis:eventOutcomeDetail>
  <premis:eventOutcomeDetailNote>CORRECT</premis:eventOutcomeDetailNote>
  </premis:eventOutcomeDetail> </premis:eventOutcomeInformation>
- <premis:linkingAgentIdentifier>
  <premis:linkingAgentIdentifierType>AgentID</premis:linkingAgentIdentifierType>
  <premis:linkingAgentIdentifierValue>LIVIT_2.6.2715</premis:linkingAgentIdentifierValue>
  <premis:linkingAgentRole>software component -
  www.libnova.es/livit</premis:linkingAgentRole> </premis:linkingAgentIdentifier>
  </premis:event> </xmlData> </mdWrap> </digiprovMD>
- <digiprovMD ID="digiprovMD_MA_MP_000263_009_0000_001_Event003">
- <mdWrap MDTYPE="PREMIS:EVENT">
- <xmlData>
- <premis:event>
- <premis:eventIdentifier> <eventIdentifierType>TYPE_ID</eventIdentifierType>
  <eventIdentifierValue>CHECKSUM_000000000001_Event001</eventIdentifierValue>
  </premis:eventIdentifier>
  <premis:eventType>Message Digest Calculation</premis:eventType>
  <premis:eventDateTime>2013-01-10T15:01:09</premis:eventDateTime>

```

```

<premis:eventDetail>Checksum generation - LibChecksum - Libnova Checksum
  generation Tool</premis:eventDetail>
- <premis:eventOutcomeInformation>
  <premis:eventOutcome>successful</premis:eventOutcome>
- <premis:eventOutcomeDetail>
  <premis:eventOutcomeDetailNote>c596c2b7afd5cb6c96fee91fba33c320</premis
:eventOutcomeDetailNote>

  </premis:eventOutcomeDetail> </premis:eventOutcomeInformation>
- <premis:linkingAgentIdentifier>
  <premis:linkingAgentIdentifierType>AgentID</premis:linkingAgentIdentifierType>
  <premis:linkingAgentIdentifierValue>LibChecksum_1.0.12</premis:linkingAgentIde
ntifierValue>
  <premis:linkingAgentRole>software component</premis:linkingAgentRole>
  </premis:linkingAgentIdentifier>
  </premis:event> </xmlData> </mdWrap> </digiprovMD>
- <digiprovMD ID="digiprovMD_MA_MP_000263_009_0000_001_Event004">
- <mdWrap MDTYPE="PREMIS:EVENT">
- <xmlData>
- <premis:event>
- <premis:eventIdentifier> <eventIdentifierType>TYPE_ID</eventIdentifierType>
  <eventIdentifierValue>ANTIVIRUSCHECK_00000000001_Event001</eventIden
tifierValue>
  </premis:eventIdentifier> <premis:eventType>Virus Check</premis:eventType>
  <premis:eventDateTime>2013-01-10T15:01:09</premis:eventDateTime>
  <premis:eventDetail>Antivirus Check - ClamAV</premis:eventDetail>
- <premis:eventOutcomeInformation>
  <premis:eventOutcome>successful</premis:eventOutcome>
- <premis:eventOutcomeDetail> <premis:eventOutcomeDetailNote>NO VIRUS
  DETECTED</premis:eventOutcomeDetailNote> </premis:eventOutcomeDetail>
  </premis:eventOutcomeInformation>
- <premis:linkingAgentIdentifier>
  <premis:linkingAgentIdentifierType>AgentID</premis:linkingAgentIdentifierType>
  <premis:linkingAgentIdentifierValue>ClamAV
  0.97.3</premis:linkingAgentIdentifierValue> <premis:linkingAgentRole>software
  component - ClamAV
  0.97.3</premis:linkingAgentRole> </premis:linkingAgentIdentifier>
  </premis:event> </xmlData> </mdWrap> </digiprovMD>
- <digiprovMD ID="digiprovMD_ED_MP_000263_009_0000_001_Event001">
- <mdWrap MDTYPE="PREMIS:EVENT">
- <xmlData>
- <premis:event>
- <premis:eventIdentifier> <eventIdentifierType>TYPE_ID</eventIdentifierType>
  <eventIdentifierValue>VALIDATION_00000000001_Event001</eventIdentifierV alue>
  </premis:eventIdentifier> <premis:eventType>VALIDATION</premis:eventType>
  <premis:eventDateTime>2013-01-10T15:01:10</premis:eventDateTime>
  <premis:eventDetail>JHOVE validation</premis:eventDetail>
- <premis:eventOutcomeInformation>
  <premis:eventOutcome>successful</premis:eventOutcome>
- <premis:eventOutcomeDetail> <premis:eventOutcomeDetailNote>Well-formed and
  valid</premis:eventOutcomeDetailNote> </premis:eventOutcomeDetail>
  </premis:eventOutcomeInformation>
- <premis:linkingAgentIdentifier>
  <premis:linkingAgentIdentifierType>AgentID</premis:linkingAgentIdentifierType>
  <premis:linkingAgentIdentifierValue>jhove_1.5</premis:linkingAgentIdentifierValue>
  <premis:linkingAgentRole>software component</premis:linkingAgentRole>
  </premis:linkingAgentIdentifier> </premis:event>
  </xmlData> </mdWrap> </digiprovMD>
- <digiprovMD ID="digiprovMD_ED_MP_000263_009_0000_001_Event002">
- <mdWrap MDTYPE="PREMIS:EVENT">
- <xmlData>
- <premis:event>

```

```

- <premis:eventIdentifier> <eventIdentifierType>TYPE_ID</eventIdentifierType>
  <eventIdentifierValue>VALIDATION_00000000001_Event002</eventIdentifierValue>
</premis:eventIdentifier> <premis:eventType>VALIDATION</premis:eventType>
<premis:eventDateTime>2013-01-10T15:01:10</premis:eventDateTime>
<premis:eventDetail>Validation LIVIT - Libnova Image Validation Indata
  Tool</premis:eventDetail>
- <premis:eventOutcomeInformation>
  <premis:eventOutcome>successful</premis:eventOutcome>
- <premis:eventOutcomeDetail>
  <premis:eventOutcomeDetailNote>CORRECT</premis:eventOutcomeDetailNote>
</premis:eventOutcomeDetail> </premis:eventOutcomeInformation>
- <premis:linkingAgentIdentifier>
  <premis:linkingAgentIdentifierType>AgentID</premis:linkingAgentIdentifierType>
  <premis:linkingAgentIdentifierValue>LIVIT_2.6.2715</premis:linkingAgentIdentifierValue>
  <premis:linkingAgentRole>software component -
  www.libnova.es/livit</premis:linkingAgentRole> </premis:linkingAgentIdentifier>
</premis:event> </xmlData>
</mdWrap> </digiprovMD>
- <digiprovMD ID="digiprovMD_ED_MP_000263_009_0000_001_Event003">
- <mdWrap MDTYPE="PREMIS:EVENT">
- <xmlData>
- <premis:event>
- <premis:eventIdentifier> <eventIdentifierType>TYPE_ID</eventIdentifierType>
  <eventIdentifierValue>CHECKSUM_00000000001_Event001</eventIdentifierValue>
</premis:eventIdentifier>
  <premis:eventType>Message Digest Calculation</premis:eventType>
  <premis:eventDateTime>2013-01-10T15:01:10</premis:eventDateTime>
  <premis:eventDetail>Checksum generation - LibCheckSUM - Libnova Checksum
  generation Tool</premis:eventDetail>
- <premis:eventOutcomeInformation>
  <premis:eventOutcome>successful</premis:eventOutcome>
- <premis:eventOutcomeDetail>
  <premis:eventOutcomeDetailNote>c596c2b7afd5cb6c96fee91fba33c320</premis:
  eventOutcomeDetailNote>
  </premis:eventOutcomeDetail> </premis:eventOutcomeInformation>
- <premis:linkingAgentIdentifier>
  <premis:linkingAgentIdentifierType>AgentID</premis:linkingAgentIdentifierType>
  <premis:linkingAgentIdentifierValue>LibCheckSUM_1.0.12</premis:linkingAgentIdentifierValue>
  <premis:linkingAgentRole>software component</premis:linkingAgentRole>
</premis:linkingAgentIdentifier>
</premis:event> </xmlData> </mdWrap> </digiprovMD>
- <digiprovMD ID="digiprovMD_ED_MP_000263_009_0000_001_Event004">
- <mdWrap MDTYPE="PREMIS:EVENT">
- <xmlData>
- <premis:event>
- <premis:eventIdentifier> <eventIdentifierType>TYPE_ID</eventIdentifierType>
  <eventIdentifierValue>ANTIVIRUSCHECK_00000000001_Event001</eventIdentifierValue>
</premis:eventIdentifier> <premis:eventType>Virus Check</premis:eventType>
  <premis:eventDateTime>2013-01-10T15:01:10</premis:eventDateTime>
  <premis:eventDetail>Antivirus Check - ClamAV</premis:eventDetail>
- <premis:eventOutcomeInformation>
  <premis:eventOutcome>successful</premis:eventOutcome>
- <premis:eventOutcomeDetail> <premis:eventOutcomeDetailNote>NO VIRUS
  DETECTED</premis:eventOutcomeDetailNote>
  </premis:eventOutcomeDetail> </premis:eventOutcomeInformation>
- <premis:linkingAgentIdentifier>
  <premis:linkingAgentIdentifierType>AgentID</premis:linkingAgentIdentifierType>
  <premis:linkingAgentIdentifierValue>ClamAV

```

0.97.3</premis:linkingAgentIdentifierValue> <premis:linkingAgentRole>**software component - ClamAV**

```

    0.97.3</premis:linkingAgentRole> </premis:linkingAgentIdentifier>
    </premis:event> </xmlData> </mdWrap> </digiprovMD>
- <digiprovMD ID="digiprovMD_PD_MP_000263_009_001_Event001">
- <mdWrap MDTYPE="PREMIS:EVENT">
- <xmlData>
- <premis:event>
- <premis:eventIdentifier> <eventIdentifierType>TYPE_ID</eventIdentifierType>
  <eventIdentifierValue>VALIDATION_00000000001_Event001</eventIdentifierValue>
  </premis:eventIdentifier> <premis:eventType>VALIDATION</premis:eventType>
  <premis:eventDateTime>2013-01-10T15:01:10</premis:eventDateTime>
  <premis:eventDetail>JHOVE validation</premis:eventDetail>
- <premis:eventOutcomeInformation>
  <premis:eventOutcome>successful</premis:eventOutcome>
- <premis:eventOutcomeDetail> <premis:eventOutcomeDetailNote>Well-formed and
  valid</premis:eventOutcomeDetailNote> </premis:eventOutcomeDetail>
  </premis:eventOutcomeInformation>
- <premis:linkingAgentIdentifier>
  <premis:linkingAgentIdentifierType>AgentID</premis:linkingAgentIdentifierType>
  <premis:linkingAgentIdentifierValue>jhove_1.5</premis:linkingAgentIdentifierValue>
  <premis:linkingAgentRole>software component</premis:linkingAgentRole>
  </premis:linkingAgentIdentifier> </premis:event>
  </xmlData> </mdWrap> </digiprovMD>
- <digiprovMD ID="digiprovMD_PD_MP_000263_009_001_Event002">
- <mdWrap MDTYPE="PREMIS:EVENT">
- <xmlData>
- <premis:event>
- <premis:eventIdentifier> <eventIdentifierType>TYPE_ID</eventIdentifierType>
  <eventIdentifierValue>VALIDATION_00000000001_Event002</eventIdentifierValue>
  </premis:eventIdentifier>
  <premis:eventType>VALIDATION</premis:eventType>
  <premis:eventDateTime>2013-01-10T15:01:10</premis:eventDateTime>
  <premis:eventDetail>Validation LIVIT - Libnova Image Validation Indata
  Tool</premis:eventDetail>
- <premis:eventOutcomeInformation>
  <premis:eventOutcome>successful</premis:eventOutcome>
- <premis:eventOutcomeDetail>
  <premis:eventOutcomeDetailNote>CORRECT</premis:eventOutcomeDetailNote>
  </premis:eventOutcomeDetail> </premis:eventOutcomeInformation>
- <premis:linkingAgentIdentifier>
  <premis:linkingAgentIdentifierType>AgentID</premis:linkingAgentIdentifierType>
  <premis:linkingAgentIdentifierValue>LIVIT_2.6.2715</premis:linkingAgentIdentifierValue>
  <premis:linkingAgentRole>software component -
  www.libnova.es/livit</premis:linkingAgentRole> </premis:linkingAgentIdentifier>
  </premis:event> </xmlData> </mdWrap> </digiprovMD>
- <digiprovMD ID="digiprovMD_PD_MP_000263_009_001_Event003">
- <mdWrap MDTYPE="PREMIS:EVENT">
- <xmlData>
- <premis:event>
- <premis:eventIdentifier> <eventIdentifierType>TYPE_ID</eventIdentifierType>
  <eventIdentifierValue>CHECKSUM_00000000001_Event001</eventIdentifierValue>
  </premis:eventIdentifier>
  <premis:eventType>Message Digest Calculation</premis:eventType>
  <premis:eventDateTime>2013-01-10T15:01:10</premis:eventDateTime>
  <premis:eventDetail>Checksum generation - LibCheckSUM - Libnova Checksum
  generation Tool</premis:eventDetail>

```

```

- <premis:eventOutcomeInformation>
  <premis:eventOutcome>successful</premis:eventOutcome>
- <premis:eventOutcomeDetail>
  <premis:eventOutcomeDetailNote>c596c2b7afd5cb6c96fee91fba33c320</premis:
  :eventOutcomeDetailNote>
  </premis:eventOutcomeDetail> </premis:eventOutcomeInformation>
- <premis:linkingAgentIdentifier>
  <premis:linkingAgentIdentifierType>AgentID</premis:linkingAgentIdentifierType>
  <premis:linkingAgentIdentifierValue>LibCheckSUM_1.0.12</premis:linkingAgentIde
  ntifierValue>
  <premis:linkingAgentRole>software component</premis:linkingAgentRole>
  </premis:linkingAgentIdentifier>
  </premis:event> </xmlData>
  </mdWrap> </digiprovMD>
- <digiprovMD ID="digiprovMD_PD_MP_000263_009_001_Event004">
- <mdWrap MDTYPE="PREMIS:EVENT">
- <xmlData>
- <premis:event>
- <premis:eventIdentifier> <eventIdentifierType>TYPE_ID</eventIdentifierType>
  <eventIdentifierValue>ANTIVIRUSCHECK_00000000001_Event001</eventIden
  tifierValue>
  </premis:eventIdentifier> <premis:eventType>Virus Check</premis:eventType>
  <premis:eventDateTime>2013-01-10T15:01:10</premis:eventDateTime>
  <premis:eventDetail>Antivirus Check - ClamAV</premis:eventDetail>
- <premis:eventOutcomeInformation>
  <premis:eventOutcome>successful</premis:eventOutcome>
- <premis:eventOutcomeDetail> <premis:eventOutcomeDetailNote>NO VIRUS
  DETECTED</premis:eventOutcomeDetailNote> </premis:eventOutcomeDetail>
  </premis:eventOutcomeInformation>
- <premis:linkingAgentIdentifier>
  <premis:linkingAgentIdentifierType>AgentID</premis:linkingAgentIdentifierType>
  <premis:linkingAgentIdentifierValue>ClamAV
  0.97.3</premis:linkingAgentIdentifierValue> <premis:linkingAgentRole>software
  component - ClamAV
  0.97.3</premis:linkingAgentRole> </premis:linkingAgentIdentifier>
  </premis:event> </xmlData> </mdWrap> </digiprovMD> </amdSec>
- <fileSec>
- <fileGrp ID="MA-GRP1" USE="Imágenes Máster">
- <file ID="FPMA000001" CREATED="2011-12-21T18:12:58"
  ADMID="techMD_MA_MP_000263_009_0000_001 rightsMD-MA-0001
  digiprovMD_MA_MP_000263_009_0000_001_Event001
  digiprovMD_MA_MP_000263_009_0000_001_Event002
  digiprovMD_MA_MP_000263_009_0000_001_Event003
  digiprovMD_MA_MP_000263_009_0000_001_Event004" MIMETYPE="image/tiff"
  SEQ="1" CHECKSUM="922c6836adc6aa938b9e62559cca91fa"
  CHECKSUMTYPE="MD5" SIZE="11712023">
  <FLocat LOCTYPE="URL" xlink:href="\LOTE
  1819\TIF_MASTER\MP_000263_009__1103015389\MP_000263_009_0000. tif"
  />
  </file> </fileGrp>
- <fileGrp ID="ED-GRP1" USE="Derivado - Imágenes editadas">
- <file ID="FPED000001" CREATED="2012-08-17T12:08:31"
  ADMID="techMD_ED_MP_000263_009_0000_001 rightsMD-ED-0001
  digiprovMD_ED_MP_000263_009_0000_001_Event001
  digiprovMD_ED_MP_000263_009_0000_001_Event002
  digiprovMD_ED_MP_000263_009_0000_001_Event003
  digiprovMD_ED_MP_000263_009_0000_001_Event004" MIMETYPE="image/jpeg"
  SEQ="1"

  CHECKSUM="c5b5535d3ab28edb9b47016731d18fa8" CHECKSUMTYPE="MD5"
  SIZE="123832">
  <FLocat LOCTYPE="URL" xlink:href="\LOTE

```

```
1819\TIF_EDITADO\MP_000263_009__1103015389\MP_000263_009_0000
.jpg" /> </file> </fileGrp>
- <fileGrp ID="PD-GRP1" USE="Derivado - Archivos PDF">
- <file ID="FPPD000001" CREATED="2012-08-27T09:08:54"
  ADMID="digiprovMD_PD_MP_000263_009_001_Event001
  digiprovMD_PD_MP_000263_009_001_Event002
  digiprovMD_PD_MP_000263_009_001_Event003
  digiprovMD_PD_MP_000263_009_001_Event004" MIMETYPE="image/tiff" SEQ="1"
  CHECKSUM="69a03d093c46f1877793a17e26a78331"
  CHECKSUMTYPE="MD5" SIZE="763412">
  <FLocat LOCTYPE="URL" xlink:href="\LOTE 1819\PDF\MP_000263_009.pdf" /> </file>
  </fileGrp> </fileSec>
- <structMap TYPE="logical" LABEL="Organización lógica de ficheros de la obra por su
  tipo">
- <div>
- <div TYPE="Imágenes Máster">
- <fptr FILEID="FPMA000001" />
  </div>
- <div TYPE="Derivado - Imágenes editadas"> <fptr FILEID="FPED000001" />
  </div>
- <div TYPE="Derivado - Archivos PDF">
- <fptr FILEID="FPPD000001" />
</div> </div> </structMap> </mets>
```

6.3 Almacenamiento de metadatos en el SGOD.

PANDORA es el Sistema de Gestión de Objetos Digitales que actualmente utiliza la BNE, tras haber migrado de Digitool, el sistema anterior con el que se ha trabajado desde el 2008.

Pandora es un repositorio de documentación para recopilar las imágenes digitales de fondos documentales permitiendo a sus usuarios la recuperación de los mismos a partir de avanzados criterios de búsqueda tales como tipo de documento, cabecera, la fecha o rango de fechas y/o términos contenidos en el texto completo de los documentos indexados.

Pandora es un instrumento que permite la recopilación, conservación y difusión de fondos documentales digitalizandos mediante una interfaz de accesos web (Internet/Intranet) de acceso simultáneo y concurrente.

Se trata de un desarrollo realizado íntegramente por Cran Consulting, S. L. que, basándose en el motor de indexación Full Text Lucene (desarrollo de dominio público bajo licencia de software libre de Apache), permite la indexación de archivos digitales, para su inclusión en una base de datos documental.

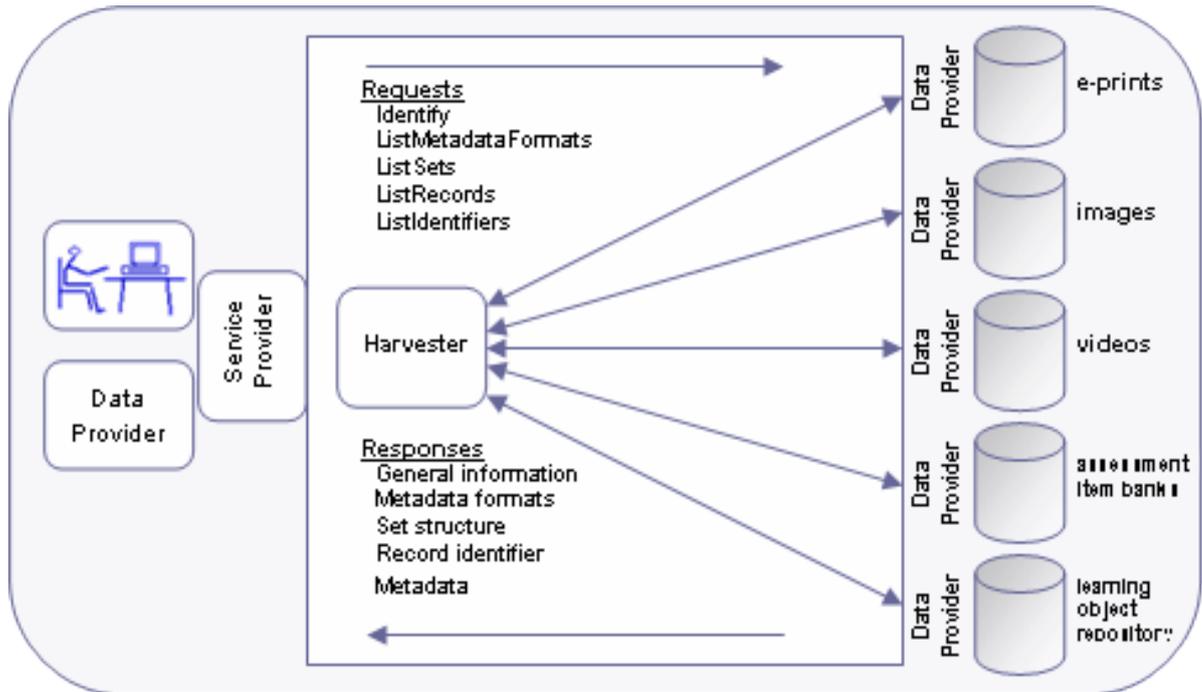
Las ventajas por las que se ha optado por migrar de Digitool a Pandora son las siguientes:

1. Se elimina la necesidad de gestionar y pagar contratos de mantenimiento y licencias de objetos digitales, lo que supone un ahorro económico y de RR.HH.
2. Se unifica así la gestión de objetos digitales dentro de la BNE, ya que Pandora es el software que gestiona la Hemeroteca Digital desde su creación en 2007.
3. Al ser un software libre, permite hacer desarrollos ulteriores para mejorar y enriquecer la experiencia del usuario. Se deja así de depender de productos comerciales cerrados.
4. Se optimizan los tiempos de respuesta/consulta de obras de grandes dimensiones (bien por número de páginas o tamaño físico de los archivos digitales) al estarse manejando archivos monopágina.
5. Se mejora la visualización y gestión de mapas y atlas, objetos que presentaban problemas de visualización y descarga en Digitool.

6.4. Exportación de metadatos. OAI-PMH

El protocolo OAI-PMH (Open Archives Initiative – Protocol for Metadata Harvesting) se utiliza para la transmisión de metadatos en Internet. Su arquitectura basada en el modelo cliente – servidor pone a disposición del público metadatos en formato Dublin Core para que puedan ser recuperados. La comunicación se realiza mediante el protocolo http y las respuestas están codificadas en XML.

En BDH contamos con un servidor OAI: <http://bdh-rd.bne.es/oai.vm?verb=ListRecords&metadataPrefix=marcxml>, a través del cual se pueden recolectar los metadatos descriptivos contenidos en BDH. Esto se puede realizar a través de comandos OAI o a través de programas del tipo MEdit, pudiéndose recuperar registros concretos, grupos de registros y los sets de OAI que hay definidos en BDH.



Modelo de exportación de metadatos OAI-PMH

7. ENTORNO TECNOLÓGICO

La BNE cuenta, de modo genérico, con las siguientes infraestructuras tecnológicas:

- o Sala de digitalización donde se encuentran instalados los escáneres necesarios.
- o Pistolas lectoras de código de barras para facilitar el nombrado de las carpetas contenedoras de las imágenes producidas durante la digitalización. Esto facilita la carga de datos en el Sistema de Preservación que actualmente está creando la Unidad de Coordinación Informática y que utiliza como identificador unívoco el ID-ITEM.
- o Acceso a Internet para la gestión y control de la herramienta de flujo de trabajo.
- o Equipos informáticos para la carga en Pandora (SGOD).
- o Servidores de almacenamiento para el volcado de archivos máster y metadatos PREMIS asociados.
- o Equipos informáticos para realizar el volcado de los archivos máster.
- o Aplicación para el control de archivos máster.

7.1 Escáneres

Las características técnicas de los escáneres utilizados son las adecuadas para el escaneado de los fondos, y que en modo alguno puedan deteriorar los originales. El modelo de escáner utilizado varía en función de la tipología del documento que se digitaliza, de acuerdo a las especificaciones técnicas de la BNE.

Los elementos que condicionan o intervienen en la elección del sistema de digitalización son:

- **Formatos:**
 1. Tamaño: El abanico de formatos comprende, al menos, desde un 8º a mayor de A1, teniendo en cuenta por tanto el porcentaje de los documentos
 2. Grosor/peso: El fondo contiene ejemplares de distintos grosores.
 3. Desplegables: relacionado con el formato del documento los desplegados suponen un añadido a tener en cuenta, tanto para la manipulación del documento como para las dimensiones del escáner.
- **Características del documento:**
 4. Original en color: La mayoría de los manuscritos contienen algún motivo coloreado de interés documental y necesario para su estudio o consulta. Esto hace necesario que el escáner ofrezca garantías de reproducción cromática fiel.
En relación con el color los códices miniados presentan una dificultad añadida en la reproducción de los dorados. No es aconsejable el modelo copibook para originales donde el color sea una característica esencial.
 5. Encuadernación: Las encuadernaciones rígidas no permiten una apertura total del libro en 180º; asimismo las encuadernaciones cerradas pueden tener pérdida de información en la parte central del documento. En ambos casos es aconsejable un escáner que permita realizar la reproducción página a página con objeto de obtener la menor pérdida de información, así como mayor planitud del documento.
 6. Material: Un porcentaje importante de manuscritos aparece en soporte pergamino. Las características específicas del soporte: hojas onduladas, pérdida de zonas en el pliego, dureza del mismo etc. determinan una manipulación especial, como inclusión de hojas que aislen las partes perdidas, y un escáner que asegure el foco en los distintos planos del documento.
- **Estado de conservación:**
 7. Fragilidad del material
 8. Falta de información: mutilaciones

- **Diversidad de fondos:** Los materiales específicos de las Secciones especiales, presenta diferentes características que motivan la diversidad de sistemas de escaneado:
 9. Códices en pergamino
 10. Papel ácido y friable
 11. Grabados
 12. Encuadernaciones históricas
 13. Dibujos
 14. Fotografías
 15. Colección de Ephemera
 16. Carteles publicitarios y grandes formatos en general

7.1.1. Tipos de escáneres

A *grosso modo*, los escáneres utilizados para la digitalización de los diferentes materiales conservados en la BNE se clasifican en:

- Escáneres tipo A: Para la digitalización en escala de grises de obras impresas (siglos XVIII a XIX)
 - CopibookHD600; i2s.
 - Bookeye 3 –R2
 - Book2net
 - ScannTECH 602i-6 ó 602i-3
- Escáneres tipo B: Para la digitalización a color de obras manuscritas o impresas encuadernadas, mayoritariamente ilustradas y de fondos gráficos en hojas sueltas (fotografías, carteles, mapas, colecciones de ephemera...)
 - Digibook Suprascan A1
 - Book2net A1
- Escáneres tipo C: Para la digitalización a color de obras que exijan una manipulación especialmente cuidadosa debido al tipo de soporte (códices, manuscritos miniados, manuscritos con tintas traspasadas, tintas ferrogálicas, encuadernaciones históricas con elementos metálicos...)
 - Metis DRS5070
- Escáneres tipo D, donde se incluye la cámara digital, para colecciones fotográficas y fondos no encuadernados de formato medio, así como el respaldo digital para originales de gran valor, especialmente delicados y que no puedan ser reproducidos mediante un escáner.

- Nikon D700 (calidad mínima)
- Nikon D3
- Respaldo digital Sinar 75 (cuatro disparos)

Se utiliza además el Robot escáner para aquellas obras cuyas características físicas y estado de conservación permite una actuación mecanizada sobre el documento sin riesgo.

Y escáneres que permiten ángulos de apertura de 60-90°, para obras que requieren este tipo de manipulación.

8. VOLCADO EN SISTEMA DE PRESERVACIÓN DIGITAL

En línea con los fines y funciones de la BNE los ficheros digitalizados se vuelcan en el sistema de preservación digital institucional. Concretamente, y como ya se ha referido se preservan: masters, másters editados, archivos de difusión y perfil de metadatos METS-PREMIS. Este sistema asocia al material una determinada política de preservación en la que se definen entre otros parámetros, el nº de copias, las políticas de acceso a los ficheros, la realización de comprobaciones periódicas...etc.

Antes de producirse la ingesta en el sistema se realizan una serie de comprobaciones que garantizan la integridad y completitud del material. El mismo sistema de preservación digital consta de un software que permite (entre otras funciones) la recuperación de copias de las imágenes de preservación con el fin de poder atender peticiones de reproducciones de los usuarios u otras instituciones. Se garantiza así preservación del material analógico (se evita repetir la reproducción digital), del material digital (los máster originales permanecen en el sistema de preservación), al tiempo que se trabaja a favor de la difusión (atender las peticiones de usuarios y/o instituciones).

9. MOTOR DE BÚSQUEDA

El objetivo principal de un motor de búsqueda en un proyecto de digitalización, es lograr que la localización y navegación sobre volúmenes inmensos de materiales digitalizados sea lo más sencilla, intuitiva y relevante posible.

Actualmente, la Biblioteca Digital Hispánica utiliza SOLR como motor de búsqueda, un software de búsqueda de código abierto, que permitirá poder extender y desarrollar funcionalidades propias, al disponer del código fuente.

A través de SOLR se indexan automáticamente los contenidos publicados en Pandora (SGOD), visibles en una interfaz de búsqueda personalizada.

Interfaz de la aplicación de búsqueda sencilla

Interfaz de la aplicación de búsqueda avanzada

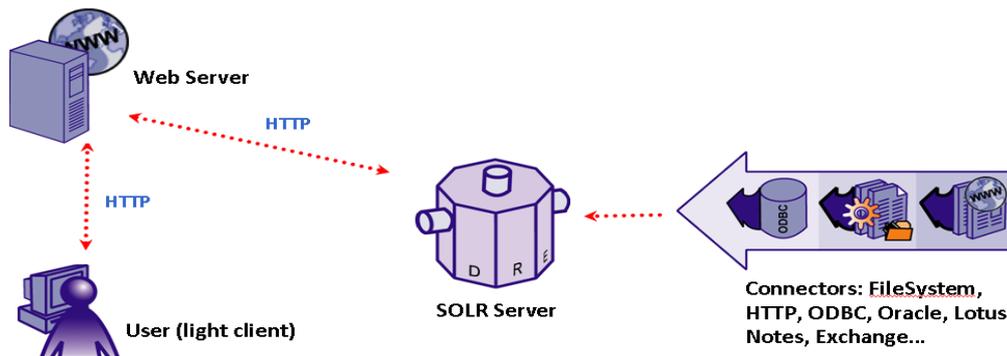
SOLR indexa a través de OAI, tanto el contenido estructurado (metadatos) como el contenido desestructurado (OCR).

Entre las funcionalidades que ofrece este motor de búsqueda a través de BDH se señalan las siguientes:

- Búsqueda básica y conceptual
- Sugerencias de búsqueda según se escribe (auto-cumplimentación)
- Búsqueda paramétrica (filtros de navegación)
- Hipervinculación (relaciones entre documentos)
- Expansión de consultas
- Resúmenes resaltando las palabras encontradas

Arquitectura SOLR en BDH

Basic Architecture



Basic Architecture:

- Front-end components: Internet Explorer, Firefox, Chrome...
- Back-end components: Solr Server
- Data Sources: DIGITool, File systems, web sites, DataBases ...

Arquitectura básica

GLOSARIO DE TÉRMINOS Y ABREVIATURAS

- **ACDsee**: Software de edición de imágenes digitales.
- **BDH**: Biblioteca Digital Hispánica
- **Bits**: El bit es la unidad mínima de información empleada en informática. Es un dígito del sistema de numeración binario, representado a través de dos valores: 0 ó 1.
- **BNE**: Biblioteca Nacional de España
- **CDU**: Clasificación Decimal Universal
- **Dpi**: Unidad de medida de la resolución de una imagen (relacionado con la calidad) de un escáner, una impresora, etc. Sirve para medir la resolución que es la cantidad de puntos (píxeles) que entran en una pulgada.
- **Dublin Core**: Es un modelo de metadatos elaborado por la DMCI (Dublin Core Metadata Initiative) una organización dedicada a fomentar la adopción extensa de los estándares interoperables de los metadatos y a promover el desarrollo de los vocabularios especializados de metadatos para describir recursos. Es el sistema de metadatos más popular en la descripción de recursos electrónicos en Internet. Define un conjunto de propiedades que se pueden usar en la descripción de un recurso (este disponible en formato electrónico o no) para facilitar su recuperación.
- **JPEG**: Es un formato de imagen para el almacenamiento y transmisión de imágenes en la Web. Los archivos de este tipo se nombran con la extensión

.jpg. Su algoritmo de comprensión permite reducir el tamaño de los ficheros, sin pérdida o con pérdida poco significativa de la calidad de imagen.

- **MARC 21:** Estándar internacional tradicionalmente utilizado por bibliotecas de todo el mundo para el intercambio de información catalográfica, con modificaciones que permiten la descripción de recursos electrónicos.
- **Marca de agua:** Es una técnica clásica utilizada para marcar papel. Una filigrana o marca al agua es una imagen formada por diferencia de espesores en una hoja de papel. Se utiliza para evitar la falsificación de documentos, para mostrar la autenticidad del origen de algún papel o impreso, como adorno o como diferenciación entre diferentes fábricas de papel.
- **Megabyte (MB):** Es una unidad de medida de cantidad de datos informáticos. Es la unidad más típica actualmente, junto al múltiplo inmediatamente superior, el gigabyte, usándose para especificar la capacidad de la memoria RAM, de las memorias de tarjetas gráficas, de los CD-ROM, o el tamaño de los programas, de los archivos grandes, etc. La capacidad de almacenamiento se mide habitualmente en gigabytes, es decir, en miles de megabytes.
- **METS:** Se denomina METS al fichero xml, que contiene los datos de un registro bibliográfico formado por varios archivos digitales (varios PDF o varios JPEG).
- **Metadatos:** Los metadatos son el conjunto de informaciones relacionadas con los objetos digitales, cuyo objetivo es facilitar la descripción, búsqueda, uso y gestión de las colecciones digitales. Son herramientas que permiten especificar la información contextual asociada a cada documento: su contenido, el historial de las transformaciones sufridas por cada objeto digital, las especificaciones de los equipos físicos necesarias para la construcción de los emuladores, los formatos de cada fichero, los programas que permitirán acceder a cada registro.
- **OAI-PMH:** El protocolo OAI-PMH (Open Archives Initiative-Protocol Metadata Harvesting), es una herramienta de interoperabilidad independiente de la aplicación que permite realizar el intercambio de información para que desde diferentes proveedores de servicio, se puedan hacer búsquedas que abarquen la información recopilada en distintos repositorios asociados. Los metadatos a transmitir vía OAI-PMH deberán codificarse en Dublin Core sin calificar con objeto de minimizar los problemas derivados de las conversiones entre múltiples formatos.
- **OCR:** Cuyo desarrollo de sus abreviaturas es Reconocimiento Óptico de Caracteres, es una tecnología que se encarga de escanear y reconocer los

caracteres de cualquier tipo de documentos.

El software OCR (Optical character recognition), transfiere esta información a formato electrónico, de forma rápida y precisa. No sólo captura y escanea los datos que contiene el documento, sino que también los almacena en un archivo o base de datos y les proporciona un formato capaz de ser reconocido y recuperado, para ser usado en otras aplicaciones. La utilización de la tecnología OCR permite la explotación de sus documentos y gestión electrónica, de forma ágil y segura. La captura de la información de los documentos o imágenes se puede efectuar manualmente desde un dispositivo, como un escaneador, que incorpora esta funcionalidad.

- **PDF (Portable Document Format):** Es un formato de almacenamiento de documentos desarrollado por la empresa Adobe Systems, especialmente adecuado para la presentación de documentos complejos (múltiples páginas, combinación de textos e imágenes de diferentes calidades). Este formato ofrece, entre otras ventajas, bastantes opciones de navegación en el documento y entre diferentes documentos, fidelidad y seguridad de la copia digital y posibilidades de búsqueda y recuperación a partir de los contenidos, incluyendo su inclusión en motores de búsqueda.
- **PhotoShop:** Software de edición de imágenes estándar para profesionales.
- **PREMIS:** Metadatos de preservación, que contienen la información que utiliza un repositorio para soportar el proceso de preservación digital.
- **SGOD:** Sistema de Gestión de Objetos Digitales.
- **Simplex:** Se denomina Simplex al fichero xml, que contiene los datos de un registro bibliográfico formado por un único archivo digital (PDF o JPEG).
- **TIFF (Tagged Image File Format):** Es un formato de ficheros para imágenes con etiquetas. Esto se debe a que los ficheros *TIFF* contienen, además de los datos de la imagen propiamente dicha, "etiquetas" en las que se archiva información sobre las características de la imagen, que sirve para su tratamiento posterior. Este formato es de aplicación generalizada a la creación de imágenes de alta calidad, produce ficheros de gran tamaño, sin pérdida, útiles como ficheros maestros pero inadecuados para la distribución y acceso público a las colecciones.
- **SYMPHONY:** Es un Sistema Integrado de Gestión Bibliotecaria, utilizado por diferentes bibliotecas universitarias. Actualmente utilizado por la BNE.

Para cualquier consulta o sugerencia dirijase a la siguiente dirección de correo electrónico: bibliotecadigital@bne.es