



Oferta digital protegida de la Biblioteca Nacional

2ª convocatoria 20/01/2010

Requisitos técnicos en ENCLAVE Editores - BNE

v3 15/01/2010

Promovido por



**FEDERACIÓN DE GREMIOS  
DE EDITORES DE ESPAÑA**



## Contenido

- 1 Objeto del proyecto 5**
  - 1.1 El proyecto ENCLAVE Editores - BNE 5
  - 1.2 Justificación 5
  - 1.3 Planteamiento inicial 5
  
- 2 Tipología de obras 6**
  - 2.1 Criterios de selección 6
  - 2.2 Otros criterios generales 6
  
- 3 Funcionalidad 7**
  - 3.1 Entorno de uso 7
  - 3.2 Recorrido del usuario e información necesaria 7
  
- 4 Acciones que llevar a cabo 9**
  - 4.1 Previas 9
  - 4.2 Editorial 9
  - 4.3 BNE 9
  
- 5 Gestión de información por las editoriales 10**
  - 5.1 Introducción 10
  - 5.2 Archivos que generar 10
    - 5.2.1 Formatos y gestión de los archivos 10
  - 5.3 Metadatos 12
    - 5.3.1 Metadatos necesarios para Enclave Editores-BNE 13
  - 5.4 Gestión de datos por parte de la editorial 15
    - 5.4.1 Obra digital completa 16
    - 5.4.2 Archivo para hojear 16
    - 5.4.3 Archivo para indexación a texto completo 16
    - 5.4.4 Metadatos 16
  
- 6 Recepción de archivos por la BNE 17**
  - 6.1 Formatos de archivos y mecanismos de envío DILVE → BNE 17
  - 6.2 Nomenclatura de archivos 18
    - 6.2.1 Convenio general 18
    - 6.2.2 Convenio específico para ODI-BNE 18
  - 6.3 Formato del archivo de metadatos 20

# 1 Objeto del proyecto

## 1.1 El proyecto ENCLAVE Editores - BNE

El proyecto *ENCLAVE Editores - BNE* (Oferta digital en la Biblioteca Nacional de España) es una iniciativa conjunta de la Biblioteca Nacional de España (BNE) y la Federación de Gremios de Editores de España (FGEE), que ha sido objeto de una subvención por parte del Ministerio de Industria, Turismo y Comercio en el marco del Subprograma Avanza Contenidos correspondiente a la segunda convocatoria de 2008.

El objetivo del proyecto es la creación y puesta en marcha de un *Centro del conocimiento estable dedicado al análisis, la definición, el desarrollo y la difusión de modelos de integración y explotación de contenidos digitales sujetos a derechos de autor en entornos de bibliotecas digitales*.

Su calendario de ejecución es de 24 meses: de noviembre de 2008 a octubre de 2010.

## 1.2 Justificación

En la Memoria técnica del proyecto se describe de esta forma el proyecto:

...

*La carencia, en la actualidad, de un modelo que establezca los términos (documentales, económicos, tecnológicos, jurídicos...) que deben regular este tipo de servicio provoca que no exista biblioteca nacional alguna que pueda incluir, entre sus servicios digitales, la posibilidad de acceso a la oferta global de libros disponibles en el mercado (sujetos a derechos de autor), aunque éstos formen parte de sus fondos.*

*Se puede crear, por tanto, la paradoja de que las enormes potencialidades que las tecnologías digitales tienen tanto para el sector del libro como para el mundo bibliotecario, se vean limitadas o puedan excluir, de facto, la consulta de la casi totalidad de la producción literaria, científica y artística contemporánea en entornos de bibliotecas digitales.*

*La búsqueda de soluciones que permitan abordar positivamente este problema va a tener una gran repercusión social, cultural y económica y permitirá dotar a los proyectos de bibliotecas digitales de servicios imprescindibles para sus usuarios.*

*ENCLAVE Editores - BNE pretende ser un centro de referencia para las bibliotecas que deseen ofrecer, en sus servicios digitales, acceso a contenidos sujetos a derechos de autor. Para ello trabajará en seis ámbitos complementarios —bibliotecario, editorial, funcional, económico, jurídico y tecnológico— e impulsará la creatividad e innovación de todos los agentes implicados en la cadena de valor del libro digital, en relación con la sociedad de la información.*

*ENCLAVE Editores - BNE se propone, en su primera fase de trabajo, un objetivo piloto concreto: la integración de contenidos digitales sujetos a derechos de autor en la Biblioteca Digital Hispánica (BDH), inaugurada en enero del 2008 por la Biblioteca Nacional de España. Esta meta servirá de hilo conductor inicial para el resto de los trabajos y acciones del Centro, al tiempo que permitirá medir, de forma objetiva, los avances logrados en cada etapa del proyecto.*

...

## 1.3 Planteamiento inicial

En su planteamiento inicial para la incorporación de obras sujetas a derecho de autor, *ENCLAVE Editores - BNE* explora un modelo en el que la BDH dispondrá, localmente, de toda

la información necesaria para permitir realizar búsquedas mediante cualquier término contenido en la obra<sup>1</sup>, integrando en los resultados ofrecidos a sus usuarios todos los datos bibliográficos completos referidos a la obra buscada así como información adicional tal como la imagen de la cubierta, un resumen, el sumario, y una muestra del contenido de la obra.

Por otra parte, a partir de la ficha de una obra sujeta a derechos de autor localizada en el catálogo de la BDH, el usuario puede entrar directamente en un entorno gestionado por la editorial correspondiente (plataforma de e-distribución), en el que esta ofrece la posibilidad de acceder al contenido completo de la obra, consultarla, descargarla, etc. de acuerdo con las condiciones comerciales particulares establecidas por cada editorial.

## 2 Tipología de obras

### 2.1 Criterios de selección

La BNE ha fijado unos criterios para la selección de obras participantes en esta fase del proyecto *ENCLAVE Editores - BNE* en un documento específico (*Criterios de selección de obras en la Biblioteca Digital Hispánica*). Entre otros, se indican tres de especial relevancia para este proyecto:

- Los textos tendrán prioridad sobre la imagen.
- Se preferirán las obras originales a los trabajos críticos.
- Se priorizarán las obras más estrechamente relacionadas con la cultura: filosofía, religión, ciencias sociales, artes, literatura, geografía, lenguas, ciencia y tecnología.

### 2.2 Otros criterios generales

- En esta fase del proyecto, solo se incluirán *obras de las cuales la editorial correspondiente detente, de forma explícita, los derechos de explotación en formato digital* y para las cuales el entorno de uso planteado en ENCLAVE Editores - BNE no colisione con las condiciones establecidas por los correspondientes derechohabientes.
- Solo se considerarán *versiones digitales de obras existentes en formato papel y que la BNE tenga catalogadas*. Por lo tanto, no se considerarán obras digitales puras, por no existir, de momento, la ficha catalográfica correspondiente de este tipo de obras en la BNE. En este sentido es importante resaltar que se manejarán manifestaciones concretas de obras, es decir, libros digitales que estarán identificados por un ISBN propio.
- Se incorporarán obras digitales que *permitan la compra de su contenido* teniendo en cuenta que estos accesos se harán, en todo caso, en entornos controlados por la editorial, o por las entidades en las que esta delegue, y bajo las condiciones técnicas y comerciales que la editorial establezca. A estos entornos se los denominará *plataformas de e-distribución* en el resto del documento.
- Cada editorial se *comprometerá a mantener operativas sus plataformas de e-distribución*, accesibles desde la BDH a través de Internet, durante el proyecto.
- La editorial se *compromete a suministrar la información y los ficheros* que se mencionan en el apartado *Gestión de información por las editoriales* de este documento. En particular, la editorial deberá facilitar a la BNE un extracto (contenido parcial) de cada obra, en

---

<sup>1</sup> Lógicamente, ENCLAVE Editores - BNE permitirá también las búsquedas habituales en entornos bibliotecarios basadas en los campos de título, autoría, materias, fecha, etc. además de las búsquedas específicas mediante términos que figuren en texto de la obra.

formato PDF, que se mostrará a los usuarios del servicio de forma libre y sin restricciones, en el entorno de la BDH.

## 3 Funcionalidad

### 3.1 Entorno de uso

El entorno de uso es el de la Biblioteca Digital Hispánica (BDH), donde las obras sujetas a derecho de autor conviven con las obras de dominio público digitalizadas por la Biblioteca Nacional de España.

En <http://bdh.bne.es/enclave> se muestran las obras incorporadas en este proyecto.

### 3.2 Recorrido del usuario e información necesaria

En este apartado se compara de forma sucinta el recorrido que, dentro de la BDH, hace un usuario para acceder al contenido de una obra digital sujeta a derechos de autor por un lado, con la información *adicional* que necesita la BDH, suministrada por la editorial, para mostrar la ficha de la obra.

La BDH basándose en los datos bibliográficos de catálogo de las versiones impresas mostrará las obras digitales incorporadas dentro del marco del proyecto *ENCLAVE Editores - BNE*.

Esta descripción es muy sintética y su objetivo no es definir las funcionalidades del interfaz de uso de la BDH sino exponer un esquema estructurado para la descripción de la información que debe suministrar la editorial para que la BDH pueda implementar estas funcionalidades, y que se presenta en posteriores apartados de este documento.

RECORRIDO DE UN USUARIO EN LA BDH	INFORMACIÓN QUE LA EDITORIAL DEBE PROPORCIONAR
<p>1. El usuario entra en la BDH y realiza una búsqueda. Entre los criterios de búsqueda puede incluir términos que deben aparecer en el cuerpo de las obras.</p> <p>El sistema de la BDH lleva a cabo la búsqueda correspondiente y ofrece un listado con los resultados. El usuario puede refinar la búsqueda, modificar el aspecto o el orden de las obras en el listado, <i>recordar</i> fichas para su descarga posterior, etc.</p> <p>Los resultados pueden incluir obras de dominio público así como obras sujetas a derechos de autor. Ambos tipos de obra se diferenciarán en los listados de las fichas correspondientes mediante iconos adecuados u otro tipo de representación visual.</p> <p>En el caso de <i>obras de dominio público su versión digital estará residente en la BDH</i> y será de acceso libre.</p> <p>En el caso de <i>obra sujeta a derechos de autor la versión digital estará residente en el entorno del e-distribuidor</i>, controlado por la editorial.</p>	<p>Para permitir las búsquedas mediante términos que aparecen en el interior de la obra, la <i>BDH necesita indexar</i> estos términos añadiéndolos a su motor de búsquedas y asociándolos a la obra correspondiente.</p> <p>Esta operación <i>requiere que la BDH disponga de un archivo con el texto completo de la obra</i>. Este archivo se utilizará exclusivamente para la indexación textual del contenido y, en ningún caso, será accesible por el usuario ni siquiera de forma parcial.</p> <p>Por otra parte, cada obra tendrá asociados una serie de términos clave de materias. La asignación de estos términos se realiza por parte de los responsables de su catalogación en la BNE.</p>
<p>2. Si el usuario selecciona una obra sujeta a derechos de autor, la BDH muestra su ficha completa incluyendo, entre otros, información bibliográfica completa (extraída del catálogo general de la BNE), la imagen de la cubierta, un resumen y, en su caso, el sumario<sup>2</sup>.</p> <p>Igualmente, y sin salir del entorno de la BDH, el usuario puede visualizar un fragmento de la obra, seleccionado y suministrado por la editorial a la BDH.</p>	<p>La información adicional incluida en la ficha de la obra: imagen de cubierta, resumen y sumario, deberá ser suministrada por la editorial, que también deberá facilitar algunos metadatos adicionales, como el ISBN de la versión digital, etc.</p> <p>De la misma forma, <i>la editorial deberá facilitar a la BDH el fragmento de la obra</i> mostrado al usuario.</p>
<p>3. Para permitir al usuario acceder al contenido completo de la obra, en la ficha de la obra presentada por la BDH figurará un hiperenlace que lleva directamente a la plataforma de e-distribución, controlada por la editorial o por la entidad en quien esta haya delegado, en la que el usuario puede hojear el interior de la obra y descargarla, leerla, etc. de acuerdo con las condiciones comerciales establecidas por la editorial en cada caso.</p>	<p>Cada editorial decidirá el entorno web en el que estará residente su oferta de obra digital para los usuarios (plataforma de e-distribución). Podrá ser una web de la propia editorial o el entorno de un e-distribuidor con el que la editorial haya llegado a un acuerdo.</p> <p>En cualquier caso, para cada obra concreta, <i>la editorial deberá suministrar a la BDH el URL de enlace a su e-distribuidor</i> para permitir construir el hipervínculo correspondiente.</p>

<sup>2</sup> Dependiendo del tipo de obra, esta información puede tener sentido o no. Por ejemplo, en el caso de un diccionario o una enciclopedia, el sumario, en general, no existirá.

## 4 Acciones que llevar a cabo

### 4.1 Previas

Las editoriales que participen en el proyecto realizarán una selección de obras de acuerdo a los criterios proporcionados por la BNE en el documento elaborado a tal efecto (en concreto, en el apartado *Tipología de obras*), así como otras posibles recomendaciones de selección en función de los avances que se produzcan en el proyecto.

Posteriormente se firmarán los correspondientes acuerdos entre la BNE, la FGEE y la editorial, para la incorporación de sus obras al catálogo de la BDH.

### 4.2 Editorial

Después de seleccionar las obras que formarán parte del proyecto *ENCLAVE Editores - BNE*, la editorial deberá generar los correspondientes ficheros, metadatos básicos y datos enriquecidos. Si la editorial ya dispone de una versión digital de la obra, deberá asegurarse de que los datos correspondientes cumplen los requerimientos particulares del proyecto *ENCLAVE Editores - BNE*. En caso contrario deberá modificarlos o completarlos.

Posteriormente, para cada obra seleccionada la editorial deberá:

1. Cargar la obra digital en la plataforma de e-distribución para su gestión, distribución y comercialización. Esta plataforma podrá estar directamente gestionada por la editorial o por la entidad (e-distribuidor) en que esta delegue tales funciones. La editorial se comprometerá a mantener la obra en esta plataforma (u otra con las mismas funcionalidades) mientras dure el proyecto *ENCLAVE Editores - BNE*. En caso de que cambie la ubicación de la obra en el transcurso del proyecto, la editorial deberá comunicar la nueva ubicación lo antes posible para evitar interrupciones en el servicio.
2. Cargar en DILVE los metadatos y datos de catálogo ricos que se describen el apartado correspondiente. Igualmente, deberá cargar en DILVE tanto el archivo con el fragmento de la obra (contenido parcial) como el archivo para indexación del texto completo.
3. Enviar a la BNE a través de DILVE > ENCLAVE, cada ficha de obra digital.

El detalle de los formatos y contenidos de estos archivos se expone más adelante.

### 4.3 BNE

Una vez recibidos los archivos y la información complementaria correspondientes a una versión digital de una obra la BNE llevará a cabo, entre otras, las siguientes operaciones, necesarias para la correcta integración de la obra en la BDH.

1. Dar de alta a la obra en el catálogo de la BDH, utilizando para ello los datos de catálogo existentes en la BNE, correspondientes a la versión impresa.
2. Incorporar a la ficha de la obra los datos de catálogo ricos (imagen de cubierta, resumen y sumario) aportados por la editorial por medio de DILVE.
3. Añadir, como recurso asociado a la obra, el archivo para indexación del texto completo, procediendo al proceso de indexación y marcando dicho recurso como *no público*.
4. Añadir, como recurso asociado a la obra, el archivo con el contenido parcial, marcando dicho recurso como *público*.
5. Establecer los mecanismos necesarios para que, en la ficha de la obra, figure el enlace al entorno externo de acceso a la misma (plataforma de e-distribución).

## 5 Gestión de información por las editoriales

### 5.1 Introducción

Para poder incorporar sus obras a *ENCLAVE Editores - BNE* las editoriales deben disponer de los archivos y demás metadatos e información de catálogo enriquecida que se exponen en este apartado.

### 5.2 Archivos que generar

En esta sección se enumeran *de forma sucinta* los archivos necesarios. En el apartado siguiente se presentan más detalles de sus características y entornos de uso.

Estos archivos son:

**a** *Obra digital en sí.*

En general, estará compuesta por un archivo en el formato adecuado y con todo el contenido de la obra, preparado para poder ser utilizado en el entorno de la plataforma de e-distribución. Este archivo contendrá toda la información adicional complementaria habitual libros digitales: referencias cruzadas, marcadores internos, etc. La editorial puede optar por comercializar la obra en varios formatos :PDF. ePUB, etc. como se expone más adelante. Estos archivos estarán residentes en la plataforma de e-distribución.

**b** *Archivo para hojear.*

Será un archivo, con todo el contenido de la obra, que permita la funcionalidad de hojear en la *plataforma de e-distribución*.

La plataforma de e-distribución decidirá el formato necesario para esta funcionalidad e implementará las reglas de negocio y mecanismos necesarios para que solo se muestren las partes de la obra que decida la editorial.

**c** *Contenido parcial de la obra.*

Será un archivo, en formato PDF, con fragmentos elegidos de la obra que, normalmente, se generará a partir del PDF con la totalidad de la obra.

**d** *Archivo para indexación*

Es un archivo con todo el contenido de la obra que se utilizará por parte de la BDH para la indexación del texto completo. La BNE necesita este archivo en formato PDF. Para poder llevar a cabo esta indexación la editorial suministrará una copia del archivo a la BDH, a través de DILVE, según se explica más adelante.

#### 5.2.1 Formatos y gestión de los archivos

Este apartado describe con más detalle los archivos enumerados en la sección anterior.

**a** *Obra digital en sí.*

Son los archivos correspondientes a la *totalidad de la obra* en su versión digital aptos para ser descargados y leídos off-line, o para su consulta on-line o las combinaciones de ambos procedimientos que determine la editorial, de acuerdo con sus planes y objetivos comerciales. Estos constituyen la obra digital en sí, el *libro digital*.

Este libro digital podrá estar compuesto por uno o varios ficheros, y estarán sujetos a transacción comercial, como cualquier otro producto editorial. Estos ficheros serán los que se descargue o consulte los usuarios finales una vez satisfechas las condiciones económicas y de acceso establecidas por la editorial correspondiente, en su entorno de e-distribución.

Algunos de los formatos más utilizados en este momento (enero 2010) son:

- PDF
- ePub
- MobiPocket

Cada editorial podrá comercializar la obra digital en más de un formato, generando estos archivos en los formatos correspondientes para llegar así a más entornos de explotación. La elección de estos formatos es competencia exclusiva de la editorial y será objeto de una decisión empresarial sopesando el coste de la generación de cada formato con el mercado potencial que pueda cubrir. La existencia de la obra digital en varios formatos, en la plataforma de e-distribución, no afecta a la operación de ENCLAVE.

En función del tipo de obras y del proyecto concreto de digitalización que la editorial establezca, las editoriales pueden considerar la generación de una versión base de la obra digital (probablemente en algún formato derivado del XML) que permita la generación, no traumática, de nuevos formatos cuando éstos sean necesarios.

En este sentido, el formato ePub, al estar basado en XML, es más flexible que el PDF y el MobiPocket. Existen herramientas que simplifican la generación de los formatos PDF y MobiPocket desde ePub. La transformación inversa es, en general, más costosa.

Sin embargo, la obtención inicial del formato ePub es, normalmente, más compleja y más cara que la generación directa de un formato PDF. También, para obras con una gran riqueza tipográfica, con esquemas e ilustraciones, etc. el formato ePub (en su versión actual) tiene limitaciones.

En cualquier caso y como se ha comentado, la elección de los formatos de la obra digital para su descarga o compra por parte del usuario final es una decisión que debe tomar la editorial, basándose en el mercado al que se dirige la obra y en su modelo de negocio digital, comparando costes y prestaciones de las distintas opciones, etc. Son decisiones similares a las que se refieren a la producción física del libro impreso: tipo de papel, tinta y tipografía, encuadernación, volumen de la tirada, etc.

La editorial debe considerar los requerimientos y prestaciones de la plataforma de e-distribución elegida, a la hora de decidir el (o los) formatos de la obra digital.

Los archivos correspondientes al producto digital en sí estarán residentes en un entorno de venta electrónica (plataforma de e-distribución) gestionado directamente por la propia editorial o por la entidad en quien delegue estas funciones.

Si un libro digital está disponible en varios formatos, cada uno de ellos se considera un producto comercial distinto, por lo que cada uno deberá tener su propio ISBN.

*No se enviarán ni a la Biblioteca Nacional (véase el punto siguiente) ni a DILVE.*

**b** *Archivo para la funcionalidad de hojear.*

Este es el archivo necesario para permitir la funcionalidad de *hojear de la obra*, con las características y limitaciones de acceso establecidas por la editorial. Esta funcionalidad se prestará en la plataforma de e-distribución ya que está íntimamente relacionada con las funcionalidades de acceso comercial a la obra digital.

Los procesos necesarios para implementar la funcionalidad de hojear dependerán de las características y la tecnología de la plataforma de e-distribución elegida por la editorial. En este entorno, y de acuerdo con el e-distribuidor, la editorial deberá decidir qué porcentaje o qué partes de la obra estarán disponibles en la función de hojear.

*No se enviará ni a la Biblioteca Nacional ni a DILVE.*

**c** *Contenido parcial.*

Este es un archivo en formato PDF correspondiente a una parte de la obra (*contenido parcial*).

Este archivo deberá estar en formato PDF, con el mismo detalle que la obra en cuanto a formateos, tipografía, etc. Su finalidad es facilitar a los usuarios, dentro del contexto de ENCLAVE Editores - BNE, una referencia respecto al tipo de contenido de la obra, de su tipografía, del nivel de tratamiento del tema, etc.

La editorial facilitará a la BNE un archivo único con todos los fragmentos seleccionados de la obra. Se recomienda que este archivo incluya el sumario y un fragmento significativo de la obra: p.ej. un capítulo, la introducción, etc. La editorial será responsable de la selección de estos fragmentos que, por otra parte, dependerán del tipo de obra. En cualquier caso se debe buscar un equilibrio entre el tamaño de este archivo (que no debería superar 1 o 2 MBytes) y su capacidad para ilustrar respecto al contenido de la obra. Aunque, como referencia, se recomienda que este archivo contenga en torno al 20% de la obra, este puede ser inferior en el caso de obra de gran extensión o de naturaleza muy homogénea (p.ej. obras de referencia, etc.).

En general, la editorial podrá generar este archivo a partir del PDF con la totalidad de la obra eliminando del mismo todas las partes que no desee hacer públicas y ajustando los enlaces internos adecuadamente.

Estos contenidos se mostrarán de *forma libre a todos los usuarios* de ENCLAVE Editores - BNE.

*Este archivo residirá físicamente en la BDH y la editorial deberá cargarlo en DILVE* junto con el resto de los metadatos y datos enriquecidos de la obra.

**d** *Archivo para indexación del texto completo de la obra.*

Para facilitar las búsquedas basadas en términos en el entorno de la BDH es preciso que la BNE disponga de un archivo que contenga todos los términos de cada obra, con el único fin de su proceso mediante los mecanismos de indexación.

La herramienta de gestión de contenidos digitales utilizada en la BDH (DigiTool de ExLibris) requiere que los archivos para indexación del texto completo estén en formato PDF.

Por lo tanto, para cada obra incorporada a ENCLAVE Editores - BNE, la editorial deberá facilitar a la BNE un archivo en formato PDF que contenga el texto completo con el fin exclusivo de su indexación.

Este archivo no será accesible por los usuarios de la BDH y en ningún caso se mostrará su contenido, quedando protegido mediante los mecanismos de restricción de acceso incorporados en la herramienta DigiTool.

No es necesario (ni conveniente) que este archivo tenga toda la riqueza tipográfica y de formateo que contienen los archivos de la obra digital, siendo suficiente con que incluya el texto puro de la misma.

*DILVE facilitará mecanismos para simplificar el envío de este archivo a la BNE por parte de las editoriales, manteniendo la confidencialidad del archivo.*

### 5.3 Metadatos

Es importante reseñar que los libros en formato digital susceptibles de ser incorporados al proyecto *ENCLAVE Editores - BNE* serán siempre versiones digitales de libros existentes en

formato impreso. Por lo tanto, tendrán un ISBN propio, y un registro específico en DILVE con sus propios metadatos y demás información de catálogo.

La versión impresa de estas publicaciones deberá estar indexada y referenciada en el catálogo general de la BNE.

Esta sección describe los metadatos específicos necesarios, además de los archivos descritos en la sección anterior, para la incorporación de obras digitales a *ENCLAVE Editores - BNE*.

En este contexto se considerará el término *metadato* en un sentido amplio, incluyendo todos aquellos *datos de carácter público sobre el libro* susceptibles de ser consultados libremente en distintos entornos y por distintas vías.

En esta categoría se incluyen, al margen de los metadatos habituales, otros datos de catálogo tales como las *imágenes de cubierta, resúmenes, tabla de contenidos o sumario, etc.* Igualmente, se incluye *el archivo de contenido parcial* que deberá cargarse en DILVE junto con el resto de metadatos.

Las editoriales cargarán en DILVE la información de catálogo de sus obras en formato digital, por las vías y en los formatos habituales: FTP, web, ONIX, CSV, etc.

Entre estos metadatos deben figurar, además de los que habitualmente estén cargando en DILVE las editoriales, aquellos establecidos por la BNE como necesarios para el proyecto *ENCLAVE Editores - BNE* y que se enumeran a continuación.

Si la editorial comercializa la obra digital en distintos formatos, solo será necesario enviar a *ENCLAVE Editores - BNE* los metadatos de una de las versiones, aunque todas ellas deberán cargarse en DILVE (cada una con su propio ISBN).

La carga de los datos de catálogo en DILVE, por parte de las editoriales, es ya un proceso normal y estandarizado, y asegura tanto la accesibilidad por parte de cualquier consumidor de información como su calidad y homogeneidad, dado que DILVE hace las oportunas comprobaciones de completitud, consistencia y calidad de los datos cargados.

### 5.3.1 Metadatos necesarios para Enclave Editores-BNE

Para su explotación en el proyecto *ENCLAVE Editores - BNE* es necesario que las editoriales carguen en DILVE, *al menos*, la siguiente información correspondiente a las versiones digitales de las obras impresas seleccionadas<sup>3</sup>, que se utilizará en el contexto del proyecto como *complemento* a la información de catálogo ya existente en la BNE:

- ISBN de la obra en formato digital
- ISBN de la versión original impresa
- Año de publicación de la versión digital<sup>4</sup>

<sup>3</sup> Para permitir la utilización de la información existente en DILVE relativa a la obra digital, al margen del contexto del proyecto ENCLAVE EDITORES – BNE, *la editorial deberá cargar todos los datos necesarios para su correcta identificación*: título, menciones de responsabilidad, precio, detalles de formato/peso de los archivos, etc.

<sup>4</sup> La obra digital, como producto editorial con entidad propia y con su propio ISBN, tendrá una fecha de edición o publicación propia. Si esta obra es simplemente una versión digitalizada de la correspondiente obra en papel, esta fecha no será muy relevante.

Sin embargo, es previsible que las obras digitales tengan una evolución editorial al margen de la obra original en papel, pudiéndose crear nuevas ediciones con contenidos adicionales o revisados, o con nuevas prestaciones (enlaces a webs temáticas, nuevas funcionalidades, etc.) por lo que, en un futuro,

- Imagen de la cubierta
- Archivo PDF con contenido parcial
- Resumen / Sinopsis
- Sumario / Tabla de contenidos si procede
- URL de acceso a la plataforma de e-distribución de la obra
- Opcionalmente, términos clave para identificación de materias

Esta lista podrá ir aumentando durante el desarrollo del proyecto *ENCLAVE Editores - BNE* de acuerdo con la evolución de las necesidades y posibilidades de los distintos agentes implicados.

Los metadatos se cargarán en DILVE por las vías habituales. DILVE proporcionará las plantillas e información necesaria para facilitar las tareas de carga a las editoriales.

Como se ha comentado anteriormente, si la editorial comercializa la obra digital en distintos formatos, solo será necesario enviar a *ENCLAVE Editores - BNE* los metadatos de una de las versiones, aunque todas ellas deberán cargarse en DILVE (cada una con su propio ISBN) para que la información correspondiente a todas las versiones esté disponible para todos los usuarios de DILVE.

A continuación se enumeran las características de los metadatos más relevantes para *ENCLAVE Editores - BNE*.

#### **Imagen de cubierta**

En RGB y formato JPEG con calidad alta. Con un tamaño del lado máximo de entre 700 y 1000 píxeles.

- Este dato se cargará en DILVE mediante el campo CSV *imagen\_cubierta*.

#### **Contenido parcial**

En formato PDF, con un 20% de la obra aproximadamente o con un porcentaje menor siempre que sea lo suficientemente representativo. Se recomienda que el peso de este archivo no supere 1 o 2 MBytes.

- Este dato se cargará en DILVE mediante el campo CSV *extracto*.

#### **Resumen**

En texto puro, sin ningún formateo. Es recomendable que su longitud no supere las 300-350 palabras o 2000 caracteres.

- Este dato se cargará en DILVE mediante el campo CSV *texto\_resumen*.

#### **Sumario, tabla de contenidos**

Es un dato opcional. Debe estar en texto puro, sin ningún formateo. Las partes o secciones de la tabla de contenidos se separarán mediante dos guiones seguidos: "--". La extensión de este campo dependerá de la estructura del libro. Si la tabla de contenidos es compleja o incluye muchos niveles anidados, se recomienda simplificarla.

Ejemplo:

---

el dato de la fecha de publicación de una versión digital puede ser un dato de especial relevancia igual que lo es el de la publicación de la versión impresa.

"1.1. La alfabetización digital en la sociedad del aprendizaje. Acciones para evitar la brecha digital -- 1.2. Las tecnologías de la información y la comunicación al servicio del aprendizaje -- 1.3. El aprendizaje permanente en la sociedad del conocimiento "

Si, por la naturaleza de la obra, el sumario no aporta información adicional significativa este campo se puede obviar. Por ejemplo, no tendrá sentido cargar un sumario del tipo: "Capítulo 1 -- Capítulo 2 -- Capítulo 3... etc".

- Este dato se cargará en DILVE mediante el campo CSV *tabla\_materias*.

NOTA: Aunque la editorial decida incluir el sumario en el archivo PDF con el *contenido parcial* de la obra para su hojear en el contexto BDH, deberá *también* enviar el sumario en formato texto según se indica en este apartado. Esta información se incluirá, junto con el resumen de la obra, en la ficha de la obra mostrada al usuario.

### **Ubicación (URL) de acceso a la plataforma de e-distribución de la obra digital**

La editorial deberá facilitar el URL de acceso a la plataforma de e-distribución de la obra digital. Este URL deberá apuntar al lugar concreto desde el cual se puede iniciar el proceso de hojear o compra de la obra correspondiente.

En esta ubicación deberá figurar en lugar visible una referencia al proyecto *ENCLAVE Editores - BNE*, de acuerdo con las especificaciones que se establecerán en su momento.

Este URL debe ser permanente, es decir, no deberá cambiar a lo largo del tiempo ni depender de las condiciones del entorno desde el cual se utilice. El URL será público y utilizable por cualquier usuario de DILVE.

Este dato se cargará en DILVE mediante el campo CSV *URL\_hojear\_producto*.

### **Relación entre obra en papel y su versión digital**

Habrá que cargar dos campos en DILVE:

- En la ficha correspondiente a la obra impresa, se cargará un campo que especifique el ISBN de la versión digital (si la obra está disponible en varios formatos digitales, deberán especificarse los ISBN de todos ellos).
- En la ficha correspondiente a la versión digital, se cargará un campo que especifique el ISBN correspondiente a la versión impresa.

#### Ficha de la obra impresa

Se cargará en DILVE el campo CSV *ISBN13\_edicion\_digital*, con el ISBN correspondiente a dicha versión de la obra (si existen varias versiones digitales, se indicarán todos sus ISBN, separados por punto y coma)

#### Ficha de la obra digital

Se cargará en DILVE el campo CSV *ISBN13\_edicion\_impresa*, con el ISBN correspondiente a dicha versión de la obra

## **5.4 Gestión de datos por parte de la editorial**

En este apartado se resumen los procesos de gestión que la editorial debe llevar a cabo con cada uno de los archivos o metadatos comentados anteriormente.

#### **5.4.1 Obra digital completa**

Los archivos correspondientes estarán residentes en los sistemas de la plataforma de e-distribución.

En general, será un solo archivo generado específicamente para este fin, en formato PDF, ePUB, etc..

Se los facilitará directamente la editorial al e-distribuidor, o la entidad responsable de su generación, siempre que así lo decida la editorial. No pasan por DILVE.

#### **5.4.2 Archivo para hojear**

El archivo correspondiente estará residente en la plataforma que implemente la funcionalidad de hojear que, normalmente, coincidirá con la de e-distribución.

Se lo facilitará directamente la editorial al e-distribuidor, o la entidad responsable de su generación, siempre que así lo decida la editorial. No pasa por DILVE.

#### **5.4.3 Contenido parcial**

Deberá estar en formato PDF, con un 20% de la obra aproximadamente o con un porcentaje menor siempre que sea lo suficientemente representativo. Se recomienda que el peso de este archivo no supere 1 o 2 MBytes. Se debe cargar en DILVE.

#### **5.4.4 Archivo para indexación a texto completo**

El archivo para indexación forma parte del corpus del contenido de la obra y no es un dato público. Este archivo se cargará en DILVE ya que la plataforma dispone de mecanismos ad-hoc que permiten canalizar su envío a la BNE, sin perder la confidencialidad de estos archivos.

#### **5.4.5 Metadatos**

Todos los metadatos se cargarán en DILVE, incluyendo el archivo PDF con el contenido parcial.

Existen mecanismos para el envío automatizado de estos datos desde DILVE a la BNE en los formatos internos acordados y mediante los mecanismos técnicos adecuados.

Este proceso de envío de los datos de cada obra a la BNE lo desencadenará la propia editorial una vez que los datos estén cargados y comprobados en DILVE.

## 6 Recepción de archivos por la BNE

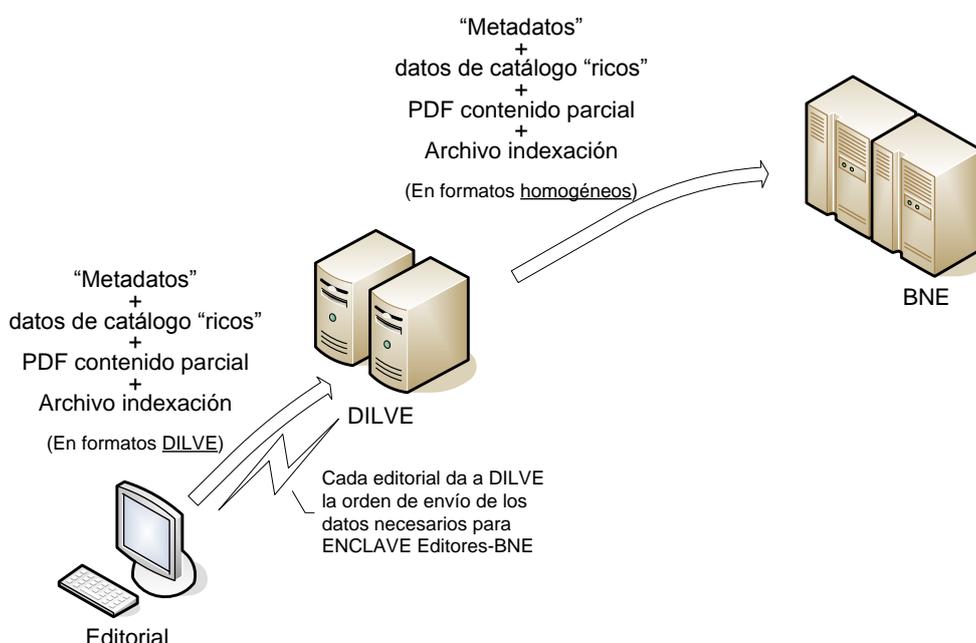
Todos los envíos de archivos a la BNE se harán vía FTP a los directorios especialmente habilitados el efecto.

Las editoriales enviarán a la BNE, desde DILVE > ENCLAVE, mediante un proceso automatizado, los metadatos y archivos asociados de las obras correspondientes a través de los mecanismos y en los formatos acordados, asegurando de esta forma la homogeneidad en formatos y vías de transmisión, así como la calidad de los datos

Al no depender de las capacidades de gestión informática de cada una de las editoriales, y tener un único interlocutor (DILVE) se simplificará notablemente la gestión del proceso por parte de la BNE.

Si la editorial comercializa la obra digital en distintos formatos, solo será necesario enviar a *ENCLAVE Editores - BNE* los metadatos de una de las versiones, aunque todas ellas deberán cargarse en DILVE (cada una con su propio ISBN).

El esquema siguiente muestra este método de acceso a los datos por parte de la BNE:



Con el fin de simplificar la identificación de los archivos recibidos se utilizará el convenio de nomenclatura de archivos expuesto en el apartado correspondiente.

### 6.1 Formatos de archivos y mecanismos de envío DILVE → BNE

El envío de metadatos desde DILVE a la BN será gestionado por cada editorial. DILVE ha desarrollado las funcionalidades necesarias para automatizar estos procesos. Existe una ubicación FTP en la BNE con dicho propósito.

Los metadatos se enviarán desde DILVE mediante cuatro archivos, por cada obra considerada:

- Un archivo en **XML** con los metadatos correspondientes en campos MARC21, incluyendo los **campos textuales**: resumen y sumario. No se descarta la utilización de un archi-

vo en formato ONIX en una fase posterior del proyecto. Lo genera automáticamente DILVE a partir de la ficha creada por la editorial.

- Un archivo en formato **JPEG** conteniendo la **imagen de la cubierta**.
- Un archivo en formato PDF con el **contenido parcial**, con la imagen de cubierta incluida automáticamente por DILVE.
- Un archivo para indexación con el contenido completo de la obra, manteniendo la confidencialidad del mismo.

## 6.2 Nomenclatura de archivos

Con el fin de simplificar el proceso de los archivos recibidos por la BNE es recomendable adoptar un convenio de nomenclatura de archivos que permita conocer, a partir de su nombre, a qué obra pertenece cada uno así como la naturaleza de su contenido.

En este sentido, se podría tomar como base la norma de nomenclatura que se utiliza en DILVE, expuesta a continuación.

### 6.2.1 Convenio general

En DILVE, se recomienda que los nombres de archivos con recursos asociados a un libro cargado en la plataforma sigan el siguiente convenio general:

```
nnnnnnnnnnnnnnn_oo_cc_t.ext
```

donde:

<b>nnnnnnnnnnnnnnn</b>	ISBN13 de la publicación. <u>En todo caso, sin guiones, espacios en blanco, ni otros separadores.</u>
<b>oo</b>	Número de la lista ONIX a la que pertenece el recurso correspondiente al archivo. P.ej. lista 33 para <i>Otro texto</i> , lista 38 para <i>Archivo de imagen/audio/video</i> .
<b>cc</b>	Código de la lista ONIX "oo" (Lista 33 o 38), que indica el elemento del libro al que se refiere el archivo..
<b>t</b>	<u>Solo para archivos de imagen y no debe figurar para el resto de los archivos.</u> Es una letra, de uso convencional en DILVE, utilizada en el caso de archivos de imagen. Indica el tamaño aproximado de la imagen (en píxeles). En principio, no se utilizará en ODI-BNE.
<b>ext</b>	Es la extensión del archivo. Indica el formato del mismo: <i>txt</i> =formato texto, <i>jpg</i> = formato JPEG, <i>png</i> = Portable Network Graphics, <i>pdf</i> = portable document format. No todos los formatos son aceptables para todos los tipos de contenidos.

### 6.2.2 Convenio específico para ENCLAVE

En el caso concreto del proyecto *ENCLAVE Editores - BNE*, solo se manejarán cuatro ficheros por cada obra digital y cada uno de ellos será siempre de la misma naturaleza, lo que permite adoptar un convenio simplificado específico para este ámbito.

Los cuatro archivos son:

- Archivo en **MARC21-XML con los metadatos** (incluyendo resumen y sumario), generado automáticamente por DILVE.
- **Imagen de cubierta**, en formato JPEG
- **Archivo de contenido parcial** en formato PDF
- **Archivo para indexación** en formato PDF

Por otra parte, se deben tener en cuenta los siguientes aspectos:

- El archivo con la imagen de cubierta se identifica con el código 04 de la Lista ONIX 38
- Un archivo con contenido parcial se identifica, genéricamente, con el código 23 de la Lista ONIX 33
- El archivo con el texto completo se identifica con el código 34 de la Lista ONIX 33
- El archivo XML con metadatos no se corresponde con ningún recurso estándar del ONIX, por lo que será objeto de una nomenclatura específica en el contexto ENCLAVE Editores - BNE.

Teniendo en cuenta los puntos anteriores se propone el siguiente convenio de nomenclatura para los archivos enviados por la editoriales a la BN:

- El archivo con la **imagen de cubierta** se denominará con el ISBN13 (sin espacios ni guiones), el sufijo **\_L38\_04\_x** (elemento 04 de la lista 38) y la extensión .jpg. En lugar de la letra "x" puede figurar otra, dependiendo del tamaño de la imagen.
- El archivo con el **contenido parcial** se denominará con el ISBN13 (sin espacios ni guiones), el sufijo **\_L33\_23** (elemento 23 de la lista 33) y la extensión .pdf.
- El archivo con el **texto para indexación** se denominará con el ISBN13 (sin espacios ni guiones), el sufijo **\_L33\_34** (elemento 34 de la lista 33) y la extensión .pdf.
- Para el archivo **MARC21-XML con los metadatos** se utilizará el siguiente convenio en el marco del proyecto ENCLAVE Editores - BNE (este archivo se genera automáticamente por DILVE en el momento del envío de metadatos a ENCLAVE):

`ISBNvd_ISBNvi_ENCLAVE Editores - BNE_vNN.xml`

dónde,

`ISBNvd` es el ISBN de la versión digital, descrita por los metadatos contenidos en el archivo.

`ISBNvi` es el ISBN de la correspondiente versión impresa

`ENCLAVE Editores - BNE` es un texto fijo

`vNN` indica la versión del formato del archivo. Empezará por v01

### Ejemplo:

Para el libro en versión digital identificado con el ISBN 9788412345678 y correspondiente a una obra impresa con ISBN 9788412348765<sup>5</sup>, se recibirán en el directorio FTP de la BNE los siguientes cuatro archivos:

- **9788412345678\_L38\_04\_n.jpg** → Imagen de cubierta
- **9788412345678\_L33\_23.pdf** → Archivo pdf para con el contenido parcial
- **9788412345678\_L33\_34.pdf** → Archivo pdf para indexación
- **9788412345678\_9788412348765\_ENCLAVE Editores - BNE\_v01.xml** → Metadatos + datos textuales "ricos"

<sup>5</sup> Estos ISBN no son reales, por lo que no cumplen las reglas de integridad marcadas para el dígito de control.

### 6.3 Formato del archivo de metadatos

Este será un archivo en formato MARC21 XML conteniendo exclusivamente los metadatos adicionales necesarios para complementar la ficha bibliográfica correspondiente a la obra impresa existente en la BNE.

A continuación se muestra un ejemplo de este archivo (se han introducido textos explicativos como contenidos de los campos, en lugar de sus valores reales):

```
<?xml version="1.0" encoding="UTF-8"?>
<collection xmlns="http://www.loc.gov/MARC21/slim"
xmlns:xsi="http://www.w3.org/2001/XMLSchema-instance"
xsi:schemaLocation="http://www.loc.gov/MARC21/slim
http://www.loc.gov/standards/marcxml/schema/MARC21slim.xsd">
  <record>
    <datafield tag="020" ind1=" " ind2=" ">
      <subfield code="a">ISBN versión digital</subfield>
    </datafield>
    <datafield tag="020" ind1=" " ind2=" ">
      <subfield code="a">ISBN versión impresa</subfield>
    </datafield>
    <datafield tag="505" ind1="0" ind2=" ">
      <subfield code="a">Sumario / Tabla de contenidos</subfield>
    </datafield>
    <datafield tag="520" ind1=" " ind2=" ">
      <subfield code="a">Resumen / Sinopsis</subfield>
    </datafield>

    <!-- El siguiente es un campo requerido, con contenido fijo -->
    <datafield tag="530" ind1=" " ind2=" ">
      <subfield code="a">Disponible también en sitio web</subfield>
    </datafield>

    <datafield tag="650" ind1=" " ind2="7">
      <subfield code="a">Materia 1</subfield>
      <subfield code="x">Materia 2</subfield>
      ...
    </datafield>
    <datafield tag="856" ind1="4" ind2="1">
      <subfield code="u">nombrefichero_contenido_parcial.pdf</subfield>
    </datafield>
  </record>
</collection>
```