



## Cargas de textos e imágenes mediante CSV

*v2 r01 01/2017*

**Enero 2017**



# Índice

<b>1</b>	<b>Introducción</b>	<b>1</b>
1.1	ONIX, DILVE y archivos CSV	1
1.2	Necesidad de estandarizar los nombres de archivos	1
<b>2</b>	<b>Campos de contenidos ricos</b>	<b>3</b>
2.1	Enumeración	3
2.1.1	Descripciones y otros textos de apoyo	3
2.1.2	Archivos de imagen, audio y vídeo	4
2.2	Campos CSV	4
2.2.1	Introducción	4
2.2.2	Nombres de los campos	4
2.2.3	Contenido de los campos CSV	5
2.2.4	Textos embebidos en campos CSV	6
<b>3</b>	<b>Gestión de los contenidos</b>	<b>8</b>
3.1	Carga física vs. Referencia (Archivos de imágenes)	8
3.1.1	Envío físico del archivo	9
3.1.2	Envío de referencias a los archivos	9
3.2	Archivos de imágenes	10
3.2.1	Formas de envío	10
3.2.2	Formato	11
3.2.3	Tamaños	11
3.3	Campos de texto	13
3.3.1	Contenido de los campos CSV	13
3.3.2	Formato de los contenidos textuales	14
<b>4</b>	<b>Nomenclatura de archivos</b>	<b>16</b>
4.1	Convenios	16
4.1.1	Convenio general:	16
4.1.2	Convenio simplificado, para archivos de imagen	17
<b>5</b>	<b>Envío de archivos CSV con archivos adjuntos</b>	<b>18</b>
5.1	Proceso	18
5.1.1	Paso 1: Agrupación de ficheros en una misma carpeta	18
5.1.2	Paso 2: Selección y empaquetado en .zip	19
5.1.3	Paso 3: Envío del archivo a DILVE	21
<b>6</b>	<b>Servicio de asistencia y soporte DILVE</b>	<b>22</b>

# 1 Introducción

## 1.1 ONIX, DILVE y archivos CSV

Al estar basada en ONIX 3.0, la estructura interna de DILVE permite la inclusión en la ficha del libro de referencias a una gran variedad de contenidos ricos: imágenes, textos, etc.

Estos contenidos se dividen en tres grandes bloques<sup>1</sup>:

- *Descripciones y otros textos de apoyo*, que se detallan en el grupo ONIX P-14. Las categorías de contenidos correspondientes se enumeran en la Lista ONIX 153.
- *Enlaces a recursos de apoyo (imagen, audio y vídeo)*, que se detallan en el grupo ONIX P-16. Las categorías de elementos de este grupo se enumeran en la Lista ONIX 158.
- *Contenido citado*, que se detalla en el grupo ONIX P-15. Las categorías de contenidos correspondientes se enumeran en la Lista ONIX 156.

Con el fin de permitir la gestión de contenidos ricos, para aquellas editoriales que no utilicen el formato ONIX, DILVE ha preparado un *subconjunto de campos* que permiten la carga y extracción de un elevado número de contenidos mediante archivos CSV.

Esos campos se enmarcan en los dos primeros bloques mencionados: *Descripciones y otros textos de apoyo*, y *Enlaces a recursos de apoyo*.

Los elementos del tercer bloque, *Contenido citado*, necesitan ONIX para su carga y requieren estar integrados en una estructura de campos jerarquizada por lo que sólo se pueden expresar, razonablemente, en formato XML-ONIX nativo.

En cualquier caso, es importante recordar, que todo elemento de contenido gestionado mediante campos en archivos CSV en DILVE, debe corresponder a algún elemento de contenido definido en ONIX, como es el caso de los expuestos en el presente documento.

En este documento se describen los campos de contenido que pueden cargarse o referenciarse mediante archivos CSV, así como los formatos requeridos por cada uno.

Este documento es un complemento del manual "Campos en archivos CSV", disponible en el web [www.dilve.es](http://www.dilve.es).

## 1.2 Necesidad de estandarizar los nombres de archivos

En general, la finalidad de los archivos cargados o referenciados en DILVE es su visualización en la pantalla de un ordenador (o dispositivos similares), fundamentalmente en entornos web. Esto entraña unas características determinadas en cuanto a su formato, y su resolución, tamaño, peso (en el caso de imágenes), etc.

Por otra parte, la explotación de los recursos de DILVE por parte de los distintos usuarios se realizará, en general, mediante métodos automáticos: descargas masivas o periódicas para alimentar otras bases de datos (p.ej. CEGAL en Red), generación automática de catálogos, etc.

---

<sup>1</sup> En esta relación se indican las Listas ONIX en las que se enumeran cada uno de los elementos de contenido, ya que estas tienen un papel importante en la creación de los nombres de los archivos correspondientes. Sin embargo, no es necesario conocer las listas. En este manual se dan todas las claves necesarias.

Dado que en DILVE coexisten muchos archivos correspondientes a distintos recursos de obras muy variadas, es fundamental *adoptar un convenio en la nomenclatura de los archivos* que simplifique su tratamiento automático, permitiendo saber, por ejemplo, de qué recurso se trata así como sus características: a qué obra corresponde, qué tamaño tiene, etc.

Efectivamente, una aplicación que realice una descarga masiva de datos de DILVE, se puede encontrar con varios miles de archivos de textos o imágenes. Sin un convenio estricto para los nombres de estos recursos, se pueden producir efectos no deseados:

- Si dos archivos distintos tienen el mismo nombre, uno de ellos "machacará" al otro. Por ejemplo, supongamos que las editoriales A y B utilizan, ambas, el nombre *cervantes.jpg* para las imágenes de las cubiertas de sus libros *Biografía de Cervantes* (Editorial A) y *Lectura crítica de los entremeses de Cervantes* (Editorial B). Si ambos libros figuran en una descarga de DILVE, sólo se puede enviar un archivo con el nombre *cervantes.jpg*, con lo que ambos libros aparecerán con la misma imagen de la cubierta. Para uno de ellos, evidentemente, esta imagen será incorrecta.
- Si en el nombre del archivo figuran caracteres no aceptables en el entorno informático que procesa las extracciones de DILVE, se producirá un error en la aplicación, resultando imposible procesar el archivo correspondiente, con lo cual se perderá la posibilidad de mostrar o procesar la imagen.

Es conveniente que, ante distintas versiones de un mismo archivo, se puedan determinar, de forma automática las características de cada uno.

En el Apartado 4 - *Nomenclatura de archivos* (pág. 16) se presenta el convenio para nombres de archivos recomendado por DILVE.

## 2 Campos de contenidos ricos

### 2.1 Enumeración

#### 2.1.1 Descripciones y otros textos de apoyo

Dado que toda la información gestionada en DILVE sigue el estándar ONIX 3.0, todos los campos cargados deben estar recogidos en este estándar. En el caso de descripciones y otros textos de apoyo los campos correspondientes se almacenan en el bloque P-14 del registro ONIX y su tipología está enumerada en la Lista de códigos ONIX número 153<sup>2</sup>.

De los tipos enumerados en esa lista estándar, DILVE permite actualmente la carga mediante archivos CSV, de los siguientes contenidos:

Contenido	Cód. lista ONIX 153
Resumen del libro / Descripción principal	03
Extracto del libro	14
Índice / Tabla de materias	04
Texto de la solapa o contracubierta	05
Comentario del autor	09
Reseña o cita de la reseña	06
Descripción para el lector	03
Descripción para las librerías	03
Descripción para los medios (prensa, etc.)	03
Titular promocional	10
Resumen breve	02

En la tabla se indica el código de la lista 153 de ONIX asignado a cada tipo de contenido, ya que este valor es necesario para crear el nombre del archivo correspondiente si el texto se envía o recibe en forma de archivo adjunto al archivo CSV. En el caso de las diferentes descripciones, la distinción de los destinatarios se establece en la etiqueta ContentAudience que se codifica con los valores de la lista 154 de códigos ONIX.

Para cada uno de los contenidos enumerados se puede indicar, además, el idioma en que está escrito.

El contenido de cada campo debe corresponder, lógicamente, al nombre del campo: no se debe incluir una *Reseña* en el campo *Índice/Tabla de materias*. DILVE no realiza ninguna comprobación en este sentido.

La finalidad última de estos campos textuales es, en general, su inclusión en entornos web. Por ello es conveniente que no sean excesivamente extensos y que su redacción permita una lectura fácil en dichos entornos. Lógicamente, para aquellos campos que se correspondan con una parte del libro (tabla de materias, primer capítulo), la extensión vendrá obligada.

<sup>2</sup> No es necesario conocer el funcionamiento interno de ONIX para gestionar estos contenidos mediante archivos CSV.

Aunque DILVE no impone ninguna norma respecto a la extensión y estilo del contenido de estos campos, es posible que alguna entidad usuaria de DILVE sí establezca determinados criterios para poder explotar estos contenidos en entornos concretos. En estos casos, esos conjuntos de criterios se harán públicos en DILVE para que las editoriales que quieran incorporar sus datos a dichas entidades puedan hacer los ajustes pertinentes en los datos cargados.

Por otra parte, la editorial que cargue cualquier contenido en DILVE debe asegurarse de que dispone de los permisos necesarios y que su publicación en DILVE, y su posterior descarga por parte de los usuarios de DILVE no infringe los derechos de autor o de propiedad intelectual.

### 2.1.2 Archivos de imagen, audio y vídeo

Por el momento, mediante archivos CSV, DILVE contempla la carga o referencia a la imagen de la cubierta, pudiéndose gestionar hasta dos archivos para cada cubierta (por ejemplo con distintos tamaños).

Contenido	Cód. lista ONIX 158
Imagen de la cubierta	01
Vídeo promocional o “book trailer” <sup>3</sup>	26

## 2.2 Campos CSV

### 2.2.1 Introducción

El proceso de envío a DILVE de archivos CSV con contenidos ricos se explica con detalle en el Apartado 5 – *Envío de archivos CSV con archivos adjuntos* (Pág. 18)

En este apartado se comentan las distintas estructuras que pueden tener estos archivos CSV. Como norma fundamental, hay que recordar que un archivo CSV es un archivo de texto, en el que cada línea corresponde a un libro y en el que la primera línea debe enumerar los nombres de los campos CSV contenidos en el archivo. Las líneas posteriores deben tener el mismo número de campos que los enumerados en la primera línea.

Estos archivos CSV se pueden generar de forma automática mediante volcados directos desde el sistema de gestión del catálogo de la editorial, si dicho sistema soporta esta opción, o mediante archivos Excel. Esta última opción es la más frecuente.

### 2.2.2 Nombres de los campos

Los nombres y características de todos los campos CSV utilizables en DILVE se recogen en la guía *Campos en archivos CSV*, que se puede descargar del web [www.dilve.es](http://www.dilve.es), y cuya lectura se recomienda. Dado que el repertorio de campos se amplía de forma permanente, es conveniente asegurarse de que se dispone de la última versión de dicha guía.

Como recordatorio, se enumeran a continuación algunos de los nombres de los campos CSV correspondientes a cada los contenidos expuestos en el presente documento:

Contenido	Campo CSV
Resumen del libro / Descripción principal	<code>texto_resumen</code>

<sup>3</sup> Los vídeos promocionales deben cargarse necesariamente como referencia a una URL.

Contenido	Campo CSV
Extracto del libro	<code>extracto</code>
Índice / Tabla de materias	<code>índice</code> ó <code>tabla_materias</code>
Primer capítulo	<code>primer_capitulo</code>
Texto de la solapa	<code>texto_solapa</code>
Texto de contracubierta	<code>texto_contracubierta</code>
Comentario del autor	<code>comentario_autor</code>
Reseña	<code>reseña</code>
Cita de reseña	<code>cita_de_reseña</code>
Descripción para el lector	<code>descripción_lector</code>
Descripción para las librerías	<code>descripción_librería</code>
Descripción para los medios (prensa, etc.)	<code>descripción_medios</code>
Titular promocional.	<code>titular_promocional</code>
Imagen de la cubierta	<code>imagen_cubierta</code>
Resumen breve	<code>resumen_breve</code>

#### NOTA

Los nombres de los campos pueden ir en cualquier combinación de mayúsculas o minúsculas, las tildes son opcionales. Al margen de esto, los nombres deben figurar exactamente como se indica en el cuadro anterior. En particular deben incluir los guiones bajos si estos figuran en el cuadro.

### 2.2.3 Contenido de los campos CSV

Los contenidos de los campos de contenido rico en un archivo CSV pueden ser de varios tipos.

- Para contenidos textuales (Descripciones y otros textos de apoyo) el campo correspondiente del archivo CSV que se envía a DILVE para su carga puede contener:
  - *El propio texto* (con la limitación que se indica en el apartado siguiente)
  - *El nombre de un archivo con el contenido textual* correspondiente (en formato texto o en formato pdf, según se detalla en el Apartado 3.3.2 *Formato de los contenidos textuales*. Pág 14 ). Este archivo, que incluye el texto, se deberá enviar a DILVE al mismo tiempo que el archivo CSV, según se indica en el Apartado 5 - *Envío de archivos CSV con archivos adjuntos* (Pág. 18).
- Para archivos gráficos (Imagen de la cubierta o vídeos promocionales) el campo CSV.
  - *El nombre de un archivo con la imagen* correspondiente (en formato jpg, según se detalla en el Apartado 3.2.2 *Formato*. Pág. 11), que se deberá enviar al mismo tiempo que el archivo CSV, según se indica en el Apartado 5 - *Envío de archivos CSV con archivos adjuntos* (Pág. 18).
  - *Una referencia a una ubicación externa (URL)* en la que se encuentra el archivo con la imagen correspondiente (debe enlazar al archivo de imagen no a una ficha del libro que contenga la imagen).

## 2.2.4 Textos embebidos en campos CSV

Si se envía a DILVE un texto embebido directamente en un campo de un archivo CSV, es importante tener en cuenta que dicho texto no puede incluir saltos de línea.

Efectivamente, en un archivo CSV cada línea debe corresponder a un mismo registro y debe tener el mismo número de campos (los indicados en la primera línea). Si uno de los valores incluidos en una línea incluye un carácter de salto de línea, se rompe la estructura del archivo y DILVE no podrá procesarlo.

La inclusión de saltos de línea en campos textuales en Excel suele producirse cuando se *copia y pega* el texto desde otra fuente.

Dependiendo del origen y tipo del archivo desde el que se hace la operación de *copiar y pegar*, el resultado en Excel puede variar:

- En Excel pueden verse *las líneas separadas*, con lo cual su localización es más sencilla<sup>4</sup>, o bien,
- Los caracteres de salto de línea *se pueden ver como "cuadraditos"*, sin que se produzca el salto de línea en el texto mostrado. En el caso de textos de gran longitud puede resultar difícil localizar estos caracteres.

En la figura siguiente se muestra textos con saltos de línea incluidos, en los resúmenes del segundo y del tercer libro. En el primer caso se trata de un salto de línea del primer tipo (el texto se ve en dos líneas en Excel), en el segundo caso el salto de línea se ve como un cuadradito (en Excel, se muestra el texto corrido).

	A	B	C	D	E
1	ISBN13	titulo	autor	resumen	precio
2	9788496479517	Venecia sin luz	Taipare, Jess	Una sola línea	24,00
3	9788496479524	Sombras venecianas	Taipare, Jess	Primera línea Segunda línea	24,00
4	9788496479500	FotoDigital	Taipare, Jess	Primera línea □ Segunda línea	30,00
5					
6					
7					
8					
9					
10					
11					

Estas dos líneas están separadas por un "salto de línea" no visible.

Este símbolo (cuadrado) indica que el texto contiene un "salto de línea".

Al pasar este Excel a formato CSV, el archivo de texto resultante es el siguiente:

```

ISBN13;      titulo;      autor;      resumen;      precio
9788496479517; Venecia sin luz;  Taipare, Jess; Una sola línea; 24,00
9788496479524; Sombras venecianas; Taipare, Jess; "Primera línea
Segunda línea";24,00
9788496479500; FotoDigital;  Taipare, Jess; "Primera línea
Segunda línea";30,00
    
```

<sup>4</sup> No deben confundirse estos saltos de línea con los que introduce Excel para visualizar un texto cuando la columna es excesivamente estrecha para mostrarlo en una sola línea. Esos saltos de línea introducidos por Excel a la hora de visualizar un texto no interfieren con la carga de textos en DILVE.

Aquí se ve que las líneas correspondientes al segundo y tercer libro se han transformado en dos líneas cada una que, lógicamente, no tienen el formato esperado en cuanto al número y tipo de campos. Si se intenta cargar este archivo en DILVE se generarán errores al procesar las líneas 2, 3, 4 y 5.

El resultado correcto debería ser:

ISBN13;	titulo;	autor;	resumen;	precio
9788496479517;	Venecia sin luz;	Taipare, Jess;	Una sola línea;	24,00
9788496479524;	Sombras venecianas;	Taipare, Jess;	Primera línea Segunda línea;	24,00
9788496479500;	FotoDigital;	Taipare, Jess;	Primera línea Segunda línea;	30,00

Normalmente el usuario no necesita ver los archivos CSV.

Sin embargo, dado que estos son archivos de texto normal se pueden abrir sin problemas con el "Bloc de notas", por ejemplo, o cualquier otro editor de textos, para intentar diagnosticar el origen de los problemas en el caso de que DILVE genere mensajes de error relativos al número o tipo de campos.

---

### IMPORTANTE

La conclusión de lo expuesto en este apartado es que, *mediante campos de archivos CSV, no se pueden cargar en DILVE textos que contengan saltos de líneas*. Por ello, en los casos en que el texto a enviar contenga saltos de línea se deberá enviar este en un archivo adjunto, según se indica en el Apartado 5 *Envío de archivos CSV con archivos adjuntos* (Pág. 18).

---

## 3 Gestión de los contenidos

### 3.1 Carga física vs. Referencia (Archivos de imágenes)

En repositorios de información sobre libros en venta, basados en ONIX, como es el caso de DILVE se pueden gestionar los contenidos ricos de dos formas:

- Cargando los archivos correspondientes en el repositorio
- Enviando la referencia a enlaces (URL) a los distintos contenidos que, en general, residirán en sitios webs controlados por las propias editoriales.

La práctica general (y recomendada por ONIX) es esta segunda opción: gestionar enlaces a los distintos contenidos referenciados en cada libro, residentes en sitios web controlados por las propias editoriales.

De esta forma, por una parte, cada editorial mantiene el control sobre sus propios recursos asegurándose siempre de que la versión ofrecida es la última y pudiendo controlar quién y cómo accede a estos recursos. Por otra parte, se reduce notablemente tanto la necesidad de almacenamiento en DILVE como el volumen del tráfico de datos.

En DILVE, se han habilitado ambas formas de gestión de contenidos *exclusivamente para archivos de imágenes*, por el momento.

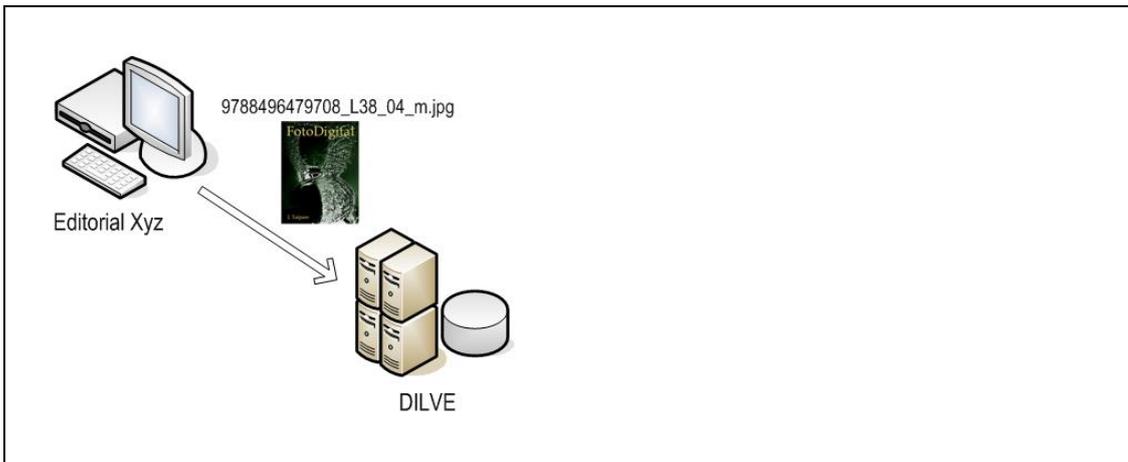
Así, por ejemplo, si la imagen de la cubierta de un determinado libro de la Editorial Xyz está en el archivo *9788496479708\_01\_m.jpg*, también accesible en el web de la editorial a través de la URL: [http://www.editorialxyz.com/recursos/cubiertas/9788496479708\\_01\\_m.jpg](http://www.editorialxyz.com/recursos/cubiertas/9788496479708_01_m.jpg), la editorial tiene dos opciones:

- *Enviar físicamente el archivo* 9788496479708\_01\_m.jpg a DILVE
- *Enviar únicamente la referencia al archivo*, mediante la URL [http://www.editorialxyz.com/recursos/cubiertas/9788496479708\\_01\\_m.jpg](http://www.editorialxyz.com/recursos/cubiertas/9788496479708_01_m.jpg)

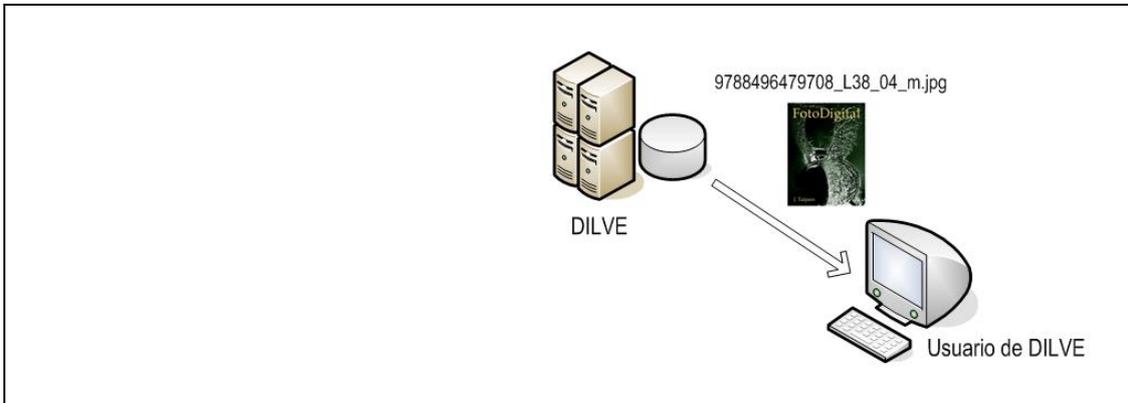
La primera forma de gestión de archivos de imágenes permite su inclusión en DILVE a aquellas editoriales que no dispongan de la capacidad de creación y gestión de sitios web para publicar dichos contenidos.

Los siguientes esquemas ilustran estas dos formas de gestionar los archivos de imágenes en DILVE.

### 3.1.1 Envío físico del archivo

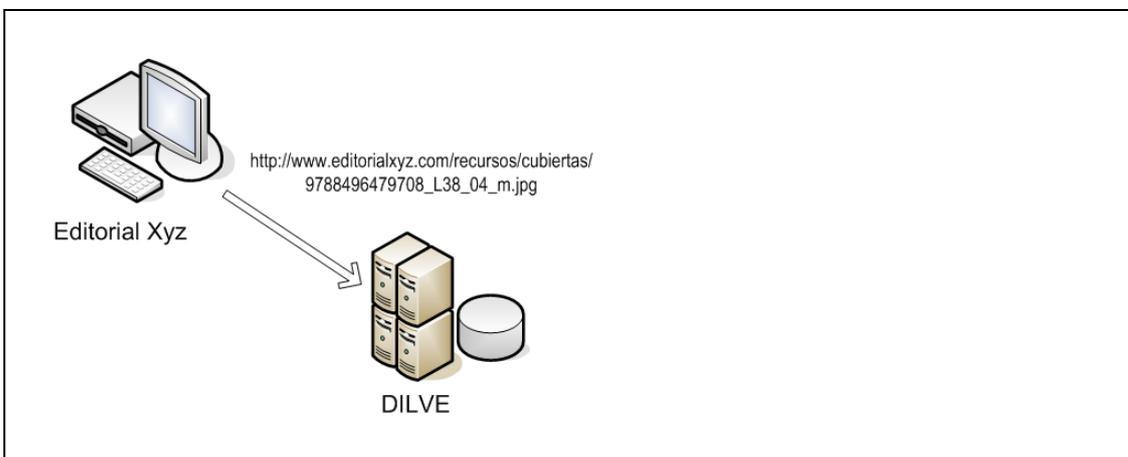


La editorial manda un archivo CSV a DILVE, enviando además una copia física de los archivos con las imágenes de las cubiertas.

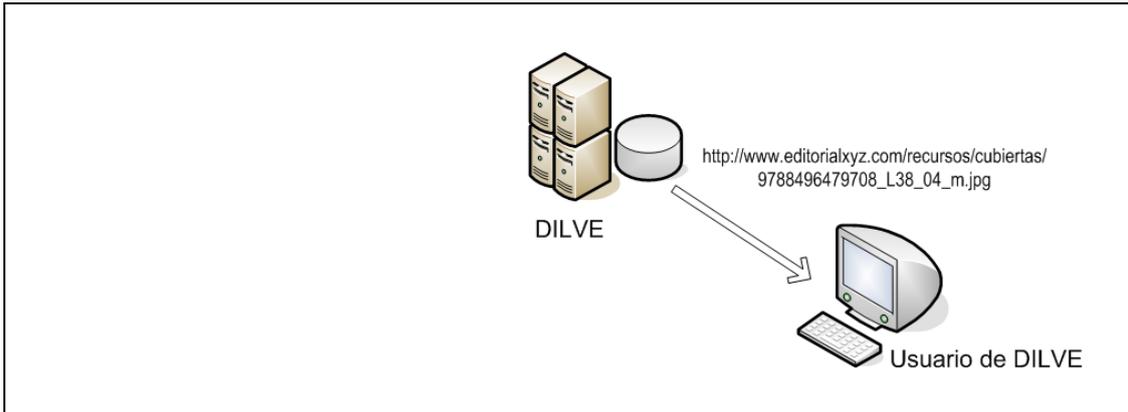


Cuando un usuario de DILVE extrae una ficha de un libro puede obtener una copia física del archivo con la imagen de la cubierta.

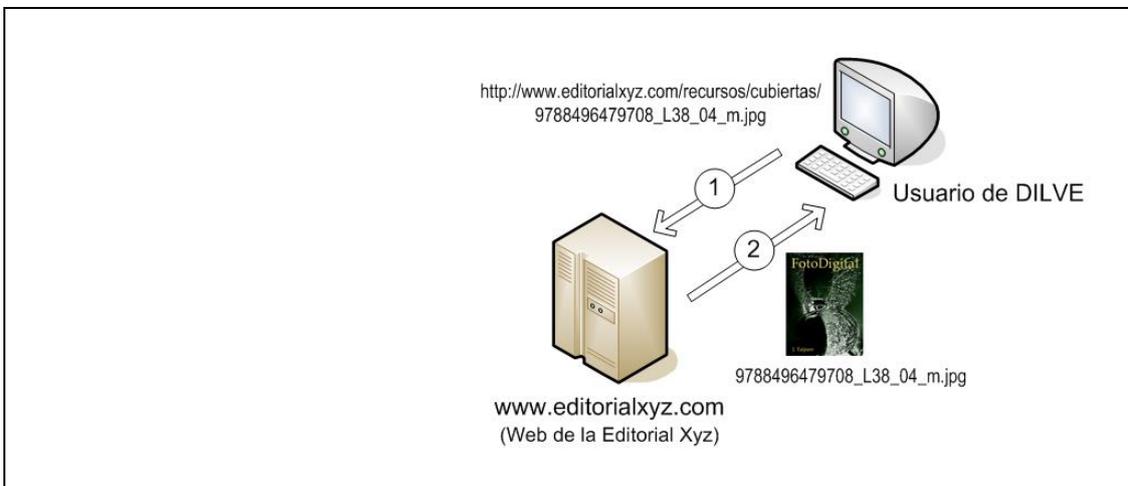
### 3.1.2 Envío de referencias a los archivos



La editorial manda un archivo CSV a DILVE, enviando las referencias (URL) de las ubicaciones de los archivos con las imágenes de las cubiertas.



Cuando un usuario extrae una ficha de un libro, obtiene la referencia de la ubicación del archivo con la imagen de la cubierta.



En cualquier momento, el usuario de DILVE puede acceder a las ubicaciones de los archivos para extraer su contenido.

## 3.2 Archivos de imágenes

### 3.2.1 Formas de envío

Según lo indicado en el Apartado 3.1, los archivos de imágenes se pueden gestionar en DILVE de dos formas:

- Mediante el envío del propio archivo a DILVE. En este caso, en el campo correspondiente del archivo CSV enviado a DILVE para su proceso se indicará el nombre del archivo con la imagen. Este archivo deberá enviarse a DILVE junto con el archivo CSV, de acuerdo con lo indicado en el Apartado 5 *Envío de archivos CSV con archivos adjuntos* (Pág. 18).

En el ejemplo siguiente, se emplea esta forma.

	A	B	C	D	E	F
1	<b>ISBN13</b>	<b>título</b>	<b>autor</b>	<b>imagen_cubierta</b>	<b>precio</b>	
2	9788496479517	Venecia sin luz	Taipare, Jess	9788496479517_04_l.jpg	24,00	
3	9788496479524	Sombras venecianas	Taipare, Jess	9788496479524_04_m.jpg	24,00	
4	9788496479500	FotoDigital	Taipare, Jess	9788496479500_04_l.jpg	30,00	
5						
6						

En este caso, al tiempo que se envía a DILVE el archivo CSV, deberán enviarse los tres archivos mencionados en la columna *imagen\_cubierta*.

- Mediante una referencia externa. En este caso, se supone que el archivo con la imagen se puede encontrar en una ubicación (URL) de la web, en general gestionada por la propia editorial. En el archivo CSV se indicará esta ubicación y no se envía ningún archivo.

	A	B	C	D	E	F
1	<b>ISBN13</b>	<b>título</b>	<b>autor</b>	<b>imagen_cubierta</b>	<b>precio</b>	
2	9788496479517	Venecia sin luz	Taipare, Jess	<a href="http://www.mieditorial.com/cubiertas/9788496479517_04_l.jpg">http://www.mieditorial.com/cubiertas/9788496479517_04_l.jpg</a>	24,00	
3	9788496479524	Sombras venecianas	Taipare, Jess	<a href="http://www.mieditorial.com/cubiertas/9788496479524_04_m.jpg">http://www.mieditorial.com/cubiertas/9788496479524_04_m.jpg</a>	24,00	
4	9788496479500	FotoDigital	Taipare, Jess	<a href="http://www.mieditorial.com/cubiertas/9788496479500_04_l.jpg">http://www.mieditorial.com/cubiertas/9788496479500_04_l.jpg</a>	30,00	
5						
6						

La editorial debe asegurarse de que las ubicaciones referenciadas (URL) son accesibles. DILVE no comprobará que las ubicaciones mencionadas son correctas ni guardará ninguna copia de los archivos correspondientes.

### 3.2.2 Formato

Los archivos de *imágenes* deberán estar en formato *JPEG* (con extensión *.jpg*), con la codificación de colores RGB.

Existe otro modo de codificación de colores en *JPEG* (*CMYK*) que se utiliza, fundamentalmente, para imágenes destinadas a imprenta y que puede dar problemas de compatibilidad con bastantes aplicaciones de gestión de imágenes en web. La codificación *CMYK* no debe utilizarse en DILVE.

También se admiten archivos en formato *png* (Portable Network Graphics).

No se admitirán imágenes en formato *gif*, ya que la gama de colores que admite este formato es mucho más reducida que el *jpg* y el *png* y, además, el formato *gif* utiliza un algoritmo de compresión de la información (Método *LZW*) patentado por *Unisys* y su utilización podría dar lugar a problemas de copyright.

### 3.2.3 Tamaños

En DILVE se admiten varios tamaños para la imagen de la cubierta, todos ellos definidos por el tamaño del *marco estándar* (ver más adelante) en el que mejor encaje la imagen correspondiente. El tamaño de la imagen quedará reflejado en el nombre del archivo que la contiene, mediante la *letra asociada* indicada en el cuadro siguiente, de acuerdo con el convenio indicado en el Apartado 4 *Nomenclatura de archivos* (Pág. 16).

Los tamaños de los *marcos estándar* son los siguientes:

Ancho x Alto (píxeles)	Nombre	Letra asociada	Peso aprox. en kBytes (**)
mayor de 1500x1500	(*)	<b>x</b>	
entre 1500x1500 y 675x675	Gigante	<b>g</b>	
entre 675x675 y 450x450	Huge (Inmenso)	<b>h</b>	200
entre 450x450 y 300x300	Large (Grande)	<b>l</b>	100
entre 300x300 y 200x200	Medium (Medio)	<b>m</b>	50
entre 200x200 y 135x135	Normal	<b>n</b>	25
entre 135x135 y 90x90	Small (Pequeño)	<b>s</b>	12
90x90 o menos	Thumbnail (Miniatura)	<b>t</b>	6

(\*) DILVE no admite imágenes cuyo archivo tenga un tamaño superior a 3 MB.

(\*\*) En la columna *Peso aprox.* se muestra un valor *normal* para cada tamaño. El peso de un archivo de imagen en formato jpg dependerá de su complejidad y del nivel de compresión utilizado. Un nivel de compresión excesivo dará lugar a archivos con poco peso pero que pueden tener una calidad insuficiente. Por el contrario, si se utiliza poca compresión el peso del archivo puede crecer bastante, sin una mejora apreciable en la calidad de la imagen final.

### Ejemplos de tamaños de imágenes:

En los siguientes ejemplos se muestran los distintos marcos y un rectángulo que representa a una imagen. Se trata de determinar en cada caso a qué marco se adapta mejor la imagen correspondiente (En los ejemplos, no se muestran los marcos "*Huge - Inmenso*" ni "*Thumbnail - Miniatura*").

#### Imagen de 190 x 260 píxeles

En la figura siguiente se ve que el marco en el que mejor encaja es el *Medio* de entre 300x300 y 200x200 píxeles.

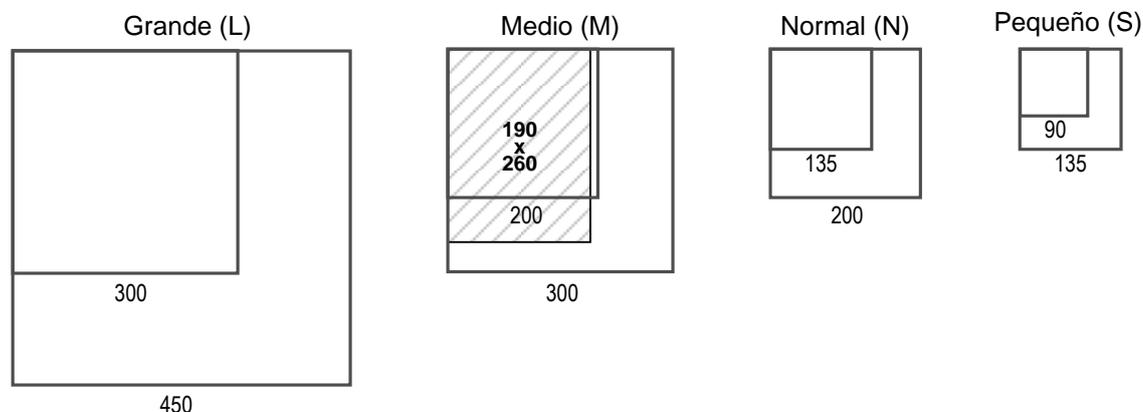
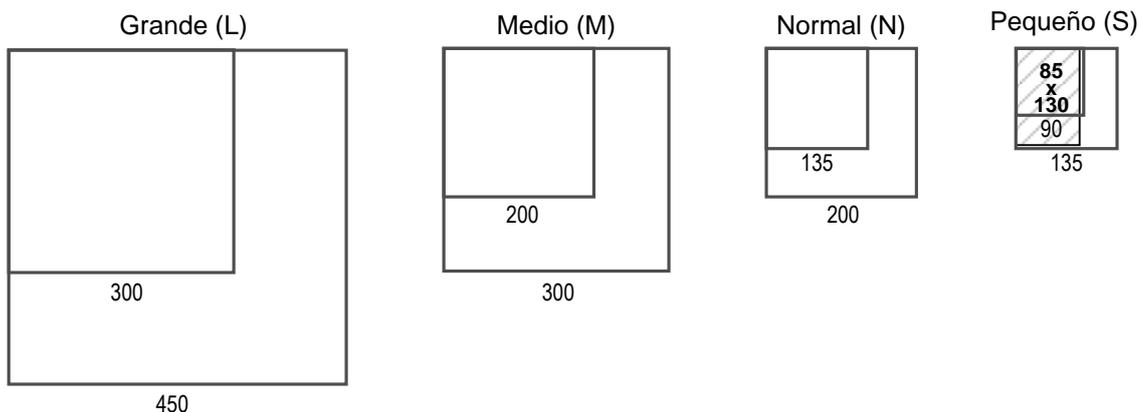
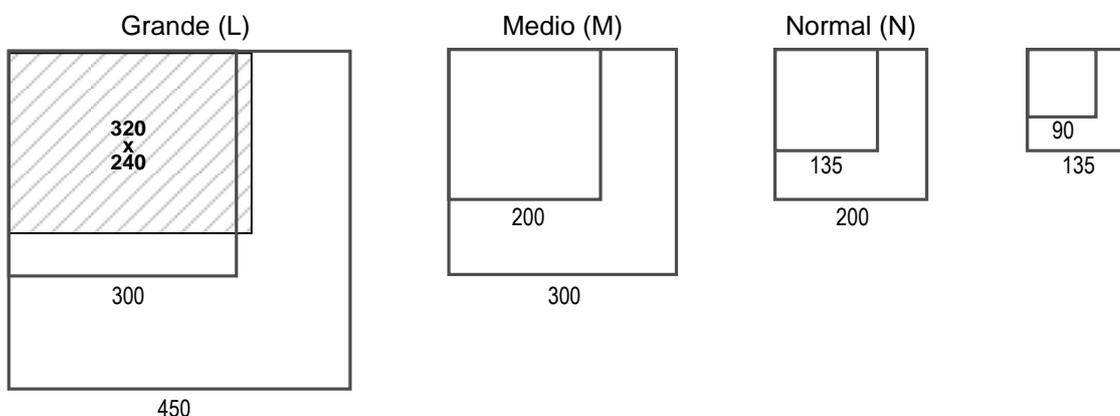


Imagen de 85 x 130 píxeles



En la figura se ve que encaja en el marco *Pequeño* de entre 135x135 y 90x90 píxeles.

Imagen apaisada de 320 x 240 píxeles



Encaja en el marco *Grande* de entre 450x450 y 300x300 píxeles.

---

**NOTA**

El tamaño del marco en el que encaja una imagen dada es una información que se incluye en el nombre del archivo correspondiente (véase el Apartado 4 *Nomenclatura de archivos* - Pág. 16).

---

### 3.3 Campos de texto

#### 3.3.1 Contenido de los campos CSV

Como se ha indicado en el Apartado 2.2.4 *Textos embebidos en campos CSV* (Pág. 6), los textos pueden incluirse directamente en el campo CSV correspondiente, si no contienen caracteres de salto de línea, o bien enviarse mediante archivos adjuntos.

En el ejemplo siguiente, los contenidos de los índices/tablas de materias se envían mediante archivos de texto adjuntos al archivo CSV, cuyos nombres figuran en los campos *tabla\_materias*.

Sin embargo, los textos de los resúmenes se incluyen en el propio archivo CSV, en el campo "resumen".

	Name Box	B	C	D
1	ISBN13	titulo	tabla_materias	resumen
2	9788496479517	Venecia sin luz	file://9788496479517_L33_01.txt	Este sería el resumen del libro 'Venecia sin luz'
3	9788496479524	Sombras venecianas	file://9788496479524_L33_01.txt	Aquí vendría el texto resumen de 'Sombras venecianas'
4	9788496479500	FotoDigital	file://9788496479500_L33_01.txt	Resumen del libro 'FotoDigital' ...
5				

### IMPORTANTE

Como se ve en la figura, para indicar que el contenido de un campo no es el propio texto, sino el nombre de un archivo adjunto, en el campo debe figurar el nombre del archivo precedido por los caracteres **file://** (es decir: letra "f", letra "i", letra "l", letra "e", dos puntos ":" y doble barra "/" / "). Estos caracteres no forman parte del nombre del archivo.

En el caso de que el texto vaya en el propio campo, se debe recordar que el texto no debe incluir saltos de línea (véase 2.2.4 *Textos embebidos en campos CSV* - Pág. 6)

### 3.3.2 Formato de los contenidos textuales

Los contenidos textuales deben estar en

- Formato texto (con posibilidad de incluir marcas de formato en XHTML)
- o bien en
- Formato PDF - *Portable Document Format* (con las especificaciones indicadas más adelante).

#### Formato texto

En este caso, los archivos deben contener *texto puro* ("plain text").

En general es el formato recomendado en DILVE para todos los campos de contenidos ya que los contextos en los que se mostrarán (o "explorarán") estos contenidos, requieren este formato.

Se pueden emplear marcas XHTML para dar formato al texto. Así, por ejemplo, el siguiente texto sería válido:

```
Este texto tiene una palabra <b>resaltada</b> y está en <br/> dos líneas.
```

Al mostrarse el texto, la palabra "resaltada" figurará en negrita y las palabras "dos líneas" aparecerán en una línea separada. Este tipo de textos, que no contienen caracteres de salto de línea explícitos sí pueden incluirse directamente en un campo CSV.

#### Formato PDF

La utilización de este formato sólo se recomienda para casos muy particulares, como pueda ser el extracto o el primer capítulo del libro, si no fuera posible enviar esta información en formato texto o XHTML.

En el caso del envío en formato PDF, no es posible incluir la información embebida en el campo CSV, por lo que siempre se enviará como archivo adjunto.

Como se expone en el Apartado 3.3.1 *Contenido de los campos CSV* (Pág. 13), el campo CSV debe contener el nombre del archivo, precedido por los caracteres `file://`. El archivo PDF se adjuntará, cuando se envíe el archivo CSV para su proceso en DILVE.

Cuando se envíen contenidos archivos en formato PDF deben tenerse en cuenta los factores siguientes:

- El archivo PDF *no debe ser el resultado directo del "escaneo" de un texto*.

En este caso, el archivo obtenido es un *archivo gráfico* en el que los textos están representados como imágenes y no son, por lo tanto, procesables en DILVE.

Si el texto se ha escaneado, el archivo PDF debe ser el resultado de un proceso de OCR - *Optical Character Recognition* (Reconocimiento óptico de caracteres) posterior al escaneo. El editor debe asegurarse de que el proceso de OCR ha reconocido correctamente todos los textos originales.

También se puede obtener un archivo PDF como transformación a este formato de un archivo preexistente, en formato textual.

- Es conveniente que el editor que vaya a utilizar el formato PDF para las cargas de contenidos, se asegure de que los usuarios de DILVE que vayan a utilizarlos puedan manejar ese formato.

Por ejemplo, *CEGAL en red* en sus consultas a DILVE requiere que el resumen del libro esté en formato texto, por lo que no podría procesar esta información si se ha suministrado en PDF.

Igualmente, los sistemas para enriquecimiento de catálogos en bibliotecas requieren, en general, campos de texto puro.

## 4 Nomenclatura de archivos

Los nombres empleados para los archivos de contenidos en DILVE deben permitir deducir la siguiente información:

- *A qué obra se refiere el archivo*
- *Qué tipo de información contiene:* imagen de cubierta, resumen, reseña, etc. En general, para indicar el tipo de información contenida, en el nombre del archivo se hará una referencia a la lista ONIX y el código correspondiente a su contenido.
- En el caso de imágenes, *tamaño aproximado de la imagen*, indicado mediante el *marco estándar* correspondiente
- *Tipo de codificación* interna del archivo (texto, jpeg, png, pdf, etc)
- *Número de orden*, para el caso en que exista más de un archivo de características similares

Por otra parte, la tipología y número de archivos de distinto tipo que se podrán cargar en DILVE irá aumentando. Los convenios de nomenclatura de archivos asociados a una obra, propuestos en este apartado, tienen en cuenta estas necesidades de crecimiento futuro.

Igualmente, dado que los usuarios finales de DILVE procesarán los archivos en una variada gama de entornos y aplicaciones informáticos, es fundamental evitar, en los nombres correspondientes, caracteres que puedan crear problemas en determinados entornos informáticos.

### 4.1 Convenios

Para los archivos con contenidos enviados a DILVE se establecen los siguientes tipos de nombre en DILVE:

#### 4.1.1 Convenio general:

Los nombres de archivos seguirán el siguiente convenio general:

`nnnnnnnnnnnnnn_cc_t_nn.ext`

donde:

- |                                   |  |
|-----------------------------------|--|
| <code>nnnnnnnnnnnnnnnnnnnn</code> | ISBN10 o el ISBN13 de la publicación. <u>En todo caso, sin guiones, espacios en blanco, ni otros separadores.</u>  |
| <code>cc</code>                   | Código de la lista ONIX (Lista 153 o 158), que indica el elemento del libro al que se refiere el archivo. Estos códigos se enumeran en el cuadro de la página 3 de este documento. |
| <code>t</code>                    | <u>Esta partícula es obligatoria 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), de acuerdo con los convenios establecidos en el apartado "Tamaños" de este documento.

- nn** Indica un número de secuencia, en el caso de que haya más de un archivo, del mismo tipo y con el mismo nombre. Opcional.
- ext** Es la extensión del archivo. Indica el formato del mismo: *txt*=formato texto, *jpg* = formato JPEG, *png* = Portable Network Graphics, *pdf* = portable document format. No todos los formatos son aceptables para todos los tipos de contenidos.

#### 4.1.2 Convenio simplificado, para archivos de imagen

En esta opción, solo aplicable a archivos de imagen, no se indica explícitamente la lista ONIX a la que corresponde el tipo de archivo ya que se supondrá, por defecto, que el archivo corresponde a un elemento de la lista 158 - *Tipo de archivo de imagen/audio/vídeo*.

El formato general de nombres de archivos de imagen es el siguiente:

*nnnnnnnnnnnnnn\_cc\_t.ext*

donde:

- nnnnnnnnnnnnnn** ISBN10 o el ISBN13 de la publicación. En todo caso, sin guiones, espacios en blanco, ni otros separadores.
- cc** Código de la lista ONIX 158, que indica el elemento del libro al que se refiere el archivo. Debe ser 01 para una imagen de la cubierta.
- t** Es una letra, de uso convencional en DILVE, que indica el tamaño aproximado de la imagen (en píxeles), de acuerdo con los convenios establecidos en el apartado "Tamaños" de este documento.
- ext** Es la extensión del archivo. Indica el formato del mismo: *jpg* = formato JPEG, *png* = Portable Network Graphics.

#### Ejemplos

- 9788496479517\_01\_n.jpg* Imagen de la cubierta del libro 978-84-96479-51-7, de tamaño aproximado de entre 200x200 y 135x135 píxeles ("n") en formato JPEG.
- 9788496479517\_03.txt* Resumen del mismo libro, en formato texto.
- 9788496479517\_03\_02.txt* Otro archivo con un resumen del mismo libro, en formato texto.
- 9788496479517\_14.txt* Extracto del mismo libro, en formato texto.
- 9788496479517\_04.txt* Tabla de materias (índice) del mismo libro, en formato texto.
- 9788496479517\_14.pdf* Primer capítulo del mismo libro, en formato pdf.

## 5 Envío de archivos CSV con archivos adjuntos

Cuando en un archivo CSV se hace referencia a archivos externos, con imágenes u otros contenidos, se deben enviar simultáneamente todos los archivos referenciados junto con el archivo CSV, para su proceso en DILVE.

Para ello, tanto el archivo CSV como los distintos archivos referenciados deben "empaquetarse" en un archivo único en formato .zip.

### 5.1 Proceso

Supongamos que se tiene un archivo CSV (que llamaremos *ejemplo\_carga\_v01.csv*) con el contenido siguiente:

	A	B	C	D
1	<b>ISBN13</b>	<b>título</b>	<b>tabla_materias</b>	<b>imagen_cubierta</b>
2	9788496479517	Venecia sin luz	file://9788496479517_04.txt	http://www.mieditorial.com/cubiertas/9788496479517_01_g.jpg
3	9788496479524	Sombras venecianas	file://9788496479524_04.txt	file://9788496479524_01_n.jpg
4	9788496479500	FotoDigital	file://9788496479500_04.txt	file://9788496479500_01_m.jpg

En este archivo CSV se hace referencia a los siguientes archivos:

- Índices / Tablas de materias
  - 9788496479517\_04.txt
  - 9788496479524\_04.txt
  - 9788496479500\_04.txt

Nótese que el prefijo "file://" no forma parte del nombre

- Imágenes de cubierta:
  - 9788496479524\_01\_n.jpg
  - 9788496479500\_01\_m.jpg

Nótese que, en el caso de archivo de imágenes de cubierta no se debe incluir el prefijo "file://"

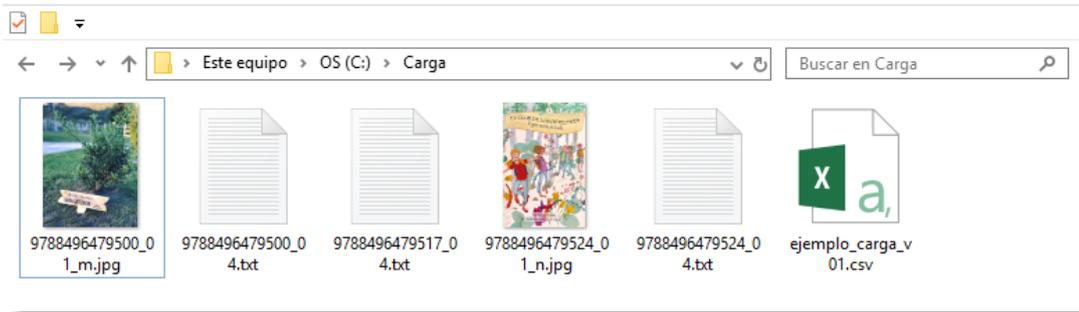
La imagen de la cubierta del primer libro, *Venecia sin luz*, se referencia mediante su URL en el web de la editorial, por lo que el archivo correspondiente no se envía a DILVE.

#### 5.1.1 Paso 1: Agrupación de ficheros en una misma carpeta

En el presente envío a DILVE van a intervenir los seis archivos siguientes:

Archivo	Contenido
ejemplo_carga_v01.csv	Archivo CSV
9788496479517_04.txt	Índice de Venecia sin luz
9788496479524_04.txt	Índice de Sombras venecianas
9788496479500_04.txt	Índice de FotoDigital
9788496479524_01_n.jpg	Imagen de la cubierta de Sombras venecianas (tamaño "normal")
9788496479500_01_m.jpg	Imagen de la cubierta de FotoDigital (tamaño "medio")

La primera operación es reunir a estos archivos en una misma carpeta.

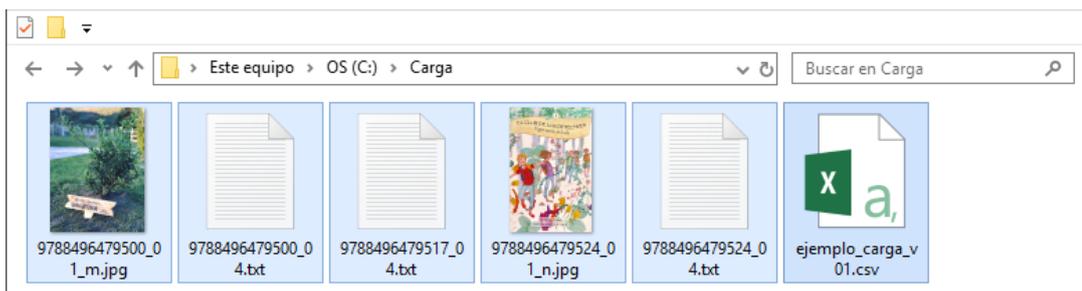


*Archivos reunidos en una misma carpeta*

Si bien esta carpeta puede contener más archivos, es conveniente (hasta que se domine el proceso) que sólo incluya a los archivos que van a formar parte del envío.

### 5.1.2 Paso 2: Selección y empaquetado en .zip

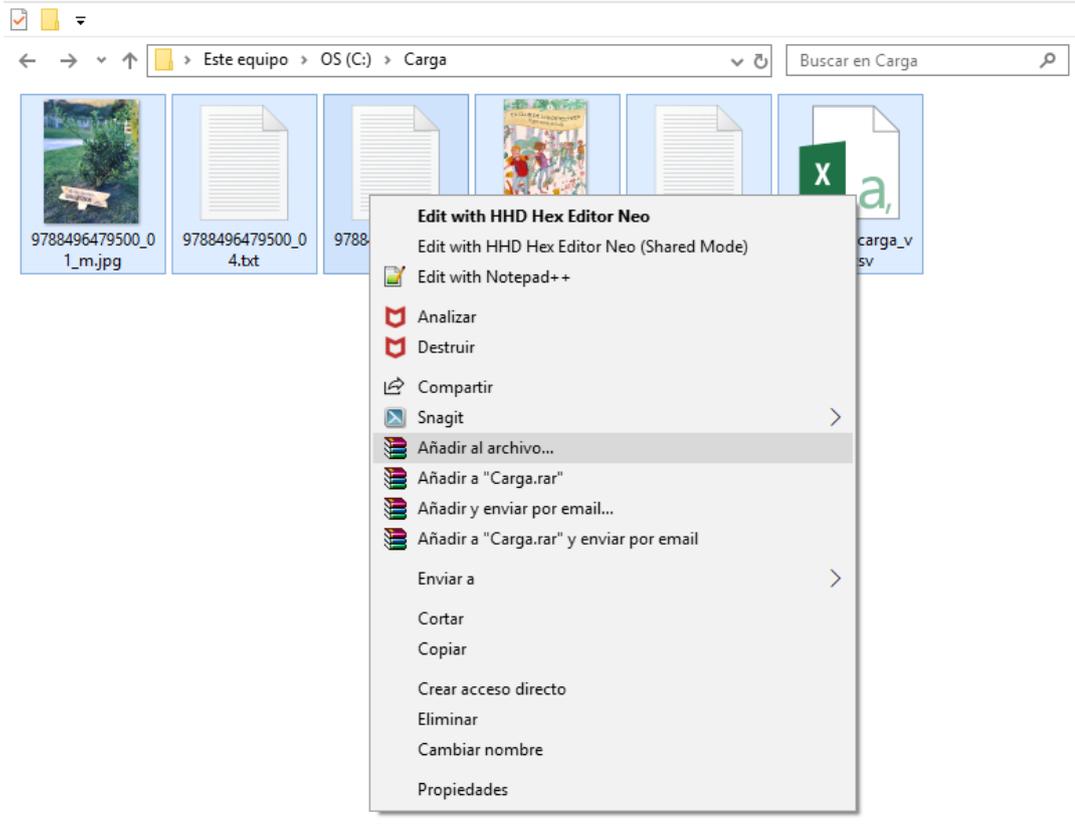
En este paso se seleccionan los archivos que intervienen en el envío, y se empaquetan en un único archivo ZIP.



*Archivos seleccionados*

El proceso de empaquetado dependerá del programa que se utilice para tal fin. Existe una gran variedad de programas de empaquetado/compresión aunque los más usuales son: *WinZip* y *WinRAR*.

Normalmente estos programas de empaquetado y compresión están accesibles pulsando el botón derecho del ratón sobre los archivos seleccionados.



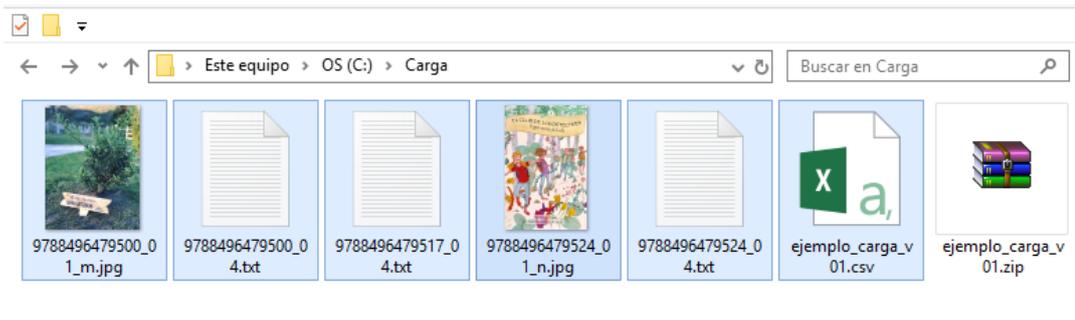
Selección del programa WinZip con el botón derecho

### IMPORTANTE

El programa *WinZip* siempre genera archivos en formato *ZIP*. sin embargo, el programa *WinRAR* permite crear archivos en formato *RAR* o en formato *ZIP*. En este caso es importante elegir el formato *ZIP* ya que *DILVE* no procesa el formato *RAR*.

Se puede elegir cualquier nombre (con las limitaciones para nombres de archivos de *DILVE*) para el archivo resultado del empaquetado de los seis archivos que se han seleccionado. Hay que tener en cuenta *que ese archivo es el que se va a enviar a DILVE para su proceso, por lo que su nombre no puede incluir caracteres especiales, acentos, espacios en blanco, etc.*

En el ejemplo se ha elegido el nombre *ejemplo\_carga\_v01.zip* para el archivo empaquetado.

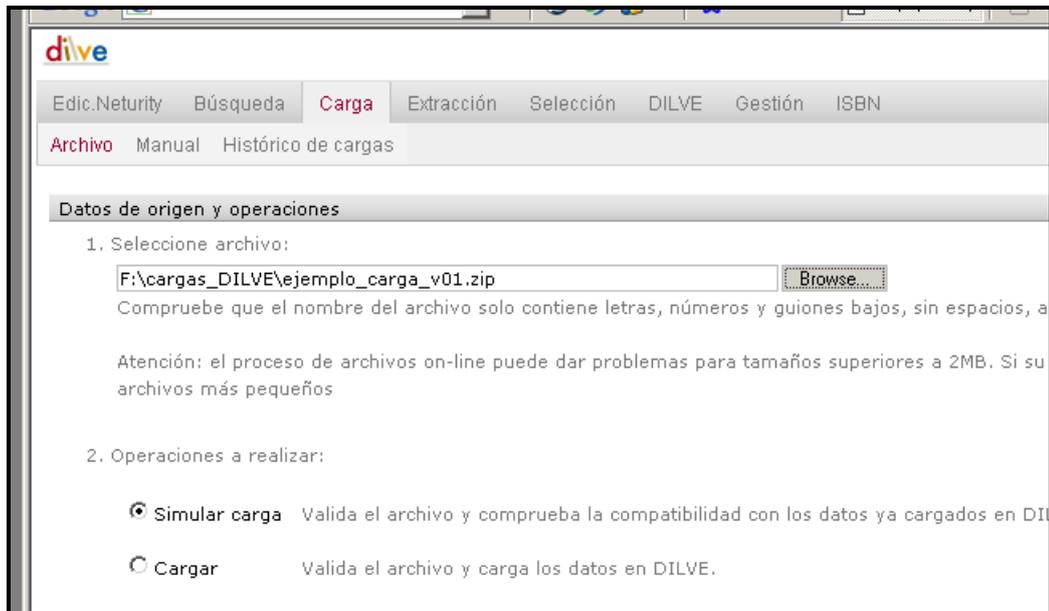


### Archivo ZIP ya creado

Este archivo, *ejemplo\_carga\_v01.zip*, es el que se debe cargar en DILVE.

### 5.1.3 Paso 3: Envío del archivo a DILVE

Este proceso es el mismo que para una archivo CSV normal. DILVE reconoce que se trata de un archivo empaquetado en ZIP y lo procesa convenientemente.



*Archivo ZIP seleccionado para su carga en DILVE*

El resto del proceso es el habitual. DILVE procesará el archivo CSV y generará los correspondientes informes de carga.

## 6 Servicio de asistencia y soporte DILVE

Para cualquier duda o consulta, recuerde que tiene a su disposición el *Servicio de Asistencia DILVE*.

Teléfono: **914 136 321**

e-mail: **asistencia@dilve.es**