Revisión de la calidad de los metadatos

Información completa y en plazo 2

¿Están cargados todos los libros? 2
Se puede buscar… 2
Se puede comparar ISBN <> DILVE… 3
Comprobar que todos están visibles 5

¿Está completa la información? 5
Comprobaciones automáticas 5
Revisar las autorías 6
Extracciones CSV a medida 6

Actualizar datos que cambian 9
Otras ayudas 10

Anexo I. Extracciones CSV de selección de ISBN 10

Anexo II. Gestión de CSV con Excel 12

Este documento describe en detalle los principales procedimientos que pueden utilizar las editoriales de DILVE para comprobar la calidad de sus metadatos y la forma de corregir o completar la información en caso de que sea necesario.

La totalidad de la metodología será de utilidad para las editoriales que hagan cargas de datos manuales o mediante archivos CSV preparados manualmente.

Para las editoriales cuyos sistemas hagan cargas automáticas a DILVE (tanto en CSV como en ONIX) los procesos descritos serán de utilidad en la parte de identificación de datos a corregir/completar pero las modificaciones deberán hacerse directamente en el sistema de la editorial para asegurarse de que los datos se completan en origen y que se mantendrán completos en futuras modificaciones.

El objetivo de los procedimientos descritos consiste en asegurarse de que:

- **están en DILVE todos los libros** que tienen que estar. Esto se consigue comprobando, por fecha, si están todas las novedades, comparando con la información que hay en la Agencia del ISBN por si alguno de los dos sistemas está incompleto o comprobando que todos los libros dados de alta en DILVE están visibles para las entidades que necesitan la información (librerías, distribuidoras, etc.),
- **todos los libros** que están activos en el mercado **están completos y con datos ricos**. Esto se consigue utilizando las comprobaciones automáticas que ofrece DILVE o haciendo extracciones CSV de metadatos concretos que permitan comprobar si falta información en alguno de los libros,
- la **información** está **perfectamente actualizada** (cambios de precios, de disponibilidad o situación en catálogo, reimpresiones, etc.).

Hay algunas herramientas específicas para esto, pero lo más importante es incorporar en los procedimientos internos de trabajo de la editorial el mantenimiento al día de los datos que salen a la cadena comercial a través de DILVE.

Información completa y en plazo

Es importante que la información cargada en DILVE esté completa, tanto en el número de libros como en los metadatos de cada uno de ellos.

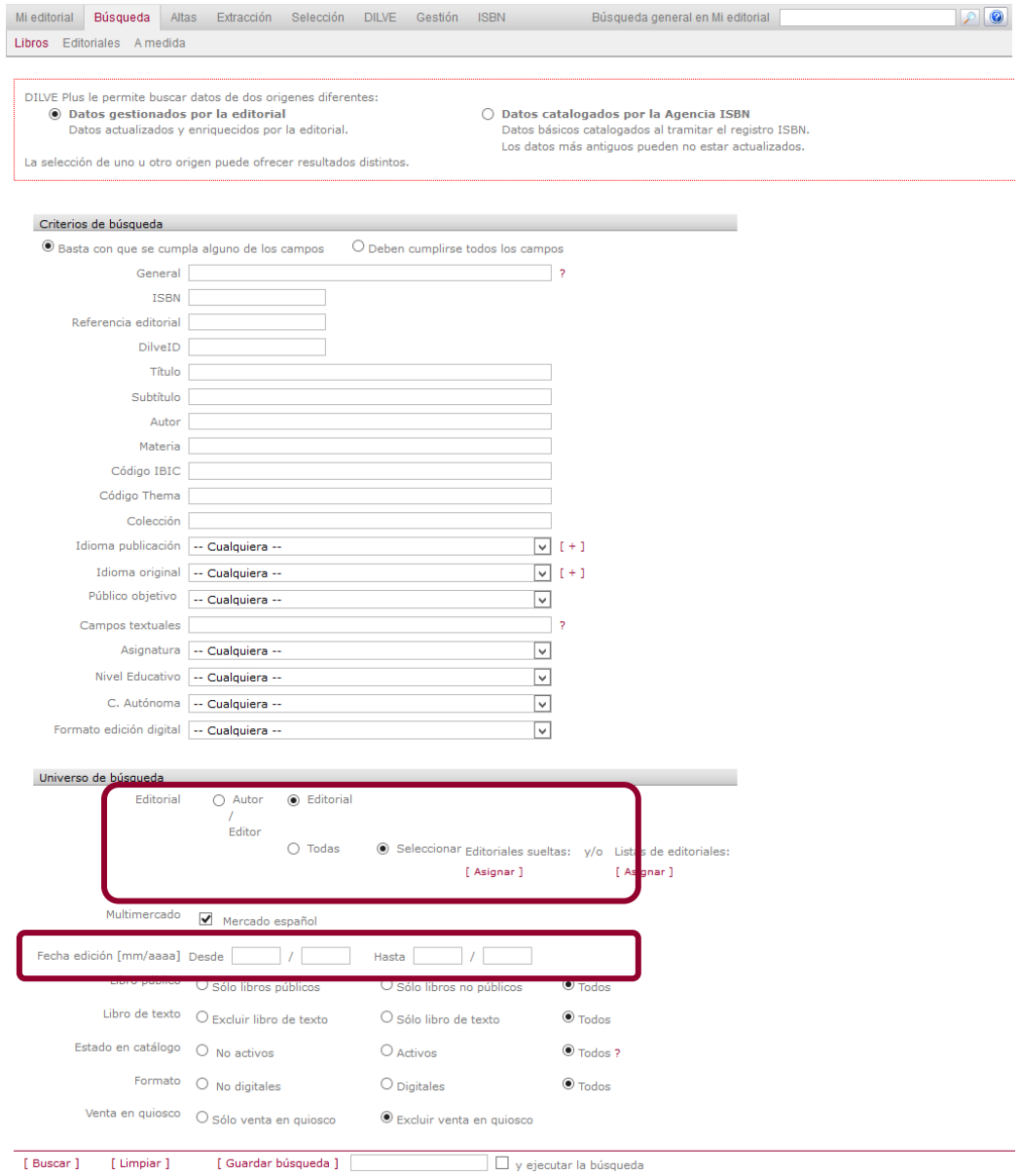
También es fundamental que la información esté disponible en DILVE a tiempo. En general, la cadena comercial necesita conocer los metadatos, bastante antes de que el libro esté en venta.

El sector de libro en España acordó que los metadatos de las novedades deberían estar disponibles en la cadena comercial al menos **1 mes antes** de su puesta a la venta.

¿Están cargados todos los libros?

Se puede buscar…

La sección de búsquedas permite buscar todos los libros de la editorial seleccionando por mes y año de edición, lo que permite comprobar, con relativa facilidad, si se han cargado en DILVE las últimas novedades:



Es importante comprobar por separado los libros públicos y no públicos ya que estos últimos son "invisibles" para todos los usuarios de DILVE, excepto los de la propia editorial.

El sector de libro en España acordó que los metadatos de las novedades deberían estar disponibles en la cadena comercial al menos **1 mes antes** de su puesta a la venta. Esto supone que estén visibles en DILVE y registrados y catalogados en la Agencia de ISBN.

Se puede comparar ISBN <> DILVE…

Para comprobar si está cargado en DILVE el catálogo completo se puede utilizar el método de comparación de la información que hay en DILVE y la que está catalogada en la Agencia del ISBN.

No es necesario cargar en DILVE absolutamente todos los libros que la editorial haya publicado a lo largo de toda su historia, pero es importante asegurarse de que están en DILVE aquellos que todavía están activos en el mercado así como aquellos que hayan dejado de estarlo recientemente y de cuya descatalogación pueda ser necesario informar a la cadena comercial.

La sección Mi editorial > DILVE<>Agencia: comparar datos permite obtener los libros catalogados en la Agencia del ISBN pero que no se han dado de alta en DILVE. Mediante esta funcionalidad se obtiene un archivo CSV que se pueden editar con Excel y completar para cargar en DILVE todos los libros en una única operación.

1. Pestaña para la consulta de los libros catalogados en la Agencia del ISBN pero no cargados en DILVE.
2. Criterios de búsqueda para acotar los resultados. Es obligatorio seleccionar al menos un prefijo.
3. Seleccionar los libros que se van a cargar en DILVE. Se pueden seleccionar de uno en uno o con las opciones de Seleccionar y Desmarcar todos en la página o en todo el listado.
4. Pulsando el botón de *Descargar*, se ejecuta el proceso que genera un CSV con la información catalogada en la Agencia del ISBN de los libros seleccionados. Unos minutos después se recibirá el archivo CSV en la dirección de correo electrónico asociada al usuario conectado a DILVE.

Una vez recibido el correo electrónico se deberá extraer de la carpeta comprimida el archivo CSV y editarlo con Excel siguiendo los pasos que se describen en el Anexo II.

Se obtendrá un Excel con los metadatos que figuran como catalogados en la Agencia del ISBN. Muchos de ellos pueden estar vacíos. Es

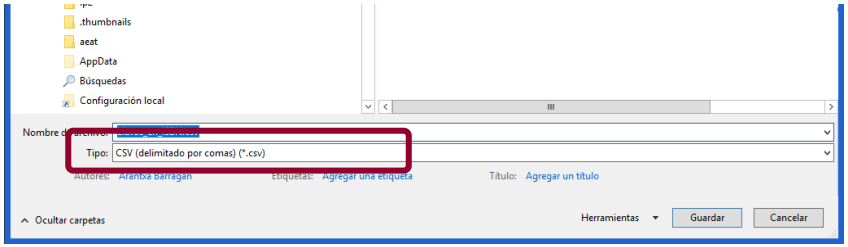
recomendable revisar y completar la información que no figure en el Excel y añadir los campos con la información comercial que no se cataloga en la Agencia del ISBN y que es fundamental para la cadena. Para incluir esta información se deben añadir al menos las siguientes columnas:

- `situ_catalogo_editorial`: Situación del libro en el catálogo de la editorial. Un valor de la lista 64 de códigos ONIX.
- `disponibilidad`: disponibilidad del libro en el mercado español. Un valor de la lista 65 de códigos ONIX.
- `precio_con_iva`: número positivo, mayor de cero y con la coma como separador de decimales.
- `iva`: número entero positivo. Actualmente sólo se admiten los valores 4 y 21.

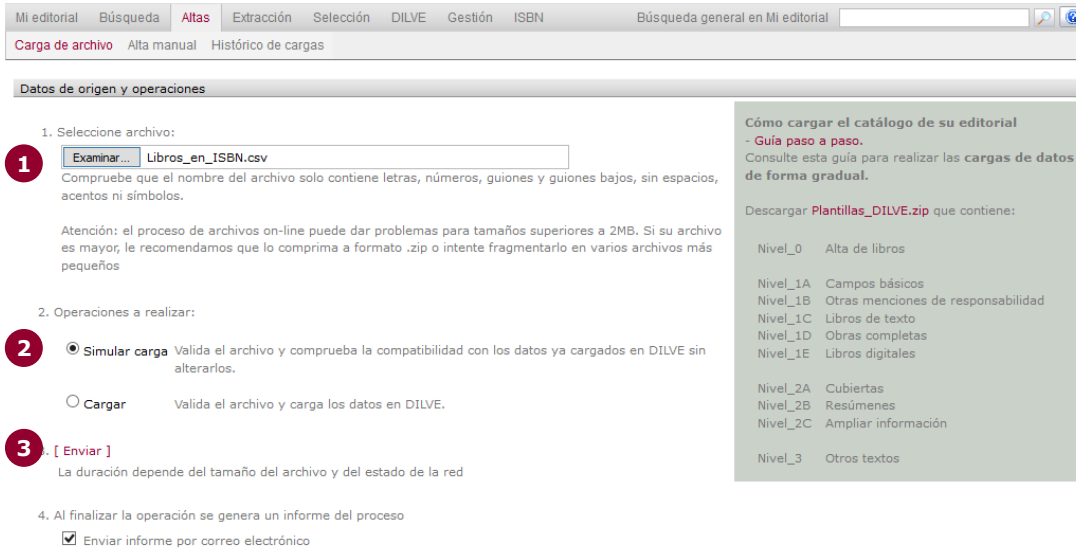
Tanto para completar la información del Excel recibido como para rellenar los campos comerciales es fundamental consultar el formato que debe de tener cada uno de los campos en el Manual de campos CSV.

Si no se indica lo contrario en el propio archivo, los libros cargados con un archivo CSV quedarán en estado no público (no visibles para los consumidores de información de DILVE). Para cargar los libros directamente como públicos se debe añadir el campo `libro_publico` en el Excel con el valor 1 en todas las filas.

Una vez completados los datos en el Excel, deberá guardarse como CSV ("CSV para Windows" si se trabaja con MAC):



El CSV guardado deberá cargarse en la sección Alta > Carga de archivo de DILVE.



1. Seleccionar el archivo CSV.
2. Se puede *simular* la carga para comprobar si la estructura del archivo es correcta pero, para actualizar realmente los datos en DILVE, debe *cargarse* el archivo.
3. Enviar el archivo para que se procese.

Si el archivo es muy grande se puede comprimir en ZIP y cargar la carpeta comprimida que tardará menos en subir. En caso de que supere el tamaño máximo permitido para las cargas desde la web (2 MB) se pueden subir el archivo de carga conectándose por FTP a ftp.dilve.es y copiando el archivo en la carpeta IN. DILVE procesará de forma automática el archivo y enviará el resultado de la carga al correo electrónico asociado al usuario que ha hecho la carga.

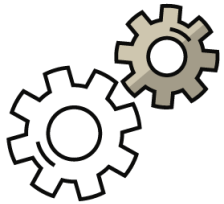
Comprobar que todos están visibles

Es importante comprobar que los libros que hay en la sección Mi editorial > Libros no públicos son realmente lo que deben ser invisibles para las librerías y distribuidoras que utilizan DILVE. Si son registros que deberían aparecer en el catálogo de la editorial deben hacerse públicos.

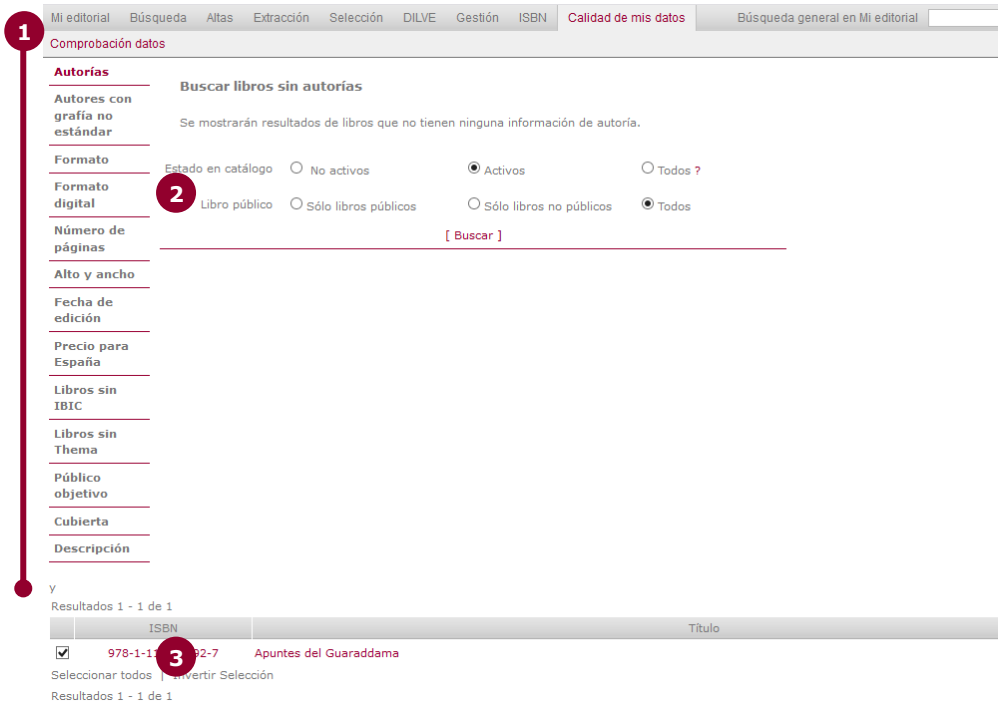
Este proceso (de "publicación" de la información) se puede hacer de tres formas: libro por libro en el formulario, con una carga CSV (campo `libro_publico`) o desde el propio listado seleccionando los libros que haya que hacer públicos y utilizando la opción de Publicar seleccionados, en la parte inferior del listado de libros no públicos.

¿Está completa la información?

Comprobaciones automáticas



Las editoriales en DILVE tienen, en la sección Calidad de mis datos herramientas que permiten hacer búsquedas de libros a los que les falte algún metadato concreto:



1. Opciones de comprobaciones específicas de libros:
   - sin autorías
   - con autores con grafía aparentemente no estándar
   - sin información correcta de formato (soporte-encuadernación)
   - digitales sin información del formato (EPUB, Kindle, PDF, etc.)
   - impresos sin información del número de páginas
   - impresos sin información de alto o ancho
   - sin fecha de publicación
   - en venta en España sin información de precio
   - sin materias IBIC
   - sin materia Thema.
   - sin información de público objetivo
   - sin imagen de cubierta
   - sin descripción o sinopsis.

2. Filtros para establecer el universo de libros a comprobar.
3. Resultado de la búsqueda según los criterios. Como en todos los listados de libros de DILVE, se permite la selección de los registros para luego hacer una extracción de la lista de ISBN. En el Anexo I de este documento se detalla el proceso para extraer datos a partir de una selección de ISBN.

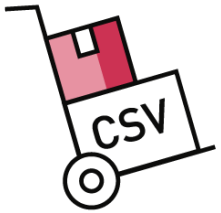
Revisar las autorías

Las editoriales que no tengan un número excesivo de libros o de autores distintos, en la sección Mi editorial>Autorías pueden comprobar los nombres de los autores (únicamente el perfil de "autor del texto") que tiene asignados a sus libros y comprobar, a primera vista, si hay alguno erróneo o más de un nombre asignado a un único bloque de autoría o más de uno que corresponda a la misma persona pero escrito de formas distintas, o que tenga información del tipo de responsabilidad como parte del nombre, etc.

Haciendo clic sobre un nombre se accede al listado de libros de la editorial que tiene asignado ese autor y es posible seleccionar los libros para hacer una extracción o editar cada registro para corregir los datos.

Para la comprobación de nombres de todos los tipos de responsables se pueden utilizar también la opción Autores con grafía no estándar de la sección Calidad de mis datos.

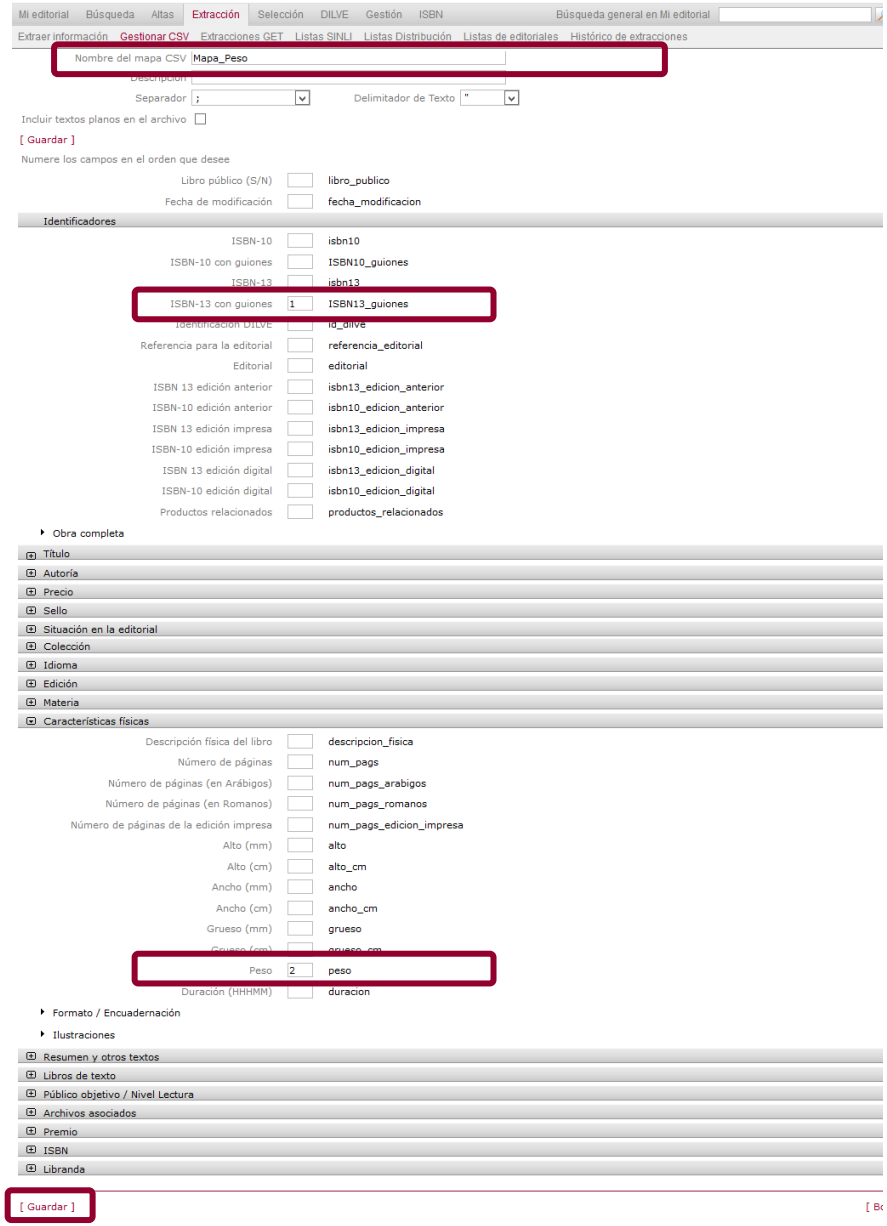
Extracciones CSV a medida



Para comprobar el estado de algún metadato concreto (que no esté incluido en las búsquedas específicas de la sección Calidad de mis datos), para todos los libros del catálogo de la editorial, es necesario hacer una extracción CSV de esa información para todos los libros y comprobar en Excel qué libros están incompletos.

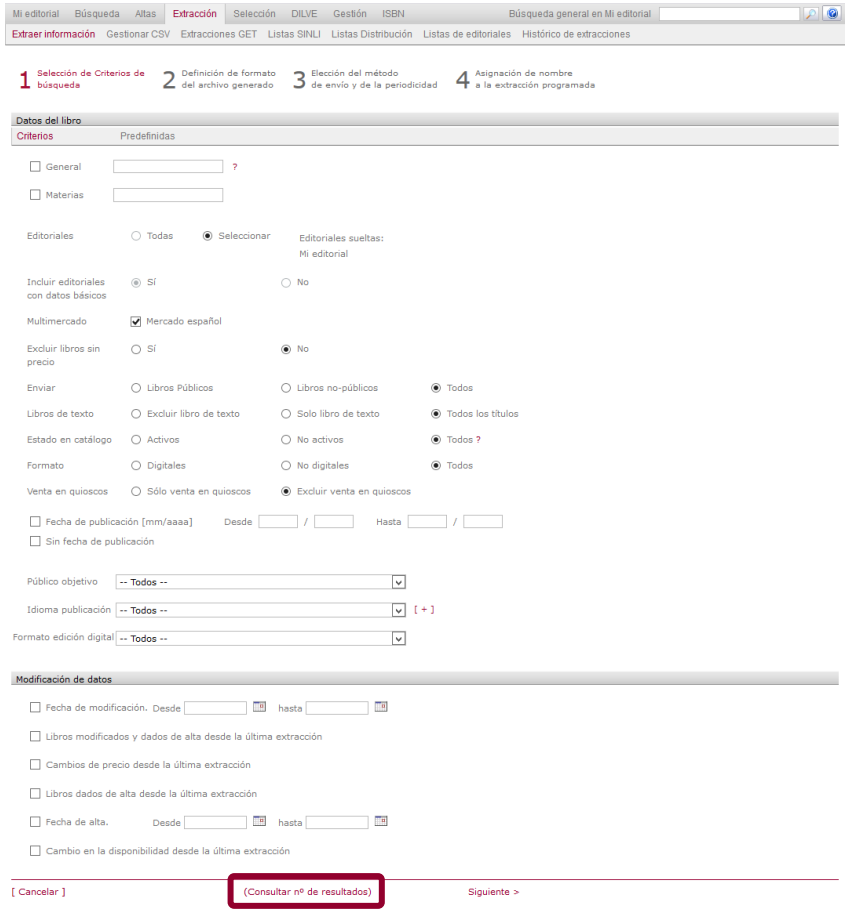
Por ejemplo, para comprobar la información sobre el peso en gramos, de los libros físicos, los pasos a seguir serían los descritos a continuación.

1. Crear un mapa CSV, en la sección Extracción > Gestionar CSV con el ISBN y el/los campo/s que se vayan a comprobar:

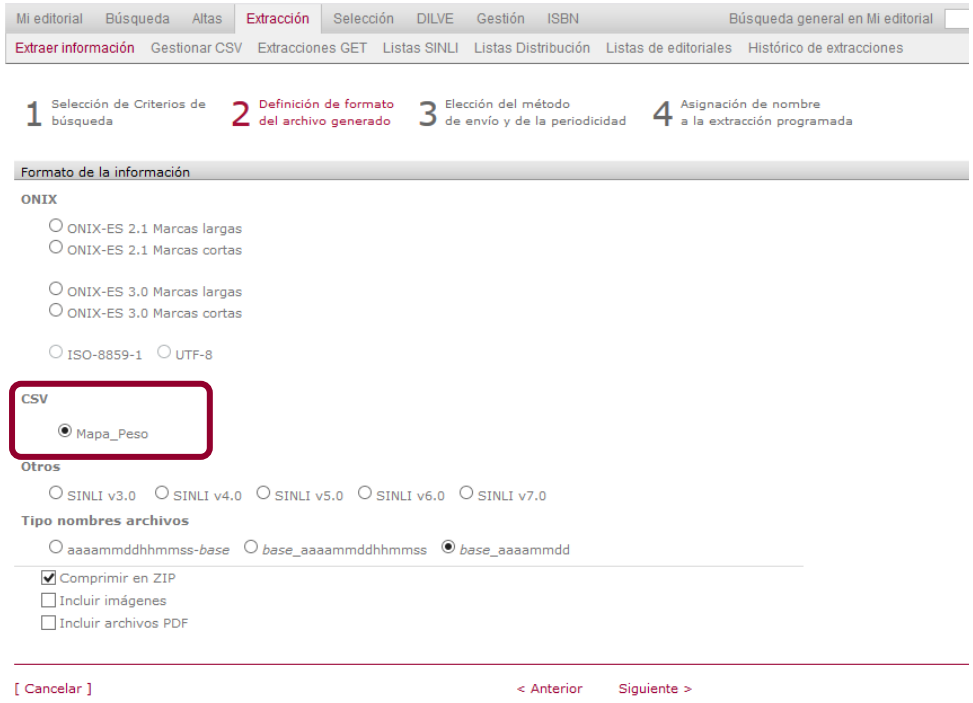


2. Crear una extracción de datos en la sección Extracción > Extraer información (Nueva programación E) siguiendo los siguientes pasos:

   1. En el paso 1, asegurarse de que los criterios que se establecen incluyen todos los libros de la editorial que se desean comprobar. Se puede utilizar el botón (Consultar nº de resultados) para confirmar que los criterios son correctos.



2. En el paso 2, seleccionar el mapa creado al comienzo del proceso.



3. En el paso 3, indicar el correo electrónico u otra forma de recibir el archivo de datos.

4. En el paso 4, asignar un nombre correcto a la extracción.

3. Por último es necesario ejecutar la extracción desde el listado y esperar a que lleguen los datos.

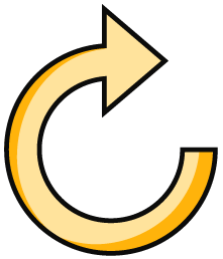
Origen	Programaciones	Periodicidad	Formato	
E	Comprobacion_Cubiertas *	Puntual	CSV Comprobacion_Cubiertas	Editar / Ejecutar / Clonar / Borrar
E	Comprobacion_Fecha_Publicacion *	Puntual	CSV Comprobacion_Fecha_Publicacion	Editar / Ejecutar / Clonar / Borrar
E	Comprobacion_Formato *	Puntual	CSV Datos_Formato	Editar / Ejecutar / Clonar / Borrar
E	Comprobacion_IBIC *	Puntual	CSV Materias IBIC	Editar / Ejecutar / Clonar / Borrar
E	Comprobacion_Imagen_Cubierta *	Puntual	CSV Comprobacion_Cubiertas	Editar / Ejecutar / Clonar / Borrar
E	Comprobacion_Peso	Puntual	CSV Mapa_Peso	Editar / Ejecutar / Clonar / Borrar
E	Comprobacion_Situacion_Disponibilidad *	Puntual	CSV Datos_Situacion_Disponibilidad	Editar / Ejecutar / Clonar / Borrar
E	Comprobacion_precios *	Puntual	CSV Precios_con_IVA	Editar / Ejecutar / Clonar / Borrar

Una vez obtenido el CSV correspondiente, se pueden encontrar y corregir fácilmente (mediante Excel), aquellos libros que no dispongan de esa información, o corregir los casos en los que sea incorrecta. En el anexo II de este documento se detalla la forma de gestionar un CSV con Excel.

Una vez completarlo el Excel únicamente es necesario guardarlo en el formato Excel por si es necesario volver a utilizarlo y guardarlo también como CSV para cargarlo como se describe en la sección Comparativa ISBN <> DILVE de este documento.

En el archivo Excel definitivo de carga se recomienda eliminar las filas de los ISBN que tengan la información correcta y dejar sólo aquellos para los que se vayan realmente a modificar la información.

Actualizar datos que cambian



Es muy importante mantener la información perfectamente actualizada en DILVE, sobre todo la comercial (disponibilidad, precio, etc.), ya que las entidades que utilizan la información de DILVE tienen extracciones diarias programadas y actualizan sus sistemas en el mismo momento de recibir la información.

No existe un método infalible para comprobar si los datos del catálogo de una editorial están actualizados o no en DILVE porque únicamente la editorial conoce la frecuencia con la que debe actualizarse su catálogo. Sin embargo, es conveniente hacer periódicamente las siguientes comprobaciones:

- **Libros públicos/no públicos**. En la sección Mi editorial > Libros no públicos comprobar si hay algún libro que debería estar público y no lo está.
- **Libros pendientes de registrar en la Agencia del ISBN**.
  - En la sección ISBN > Pendientes comprobar si hay algún ISBN rechazado por la Agencia cuyos datos deban depurarse.
  - En la sección Mi editorial > DILVE <> Agencia: comparar datos, en la pestaña Libros en DILVE y NO en Agencia ISBN comprobar qué libros están pendientes de enviar a registro y hacer el envío.
- **Libros modificados**. En la sección Mi editorial > Libros públicos hacer clic sobre la columna *F.Modificación* para conocer la última fecha en la que se modificaron los datos del catálogo para saber si está o no actualizado.
- **Novedades**. Hacer una búsqueda de los libros de la editorial por fecha de edición para comprobar que las últimas novedades está dadas de alta en DILVE y con información rica.
- **Cambios de precio**. Crear una extracción que permita extraer los libros que hayan podido cambiar de precio para obtener un Excel que permita hacer una carga masiva.

- **Búsquedas de libros activos y no activos**. Realizar búsquedas periódicas distinguiendo entre libros activos y no activos para comprobar que el estado de cada uno de ellos es correcto.

Otras ayudas

DILVE dispone de algunas herramientas que permiten automatizar algunas comprobaciones:

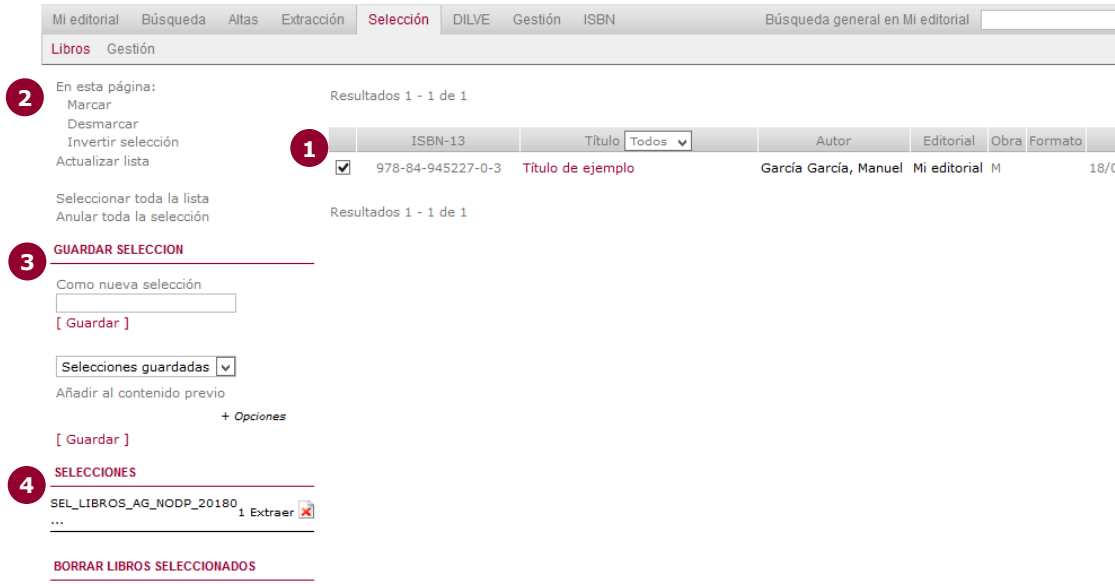
- **Actualización automática de Estado en catálogo y Disponibilidad**. Se debe solicitar el alta al servido de asistencia (asistencia@dilve.es) y esta funcionalidad se encarga de cambiar a estado "Activo" los libros que tengan estado "Próxima aparición" cuando se llegue a la fecha de publicación indicada y a "Disponible. Sin detalles." (en el mercado español) los libros que tengan "No disponible todavía" o "En espera de stock" y se llegue a la fecha de disponibilidad.
- **Servicio de avisos** que diariamente **comprueba y envía un correo** si existen:
  - Libros no públicos.
  - Libros que, por sus metadatos, deberían pasar a estado "Activo" o a "Disponible" y no se ha hecho la actualización.

Anexo I. Extracciones CSV de selección de ISBN

En algunos puntos de este documento se ha mencionado la posibilidad de seleccionar libros en los listados DILVE para posteriormente extraer algunos de sus metadatos. Aquí se describe en detalle cómo se utilizan las selecciones de libros en DILVE.

La selección de los libros se realiza desde los listados de libros, marcando los registros uno por uno o con las opciones de Seleccionar todos o Invertir selección que se incluyen debajo de todos los listados de libros.

Los libros que se tienen seleccionados en cada momento constituyen la "Selección actual", que se puede consultar y gestionar en la sección Selección > Libros.



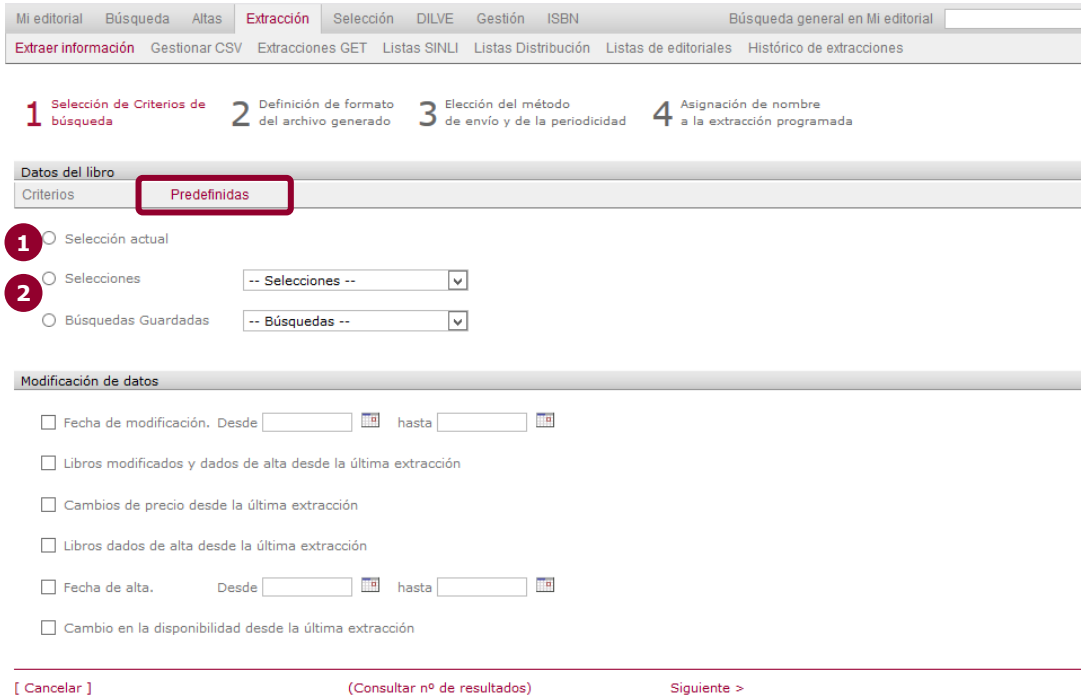
1. Libros seleccionados en la "Selección Actual".

2. Acciones que se pueden hacer para modificar la selección actual.
3. Opciones de guardado de la selección. Puede guardarse en una nueva selección independiente o hacer operaciones de conjuntos con otras selecciones de ISBN.
4. Listado de selecciones guardadas.

La gestión de las selecciones ya guardadas se realiza en Selección > Gestión:

Se recomienda consultar la guía de uso para editoriales para conocer en detalle las opciones que ofrece la gestión de selecciones.

Para realizar una extracción de los libros de la selección actual o de una selección guardada, debe indicarse en el paso 1 de las extracciones, en la pestaña de Predefinidas.



1. Selección actual. Es importante tener en cuenta que la selección actual de libros se mantiene mientras el usuario siga conectado a DILVE por lo que, si es probable que vaya a ser necesario repetir la extracción en otro momento, es recomendable guardar la selección asignándole un nombre.
2. Desplegable con las selecciones de ISBN guardadas por el usuario conectado

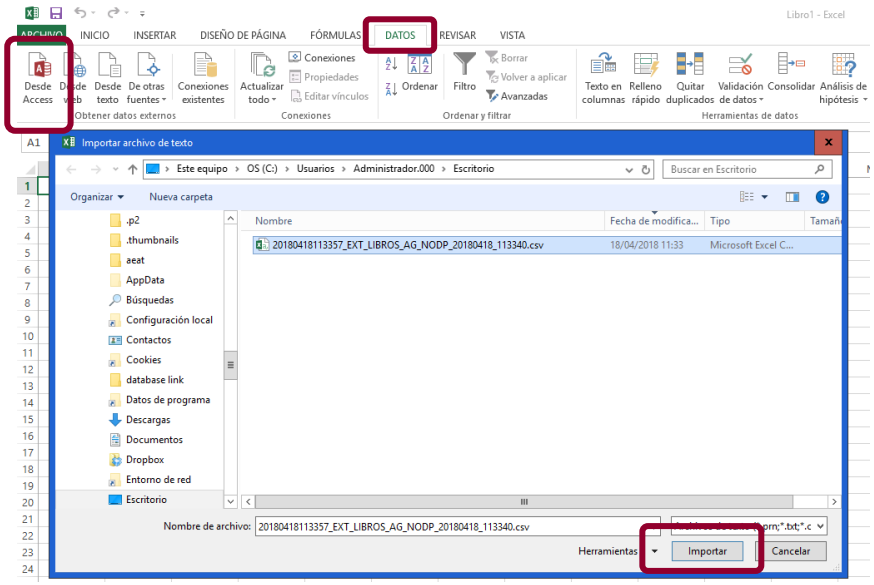
El resto de parámetros y la funcionalidad de la extracción de metadatos son idénticos a las descritas en otros puntos de este documento.

Anexo II. Gestión de CSV con Excel

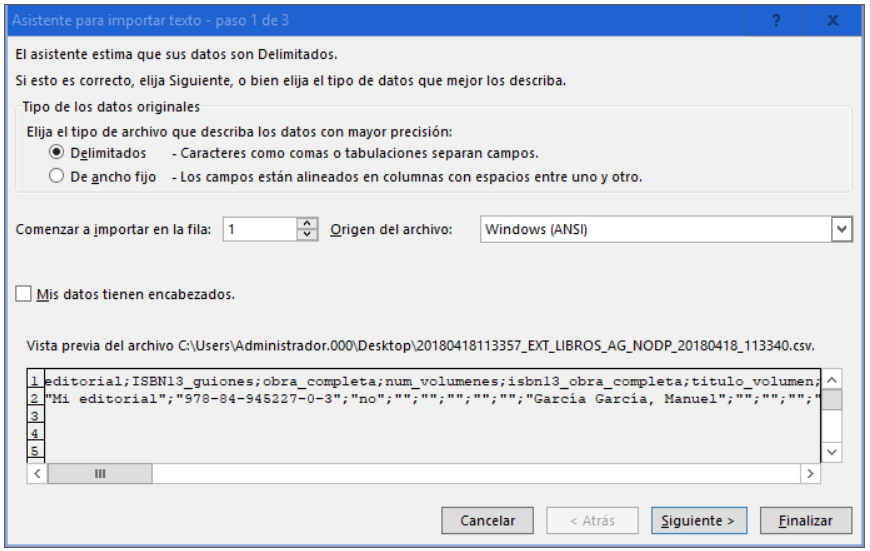
En este anexo se describe la forma óptima de abrir un CSV con Excel para evitar que la interpretación de la información que hace Excel de problemas al guardar el archivo como CSV y cargarlo en DILVE.

Los pasos a seguir serían los siguientes:

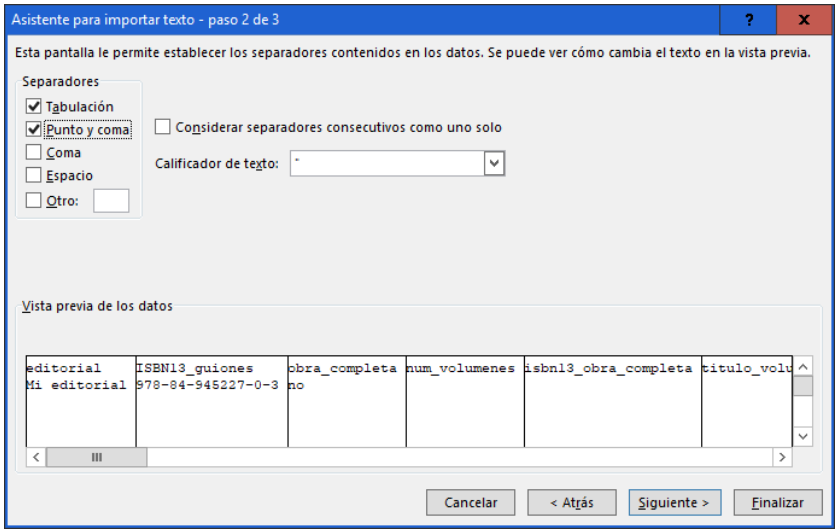
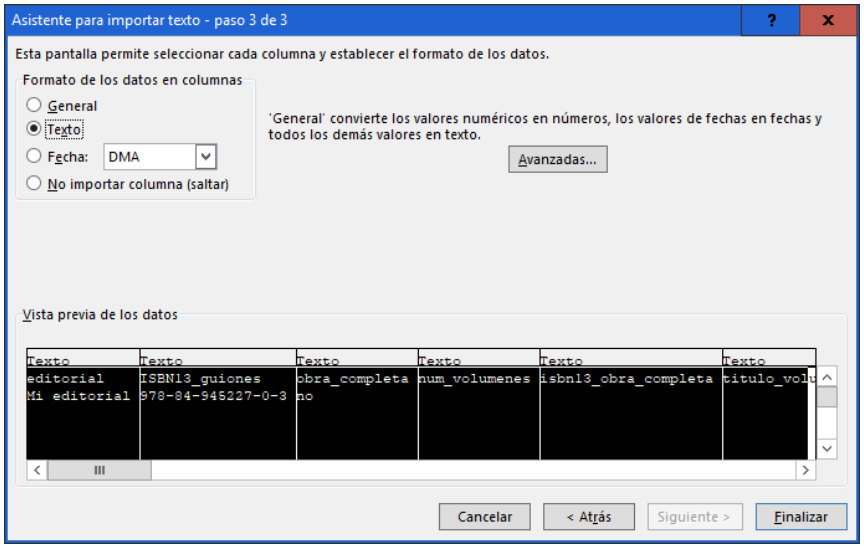
1. Abrir un libro Excel nuevo (vacío) y desde la sección Datos > desde texto cargar el CSV:



2. Seleccionar las opciones de *Delimitados* y *Origen de archivos Windows (ANSI)* que se muestran en la siguiente imagen y hacer clic en *Siguiente*:



3. Seleccionar como separador *Punto y coma* (o el separador seleccionado en DILVE) y como *Calificador de texto* las dobles comillas y hacer clic en *Siguiente*:



4. Indicar que todas las columnas deben importarse con formato *Texto* y hacer clic en *Finalizar*: