

PR24 – Limpieza de documentos

Clasificación de la Información:

Nivel del Documento	Procedimiento
Nombre del Fichero	PR24 – Limpieza de documentos
Tipo	Información No Clasificada
Ámbito de Difusión	Usuarios Universitat de València
Responsable	Responsable de Seguridad de la Información

CONTROL DE MODIFICACIONES		
Descripción	Versión	Fecha
Versión preliminar del documento.	0.1	24/01/2023
Revisión RD311/2022. Versión definitiva del documento	1.0	12/01/2024
Revisión eliminación Exiftool + qpdf	1.1	31/07/2024

ELABORADO:	REVISADO:	APROBADO:
Servei d'Informàtica	Responsable de Seguridad	Comité de Seguridad

ÍNDICE DE CONTENIDO

ÍNDICE DE CONTENIDO	3
1. OBJETO	4
2. ALCANCE	4
3. DESARROLLO	4
4. RESPONSABILIDADES	21
5. FICHAS	21
6. ANEXOS	21

1. OBJETO

El objeto del documento es proporcionar unas buenas prácticas para realizar la inspección y borrado tanto de metadatos, como de otros datos ocultos asociados a los documentos electrónicos, que facilite la implementación de un proceso de limpieza de documentos adecuado para la organización. Para ello, se tomará como referencia la Guía de Seguridad de las TIC CCN-STIC 835 Borrado de Metadatos y Blog publicado por INCIBE “Qué son los metadatos y como eliminarlos” (2022).

2. ALCANCE

El alcance de la presente guía abarca la inspección y borrado de los metadatos y otros datos ocultos existentes en los documentos, incorporados de forma automática por los programas de generación y tratamiento de estos documentos, o por los propios usuarios de la organización.

El ENS establece una serie de medidas de seguridad en su Anexo II, que se aplicarán “para lograr el cumplimiento de los principios básicos y requisitos mínimos establecidos” y que “serán proporcionales a las dimensiones de seguridad relevantes en el sistema a proteger y a la categoría del sistema de información a proteger”.

Dentro de las medidas de protección, la [mp.info.6] se refiere a la “Limpieza de documentos”, y es una medida destinada a la protección de la Confidencialidad de la Información, que aplica por igual a todos los Sistemas, cualquiera que sea su nivel de seguridad exigido (Alto, Medio o Bajo).

3. DESARROLLO

Se consideran como documentos electrónicos, los archivos de imagen (por ejemplo, fotografías digitales), los archivos multimedia, los documentos PDF, las hojas de cálculo, las presentaciones, gráficos, los documentos de texto formateado, etc.

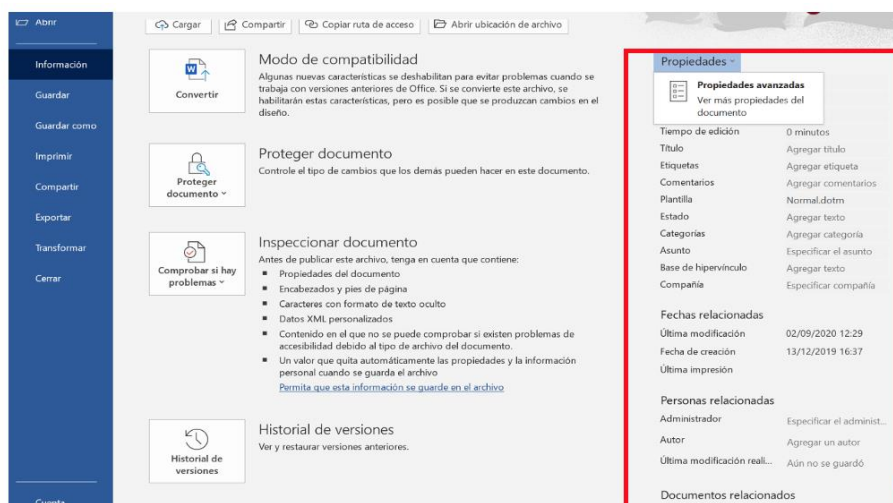
La medida de protección [mp.info.6] determina que “En el proceso de limpieza de documentos, se retirará de estos toda la información adicional contenida en campos ocultos, metadatos, comentarios o revisiones anteriores, salvo cuando dicha información sea pertinente para el receptor del documento.

Cualquier extensión que no quede reflejada a continuación, se consultará en la Guía de Seguridad de las TIC CCN-STIC 835 Borrado de Metadatos.

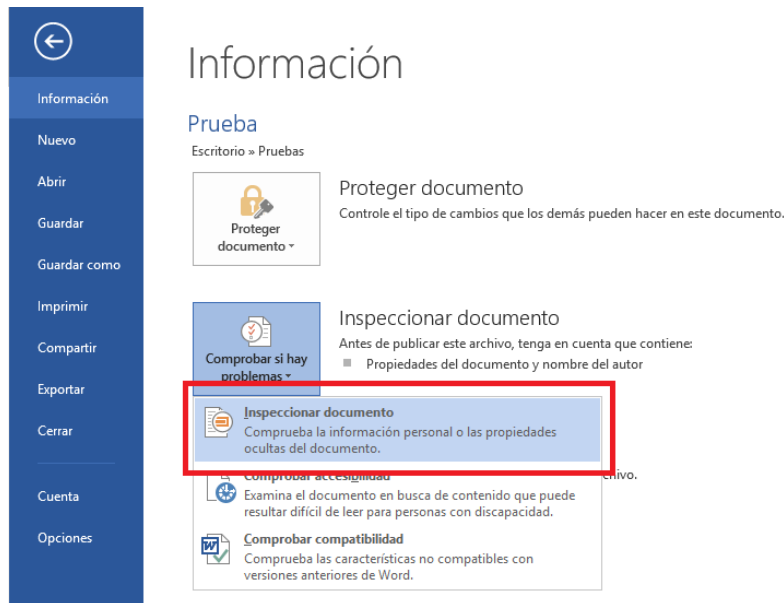
3.1 Limpieza de metadatos en documentos Microsoft Office (Word, Excel y PowerPoint)

Los documentos de Microsoft Office contienen metadatos en las Propiedades del documento que incluyen detalles sobre el archivo para describirlo e identificarlo, como el título, nombre del autor, asunto y etiquetas para identificar el contenido del documento y poder filtrar en las búsquedas.

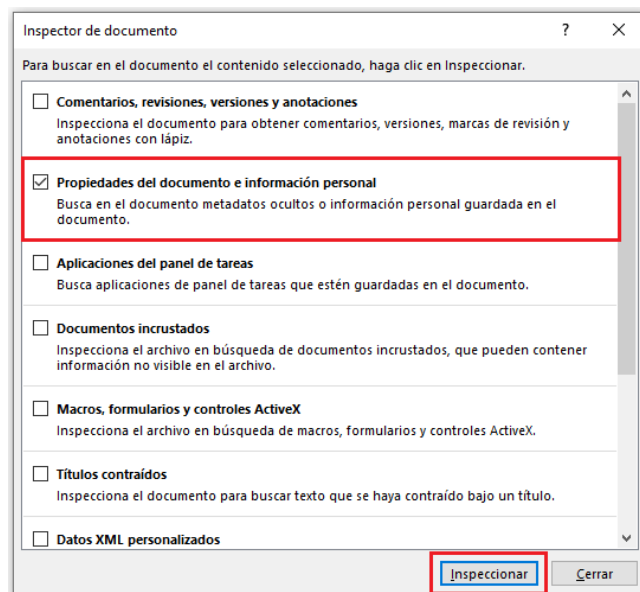
- a) Para visualizar la información de propiedades de un documento de Office nos dirigimos a la pestaña Archivo y a continuación hacemos clic en Información para ver las propiedades del documento, que aparecerán en el lado derecho de la pantalla.



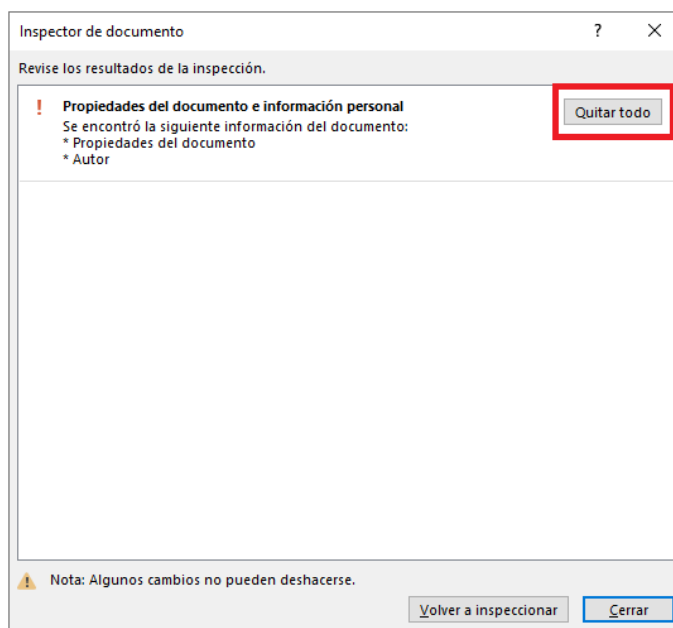
- b) Para eliminar los metadatos del documento guardado hacemos clic en comprobar si hay problemas y luego en Inspeccionar documento.



- c) Luego nos aparecerá un cuadro de diálogo con diferentes opciones, hacemos clic “Propiedades del documento e información personal” y, a continuación, pulsa en el botón “Inspeccionar”.



- d) Después de inspeccionar nuestro documento nos mostrara una ventana indicándonos si posee información personal en los metadatos, para eliminarlo solo hacemos clic en “Quitar todo” y en “Cerrar” para finalizar.

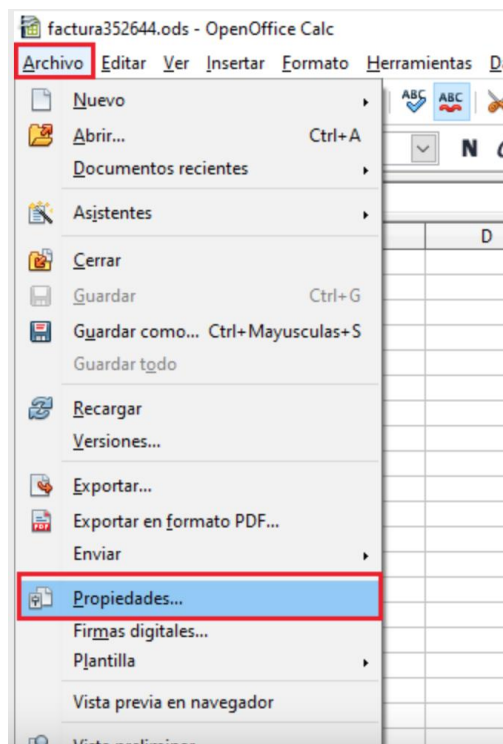


Finalmente, nuestro documento Word quedara limpio de metadatos. La persona que publica o comparte el documento será la encargada de asegurarse que el documento no contiene metadatos.

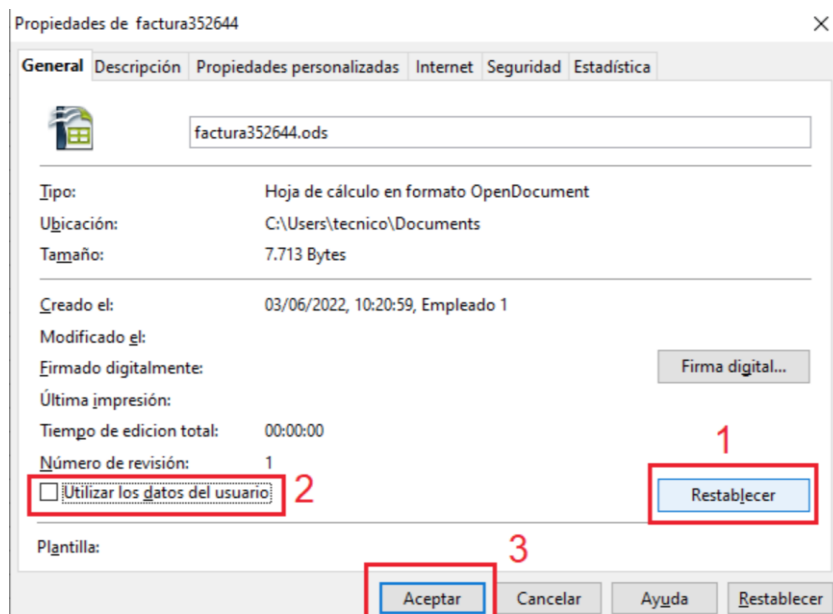
En las auditorías que se realizan sobre la web se comprueban los metadatos que puedan existir.

3.2 Limpieza de metadatos en documentos ofimáticos de LibreOffice.

- a) Para eliminar los metadatos en un archivo creado en LibreOffice, primero guardamos el documento, luego hacemos clic en el menú “Archivo” y seguidamente en “Propiedades”.



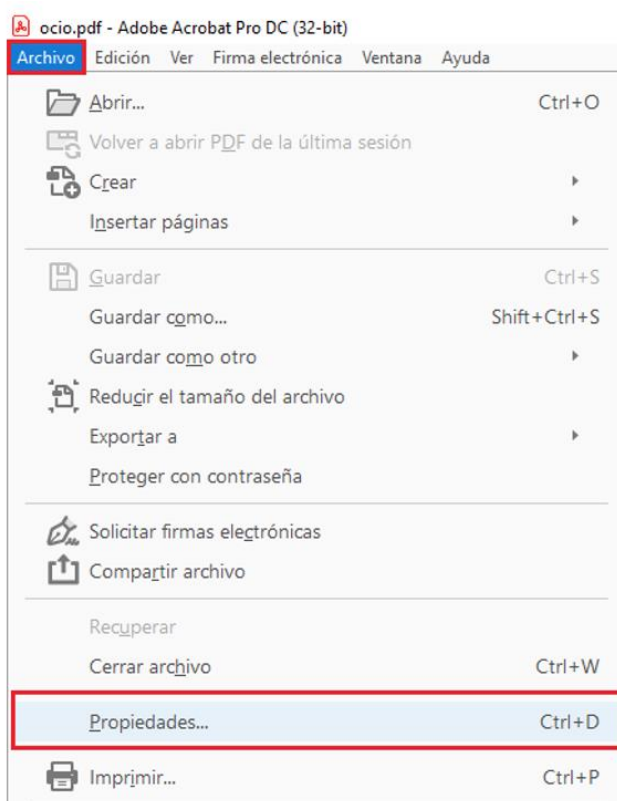
- b) Se abrirá una nueva ventana. Nos dirigimos a la pestaña “General”, hacemos clic en “Restablecer” y luego **deseleccionamos** la casilla “Utilizar los datos del usuario”, y finalmente damos clic en aceptar.



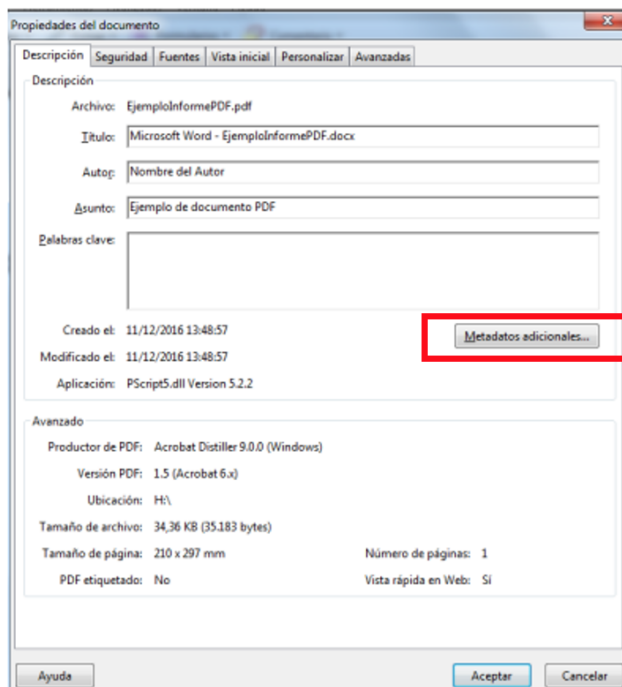
3.3 Limpieza de metadatos en documentos PDF- Adobe Acrobat Professional

Al igual que los documentos Office, los documentos PDF disponen también de propiedades del documento donde se almacenan metadatos insertados de forma automática por la aplicación como son el Título, Autor, Asunto, Palabras clave, Fechas de creación y modificación o Aplicación con la que se creó el documento PDF o metadatos estándar y personalizados insertados por el usuario o la organización.

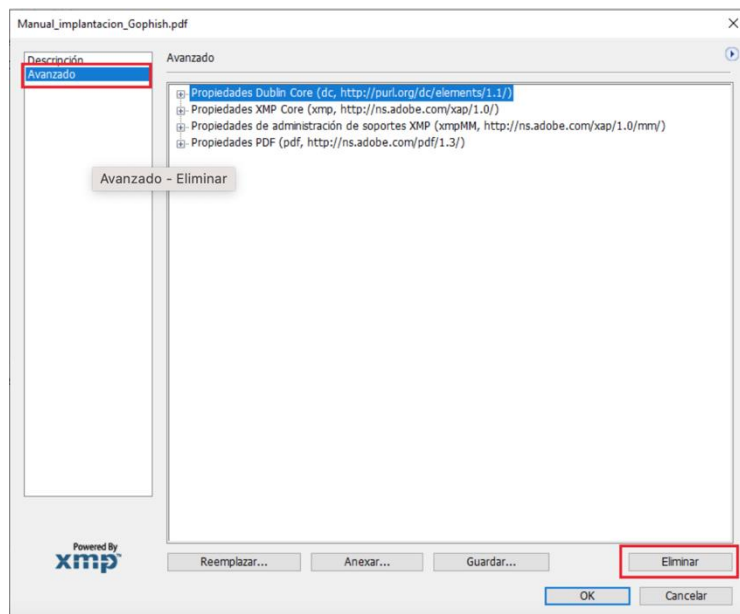
- a) Para visualizar los metadatos se accede desde la opción de menú Archivo, seleccionando Propiedades.



- b) Para eliminar los metadatos de nuestro pdf hacemos clic en metadatos adicionales, mostrado en la imagen anterior,

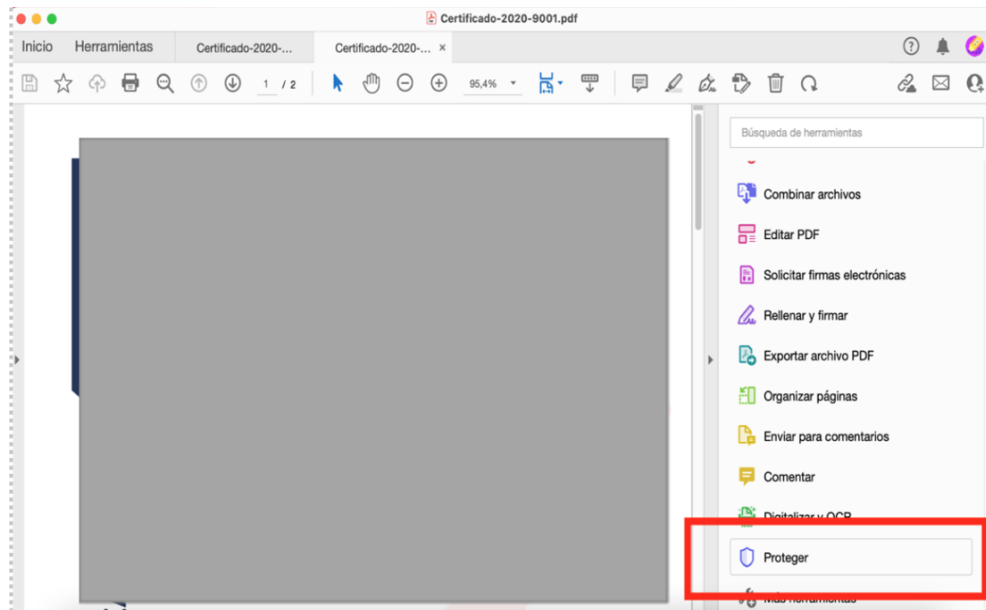


- c) Luego se nos aparecera una nueva venta, seleccionamos “Avanzado” y nos aparecera una listado de líneas en la derecha de la ventana , hacemos clic en una de ellas y pulsamos en el boton eliminar, luego repetir este proceso por cada una de ellas.

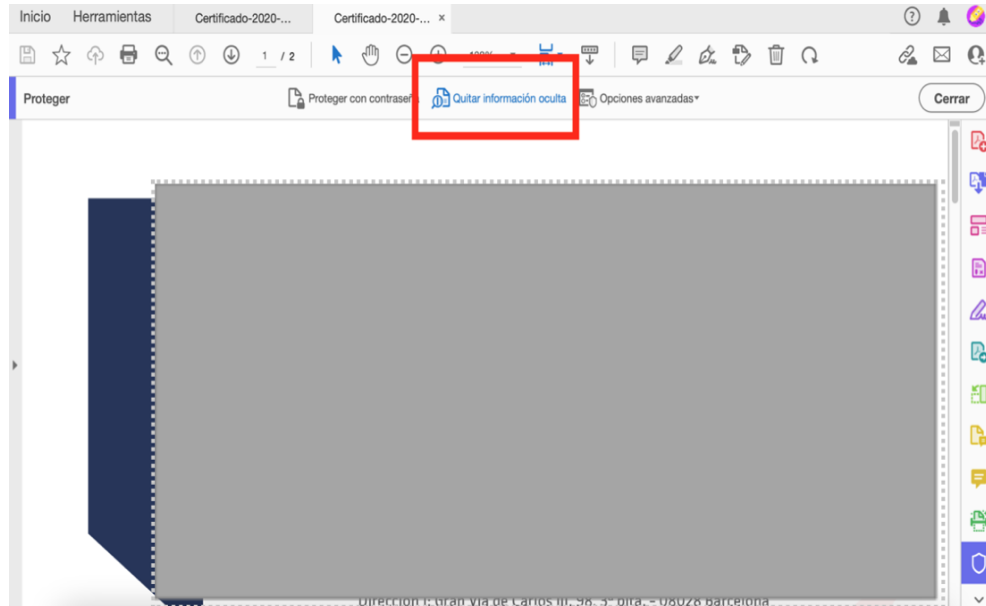


Finalmente, nuestro documento PDF quedara limpio de metadatos. Otra opción dentro de Adobe Acrobat Pro para eliminar toda informacion oculta y medatados adicionales es utilizar la herramienta PROTEGER.

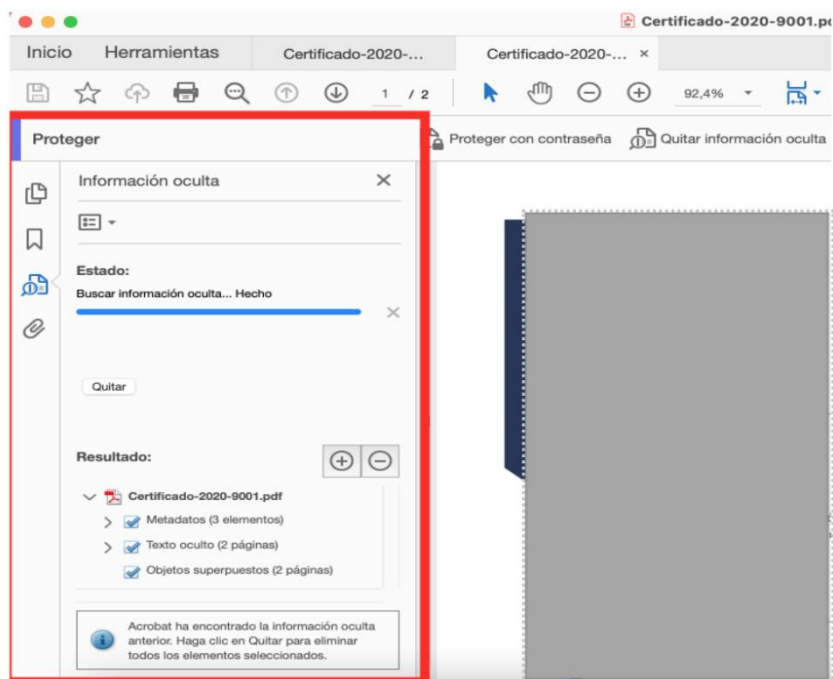
a) Abrimos nuestro documento y hacemos clic en la herramienta “Proteger”.



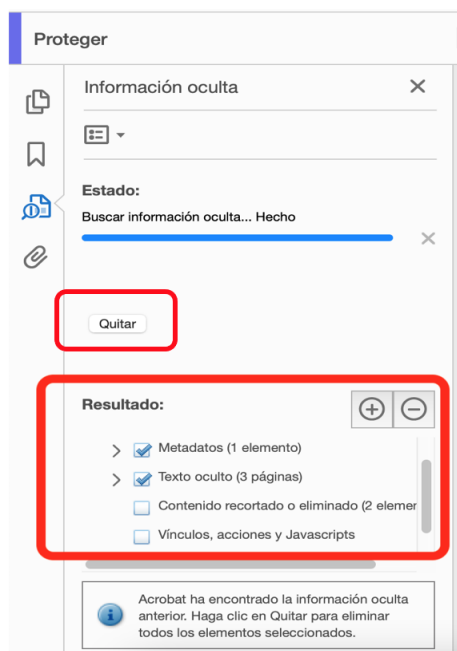
b) En la parte superior aparecerá una barra de menú secundaria y damos clic en “Quitar información oculta”.



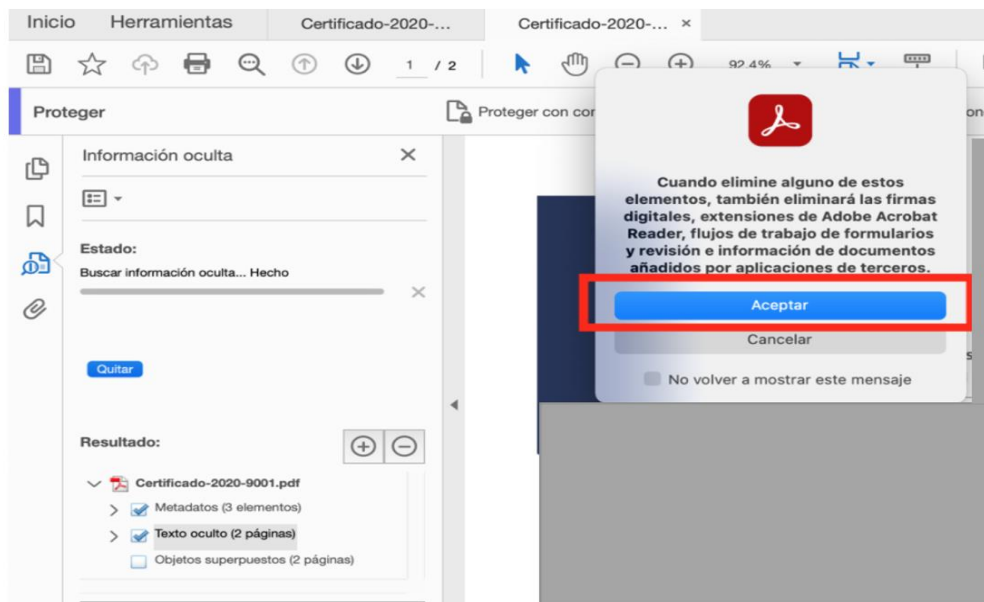
c) Se nos abrirá una barra lateral izquierda, donde empezará la buscar la información oculta que pueda contener el documento.



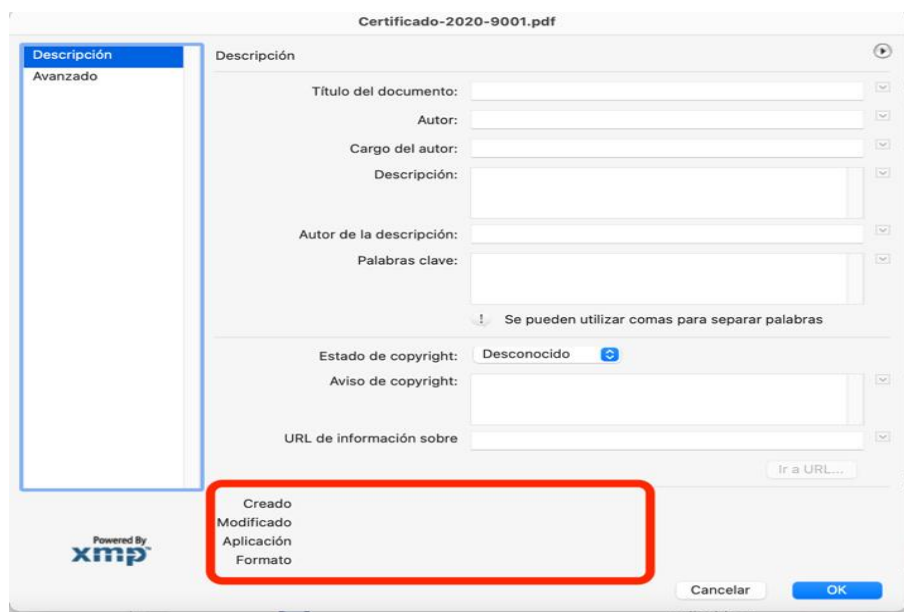
- d) Después de terminar la búsqueda, nos mostrará el resultado con el tipo de información encontrada. Debemos marcar las casillas de la información que queremos eliminar del documento.

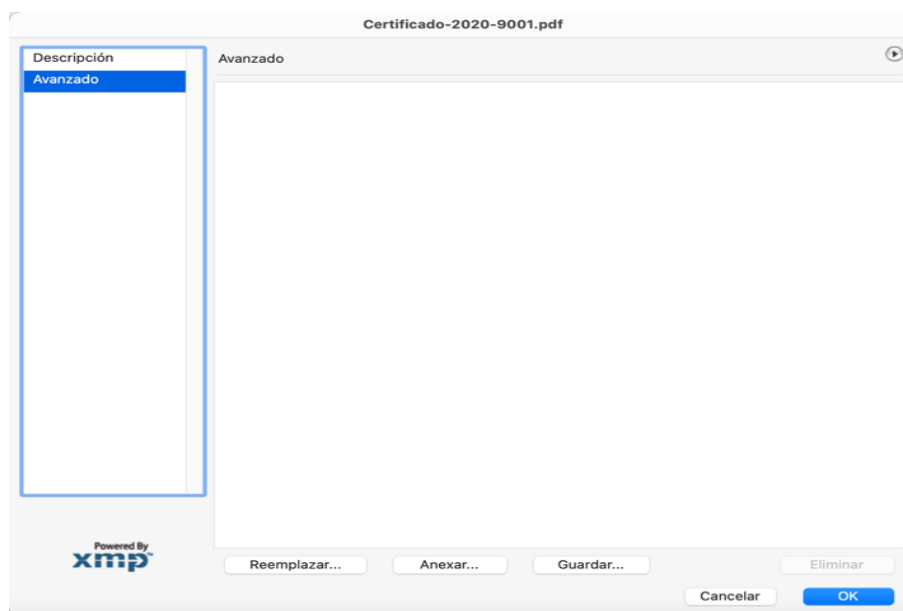


- e) Hacemos clic en “Quitar” para comenzar la eliminación. Luego nos aparecerá una ventana emergente y damos clic en “Aceptar”.



- f) Finalmente, comprobamos si nuestro documento está libre de metadatos ingresando de Archivo -> propiedades y guardamos el documento.





La persona que publica o comparte el documento será la encargada de asegurarse que el documento no contiene metadatos.

En las auditorías que se realizan sobre la web se comprueban los metadatos que puedan existir.

3.3 Limpieza de metadatos en imágenes

Los documentos de imagen (a los que nos referiremos como imágenes digitales), emplean múltiples formatos de fichero. Los más comunes son: TIFF, JPEG, PSD y RAW. Cada uno de estos formatos, tiene sus propias reglas sobre cómo almacenar los metadatos asociados.

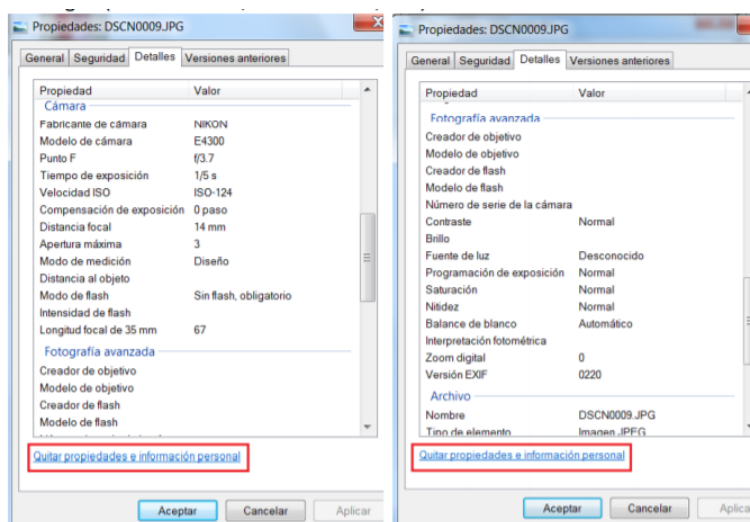
Los dispositivos móviles, tanto iOS como Android, permiten añadir información de la localización geográfica a las fotografías realizadas con la cámara del dispositivo. Esta información es conocida como geoetiquetas (o geotags). Al realizar la fotografía se añaden a la cabecera EXIF de la imagen las coordenadas GPS en el momento de realizar la captura.

En lo que respecta a las cámaras digitales, algunas disponen de opción para deshabilitar las coordenadas GPS de localización geográfica, y otras no. Dependerá del fabricante y modelo, y para modificar la configuración será necesario consultar el Manual de instrucciones de la cámara.

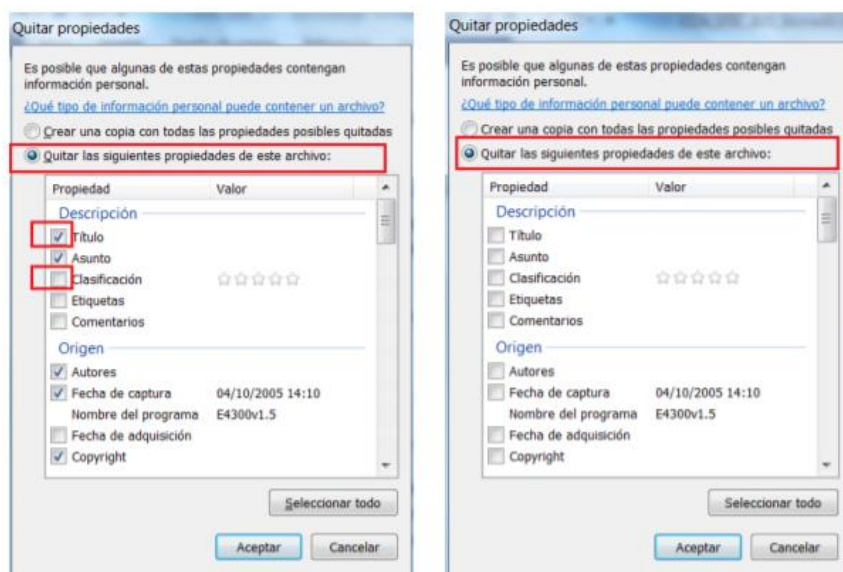
Se recomienda no hacer uso de esta funcionalidad que incluye la información de localización en fotografías, especialmente para su publicación o distribución en Internet, salvo que se quiera hacer uso explícito e intencionado de estas capacidades. En caso contrario, las imágenes revelarán los detalles exactos de dónde han sido tomadas.

Microsoft Windows permite visualizar y eliminar metadatos asociados a una imagen digital a través de las siguientes utilidades.

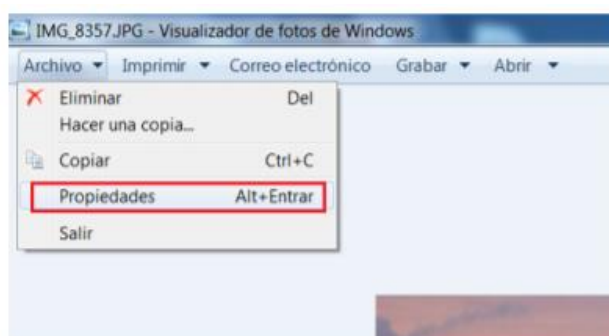
- Abrir el explorador de archivos y navegar hasta la imagen. Sobre ella, pulsar botón derecho del ratón, seleccionar Propiedades y seleccionar la pestaña Detalles. Windows detecta dos categorías de metadatos EXIF: “Cámara” y “Fotografía avanzada”. Luego hacemos clic en quitar propiedades e información personal.

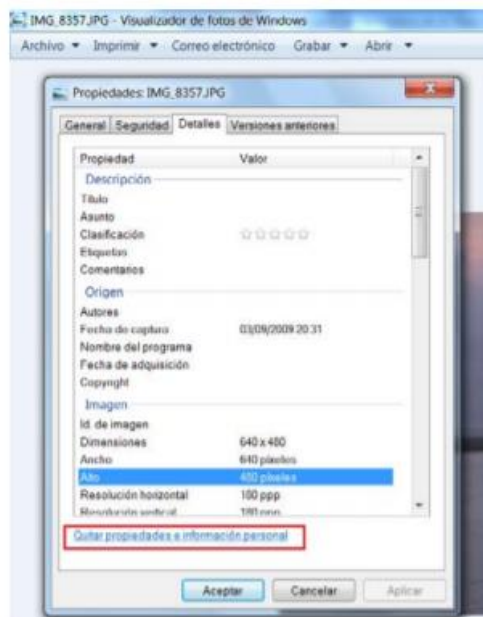


- Luego se nos aparecerá una ventana para seleccionar los datos que queremos borrar (todos, o un conjunto de ellos marcando su correspondiente check).



- c) Otra opción es con Visualizador de fotos de Windows. Si abrimos la imagen digital con el visualizador por defecto de Windows, al seleccionar Archivo y a continuación Propiedades, se abre la misma ventana que desde el explorador de ficheros. Permite por lo tanto visualizar y eliminar los metadatos tal y como se ha explicado en el apartado anterior.





Antes de publicar cualquier imagen en la web será necesario eliminar los metadatos que pudiera contener.

3.4 Eliminar metadatos con la herramienta ExifTool + qpdf

Para utilizar las herramientas ExifTool y qpdf, requieren de un conocimiento técnico más avanzando que los ejemplos anteriores. Primero tenemos que descargar los programas <https://exiftool.org/> y <https://github.com/qpdf/qpdf/releases>, si tienes problemas con la instalación, puedes solicitar el apoyo de nuestro equipo de sistemas.

Exiftool:

ExifTool by Phil Harvey

Read, Write and Edit Meta Information!

Also available --> [Utility to fix Nikon NEF images corrupted by Nikon software](#)

Note: If exiftool.org goes down, it is because of the crappy DreamHost web hosting which disables an "unlimited traffic" web site if a single bot hammers the site with a moderate load. An alternate ExifTool homepage is available at <http://exiftool.sourceforge.net/>

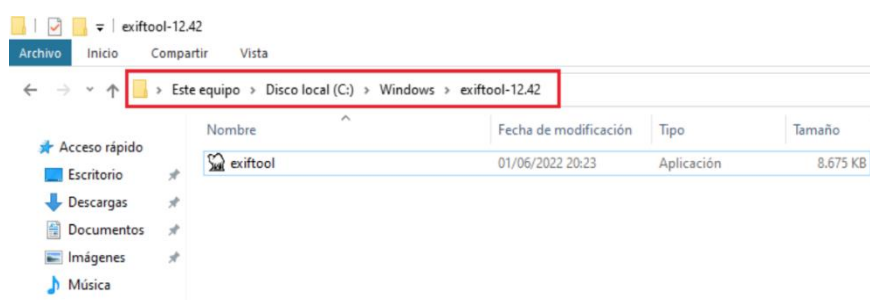
Installing	Tag Names	Resources	History	Forum	FAQ
<p>Download Version 12.44 (4.8 MB) - July 21, 2022</p> <p>ExifTool is a platform-independent Perl library plus a command-line application for reading, writing and editing meta information in a wide variety of files. ExifTool supports many different metadata formats including EXIF, GPS, IPTC, XMP, JFIF, GeoTIFF, ICC Profile, Photoshop IRB, FlashPix, ACP and ID3. Lyrics3, as well as the maker notes of many digital cameras by Canon, Casio, DJI, FLIR, FujiFilm, GE, GoPro, HP, JVC/Victor, Kodak, Leaf, Minolta/Konica-Minolta, Motorola, Nikon, Nintendo, Olympus/Epson, Panasonic/Leica, Pentax/Asahi, Phase One, Reconyx, Ricoh, Samsung, Sanyo, Sigma/Foveon and Sony.</p> <p>ExifTool is also available as a stand-alone Windows executable and a MacOS package: (Note that these versions contain the executable only, and do not include the HTML documentation or other files of the full distribution above.)</p> <p>Windows Executable: exiftool-12.44.zip (6.4 MB)</p> <p>The stand-alone Windows executable does not require Perl. Just download and un-zip the archive then double-click on "exiftool(-k).exe" to read the application documentation, drag-and-drop files and folders to view meta information, or rename to "exiftool.exe" for command-line use. Runs on all versions of Windows.</p> <p>(Note: Oliver Betz provides an alternate ExifTool Windows installer that avoids some problems of the self-extracting archive version above. Please post here if you have any problems/comments with this version.)</p> <p>MacOS Package: ExifTool-12.44.dmg (3.1 MB)</p>					
<p>Features User Comments Supported File Type System Requirements Running ExifTool Example Output Tag Names Explained Tag Groups Writing Information Writer Limitations Known Problems Security Issues Date/Time Shift Renaming Files Performance ExifTool Library Additional Resources New Discoveries Acknowledgements License Donate Contact Me</p>					

Consulte la sección correspondiente a continuación con instrucciones para instalar o desinstalar ExifTool en su plataforma específica:

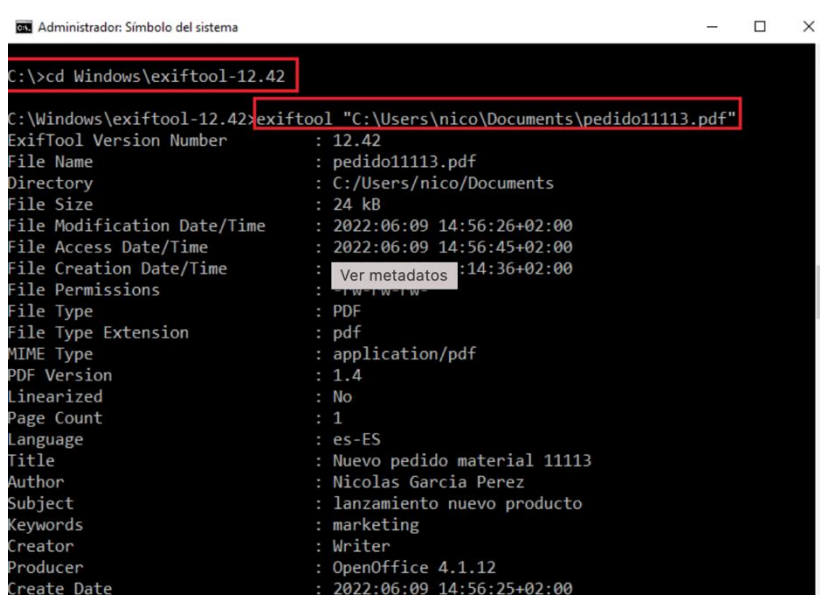
- [Windows](#)
- [MacOS](#)
- [Unix Platforms](#)

Pasos de instalación para Windows

- Descargamos el archivo de la página oficial, nos dirigimos a descargas y descomprimos el archivo, luego abrimos la carpeta que acabamos de descomprimir donde se encuentra el ejecutable del programa.
- Cambiar el nombre de " exiftool(-k).exe" a " exiftool.exe". (o " exiftool(-k)" a " exiftool" si las extensiones de nombre de archivo están ocultas en su sistema)
- Mueve "exiftool.exe" al directorio "C:\WINDOWS" (o a cualquier otro directorio de su ruta), como se muestra en la siguiente imagen.



- Luego abrimos la consola "cmd", accedemos a la ruta donde guardamos el programa ExifTool. Ahora podemos visualizar los metadatos del archivo con el comando "exiftool", seguido de la ruta del archivo junto con su extensión entre comillas.



- e) Para eliminar los metadatos mostrados del documento, igual que el paso anterior escribimos el comando `exiftool -all= ruta y nombre del archivo sin comillas`.

```
Administrador: Símbolo del sistema
C:\Windows\exiftool-12.42>exiftool -all= C:\Users\nico\Documents\pedido11113.pdf
Warning: [minor] ExifTool PDF edits are reversible. Deleted tags may be recovered! -
C:/Users/nico/Documents/pedido11113.pdf
1 image files updated
C:\Windows\exiftool-12.42>
```

Una vez eliminados los metadatos podemos comprobar que el documento original sigue guardado con el terminador “_original” y adicional tenemos una copia del documento sin metadatos.

Nota: En el caso de documentos .pdf, es importante reseñar que con la herramienta Exiftool se pueden recuperar los metadatos eliminados con ella, y es por ello necesario a continuación, hacer uso de la segunda herramienta **qpdf**, que produce la linealización del archivo, impidiendo la recuperación de los metadatos eliminados.

qpdf:

github.com/qpdf/qpdf/releases

Product Solutions Resources Open Source Enterprise Pricing

qpdf / qpdf Public

Code Issues 103 Pull requests 22 Discussions Actions Projects Wiki Security Insights

Releases Tags Find a release

Jun 7
jberkenbilt
v11.9.1
c03ee7a
Compare

qpdf 11.9.1 Latest

This is qpdf version 11.9.1. This is an accumulation of minor fixes and performance enhancements since 11.9.0. No new functionality has been added.

For a full list of changes from previous releases, please see the [release notes](#). See also [README-what-to-download](#) for details about the available source and binary distributions.

Assets 17

qpdf-11.9.1-bin-linux-x86_64.zip	3.16 MB	Jun 7
qpdf-11.9.1-doc.zip	6.58 MB	Jun 7
qpdf-11.9.1-mingw32.exe	8.6 MB	Jun 7
qpdf-11.9.1-mingw32.zip	12.5 MB	Jun 7
qpdf-11.9.1-mingw64.exe	8.5 MB	Jun 7
qpdf-11.9.1-mingw64.zip	12.5 MB	Jun 7
qpdf-11.9.1-msvc32.exe	7.82 MB	Jun 7
qpdf-11.9.1-msvc32.zip	12.1 MB	Jun 7
qpdf-11.9.1-msvc64.exe	8.79 MB	Jun 7
qpdf-11.9.1-msvc64.zip	13.8 MB	Jun 7
Source code (zip)		Jun 7
Source code (tar.gz)		Jun 7

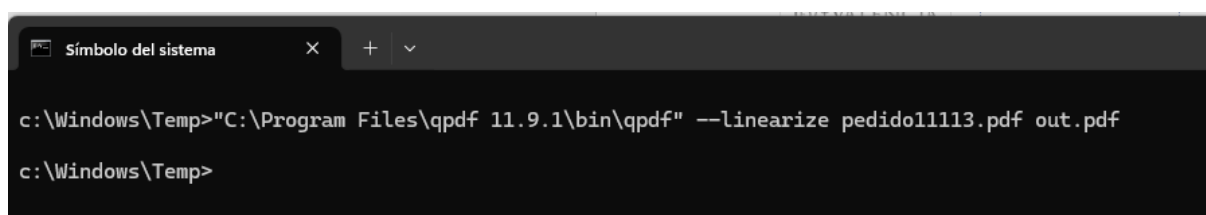
La linealización de PDF es un proceso que reordena las estructuras de datos internas en un archivo PDF. El objetivo de esta reordenación es doble: colocar los datos de la primera página al principio del archivo PDF y asegurarse de que el índice de ubicación de los datos también se inserte al principio del archivo. El resultado en cualquier visor de PDF es exactamente el mismo: no se agrega ni se elimina nada.

La linealización se puede realizar en cualquier archivo PDF, independientemente del software que se utilizó para crear el archivo PDF. También conocida como «optimización web» o «vista web rápida», la linealización debe realizarse antes de publicar el archivo desde el servidor, porque implica una optimización.

Pasos de instalación para Windows

- f) Descargamos el archivo de la página oficial, nos dirigimos a descargas y ejecutamos el archivo, la instalación asistida iniciará y al finalizar tendremos el ejecutable en la ruta de instalación (por defecto *C:\Program Files\qpdf 11.9.1\bin*).
- g) Luego abrimos la consola "cmd", podemos ejecutar (el comando que se incluye es para ejecutarlo en la ruta por defecto de instalación, cambiar la ruta en el caso de que se haya instalado en otra) el programa qpdf en el mismo directorio donde se encuentra el archivo pdf a linealizar mediante el comando "*C:\Program Files\qpdf 11.9.1\bin\qpdf" --linearize pedido11113.pdf out.pdf*"

Donde *pedido11113.pdf* es el **fichero resultado del apartado e)** tras aplicarle la eliminación de metadatos con el comando de exiftool y *out.pdf* es el resultado final del **archivo limpio de metadatos e irrecuperables**.



```
Símbolo del sistema
c:\Windows\Temp>"C:\Program Files\qpdf 11.9.1\bin\qpdf" --linearize pedido11113.pdf out.pdf
c:\Windows\Temp>
```

En resumen, para la eliminación completa de los metadatos en un archivo pdf, es necesario aplicar ambos programas, primero exiftool y luego qpdf:

- 1) proceso indicado en el apartado e)
- 2) al resultado de 1) aplicar el proceso indicado en el apartado g)

4. RESPONSABILIDADES

Los roles específicos del ENS para este proceso son los siguientes:

- **Responsable de Información:** revisara si es efectivo el procedimiento establecido.
- **Responsable de seguridad:** Se encarga de asegurar que se conozca y se cumpla el procedimiento.
- **Personal de la Universitat de València que publique o comparta documentos e imágenes:** Se encarga de eliminar los metadatos

5. FICHAS

N/A.

6. ANEXOS

Guía de Seguridad de las TIC CCN-STIC 835 Borrado de Metadatos.