

Guía de Seguridad de las TIC CCN-STIC 835

Esquema Nacional de Seguridad Borrado de Metadatos



Marzo 2017

Edita:



© Centro Criptológico Nacional, 2017



CENTRO CRIPTOLOGICO NACIONAL
2.5.4.13=Qualified Certificate: AAPP-SEP-M-SW-KPSC,
ou=sello electrónico, serialNumber=S2800155J,
o=CENTRO CRIPTOLOGICO NACIONAL, cn=CENTRO
CRYPTOLOGICO NACIONAL, c=ES
2017.06.26 20:51:07 +02'00'

NIPO: 785-17-011-6

Fecha de Edición: junio de 2017

El Ministerio de Hacienda y Función Pública ha financiado el desarrollo de la presente guía.

Isdefe ha participado en la elaboración y modificación del presente documento y sus anexos.

LIMITACIÓN DE RESPONSABILIDAD

El presente documento se proporciona de acuerdo con los términos en él recogidos, rechazando expresamente cualquier tipo de garantía implícita que se pueda encontrar relacionada. En ningún caso, el Centro Criptológico Nacional puede ser considerado responsable del daño directo, indirecto, fortuito o extraordinario derivado de la utilización de la información y software que se indican incluso cuando se advierta de tal posibilidad.

AVISO LEGAL

Quedan rigurosamente prohibidas, sin la autorización escrita del Centro Criptológico Nacional, bajo las sanciones establecidas en las leyes, la reproducción parcial o total de este documento por cualquier medio o procedimiento, comprendidos la reprografía y el tratamiento informático, y la distribución de ejemplares del mismo mediante alquiler o préstamo públicos.

PRÓLOGO

El uso masivo de las tecnologías de la información y las telecomunicaciones (TIC), en todos los ámbitos de la sociedad, ha creado un nuevo espacio, el ciberespacio, donde se producirán conflictos y agresiones, y donde existen ciberamenazas que atentarán contra la seguridad nacional, el estado de derecho, la prosperidad económica, el estado de bienestar y el normal funcionamiento de la sociedad y de las administraciones públicas.

La Ley 11/2002, de 6 de mayo, reguladora del Centro Nacional de Inteligencia (CNI), encomienda al Centro Nacional de Inteligencia el ejercicio de las funciones relativas a la seguridad de las tecnologías de la información en su artículo 4.e), y de protección de la información clasificada en su artículo 4.f), a la vez que confiere a su Secretario de Estado Director la responsabilidad de dirigir el Centro Criptológico Nacional (CCN) en su artículo 9.2.f).

Partiendo del conocimiento y la experiencia del CNI sobre amenazas y vulnerabilidades en materia de riesgos emergentes, el Centro realiza, a través de su Centro Criptológico Nacional, regulado por el Real Decreto 421/2004, de 12 de marzo, diversas actividades directamente relacionadas con la seguridad de las TIC, orientadas a la formación de personal experto, a la aplicación de políticas y procedimientos de seguridad, y al empleo de tecnologías de seguridad adecuadas.

El Real Decreto 3/2010, de 8 de enero, por el que se regula el Esquema Nacional de Seguridad en el ámbito de la Administración Electrónica (ENS, en adelante), al que se refiere el apartado segundo del artículo 156 de la Ley 40/2015, de 1 de octubre, de Régimen Jurídico del Sector Público, establece la política de seguridad en la utilización de medios electrónicos que permita una protección adecuada de la información.

Precisamente el Real Decreto 3/2010 de 8 de Enero, actualizado por el Real Decreto 951/2015, de 23 de octubre, fija los principios básicos y requisitos mínimos así como las medidas de protección a implantar en los sistemas de la Administración, y promueve la elaboración y difusión de guías de seguridad de las tecnologías de la información y las comunicaciones (STIC) por parte de CCN para facilitar un mejor cumplimiento de dichos requisitos mínimos.

En definitiva, la serie de documentos CCN-STIC se elabora para dar cumplimiento a los cometidos del Centro Criptológico Nacional y a lo reflejado en el Esquema Nacional de Seguridad, conscientes de la importancia que tiene el establecimiento de un marco de referencia en esta materia que sirva de apoyo para que el personal de la Administración lleve a cabo su difícil, y en ocasiones, ingrata tarea de proporcionar seguridad a los sistemas de las TIC bajo su responsabilidad.

Marzo de 2017



Félix Sanz Roldán
Secretario de Estado
Director del Centro Criptológico Nacional

ÍNDICE

1. INTRODUCCIÓN	6
2. OBJETO	7
3. ALCANCE.....	7
4. METADATOS	8
4.1 DEFINICIÓN Y GENERALIDADES.....	8
4.2 TIPOS DE METADATOS	10
4.3 RIESGOS Y AMENAZAS	13
5. MEDIDAS DE SEGURIDAD EN EL ENS.....	17
6. HERRAMIENTAS DE INSPECCIÓN Y BORRADO DE METADATOS	21
6.1 TIPOS DE HERRAMIENTAS.....	22
6.2 CARACTERÍSTICAS DE LAS HERRAMIENTAS	25
ANEXO A. METADATOS EN DOCUMENTOS MICROSOFT OFFICE	28
1. TIPOS DE METADATOS E INFORMACIÓN OCULTA	28
2. CONFIGURACIÓN DE SEGURIDAD	32
3. INSPECCIÓN Y BORRADO DE METADATOS E INFORMACIÓN OCULTA	33
ANEXO B. METADATOS EN DOCUMENTOS OPENOFFICE.....	40
1. TIPOS DE METADATOS E INFORMACIÓN OCULTA	40
2. CONFIGURACIÓN DE SEGURIDAD	42
3. INSPECCIÓN Y BORRADO DE METADATOS E INFORMACIÓN OCULTA	44
ANEXO C. METADATOS EN DOCUMENTOS PDF	46
1. TIPOS DE METADATOS E INFORMACIÓN OCULTA	46
2. CONFIGURACIÓN DE SEGURIDAD	49
3. INSPECCIÓN Y BORRADO DE METADATOS E INFORMACIÓN OCULTA	49
ANEXO D. METADATOS EN IMÁGENES.....	53
1. TIPOS DE METADATOS	53
2. CONFIGURACIÓN DE SEGURIDAD	57
3. INSPECCIÓN Y BORRADO DE METADATOS.....	59
3.1. UTILIDADES PARA ORDENADOR	59
3.2. UTILIDADES PARA DISPOSITIVOS MÓVILES	64
ANEXO E. DEFINICIONES	66
ANEXO F. REFERENCIAS	68

TABLAS

Tabla 1. Formatos de metadatos comunes según los tipos de Documento.....	13
Tabla 2. Metadatos y datos ocultos en los documentos y sus riesgos asociados.....	17
Tabla 3. Propiedades de los documentos Microsoft Office 2010 (Word, Excel y PowerPoint).	29
Tabla 4. Información oculta en los documentos Office 2010/2013/2016.	32
Tabla 5. Información que inspecciona y elimina el Inspector de Documentos para Office 2010/2013/2016.....	37
Tabla 6. Información que revisa, pero no elimina el Inspector de Documentos en Office 2010/2013/2016.....	38
Tabla 7. Metadatos en Propiedades de documentos OpenOffice.org (versión 4)	41
Tabla 8. Información que inspecciona y elimina la utilidad “Eliminar Información oculta” de Adobe Acrobat X.	52
Tabla 9. Detalle de Metadatos comunes o exclusivos en cada estándar.....	56

FIGURAS

Figura 1. Ejemplo de un documento que revela información sensible a través de los metadatos de las propiedades del documento.	13
Figura 2. Medidas del ENS relacionadas.....	19
Figura 3. Diagrama ejemplo del funcionamiento de una herramienta de inspección y borrado de metadatos para correo electrónico.	24
Figura 4. Propiedades del documento en Adobe Acrobat 9.	46
Figura 5. Metadatos XMP en documento Adobe Acrobat 9.	47
Figura 6. Utilidad de Examinar documento en Acrobat 9.	50
Figura 7. Ejemplo de Metadatos EXIF.....	53
Figura 8. Ejemplo de Metadatos IPTC.	54
Figura 9. Ejemplo de Metadatos XMP.....	55
Figura 10. Metadatos comunes o exclusivos en cada estándar.....	55
Figura 11. Mapeo de metadatos a XMP.	56
Figura 12. Tipos de Metadatos en ficheros JPEG.	56
Figura 13. Deshabilitar localización en cámara iOS 10.x.	58
Figura 14. Deshabilitar localización en cámara Android 5.x.	59
Figura 15. Metadatos mostrados por explorador de ficheros Windows.	60
Figura 16. Eliminar metadatos con Explorador de ficheros Windows.	61
Figura 17. Visualizador de fotos de Windows.	61
Figura 18. Propiedades con Visualizador de fotos de Windows.	62
Figura 19. Metadatos visualizados en GIMP 2.8.20.	62
Figura 20. Eliminar metadatos en GIMP 2.8.20. Paso 1.	63
Figura 21. Eliminar metadatos en GIMP 2.8.20. Paso 2.	63
Figura 22. Eliminar metadatos en GIMP 2.8.20. Paso 3.	63
Figura 23. Ejemplos de visualización de metadatos en iPhone.	64
Figura 24. Ejemplo de borrado de metadatos en iPhone.	65
Figura 25. Ejemplos de compartir imagen con o sin metadatos en iPhone.....	65

1. INTRODUCCIÓN

1. El Real Decreto 3/2010, de 8 de enero, por el que se regula el Esquema Nacional de Seguridad (ENS) en el ámbito de la Administración Electrónica, modificado a su vez por el Real Decreto 951/2015, de 23 de octubre, establece una serie de medidas de seguridad en su Anexo II, entre las que se encuentra la [mp.info.6] sobre la “**Limpieza de documentos**”, destinada a la protección de la Confidencialidad de la Información.
2. El Real Decreto 4/2010, de 8 de enero, por el que se regula el Esquema Nacional de Interoperabilidad en el ámbito de la Administración Electrónica, define **documento electrónico** como “información de cualquier naturaleza en forma electrónica, archivada en un soporte electrónico según un formato determinado y susceptible de identificación y tratamiento diferenciado”.
3. Para su uso en esta guía, se tienen en cuenta las siguientes definiciones:
 - Se definen **programas de generación y tratamiento de documentos**, como aquellos programas de ordenador destinados a la generación de documentos en cualquier formato (por ejemplo, en formato PDF). Dentro de estos programas, se encuentran los **programas ofimáticos** que se definen como un conjunto de programas básicos para su uso en oficinas, con un interfaz y funciones comunes y cuyo objetivo será el tratamiento de textos, hojas de cálculo, presentaciones, gráficos, tablas, etc. Dos ejemplos de programas ofimáticos muy conocidos, son Microsoft Office y Apache OpenOffice.
 - Se definen **documentos ofimáticos** como un tipo de documento electrónico generado por los programas ofimáticos.
 - Se define **metadato** como información estructurada que describe, explica, localiza y además hace más fácil recuperar, utilizar o gestionar un recurso de información. Los metadatos son comúnmente llamados “datos sobre los datos” o “información sobre la información”.
 - Se define **información o datos ocultos** como aquellos datos existentes en el contenido de los documentos electrónicos, que no son visibles con la configuración estándar o configuración por defecto de los programas utilizados para su creación y tratamiento, siendo necesario aplicar alguna opción específica dentro de la configuración de estos programas, para su visualización. Un ejemplo de datos ocultos es el texto oculto, filas o columnas ocultas, comentarios o información del documento, etc.
4. Los metadatos y los datos ocultos pueden contener información sensible, y esto representa un riesgo que las organizaciones y sus usuarios deben entender. Este riesgo está asociado a la posibilidad de revelar información sensible cuando el documento sea compartido y salga fuera de su dominio de seguridad.
5. El impacto causado, por un lado, podría conducir a la organización a una pérdida de confidencialidad, enojo por parte de los clientes, acciones disciplinarias, daño

en su reputación, escándalo y sanciones legales. Pudiendo ser más grave en caso de que esta información contenga datos personales identificables (PII) que podrían estar sujetos a la Ley Orgánica 15/1999, de 13 de diciembre, de Protección de Datos de Carácter Personal.

6. Por otro lado, los metadatos en los archivos electrónicos, resultan un medio muy útil para la ingeniería social. A través de ellos se puede obtener cierta información de forma sencilla, costosa de conseguir por otros medios.
7. Para mitigar el riesgo asociado a los metadatos y datos ocultos, el Esquema Nacional de Seguridad (ENS) dispone la medida de protección [mp.info.6], que establece la necesidad de un Proceso de Limpieza de Documentos. No disponer de un proceso de limpieza de documentos adecuado, tal y como se indica en el apartado 5.7.6 del Anexo II del Esquema Nacional de Seguridad, puede perjudicar:
 - a) Al mantenimiento de la confidencialidad de información que no debería haberse revelado al receptor del documento.
 - b) Al mantenimiento de la confidencialidad de las fuentes y orígenes de la información, que no debe conocer el receptor del documento.
 - c) A la buena imagen de la organización que difunde el documento por cuanto demuestra un descuido en su buen hacer.

2. OBJETO

8. Este documento forma parte del conjunto de normas desarrolladas por el Centro Criptológico Nacional para la implementación del Esquema Nacional de Seguridad (CCN-STIC-800), siendo de aplicación para el Sector Público y teniendo como objeto la protección de los servicios prestados a los ciudadanos y entre las diferentes administraciones.
9. El objeto del documento es proporcionar una guía de 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.
10. Establece unas pautas de carácter general que puedan resultar de aplicación a entidades de distinta naturaleza, dimensión y sensibilidad, sin entrar en casuísticas específicas. Por ello, es de esperar que cada organización las particularice para adaptarlas a su entorno singular.

3. ALCANCE

11. El Esquema Nacional de Interoperabilidad en el ámbito de la Administración Electrónica (ENI), regulado por el RD 4/2010, de 8 de enero, y la Norma Técnica de Interoperabilidad de Documento Electrónico, aprobada por la resolución de 19 de julio de 2011, de la Secretaría de Estado para la Función Pública, establecen un esquema de metadatos con metadatos mínimos obligatorios que deben estar necesariamente presentes en cualquier proceso de intercambio de documentos

electrónicos, entre órganos de la Administración y Entidades de Derecho Público vinculadas o dependientes de aquella, y con el ciudadano.

12. El alcance de la presente guía, no abarca la inspección y borrado de estos metadatos pertenecientes al esquema de metadatos especificado en el ENI y en la Norma Técnica de Interoperabilidad de Documento Electrónico. Será cada organización, en función de los métodos empleados para la implementación del esquema de metadatos del ENI y de sus posibles escenarios de almacenado e intercambio de documentos electrónicos, la que deberá realizar el análisis correspondiente y determinar las medidas más adecuadas que minimicen los riesgos asociados a la información incluida en estos metadatos.
13. El alcance de la presente guía abarca la inspección y borrado de los metadatos y otros datos ocultos existentes en los documentos electrónicos, incorporados de forma automática por los programas de generación y tratamiento de estos documentos, o por los propios usuarios de la organización.
14. En esta guía no se hace referencia a ninguna herramienta gratuita ni comercial para la inspección y borrado de metadatos. No obstante, establece las funcionalidades y características recomendables para este tipo de herramientas, con objeto de facilitar la evaluación y selección de la herramienta más adecuada para cada organización. En la Guía CCN-STIC-818 de Herramientas de Seguridad en el ENS, se incluyen en su Anexo A varias herramientas orientativas para análisis y limpieza de metadatos.
15. Así mismo, se indican en esta guía las utilidades para inspección y borrado de metadatos, de las que disponen de forma integrada determinados programas de generación y tratamiento de documentos, por considerarlas muy extendidas en las organizaciones. Estas aplicaciones son Microsoft Office, OpenOffice.org y Adobe Acrobat.

4. METADATOS

4.1 Definición y Generalidades

16. La Sociedad Española de Documentación e Información Científica (SEDIC), define metadato como “toda aquella información descriptiva sobre el contexto, calidad, condición o características de un recurso u objeto de información que tiene la finalidad de facilitar su recuperación, autenticación, evaluación, preservación y/o interoperabilidad”.
17. En el caso de fotografías digitales, por ejemplo, la propia cámara a la vez que captura las imágenes puede ir guardando en forma de metadatos, información de cómo fue tomada la fotografía: fecha, hora, diafragma, velocidad, uso de flash, modo de captura o localización.
18. En el caso de archivos de audio o vídeo, los metadatos pueden almacenar información como el título de la obra, álbum, año, autor, carátula o género.
19. En el caso de documentos ofimáticos, los metadatos pueden almacenar

información de quién lo creó, quién lo modificó, quien realizó el último acceso al documento y las fechas correspondientes, tiempo que ha tardado en editarse el documento, dispositivo o software utilizado para la creación del documento, o compañía y departamento al que pertenece.

20. La principal razón para crear metadatos es facilitar la búsqueda de información relevante utilizando diversos criterios de búsqueda. Los metadatos pueden ayudar a organizar los documentos electrónicos, facilitar la interoperabilidad entre organizaciones, proveer la identificación digital y dar soporte a la gestión del ciclo de vida de los documentos.
21. Un esquema de metadatos, es una colección de elementos de metadatos diseñados para un propósito específico, como describir un tipo particular de documento electrónico. A la definición de cada elemento se le llama semántica del esquema y a los valores asignados a cada elemento se les llama contenido. Un esquema de metadatos normalmente especifica los nombres de los elementos y su semántica. También puede especificar reglas de contenido para indicar como debe formularse el contenido y reglas de sintaxis para indicar cómo deben codificarse los elementos y su contenido.
22. Normalmente las organizaciones disponen de un esquema de metadatos, que responderá a las particularidades y necesidades específicas de gestión de los documentos electrónicos de cada organización.
23. Un ejemplo de esquema de metadatos, es el Esquema de Metadatos para la Gestión del Documento Electrónico (e-EMGDE), disponible en el Centro de Interoperabilidad Semántica, que incluye los metadatos mínimos obligatorios, definidos en las Normas Técnicas de Interoperabilidad de Documento electrónico y Expediente electrónico, así como otros metadatos complementarios pertinentes en una política de gestión y conservación de documentos electrónicos.
24. Cada organización, dentro de las políticas implantadas para el desempeño de sus actividades, dispondrá de una Política de Gestión Documental en la que se establecerán los criterios y normas en relación con la gestión de los documentos electrónicos. Dentro de esta política se especificará el esquema de metadatos asociados a los documentos electrónicos para asegurar la gestión, recuperación y conservación de los mismos durante todo su ciclo de vida
25. Por otro lado, es importante indicar que tanto los dispositivos (ordenadores o cámaras, por ejemplo), como muchos de los programas de generación y tratamiento de documentos, insertan sus propios metadatos sin que en muchos casos el usuario sea consciente de ello.
26. Los metadatos normalmente se encuentran ocultos y no son visibles usando la configuración estándar de la aplicación con la que estemos trabajando sobre el archivo. Para visualizarlos es necesario establecer una configuración específica o incluso utilizar un software específico para revelar esos datos ocultos.
27. En el caso de los documentos ofimáticos, además de los metadatos, también puede existir otro tipo de información oculta en el propio contenido del

documento, como por ejemplo texto y objetos formateados como invisibles, datos fuera del área de visión del documento, o información relativa a comentarios y cambios de revisión e identidad de quién realizó cada uno de ellos. Esto normalmente se conoce como Información oculta o Datos ocultos y puesto que no es visible a simple vista, el usuario puede no ser consciente de su existencia y supone igualmente un riesgo en caso de que el documento sea distribuido a personal ajeno a la organización.

28. Por lo tanto, y dado que el canal de difusión de los archivos de información es cada vez mayor, las personas y las organizaciones deben establecer medidas para proteger su información privada y confidencial. Parte de esas medidas de protección exigen procedimientos y herramientas de revisión y limpieza de documentos y archivos, para minimizar el riesgo de que información sensible se revele a través de metadatos o datos ocultos.

4.2 Tipos de Metadatos

29. Tipos de metadatos según **su propósito**:

- a) *Metadatos descriptivos*, que describen aspectos del documento electrónico tales como título, asunto, autor o fechas de creación.
- b) *Metadatos estructurales*, que indican cómo se construye el documento a partir de sus componentes (por ejemplo, un fichero de texto principal y varios ficheros de textos complementarios o anexos).
- c) *Metadatos administrativos*, que proveen información que permitirá la gestión, recuperación y conservación de los documentos a lo largo de todo su ciclo de vida.

30. Tipos de metadatos según **sea su origen**:

- a) *Metadatos corporativos*, que son agregados a los documentos como resultado de aplicación del esquema de metadatos de la organización, definido en la Política de Gestión Documental para la gestión, recuperación y conservación de sus documentos y otros objetos de información a lo largo de todo su ciclo de vida.
- b) *Metadatos de usuario*, que son agregados a los documentos por el propio usuario que crea, modifica o revisa el documento para su propia gestión y seguimiento.
- c) *Metadatos de la aplicación*, que son agregados a los documentos de forma automática por los programas de generación y tratamiento de estos documentos y que en algunos casos se pueden modificar o eliminar y en otros casos no.

31. Tipos de metadatos según **sea el método mediante el que se asocian al documento**:

- a) *Metadatos embebidos en el documento*, que no pueden separarse del documento (por ejemplo, metadatos embebidos en documentos HTML, en cabeceras de archivos de imagen o en las propiedades de documentos). Este método de asociación tiene como ventajas asegurar que los metadatos no se van a perder, evitar los problemas derivados de enlazar los metadatos y los datos y asegurar que los metadatos y el documento serán actualizados simultáneamente.
- b) *Metadatos separados del documento*, que pueden ir en archivos adicionales o en bases de datos enlazadas con el documento. Este método de asociación tiene como ventaja que se simplifica la gestión de los metadatos y se facilita la búsqueda y recuperación de los documentos.

32. Tipos de metadatos según **el tipo de documento al que van asociados**:

- a) *Metadatos en Documento ofimáticos y PDF*

Este tipo de documentos, contienen metadatos embebidos a través de las propiedades del documento. Estos metadatos contienen información como: título, asunto, comentarios, etiquetas, autor, fechas de creación y modificación, fecha de impresión, último usuario que modificó el documento, tiempo de edición, estadísticas, etc.

Los metadatos en las propiedades de los documentos pueden ser estándar (metadatos prefijados por el programa) cumplimentados por el programa de forma automática, o de forma manual por el usuario o la organización. También pueden ser metadatos personalizados, que son tipos específicos de metadatos que el usuario o la organización crean y cumplimentan.

Estos documentos, además de contener metadatos en sus propiedades, pueden llevar asociados metadatos más específicos en diversos formatos (XMP, RDF, etc.), bien embebidos en el documento, bien separados de él (por ejemplo, en ficheros aparte que se denominan “sidecar files”).

- b) *Metadatos en Documentos de Imagen*

Los documentos que consisten en imágenes incluyen metadatos:

- *Descriptivos*, que serán introducidos por el que genera o gestiona las fotografías o imágenes para incluir información sobre el contenido de la imagen, como título, etiquetas, lugar de captura de la imagen, etc.
- *Técnicos*, que son generados por los dispositivos de captura de imagen (como las cámaras digitales). Son datos sobre la configuración usada en el dispositivo para la captura de la imagen (por ejemplo, tiempo de exposición, distancia focal, modo de flash, velocidad ISO, etc.).

- *Administrativos*, que incluyen información administrativa como licencias, propietario del copyright, restricciones de uso de la imagen, información de contacto, etc.

Los metadatos se pueden incluir embebidos formando parte del archivo de imagen o almacenados en un fichero aparte (sidecar file).

c) *Metadatos en Documentos multimedia*

Los metadatos en los documentos multimedia (por ejemplo audios o vídeos) se utilizan normalmente para su catalogación, ya que la clasificación únicamente por carpetas y nombres de fichero es insuficiente para grandes colecciones y es necesario disponer de más criterios de clasificación y búsqueda, que proporcionan los metadatos a través de información como: título, artista, álbum, discográfica, número de pista, compositor, copyright, propietario del copyright, propietario de los derechos de publicación, fecha de edición, género, idioma, y muchos más datos.

33. Hay multitud de esquemas, modelos o estándares de metadatos según **el estándar o formato** utilizado. Se incluye a continuación una tabla en la que se indica para cada tipo de documento, los formatos de metadatos más empleados.

TIPO DE DOCUMENTO	EXTENSIÓN ARCHIVO	FORMATO DE METADATOS
DOCUMENTOS OFIMÁTICOS		
Documentos Microsoft Office: Word®, Excel®, PowerPoint®	.doc /.docx /.docm .xls /.xlsx /.xlsm .ppt /.pptx /.pptm	Propiedades del Documento
Documentos OpenDocument: texto, hojas de cálculo, presentaciones, gráficos	.odt /.ods /.odp /.odg	Propiedades del Documento
DOCUMENTOS PDF		
Documentos PDF	.pdf	Propiedades del Documento Metadatos XMP
IMÁGENES		
Imagen JPEG, TIFF, PSD, Raw	.jpeg /.jpg .tiff /.tif .psd Formatos raw	IPTC (IIM / Core /Extension)
		EXIF
		XMP
Imagen PNG	.png	XMP
ARCHIVOS MULTIMEDIA		
Vídeo AVI	.avi	INFO; XMP
Audio MP3	.mp3	ID3v1 tag; ID3v2 tag;
		APE tag; XMP
Archivo MP4	.mp4 /.m4a /.m4v	XMP
Archivo F4V	.f4v	XMP
Audio WAVE / BWF	.wav	INFO; XMP; ID3v2 tag;

TIPO DE DOCUMENTO	EXTENSIÓN ARCHIVO	FORMATO DE METADATOS
		Broadcast Audio Extension; iXML; Cart Chunk; ISRC
Audio AIFF	.aif	ID3v2 tag; XMP
Audio WavPack	.wv	ID3v1 tag; APE tag
Audio Monkey's	.ape	APE tag
Audio Musepack	.mpc /.mpp /.mp+	APE tag
Audio OptimFROG	.ofr /.ofs	APE tag
Audio comprimido Tom's Audio	.tak	APE tag
Extensible Metadata Platform File	.xmp	XMP

Tabla 1.- Formatos de metadatos comunes según los tipos de Documento.

4.3 Riesgos y Amenazas

34. Como se ha comentado ya en apartados anteriores, los metadatos son una fuente de riesgo, ya que pueden contener información sensible que no debe revelarse a personal ajeno a la organización. Por ello, es necesario que las organizaciones y los usuarios sean conscientes del riesgo que supone la fuga de esta información sensible, como datos de clientes, propiedad intelectual, detalles financieros o cualquier otra información que dar a conocer resulte un inconveniente para la organización.
35. La siguiente figura muestra un ejemplo del impacto que podría causar el exponer cierta información almacenada en los metadatos de un documento.



Figura 1.- Ejemplo de un documento que revela información sensible a través de los metadatos de las propiedades del documento.

36. Como vemos en la figura, las implicaciones y la severidad del riesgo, varía dependiendo del tipo de información que pueda ser desvelada o deducida. En el mejor de los casos, sólo dañará la reputación de la organización (por ejemplo, en caso de que el cliente deduzca que ha recibido un documento cuyo contenido ha sido copiado de otro). En el peor de los casos, podría conducir a invalidar contratos, pleitos, sanciones o causar serios perjuicios a la organización.
37. La ingeniería social utiliza multitud de métodos y técnicas, y los metadatos y datos ocultos, son un medio muy útil para estos fines, ya que de forma relativamente sencilla se puede extraer gran cantidad de información valiosa sobre la organización para utilizar en posteriores ataques.
38. La ingeniería social en el contexto de la seguridad de la información, se puede definir como el arte de averiguar información sensible y/o manipular a los individuos para realizar ciertas acciones, que resulten en una brecha en la seguridad de la organización.
39. En el caso de los documentos de imagen, los metadatos pueden contener información sobre el dispositivo que ha tomado la fotografía (por ejemplo, el tipo de smartphone y su sistema operativo) o la localización geográfica del lugar donde se tomó la fotografía (en el caso de los dispositivos que disponen de GPS). La localización es una información muy sensible, ya que si la imagen se publica a través de algún medio en Internet podría desvelarse la ubicación y dar publicidad, por ejemplo, a alguna infraestructura crítica de la organización.
40. En el caso de los documentos ofimáticos y PDF, los metadatos y datos ocultos pueden contener información como: nombre, iniciales o incluso nombre de usuario (username) que ha creado o modificado el documento, nombre del ordenador, su sistema operativo y el programa que ha creado el documento, direcciones de correo electrónico, etc. De esta manera, podrían utilizarse estos datos para realizar diferentes acciones:
- A través de los nombres de empleado y complementando con la búsqueda en redes sociales (por ejemplo, LinkedIn), se puede obtener todo un listado de empleados de la organización, sus cargos, e incluso sus correos electrónicos, lo cual puede servir para ataques de phishing.
 - A través del sistema operativo y las aplicaciones que utilizan los ordenadores, se puede conocer el entorno tecnológico de la organización y realizar ataques dirigidos más efectivos.
 - A través de los nombres de usuario (usernames), se puede deducir la convención de nombres empleada en la organización y componer direcciones de correo para ataques de phishing o intentar ataques de fuerza bruta.
41. A continuación, se incluye una tabla con algunos metadatos y datos ocultos que pueden estar presentes en los documentos y sus riesgos asociados.

TIPO DE METADATO O DATO OCULTO	DESCRIPCIÓN	RIESGOS
Propiedades Estándar	Datos que contienen información descriptiva del documento, como Título, Asunto, Etiquetas, Categoría, Comentarios, Autor, Administrador, Compañía, Plantilla, etc.	Uno de los riesgos asociados consiste en revelar información sobre el documento que pueda resultar sensible. Por ejemplo, los nombres de usuario (Autor, Administrador o última persona que modificó el documento) son información personal que afecta a la privacidad. También supone un riesgo que el documento cuente con información obsoleta o incorrecta en estos datos y el documento se reutiliza para otros fines. Finalmente está el riesgo asociado a la ingeniería social y sus capacidades de obtener información a través de estos datos.
Propiedades Personalizadas	Datos insertados a medida por los usuarios o por la organización. Por ejemplo, Cliente, Departamento, División, Identificador del documento, Proyecto, etc.	Normalmente se utilizan para incluir metadatos corporativos especificados por la Política de Gestión documental de la organización y sirven para posibilitar la gestión del documento a lo largo de su ciclo de vida. El riesgo asociado a estos metadatos depende de su contenido y de la información sensible que incluyan y no deba ser revelada, pero sí representan un elevado riesgo en relación con la ingeniería social, ya que de ellos se puede deducir información muy valiosa sobre la organización.
Estadísticas	Datos que contienen información de estadísticas del documento como número de páginas, líneas, palabras, tiempo de edición, etc.	El riesgo asociado es muy bajo, ya que las estadísticas almacenan detalles de edición del documento y no revelan información sensible.
Fechas	Datos que contienen fechas asociadas al documento: Fecha de creación, de última modificación, de última impresión, etc.	El riesgo asociado consiste en que revelar información asociada a las fechas del documento, puede resultar en algunas circunstancias un gran inconveniente para la organización.
Versiones	Las versiones de un documento se pueden almacenar en un mismo fichero. Esto permite recuperar una versión anterior del documento.	El riesgo consiste en dejar accesibles versiones anteriores del documento que no fueron creadas para su distribución.

TIPO DE METADATO O DATO OCULTO	DESCRIPCIÓN	RIESGOS
Rutas de archivos	Los documentos pueden contener otros archivos insertados (por ejemplo archivos multimedia) y almacenar la ruta completa del archivo.	El riesgo de que el documento contenga la ruta de red a un fichero, supone proporcionar una vista de la topología de red de la organización y desvelar nombres de carpetas que podrían contener información sensible (por ejemplo, nombres de clientes). Este dato es muy valioso para la ingeniería social, ya que revela información sensible sobre el entorno tecnológico de la organización permitiendo diseñar ataques dirigidos.
Histórico de Autores del documento	Los documentos pueden contener los nombres de los últimos autores que guardaron el documento, almacenados en un área del documento inaccesible a través de la aplicación.	El riesgo asociado es la exposición de información personal (normalmente nombres de usuario) y la visibilidad de rutas locales o de red que indiquen dónde se almacenó cada versión del documento. Claramente, esta información es muy valiosa para la ingeniería social.
Comentarios y registros de revisión	Los documentos pueden contener comentarios y registros de cambios, insertados por las personas que han llevado a cabo la revisión del documento.	En función del tipo de comentario o registro de revisión, el riesgo puede resultar menor (si por ejemplo el objetivo es clarificar algún texto) o mayor (si se trata de discusiones internas).
Conexiones a Bases de Datos	Los documentos pueden contener conexiones a Bases de Datos desde las que se importan datos. Estas conexiones pueden contener los datos del Servidor de Base de datos, nombres y contraseñas de conexión y sentencias de consulta a la Base de Datos.	Los datos de la conexión son claramente datos sensibles, ya que posibilitan a un usuario no autorizado, realizar consultas de forma independiente a una Base de Datos que puede contener datos sensibles. Además, las sentencias de consulta también tienen el riesgo de ser utilizadas para deducir la estructura de la Base de Datos.
Objetos Insertados	Los documentos pueden contener otros documentos insertados.	Los documentos insertados (archivos origen) pueden traer consigo metadatos no visibles en el documento en el que se insertan (archivo destino). El riesgo radica en que estos metadatos contengan información sensible.
Objetos Ocultos	Los documentos pueden contener objetos ocultos, como imágenes, gráficos, textos, hojas de cálculo, diapositivas, etc.	Normalmente el documento contiene objetos ocultos porque el usuario los ha ocultado intencionadamente. Por lo tanto contienen información sensible que debe ser eliminada antes de distribuir el documento.

TIPO DE METADATO O DATO OCULTO	DESCRIPCIÓN	RIESGOS
Macros y Código	Los documentos pueden contener macros u otro tipo de código (por ejemplo en Visual Basic).	Las macros y otro tipo de código, además de la posibilidad de distribuir código malicioso, también pueden contener información sensible dentro del código (nombres de usuario, comentarios y líneas de código confidenciales para acceder a los recursos corporativos).
Información de Impresoras	Los documentos pueden almacenar información de la impresora de red utilizada.	El riesgo asociado consiste en que se está revelando información de la ruta de la impresora de red y puede que también datos de la impresora (fabricante y modelo). Al igual que otras rutas, estos son datos valiosos para la ingeniería social, que puede deducir la topología y entorno tecnológico de la organización.
Hipervínculos	Los documentos pueden contener hipervínculos.	El riesgo asociado es que los hipervínculos sean enlaces a recursos ubicados en la red interna de la organización, y se pueda desvelar a través de la ruta, la topología de red de la organización o información sensible de nombre de carpetas.

Tabla 2.- Metadatos y datos ocultos en los documentos y sus riesgos asociados.

5. MEDIDAS DE SEGURIDAD EN EL ENS

42. El Real Decreto 3/2010, de 8 de enero, por el que se regula el Esquema Nacional de Seguridad en el ámbito de la Administración Electrónica (ENS), modificado a su vez por el Real Decreto 951/2015, de 23 de octubre, es de aplicación por las Administraciones Públicas y tiene como objetivo asegurar el acceso, confidencialidad, integridad, autenticidad, trazabilidad, disponibilidad y conservación de los datos, informaciones y servicios utilizados en los medios electrónicos que gestionen el ejercicio de sus competencias.
43. 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”.
44. Dentro de estas medidas de seguridad definidas en el ENS, las medidas de protección [mp] “se centran en proteger activos concretos, según su naturaleza y la calidad exigida por el nivel de seguridad de las dimensiones afectadas”.
45. 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).

46. 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.
47. 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”*.
48. El primer requisito que se extrae por lo tanto de la medida, es la necesidad de que exista en la organización un **proceso de limpieza de documentos** en el que deberán tenerse en cuenta no sólo los metadatos, sino también otro tipo de información o datos ocultos que puedan contener los documentos. Y que, tal y como se indica en la [mp.info.6], dicho proceso deberá aplicarse especialmente “cuando el documento se difunde ampliamente, como ocurre cuando se ofrece al público en un servidor web y otro tipo de repositorio de información”.
49. El Proceso de limpieza de documentos deberá indicar las herramientas a emplear en función del tipo de documento, y cómo se utilizará cada una de ellas para eliminar el metadato o dato oculto deseado. Deberá incluir el detalle de las acciones a llevar a cabo para completar la limpieza del documento, y también las acciones necesarias para verificar que la limpieza se ha completado de forma efectiva.
50. La medida de protección [mp.info.6] no especifica, sin embargo, cuáles son los metadatos o datos ocultos que deben eliminarse. Esto dependerá, por un lado, de las circunstancias en las que el documento va a ser almacenado o distribuido y, por otro lado, también dependerá de las necesidades, criterios y normativa específica de la organización. Deberá ser la Política de Gestión Documental de la organización la que defina por lo tanto qué metadatos y datos ocultos no deben estar presentes en los documentos según los escenarios siendo el procedimiento de limpieza de documentos el que especifique cómo eliminar estos metadatos.
51. En muchos casos puede no resultar conveniente mantener metadatos relacionados con información personal (PII), como el autor, o el usuario o usuarios que realizaron cambios en el documento, especialmente cuando tales datos hayan sido introducidos de forma automática por los dispositivos o los programas, sin intervención del usuario o de la organización. Serán este tipo de metadatos los que exijan una especial revisión dentro del proceso de limpieza de documentos de la organización.
52. Por otro lado, existen metadatos, como los relacionados con la información de copyright, que no deben eliminarse. En la advertencia de copyright, el propietario reivindica que el documento no puede usarse sin autorización por ninguna otra persona u organización. Los metadatos de copyright, por lo tanto, deben

permanecer en los documentos para indicar a otros, los derechos de uso del archivo.

53. Además de la medida de protección de la información [mp.info.6], que exige un proceso de limpieza de documentos en la organización, existen también otras medidas dentro del ENS que emanan de ésta o se relacionan con ella, tal y como se indica en la siguiente figura.

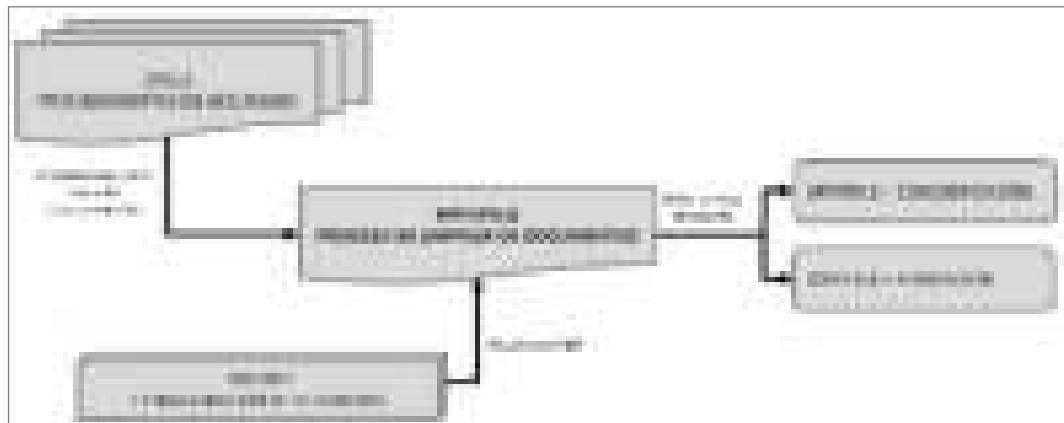


Figura 2.- Medidas del ENS relacionadas.

54. Medida [op.exp.2] - Configuración de seguridad.

Esta medida de seguridad, perteneciente al conjunto de medidas del marco operacional definidas en el ENS, y de aplicación a todos los sistemas independientemente de su categoría, establece entre sus directrices la necesidad de aplicar las reglas de “seguridad por defecto” y “mínima funcionalidad” a los sistemas. Si extendemos estas reglas a los documentos electrónicos, se derivan los siguientes requisitos:

- Los documentos deberán contener únicamente la información requerida para que la organización alcance sus objetivos, y ninguna otra información.
- Se eliminará o desactivará mediante control de la configuración, el almacenado de aquella información que no sea de interés, no sea necesaria, e incluso sea inadecuada para el fin que se persigue.
- Las medidas de seguridad activadas sobre los documentos serán respetuosas con el usuario y protegerán a éste, salvo que se exponga conscientemente a un riesgo.
- Para reducir la seguridad en los documentos, el usuario tiene que realizar acciones conscientes.
- El uso natural, en los casos que el usuario no ha consultado el procedimiento, será un uso seguro.

Estos requisitos sobre los documentos electrónicos, están alineados con el objetivo que se persigue con el proceso de limpieza de documentos especificado en la [mp.info.6].

De ellos se deriva la necesidad de establecer la configuración de seguridad en los dispositivos y programas de generación y tratamiento de documentos, que favorezca el cumplimiento de estos requisitos en los documentos electrónicos y ayude, por lo tanto, al cumplimiento de la [mp.info.6].

55. Medida [org.3] - Procedimientos de seguridad.

Esta medida de seguridad, perteneciente al conjunto de medidas del marco organizativo definidas en el ENS y de aplicación a todos los sistemas independientemente de su categoría, establece que la organización deberá disponer de una serie de procedimientos de seguridad y los requisitos que éstos deberán cumplir.

Puesto que de las medidas de Limpieza de Documentos [mp.info.6] y Configuración de Seguridad [op.exp.2] se derivan dos (2) procedimientos, éstos deberán formar parte de los procedimientos de seguridad y deberán cumplir los requisitos establecidos en esta medida, por lo que “deberán detallar de forma clara y precisa”:

- a) Cómo llevar a cabo las tareas habituales.
- b) Quién debe hacer cada tarea.
- c) Cómo identificar y reportar comportamientos anómalos.

56. Medida [mp.per.3] - Concienciación y [mp.per.4] - Formación

Estas medidas de seguridad, pertenecientes al conjunto de medidas de protección definidas en el ENS y relacionadas con la gestión del personal, establecen requisitos sobre la concienciación y la formación continua y regular del personal. Para lograr el cumplimiento de estos requisitos, deberá tenerse en cuenta la limpieza de documentos especificada en la medida [mp.info.6].

Dentro de las acciones de concienciación, deberá incluirse la concienciación a los usuarios en la necesidad de limpieza de documentos, y en los riesgos asociados con revelar información sensible o confidencial, a través de metadatos incluidos en documentos que no han sido saneados adecuadamente. Tal y como se indica en el ENS, “esta medida es especialmente relevante cuando el documento se difunde ampliamente, como ocurre cuando se ofrece al público en un servidor web y otro tipo de repositorio de información”.

Igualmente, dentro de las acciones de formación, es necesario llevar a cabo acciones de formación relativas a la adecuada gestión de los documentos electrónicos. Especialmente deberán conocer el proceso de limpieza de documentos y cómo y con qué herramientas deberán llevarse a cabo.

57. Otras acciones recomendables que pueden ayudar a lograr el objetivo de limpieza de documentos, y con ello a mitigar el riesgo asociado a revelar información a

través de los metadatos, son las siguientes:

- Utilizar herramientas evaluadas para realizar la inspección y borrado de metadatos.
- Configurar en las herramientas de seguridad de la organización, tales como cortafuegos, IPS/IDS, etc., las reglas adecuadas para detectar y bloquear los escaneos cuyo fin sea la extracción y análisis de metadatos.
- Realizar supervisiones periódicas de las actividades realizadas por el personal y de los documentos generados, para verificar el cumplimiento y la efectividad del proceso de limpieza, siempre que esté permitido por la legislación o las normativas vigentes.
- Existe una categoría especial de herramientas de inspección de metadatos, cuyo objetivo es recopilar y examinar los metadatos de los archivos de información que se le indiquen y sacar conclusiones a partir de ellos. Las llamadas Herramientas de Análisis de Metadatos del Apartado 6 de esta guía.

58. Es necesario hacer énfasis en que el equipo del usuario, ya sea un ordenador o dispositivo móvil, debe cumplir con la Normativa de seguridad vigente en la organización de tal manera que implemente una configuración de seguridad apropiada que permita un uso seguro de los equipos y dispositivos empleados. Por otro lado, hay que tener en cuenta las recomendaciones recogidas en las guías CCN-STIC pudiendo ser de aplicación la CCN-STIC-827 sobre Gestión y uso de dispositivos móviles o las series CCN-STIC-850 y 899 sobre la implantación del ENS para sistemas MS Windows.
59. Por último, son de aplicación las recomendaciones realizadas en la presente guía, relativas a la configuración de seguridad apropiada en los dispositivos y programas de generación y/o tratamiento de documentos electrónicos, en relación con el almacenamiento automático de metadatos.

6. HERRAMIENTAS DE INSPECCIÓN Y BORRADO DE METADATOS

60. En general, los dispositivos y los programas dificultan la visualización, edición o borrado de los metadatos adjuntos en los documentos, ya que muchas veces estos metadatos se encuentran en lugares que no son fáciles de alcanzar por los usuarios habituales.
61. Algunos sistemas operativos, como por ejemplo Microsoft Windows, permiten visualizar determinados metadatos contenidos en los documentos de forma sencilla, basta con seleccionar el archivo, hacer clic con el botón derecho del ratón y seleccionar Propiedades.
62. Algunos programas incluyen utilidades para visualizar, editar o eliminar metadatos (por ejemplo, el Inspector de Documentos de Microsoft Office para Windows a partir de la versión 2007 o la utilidad de Eliminar Información oculta de Adobe Acrobat X).
63. También existe una variedad bastante amplia, de programas gratuitos, y programas

comerciales diseñados para inspeccionar, editar y eliminar metadatos en diversos formatos para multitud de tipos de documentos. La Guía CCN-STIC-818 de Herramientas de Seguridad en el ENS, en su Anexo A incluye varias de estas herramientas.

64. Normalmente, estas herramientas de inspección y borrado de metadatos son automáticas, es decir, permiten aplicar una configuración específica de manera uniforme a toda la organización, para cumplir con los requisitos establecidos en la Política de Gestión Documental sobre la presencia de metadatos, y no requieren de la intervención de los usuarios.
65. El uso de las herramientas de inspección y borrado de metadatos puede reportar grandes beneficios a la organización:
- Reducción del riesgo, al depurar los metadatos de los documentos de forma automática antes de que puedan ser distribuidos fuera de la organización evitando costes financieros o daño en la reputación.
 - Incremento de la seguridad, ya que previenen de revelar información privada o sensible.
 - Ahorro de tiempo, ya que al ser automáticas evitan repetir las actividades que supone depurar de forma manual los documentos.
 - Cumplimiento de normativas y regulaciones y cumplimiento de la Política de Gestión Documental de la organización.

6.1 Tipos de Herramientas

66. Tipos de herramientas según **su propósito**:

- a) *Inspección y borrado de metadatos*, cuyo objetivo es inspeccionar los metadatos asociados a los documentos, visualizarlos, modificarlos o eliminarlos.
- b) *Búsqueda de Metadatos*, cuyo objetivo es buscar documentos a partir de los valores de sus metadatos.
- c) *Análisis de Metadatos*, cuyo objetivo es analizar los metadatos de los documentos para inferir datos a partir de ellos.
- d) Dentro de estas herramientas, merecen especial mención aquellas especializadas en la *búsqueda, inspección y análisis de metadatos* de documentos publicados en Internet. Son herramientas especializadas en usar varios buscadores para encontrar y descargar todos los documentos de un dominio web, extraer sus metadatos y llevar a cabo el análisis para deducir información útil, como por ejemplo documentos creados desde un mismo equipo, y qué servidores y clientes se pueden inferir de ellos.

Es habitual que una misma herramienta cumpla varios propósitos de los anteriores. La organización debe disponer al menos, de herramientas de inspección y borrado de metadatos, para llevar a cabo las actividades de revisión y limpieza de metadatos de sus documentos, necesarias para el proceso de limpieza de documentos.

67. Tipos de herramientas según el tipo de documento y formato de metadatos que tratan:

- a) *Metadatos en documentos ofimáticos y PDF.* Herramientas especializadas en el tratamiento de metadatos de documentos ofimáticos (generados por programas ofimáticos como Microsoft Office o Apache OpenOffice) y PDF, tanto incrustados en el propio contenido del documento, como separados en otro archivo.

Estas herramientas normalmente además de los metadatos, también inspeccionan y eliminan información o datos ocultos en los documentos, e incluso pueden revisar, modificar o censurar su contenido siguiendo patrones y reglas.

- b) *Metadatos en documentos de imagen.* Herramientas especializadas en el tratamiento de metadatos de documentos de imágenes con formatos de metadatos especializados como EXIF, XMP, IPTC-IMM, IPTC Core & Extensión, etc.
- c) *Metadatos en documentos multimedia.* Herramientas especializadas en el tratamiento de metadatos en documentos multimedia como audio o vídeo con formatos de metadatos especializados como XMP, APE tag, ID3 tag, etc.

Es habitual que una herramienta de tratamiento de metadatos trate varios tipos de documentos y formatos de metadatos.

68. Según **su arquitectura**:

- a) *Basadas en Cliente (Client Based).* Herramientas que se instalan y operan en la parte cliente, es decir, localmente en los equipos de los usuarios.
- b) *Basadas en Servidor (Server Based).* Herramientas que se instalan y operan en la parte servidor (servidores de archivos, servidores de correo electrónico, etc.) y que dan servicio de forma centralizada a todos los usuarios.

69. Según **su modo de operación**:

- a) *Inspección de documentos almacenados (data at rest).* Herramientas que inspeccionan documentos que se encuentran almacenados (en servidores de ficheros, estaciones de trabajo, sistemas de almacenamiento, etc.), permitiendo seleccionarlos para buscar y eliminar metadatos. Algunas de ellas tienen capacidades para procesar varios documentos simultáneamente (lo que se llama procesamiento *batch*).

Estas herramientas pueden funcionar de forma automática eliminando todos los metadatos y datos ocultos que encuentren en su inspección, y según las

reglas que se hayan configurado en la herramienta, o bien pueden informar al usuario o administrador de los datos encontrados y requerir su intervención para proceder a eliminarlos.

- b) *Inspección de documentos en tránsito (data in transit)*. Herramientas que se instalan habitualmente en los servidores de correo electrónico de la organización (basadas en servidor) y permiten configurar reglas para realizar la inspección y borrado de los metadatos incluidos en los documentos adjuntos de todos y cada uno de los correos electrónicos de la organización, antes de realizar su envío al destino.

Estas herramientas se configuran de acuerdo a un conjunto de reglas establecidas por la organización de forma centralizada dictadas por su Política de Gestión Documental.

Existen otras herramientas con la misma finalidad que se instalan en el cliente de correo electrónico (basadas en cliente), de forma que cuando el usuario envía un correo electrónico que lleva adjunto un documento, la herramienta lo detecta e informa al usuario ofreciendo la posibilidad de inspeccionar el documento en busca de metadatos para depurarlo.



Figura 3.- Diagrama ejemplo del funcionamiento de una herramienta de inspección y borrado de metadatos para correo electrónico.

- c) Ambas herramientas pueden funcionar de forma automática, realizando la inspección y depuración del documento adjunto según las reglas configuradas, o bien pueden funcionar de forma manual y avisar al usuario de la presencia del documento adjunto y requerir su intervención para realizar la inspección y eliminación de los metadatos.
- d) Estas herramientas también pueden admitir excepciones y disponer, por ejemplo, de listas blancas, de forma que cuando los mensajes vayan dirigidos a determinados destinatarios no se realice la inspección o permitir que los

usuarios impidan la inspección de ciertos correos especificando algún tipo de etiqueta en el asunto del correo.

70. Según el tipo de dispositivo:

- a) *Herramientas para Ordenadores*. Herramientas diseñadas para funcionar en PC o estaciones de trabajo (herramientas basadas en cliente) o servidores (herramientas basadas en servidor). Constituyen la mayor parte de las herramientas de inspección y borrado de metadatos y pueden tratar muchos formatos de metadatos y documentos.
- b) *Herramientas para Smartphones y Tablets*. Existen muchas herramientas de inspección y borrado de metadatos para documentos de imagen, disponibles para iOS (desde iTunes) y Android (desde Google Play). Algunas son gratuitas y otras son comerciales. Para la inspección de documentos ofimáticos y PDF, existen herramientas para ordenadores que disponen de clientes para su uso en dispositivos móviles.

6.2 Características de las Herramientas

71. A continuación, se relacionan una serie de características recomendables, que las organizaciones deben tener en cuenta a la hora de seleccionar una herramienta de inspección y borrado de metadatos adecuada para el Proceso de limpieza de documentos.

72. Integración con el Sistema de Gestión documental de la organización.

Es recomendable que la herramienta pueda integrarse con el sistema de gestión documental que se utilice en la organización, para la gestión del ciclo de vida de los documentos.

De esta forma, la herramienta participará en el flujo de trabajo del documento y así, por ejemplo, cuando un documento pase de borrador a versión definitiva, cuando sea exportado del sistema o cuando finalice su ciclo de vida, será automáticamente inspeccionado y sus metadatos serán depurados según los requisitos especificados en la Política de Gestión Documental, que definirán múltiples criterios en función del tipo y clasificación de los documentos, o de los usuarios que lleven a cabo determinadas acciones.

73. Flexibilidad de reglas.

Es recomendable que la herramienta tenga flexibilidad, a la hora de configurar las reglas bajo las que realizará la inspección y depuración de los metadatos del documento.

Debe permitir la implementación de reglas sofisticadas para ofrecer mayores posibilidades de satisfacer los requisitos de la Política de Gestión Documental, siempre manteniendo el equilibrio con la simplicidad y usabilidad que faciliten el uso de la herramienta por parte de los usuarios y administradores.

74. Flexibilidad en Formatos de documentos.

Es recomendable que la herramienta pueda realizar la inspección y depuración de metadatos de los formatos de documentos más extendidos (al menos Microsoft Office, OpenOffice y PDF). También es recomendable que disponga de compatibilidad con versiones obsoletas (por ejemplo, Microsoft Office 2000) ya que existirán muchos documentos creados en el pasado que pueden necesitar ser inspeccionados y depurados. También es recomendable que tenga capacidad de detectar metadatos, aunque los documentos se encuentren en formato comprimido (en un ZIP, por ejemplo).

75. Flexibilidad en Formatos de Metadatos.

Es recomendable que la herramienta pueda realizar la inspección y depuración de metadatos tanto incrustados en los documentos a examinar, como introducidos en ficheros separados y en múltiples formatos: EXIF, IPTC, XMP, Dublin Core, etc.

76. Soluciones basadas en Servidor.

Es recomendable que la solución se base en Servidor, de forma que no sea necesaria la instalación de software en cada uno de los equipos de los usuarios, lo cual además de añadir complejidad a su instalación y despliegue, podría elevar el coste con mucha probabilidad.

77. Independencia de la Plataforma.

Es recomendable que la herramienta sea independiente de la plataforma, y sea compatible con los sistemas operativos más extendidos para equipos y estaciones de trabajo de usuarios, sistemas de virtualización de escritorio, servidores o sistemas de correo electrónico (incluidos webmails).

78. Simplicidad y Usabilidad.

Es recomendable que la herramienta sea sencilla de manejar e intuitiva, lo que facilitará su manejo a los usuarios y administradores y reducirá costes y tiempos de formación.

79. Inspección de información o datos ocultos.

Es recomendable que la herramienta tenga capacidad para además de los metadatos, inspeccionar y depurar también información o datos ocultos, como texto u objetos invisibles, capas ocultas, comentarios, rutas de archivos, etc.

80. Tratamiento masivo de documentos y escalabilidad.

Es recomendable que la herramienta disponga de capacidades para el tratamiento masivo de documentos y que sea escalable, de forma que permita desde inspeccionar y depurar un solo documento solicitado por un usuario, hasta inspeccionar un número elevado de documentos simultáneos solicitados por el administrador del sistema.

También deberá ser capaz de inspeccionar y tratar documentos de gran tamaño, al menos del máximo tamaño de documento usado en la organización.

81. Automática.

Es recomendable que la herramienta tenga capacidad para funcionar de forma automática gobernada por las reglas configuradas, sin necesidad de intervención del usuario.

82. Análisis de resultados y elaboración de informes.

Es recomendable que la herramienta disponga de capacidad de análisis de los resultados de la inspección realizada sobre los documentos, indicando al menos los metadatos encontrados y el riesgo asociado a cada uno de ellos y permitiendo exportar estos resultados a Informes.

83. Presentación de mensajes a los usuarios.

Es recomendable que la herramienta tenga capacidad de mostrar mensajes a los usuarios. De esta forma, por ejemplo, se puede enviar un mensaje recordatorio de la necesidad de depurar los metadatos de un documento antes de enviarlo por correo electrónico, cuando haya detectado de forma automática un correo con archivo adjunto, o para transmitirles cualquier otra información relevante.

84. Inspección de contenido.

Otra característica deseable es que la herramienta tenga capacidades de inspección del contenido de documentos, de forma que se puedan especificar expresiones regulares (palabras, frases o párrafos) que supongan información sensible, para que la herramienta las detecte de forma automática y active la correspondiente alarma.

ANEXO A. METADATOS EN DOCUMENTOS MICROSOFT OFFICE

85. Toda la información recogida en este anexo, aplica a los documentos generados por los programas de Microsoft Office: Word, Excel y PowerPoint en versiones 2010, 2013 y 2016. Serán estos documentos a los que se haga referencia como “documentos Office” a lo largo del anexo.

1. Tipos de Metadatos e Información oculta

86. 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.

87. La información de Propiedades del documento contiene los siguientes datos:

- a) *Metadatos generados de forma automática por los programas de Office.* Incluyen datos y estadísticas que se generan y mantienen de forma automática por los programas Office, tanto características propias del fichero como son el tamaño, fechas de creación y de última modificación, o localización del documento (para versiones de Excel 2013 o superiores), como estadísticas (número de palabras o de páginas).

Esta información no puede ser modificada por el usuario y puede resultar especialmente sensible ya que, al ser generada de forma automática, el usuario puede no ser consciente de su existencia y con ella puede revelar información que no desea difundir (por ejemplo, el Autor del documento).

- b) *Metadatos generados manualmente.* Son propiedades cuyo valor se rellena de forma manual. Incluyen por un lado metadatos estándar predefinidos por la aplicación (como por ejemplo Título, Asunto, Etiquetas, Comentario, etc.). Por otro lado, también incluyen metadatos personalizados, que definen un tipo específico de metadatos representados por una etiqueta y un valor.

Estos metadatos son creados por el usuario o por la organización en aplicación de la Política de Gestión Documental para la gestión, recuperación y conservación de los documentos. Pueden contener información como el identificador del documento, departamento, unidad, código de la oficina, número de expediente asociado, advertencias de seguridad, etc.

88. A continuación, se muestra una tabla con los metadatos existentes en las propiedades del documento, indicando si se generan de forma automática por el programa, o de forma manual por el usuario u organización.

	Generado de forma Automática	Generado de forma Manual
Tamaño	X	
Número de Páginas ⁽¹⁾	X	
Número de Palabras ⁽¹⁾	X	
Diapositivas ⁽¹⁾	X	
Diapositivas ocultas ⁽¹⁾	X	

	Generado de forma Automática	Generado de forma Manual
Notas ⁽¹⁾	X	
Clips Multimedia ⁽¹⁾	X	
Formato de Presentación ⁽¹⁾	X	
Tiempo de Edición ⁽²⁾	X	
Título		X
Etiquetas		X
Comentarios		X
Plantilla ⁽²⁾	X	
Estado		X
Categorías		X
Asunto		X
Base de hipervínculo		X
Compañía		X
F/H última modificación	X	
F/H creación	X	
F/H última impresión	X	
Administrador		X
Autor	X ⁽³⁾	
Último modificador	X	
Número de Revisiones ⁽²⁾	X	

(1) Estas propiedades estarán presentes o no, dependiendo del tipo de documento Office del que se trate. Por ejemplo, el Número de páginas sólo estará presente en documentos Word.

(2) Estas propiedades no aplican a documentos Excel.

(3) El Autor es una propiedad creada automáticamente por los programas Office, pero que puede ser modificada o incluso eliminada por el usuario.

Tabla 3. Propiedades de los documentos Microsoft Office 2010 (Word, Excel y PowerPoint).

89. La información de Propiedades de un documento Office, se puede visualizar de la siguiente forma:

- Con el archivo abierto, hacer clic en la pestaña Archivo y a continuación clic en Información para ver las propiedades del documento, que aparecerán en el lado derecho de la pantalla.
- Para ver más propiedades, hacer clic en Mostrar todas las propiedades. Para ver menos propiedades, hacer clic en Mostrar menos propiedades.
- Para ver las propiedades personalizadas, hacer clic en Propiedades y seleccionar Propiedades avanzadas. Se abrirá una ventana en la que se muestran todas las propiedades del documento agrupadas en varias pestañas. Las propiedades personalizadas estarán en la pestaña Personalizar.
- En caso de que se hayan añadido metadatos particulares de la organización o si el documento ha sido almacenado en un servidor de documentos, posiblemente habrá más vistas adicionales disponibles.

90. Por otro lado, dentro del propio documento también puede existir información

oculta, es decir, información que no es visible a simple vista y de la cual el usuario puede no ser consciente. Es por ello que este tipo de información debe ser objeto de especial revisión, ya que podría descubrir a personas ajenas datos que el responsable del documento no desea revelar.

91. La información oculta del documento puede contener los siguientes datos:

a) **Comentarios, revisiones, versiones y anotaciones.**

Comentarios y cambios insertados en el documento por personas que colaboran en la revisión del mismo. Esta información normalmente es visible, pero puede ser ocultada y contiene nombres o iniciales de los participantes en la revisión del documento, sus comentarios y los cambios que realizaron.

b) **Encabezados, Pies de página y marcas de agua.**

Los documentos pueden contener información dentro de los encabezados y pies de página o en las marcas de agua, que pueden no ser visibles a simple vista.

c) **Texto, Filas y Columnas, Hojas ocultas.**

Los documentos Word pueden contener texto que ha sido formateado como texto oculto. Los documentos Excel pueden contener filas, columnas y hojas de cálculo ocultas.

d) **Contenido invisible.**

Los documentos pueden tener objetos invisibles (imágenes, formas, cuadros de texto, gráficos, tablas, etc.) que pueden haber sido marcados como no visibles.

e) **Contenido externo a las diapositivas.**

Los documentos PowerPoint pueden contener objetos que no son inmediatamente visibles porque han sido colocados en el área externa a la diapositiva. Estos objetos pueden ser cajas de texto, imágenes, gráficos y tablas.

f) **Notas de Presentación.**

Los documentos PowerPoint tienen una sección para notas sobre la presentación en la que se puede colocar texto. Normalmente estas notas están escritas únicamente para la persona que va a llevar a cabo la presentación y el autor no desea que sea compartido por nadie más.

g) **Datos XML personalizados.**

Los documentos pueden contener también datos XML que no son visibles desde el propio documento.

h) **Vínculos externos.**

Las hojas de cálculo de un documento Excel, pueden contener vínculos a los datos de hojas de cálculo de otros documentos Excel (vínculos externos). Los nombres de esas hojas de cálculo que contienen los datos a los que el documento está vinculado, se almacenan junto con el documento, pero puede que no sean visibles.

i) **Archivos incrustados u objetos.**

Los documentos pueden tener incrustados archivos u objetos que podrían tener metadatos asociados que no son visibles. Por ejemplo, si copiamos un gráfico Excel en una diapositiva PowerPoint, estamos copiando el gráfico y sus datos asociados, los cuales no son visibles, pero están en la caché.

j) **Macros de código VBA.**

Los documentos pueden contener macros, módulos VBA (Visual Basic para Aplicaciones), controles ActiveX o COM, formularios de usuario o funciones definidas por el usuario (UDF) que pueden contener datos ocultos.

k) **Elementos que pueden tener datos en caché.**

Algunas características de los programas Office utilizan datos en caché. Esto puede ser un problema, ya que estos datos en cache pueden contener metadatos y pueden quedar asociados al documento de forma oculta. Un ejemplo es el uso de tablas y gráficos dinámicos en Excel, que almacena los datos en caché para poder realizar de forma rápida el cálculo dinámico y aunque se elimine posteriormente la tabla o el gráfico, esos datos de caché quedan asociados al documento de forma no visible.

l) **Encuestas de Excel.**

Los documentos Excel (a partir de 2013) pueden contener de forma oculta preguntas de las encuestas Excel que hayan sido introducidas desde Excel Online, ya que han sido guardadas con el documento, pero no son visibles.

m) **Escenarios del Administrador de escenarios.**

Los documentos Excel pueden contener escenarios definidos a través del Administrador de escenarios. Estos escenarios pueden contener datos ocultos o en caché.

n) **Filtros.**

Los documentos Excel pueden contener filtros automáticos o filtros de tabla activos que podrían provocar que se almacenen datos ocultos o en caché en el documento. En caso de que, por ejemplo, apliquemos un filtro a una columna y después borremos algunas de sus celdas, los valores borrados pueden continuar en caché como parte del filtro, mientras que ya no aparecen de forma visible en el documento.

o) **Nombres ocultos.**

Los documentos Excel pueden contener nombres ocultos que podrían ser el origen de parte de los datos ocultos.

	Word	Excel	PowerPoint
Comentarios, revisiones, versiones y anotaciones	X	X	X
Propiedades del Documento e Información personal	X	X	X
Encabezados, Pies de página y marcas de agua	X	X	
Texto oculto	X		
Filas y Columnas ocultas		X	
Hojas de cálculo ocultas		X	

	Word	Excel	PowerPoint
Contenido Invisible	X	X	X
Contenido externo a las diapositivas			X
Notas de Presentación			X
Datos XML personalizados	X	X	
Vínculos externos		X	
Archivos incrustados y objetos	X	X	X
Macros de código VBA	X	X	X
Elementos que pueden tener datos en caché	X	X	X
Encuestas de Excel		X	
Escenarios del Administrador de escenarios		X	
Filtros		X	
Nombres ocultos		X	

Tabla 4.- Información oculta en los documentos Office 2010/2013/2016.

92. Finalmente, además de los metadatos en Propiedades del documento y de la Información oculta en documentos Office, los documentos también pueden contener:

- a) *Metadatos con propiedades del Servidor de Documentos.* En caso de que el documento haya sido almacenado en un servidor de gestión de documentos, como un espacio de trabajo de documentos o una librería basados en Microsoft SharePoint Server, el documento puede contener propiedades adicionales o información relacionada con el servidor de documentos.
- b) *Otros Metadatos.* En caso de utilizar características especiales, el documento puede contener tipos adicionales de información personal, como cabeceras de correo electrónico (incrustadas en el documento al usar ciertas características de flujos de revisión), información enviada para revisión o listas de distribución del documento.

2. Configuración de Seguridad

93. Los programas Microsoft Office ofrecen varias opciones para llevar a cabo una configuración de seguridad que ejerza el control sobre los datos personales que se almacenan en el documento.

- a) Especificar la información personal que aparece en todos los documentos de Office

Es una buena práctica especificar de forma apropiada la información personal que aparecerá en todos los documentos Office que un usuario vaya a crear o modificar.

Esta información se encuentra en las opciones de configuración de Office y puede ser editada o eliminada, especificando con ello la información que se mostrará en Autor del documento, Última persona que ha realizado modificaciones y Autor de comentarios de revisión del documento.

Una vez que se actualiza la información de una aplicación de Office, la información se actualiza automáticamente para el resto de aplicaciones de Office.

La forma de acceder a esta información es la siguiente:

- Con el archivo abierto, hacer clic en Archivo y a continuación hacer clic en Opciones. Se abrirá la ventana de Opciones de la aplicación, seleccionar General.
- En el área de Personalizar la copia de Microsoft Office aparece el Nombre de usuario (con el que Office rellenará de forma automática el Autor y Última persona que ha realizado modificaciones en el documento) y también las Iniciales (que se emplearán para identificar el Autor de comentarios de revisión de documento).

b) No guardar la información personal de un documento Office

Otra buena práctica es evitar que se archive información personal cuando se guarda un documento. Office permite seleccionar esta opción, de forma que cada vez que el documento se guarde, no se almacenará ningún metadato relacionado con información personal (Autor, Administrador, Última persona que ha realizado modificaciones en el documento, Compañía o Autor de comentarios de revisión del documento). Esta selección debe realizarse en cada documento.

La forma de acceder a esta información es la siguiente:

- Con el archivo abierto, hacer clic en Archivo y a continuación hacer clic en Opciones. Se abrirá la ventana de Opciones de la aplicación, seleccionar Centro de Confianza y pulsar en Configuración del Centro de Confianza. Se abre la ventana de Centro de Confianza.
- Seleccionar Opciones de privacidad y en el cuadro destinado a Configuración específica del documento aparecerá la opción “Quitar Información personal de las propiedades del archivo al guardarlo”. Esta opción sólo podrá seleccionarse cuando previamente se haya eliminado toda la información personal del documento y hace que cada vez que el documento se guarde, se elimine la información personal.

3. Inspección y Borrado de Metadatos e Información oculta

94. Existen multitud de aplicaciones, tanto gratuitas, como comerciales, que realizan la inspección y borrado de metadatos para documentos Microsoft Office. Aparte de estas soluciones, Microsoft Office dispone de varias utilidades que permiten eliminar metadatos de las propiedades de los documentos.

95. Una de ellas, es la opción de “Eliminar Propiedades e Información personal”, a la que se accede haciendo clic sobre el nombre del archivo, botón derecho del ratón, seleccionar Propiedades, seleccionar Detalles y pulsar la opción al final de la ventana “Eliminar Propiedades e Información personal”.

96. Se recomienda el uso del Inspector de Documentos, que es una utilidad incluida en

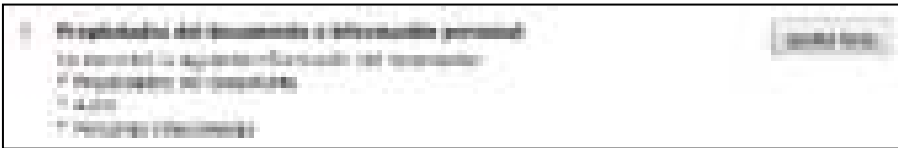
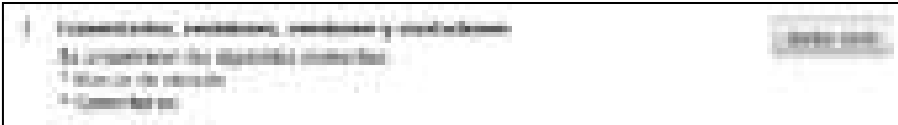
Microsoft Office para revisar documentos a partir de la versión de Office 2007. Esta utilidad permite inspeccionar los documentos para revisar, modificar o eliminar no sólo sus metadatos, sino también información y datos ocultos que pueda contener el documento.

97. Es importante indicar que Microsoft no permite el uso del Inspector de Documentos para revisar y eliminar información en documentos que han sido protegidos o firmados. Tampoco en aquellos documentos que utilicen Information Rights Management (IRM), que es una utilidad de Microsoft que permite restringir los permisos de un documento para evitar que personas no autorizadas impriman, reenvíen o copien información confidencial. Por lo tanto, antes de proteger o firmar el archivo será cuando deba realizarse la revisión con el Inspector de Documentos.
98. Asimismo, cuando se eliminen metadatos y datos ocultos a través del Inspector de Documentos, es muy probable que éstos no puedan ser recuperados de nuevo con el comando Deshacer, por lo que es recomendable realizar una copia del documento original antes de eliminar los datos.
99. Funcionamiento del Inspector de Documentos
 - a) Abrir el documento que se quiere inspeccionar.
 - b) Realizar una copia del documento original, si procede.
 - c) Abrir el Inspector de Documentos. Se accede haciendo clic en Archivo, seleccionar Información, hacer clic en Comprobar si hay problemas y seleccionar Inspeccionar Documento.
 - d) Se abre una caja de diálogo en la que se deberá seleccionar el tipo de metadatos e información oculta que se desea revisar en el documento. Una vez seleccionados hacer clic en Inspeccionar.
 - e) Una vez que el Inspector de Documentos finaliza su inspección, muestra los resultados de cada uno de los módulos inspeccionados en una ventana. Si ha encontrado datos en un módulo determinado, la ventana incluirá el botón “Quitar todo”, en el que podrá pincharse para borrar los datos encontrados en ese módulo. Si no ha encontrado datos en el módulo, la ventana mostrará un mensaje indicándolo.
 - f) En caso de haber seleccionado “Quitar todo” para un módulo, al finalizar se mostrará un texto indicando si la operación se ha realizado con éxito o no. Si el Inspector de Documentos encuentra errores en la operación, marcará con un aviso el módulo y mostrará un mensaje de error, no modificando los datos de ese módulo en el documento.

100. Información que inspecciona y elimina el Inspector de Documentos.

El Inspector de Documentos en sus módulos destinados a la inspección y borrado de contenido invisible, sólo detectará aquel contenido que haya sido formateado explícitamente como oculto o invisible. Por ejemplo, datos de un documento Excel que hayan sido colocados en la fila 10.000 y que por lo tanto no se encuentran dentro del área visible del documento, pueden considerarse por ello ocultos, pero no son detectados por el Inspector de Documentos, que los considerará como datos válidos.

Como excepción a esto, PowerPoint posee un módulo de Inspector de Documentos que sí permite detectar contenido fuera de la diapositiva, pero al igual que en Excel y Word, el Inspector de Documentos no detectará objetos no formateados como invisibles, pero ocultos por otros métodos (por ejemplo, formas sin contorno ni relleno, que por lo tanto no son visibles).

	Word	Excel	PowerPoint
Propiedades del Documento e Información personal - Metadatos generados de forma automática por los programas Office - Metadatos generados manualmente por el usuario o por la organización - Propiedades del Servidor de documentos - Otras propiedades  <i>Fragmento de la Ventana del Inspector de Documentos relativo a Propiedades en un documento Microsoft Word 2010.</i>	X	X	X
Comentarios, revisiones, versiones y anotaciones - Comentarios ⁽¹⁾ - Marcas de revisión de los cambios realizados - Información de versión del documento - Anotaciones con lápiz ⁽¹⁾  <i>Fragmento de la Ventana del Inspector de Documentos relativo a Comentarios y revisiones, en un documento Microsoft Word 2010.</i>  <i>Fragmento de la Ventana del Inspector de Documentos relativo a Comentarios y revisiones, en un documento Microsoft Excel 2010.</i>	X	X Sólo (1)	X Sólo (1)

	Word	Excel	PowerPoint
Encabezados, Pies de página y marcas de agua - Información de encabezados de los documentos ⁽¹⁾ - Información de los pies de página de los documentos ⁽¹⁾ - Marcas de Agua  Fragmento de la Ventana del Inspector de Documentos relativo a encabezado y pie de página en un documento Microsoft Word 2010.  Fragmento de la Ventana del Inspector de Documentos relativo a encabezado y pie de página en un documento Microsoft Excel 2010.	X	X Sólo (1)	-
Texto oculto - Texto cuyo formato es de texto oculto - No detecta texto ocultado por otros métodos (por ejemplo, en color blanco)  Fragmento de la Ventana del Inspector de Documentos relativo a texto oculto en un documento Microsoft Word 2010.	X	-	-
Filas y Columnas ocultas - Filas ocultas - Columnas ocultas que contienen datos o que no contienen datos en el caso de que estén situadas entre columnas que sí los contengan. - No detecta formas, gráficos, controles, objetos o controles de Microsoft ActiveX, imágenes ni elementos gráficos SmartArt que puedan estar situados en columnas ocultas. - No elimina filas o columnas ocultas si forman parte de un encabezado de tabla, de la lista o una tabla dinámica.  Fragmento de la Ventana del Inspector de Documentos relativo a filas y columnas ocultas en un documento Microsoft Excel 2010.	-	X	-
Hojas de cálculo ocultas  Fragmento de la Ventana del Inspector de Documentos relativo a hojas ocultas en un documento Microsoft Excel 2010.	-	X	-

	Word	Excel	PowerPoint
Contenido Invisible - Objetos que se han seleccionado como no visibles - No detecta objetos ocultos por otros métodos (por ejemplo, cubiertos por otros objetos)  Fragmento de la Ventana del Inspector de Documentos relativo al contenido invisible en un documento Microsoft Excel 2010.  Fragmento de la Ventana del Inspector de Documentos relativo al contenido invisible en un documento Microsoft PowerPoint 2010.	X	X	X
Contenido externo a las diapositivas - Detecta contenido y objetos ubicados fuera del área de la diapositiva: imágenes, cuadros de texto, gráficos, tablas, etc. - No detecta objetos externos a la diapositiva con efectos de animación.  Fragmento de la Ventana del Inspector de Documentos relativo al contenido externo a diapositivas en un documento Microsoft PowerPoint 2010.	-	-	X
Notas de Presentación - Detecta y elimina texto que se haya agregado a la sección Notas de una presentación. - No elimina imágenes agregadas a la sección Notas de una presentación.  Fragmento de la Ventana del Inspector de Documentos relativo a notas de presentación en un documento Microsoft PowerPoint 2010.	-	-	X
Datos XML personalizados  Fragmento de la Ventana del Inspector de Documentos relativo a notas de Datos XML personalizados en un documento Microsoft Word 2010.	X	X	X

Tabla 5. Información que inspecciona y elimina el Inspector de Documentos para Office 2010/2013/2016.

101. Información que inspecciona, pero no elimina el Inspector de Documentos

Hay datos que el Inspector de Documentos revisa y detecta, pero que no puede eliminar ya que su desaparición podría provocar que el documento no funcionase correctamente.

Por ejemplo, no puede eliminar el nombre de la plantilla de un documento, ya que si lo hiciese el documento al abrirse no sabría que plantilla usar. O no puede eliminar un código VBA (Visual Basic for Applications), ya que no puede determinar si eso eliminaría datos críticos o no. Es el usuario quien debe examinar de forma manual los elementos encontrados y eliminarlos o reemplazarlos si corresponde.

	Word	Excel	PowerPoint
Propiedades - Nombre y ruta de la plantilla - Nombre y ruta del fichero	X	X	X
Vínculos externos Llamadas a otros documentos Excel, que se usan en: - Celdas de hojas de cálculo - Nombres - Objetos - Títulos y series de datos de gráficos	-	X	-
Archivos u objetos incrustados	X	X	X
Macros de código VBA - Macros y Módulos VBA - Controles COM o ActiveX - Formularios de usuario - Funciones definidas por el usuario (UDF)	X	X	X
Elementos que podrían contener datos en caché Por ejemplo: Tablas y gráficos dinámicos, Segmentaciones y escalas de tiempo y Fórmulas de cubo	X	X	X
Encuestas de Excel Preguntas de la Encuesta de Excel (Excel Online)	-	X	-
Escenarios del Administrador de escenarios Escenarios definidos con el Administrador de escenarios que podrían almacenar datos en caché y ocultos	-	X	-
Filtros Filtros automáticos o filtros de tabla aplicados, que podrían almacenar datos en caché y ocultos.	-	X	-
Nombres ocultos Nombres ocultos que podrían almacenar datos ocultos.	-	X	-

Tabla 6. Información que revisa, pero no elimina el Inspector de Documentos en Office 2010/2013/2016.

102. Metadatos fuera del alcance de los módulos estándar del Inspector de Documentos

Para aquellos metadatos específicos adscritos al documento, que el Inspector de Documentos deba revisar y eliminar y que no se encuentren dentro de sus módulos disponibles de forma integrada, se podrá personalizar el Inspector de Documentos mediante programación, ya que Microsoft ofrece un interfaz para acceder a los métodos de los objetos del Inspector de Documentos.

Las organizaciones, por lo tanto, pueden crear e instalar módulos que utilicen el modelo de objetos del Inspector de documentos para inspeccionar y eliminar información de sus metadatos corporativos y personalizados.

ANEXO B. METADATOS EN DOCUMENTOS OPENOFFICE

103. El Formato de Documento Abierto para Aplicaciones Ofimáticas (OpenDocument Format, ODF) de OASIS, es un formato abierto y estándar para archivos de documentos ofimáticos. Esto incluye documentos de texto (extensión .odt), libros de cálculo (extensión .ods), presentaciones (extensión .odp), dibujos (extensión .odg), gráficos (extensión .odc), fórmulas matemáticas (extensión .odf) e imágenes (extensión .odi). Este formato ha sido aprobado por ISO e IEC como estándar internacional y está basado en XML.
104. Existen varias aplicaciones de código abierto (open source) que soportan el formato ODF. De entre todas ellas la más extendida es OpenOffice.org que dispone de una suite de aplicaciones ofimáticas para procesar textos, hojas de cálculo y presentaciones. Es un software gratuito y se encuentra publicado por la Fundación Software Apache, disponible en varios idiomas y para varias plataformas.
105. Toda la información recogida en este anexo, aplica a los documentos generados por los programas de OpenOffice: Writer, Calc e Impress en versión 4. Serán estos documentos a los que se haga referencia como “documentos OpenOffice” a lo largo del Anexo.

1. Tipos de Metadatos e Información oculta

106. Un documento OpenOffice es un paquete zip que contiene varios ficheros, existiendo una clara separación entre el contenido, la disposición de éste en el documento, y los metadatos. De entre todos los ficheros, existe uno que contiene los metadatos y es el fichero meta.xml. En este fichero se registran entre otros, los metadatos que se visualizan desde la Propiedades del documento en OpenOffice.

107. Metadatos en Propiedades del Documento.

Al igual que los documentos Office, los documentos OpenOffice contienen metadatos en las propiedades del documento. Estos metadatos pueden ser añadidos automáticamente por la aplicación, o pueden ser metadatos estándar y personalizados añadidos por el usuario o por la organización.

Se accede a ellos desde la opción del menú Archivo y seleccionando Propiedades. En la ventana de propiedades del documento, se podrán visualizar los metadatos agrupados en varias pestañas: General, Descripción y Propiedades personalizadas.

A continuación, se muestra una tabla con los metadatos existentes en las propiedades del documento, indicando si se generan de forma automática por el programa, o de forma manual por el usuario u organización.

Metadato	Generado de forma Automática	Generado de forma Manual
Tipo de archivo	X	
Herramienta con la que se creó el archivo	X	

Metadato	Generado de forma Automática	Generado de forma Manual
Tamaño	X	
Fecha Hora de Creación	X	
Usuario de creación	X	
Fecha Hora de última Modificación	X	
Usuario que realizó la última modificación	X	
Fecha Hora de última Impresión	X	
Usuario que realizó la última impresión	X	
Tiempo de edición total del documento	X	
Número de revisión del documento	X	
Plantilla	X	
Título		X
Tema		X
Palabras clave		X
Comentarios		X
Propiedades Personalizadas		X
Estadísticas del documento (número de páginas, tablas, imágenes, objetos OLE, párrafos, palabras, caracteres y líneas)	X	

Tabla 7. Metadatos en Propiedades de documentos OpenOffice.org (versión 4)

108. Metadatos incrustados en el documento.

Se pueden insertar metadatos también en el contenido del documento para documentos de texto (Writer). Deben ser metadatos definidos dentro de las propiedades del documento, o bien estándar o bien personalizados, y contenidos por lo tanto en el fichero meta.xml.

Estos metadatos se insertan desde la opción del menú Insertar, seleccionando Campos. Se despliega una ventana donde se listan metadatos estándar que se pueden insertar. Si seleccionamos Otros y en la pestaña de Información del Documento, se listan todos los metadatos estándar y personalizados que se pueden insertar en el contenido del documento.

109. Metadatos RDF.

RDF (Resource Description Framework) es un formato de metadatos utilizado por ODF 1.2 y es un estándar W3C. Los documentos OpenOffice pueden contener metadatos en formato RDF. Estos metadatos se pueden asociar a entidades del documento de texto, como párrafos, encabezados, tablas, filas o columnas de tablas, secciones, etc. y se almacenan en repositorios RDF desde los cuales, se referencian en la entidad del documento.

110. Los documentos OpenOffice también pueden contener información oculta, es decir, información que no es directamente visible en el documento a través de la configuración estándar de la aplicación, sino que es necesario seleccionar una configuración determinada para ver estos datos.

a) **Comentarios y Cambios.**

Los comentarios y cambios insertados en el documento por personas que colaboran en la revisión del mismo, contienen nombres o iniciales de los participantes en la revisión, sus comentarios y los cambios que realizaron. Para hacer visibles los cambios, acceder desde la opción de menú Editar, seleccionando Cambios y seleccionando Mostrar para que se hagan visibles todos los cambios.

Respecto a los comentarios, acceder desde la opción de menú Ver, y seleccionar Comentarios para que se hagan todos ellos visibles.

b) **Información de Versiones del documento.**

La información de versiones del documento solo es visible desde la opción del menú Archivo, seleccionando Versiones. Esto abre una ventana donde pueden gestionarse las versiones y añadir, revisar, comparar o eliminar versiones.

c) **Párrafos Ocultos.**

En los documentos de Writer, se pueden insertar párrafos ocultos que no serán visibles a no ser que acceda a la opción de menú Ver, y seleccionando Párrafos ocultos.

d) **Texto Oculto.**

En los documentos de Writer, se puede insertar texto oculto que no será visible a no ser que acceda a la opción de menú Ver, y seleccionando Caracteres no imprimibles.

e) **Secciones Ocultas.**

En los documentos de Writer, se pueden ocultar secciones que no serán visibles a no ser que acceda a la opción de menú Formato, seleccionando Secciones y en la ventana de Modificar secciones hacer clic en cada una de ellas para revisar está activo el check de ocultar, en cuyo caso habrá que desactivarlo para hacer la sección visible.

f) **Filas, Columnas, celdas y Hojas ocultas.**

Las hojas de cálculo de Calc, pueden contener filas, columnas, celdas y hojas de cálculo ocultas. Para mostrar las hojas, filas y columnas ocultas, acceder desde la opción del menú Formato, seleccionando fila, columna, hoja y a continuación Mostrar.

Para mostrar una celda oculta, acceder desde el menú Formato y seleccionar la opción Celdas. En la ventana de Formato de Celdas, ir a la pestaña de Protección de celda y desactivar las opciones de ocultar.

2. Configuración de Seguridad

111. Los programas OpenOffice ofrecen varias opciones para llevar a cabo una configuración de seguridad que ejerza un cierto control sobre los datos personales que se almacenan en el documento.

- a) Especificar la información personal que aparece en todos los documentos de OpenOffice.

Es una buena práctica especificar de forma apropiada la información personal que aparecerá en todos los documentos OpenOffice que un usuario vaya a crear o modificar.

Esta información se encuentra en las opciones de configuración y puede ser editada o eliminada, especificando con ello la información que se mostrará en Autor del documento, Última persona que ha realizado modificaciones, Autor de comentarios del documento y última persona que ha impreso el documento.

Una vez que se actualiza esta información de una aplicación de OpenOffice, la información se actualiza automáticamente para el resto de aplicaciones de OpenOffice.

La forma de acceder a esta información es la siguiente:

- Con el archivo abierto, desde el menú Herramientas, seleccionar Opciones. Se abrirá la ventana de Opciones de la aplicación.
- En el área de OpenOffice seleccionar Datos del Usuario. Aparecerá el Nombre de usuario (con el que OpenOffice rellenará de forma automática el Autor, Última persona que ha realizado modificaciones en el documento, última persona que ha impreso el documento y autor de cambios y comentarios del documento).

- b) No guardar la información personal en un documento OpenOffice.

Otra buena práctica es evitar que se archive información personal cuando se guarda un documento. OpenOffice permite seleccionar esta opción, de forma que cada vez que el documento se guarde, no se almacenará la información del usuario que crea, modifica o imprime el documento. Esta configuración debe realizarse en cada documento.

La forma de acceder a esta información es la siguiente:

- Con el archivo abierto, desde la opción del menú Herramientas, seleccionar Opciones. Se abrirá la ventana de Opciones de la aplicación.
- En el área de OpenOffice, seleccionar Seguridad, y pulsar el botón Opciones en el área de Opciones de seguridad y alertas. Se abre la ventana de seguridad con las siguientes opciones que habrá que seleccionar:
 - Advertir si el documento contiene cambios, versiones, información oculta o notas al guardar o enviar, al firmar, al imprimir y al crear archivos PDF.
 - Eliminar la información personal al guardar el documento.
- Desde la opción de menú Archivo, seleccionar Propiedades. En la pestaña General, deseleccionar el check "Utilizar los datos del usuario" y pulsar Restablecer para que elimine los datos actuales si los tiene.

3. Inspección y Borrado de Metadatos e Información oculta

112. Existen multitud de aplicaciones, tanto gratuitas, como comerciales, que realizan la inspección y borrado de metadatos para documentos OpenOffice.
113. El Inspector de Documentos de Microsoft Office, también se puede aplicar sobre documentos OpenOffice para eliminar metadatos e información oculta del mismo modo que se indica en el Anexo A. Únicamente indicar, que debe ejecutarse el Inspector de Documentos cada vez que se guarde el archivo en formato de OpenDocument.
114. Aparte de estas soluciones, existen varias opciones para eliminar metadatos de los documentos OpenOffice:

a) Modificar o Eliminar el archivo meta.xml.

Dado que un documento OpenOffice no es realmente un archivo individual, sino que son varios archivos comprimidos en un paquete zip, se puede descomprimir el documento (por ejemplo, cambiándole la extensión a zip y usando un descompresor) y del listado de ficheros resultantes, eliminar o modificar el fichero meta.xml, que es el que contiene los metadatos del documento.

Una vez eliminado o modificado este fichero, se deberá volver a realizar la compresión para obtener el documento sin metadatos (cambiando la extensión zip resultante por la del documento). Esta acción habría que realizarla cada vez que se guarde el archivo, ya que el fichero meta.xml se regenera y es necesario eliminarlo o modificarlo de nuevo.

Esta opción, sin embargo, no elimina otra información oculta del documento, como por ejemplo los Cambios y comentarios, información de versiones, etc. Para ello será necesario Editar el documento y eliminarlos manualmente.

b) Borrado manual de metadatos.

Se pueden también borrar los metadatos de forma manual de la siguiente forma:

- A través de la utilidad “Restablecer” se pueden eliminar datos personales. Se accede desde la opción de menú Archivo, seleccionando Propiedades y en la pestaña General, abajo a la derecha figura el botón “Restablecer”.

Si lo activamos, se eliminan los siguientes datos: Fecha de creación (la reinicia a la Fecha y Hora actual), Fecha y usuario que realizó la última modificación, Fecha y usuario que realizó la última impresión, tiempo total de edición y número de revisión. En caso de que también se quiera eliminar el usuario de creación, deseleccionar el check “Utilizar los datos del usuario” y pulsar “Restablecer”.

- Para borrar el resto de metadatos estándar y personalizados, se deberá acceder a las pestañas Descripción y Propiedades personalizadas y borrar los metadatos de forma manual.

- Para eliminar otra información oculta del documento, como cambios o versiones, habrá que editarlo y eliminar esta información de forma manual.

ANEXO C. METADATOS EN DOCUMENTOS PDF

115. Toda la información recogida en este anexo, aplica a los documentos generados por los programas Adobe Acrobat X, XI, DC. Serán estos documentos a los que se haga referencia como “documentos PDF” a lo largo del anexo.

1. Tipos de Metadatos e Información oculta

116. Metadatos en Propiedades del Documento.

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 o metadatos estándar y personalizados insertados por el usuario o la organización.

A estos metadatos se accede desde la opción de menú Archivo, seleccionando Propiedades.

Los metadatos estándar se encuentran en la pestaña Descripción y serán propiedades ya predefinidas 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. Los metadatos personalizados se encuentran en la Pestaña Personalizar y son creados por el usuario o la organización añadiendo una etiqueta, un tipo de dato y un valor.



Figura 4.- Propiedades del documento en Adobe Acrobat 9.

117. Metadatos XMP.

A partir de la versión 5.0 de Adobe Acrobat y PDF 1.4, los documentos PDF pueden contener metadatos XML usando el formato XMP (Extensible Metadata Platform). Desde un documento PDF se pueden guardar, exportar e importar metadatos en formato XMP, para poder compartir metadatos entre varios documentos PDF.

Estos metadatos se pueden ver desde la opción de menú Archivo, seleccionando Propiedades, seleccionando la pestaña Descripción, pinchar en Metadatos adicionales seleccionar Avanzado. Se mostrarán los metadatos incrustados en el documento y según el esquema XMP, es decir, en grupos predefinidos de información relacionada.

Desde aquí también se podrán añadir o reemplazar metadatos XMP utilizando la opción de Anexar o Reemplazar y seleccionando un archivo XMP.



Figura 5.- Metadatos XMP en documento Adobe Acrobat 9.

118. Metadatos en plantilla.

Adobe Acrobat también permite exportar a una plantilla de metadatos, los metadatos de un documento para poder reutilizarlos en otros documentos PDF. La exportación e importación de este tipo de metadatos, se hace desde la opción de menú Archivo, seleccionando Propiedades, pestaña Descripción, hacer clic en Metadatos adicionales, seleccionar Avanzado y Guardar plantilla de metadatos.

119. Metadatos de Objetos dentro de un documento.

Además de los metadatos asociados al documento PDF, hay ciertos objetos, etiquetas e imágenes dentro de un PDF que también pueden tener sus propios metadatos.

120. Los documentos PDF también pueden contener información oculta, es decir, información que no es directamente visible en el documento a través de la configuración estándar de la aplicación, sino que es necesario seleccionar una configuración determinada para ver estos datos.

- a) **Archivos adjuntos.** Archivos en cualquier formato pueden ir asociados a un documento PDF como adjuntos. En algunos casos puede resultar inconveniente que el documento lleve ciertos archivos adjuntos por lo que deben revisarse.
- b) **Marcadores.** Los marcadores son enlaces en algún texto representativo, que conducen a ciertas secciones del documento. Dada su función, no suelen representar ningún riesgo, salvo que el título del marcador sea comprometedor.
- c) **Comentarios.** Los documentos PDF pueden contener comentarios que han sido realizados a través de la herramienta de comentarios y marcas. Pero estos comentarios además pueden estar ocultos y no ser visibles para el usuario, por lo que es imprescindible revisar el documento para evitar que información sensible contenida en los comentarios sea distribuida. Los comentarios pueden contener también archivos adjuntos.
- d) **Campos de Formularios.** Los documentos PDF pueden contener campos de formularios, en los que se pueden incluir campos de firma o cualquier tipo de acción o cálculo asociados a ellos. Es por lo tanto necesario revisar el contenido de estos campos antes de la distribución del documento.
- e) **Texto oculto.** Los documentos PDF pueden contener texto que haya sido ocultado por diversos métodos: texto transparente, del color del fondo o cubierto por otro texto y objetos. Es prioritario examinar el documento en busca de texto oculto, ya que al no ser visible puede contener información que no debe ser distribuida.
- f) **Capas ocultas.** Los documentos PDF pueden contener varias capas, que pueden estar visibles u ocultas. Las capas ocultas pueden contener texto u objetos sensibles que no se ven a simple vista y que hay que examinar antes de distribuir el documento.
- g) **Índices de Búsqueda.** Los índices de búsqueda se utilizan en los documentos PDF para realizar búsquedas en el documento con mayor rapidez. No suelen representar un riesgo.
- h) **Contenido eliminado.** Los documentos PDF pueden en algunos casos retener contenido que haya sido eliminado y no sea ya visible, como páginas eliminadas o cortadas, o imágenes borradas.

Al tratarse de información que ha sido eliminada, supone un contenido que no se quiere mostrar, por lo que el riesgo de que sea distribuido es elevado y debe revisarse el documento para buscar este tipo de contenido.

- i) **Enlaces (links), Acciones y Scripts de Java.** Los documentos PDF pueden contener enlaces a páginas web y acciones añadidas a través de la utilidad Acciones y scripts de Java (por ejemplo, lanzar un mensaje mediante un script de java cuando se realice la acción de cerrar el documento). Los enlaces son difíciles de encontrar en un documento ya que pueden estar asociados a múltiples objetos.
- j) **Objetos solapados.** Los documentos PDF pueden contener objetos que cubren o solapan a otros objetos. Estos objetos pueden ser imágenes, gráficos, patrones, etc.

Los objetos solapados pueden ser información obsoleta o información sensible, por lo que deben ser revisados antes de distribuir el documento.

2. Configuración de Seguridad

121. En el caso de que los documentos PDF no tengan que llevar ningún metadato o información oculta, sino que deban ser documentos PDF planos, se puede automatizar el proceso de eliminación de estos ítems a través de las preferencias de Adobe Acrobat.
122. Para ello, dentro de las preferencias seleccionar la categoría Documento y seleccionar las opciones correspondientes dependiendo de la versión de Adobe Acrobat.
123. En Acrobat 9, dentro del área examinar documento estarán las opciones:
 - Examinar documento al cerrar documento.
 - Examinar documento al enviarlo por correo electrónico.
124. En Acrobat X, dentro del área de Información oculta:
 - Eliminar la información oculta, cuando el documento es guardado.
 - Eliminar la información oculta, cuando el documento es enviado por correo electrónico.

3. Inspección y Borrado de Metadatos e Información oculta

125. Existen multitud de aplicaciones, tanto gratuitas, como comerciales, que realizan la inspección y borrado de metadatos para documentos PDF. Aparte de estas soluciones, Adobe Acrobat dispone de varias utilidades que permiten eliminar metadatos de las propiedades de los documentos.

126. Censurar contenido.

Para eliminar o censurar información sensible o privada del contenido de un PDF que no queremos que se muestre cuando el documento sea distribuido, Adobe Acrobat dispone de utilidades que eliminan de forma permanente estas palabras, frases, párrafos, gráficos o imágenes del documento, permitiendo sustituirlo por un área en blanco, por cajas de color (negro, por ejemplo) o por otro texto y objetos.

La ubicación de esta utilidad depende de la versión de Adobe Acrobat. En Acrobat DC, por ejemplo, se encuentra en Herramientas – Redacción.

127. Borrado manual de metadatos.

Los metadatos que se encuentran en las Propiedades de documento, se pueden modificar y eliminar de forma manual con Adobe Acrobat accediendo a ellos y modificando o borrando su contenido.

128. Inspección y borrado automático de metadatos: “Eliminar Información oculta”

Adobe Acrobat dispone de una utilidad para inspeccionar y eliminar todos los metadatos e información oculta de un documento: en Acrobat 9, es la utilidad de “Examinar Documento” y en Adobe X, XI y DC es la utilidad de “Eliminar Información oculta”.

Una vez que se ejecuta la utilidad, se despliega el listado ítems de metadatos y datos ocultos que ha encontrado sobre el documento, dando la opción de pre visualizarlos para poder verificar de que datos se trata. La utilidad permite seleccionar qué ítems se quieren eliminar.



Figura 6.- Utilidad de Examinar documento en Acrobat 9.

La utilidad de “Eliminar Información oculta”, debe ser usada con minuciosidad, ya que puede también eliminar cualquier otro objeto adjunto al documento, como pueden ser firmas digitales, información añadida a través de plugins y aplicaciones de terceras partes o características especiales de Acrobat Reader, a través de las cuales los usuarios revisan, firman y rellenan documentos PDF.

a) **Funcionamiento de “Eliminar Información oculta” para Adobe Acrobat X:**

- Abrir el documento que se quiere inspeccionar con Adobe Acrobat.
- Desde el panel de Herramientas, desplegar el panel de Protección y seleccionar Eliminar Información oculta.
- Se abre una ventana y comienza el análisis del documento. Seleccionar Expandir todo, para poder ver los ítems de información oculta que va encontrando.
- Una vez finaliza, seleccionar los ítems de metadatos e información oculta que queremos eliminar y pulsar Eliminar.
- Los cambios al documento no se aplicarán hasta que se guarde. Dado que una vez guardado los cambios son permanentes y no se podrá recuperar la información y metadatos eliminados, es recomendable haber realizado una copia de seguridad previa.

La utilidad permite previsualizar los ítems de metadatos e información oculta previamente a seleccionarlos para su borrado. De esta forma podemos acceder a los ítems para modificarlos (por ejemplo, en el caso de metadatos), y desactivarlos en caso de que una vez modificados ya no nos interese eliminarlos.

b) **Metadatos e Información oculta que inspecciona y elimina la utilidad “Eliminar Información oculta” para Adobe Acrobat X:**

Tipo de Metadato o Información oculta	“Eliminar Información oculta”
Metadatos en Propiedades del documento	Sí
Metadatos XMP	Sí
Metadatos importados con plantilla Adobe	Sí
Metadatos en Objetos dentro del documento	Sí
Archivos Adjuntos	Sí
Marcadores	Sí
Comentarios	Sí
Texto Oculto	Sí

Tipo de Metadato o Información oculta	“Eliminar Información oculta”
Capas ocultas	Sí, pero si se eliminan las capas ocultas, se formatean las demás capas visibles en una sola capa.
Enlaces a URL y Acciones y Scripts de Java	Sí
Objetos solapados	Sí
Contenido eliminado	Sí
Índices de Búsqueda	Sí, pero si se eliminan estos índices, se reduce el tamaño del archivo pero se incrementan los tiempos de búsqueda.
Campos de Formulario	Sí, pero si se eliminan estos campos, se dejan formateados y no se podrán rellenar, editar o firmar.

Tabla 8.- Información que inspecciona y elimina la utilidad “Eliminar Información oculta” de Adobe Acrobat X.

[illegible]

Figura 9.- Ejemplo de Metadatos XMP.

136. Algunos tipos de metadatos, son creados y utilizados en exclusiva por alguno de los estándares. En cambio, otros tipos de metadatos son implementados por más de un estándar.

La siguiente figura y la siguiente tabla, muestran metadatos compartidos en los tres estándares y metadatos usados en exclusiva por alguno de ellos.



Figura 10.- Metadatos comunes o exclusivos en cada estándar.

Metadato	Descripción	EXIF	XMP	IPTC
Descripción	Descripción de la imagen	X	X	X
Fecha Creación	Cuándo la imagen fue tomada o digitalizada	X	X	X
Copyright	Información del copyright	X	X	X
Autor	Autor de la imagen	X	X	X
Etiquetas	Etiquetas para identificar la imagen y poder filtrar en las búsquedas.		X	X
Datos textuales sobre localización	Datos de texto sobre la localización: Región del mundo, País, Provincia, Ciudad, Zona.		X	X
Clasificación	Puntuación de la imagen (por ejemplo, de 1 a 5 estrellas)		X	

Localización GPS	Coordenadas GPS creadas automáticamente por el dispositivo de captura de imagen.	X		
Tipo de dispositivo	Tipo de dispositivo con el que se ha capturado la imagen: fabricante, modelo, número de serie, etc.	X		
Configuración del dispositivo	Configuración del dispositivo con el que se ha capturado la imagen: tiempo de exposición, número F o abertura, distancia focal, modo de flash, etc.	X		

Tabla 9.- Detalle de Metadatos comunes o exclusivos en cada estándar.

137. Por otro lado, y como se muestra en la siguiente figura, las últimas versiones de XPM ya permiten mapear casi todos los metadatos EXIF e IPTC.

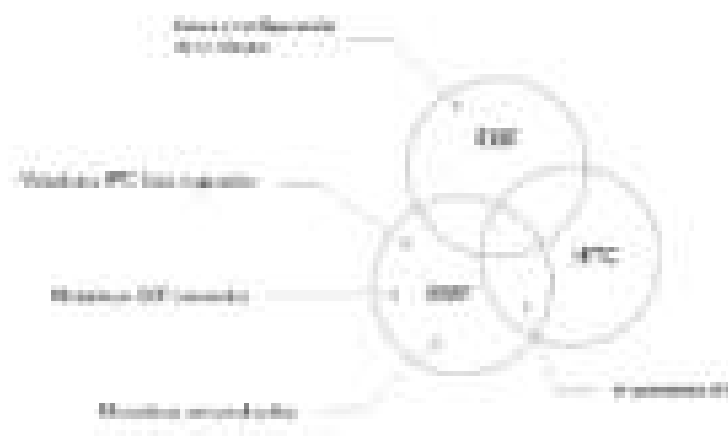


Figura 11.- Mapeo de metadatos a XMP.

138. Un archivo de imagen digital (por ejemplo, JPEG, TIFF o PSD) puede almacenar metadatos en varios formatos (EXIF, XPM o IPTC). Cada tipo de metadato deberá ir en el segmento apropiado del fichero. Por ejemplo, en un archivo JPEG los segmentos APP1 son los destinados a metadatos EXIF y XMP. El segmento APP13 se destina a información “no gráfica” de Adobe Photoshop y en este segmento irán los metadatos IPTC.

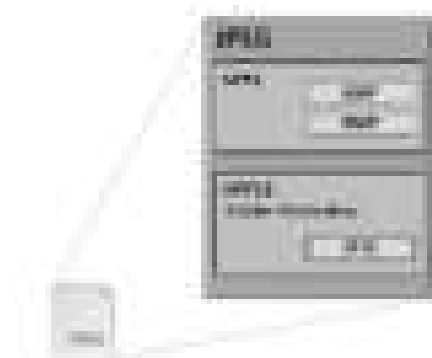


Figura 12.- Tipos de Metadatos en ficheros JPEG.

2. Configuración de Seguridad

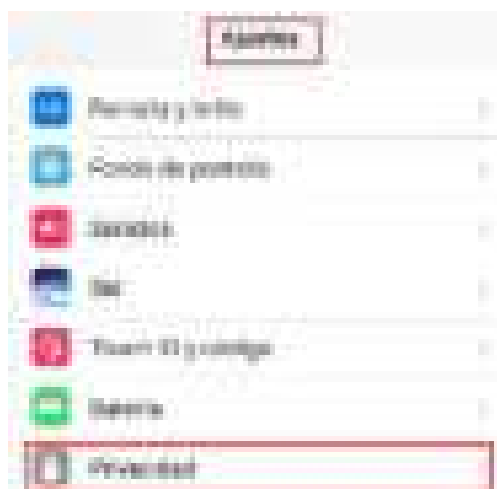
139. Los dispositivos móviles, tanto iOS como Android, permiten añadir información de la localización geográfica a las fotografías realizadas con la cámara del dispositivo. Esta información es conocida como geoetiquetas (o geotags). Al realizar la fotografía se añaden a la cabecera EXIF de la imagen las coordenadas GPS en el momento de realizar la captura.
140. 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.
141. 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.
142. **Dispositivos iOS.**

Por defecto, iOS no tiene definido si la información de localización se incorporará a las fotografías tomadas con el dispositivo. La primera vez que se hace uso de la cámara, iOS solicita permiso al usuario para incorporar la información de localización a las fotografías.

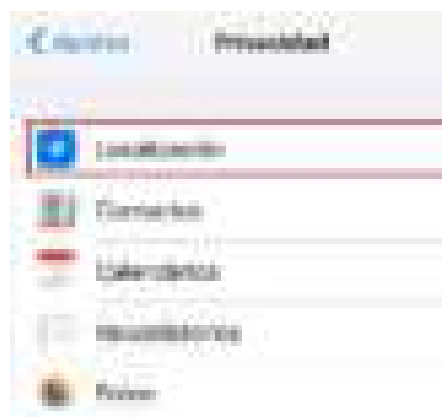
Esta funcionalidad es utilizada automáticamente por iOS una vez la aplicación "Cámara" ha sido añadida al centro de gestión de localización de iOS.

A través del menú de configuración del centro de gestión de localización, es posible habilitar o deshabilitar la incorporación de esta información a las fotografías que se realicen con la cámara. En iOS 10.x, se realizará siguiendo estos pasos:

- Acceder a Ajustes, seleccionar privacidad.
- Seleccionar Localización. Esto nos da acceso al centro de gestión de localización.
- De entre las aplicaciones disponibles en el centro, seleccionar Cámara.
- Dentro de las opciones de Permitir acceso a ubicación, seleccionar Nunca.



Paso 1



Paso 2



Paso 3



Paso 4

Figura 13. Deshabilitar localización en cámara iOS 10.x.

Para comprobar si una imagen de la galería de imágenes tiene asociados metadatos de localización, se debe seleccionar y pulsar Detalles. Ahí se podrá observar si se encuentran los metadatos de la localización donde fue tomada la imagen, e incluso si dispone de estos metadatos la situará en un plano.

143. Dispositivos Android.

En Android la primera vez que se ejecuta la aplicación "Cámara", se pregunta al usuario si desea etiquetar las fotografías y vídeos con la información de la ubicación donde se han capturado (opción habilitada por defecto).

Esta funcionalidad puede ser habilitada y deshabilitada en cualquier momento a través del menú de configuración de la cámara.

En Android 5.x, se realizará siguiendo estos pasos:

- Abrir la cámara.

- Seleccionar configuración.
- Dentro de la ventana de configuración, ir a la pestaña de ajustes.
- Deshabilitar la opción “Etiquetado Geográfico”.



Figura 14. Deshabilitar localización en cámara Android 5.x.

Cuando seleccionamos una imagen de la galería de imágenes, al pulsar Información observamos si se encuentran los metadatos de la localización donde fue tomada la imagen, e incluso la situará en un plano.

144. Deshabilitar la función de localización de la cámara puede no evitar que otras aplicaciones del dispositivo que tengan acceso a ella, y que tengan habilitada la localización (como, por ejemplo, WhatsApp, Facebook, Instagram, etc.), sean todavía capaces de almacenar los datos de localización en la imagen digital que se capture directamente con ellas. Esto dependerá de las versiones de sistema operativo del teléfono y de las versiones de dichas aplicaciones.
145. Es por lo tanto conveniente capturar primero la imagen con la cámara del dispositivo, que tendrá deshabilitada la captura de datos de localización, antes de proceder a compartirla con estas aplicaciones.

3. Inspección y Borrado de Metadatos

3.1. Utilidades para Ordenador

146. Cuando la imagen digital se encuentra en el ordenador, se pueden emplear varios métodos para visualizar y eliminar los metadatos que contiene. Por un lado, existen utilidades propias del sistema operativo del equipo, y, por otro lado, existen multitud de aplicaciones, tanto gratuitas como comerciales, que permiten

visualizar y borrar los metadatos asociados a las imágenes digitales.

147. Utilidades de Microsoft Windows.

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

a) *Explorador de Archivos.*

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”. En ellas se recogen las características de la cámara (fabricante, modelo, software, etc.), y la configuración en el momento de captura de la imagen (distancia focal, velocidad ISO, etc.).

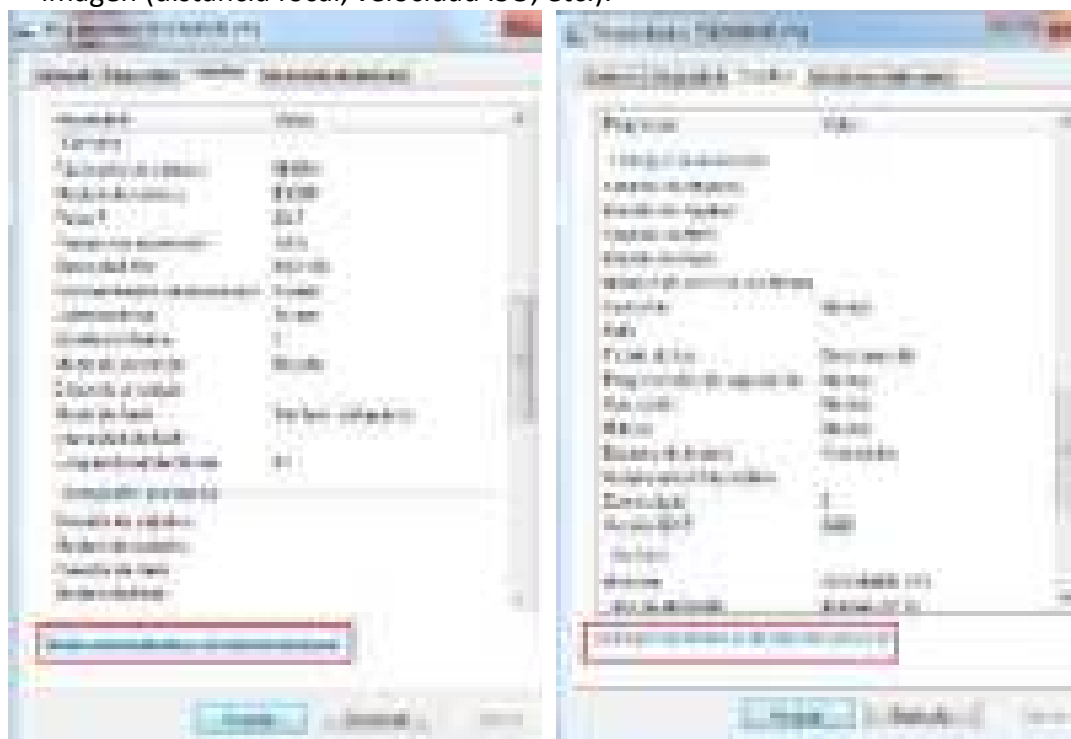


Figura 15. Metadatos mostrados por explorador de ficheros Windows.

En la parte baja de la ventana se encuentra la opción de “Quitar propiedades e información personal”. Al pulsar en esta opción, se abre la herramienta de Windows para eliminar los datos EXIF. Esta herramienta permite dos opciones:

- Seleccionar los datos que queremos borrar (todos, o un conjunto de ellos marcando su correspondiente check).
- Dejar que Windows cree una copia de la imagen eliminando todos los datos que sea posible.



Figura 16. Eliminar metadatos con Explorador de ficheros Windows.

b) *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.

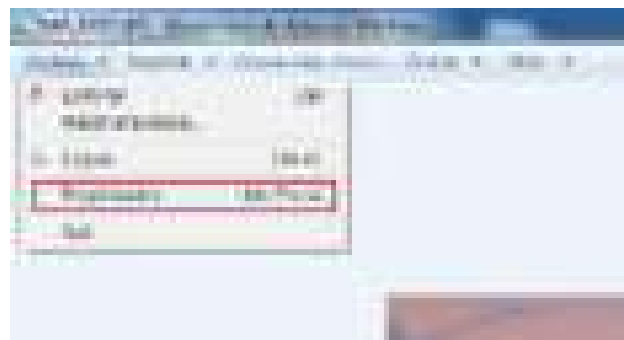


Figura 17. Visualizador de fotos de Windows.

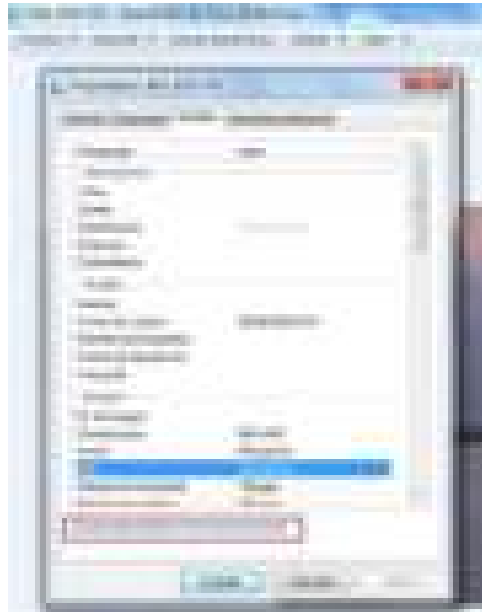


Figura 18. Propiedades con Visualizador de fotos de Windows.

148. GIMP (GNU Image Manipulation Program)

GIMP es un software libre de edición de imágenes compatible con múltiples plataformas. Se incluye en esta guía como ejemplo para ilustrar el borrado de metadatos, pero existen muchos otros.

Desde GIMP se pueden visualizar, agregar, editar o borrar los metadatos asociados a las imágenes digitales. A continuación, se indican los pasos a seguir con la versión 2.8.20 para eliminar metadatos.

- Desde GIMP, abrir la imagen digital que queremos modificar. Si seleccionamos Archivo y a continuación Propiedades, se pueden observar los metadatos que contiene.



Figura 19. Metadatos visualizados en GIMP 2.8.20.

- Para eliminar los metadatos, seleccionar la opción de menú Archivo y a continuación seleccionar Exportar. Indicar el nuevo nombre de la imagen que vamos a crear sin metadatos. Seleccionar la carpeta en la que vamos a guardar la imagen exportada.

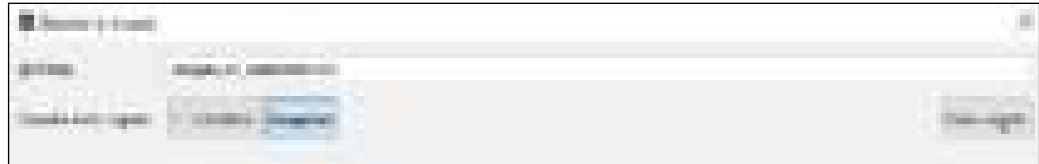


Figura 20. Eliminar metadatos en GIMP 2.8.20. Paso 1.

- Pulsar Exportar. Esto abre una ventana con las opciones para exportar. Expandir las opciones seleccionando el panel de Opciones Avanzadas.

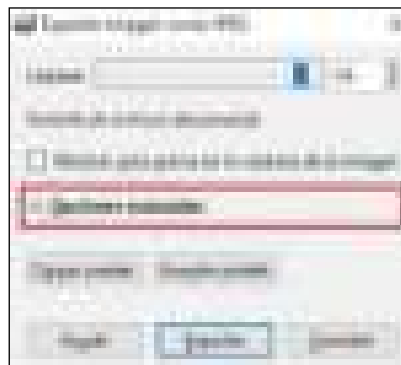


Figura 21. Eliminar metadatos en GIMP 2.8.20. Paso 2.

- Quitar el check de "Guardar datos EXIF" y "Guardar datos XMP". Seleccionar el resto de opciones según corresponda, y pulsar Exportar.



Figura 22. Eliminar metadatos en GIMP 2.8.20. Paso 3.

3.2. Utilidades para Dispositivos móviles

149. En caso de que las imágenes digitales se encuentren en un dispositivo móvil (por ejemplo, Smartphone o Tablet), puede ser más apropiado emplear una aplicación que elimine los metadatos directamente en el dispositivo, en lugar de trasladar la imagen digital al ordenador.

150. Existen multitud de aplicaciones en la tienda de aplicaciones, tanto de Apple (App Store), como de Android (Play Store), que se pueden descargar de forma gratuita, y que permiten abrir una imagen digital del álbum del dispositivo y realizar varias funciones como las siguientes.

- Visualizar sus metadatos. En la siguiente figura se muestran dos ejemplos de cómo se pueden visualizar los metadatos de una imagen digital, en aplicaciones gratuitas para iPhone.



Figura 23. Ejemplos de visualización de metadatos en iPhone.

- Eliminar sus metadatos. En las siguientes figuras se muestra un ejemplo de cómo se pueden eliminar los metadatos de una imagen digital, con una aplicación gratuita para iPhone. La primera figura representa la imagen original con los metadatos EXIF, incluida la localización. La segunda es la copia de la imagen original que ha creado la aplicación, con los metadatos eliminados.



Figura 24. Ejemplo de borrado de metadatos en iPhone.

- Compartir imagen. Algunas aplicaciones ofrecen la opción de compartir la imagen con o sin metadatos. Otras, no permiten compartir la imagen y generan un aviso para indicar que deben eliminarse primero sus metadatos sensibles.
- La siguiente figura muestra un ejemplo de ambos casos, con aplicaciones gratuitas para iPhone.

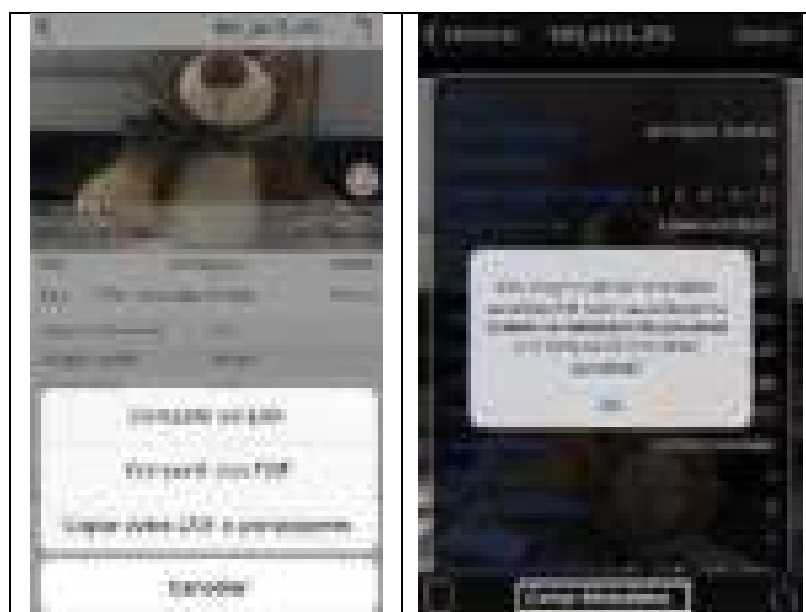


Figura 25. Ejemplos de compartir imagen con o sin metadatos en iPhone.

ANEXO E. DEFINICIONES

Archivo electrónico. Un tipo de objeto de información electrónica perdurable, normalmente generado por un programa de ordenador y disponible para su uso por otros programas.

Documentos Ofimáticos. Son un tipo de documento electrónico generado por los ‘programas ofimáticos’.

EXIF. Exchangeable Image File Format. Especificación para formatos de archivos de imagen usado por las cámaras digitales. Esta especificación usa los formatos de archivos de imagen más comunes, como JPEG, TIFF Rev. 6.0, RIFF y agrega ítems específicos de metadatos.

Información o datos ocultos. Son aquellos datos existentes en el contenido de los documentos electrónicos, que no son visibles con la configuración estándar o configuración por defecto de los programas office, siendo necesario aplicar alguna opción específica dentro de la configuración de estos programas para su visualización. Un ejemplo de datos ocultos es el texto oculto, filas o columnas ocultas, comentarios o información del documento, etc.

IPTC (International Press Telecommunication Council). Organización que crea y mantiene estándares de metadatos.

IPTC – IIM (IPTC Information Interchange Model). Estándar de formato de metadatos multimedia creado por IPTC.

IPTC Core. Estándar de formato de metadatos para imágenes digitales basado en XMP, creado por IPTC.

IRM (Information Rights Management). Utilidad de Microsoft que permite restringir los permisos de un documento para evitar que personas no autorizadas impriman, reenvíen o copien información confidencial.

JPEG (Joint Photographic Experts Group). Formato de ficheros digitales, ampliamente utilizado para ficheros de imagen y fotografía.

Metadatos. Información estructurada que describe, explica, localiza y además hace más fácil recuperar, utilizar o gestionar un recurso de información. Los metadatos son comúnmente llamados “datos sobre los datos” o “información sobre la información”.

ODF (OpenDocument Format). Formato de documento abierto. Es un formato abierto y estándar para archivos de documentos ofimáticos. Esto incluye documentos de texto (extensión .odt), libros de cálculo (extensión .ods), presentaciones (extensión .odp), dibujos (extensión .odg), gráficos (extensión .odc), fórmulas matemáticas (extensión .odf) e imágenes (extensión .odi). Este formato ha sido aprobado por ISO e IEC como estándar internacional y está basado en XML.

PII (Personally Identifiable Information). Información personalmente identificable. Información que puede usarse para identificar, contactar o localizar a una persona en concreto, o puede usarse junto con otras fuentes de información para hacerlo.

Política de Gestión documental. Dentro de las Políticas implantadas en la organización para el desempeño de sus actividades, la Política de Gestión documental establecerá unos criterios y normas en relación con la gestión de los documentos. En ella estarán especificados y detallados los metadatos que deben contener los documentos para asegurar la gestión, recuperación y conservación de los mismos durante todo su ciclo de vida.

Proceso de Configuración Segura de las aplicaciones. Proceso que describe cómo establecer la Configuración de Seguridad en las aplicaciones o programas que realicen el tratamiento y gestión de documentos electrónicos, de forma que se eviten o al menos se limiten los metadatos e información oculta que dichos programas puedan almacenar en los documentos.

Proceso de limpieza de documentos. Proceso perteneciente a los Procedimientos de Seguridad de la organización, cuyo objetivo es especificar las actividades a realizar para eliminar los metadatos e información oculta de los documentos. Deberá indicar las herramientas a emplear en función del tipo de documento, y cómo se utilizará cada una de ellas para eliminar el metadato o dato oculto deseado, incluyendo el detalle de las acciones a llevar a cabo para completar la limpieza del documento y también las acciones necesarias para verificar que la limpieza se ha completado de forma efectiva.

Programas de generación y tratamiento de documentos. Programas de ordenador destinados a la generación de documentos en cualquier formato (por ejemplo, Adobe Acrobat).

Programas o Aplicaciones Ofimáticas. Dentro de los programas de generación y tratamiento de documentos, los programas ofimáticos son un conjunto de programas básicos para su uso en oficinas, con un interfaz y funciones comunes y cuyo objetivo será el tratamiento de textos, hojas de cálculo, presentaciones, gráficos, tablas, etc. Dos ejemplos de programas ofimáticos muy conocidos, son Microsoft Office y Apache OpenOffice.

PSD. Formato de fichero nativo de Adobe Photoshop.

Sidecar files. Ficheros que almacenan datos, normalmente metadatos, no soportados por el formato del documento fuente. Cada documento puede llevar asociados uno o más “sidecar files” con metadatos en diversos formatos.

TIFF (Tagged Image File Format). Formato de fichero para almacenamiento de imágenes y fotografías digitales.

Usuario. Individuo que crea, modifica, almacena o distribuye un documento electrónico.

XMP. Plataforma Extensible de Metadatos (Extensible Metadata Platform). Es un tipo de lenguaje especificado extensible de marcado (eXtensible Markup Language) introducido por Adobe System y usado en los documentos PDF y aplicaciones de fotografía o de retoque fotográfico para introducir metadatos específicos.

ANEXO F. REFERENCIAS

- Real Decreto 3/2010, de 8 de enero, por el que se regula el Esquema Nacional de Seguridad en el ámbito de la Administración Electrónica.
- Real Decreto 951/2015, de 23 de octubre, de modificación del Real Decreto 3/2010, de 8 de enero, por el que se regula el Esquema Nacional de Seguridad en el ámbito de la Administración Electrónica.
- Real Decreto 4/2010, de 8 de enero, por el que se regula el Esquema Nacional de Interoperabilidad en el ámbito de la Administración Electrónica.
- Resolución de 19 de julio de 2011, de la Secretaría de Estado para la Función Pública, por la que se aprueba la Norma Técnica de Interoperabilidad de Documento Electrónico.
- CCN-STIC-818 de Herramientas de Seguridad en el ENS.
- PAe – Portal de administración electrónica – Archivo electrónico.
- https://administracionelectronica.gob.es/pae_Home/pae_Estrategias/Archivo_electronico.html#.WFK2dmfmqic
- Sociedad Española de Documentación e Información Científica (SEDIC) - <http://www.sedic.es/>
- Oracle White Paper – March 2007 – “The Risks of Metadata an Hidden Information”
- NISO (National Information Standards Organization) booklet “Understanding Metadata”
- Guidelines for Handling Image Metadata. Version 2. November 2010. Metadata Working Group.
- Web de Soporte de Microsoft - <https://support.office.com>
- Web de soporte Apache OpenOffice - <https://wiki.openoffice.org>
- Web de soporte Adobe Acrobat - <https://helpx.adobe.com>