

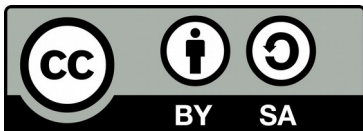
# **Software libre para la gestión documental**

Módulo 2

## **Herramientas libres para la gestión documental**

Guillermo Castellano





Herramientas libres para la gestión documental

Guillermo Castellano Casas

Septiembre 2017

Maquetado por Guillermo Castellano Casas con LibreOffice 5.0 en Linux Mint 17.3.

Fuente Alegreya Sans de Juan Pablo del Peral.

Imagen de portada de [Pexels](#).

Infografías diseñadas por Ana Amelia Patiño Esteo con Inkscape en Ubuntu 16.04.

Puedes compartir, es decir, copiar y distribuir el material en cualquier medio o formato y adaptar, es decir, mezclar, transformar y crear a partir del material. Las condiciones para esto son reconocer adecuadamente la autoría, proporcionando un enlace a la licencia e indicando si se han realizado cambios (haciéndolo de cualquier forma razonable, sin sugerir que se tiene el apoyo del licenciador o lo recibe por el uso que hace), y compartiendo con la misma licencia si se mezcla, transforma o crea a partir del material.

# Herramientas libres para la gestión documental

## 1. Ventajas del software libre para la gestión documental

### 1.1. Menor coste total de la solución

Aunque no es la gratuidad, sino la libertad lo que caracteriza el software libre, muchos usuarios descubren este tipo de herramientas buscando alternativas más económicas al software privativo. Como vimos en el módulo anterior, el modelo del software libre y de código abierto permite obtener de forma completamente gratuita el programa y pagar únicamente por los servicios que precisemos, como la implementación, la formación o el soporte técnico. Esto conlleva que las organizaciones con menos recursos, como pueden ser un centro de archivo, una biblioteca o una pequeña empresa, también puedan acceder a soluciones de gestión documental profesionales.

## 5 Falsos mitos sobre el software libre

<b>Es menos seguro</b> El software libre no solo no compromete la seguridad sino que la hace accesible a todo el mundo.	<b>Significa software gratuito</b> No es la gratuidad lo que define al software libre, sino la libertad de ejecutarlo, estudiarlo, distribuirlo y mejorarlo.	<b>Es más difícil de usar</b> La complejidad del programa no depende de que su código sea abierto o cerrado. Software de uso diario como Android, WordPress o Firefox es libre.	<b>Es menos compatible</b> El software libre fomenta los formatos abiertos, que favorecen la interoperabilidad y la preservación de la información digital.	<b>No vale para la empresa</b> Un 75% de las empresas españolas utiliza alguna tipología de software libre.
				

5 falsos mitos del software libre.

Fuente: Elaboración propia.

Algunos desarrolladores de software privativo argumentan que el coste de la licencia sólo representa una pequeña parte del coste total de su programa, en el cual están incluyendo ya la formación, la implementación y otros servicios. Según ellos, cuando sumamos estos servicios agregados al coste del producto, la opción con software libre no siempre es más económica que la opción con software privativo. Un ejemplo de este razonamiento lo ofrece el gerente de MTC Soft, Fernando Moreno-Torres (2016), quien afirma en su blog que, “cuando una empresa

## Módulo 2. Herramientas libres para la gestión documental

decide instalar un programa de gestión documental, el coste de la licencia llega a ser casi irrelevante. Es mucho más costoso el trabajo de instalación, configuración inicial, formación de los usuarios y soporte técnico a lo largo de la vida útil del software. Y si añadimos a la lista el coste de digitalizar un archivo con 200 ó 300 mil documentos, la parte correspondiente a la licencia es muy pequeña”.

Ciertamente, al valorar el coste de un proyecto de implementación de un software de gestión documental debemos tener en cuenta todas las partidas del presupuesto, y no solamente el coste de las licencias. Puede ocurrir que el coste total del proyecto sea mayor con una solución de software libre que con una solución privativa, pero no es lo normal. Fernando Moreno-Torres comete un error de metodología al comparar el coste total de implementar su software de gestión documental, Archivitech, con el coste total de implementar Alfresco, que es una herramienta de gestión documental mucho más completa y, por lo tanto, más compleja tanto de instalar y parametrizar como de enseñar a utilizar..

Dadas las mismas funcionalidades, el software libre tiende a ser más económico que el software privativo porque el desarrollo distribuido en comunidades abiertas abarata los costes de desarrollo y la inexistencia de un monopolio del código fuente posibilita que varias empresas compitan en el mercado por ofrecer un servicio más competitivo. Muchas de las prácticas abusivas en el sector del software se deben al monopolio ejercido por las empresas propietarias del código, que pueden fijar los precios no solamente sobre el programa, sino también sobre los servicios asociados al programa. Cuanto más proveedores de software trabajen con un determinado software de código abierto, más ofertas tendrá el potencial comprador, que podrá descartar a los proveedores con precios menos interesantes e, incluso, decidir encargar el proyecto a su propio personal si no le convence ninguna de las opciones que encuentra en el mercado.

Alfresco (2008) publicó un estudio comparativo del coste total de su software de gestión documental y el coste total de OpenText, Documentum y SharePoint. El resultado de este estudio fue que la alternativa de software libre resultaba entre un 42% y un 96% más económica que un ECM privativo. Aunque ha transcurrido casi una década desde la publicación de esos datos y existen más actores en el mercado de los ECM de código abierto que Alfresco, la realidad que me he encontrado en mi trabajo como consultor de gestión documental no es muy diferente a esas cifras. Como expliqué recientemente en Archivamos (Castellano 2017, p. 32), “en algunos proyectos en los que he trabajado, implementar un software privativo hubiera costado diez veces más que la alternativa libre”.

### 1.2. Eficiencia

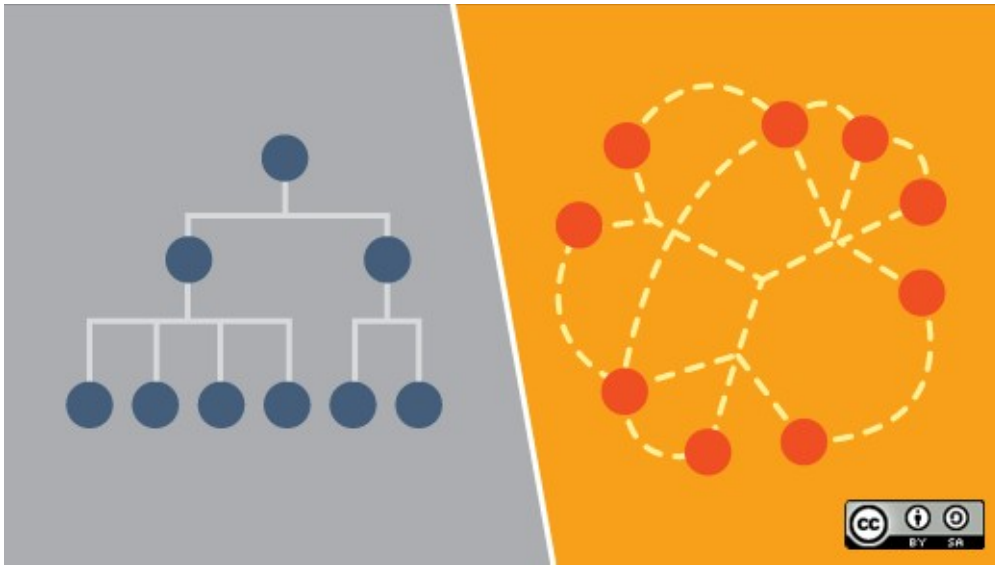
Con todo, el coste no es el principal aspecto en el que se fijan ni las empresas privadas ni las administraciones públicas a la hora de seleccionar una herramienta informática. Uno de los datos más importantes de un proyecto de este tipo es el denominado retorno de inversión o ROI, que es el tiempo que transcurre desde la adquisición de la herramienta hasta que el ahorro y/o los nuevos ingresos logrados gracias al software de gestión documental igualan el coste total del proyecto. El ROI nos da el momento en el que volveremos a tener disponible el dinero invertido en la herramienta, así como el ahorro y/o las ganancias que nos generará a partir de ese momento o, dicho de otra manera, su eficiencia, su valor.

El estudio sobre el uso de software libre en la administración pública estatal realizado por el CENATIC (2011, p. 56) muestra que, “si estableciéramos un ranking de los diferentes criterios que guían las políticas y prácticas de adquisición de software seguidas por los organismos públicos entrevistados, encontramos que cumplimiento de expectativas y rendimiento obtenido serían los elementos claves e irrenunciables en el proceso”.

La eficiencia es, precisamente, uno de los aspectos que más han destacado sobre el software libre los defensores del modelo de desarrollo tipo bazar que vimos en el primer módulo. Recordemos que la Ley de Linus afirma que “dada una base suficiente de desarrolladores asistentes y beta-testers, casi cualquier problema puede ser caracterizado rápidamente, y su solución ser obvia al menos para alguien” (Raymond 1998). El desarrollo distribuido por una comunidad abierta no solo reduce los costes del desarrollo, sino que también agiliza las iteraciones, lo que permite que los errores se corrijan antes y que los usuarios tengan antes las mejoras que necesitan.



## Módulo 2. Herramientas libres para la gestión documental



Modelo catedral y modelo bazar.  
Fuente: opensource.com.

Las empresas desarrolladoras de software de gestión documental privativo argumentan que las alternativas libres tienen capacidades limitadas. Señala Kyocera (2017) que estos programas se integran peor con aplicaciones de terceros, limitan el número de documentos que se puede subir y ponen marcas de agua en los documentos.

Ninguno de los ECM open source del mercado limita el número de documentos que se puede subir. Al igual que ocurre con los programas privativos, el volumen de documentos que podrá manejar el usuario dependerá solamente del espacio reservado para el repositorio en el servidor. Se trata, por lo tanto, de una cuestión relacionada con el hardware del que disponga la organización (o el servicio que haya contratado si se ha decantado por una solución de gestión documental en la nube), no con la aplicación escogida.

Tampoco es cierto que el software de gestión documental libre ponga marcas de agua en los documentos contra la voluntad del usuario. Las herramientas que hacen eso son software privativo que el desarrollador ofrece de forma gratuita con limitaciones como ésta (freeware) para que el usuario se vea forzado a comprar la versión completa del software. Una limitación así no tendría sentido en un gestor documental libre: al disponer de la libertad de ejecutar, estudiar, mejorar y distribuir el código, nada impediría que los usuarios eliminaran las marcas de agua obligatorias y subieran un fork del programa sin esa limitación tan molesta.

### 1.3. Seguridad

El segundo de los inconvenientes que señala el artículo de Kyocera (2017) es que los programas de gestión documental libres están por detrás de los privativos en materia seguridad. Se trata de un argumento muy utilizado por los vendedores de software privativo, plenamente conscientes de la importancia cada vez mayor que los usuarios le otorgan. Y no es para menos. Pero, ¿realmente compromete el software libre la seguridad?

Aunque es natural creer que un programa cuyo código está publicado en internet es menos seguro que un programa cuyo código sólo conocen sus desarrolladores, si mantener en secreto el código fuente del software lo hiciera seguro, entonces no sería necesario utilizar herramientas de seguridad en Microsoft Windows, el sistema operativo privativo por excelencia. La seguridad por oscuridad, como se conoce al principio de seguridad que utiliza el secreto de diseño y de implementación para garantizar la seguridad, no es eficaz. Y, ello, por varios motivos:

1. Al ocultar el código, la empresa desarrolladora oculta deliberadamente posibles vulnerabilidades a los usuarios, que no pueden protegerse de lo que no conocen.

## Módulo 2. Herramientas libres para la gestión documental

2. La falta de libertad para estudiar el código hace que otros desarrolladores no puedan contribuir a arreglar esas vulnerabilidades o a detectar errores que han pasado desapercibidos a la propia empresa.
3. Parte o la totalidad del código se puede conocer por medio de la ingeniería inversa o las filtraciones.

Desarrollar en código abierto no garantiza que el programa no tenga vulnerabilidades, pero permite que la comunidad de desarrolladores y usuarios las encuentre y arregle antes. El énfasis de las comunidades de desarrollo de software libre en la seguridad es bien conocido en sectores que dependen especialmente de ella, como el bancario. Como afirmó en una entrevista Eneko Astigarraga, presidente de ESLE, “los grandes bancos también utilizan el software libre precisamente por seguridad, porque hay miles de desarrolladores mirando la plataforma y es más fácil detectar los problemas” (Mariscal 2016).

El artículo afirma abiertamente que “la mayoría de los sistemas de gestión open source no encriptan los documentos” (Kyocera, 2017). Sin embargo, todos los ECM que vamos a analizar en este curso permiten encriptar la información sin software adicional.



El software libre sigue un modelo de desarrollo colaborativo  
Fuente: opensource.com.

### 1.4. Otras ventajas del software libre: interoperabilidad, transferencia del conocimiento y preservación digital

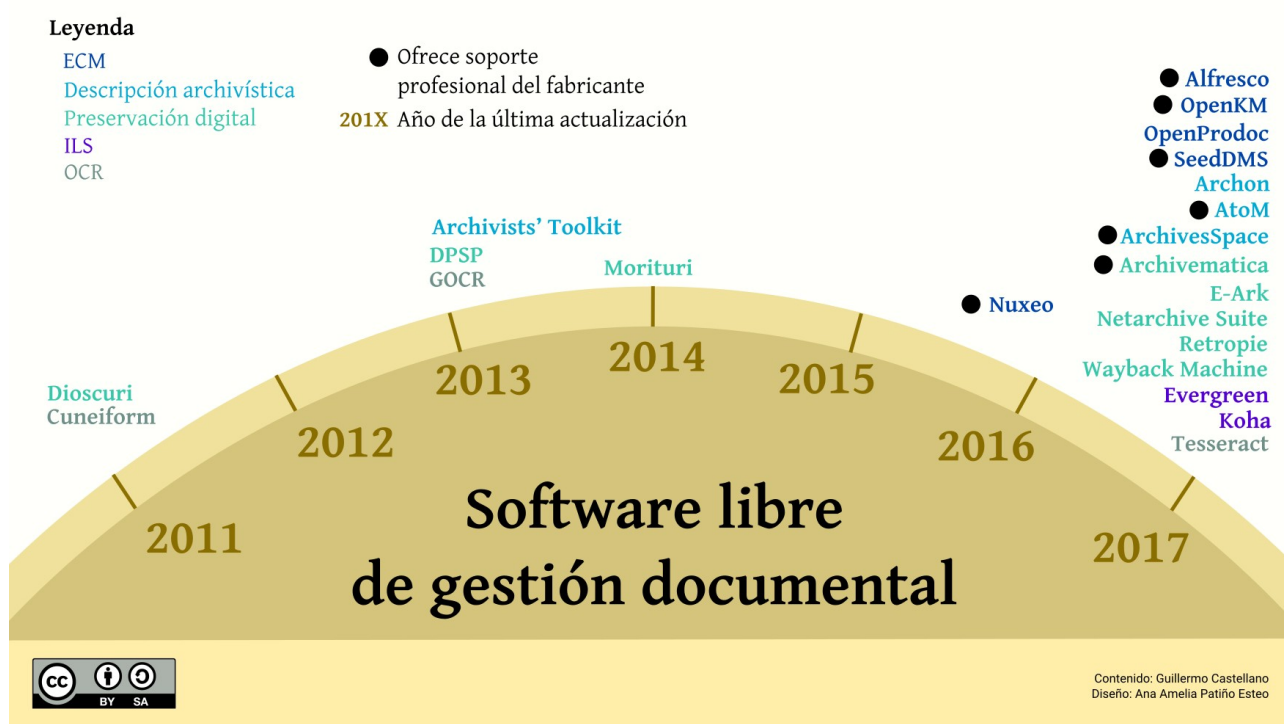
Una de las principales ventajas de este tipo de soluciones es su interoperabilidad. Como explicaremos con más detalle en el cuarto módulo, estas herramientas están construidas para integrarse fácilmente con múltiples aplicaciones, repositorios y sistemas. Suelen estar desarrolladas en Java, lo que posibilita su ejecución como servicio en servidores y contenedores de aplicaciones Java como Tomcat, Jetty, JBoss y OSGi, así como su compatibilidad con un amplio ecosistema de software de código abierto. De hecho, estas aplicaciones están concebidas como sistemas modulares que conectan diferentes programas de terceros (Lucene, Solr, Tesseract, jBPM, Bonita BPM, Hibernate, PostgreSQL, LibreOffice, etc.) con un núcleo, al que el usuario accede por medio de una interfaz gráfica. Algunas de ellas también soportan CMIS (siglas en inglés de Content Management Interoperability Services), un estándar abierto aprobado por OASIS que permite la interoperabilidad entre diferentes sistemas de gestión de contenidos.

Otra ventaja del software libre para la gestión documental es que las organizaciones pueden transferir a la sociedad las mejoras desarrolladas sobre sus aplicaciones. Este modelo de cooperación es lo que se conoce como un win-win (“ganar-ganar”): los recursos que invierten otros usuarios de la comunidad en mejorar el programa nos benefician a nosotros y los recursos que invertimos nosotros benefician a toda la comunidad<sup>1</sup>.

Finalmente, el software libre también favorece la conservación a largo plazo de la información, como veremos en los dos últimos módulos de este curso.

## 2. Panorámica del software libre de gestión documental

En el módulo anterior vimos que a finales de la década de los noventa importantes empresas del sector informático comenzaron a invertir en software libre y que desde algunos sectores del movimiento del software libre se promovió este acercamiento de las empresas, creando la denominación de “software de código abierto”. A comienzos y mediados de la década del 2000, surgieron las primeras alternativas libres a las herramientas privativas que entonces dominaban el mercado del software libre de gestión documental: MyDMS (2002), Nuxeo CPS (2003), Alfresco (2005), OpenKM (2006) e ICA-AtoM, Archon y Archivists' Toolkit (2006). Hoy disponemos de una veintena de aplicaciones libres para gestionar archivos y bibliotecas, sin contar programas auxiliares como los convertidores de formato y los antivirus.



Software libre de gestión documental.

Fuente: Elaboración propia.

Cuando hablamos de software libre de gestión documental, normalmente nos vienen a la cabeza los ECM (Enterprise Content Management). Estos programas sirven para gestionar los documentos y otros contenidos generados por una organización en el desarrollo de su actividad. A diferencia de los repositorios documentales, no se limitan a almacenar y servir documentos, sino que permiten controlar las versiones, establecer reglas de acceso y de conservación, definir flujos de trabajo, buscar mediante metadatos, automatizar procesos de negocio, etc. Las alternativas libres más conocidas son Alfresco y Nuxeo, pero también merecen nuestra atención OpenKM, OpenProdoc y SeedDMS.

Tenemos también herramientas específicas de descripción archivística, que sirven tanto para describir y como para difundir en la red los fondos de los archivos históricos. La más conocida de todas en nuestro entorno es AtoM (originalmente, ICA-AtoM), desarrollada a instancias de la UNESCO por el ICA, en colaboración con la Escuela Holandesa de Archivos y la compañía canadiense Artefactual Systems. Existe ArchivesSpace, resultado de la fusión de Archon y Archivists' Toolkit, herramientas más conocidas en el ámbito estadounidense.

## Módulo 2. Herramientas libres para la gestión documental

Si lo que necesitamos es conservar a largo plazo nuestros documentos electrónicos, disponemos de una amplia variedad de programas de preservación digital, entre los que destacaría Archivematica, las herramientas del proyecto E-Ark y DPSP (Digital Preservation Software Platform) y la plataforma NetarchiveSuite, que integra las principales herramientas desarrolladas por la comunidad para el archivo de páginas web (entre ellas, Wayback Machine). Podemos mencionar también RetroPie y Dioscuri, dos emuladores de videojuegos, y Morituri, que permite digitalizar pistas de audio de CD sin pérdida de información. Hay más programas susceptibles de ser utilizados dentro de una estrategia de preservación digital, como los convertidores de formato y los antivirus, como veremos en los dos últimos módulos.

Las tecnologías de reconocimiento óptico de caracteres, más conocidas como OCR (Optical Character Recognition), permiten automatizar procesos que impliquen captura de datos, como la tramitación de un pedido recibido por correo electrónico o el archivo de documentación. El OCR open source más preciso es Tesseract, creado por HP y desarrollado desde 2006 por Google, pero existen también Cuneiform y GOCR.

Aunque no pertenecen a la categoría de software libre de gestión documental, en la infografía están incluidos también los ILS (Integrated Library System) Evergreen y Koha para mostrar que las alternativas de código abierto se están consolidando también en el sector de las bibliotecas.

El mercado del software de gestión documental goza de buena salud a día de hoy: lo vemos en el hecho de que el 75% del software libre de gestión documental ha recibido actualizaciones en 2017 ó 2016 y disponemos de varias opciones si precisamos soporte del fabricante, sobre todo en el campo de los ECM. Como vamos a explicar en el siguiente módulo, las funcionalidades y el coste no lo son todo a la hora de seleccionar una herramienta informática y debemos evaluar siempre factores como las actualizaciones y el soporte que recibiremos.

## Resumen

Las principales ventajas del software libre para la gestión documental son:

1. Su menor coste.
2. Su mayor eficiencia.
3. Su mayor seguridad.
4. Su mayor personalización e interoperabilidad.
5. Que favorece la transferencia de conocimiento a la sociedad y la cooperación.
6. Que favorece la preservación digital.

En el mercado del software libre de gestión documental disponemos de una veintena de aplicaciones, entre las cuales destacan las siguientes:

1. ECM: Alfresco, Nuxeo, OpenKM, OpenProdDoc y SeedDMS.
2. Descripción archivística: AtoM, Archon, ArchivesSpace.
3. Preservación digital: Archivematica, E-Ark, DPSP, Dioscuri, Morituri, convertidores de formato y antivirus.
4. OCR: Tesseract, Cuneiform y GOCR.
5. ILS: Koha y Evergreen.

El 75% de estas herramientas está actualizado y hay varias opciones con soporte del fabricante, notablemente entre los ECM.



## Bibliografía

- Alfresco. *Total cost of ownership for Enterprise Content Management* [en línea]. 2008 [consulta 05-05-2017]. Disponible en: [http://876solutions.com/sites/default/files/white-papers/Alfresco White Paper TCO for ECM.pdf](http://876solutions.com/sites/default/files/white-papers/Alfresco%20White%20Paper%20TCO%20for%20ECM.pdf).
- ARAÑA Cruz, J.; HERRERA Tejada, C. "ICA-AtoM, una buena herramienta para la difusión de los archivos en la web". En: *Ateneo de Madrid* [en línea]. Marzo 2011 [consulta 05-05-2017]. Disponible en: <https://www.ateneodemadrid.com/index.php/content/download/9985/103607/file/ICA-AtoM,%20una%20buena%20herramienta%20para%20la%20difusi%C3%B3n%20de%20los%20archivos%20en%20la%20web.pdf>.
- CASTELLANO Casas, G. "Archivando con software libre: aplicaciones de código abierto para la gestión documental". En: *Archivamos*. 2017, n. 103, pp.103-104. ISSN 1576-320X.
- CENATIC. "El software libre en los organismos públicos del ámbito estatal: informe de resultados de la encuesta sobre el software de fuentes abiertas en la administración pública (ESFA-AGE) 2011". En: *CENATIC* [en línea]. 4 enero 2012 [consulta 05-05-2017]. Disponible en: <http://www.cenatic.es/dossier/panel-informe-age2011>.
- Kyocera. "Ventajas e inconvenientes de una gestión documental open source". En: *El blog de Soluciones para la transformación digital en procesos, data, documentos e impresión para CIOs* [blog]. 1 julio 2017 [Consulta 11-09-2017]. Disponible en: <http://smarterworkspaces.kyocera.es/blog/ventajas-e-inconvenientes-una-gestion-documental-open-source/>.
- MARISCAL, M. "Muchas empresas optan por el software libre porque es más innovador y garantiza la seguridad". En: *Noticias de Gipuzkoa* [en línea]. 6 noviembre 2016 [consulta 05-05-2017]. Disponible en: <http://www.noticiasdegipuzkoa.com/2016/11/06/economia/muchas-empresas-optan-por-el-software-libre-porque-es-mas-innovador-y-garantiza-la-seguridad>.
- MORENO-TORRES Camy, F. "¿Software libre para todos?" En: *Gestión documental para gente (casi) normal* [blog]. 7 octubre 2013 [consulta 05-05-2017]. Disponible en: <https://gestiondocumentalparagentenormal.com/2013/10/07/software-libre-para-todos/>.
- RAYMOND, E. *La catedral y el bazar* [en línea]. 1998 [consulta 05-05-2017]. Disponible en: <http://biblioweb.sindominio.net/telematica/catedral.html>.