

PROSPECTIVA DE LOS SISTEMAS GESTORES DE DOCUMENTOS: HOMENAJE A PEDRO MANUEL DÍAZ ORTUÑO

Pedro Manuel Díaz Ortúño

Departamento de Información y Documentación. Universidad de Murcia.

<https://orcid.org/0000-0002-2975-766X>

Resumen: Este estudio examina la evolución y perspectivas de los sistemas de gestión documental (SGD) en el marco de la transformación digital. Analiza la integración de la gestión documental en los procesos organizativos, conforme a las normas ISO 15489 e ISO 16175, que garantizan autenticidad, integridad y disponibilidad de los documentos. Se describen las fases del ciclo de vida documental —captura, clasificación, retención, preservación y acceso— y la función de los SGD en la mejora de la eficiencia, trazabilidad y cumplimiento normativo. El estudio compara las soluciones de software abierto y propietario, valorando su grado de flexibilidad, personalización y soporte técnico. En la parte final, el autor concluye que los gestores documentales abiertos, como Alfresco, OpenKM y OpenProdoc, ofrecen una sólida base para la implantación de políticas de gestión documental interoperables y sostenibles, al tiempo que promueven la transparencia, la colaboración y la preservación digital a largo plazo en las organizaciones.

Palabras clave: Gestión documental; Sistemas Gestores de Documentos; Norma ISO 15489; Norma ISO 16175; Transformación digital

Title: PROSPECTS OF DOCUMENT MANAGEMENT SYSTEMS: A TRIBUTE TO PEDRO MANUEL DÍAZ ORTUÑO.

Abstract: This study examines the evolution and future prospects of Document Management Systems (DMS) within the framework of digital transformation. It analyzes the integration of document management into organizational processes in accordance with ISO 15489 and ISO 16175 standards, which ensure the authenticity, integrity, and availability of documents. The phases of the document life cycle—capture, classification, retention, preservation, and access—are described, as well as the role of DMS in enhancing efficiency, traceability, and regulatory compliance. The study compares open-source and proprietary software solutions, assessing their degree of flexibility, customization, and technical support. In its final section, the author concludes that open-source document managers such as Alfresco, OpenKM, and OpenProdoc provide a solid foundation for implementing interoperable and sustainable document management policies, while promoting transparency, collaboration, and long-term digital preservation within organizations.

Keywords: Document Management; Document Management Systems; ISO 15489 Standard; ISO 16175 Standard; Digital Transformation.

Copyright: © 2025 Servicio de Publicaciones de la Universidad de Murcia (Spain). Este es un artículo de acceso abierto distribuido bajo los términos de la licencia Creative Commons Reconocimiento 4.0 Internacional (CC BY 4.0).

Datos de edición: Este artículo es una adaptación de un capítulo de libro del mismo título incluido en la obra colectiva Vera Baceta, M.A. y Martínez Méndez, F.J. (eds), (2022). *La transformación digital de entidades del Tercer Sector de Acción Social: un marco para la gestión documental.* CEPAIM-Universidad de Murcia. <http://hdl.handle.net/10201/126388>

1 INTRODUCCIÓN

Los documentos, en papel o electrónicos, forman parte integral de cualquier entidad, organización empresa o institución. Dentro de los documentos más comunes que se generan en una organización encontramos actas, informes, contratos, facturas, acuerdos, albaranes, presupuestos, impuestos, proyectos, memorias, etc. y hacen de su correcta gestión un desafío de carácter permanente. La gestión documental hace referencia al conjunto de normas y técnicas empleadas para administrar cualquier documento de una organización, permitiendo poder recuperarlo en cualquier momento y controlar el tiempo que debe guardarse.

Las organizaciones son conscientes de la importancia de gestionar adecuadamente el flujo de documentos asociados a sus procesos de funcionamiento, desde la creación o captura del documento, hasta su eliminación final. Bustelo-Ruesta (2011, 130) afirma que “el principal objetivo de cualquier modelo de gestión de documentos debe ser la integración de los procesos y controles documentales en los procesos de trabajo”. La Norma ISO 15489 para la gestión de documentos propone “la integración de la gestión de documentos en los sistemas y los procesos de la organización” considerando que, por la vinculación entre procesos y flujo de documentos asociados, “los sistemas de información, las aplicaciones de negocio y los sistemas de comunicación, así como los procesos de gestión que estos respaldan, se deberían diseñar, modificar o revisar para que la creación e incorporación de los documentos apropiados sea una tarea habitual dentro de las actividades de la organización”.

El software de gestión documental facilita el almacenamiento seguro de los documentos y ayuda a encontrar el que se necesita. Un sistema de gestión de documentos (SGD) tiene como objetivo la gestión ordenada del capital intelectual de una empresa mediante la documentación de todo conocimiento presente dentro de ella. La necesidad de organizar los documentos electrónicos es un factor importante para las organizaciones, pasando a un segundo plano las copias en papel. Pretende la captura, almacenado, actualización y rastreo de la documentación, con el fin de reducir el riesgo de pérdida por no documentación, renuncia o despido de trabajadores.

La gestión de contenidos mediante sistemas de gestión documental tiene como objetivo hacer que la gestión de la información corporativa sea más eficiente y efectiva por la simplificación del almacenamiento, seguridad, control de versiones, y la retención. Entre los beneficios para una organización destacan una mayor eficiencia y una mejor control y reducción de costos. No es suficiente la “gestión” del contenido, es importante la capacidad de acceso a la versión adecuada de un documento o registro: “*ECM cubre la gestión de la información en todo el ámbito de una empresa si esa información se presenta en forma de un documento en papel, archivo electrónico, flujo de impresión de base de datos, o incluso un correo electrónico*” (AIIM, 2021).

Según AIIM (2021) la gestión de contenido organizacional es fruto de la recopilación y organización sistemáticas de información que va a ser utilizada por una audiencia designada: ejecutivos de negocios, clientes, etc. No es una tecnología única, ni una metodología ni un proceso, es una combinación dinámica de estrategias, métodos y herramientas utilizadas para capturar, gestionar, almacenar, preservar y entregar información que respalda los procesos organizacionales clave a lo largo de todo su ciclo de vida:

- La captura se reduce a la entrada del contenido en el sistema. Implica la conversión de la información de documentos en papel a formato electrónico a través de la exploración, escaneo o digitalización. Las tecnologías de captura también abarcan la creación de metadatos que describen las características de un documento para su fácil localización a través de la tecnología de búsqueda.
- Gestionar es hacer que el contenido pueda ser encontrado y utilizado por cualquier persona a la que esté destinado. Esta categoría incluye cinco áreas de aplicación tradicionales:
 - Gestión de documentos.
 - Gestión de contenidos web (CMS).
 - Gestión de documentos de archivo (‘Record Management’).
 - Gestión de flujos de trabajo y procesos de gestión (BPM)
 - Colaboración (‘groupware’).
- Almacenar significa encontrar un lugar apropiado en la infraestructura en un sistema de gestión de contenido formal u otra solución (sistemas de archivos, gestor de contenidos, bases de datos, depósito de datos o ‘datawarehouse’, etc.).
- Preservar se refiere a la atención a largo plazo, a “archivar”, a la práctica de proteger el contenido para que pueda utilizarse cuando la organización necesite que esté disponible. La conservación se realiza generalmente teniendo en cuenta las características de gestión de documentos de archivo y el cumplimiento de regulaciones gubernamentales y corporativas.
- Y con entregar, se trata de poner la información en las manos de las personas adecuadas cuando la necesitan.

Las herramientas de apoyo a la gestión documental se orientan por tanto a: la gestión de contenidos empresariales; el almacenamiento, preservación, búsqueda, visualización de documentos electrónicos en diferentes formatos; flujos de trabajo, asignación de metadatos, administración de usuarios, gestión de versiones de documentos, y edición de documentos con herramientas ofimáticas.

Los documentos de archivo son los acumulados por la actividad de una organización que adquieren una función informativa después de cumplir con la función principal para la cual fueron producidos. La gestión de documentos de archivo forma parte de los sistemas de gestión documental, y permite implementar las políticas y procedimientos para la retención, control y disposición de los documentos de archivo para el cumplimiento normativo y corporativo.

Los sistemas de gestión de procesos (BPMS) soportan el ciclo de vida de los procesos de gestión: modelado, ejecución, monitorización, medición, análisis, auditoría y optimización. Automatizan procesos complejos, pero no destacan por una visión integrada de procesos y documentos con metadatos generados. Por este motivo, suelen disponer de interfaces de integración con sistemas ECM para el intercambio de archivos. La notación BPMN 2.0 y CMIS permiten mecanismos de interoperabilidad flexibles y facilitan la integración de sistemas de planificación de procesos con los sistemas de gestión documental.

2 LAS NORMAS ISO VINCULADAS A LA GESTIÓN DOCUMENTAL

La documentación que genera una entidad u organización para realizar su actividad a través de aplicaciones de software requiere contar con normas y estándares que permitan garantizar la autenticidad, fiabilidad, integridad y disponibilidad. ISO ('International Organization for Standardization') junto con el Consejo Internacional de Archivos (ICA), creó la Norma ISO 16175 en el año 2006 (adoptada por la organización de normalización española AENOR bajo el título *Principios y requisitos funcionales para documentos en entornos de oficina electrónica*) con el propósito de desarrollar unos requisitos funcionales para productos de software utilizados en las organizaciones para la gestión de documentos, en coherencia con la Norma ISO 15489 para la gestión de documentos.

La parte 1 de esta última norma define los conceptos y principios desde los que desarrollar los enfoques para la creación, captura y gestión de documentos, describiendo conceptos y principios relativos a:

- a) los documentos, los metadatos de gestión de documentos y los sistemas de gestión documental;
- b) las políticas, la asignación de responsabilidades, la supervisión y la formación como apoyo a la gestión efectiva de documentos;
- c) el análisis del contexto de la organización y la identificación de los requisitos de gestión de documentos;
- d) los instrumentos de gestión de documentos;
- e) los procesos para crear, capturar y gestionar documentos.

Los SGD comprenden un número de elementos combinados de manera que puedan satisfacer los requisitos identificados en un entorno determinado. La Norma ISO 16175-1 proporciona un modelo, requisitos funcionales de alto nivel, con información explicativa y orientación de uso, para cualquier aplicación de software destinada a gestionar documentos electrónicos. Esta serie de requisitos se centra en lo necesario para asegurar que los documentos se gestionan de forma correcta. Se estructuran en cuatro áreas clave:

- a) Captura y clasificación de documentos: capturar y/o importar/ingresar evidencias de las actividades u operaciones de gestión. Esto implica identificar conjuntos de información digital para que sirvan como documentos y vincularlos a su contexto de gestión mediante metadatos.
- b) Retención y disposición de documentos: los documentos deben conservarse y mantenerse accesibles para las personas autorizadas durante el tiempo que sea necesario para dar respuesta a las necesidades legislativas y corporativas. De conformidad con las autoridades de disposición autorizadas, los documentos deberían conservarse y eliminarse de forma gestionada, sistemática y auditável.
- c) Integridad y mantenimiento de documentos: poder registrar cualquier interacción o cambio en los documentos. Los metadatos que registran estas interacciones y modificaciones deberían incluir el nombre de la persona responsable del documento, una marca de tiempo y detalles de cualquier cambio. Los metadatos deberían estar vinculados de forma permanente a los documentos mientras existan.
- d) Búsqueda, uso e intercambio de documentos: las aplicaciones de gestión documental deben permitir la búsqueda, recuperación, uso, intercambio e interoperabilidad entre plataformas a lo largo del tiempo.



Figura 1: Requisitos de los sistemas gestores de documentos agrupados por área. Fuente: Elaboración propia.

Tabla I: Áreas de requisitos clave

R1 Captura y clasificación	Obligación
R1.1 Creación, captura e importación de documentos	
R1.1.1 captura de documentos y los metadatos de gestión de documentos asociados; R1.1.2 cuando una aplicación de gestión no pueda proporcionar la capacidad de ingresar documentos y metadatos asociados, otros componentes deben proporcionar esa capacidad de ingresar documentos y/o metadatos asociados desde otras aplicaciones o repositorios externos relacionados; R1.1.3 capturar los archivos adjuntos y objetos incrustados junto con los mensajes digitales como documentos vinculados o como un documento único compuesto cuando la aplicación cree o reciba documentos vía sistemas de mensajería electrónica (correo electrónico, mensajería de texto o mensajería instantánea); R1.1.4 permitir que la aplicación de gestión importe documentos electrónicos y metadatos asociados directamente desde una aplicación de gestión externa, ya sea de forma masiva o individual, asegurando la integridad del contenido y estructura de los documentos. Cuando los documentos que se vayan a capturar tengan criterios adicionales de verificación de integridad (firmas electrónicas, ' <i>checksum</i> ', etc.), el sistema que los capture debe poder verificar la validez de acuerdo con estos criterios (al menos en el momento de la captura).	Debe
R1.1.5 soportar la captura de documentos en formatos de archivo de uso común y/o en sus formatos nativos.	Debería
R1.2 Captura de metadatos de gestión de documentos	
R1.2.1 habilitar la captura y mantenimiento de metadatos para la gestión de documentos en cualquier momento de la existencia del documento, de acuerdo con uno o más esquemas de metadatos predeterminados; R1.2.2 poder asignar y vincular persistentemente identificadores únicos a cada documento y agregación de documentos;	Debe

R1.2.3 soportar la capacidad de detectar y capturar automáticamente metadatos preexistentes para la gestión de documentos procedentes de sistemas y procesos de gestión asociados; R1.2.4 poder capturar metadatos ingresados manualmente.	
R1.2.5 poder validar los valores de los metadatos frente a esquemas predeterminados y/o estándares sintácticos; R1.2.6 admitir formatos comerciales comunes (por ejemplo, XML) o combinaciones de formatos para elementos/propiedades de metadatos; R1.2.7 poder capturar y vincular de manera persistente los procesos de flujo de trabajo asociados con los documentos, incluidas las aprobaciones.	Debería
R1.2.8 permitir que los agentes autorizados agreguen anotaciones o notas a los documentos como metadatos vinculados de acuerdo con las reglas y políticas de negocio (por ejemplo, los agentes autorizados pueden incluir a personas que son objeto de los documentos).	Puede
R1.3 Clasificación de documentos	
R1.3.1 poder asociar documentos a nivel de objeto individual y/o agregación en su contexto de negocio; R1.3.2 respaldar la documentación de los cambios en el contexto de la gestión a lo largo del tiempo, conservando la capacidad de vincular el contexto de negocio con precisión a lo largo del tiempo.	Debe
R1.4 Gestión de cuadros de clasificación	
R1.4.1 gestionar y mantener un cuadro de clasificación aprobado.	Puede

R2 Retención y disposición	Obligación
R2.1 Retención, revisión, transferencia y destrucción de documentos	
R2.1.1 poder asignar un período de retención y disposición apropiado en la solicitud para cada documento y agregación de documentos y también permitir la actualización de estos períodos de retención y disposición si fuera necesario, debido a cambios en los requisitos legales o de negocio; R2.1.2 conservar los metadatos clave (según lo estipulado por los estándares locales/jurisdiccionales pertinentes), incluidos los metadatos que documenten la autorización de disposición, para el documento y/o agregación después de la disposición del documento y/o agregación; R2.1.3 almacenar el estado del documento (y/ o de la agregación de documentos) como «eliminado» o «transferido a X» desde la fecha de disposición del documento; R2.1.4 tener la capacidad de pausar la disposición, evitando que los documentos sean eliminados; R2.1.5 informar sobre el estado y la actividad de disposición cuando se realiza la disposición de los documentos dentro de la aplicación; R2.1.6 si lo requiere el calendario de conservación, asegurar que la destrucción de los documentos resulte en la eliminación total o inaccesibilidad del contenido de todos los documentos (incluidos todos los componentes de cada documento y cualquier copia de seguridad) según lo autorizado, y que no se puedan restaurar mediante funciones del sistema operativo o técnicas especializadas de recuperación de datos; R2.1.7 contar con la capacidad de transferir documentos de valor continuo junto con sus metadatos a un archivo interno de la organización o a un servicio de archivo de terceros autorizado.	Debe
R2.1.8 marcar automáticamente los documentos como elegibles para disposición y eliminarlos una vez que hayan expirado sus períodos de retención y, si así lo requiere el calendario de conservación, cuando los documentos hayan sido revisados por un agente/usuario autorizado; R2.1.9 restringir la operación de disposición a un administrador de la aplicación de gestión; admitir una variedad de factores desencadenantes de disposición (fecha de creación del documento, fecha de su última recuperación, fecha de apertura o cierre de una revisión y/o agregación de documentos, etc.).	Debería
R2.2 Migración y exportación de documentos	
R2.2.1 poder migrar/exportar documentos y metadatos asociados y, cuando corresponda, llevar a cabo agregaciones de documentos aplicables hacia un formato de archivo más apropiado o actualizado para los documentos para respaldar la retención y uso continuos, otra aplicación de gestión dentro de la	Debe

<p>organización, una aplicación de gestión de sustitución cuando la aplicación de origen se desmantela (“decomisiones”); y/o un sistema en una organización diferente;</p> <p>R2.2.2 asegurar que cualquier acción de migración/exportación pueda incluir:</p> <ul style="list-style-type: none"> todos los documentos y, en su caso, agregaciones de documentos; los metadatos asociados a los documentos y agregaciones de documentos exportados; y los metadatos del historial de eventos asociados con los documentos exportados; <p>R2.2.3 poder migrar/exportar documentos electrónicos y, en su caso, agregaciones de documentos de manera que:</p> <ul style="list-style-type: none"> no se degrada el contenido y la estructura de los documentos ni de las agregaciones de documentos; se mantienen las asociaciones entre los documentos exportados y sus metadatos asociados; y se mantienen las relaciones entre los componentes exportados de un documento, entre documentos exportados y, en su caso, entre agregaciones de documentos, para que sus vínculos estructurales se puedan reconstruir en la aplicación receptora; <p>R2.2.4 poder testear y verificar que la migración o exportación no ha degradado la integridad de los documentos y de los metadatos clave por debajo de los estándares mínimos establecidos por el negocio y/o por la jurisdicción;</p> <p>R2.2.5 permite migrar/exportar documentos más de una vez.</p>	
<p>R2.2.6 poder implementar la destrucción de los documentos de origen después de su migración/exportación exitosa a la nueva aplicación;</p> <p>R2.2.7 estar diseñado de tal manera que los agentes autorizados puedan acceder a las tablas de datos subyacentes y al contenido de la aplicación después de su desmantelamiento (“decomisión”) a través de un visor o usando tecnología de virtualización, cuando los documentos inactivos/cerrados solo se requieren para la retención a corto plazo y, por lo tanto, no para ser migrados;</p> <p>R2.2.8 soportar el control de acceso al contenido de los documentos aplicado a nivel de documento y la persistencia de ese control cuando el contenido se exporta/transmite/migra desde la aplicación de origen.</p>	Debería

R3 Integridad y mantenimiento	Obligación
R3.1 Autenticación y seguridad en la gestión de documentos	
<p>R3.1.1 asegurar que el contenido de los documentos pueda ser reparado o protegido de alteraciones no autorizadas; permitir controles sobre la alteración o edición de metadatos de acuerdo con las reglas de negocio estipuladas por la organización;</p> <p>R3.1.2 poder generar ‘checksums’, ‘hashes’ o implementar otros mecanismos para respaldar la verificación de integridad en un momento dado;</p> <p>R3.1.3 autenticar rutinariamente a cualquier usuario antes de permitir el acceso a la información si los documentos requieren un cierto nivel de permiso;</p> <p>R3.1.4 capturar y mostrar automáticamente los detalles de todas las operaciones de autenticación y relacionadas con la seguridad.</p>	Debe
<p>La aplicación de gestión debería, ya sea sola o junto con otras aplicaciones:</p> <p>R3.1.5 poder capturar y almacenar de forma persistente metadatos que documenten el uso de firmas digitales (fecha, hora y validación) con el documento;</p> <p>R3.1.6 capturar cualquier otro detalle de confirmación para firmas digitales de tal manera que puedan ser recuperados con el documento, pero sin comprometer la integridad de una clave privada;</p> <p>R3.1.7 apoyar la transmisión segura de documentos, incluido el cifrado de documentos para una transmisión segura cuando sea apropiado.</p>	Debería
R3.2 Almacenamiento, generación de informes y gestión de metadatos	
<p>R3.2.1 garantizar que los documentos y los metadatos asociados controlados por la aplicación se almacenan de forma persistente y segura y que permanecen accesibles y recuperables a lo largo del tiempo durante sus períodos mínimos de retención para los agentes autorizados;</p> <p>R3.2.2 poder producir informes sobre captura, uso y disposición de los documentos;</p> <p>R3.2.3 poder informar de las acciones realizadas en los documentos, ya sea por la propia aplicación o por agentes y administradores autorizados.</p>	Debe

<p>La aplicación de gestión debería, ya sea sola o junto con otras aplicaciones:</p> <p>R3.2.4 poder administrar, mantener y/o vincular a perfiles o esquemas de metadatos aprobados a lo largo del tiempo de manera que respalden la captura automatizada y el mantenimiento de valores de metadatos validados para documentos y agregaciones de documentos.</p>	Debería
R4 Descubrimiento, uso e intercambio	Obligación
R4.1 Búsqueda, recuperación, presentación, uso e interoperabilidad	
<p>R4.1.1 proporcionar herramientas para buscar y recuperar documentos y metadatos;</p> <p>R4.1.2 extraer y procesar documentos en un formato utilizable;</p> <p>R4.1.3 poder integrarse e interoperar con otros sistemas de información que tengan la funcionalidad adecuada cuando la aplicación de gestión no pueda por sí misma realizar procesos de gestión de documentos.</p> <p>R4.1.4 soportar el uso del sistema a través de dispositivos móviles;</p> <p>R4.1.5 permitir a los agentes configurar la aplicación para crear vistas de documentos, búsquedas o servicios de presentación personalizados;</p> <p>R4.1.6 permitir la colaboración con terceros externos y para el intercambio de documentos y/o metadatos con plataformas externas, redes, servicios y espacios de trabajo colaborativos;</p> <p>R4.1.7 poder hacer que los datos aprobados almacenados dentro de la aplicación estén disponibles para ser recolectados por servicios y aplicaciones externos como datos abiertos vinculados ('linked open data' o LOD) a través de una interfaz de protocolo de aplicación utilizando protocolos como OAI-PMH o CMIS.</p>	Debe
R4.2 Restricciones y permisos de acceso	Puede
<p>R4.2.1 aplicar restricciones de seguridad y acceso asegurando que solo los agentes autorizados pueden acceder a los documentos adecuados según sus derechos de acceso;</p> <p>R4.2.2 aplicar protocolos de seguridad y acceso para proteger el contenido de los documentos y de sus metadatos del acceso no autorizado, alteración o destrucción;</p> <p>R4.2.3 crear y mantener metadatos de acceso, uso y seguridad, generando registros de eventos seguros que documenten el acceso y uso de los documentos para cada documento/agregación de documentos específico.</p>	Debe
R4.3 Duplicación, extracción y redacción	
<p>R4.3.1 contar con mecanismos que permitan la duplicación autorizada de documentos o extractos de documentos para su uso dentro de la propia aplicación o en otras aplicaciones y/u organizaciones, de acuerdo con las reglas de negocio predeterminadas.</p> <p>R4.3.2 permitir la creación de un extracto de un documento, mediante el que la información sensible se elimina u oculta a la vista en el extracto, mientras que el documento de origen permanece intacto, asegurando que los metadatos que documenten la acción de extracción se generan y capturan.</p>	Debe Debería

3 LOS SISTEMAS DE GESTIÓN DE DOCUMENTOS

Las organizaciones necesitan disponer de una estructura organizada y segura para sus documentos. Para ayudar con esta organización y establecer restricciones de acceso a documentos al personal no autorizado autorizaciones deciden dotarse de un sistema de gestión documental (SGD). La función principal de cualquier gestor documental es manejar los documentos y expedientes necesarios para una organización o equipo de personas. Trata de organizar y clasificar los documentos, expedientes, archivos e imágenes digitales en una localización centralizada a la que los miembros de la organización puedan acceder de forma fácil y sencilla.

Este tipo de sistemas está orientado al almacenamiento y gestión de documentos digitales que proporcionan servicios de: control del flujo de documentos, almacenamiento, control de versiones, validación de datos, asignación de tipologías (para poder gestionar su ciclo de vida particular) y metadatos (fichas descriptivas), agrupación en expedientes o estructuras jerárquicas, asignación de niveles de seguridad (limitar qué puede hacer cada usuario o grupo de usuarios), distribución, búsqueda y recuperación documentos, y definición de un ciclo de vida para que, por ejemplo, los

documentos se borren automáticamente al cabo de los años. Soporta tanto los formatos de documentos más utilizados como XML, .doc, .odf o .pdf, como formatos más técnicos y específicos de determinados sectores. Una de las funciones importantes es la del reconocimiento óptico de caracteres (OCR) para la digitalización de documentos en papel y la identificación del contenido, adaptándolo a la estructura de organización y almacenamiento.

3.1 Características

Las principales características de los SGD existentes en el mercado son:

1. Localización central. El amplio número de canales a través de los que la información llega a las organizaciones puede ocasionar que grandes volúmenes de información queden desestructurados y repartidos entre distintos sistemas. Un sistema de gestión documental almacena y organiza toda la información proveniente del trabajo diario de la organización en una ubicación central (repositorio). La organización determinará qué empleados pueden tener acceso a los documentos alojados en dicha localización gracias al mecanismo de concesión de permisos. Esta centralización de la información supone terminar con la búsqueda de documentos por las redes de carpetas de la organización agilizando de esta forma el trabajo.
2. Organización y gestión. Un SGD, normalmente con un sistema de almacenamiento compartido, permite clasificar y organizar los documentos en carpetas y subcarpetas, identificadas únicamente por el nombre. Los documentos quedan agrupados en una jerarquía de contenidos (carpetas o expedientes), con metadatos asociados que permiten la estructuración de los contenidos y la realización de búsquedas más específicas dentro de la estructura, y facilitan la aplicación de ciclos de vida a los documentos contenidos. Cuando almacenamos un documento en un contenedor asignamos un tipo de documento. Cada tipo documental se asocia a metadatos (características de los documentos que pueden ser obligatorias, opcionales, con valores controlados por listas o tesauros, etc.) que identifican al documento y permiten la recuperación de acuerdo a diversos criterios. Los tipos de documento y sus metadatos característicos facilitarán la búsqueda de todos los documentos de un tipo dado, posibilitarán la determinación de tratamientos específicos sobre los tipos documentales (copias automáticas, avisos a usuarios, borrado en una determinada fecha, etc.).
3. Búsqueda por contenido. Sin un motor interno de búsqueda el repositorio puede ser un lugar perfecto para esconder documentos. El motor de búsqueda permite facilitar la localización de documentos en el repositorio utilizando diferentes técnicas desde las más tradicionales, a la búsqueda en lenguaje natural o utilizando técnicas de inteligencia artificial. Generalmente extraen el contenido de los documentos independientemente del formato de archivo y lo indexa para poder buscar por combinaciones de palabras del contenido de los documentos, además de los metadatos. Este análisis e indexación además puede realizarse en diferido (programado) para no interferir en el trabajo habitual, aplicando técnicas avanzadas de recuperación de información para implementar la búsqueda por palabras y sus derivados.
4. Mejorar el flujo de trabajo. En un SGD es posible definir flujos de trabajo, enrutamiento de tareas, gestión de formularios, notificaciones, aprobaciones, etc. y posibilita convertir estos flujos en procesos eficientes y productivos. Mediante la automatización de funciones y tareas, el sistema proporcionará una imagen global de los procesos de trabajo dentro de la organización. El control de procesos permitirá el seguimiento de tareas completas o incompletas, o la automatización de tareas repetitivas. Puesto que uno de los objetivos es la automatización de las tareas, la mayoría cuentan con algún tipo de herramienta de programación de tareas, motor de flujo de trabajo o BPM que permite la automatización de la gestión de documentos o expedientes (por ejemplo, notificación de fechas de cumplimiento, eliminación de expedientes obsoletos o realización de copias).
5. Colaboración documental. Múltiples personas pueden trabajar en un documento al mismo tiempo, mientras éste está alojado en una localización central. Con ello, los empleados podrán acceder en todo momento a una visión general de un documento, pudiendo modificarlo si se considera necesario. La colaboración documental permite, de esta forma, compartir ideas e información entre los empleados de una organización de forma más sencilla.
6. Control de versiones. Como un objetivo de los SGD es el trabajo en equipo, permiten la realización de modificaciones de los documentos, pero controlando y conservando las versiones realizadas. Un documento realizado por un grupo de trabajo puede ser modificado infinidad de veces, y los gestores documentales tratan de controlar las diferentes versiones a través de un historial de versiones, con la posibilidad de acceder a cualquier versión para comparar versiones, recuperar información, eliminarla o añadirla. La mayoría de los gestores documentales permiten la monitorización de cada actividad o los cambios realizados en cualquier documento o qué usuario accedió a qué documento a través de las herramientas de control de versiones. Un

documento se podrá elaborar con diferentes versiones de este, realizadas por uno o varios usuarios consecutivamente. Las opciones de versionado son variables, tanto en la forma de bloquear los documentos como en la gestión de versiones y la visibilidad (en algunos casos se puede ver la versión en elaboración y en otros solo la publicada). El control de versiones realiza un seguimiento de los cambios mediante la creación de nuevos archivos con cambios en el transcurso del tiempo para que pueda ver cualquier versión editada que desee.

7. Compartir documentos. Los sistemas de gestión documental facilitan la tarea de compartición de documentos. A través de la creación de grupos o con accesos ilimitados a la localización central, los documentos pueden ser distribuidos tanto de forma interna como externa. Los gestores documentales permiten a diferentes grupos externos a la organización el acceso a documentos necesarios.
8. Seguridad de la información y recuperación. Aún son muchas las organizaciones que mantienen sus documentos almacenados en archivos físicos y la posibilidad de que un documento termine perdido o deteriorado es alta. Estableciendo copias digitales de documentos en papel, el riesgo de pérdida disminuye de forma considerable. De esta forma, la organización trabaja con la seguridad de que toda su información está ubicada en un mismo sistema, sistema que puede incluir entre otras características, la recuperación de datos en caso de desastre. La mayoría de los SGD ofrecen opciones de copia de seguridad. Si todos los datos desaparecen por cualquier mal funcionamiento, siempre tendremos la posibilidad de restaurarlos desde las copias de seguridad. Los documentos electrónicos se pueden restaurar en caso de un mal funcionamiento en el sistema de gestión documental porque la mayoría de ellos permiten crear un archivo de respaldo para sus datos.

3.2 El software de gestión de documentos

El software de los SGD está disponible en varias formas, desde la instalación en servidores propios o de terceros hasta las versiones de software en la nube. Hay software gratuito o de pago, paquetes de software de fuentes abiertas o privativas, o softwares especiales para sectores concretos. No todos los paquetes son adecuados para ser utilizados por todos los tamaños de organizaciones y esta adecuación se refleja en la diversidad de la oferta de software de gestión documental. Algunos sistemas están especialmente diseñados para pequeñas organizaciones, mientras que otros son interesantes para medianas o para grandes. Los servicios de consultoría y desarrollo permiten adaptar la funcionalidad del sistema gestor a los procedimientos y especificidades de la organización. Donde antes se implantaban sistemas complejos de empresas de cierto renombre, ahora también existen softwares sencillos, pequeños y flexibles de marcas no tan conocidas. La mayoría de los proveedores están dirigidos a las pymes en España, y pueden ofrecer una solución para la gestión de documentos.

3.2.1 Software código abierto

En el mercado hay variantes ‘freeware’ (distribución gratuita). Estos sistemas pueden descargarse, usarse y distribuirse sin costes. Estas aplicaciones no permiten hacer cambios en el código fuente. Estas modificaciones si se pueden realizar en los softwares de código abierto. Estas adaptaciones y modificaciones en el código fuente consiguen que un sistema pueda llegar a cumplir los requisitos específicos de una organización. Esta posibilidad implica que el personal encargado de su realización tenga un gran dominio del lenguaje de programación y del propio SGD a adaptar. Es fundamental trabajar con proveedores que ofrezcan garantías en cuanto al mantenimiento y las actualizaciones del software. También es importante que el sistema elegido cuente con una amplia comunidad de desarrolladores que favorezca el futuro del programa. Este tipo de software posee las siguientes ventajas:

1. Flexibilidad de la solución: el acceso al código permite a las organizaciones el control del producto y su modificación para adaptarlo a las necesidades de la organización.
2. Conocimiento del sistema: con el acceso al código fuente va implícito un entendimiento superior del funcionamiento del programa, por la organización que lo adquiere o por el proveedor que se contrate para la implantación.
3. Actualizaciones: las comunidades de desarrolladores que trabajan detrás de estos productos de código abierto mantienen el software en constante estado de evolución y mejora.

Por el contrario, presenta las siguientes desventajas:

1. La ayuda técnica es, normalmente, proporcionada por la comunidad. Existen escenarios en los que una organización experimenta un problema en su gestor documental pero no existe un equipo de técnicos capaces de encontrar soluciones. La organización tiene que recurrir a la comunidad de desarrolladores detrás de la herramienta para recibir ayuda.

2. Carencia de habilidades para la administración: aunque el conocimiento del sistema puede ser una ventaja, la carencia de las habilidades necesarias para la administración del sistema se convierte en una desventaja. En muchos casos, la contratación de expertos que formen a los administradores de sistemas de la organización es también la mejor opción.

3.2.2 Software de gestión documental propietario

Un gran número de organizaciones optan por hacer uso de un SGD propietario. A las organizaciones con necesidades específicas se les ofrece sistemas mucho más a medida, y ven en este tipo de software una forma de satisfacer sus necesidades. Las funcionalidades adicionales hacen que los proveedores de software y hardware se decanten por sistemas propietarios. El software se puede comprar e implementar en el servidor de la organización, pero también en la nube del proveedor. El proveedor se encarga de la gestión, el mantenimiento y las actualizaciones del software. La mayoría de los pagos son por licencia y funcionalidades. Presenta las siguientes ventajas:

1. Soporte técnico. El proveedor proporciona la posibilidad de contratar el servicio técnico del sistema para solucionar cualquier inconveniente que la organización sufra con su gestor documental.
2. Documentación y Formación. Disponibilidad de documentación y diversidad de servicios de formación (cursos on-line, presenciales, etc.). Las organizaciones necesitan que no sólo los administradores de sistemas puedan entender los programas con los que trabajan, sino todos los miembros que lo necesiten.

En el sentido contrario, presenta las siguientes desventajas:

1. El coste de este tipo de SGD suele ser más alto que las soluciones basadas en software ‘open source’.
2. Dependencia de la organización respecto al proveedor. Las actualizaciones, nuevas funcionalidades o las soluciones a errores son lanzadas en el momento en el que el productor estima necesario.
3. Las organizaciones adquieren una licencia para utilizar un SGD, pero no pueden ver o transformar el código. Tienen restringidos el uso, modificación y copia del software por motivos técnicos o legales, lo que limita las posibilidades de personalización y adaptación.

3.3 El dilema entre software de código abierto o el software propietario

El código abierto implica ciertas ventajas, tales como que permite trabajar con tecnología que es mejorada constantemente por la comunidad de desarrolladores, que la utilizan por su fácil uso y las posibilidades de personalización y adaptación, tanto desde la perspectiva del usuario final como del personal técnico. Por otro lado, el software propietario ofrece servicios profesionales de implantación, soporte y actualización de versiones. Tanto en el caso de proyectos basados en código abierto como soluciones privativas, se deben tener en cuenta, a efectos de un cribado inicial, los siguientes criterios:

- a) Consolidación: hay que valorar si el sistema evaluado es un proyecto nuevo aún no consolidado, un proyecto consolidado con una larga trayectoria, o un proyecto que en los últimos años ha tenido pocas actualizaciones o ninguna.
- b) Es conveniente ponderar si el sistema evaluado tiene una comunidad de usuarios y desarrolladores relevante y activa. La observación del movimiento en foros públicos y los tiempos de respuesta de la comunidad puede ser reveladora.
- c) Hay que estudiar si se ofrecen servicios profesionales relacionados con la implantación, soporte y mantenimiento.
- d) Un dato importante a tener en cuenta es la periodicidad de las actualizaciones de la aplicación.
- e) Es conveniente valorar el tiempo que se tarda en proporcionar un parche que solucione alguna vulnerabilidad grave del sistema. También es interesante averiguar si existe algún formulario o correo electrónico de contacto, cuánto tiempo tardan en la respuesta. Estos aspectos dan una idea de la calidad del servicio.
- f) También es conveniente localizar la documentación de la aplicación y echarle un vistazo rápido con el fin de vislumbrar si va a ser suficiente en caso de tener que consultarla.

3.4 El formato del sistema de gestión documental

Alfresco se puede instalar en un servidor propio, alojarse en un servidor propio o en una nube (en forma de servicio SaaS).

Ventajas y desventajas del SGD instalado en servidor:

- **Control:** decidimos la infraestructura y el entorno de implantación del SGD.

- **Personalización:** si se utiliza un SGD ‘open source’ tendremos más posibilidades de adaptar la aplicación.
- **Escalabilidad:** el SGD viene con un conjunto de funcionalidades que se pueden incrementar en función de las actualizaciones de nuestro proveedor. Por otro lado, como administradores de la plataforma podemos habilitar o deshabilitar funcionalidades para adecuar el SGD a cada escenario de uso dentro de la organización.
- **Inversión inicial:** elevada, sobre todo si la comparamos con el uso de SGD tipo SaaS. A cambio, nos aporta independencia y autonomía.
- **Periodo de implantación:** será más prolongado que en el caso del SGD como servicio. Deberemos tener en cuenta la instalación y el proceso de configuración.
- **Costes de implantación, mantenimiento y actualización:** a no ser que el proveedor incluya el proceso de implantación, éste será un coste adicional, además del mantenimiento del servidor, y las actualizaciones del software.

Ventajas y desventajas del SGD alojado en servidores de terceros:

- **Proceso de implantación:** más rápido que en la versión ‘on premise’ porque el software estará instalado y configurado, al menos a nivel standard.
- **Escalabilidad:** asistida por el proveedor de la solución SGD.
- **Inversión inicial:** menor que en el caso de instalación en servidor propio y siempre puedes alquilar el espacio no utilizado a terceros.
- **Vulnerabilidades:** la seguridad y escalabilidad dependen del proveedor del software, o del servidor en el que está alojado el SGD.

Ventajas y desventajas del SGD en formato SaaS:

- **Sin instalación:** el SGD ya está instalado y listo para usar. Estos servicios son gestionados también por el proveedor.
- **Pago por servicio:** las plataformas SaaS no suelen tener compromisos de permanencia. Se paga por uso, lo que permite cambiar de modelo de gestión en el momento que se deseé. Sin costes de mantenimiento: todos los gastos de mantenimiento o seguridad son a cargo del proveedor.
- **Escalabilidad:** a demanda de la organización, que es el cliente del proveedor.
- **Variedad de funcionalidades:** suelen ofrecer opciones de personalización del SGD y adaptación de la gestión a la medida de la organización.
- **Personalización:** usar el software como servicio, en ocasiones implica que no se pueden hacer cambios drásticos en la configuración.
- **Actualizaciones regulares:** los proveedores se encargan de los fallos de seguridad y de las nuevas funcionalidades para mantener los niveles de satisfacción y fidelidad del cliente.
- **Seguridad:** es gestionada por el propietario del software en modo SaaS, desde las copias de seguridad a la resolución de problemas de seguridad o desempeño. Además, el cliente puede establecer procesos adicionales que incrementen la seguridad.
- **Pérdida de control:** el uso de la aplicación como servicio, supone dependencia por parte del proveedor en materia de seguridad. Si gestionamos información sensible, puede suponer una pérdida de control sobre la seguridad.
- **Nivel de soporte:** la mayoría de los softwares en modo SaaS, ofrecen soporte técnico incluido con gestión de copias de seguridad.

4 APlicativos Software de Gestión Documental

Los desarrollos que se le hacen a un software de gestión documental se realizan de forma rápida y consecutiva. Es importante que una empresa esté bien informada antes de que inicie el proceso de selección de software de gestión documental. En España, hay muchos proveedores diferentes con varios tipos de software de gestión documental. Mientras que un paquete tiene como punto fuerte el archivo automático, otros paquetes ofrecen una calidad muy alta de colaboración entre compañeros. *tic.PORTAL*, plataforma perteneciente al Centro Europeo del Conocimiento EKCIT, investiga todos los tipos de software de gestión de documentos en España y comparte sus conocimientos con el lector de forma gratuita. En la *Guía de Gestión Documental 2022* se encuentra información sobre los proveedores, los paquetes de gestión documental y los posibles precios.

En nuestro estudio hemos seleccionado tres de los SGD que consideramos de interés para su implementación en organizaciones grandes, medianas y pequeñas, y que cuentan con distribuciones en código abierto: Alfresco, OpenKM y Openprodoc.

4.1 Alfresco

Sistema creado en junio de 2005 por John Newton, cofundador de Documentum y John Powell, antiguo ejecutivo de Business Objects, quienes trajeron consigo una gran experiencia respecto a las necesidades de una aplicación orientada a la gestión de contenido empresarial. Es un sistema de gestión de documentos de escala organizacional desarrollado en java, basado en estándares abiertos y para sistemas operativos tipo windows y linux. Reúne un gran número de elementos técnicos muy reconocidos (Hibernate, Lucene, etc.) procedentes del sector de código abierto. Proporciona un juego de interfaces de programación que soportan múltiples protocolos con los que los desarrolladores pueden crear a medida y otras. Con Alfresco todos los documentos son objetos a los que se puede aplicar una serie de métodos como la previsualización, la gestión de versiones, la gestión de permisos, etc.

Alfresco es, en muchos aspectos, bastante similar a otros SGD. Está formado por un repositorio y un servidor de aplicaciones que permite administrar contenidos, metadatos, asociaciones e indexación y búsquedas con el motor Solr-Lucene. Incluye también herramientas para la gestión de un portal web para el acceso al contenido, y gestión de flujos de trabajo con el estándar JBPM. Proporciona soluciones estándar como gestión de documentos, la gestión de registros y la gestión de contenidos para la web. Entre las funcionalidades principales, cabe destacar la gestión de documentos y los metadatos asociados (documentos ofimáticos, presentaciones, imágenes, XML, multimedia, etc.), gestión del ciclo de vida de los contenidos (crear, compartir, versionar, aprobar, publicar, archivar, retirar), trabajo colaborativo (foros, notificaciones, RSS, blogs, wikis, etc.) y gestión avanzada de flujos de trabajo, notificaciones y circuitos de validación.

4.1.1 Objetivos de Alfresco

- Gestiona cualquier tipo de información corporativa, independientemente del soporte (electrónico o en papel), del formato o de la aplicación que se usó para crearla.
 - Por ejemplo, permite a los usuarios guardar y gestionar ficheros multimedia (fotografías, vídeos, ficheros de audio) en una base de datos de contenido común y proporciona vistas dedicadas para manejar los diversos formatos de ficheros de medios enriquecidos.
- Ofrece a los usuarios las herramientas que necesitan para colaborar de forma segura con compañeros a nivel interno, y con socios de negocio (clientes, clientes potenciales, consultores y agencias) a nivel externo.
 - Con máximo control y herramientas de colaboración como wikis, foros, enlaces, blogs y listas.
- Ayuda a publicar documentos o contenidos. En concreto, la publicación social permite que la publicación automática del contenido en sitios de medios sociales (como twitter, facebook o youtube) para compartirlo externamente.
- Ofrece las funcionalidades de un programa de gestión de documentos ya que permite almacenar, recuperar y compartir documentos y contenidos.
- Edita directamente documentos almacenados en la plataforma con Google Docs o Microsoft Office.
- Automatiza flujos de trabajo y gestiona los procesos empresariales mediante un motor de procesos y flujos de trabajo integrados. También permite la creación de reglas de fácil configuración.
 - Por ejemplo, iniciar controlar los procesos de forma automática. El flujo de trabajo híbrido, a su vez, permite la sincronización de procesos en la nube.
- Ofrece un control de acceso total para garantizar la seguridad y la integridad de los documentos.
- Incluye un potente motor de búsqueda que se puede basar tanto en los contenidos como en los metadatos.
- Controla versiones y cambios de contenido.
- Extrae e indexa automáticamente metadatos.
- Permite la gestión archivística de los documentos mediante la creación e implantación de cuadros de clasificación, cuadros de acceso, calendarios de transferencia, de conservación y de destrucción de documentos.
- Gestiona evidencias y elabora informes de auditoría.
- Permite el acceso y la sincronización desde distintos dispositivos.

4.1.2 Distribuciones

Existen tres distribuciones de Alfresco:

- Community Edition: licencia LGPL ('Library general public license') que la define como software libre, de código y estándares abiertos.
- Enterprise Edition: distribución bajo licencia de código y estándares abiertos, con la posibilidad de soporte comercial.
- Cloud Edition (nube): versión SaaS ("software como servicio").

La versión empresarial-comercial ofrece funciones especiales como la gestión de grupos, LDAP, SSO, y niveles de soporte especializado. La versión Community es la alternativa de código abierto para la gestión de contenido empresarial. Se distribuye con la licencia LGPL v3, pero no cuenta con soporte técnico por lo que es más adecuado para desarrolladores y para entornos empresariales no críticos. Permite a las organizaciones gestionar cualquier tipo de contenido, desde documentos administrativos sencillos a imágenes escaneadas, fotografías, planos o grandes ficheros de vídeo, a través de sistemas de gestión de documentos y repositorios de cumplimiento CMIS. Se puede descargar gratuitamente, pero se debe tener en cuenta que no incluye:

- Soporte, correcciones de fallos, indemnización o garantía.
- Soporte para bases de datos propietarias e integraciones con software propietario.
- Consola de administración para configurar y reiniciar los subsistemas de Alfresco.
- Capacidad de definir políticas de almacenamiento de contenido para satisfacer los requisitos del ciclo de vida del contenido.
- Posibilidad de sincronización de contenido y procesos entre Alfresco a nivel local y Alfresco en la nube.
- Cifrado de contenido.
- Escalabilidad de la aplicación a través de un clúster.
- Certificación de clústeres / alta disponibilidad.

4.1.3 Estructura y funcionalidades de Alfresco

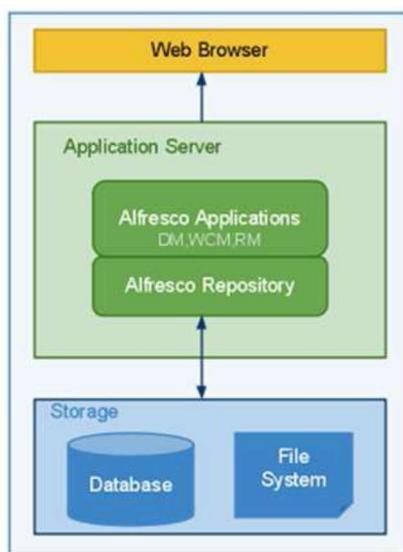


Figura 2: Arquitectura de Alfresco. Fuente: Alfresco, Inc.

Su arquitectura es coincidente con la arquitectura de una aplicación web basada en java: un diagrama típico de una arquitectura web. Un servidor de aplicaciones que alberga la lógica tanto la interfaz de usuario como del modelo de contenido. Los servidores de aplicación incluyen tanto las aplicaciones de gestión de documentos (DM), gestión de contenidos web (WCM) y de gestión de documentos de archivo (RM) como el repositorio. Alfresco separa a nivel lógico el repositorio (lugar donde se almacenan los contenidos) de las aplicaciones que acceden a él. El almacenamiento de datos y contenido es proporcionado por soluciones 'back-end' como bases de datos y sistemas de archivos. Los navegadores web pueden conectarse al servidor de aplicación sin necesidad de instalar ningún software específico en el cliente. Los usuarios acceden a la interfaz de usuario de Alfresco (gestión documental) o Share (entorno colaborativo).

Estas interfaces pueden variar en función de los componentes instalados ('Web Content Manager', 'Records Management', etc.) para adecuarse a la lógica de estos.

El repositorio ofrece un conjunto de servicios de gestión reutilizables y transversales que pueden ser utilizados por múltiples aplicaciones: almacenamiento de contenido, consulta, control de versiones y transformación de contenidos. Todos los documentos subidos al repositorio de Alfresco con toda su información adicional (metadatos) persisten (almacenan) en la capa de almacenamiento según su naturaleza: almacena los metadatos en una base de datos y el contenido en un sistema de archivos. Con la base de datos se obtienen los beneficios del soporte transaccional y escalabilidad de los sistemas de gestión de bases de datos. El contenido se almacena en el sistema de archivos para permitir contenidos de tamaño grande, acceso aleatorio, streaming y opciones para diferentes dispositivos de almacenamiento.

A nivel lógico sería posible no emplear de la interfaz de Alfresco y programar una interfaz propia o acceder al repositorio a través de una tercera aplicación. Esto es posible gracias a los '*Foundation Services*' de Alfresco, los cuales permiten acceder al contenido del repositorio y realizar acciones sobre él. Se dispone entre otros de servicios de autenticación, búsquedas, permisos, contenido, etc. El listado completo de los servicios lo podemos obtener dentro del fichero public-services-context.xml

4.1.4 Entornos de trabajo

Alfresco presenta diferentes entornos especializados: Share y Records Management. El primero permite establecer espacios colaborativos y todo aquello relacionado con la gestión de documentos y la gestión de contenidos. Es un entorno más enfocado a los encargados de la gestión de la información. El segundo entorno, Alfresco Records Management, facilita a los usuarios finales trabajar con ficheros y declararlos documentos de archivo en el momento oportuno, sin cambiar su forma de trabajar. Permite un archivado sencillo de *registros*. Los usuarios pueden cargar documentos de archivo desde su navegador web, correo electrónico o desde su ordenador de escritorio.

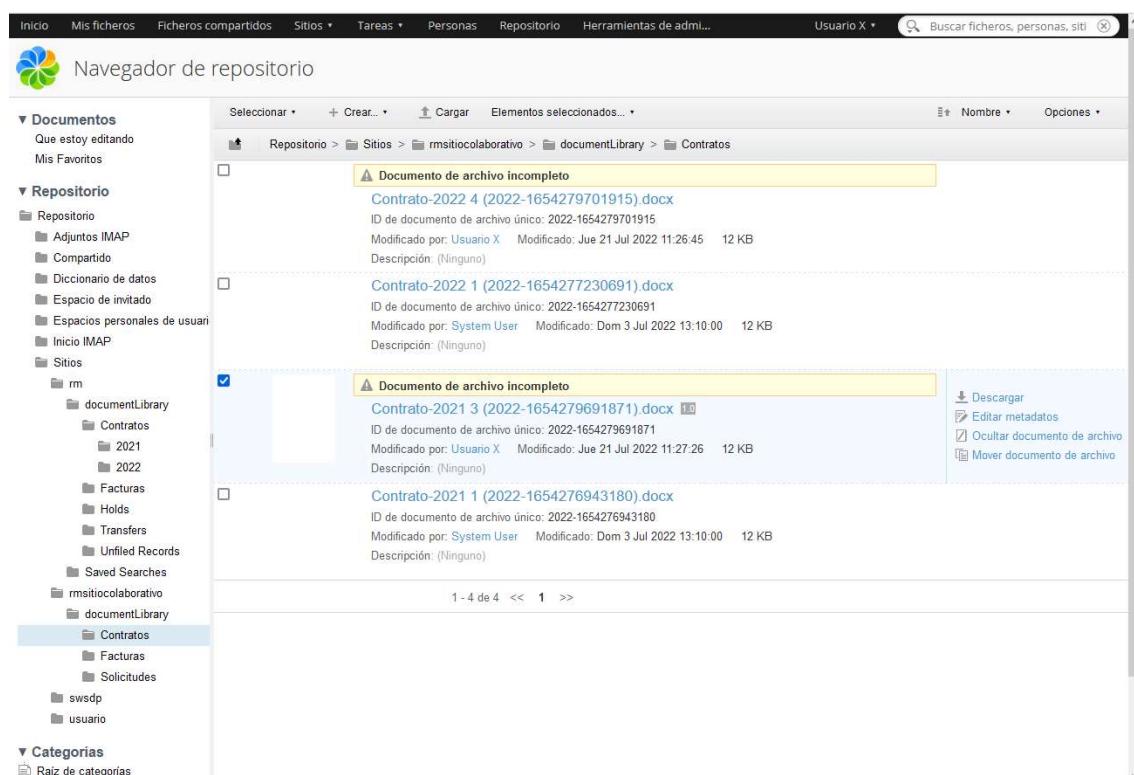


Figura 3: Visualización del repositorio en Alfresco. Fuente: elaboración propia.

- **Repositorio.** Alfresco como núcleo es un repositorio que permite almacenamiento de contenido, y una amplia gama de servicios que pueden ser utilizados por aplicaciones de gestión. El repositorio proporciona almacenamiento para documentos y otros contenidos. Como ya hemos comentado, los metadatos de contenido se almacenan en una base de datos relacional, mientras que el contenido en sí se almacena directamente en el sistema de archivos. Las relaciones entre los elementos de contenido, y sus diversas propiedades (metadatos) se definen en uno o más modelos de contenido. Los modelos de contenido pueden ser considerados como la descripción de los tipos de contenido. El repositorio se construye a partir de los conceptos básicos de sitio (espacio equivalente a una carpeta o directorio, pero con reglas de contenido y de seguridad) y contenido (cualquier archivo almacenado). El repositorio se estructura en forma de árbol (jerarquía) a partir del espacio raíz, del que cuelgan todos los espacios y contenidos que se crean. Las tres acciones típicas que realizaremos serán las de subir, bajar y buscar documentos. Estas acciones se realizan mediante tres servicios que se apoyan en elementos ya probados dentro del mundo open source como Hibernate o Solr/Lucene.
- **NodeService.** Todos los elementos son tratados como nodos, pero con diferentes propiedades. El servicio encargado de trabajar con los nodos será el ‘NodeService’ y toda la información sobre los nodos, se almacenará en base de datos. Por ejemplo, un archivo es un nodo con ciertas propiedades o metadatos como título, autor, fecha de creación e incluso el propio contenido no es más que un metadato. A su vez un sitio es también un nodo que tiene una asociación del tipo “contiene a” con otros nodos.
- **ContentService.** Este servicio se encarga de leer o escribir el contenido en el repositorio, así como de transformarlo de un tipo MIME a otro. El contenido suele estar referido a los archivos binarios que subimos o creamos dentro de Alfresco, es decir un archivo PowerPoint, Word o PDF o cualquier tipo MIME. Estos archivos binarios se almacenarán en el sistema de archivos del servidor (‘ContentStore’) siguiendo una estructura de carpetas ordenada por fecha.
- **SearchService.** Cada vez que se sube contenido al sistema, este es indexado de forma automática (tanto sus metadatos como el contenido). Este servicio utiliza Solr/Lucene para la indexación y consultas. Los índices generados se almacenan en el sistema de archivos.

Además de la funcionalidad básica de almacenamiento de contenido, el servidor Alfresco ofrece otros servicios relacionados como el servicio de carpetas de documentos y el servicio de control de versiones. Hay servicios de más alto nivel, como el servicio para la creación de imágenes, el servicio de sitios para crear y administrar sitios en la aplicación ‘Share’, el servicio de etiquetado que proporciona la posibilidad de etiquetar contenido con palabras clave:

- **Sitios de trabajo.** Es la unidad fundamental de organización de Alfresco Share y se podría definir como un lugar donde poder incorporar contenido, documentos, reglas, flujos de trabajo, usuarios y, en general colaborar con otros con el fin de desarrollar un proyecto o una función concreta. Alfresco funciona mediante sitios de trabajo superpuestos, de forma que podemos incluir un sitio de trabajo en un directorio que puede contener un determinado número de documentos y otros sub-sitios. Contiene archivos y sub-espacios y se gestiona a través de aspectos/propiedades: metadatos, permisos y reglas. Hablamos de sitios inteligentes en la medida en que podemos asociar a cada sitio un determinado número de reglas o tratamientos sobre los documentos que entran y salen del espacio. Por ejemplo, podemos convertir automáticamente y de manera transparente para el usuario un documento a PDF, en el mismo sitio o en un sitio específico.
- En Alfresco es posible definir **modelos de sitios** que permitirán, cuando se cree un nuevo espacio, recrear de forma automática un diagrama arbóreo concreto. Se puede usar una plantilla de sitio creada o que nos han proporcionado para optimizar la creación de un nuevo sitio y además asegurar la consistencia de los sitios de esta forma.
- **Aspectos y propiedades.** Los aspectos son características de los documentos (es un concepto más amplio que el de metadatos). Asignar aspectos a un documento es, por ejemplo, decir que el documento es versionable (introducirá el control de versiones) o clasificable. Asignando aspectos también podemos hacer que un documento quede sujeto a un esquema de metadatos estándar o a uno propio de nuestra organización. En Alfresco, las propiedades son los metadatos: es el conjunto de informaciones utilizadas para caracterizar un documento. Los conjuntos de metadatos se crean a través de archivos XML, denominados aspectos. Los metadatos están caracterizados por defecto (fecha, entero, cadena, usuarios, etc), pueden estar sujetos a obligatoriedades, intervalos o listas de valores, expresiones regulares, etc. Se puede consultar una ficha detallada de un documento que reúne el conjunto de informaciones propias del documento y de sus versiones. Cada aspecto puede ser aplicado posteriormente a un documento. También se pueden utilizar ‘dashlets, pequeños programas o aplicaciones que podemos usar a voluntad para, por ejemplo, añadir un calendario, un foro o una wiki.

- **Contenido y formatos de archivo.** Como comentamos anteriormente, el contenido comprende la información dentro de una carpeta. Puede tratarse de información estructurada o no estructurada y permite múltiples formatos. Es capaz de gestionar todo tipo de documentos, desde archivos de texto simples a archivos binarios pasando por documentos ofimáticos. El sistema gestor permite crear sobre la marcha documentos de texto y HTML directamente utilizando el editor TinyMCE que viene integrado. Con ciertos tipos de documentos puede analizar el archivo para recuperar metadatos (título, descripción), etc. que serán archivados en los metadatos del documento Alfresco. Se realiza esta extracción automática de metadatos con archivos word, excel, powerpoint, pdf, etc. Desde un punto de vista técnico, la aplicación utiliza herramientas como Open Office que permiten leer, escribir y convertir los archivos en formatos diferentes. Podremos convertir automáticamente, por ejemplo, cualquier documento office en PDF. Existe la opción de crear enlaces a contenidos. Son un tipo especial de fichero que referencia a otro fichero. Similar a los enlaces simbólicos de Unix. Solo contiene una ruta simbólica al contenido almacenado en algún otro lugar del repositorio de Alfresco.
- **Categorías.** Sirven para realizar una clasificación de los elementos (sitios o contenidos), aparte de la propia clasificación que se establece por la organización del repositorio como el árbol de sitios. Esto permitirá hacer búsquedas más restrictivas y con mejores resultados. Los elementos, se pueden asignar a una o varias categorías. Las categorías se definen según las necesidades del sistema: un usuario o grupo de usuarios puede solicitar al administrador la creación de una categoría para clasificar sus documentos. Se puede decir que Alfresco entiende como categoría la clasificación archivística, aunque emplea la misma palabra en dos contextos distintos: (1) categoría en el módulo de gestión de archivo (RM) para una entrada del cuadro de clasificación (fondo, sección, grupo de series, serie o subserie) y (2) categoría fuera del módulo de RM (aquella que no tiene ninguna consecuencia archivística más allá de establecer un criterio organizativo).
- **Reglas de contenidos.** Se trata de los diferentes tipos de tratamiento que se pueden aplicar al conjunto de documentos que salen o entran de un espacio/sitio. Durante la creación de una regla de contenido, debemos definir el tipo de acción a ejecutar y las condiciones de activación de la regla. Todos los procesos de producción y de validación se engloban en estas reglas de contenido. Especialmente permiten:
 - Convertir el documento sobre la marcha en un formato alternativo: word en pdf, por ejemplo.
 - Activar un almacenamiento del contenido.
 - Extraer los campos de metadatos clásicos del elemento.
 - Añadir/eliminar un aspecto del elemento.
 - Asociar el elemento a una categoría.
 - Convertir, desplazar o copiar el documento/elemento/imagen a otro espacio, proyecto web o destino.
 - Añadir un flujo de trabajo básico al elemento.
 - Enviar una alerta a uno o varios usuarios.
 - Ejecutar un script.
- **Flujos de trabajo.** Alfresco integra el motor JBPM creado por jBoss que permite aplicar fácilmente flujos de trabajo profesionales avanzados. JBPM puede utilizarse para realizar circuitos de validación complejos. Los flujos de trabajo permiten definir tareas en serie y en paralelo. Además, se pueden establecer "alarmas" para gestionar la finalización de determinadas tareas.
- **Estándares y protocolos de comunicación.** Alfresco se basa en un determinado número de elementos de código abierto reconocidos y sólidos. Está diseñado para respetar los estándares establecidos por el W3C. Entre ellos destacan LDAP, SOAP, JSR 168, WebDAV/CIFS y CMIS. El repositorio de Alfresco soporta por tanto una variedad de protocolos de comunicación (CIFS, WebDAV, FTP, etc.) que implementan el paradigma de carpetas y archivos, y que se integran perfectamente con las carpetas y los nodos de archivos del repositorio.
 - **CIFS** transforma el repositorio de Alfresco en un sistema de archivos estándar. Así, cualquier herramienta que está diseñada para leer y escribir en un sistema de archivos también sabe leer y escribir directamente del repositorio de Alfresco. Se recomienda la utilización del protocolo CIFS que resulta más sólido. Simplifica enormemente el uso diario de su solución de gestión documental.
 - **WebDAV** es un estándar que describe cómo realizar acciones de gestión de archivos (tales como escribir, copiar, eliminar o modificar) a través del protocolo HTTP. Cualquier aplicación WebDAV realiza estas tareas con documentos en Alfresco. Hay que tener en cuenta que los archivos guardados mediante WebDAV no tienen ni descripción ni autor, ni se les aplicarían las reglas si las hubiera.
 - **'Plugins' para Microsoft Office:** permite interactuar integralmente con la herramienta de gestión documental, ejecutando las principales funcionalidades (creación, edición, revisión, publicación y

- búsqueda) directamente desde Office sin necesidad de instalar nada en el sistema del cliente. Alfresco es compatible con SharePoint de Microsoft, por lo que puede emular a un servidor SharePoint.
- ‘**Plugins**’ para OpenOffice: con las mismas funcionalidades que el ‘plugin’ para Office, permitiendo a los usuarios modificar documentos de Alfresco directamente desde la interfaz de usuario de OpenOffice y realizar las operaciones típicas (leer, editar, crear y guardar), el control de versiones y la visualización de los metadatos.
 - **Integración.** Cada parte del repositorio de Alfresco es un componente o un servicio. Un componente proporciona una funcionalidad o capacidad específica. Un servicio es un punto de entrada de interfaz para un cliente para unirse y utilizarlo. Este enfoque permite que componentes existentes puedan adaptarse a nuevas implementaciones, añadir nuevos componentes con facilidad y que los clientes puedan conectarse y utilizar los servicios sin conocer como han sido implementados. El servidor de Alfresco es extensible: estos servicios se implementan en java, y exponen una API. Se pueden desarrollar extensiones en java y javaScript, y aplicaciones cliente en cualquier idioma que usen los servicios REST. Esta extensibilidad y capacidad de configuración se logra mediante el framework Spring, un marco de código abierto estándar de facto. Alfresco ha hecho de esta utilidad la base fundamental de su arquitectura. Los componentes de Alfresco se configuran y se integran mediante los mecanismos de configuración de Spring de forma declarativa. Las características de programación orientada a aspectos de Spring permiten construir Transacciones y Seguridad en los diferentes componentes sin duplicar funcionalidad. Si hay una característica de Alfresco que no resulta necesaria, se puede desactivar a través de configuración, proporcionando una aplicación más ligera y posiblemente más rápida. Si hay una característica que se decide incorporar, se puede intercambiar fácilmente a través de la configuración de Spring. Dado que Alfresco usa el marco de desarrollo de ‘Spring’, cada uno de estos servicios y componentes está configurado mediante ficheros XML.

4.2 OpenKM

Se trata de un SGD escrito en java por la empresa Open Document Management System S.L. con sede en Palma de Mallorca (<https://www.openkm.com/es/sobre-nosotros.html>). Fue un proyecto respaldado por el gobierno español como parte del programa “Project Profit” y su primera versión fue lanzada a finales del 2006. La traducción de su herramienta a 35 lenguajes distintos ha permitido que pueda ser utilizado a nivel mundial. Es una aplicación web de gestión documental que utiliza estándares y tecnologías de código abierto y compatible con la mayoría de los navegadores como Explorer, Firefox, Safari o Chrome. Permite gestionar documentos, documentos de archivo, correos electrónicos, imágenes y ofrecer un entorno de desarrollo colaborativo de contenidos. Integra todo lo esencial para gestión de documentos, colaboración y funcionalidades de búsqueda avanzada, en una solución fácil de usar. Provee de una interfaz web para la gestión de archivos que incluye un repositorio de contenidos. Hace uso del motor de búsqueda Solr/Lucene como medio para lograr la indexación de archivos y de jBPM como motor de flujo de trabajo. La aplicación incluye herramientas administrativas para definir los roles de los distintos usuarios, cuotas para cada usuario, seguridad a nivel de documento, un completo registro de la actividad y la configuración de tareas automáticas. Se diferencia de otras soluciones en que está optimizado para la gestión documental pura y dura y presenta menos funcionalidades adicionales. Cuenta con módulos muy interesantes como el de archivo de correos electrónicos y el de factura electrónica, así como con un motor de flujos de trabajo y un entorno de desarrollo (‘framework’) de aplicaciones.

4.2.1 Distribuciones

OpenKM está disponible en tres versiones: ‘community’, profesional y ‘cloud’. La primera está soportada por desarrolladores y técnicos de la comunidad de usuarios, pero no directamente por el equipo de desarrolladores de OpenKM. Mediante la licencia GNU se garantiza la libertad de utilizar, estudiar, compartir y modificar el software.

4.2.2 Estructura y funcionalidades

Su arquitectura está dividida en cuatro niveles o capas, algunas compuestas por dos o más módulos.

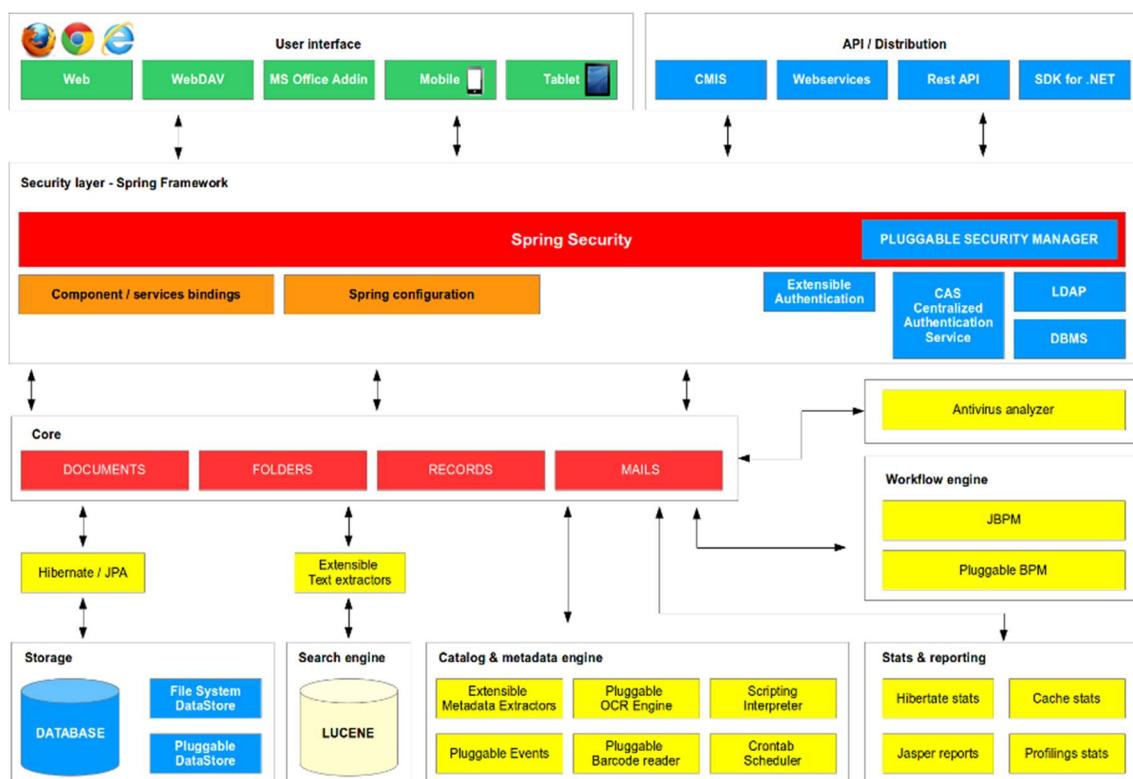


Figura 4. Fuente: <https://www.openkm.com/es/arquitectura.html>

Este software presenta los siguientes módulos:

- **Multimpresa.** Se organiza como una instancia o sistema único por cada empresa que emplea el programa. Ello posibilita el que varios clientes personalizados puedan trabajar en el mismo entorno, y así compartir el coste del servidor, al igual te los costes de estructura y mantenimiento del personal.
- **Archivo de correos electrónicos.** Emplea el propio repositorio de gestión documental como sistema para el almacenamiento de mensajes de correo electrónico.
- **Factura electrónica.** Este módulo extrae las facturas electrónicas en formato XML, y además permite firmarlas digitalmente, descargarlas en PDF, y enviar notificaciones.
- **Motor de ‘workflow’ integrado.** Automatiza el flujo de trabajo establecido dentro del sistema.
- **Generador de informes.** Puede generar estos de forma automatizada una vez introducidos los datos en el sistema.
- **Códigos de barras.** Incluye un módulo para identificar y leer estos códigos, incluyendo, por supuesto, los códigos QR. A ello hay que añadir que la propia estructura de ‘plugins’ de OpenKM permite integrar nuevos formatos de estos códigos en el propio sistema.
- **CMIS.** Este módulo permite interoperar diferentes sistemas de gestión por medio de internet. Así podemos trabajar con diversos repositorios de datos, y trasladar la información, publicarla y buscarla de unos a otros, utilizarla en aplicaciones colaborativas, etcétera.
- **Apps.** Incorpora un conjunto de aplicaciones para diferentes usos: sincronización de carpetas locales, extensión de las funcionalidades del paquete ofimático Microsoft para interactuar con OpenKM, firma electrónica, captura de imágenes, escaneo, impresora virtual, y edición rápida de documentos.
- **Capa de seguridad.** OpenKM es una aplicación escrita en java que utiliza el marco de desarrollo Spring. El módulo más relevante es la capa de seguridad que centraliza la gestión de los accesos permitidos a los usuarios, en función de sus credenciales. OpenKM permite implementar lógicas de seguridad a medida. El proceso de autenticación en OpenKM puede llevarse a cabo mediante un servicio centralizado de autenticación LDAP o por medio de una base de datos donde se encuentren registrados los usuarios.

- **Repositorio.** El núcleo de OpenKM centraliza e implementa la lógica de gestión y procesamiento de los distintos tipos de objetos que se almacenan en el repositorio. Estos objetos son nodos de tipo documento, carpeta, correos electrónico y registros, así como el conjunto formado por las estructuras de metadatos.
- **Almacenamiento.** OpenKM utiliza Hibernate para el mapeo de datos ORM ('Object Relation Mapping'), soportando distintos sistemas gestores de bases de datos relacionales como PostgreSQL, MySQL, Oracle, MS SQL Server, DB2 entre otras. El conjunto de la capa de metadatos se almacena en una base de datos, mientras que los objetos binarios (documentos) en función del tipo de almacenamiento seleccionado, se almacenarán bien en el sistema de ficheros, en una base de datos o en una implementación específica de ese almacenamiento (una de las ventajas de esta aplicación).

4.2.3 Cliente

Mediante un navegador web se puede acceder a los documentos desde cualquier lugar, no se precisa instalar software cliente. Permite la administración, previsualización de ficheros multimedia, 'drag & drop' desde el escritorio. Open KM Implementa una interfaz de usuario Web 2.0 rápida y usable en el marco de desarrollo GWT ('Google Web Toolkit') que soporta los navegadores Firefox, Safari, Chromium y Chrome, así como las últimas versiones de Opera. También dispone de una interfaz de usuario adaptada a dispositivos móviles.

4.2.4 Escritorio

En el acceso al sistema, el escritorio se visualiza con las opciones del menú y tres áreas de trabajo diferentes: (1) árbol de carpetas; (2) navegador de documentos y (3) propiedades de documentos y carpetas, además de la información del usuario que está conectado en ese momento. Las áreas de trabajo permiten cambiar entre la búsqueda y la lista de carpetas y documentos.

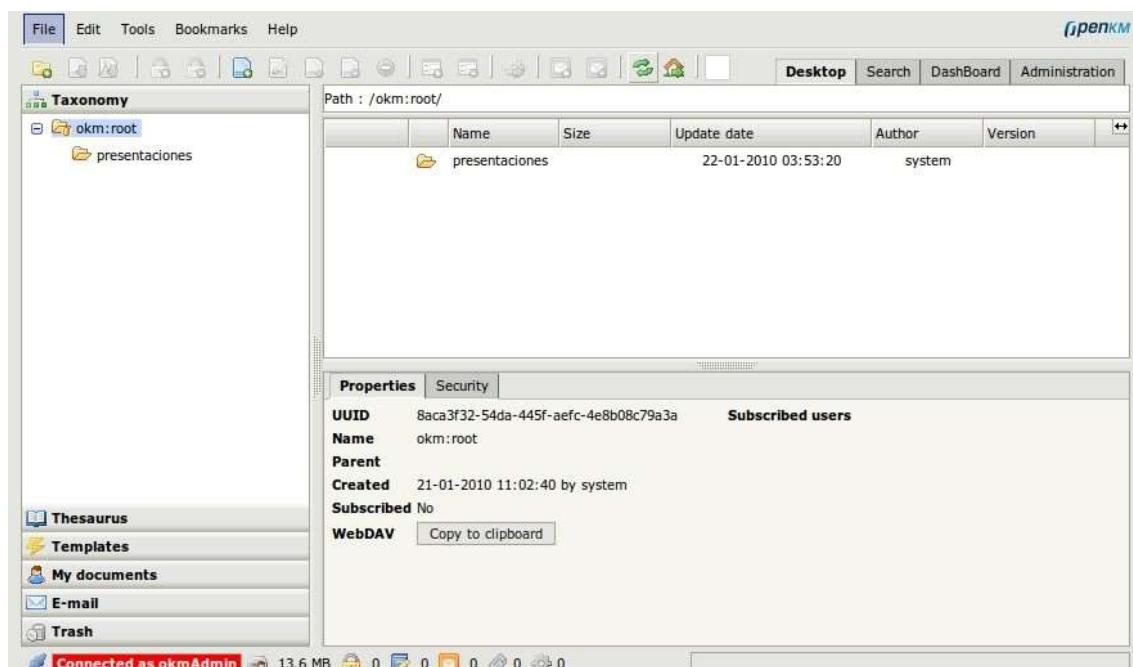


Figura 5: Vista del escritorio de OpenKM. Fuente: elaboración propia.

En el escritorio se dispone de cuatro pestañas para permitir cuatro modos de trabajo:

- Escritorio ('desktop'): para completar las tareas principales con carpetas y documentos (crear, añadir, mover, descargar, etc.).
- Búsqueda ('search'): para buscar documentos.
- Tablero de mando ('dashboard'): para acceder al flujo de trabajo, a la vista de documentos de usuario, a la vista general de documentos y el servicio de notificaciones de noticias.

- Administración ('administration'): utilidades de administración disponibles para usuarios con privilegios de usuario administrador.

A la izquierda del escritorio, está el árbol de carpetas o taxonomía, definido como un árbol a través del cual se organizan los contenidos del repositorio por categorías y subcategorías. La sección de panel del escritorio ofrece una vista de los conceptos de categorías, organizadas en una estructura de carpetas jerárquicas. No solo es una forma especial de ver el repositorio: el repositorio está organizado físicamente por la vista de taxonomía. En la sección superior puede verse el navegador de documentos, que permite ver la lista de documentos y carpetas disponibles en un nodo del árbol específico (taxonomía). OpenKM dispone de plantillas de documentos que pueden ser usadas por los usuarios para crear nuevos documentos. También existen plantillas para las estructuras de "árbol de carpetas" que ayudan al usuario a organizar esos documentos. Así, la organización crea nuevos documentos a partir de documentos dentro del panel de plantillas. Los documentos personales del usuario forman una estructura de árbol disponible para que cada usuario pueda guardar algún tipo de documento solo accesible por el mismo usuario.

La vista de tesoro ofrece una visualización de los términos conceptuales, organizados en una estructura de carpetas SKOS ('Simple Knowledge Organization System'). El tesoro está definido como un árbol a través del cual se organiza la información como una vista del repositorio organizado por conceptos/términos. Los conceptos SKOS pueden tener relaciones de muchos a muchos, pero son representados como una estructura de árbol jerárquico. La sección de panel del escritorio Las relaciones jerárquicas entre los términos dependen de cómo el tesoro está definido. Algunos documentos son enlazados con muchas carpetas SKOS al mismo tiempo. Cada vez que alguna carpeta de estas carpetas es seleccionada, todos los documentos relacionados serán mostrados en el panel del explorador de documentos.

El tablero de mando es un servicio que ofrece información importante y de gran utilidad al usuario (documentos editados, bloqueados, descargados, suscritos, últimos modificados o subidos), documentos generales (documentos más vistos / subidos / modificados) y correos electrónicos (nuevos correos y documentos adjuntos), documentos vistos en consultas específicas, un servicio de búsquedas actualizadas (agente de búsqueda, noticias), flujos de trabajo, nube de palabras claves (nube de etiquetas) y servicio RSS.

4.2.5 Motor de workflow

El software de gestión documental de OpenKM, incorpora el motor de flujo de trabajo ('workflow') JBPm para la definición y la reutilización de la lógica de negocio, simplificando los procesos de gestión y ayudando en una mejor coordinación. La arquitectura java implementada permite que pueda conectar con otros motores de flujos de trabajo de forma transparente. Todos sus contenidos se almacenan de forma que tengan sentido para las necesidades de sus flujos de negocio y puedan ser fácilmente recuperables mediante búsquedas globales inteligentes.

- Creación de flujos de trabajo complejos.
- Flujos de trabajos de revisión, aprobación y validación.
- Soporte para flujos de trabajo lineales o en paralelo.
- Asignación de tareas a grupos de usuarios.
- Monitorizar las tareas de los flujos de trabajo, el estado y los procesos.
- Mensajes de notificación por correo electrónico configurables.

Dado un evento específico, puede aplicar reglas de validación y ciertas acciones se realizan sobre los documentos sin intervención humana:

- Captura de metadatos.
- Clasificación de documentos.
- Ejecución automática de un flujo de trabajo.
- Firma electrónica.

4.2.6 Motor de búsqueda, catalogación y metadatos

OpenKM utiliza Solr/Lucene como motor de búsqueda. Todos los objetos con los que se trabaja sean o no binarios, son indexados por el motor de búsqueda. En el caso de los objetos binarios tales como documentos de Office, PDF o imágenes (entre otros), son añadidos a una cola de indexación. De forma previa a ser procesados por Lucene, los documentos son analizados por extractores de texto. En el caso de las imágenes, se procesan por medio de un motor de OCR para identificar cadenas de texto, las cuales serán utilizadas durante el proceso de indexación. Los usuarios sólo pueden acceder a aquella información sobre la que tienen privilegios. El motor de códigos de barras permite identificar

y leer códigos de barras en los documentos. La arquitectura java implementada en OpenKM, permite que se puedan extender la capacidad de lectura y procesamiento de los formatos de códigos de barras. OpenKM permite planificar, ejecutar y controlar el proceso de captura automática de metadatos, así como automatizar procesos complejos de una forma totalmente transparente para el usuario. Se integra con diversas soluciones de OCR de código abierto (como Tesseract) y comerciales (como Abby, Kofax o Cognitive).

4.2.7 Estadísticas e informes

Este sistema proporciona a los administradores una potente fuente de información, mediante la cual controlar el estado de la aplicación y adelantarse a problemas que implican la ampliación del hardware. Esta información ayuda en la toma de decisiones para establecer los valores óptimos para los objetos en la cache de segundo nivel, la parametrización de los recursos utilizados por el sistema gestor de la base de datos y cómo son utilizados.

4.2.8 Gestión de documentos de archivo ('record management')

La gestión de documentos de archivo ayuda en la gestión eficaz y sistemática de la creación, la recepción, el mantenimiento, el uso y la disposición de documentos, incluye los procesos para incorporar y mantener, en forma de documentos, la información y prueba de las actividades y operaciones de la organización. El sistema de gestión de documentos de archivo de OpenKM incorpora, gestiona y facilita el acceso a los documentos a lo largo del tiempo. El proceso de trazabilidad de las acciones de OpenKM permite controlar y establecer etapas y plazos en las acciones a quienes las llevan a cabo. La disposición en OpenKM es una serie de procesos asociados con la aplicación de decisiones de transferencia, destrucción o conservación de documentos, que se documentan en los calendarios de conservación u otros instrumentos. La destrucción es el proceso de eliminación o borrado de documentos sin que sea posible su reconstrucción. El plan de archivo de OpenKM es un conjunto de reglas que controla, para cada tipo de documento de archivo de la empresa; dónde tienen que conservarse los documentos de archivo, las políticas que se aplican a ellos, cuánto tiempo deben conservarse, cómo deben ser eliminados y quién es el responsable de su gestión.

Para el desarrollo de OpenKM, se utilizaron herramientas pertenecientes a tecnologías abiertas y basadas en estándares java y la aplicación de servidor JBoss, lo que le posibilita al sistema ser instalado y utilizado en diferentes plataformas. Proporciona una plataforma interoperable para sacar provecho a través de las aplicaciones y los repositorios de la empresa. En el módulo de distribución, el gestor implementa el protocolo de Servicios de Interoperabilidad de Gestión de Contenidos (CMIS), un estándar abierto que define una capa intermedia que permite la interconexión y control de distintos sistemas de gestión documental y repositorios utilizando protocolos web. Adicionalmente, expone una completa API mediante servicios web (SOAP y REST), que puede utilizarse como punto de integración con terceras aplicaciones. Para el desarrollo de aplicaciones se encuentran disponibles los SDK's (Software Development Kit) para java, .NET y PHP, que encapsulan toda la API de OpenKM. Dispone de conectores para las aplicaciones DICOM, Autocad, Google Drive, MS Office, Joomla, Wordpress, BonitaSoft, Kofax, Abby Flexicapture.

4.3 OpenProdoc

SGD de código abierto desarrollado por Joaquín Hierro (https://jhierrt.github.io/openprodoc/index_ES.html) con el propósito de que todo el mundo pueda acceder a un gestor documental profesional. Reúne las funcionalidades habituales de la gestión documental en un motor que se puede ejecutar en la mayoría de sistemas informáticos. A diferencia de otros gestores documentales de código abierto, absolutamente todo su código está liberado. Es el que menos funcionalidades ofrece, pero en compensación es el más sencillo de administrar y el que menos recursos consume. Esto lo convierte en una opción para las organizaciones que busquen un software de gestión documental sencillo pero sólido. Dispone de una versión portable y de una versión multiusuario/empresarial. OpenProdoc, en su versión 3.0.3, ofrece un conjunto de herramientas que permiten manejar documentos y compartirlos entre los usuarios de la empresa, incluyendo una ficha descriptiva de cada documento para poder describirlo y localizarlo. Permite definir tipos y subtipos de documentos y de expedientes, cada uno de ellos con sus metadatos, gestionar el ciclo de vida de los documentos, generar informes, gestionar usuarios, exportar e importar el contenido del repositorio, crear varios tesauros multiidioma e intercambiarlos en formato SKOS o gestionar referencias bibliográficas e intercambiarlas en formato RIS. Permite importar documentos digitalizados con las aplicaciones Kofax Capture o Abby FineReader. Ofrece las funcionalidades esenciales para gestionar la documentación de una organización: gestión de reglas de acceso, definición de tipos de documentos y carpetas, gestión de tesauros, descripción y clasificación de documentos, control de versiones, programación de tareas, búsqueda por metadatos y texto libre, importación y exportación de documentos, generación de informes, auditoría y encriptación de repositorios.

4.3.1 Distribuciones

Existen varias distribuciones de OpenProdoc, todas ellas incluyen básicamente las mismas funciones. Las definiciones, configuraciones y documentos introducidos pueden importarse y exportarse entre instalaciones, por lo que el trabajo realizado en una puede exportarse e importarse en otra. El sistema se ha diseñado utilizando un mínimo de instrucciones y evitando características propietarias, por lo que debería funcionar en la mayoría de los entornos.

- **Portable:** tiene todas las funciones y es portable, no es necesaria la instalación. Dispone de un servidor embebido que pueden utilizar simultáneamente varios usuarios desde un navegador. Incluye un cliente ‘swing’ (“embebido”) para realizar operaciones masivas. Versión adecuada para hacer pruebas de modelado de tipos documentales o creación de tesoros.
- **Instalación base y cliente swing:** utiliza un interfaz de usuario idéntico a la opción portable o web, pero permite el trabajo multiusuario simultáneo conectándose a servidores de base de datos y de archivos. Adecuada para administradores o para operaciones masivas que no son posibles con un interfaz web (como importar o exportar una estructura de carpetas y documentos en disco).
- **Cliente Web:** ofrece un interfaz web que permite realizar desde un navegador la mayor parte de las funciones (exceptuando algunas de importación/exportación masiva).

Las dos últimas opciones son compatibles, es decir, unos usuarios pueden usar el cliente web mientras otros utilizan el cliente ‘swing’.

4.3.2 Estructura y funcionalidades

Para utilizar OpenProdoc lo mínimo imprescindible que se necesita es disponer de una base de datos y un entorno java JVM 1.6 o superior. En la base de datos se almacenará toda la información de la configuración, seguridad y metadatos de los documentos y carpetas. Dependiendo de donde se desee almacenar los documentos podrá optarse por una carpeta del servidor o el uso de esta base de datos para almacenar los documentos. En cuanto al acceso, para el cliente ‘swing’ (tanto para acceso a documentos como administración) no es necesario nada adicional. Para el uso del cliente web es necesario disponer de un servidor de aplicaciones. Se trata de una herramienta multiplataforma, siendo compatible con distintos:

- Sistemas operativos (linux, windows y macOS).
- Servidores de aplicaciones (Tomcat, JBoss/WildFly, GlassFish, Jetty, etc.).
- Bases de datos (MySQL, MariaDB, PostgreSQL, Oracle, MS SQL Server, etc.).
- Navegadores (Firefox, Chrome, Edge, Opera, Safari, etc.).
- **java:** java 1.8, 1.9, 1.10, 1.13, OpenJDK 1.8

Como otros gestores documentales tiene una estructura compuesta por tres elementos:

- El componente principal es un servidor de aplicaciones java J2EE: atiende las peticiones, presenta pantallas y coordina los procesos. Se descompone en varios módulos interconectados para atender las distintas funciones del gestor.
- Un servidor de base de datos para almacenar la configuración, los metadatos de los documentos o expedientes que se introducen y referencias a los documentos almacenados.
- Un sistema de almacenamiento de los documentos.

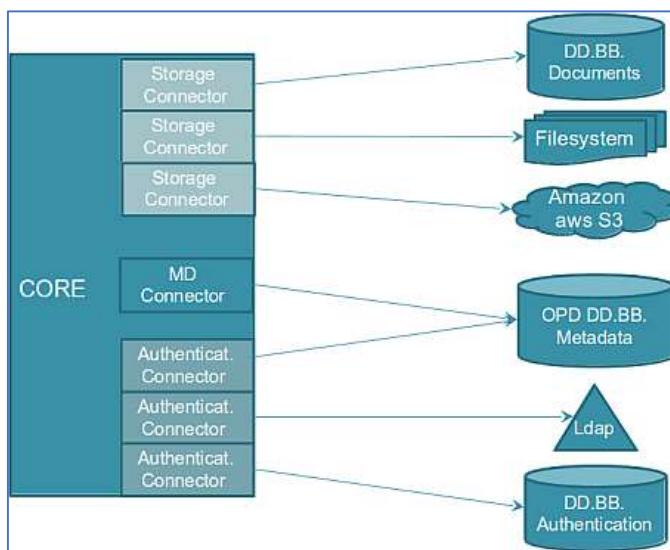


Figura 6: Arquitectura de OpenProdoc. Fuente: <https://jherrot.github.io/openprodoc/help/ES/Architect.html>

4.3.3 *Repositorios de documentos*

Pueden crearse un número ilimitado de repositorios de cada tipo, según las necesidades de la entidad o el proyecto. Los metadatos de un documento se almacenan siempre en la base de datos. Para los documentos pueden utilizarse diversos tipos de almacenamiento de los documentos. Entre los sistemas disponibles (conectores) contamos con el almacenamiento en:

- una carpeta de sistema (o carpeta de red) accesible para los clientes web o ‘swing’.
- una tabla de base de datos como atributo BLOB (recomendable para archivos pequeños o tipos documentales con pocos ejemplares)
- un servidor FTP (recomendable para tipos documentales con poco acceso).
- nube de Amazon S3 (permite almacenar información en todo el mundo con seguridad y con la velocidad y costes contratados).

Cada tipo de repositorio tendrá características propias. Por ejemplo, aunque suele ser más eficaz almacenar los documentos en una carpeta de archivos, si los documentos son muy pequeños puede ser más eficaz almacenarlos como BLOB en una base de datos.

4.3.4 *Interfaz de usuario*

La ventana principal de OpenProdoc tiene dos zonas principales de trabajo:

- El árbol de carpetas a la izquierda.
- Zona principal a la derecha.

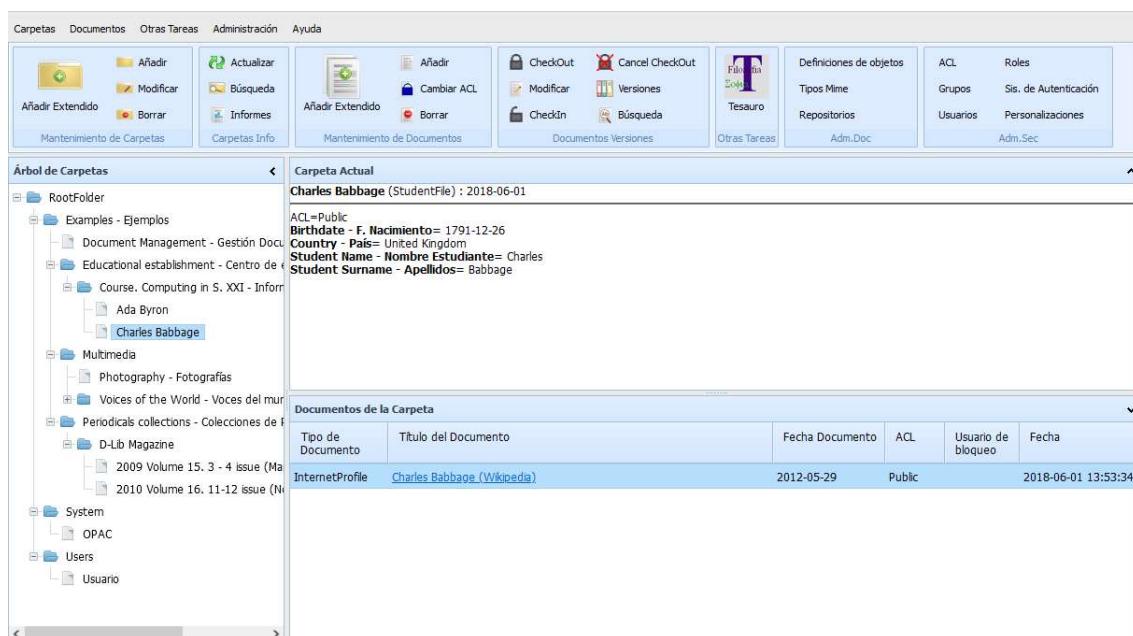


Figura 7: Ventana principal de OpenProdoc. Fuente: elaboración propia.

En la parte superior se dispone de unos menús desplegables para realizar operaciones sobre la carpeta actual o sobre el documento actual, así como otra serie de opciones. Las opciones disponibles dependerán de los permisos y perfil del usuario con el que nos conectamos.

La seguridad de documentos y carpetas/expedientes se basa en tres elementos: (1) los usuarios y grupos; (2) los perfiles/roles y (3) las ACL ('access control lists'). Todos los documentos y carpetas tienen asignado una ACL que determina las operaciones que puede realizar un usuario. Cuando un usuario pretende realizar cualquier operación (incluida la visualización) sobre un documento o carpeta, OpenProdoc consulta la ACL del elemento y verifica los permisos asignados a cada uno de los usuarios o grupos. El árbol de carpetas presenta una estructura con iconos y nombres de carpeta. La estructura de carpetas depende de las estructuras creadas y de los permisos (sólo aparecen las carpetas y documentos sobre los que el usuario tiene permisos). Permite navegar por las distintas estructuras que componen el repositorio:

- Rama principal (“rootFolder”).
- Carpeta “users”: carpetas personales de usuario. Según los permisos, podrá verse solo la carpeta personal propia u otras carpetas de usuarios sobre las que tenga autorizado el acceso.
- Carpeta “system”: documentos y elementos de configuración.

No hay límite al número de niveles de carpetas o número de documentos que puede insertarse en una carpeta, aunque es conveniente mantener una estructura jerárquica que pueda ayudar a organizar la información. Cuando se navega por las carpetas, la zona derecha contendrá a su vez dos subzonas:

- El detalle de la carpeta actual, con todos los metadatos de la carpeta.
- La tabla de documentos contenidos en la carpeta actual, mostrando título, tipo, fecha, usuario que tiene bloqueado el documento para modificarlo y la fecha en que se creó o modificó el documento en OpenProdoc.

Podemos añadir un documento del tipo básico (denominado PD_DOCS) introduciendo solo información de los metadatos *Título* y *Fecha del documento*. Los tipos de documentos disponibles para un usuario aparecerán en un desplegable. Además de los tipos documentales “normales” (con documento y metadatos almacenados en el repositorio), pueden definirse tipos de documento “referenciales” con una referencia (URL) a los mismos almacenada en el repositorio. Al elegir el tipo de documento, automáticamente aparecen los metadatos que se han definido para ese tipo documental distribuidos en un formulario (único o con varias solapas). Cada metadato tiene un nombre y un texto descriptivo que podrá contener recomendaciones o normas de introducción del metadato. Por defecto los permisos asignados a un documento son los mismos de la carpeta/expediente que lo contiene, pero esos permisos pueden cambiarse al insertar en el “modo extendido”.

Al finalizar la entrada, el documento se añadirá al repositorio con el tipo documental y los metadatos introducidos, añadiendo además un identificador único del documento, información de permisos (ACL) y datos de trazabilidad (auditoría) como el nombre del usuario que ha introducido el documento y la fecha/hora de inserción. Tras la inserción estará accesible para todos los usuarios con permisos sobre el mismo. Si un usuario tiene permisos de escritura o de borrado podrá elaborar una nueva versión del documento. Para ello se deberá bloquear el documento que quedará reservado hasta que publique la nueva versión o se anulen los cambios. Tras bloquear un documento, el usuario de bloqueo aparecerá indicado en la lista de documentos de forma que todos puedan ver los documentos en proceso de modificación.

4.3.5 Tesauros

OpenProdoc permite crear y utilizar múltiples tesauros (multilingües o monolingües) utilizando el estándar SKOS, importar tesauros ya creados o exportarlos para su uso en otras entidades o entornos. Los tesauros pueden utilizarse de forma aislada, como un gestor de tesauros, o integrados con los tipos documentales o de carpetas. Pueden ser tesauros “formales” completos, con toda la jerarquía y relaciones habituales o simplemente listas controladas de valores con estructura plana (entrada de valores de metadatos asociada a un tesauro).

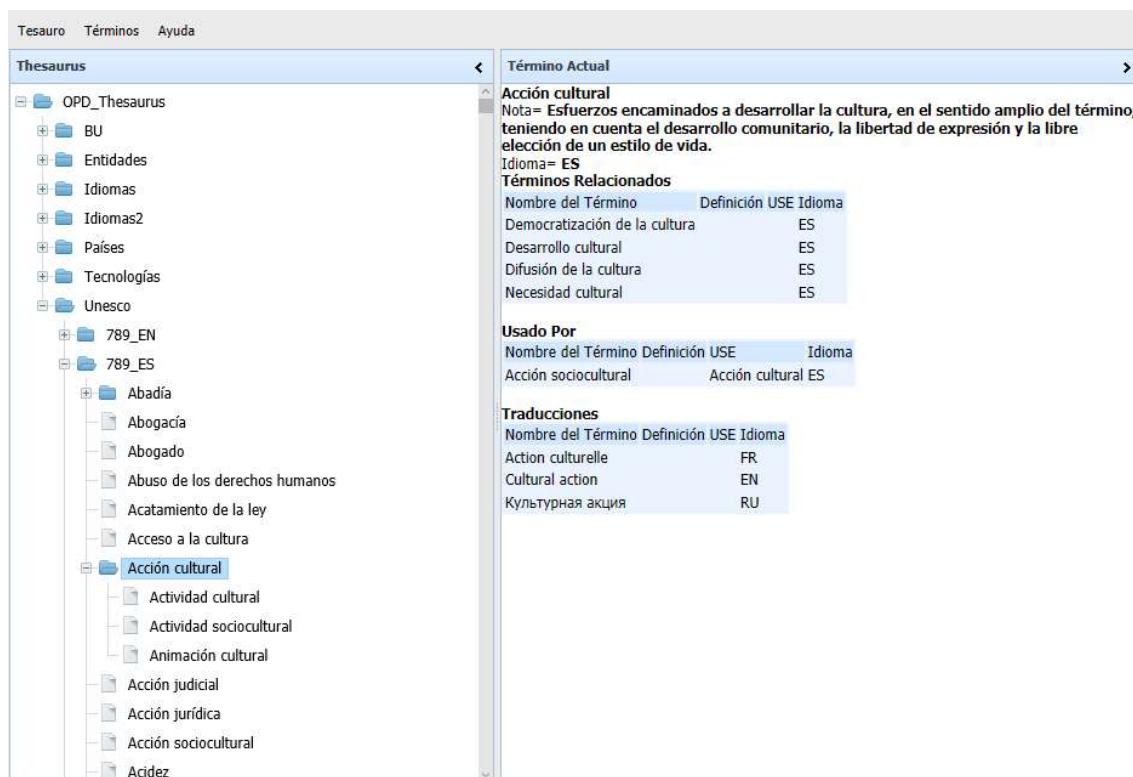


Figura 8: Vista de tesauro en OpenProdoc. Fuente: elaboración propia.

4.3.6 Tareas ligadas a eventos

Las tareas se desencadenan cuando se produce un evento (creación, modificación o borrado) sobre una clase de objeto documental (tipo carpeta o tipo documento). Para cada tipo documental y evento las tareas deben definirse con orden de ejecución por medio de un número de secuencia. Se aplican a un tipo documental y todos sus subtipos, según el orden de descendencia y orden de secuencia.

4.3.7 Indización y búsqueda de documentos

Para la indización (indexación) y búsqueda por contenido de los documentos se utiliza el motor Lucene de forma que la sintaxis de búsqueda y comportamiento son los habituales de este sistema. El repositorio se puede configurar para

activar la indización por texto completo de tipos documentales (algunos o todos) y para una parte o toda la estructura de carpetas. La inserción de nuevos documentos producirá un análisis del documento y la indización por todas las palabras, incluyendo metadatos internos del documento (propiedades como autor, asunto o palabras clave, etc.) de formatos como PDF o Word). Si hay documentos no indizables, como imágenes (al no haberse aplicado OCR) o formatos no soportados, será posible recuperarlos por los metadatos o navegando por la jerarquía de carpetas, pero no por el contenido. Las búsquedas se realizan mediante un formulario de búsqueda (“*Documentos → Búsqueda de Documentos*”). Una vez localizados, esto permitirá realizar diversas operaciones sobre los documentos. El formulario solicita la selección del tipo de documento sobre el que se desea realizar la búsqueda mediante una lista desplegable de tipos documentales. Al elegir el tipo de documento, se actualiza el formulario con los metadatos propios de ese tipo documental. Además de la búsqueda sobre metadatos, hay cuatro criterios de búsqueda comunes a todos los tipos documentales: por texto libre/contenido de los documentos, en la jerarquía actual de carpetas, sobre versiones y sobre subtipos. Pueden combinarse estas opciones con los criterios de búsqueda sobre metadatos.

The screenshot shows a window titled "OpenProdoc" with a sub-titile "Búsqueda de Documentos". At the top, there is a dropdown menu labeled "Tipo de Documento" set to "Base Document". Below the title bar are three tabs: "Búsqueda de Documentos" (selected), "Resultados de la búsqueda", and "Generación de Informes". The main area is titled "Búsqueda de Documentos" and contains several search criteria:

- Subtipos:** Checkboxes for "SubCarpetas" and "Versiones".
- Búsqueda por texto completo:** A search input field.
- Filter fields:** A grid of 12 input fields for various metadata fields, each with an equals sign (=) and a dropdown arrow. The fields are:
 - PId
 - Fecha Documento
 - Título del Documento
 - Usuario de bloqueo
 - Fecha de Expurgo
 - ACL del Documento
 - Repositorio
 - Tipo Mime
 - Nombre Archivo
 - Carpeta
 - Versión
 - Autor
 - Fecha
- Buttons:** "Aceptar" (Accept) and "Cancelar" (Cancel).

Figura 9: Pantalla de búsquedas en Openprodoc. Fuente: elaboración propia.

4.3.8 Importación y exportación de documentos y carpetas

Las opciones de exportación permiten volcar a disco carpetas y documentos con sus metadatos y tipos documentales. Esto puede utilizarse para compartir información con entidades o departamentos que no comparten la misma instalación de OpenProdoc. Permitirá importar documentos y carpetas sin necesidad de tener que describir de nuevo toda la información y metadatos de cada elemento. También puede utilizarse dentro de una misma entidad para crear copias dentro de la misma instalación de OpenProdoc o entre entornos (entorno de pruebas y entorno de producción). La versión cliente ‘swing’ puede utilizarse para importar documentos digitalizados, clasificados y catalogados con aplicaciones como Kofax Capture y Abbyy FineReader. Los documentos almacenados en una estructura de carpetas, junto con sus metadatos y clasificación con el formato estándar respectivo, se importan como carpetas y documentos OpenProdoc del tipo documental indicado en el fichero de metadatos. En el momento de importar deben existir las definiciones de tipos documentales y tipos de carpeta (con, al menos, los mismos metadatos) y tipo Mime.

4.3.9 Integración

La API java (el núcleo de OpenProdoc) incluye todas las funciones y servicios, excepto el interfaz de usuario y el almacenamiento, disponibles en OpenProdoc, incluyendo la gestión de usuarios, grupos, repositorios, definición de objetos, tareas, etc. Permite que otras aplicaciones puedan invocar servicios del gestor y puedan automatizarse algunas tareas. Este componente, el núcleo, puede embeberse en cualquier programa y se dispone de un gestor documental completo (excepto la interfaz de usuario). La API REST forma parte del servidor OpenProdoc. Publica un conjunto de métodos/servicios que pueden ser llamados desde cualquier lenguaje de programación que soporte el estándar REST (javaScript, java, Python, php, .Net, etc.). Está estructurada en cuatro grupos de servicios: servicios de inicio de sesión/sesión, servicios de carpetas, servicios de documentos y servicios de tesoros. Todas las necesidades habituales relacionadas con los documentos, las carpetas y los metadatos controlados por el Tesoro están incluidas.

5 CONCLUSIONES

La gestión de documentos en una entidad u organización abarca estrategias, métodos y herramientas que se utilizan a lo largo del ciclo de vida del contenido (capturar y crear, organizar, colaborar, buscar y recuperar, publicar y entregar). Debemos tener una visión holística del entorno de la información de toda la organización. No sólo se debe tener en cuenta a la tecnología, sino además a las personas, los procesos y la información existente y accesible.

La gestión de documentos debe ser vista con un enfoque estratégico para la gestión de información, pero también debe analizarse el conjunto de herramientas de software que ayudan a su implementación y que pueden implementarse como un paquete completo o utilizarse como herramientas específicas por separado.

Las tres plataformas tecnológicas analizadas, ayudarían a desarrollar una adecuada implementación del sistema de gestión documental, cada una de ellas en un escenario diferente (grandes, medianas o pequeñas organizaciones). Alfresco es una herramienta que proporciona todo el valor que ofrecen las grandes plataformas ECM tradicionales: gestión de contenidos, gestión del ciclo de vida de los contenidos, gestión de portales web, trabajo colaborativo, etc. Con toda la complejidad del modelo documental y la ineludible dificultad de aprendizaje y uso que implica. OpenKM es más sencillo de utilizar que Alfresco. Está optimizado para la gestión documental y presenta menos funcionalidades adicionales (que muchas veces no llegan a emplearse en la práctica). Como Alfresco, también dispone de un módulo específico de gestión de documentos de archivo. Un aspecto a tener especialmente en cuenta es que la combinación de gestión documental y gestión de archivo que proporcionan Alfresco y OpenKM puede ayudar a las organizaciones a cumplir con los requisitos funcionales indicados en la Norma ISO 16175. OpenProdoc, como herramienta completamente de código abierto y gratuita, se presenta como una alternativa adecuada para pequeñas organizaciones con limitaciones presupuestarias. Como OpenKM, se ha desarrollado también con el objetivo de que sea muy sencilla de utilizar y capaz de cubrir eficazmente todas las funciones nucleares de la gestión documental sin sobrecargar con funciones usadas esporádicamente. No incluye funciones como la planificación de procesos, flujos de trabajo o CMIS, aunque está previsto su soporte en versiones futuras. Si proporciona la posibilidad de definir y asignar tareas controladas por eventos (procesos simples, que pueden programarse).

Las tres plataformas disponen de versiones de fuentes abiertas, utilizan estándares abiertos y cuentan con una comunidad de usuarios. Los usuarios tienen la opción de realizar ajustes en el código fuente, consiguiendo que un sistema de gestión documental cumpla con los requisitos específicos. La tecnología puede permitir una gestión de los contenidos, pero la estrategia y el proyecto subyacente debe ser lo primero y más importante a definir. Recopilar las necesidades, desarrollar una estrategia, y poner en práctica el plan, con la definición de las políticas de información, la configuración, adaptación, implementación y pruebas del sistema, y las acciones de formación adecuadas para la implantación.

6 BIBLIOGRAFÍA

- AIIM. (2021). *What is Enterprise Content Management (ECM)?* <https://www.aiim.org/resources/glossary/enterprise-content-management>
- Alfresco Docs. (2022). <https://docs.alfresco.com>
- Alfresco Connect. (2022). <https://hub.alfresco.com>

- Bustelo Ruesta, Carlota. (2011). Los grandes temas relacionados con la gestión de documentos: Desafíos y oportunidades. *El profesional de la información*, 20(2), 129-134. <https://doi.org/10.3145/epi.2011.mar.01>
- Hierro, Joaquín. (2021). Índice Ayuda OpenProdoc. <https://jhierrrot.github.io/openprodoc/help/ES/HelpIndex.html>
- Hierro, Joaquín. (2018). *Introducción a los gestores documentales con OpenProdoc*. <https://jhierrrot.github.io/openprodoc/Docs/Introducci%C3%B3nG.D.OpenProdoc.pdf>
- Hierro, Joaquín. (2021). OpenProdoc: gestor documental. Preguntas Frecuentes. https://jhierrrot.github.io/openprodoc/FAQ_ES.html
- OpenKM. (2022). *Blog de OpenKM*. <https://www.openkm.com/es/blog>
- OpenKM. (2022). *Documentación de OpenKM 7.1*. <https://docs.openkm.com/kcenter/view/okm-7.1/>
- Pastor Bermúdez, Andrés. (2020). El repositorio, dónde viven los documentos electrónicos. En: Bustos Pretel, Gerardo. *La gestión del documento electrónico* (2^ªed.). Madrid, Wolters Kluwer, 619-633.
- European Knowledge Center for Information Technology (EKCIT)/ tic.PORTAL (2022). *Guía de Gestión Documental: toda la información esencial*. <https://www.ticportal.es/guias/guia-gestion-documental>
- UNE-ISO 16175-1:2021 (2021). *Información y Documentación. Principios y requisitos funcionales para documentos en entornos de oficina electrónica. Parte I: Generalidades y declaración de principios*. Madrid: Asociación Española de Normalización.