

Odilo, más archivo integral en universidades: hacia la conservación a largo plazo de la documentación y los datos científicos

Francisco José Valentín Ruiz

Director del Área de Archivos y Preservación Digital

Odilo



Odilo, más archivo integral en universidades: hacia la conservación a largo plazo de la documentación y los datos científicos

10/10/2019

Francisco José Valentín Ruiz
fjvalentin@odilotid.es

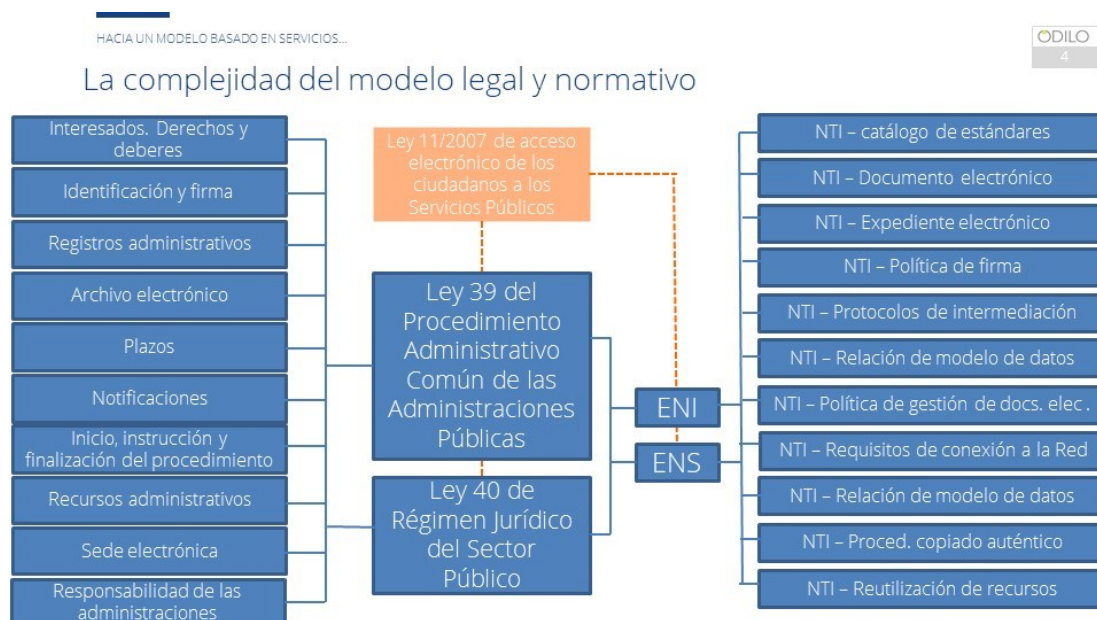


1. Hacia un modelo basado en servicios para la gestión del archivo y la preservación digital

Para entender el título de esta comunicación es necesario retrotraernos un año, a 2018, a la celebración de estas mismas Jornadas CAU en Salamanca. En aquella ocasión presentamos un modelo de archivo integral aplicado al caso de la Universidad de Castilla-La Mancha. Por aquel entonces, Odilo estaba llevando a cabo un proceso de implantación de un sistema integral de archivo físico, electrónico y de preservación digital con esta Universidad.

De modo que la referencia «más archivo integral en universidades» recoge aquella propuesta y la amplía con algunos elementos adicionales. Merece la pena, sin embargo, recordar algunas de las principales conclusiones de la comunicación «Hacia un modelo basado en servicios para la gestión del archivo y la preservación digital» presentada en 2018:

- **Un sistema administrativo garantista y complejo: la materialización de la administración electrónica y su traslado a los archivos.** Los documentos siguen siendo el núcleo de la actividad administrativa y siguen representando los derechos y obligaciones de los ciudadanos y las administraciones, aunque con el desarrollo de la administración electrónica se haya sustituido su soporte (ahora son documentos electrónicos). Además, el sistema administrativo se ha dotado de una serie de estrategias y elementos para garantizar la validez jurídica, integridad, autenticidad y accesibilidad de los documentos. Estos se han trasladado a través de una estructura normativa compleja que ha impuesto un reto adicional a los profesionales de la archivística encargados de implantar el modelo en sus archivos:





- **Unas normas de interoperabilidad que están llamadas a resolver algunos de los problemas tradicionales del archivo.** La construcción de estructuras predefinidas para el intercambio de expedientes y documentos es un hecho de enorme trascendencia para los archivos: solo los expedientes y documentos que cumplan con una serie de condicionantes mínimos (definidos principalmente en las normas técnicas de interoperabilidad de documento y expediente electrónicos) tendrán cabida en el archivo. Y de ahí la importancia de que los criterios archivísticos se apliquen en las etapas previas al archivo, desde la misma génesis de los documentos. Como en la imagen, con la llegada de la administración electrónica no cualquier cosa puede ingresar en el archivo (al menos en teoría):



- **Pisto manchego y metadatos.** El pisto manchego tradicional se elabora con pimiento y tomate, la adición de otras hortalizas (como la berenjena o el calabacín) implica adulterar la receta tradicional (al menos para los más puristas). Sin embargo, para muchos, entre los que se encuentra el autor de estas líneas, la receta adulterada es mejor que la original (en cualquier caso, no todo puede considerarse pisto manchego y, desde luego, el ejemplo de la imagen no lo sería). Con los metadatos ocurre algo muy parecido: los expedientes y documentos electrónicos puros creados a través del cumplimiento estricto de las normas técnicas de interoperabilidad a base de un conjunto de metadatos mínimos obligatorios serían como un pisto con pimiento y tomate. Para mejorar la receta es necesario incorporar metadatos de gestión documental (eEMGDE es el estándar de facto al que aspiramos) y metadatos de negocio (eEMC se postula como el mejor candidato). Sin estos conjuntos adicionales de metadatos los archivos no tendrán nada.



- **La insuficiencia del archivo electrónico único.** Mientras que la implantación de la administración electrónica conseguirá digitalizar todos los procesos documentales y, en consecuencia, la actividad administrativa dejará de consumir papel, en el archivo, que es el reflejo de la actividad administrativa y guardián de la memoria institucional e histórica, el papel no desaparecerá. En este contexto, además, hay que tener en cuenta que la actividad administrativa no es la única que realizan las administraciones y entidades del sector público (solo sobre esta actividad documental aplica el entramado de las leyes 39/2015, 40/2015, ENI, y normas técnicas de interoperabilidad). De modo que el archivo debe atender a otras necesidades adicionales que van más allá del «archivo electrónico único» de la Ley 39/2015: gestión de la documentación en soportes físicos o convencionales y gestión de documentación electrónica no administrativa (de carácter patrimonial, por ejemplo) o generada en etapas previas a la existencia de la normativa actual.

La insuficiencia del Archivo Electrónico Único (I)

Tramitación



Archivo



La insuficiencia del Archivo Electrónico Único (II): su alcance



- **La tradición archivística y el agua del botijo.** De modo que la actividad de los archivos y el tratamiento convencional no son sustituidos, sino que son ampliados con nuevas funciones, estableciéndose una convivencia de actividades tradicionales y nuevos enfoques relacionados con la administración electrónica. Esta tradición archivística, que ha demostrado su validez durante décadas, debe mantenerse y adaptarse a los nuevos medios. Como el botijo, lo que funciona, no debe ser suprimido.



- **Aislamiento y falta de cooperación de los archivos.** Tradicionalmente muchos archivos habían aplicado modelos propios, únicos, con especificidades. Las circunstancias para llegar a esa situación han sido diversas: falta de recursos, aislamiento, etc. En un contexto interconectado estas prácticas ya no son admisibles y dificultan enormemente la implantación de sistemas tecnológicos.



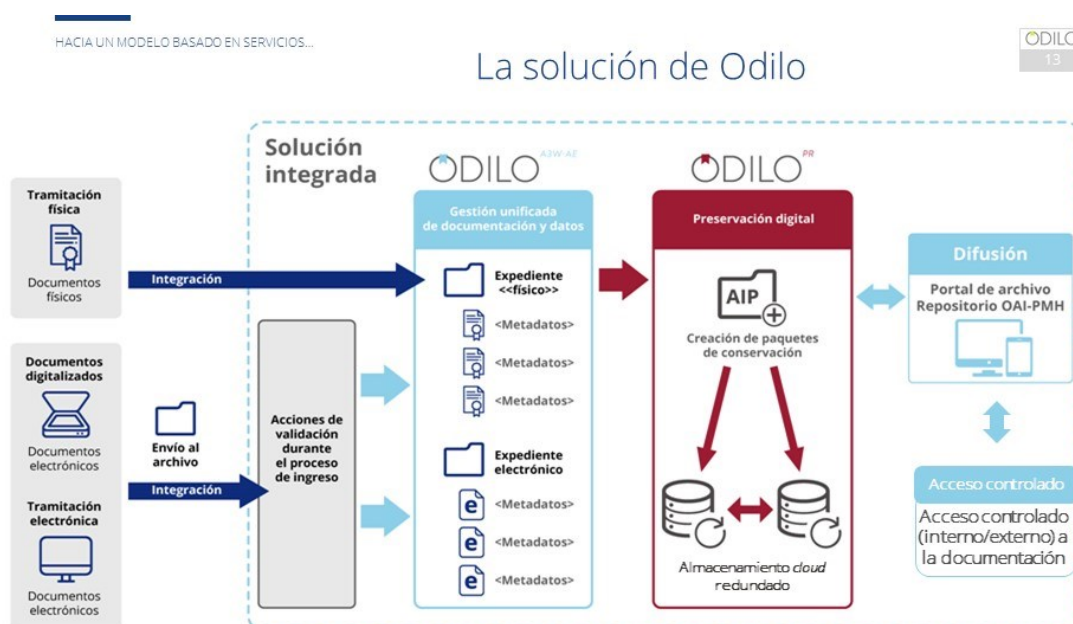
- Promesas cumplidas.** Centrándonos en la actividad de Odilo, también adelantábamos en 2018 que los siguientes objetivos eran la actualización y pleno funcionamiento de un portal de archivos adaptado a las nuevas tendencias y demandas sociales y con capacidad de proporcionar acceso público y, al mismo tiempo, controlado a la documentación (para funcionarios y ciudadanos). Son varios los ejemplos de instituciones en las que están implantadas las soluciones de archivo y preservación de Odilo que, durante el último año, han actualizado sus portales de archivo. La más reciente, la Universidad de Zaragoza:

El catálogo del Archivo Universitario de Zaragoza te permite la búsqueda y localización de los fondos documentales que conserva. Te presentamos dos accesos diferenciados: **Catálogo Archivo Universitario**, que aloja la descripción del fondo universitario propiamente dicho, fondos de escuelas profesionales y colecciones de materiales gráficos y **Catálogo Tesis doctorales**, que conserva tesis doctorales leídas en la Universidad de Zaragoza desde 1998 hasta la actualidad.

El fondo sin migrar del Archivo Histórico queda como un enlace a su antiguo catálogo. Dicho fondo se irá migrando progresivamente al Catálogo Archivo Universitario, por lo que recomendamos a los investigadores que realicen las búsquedas tanto en el Catálogo Archivo Universitario como en el enlace al Archivo Histórico hasta su fusión definitiva.

<https://albarracin.unizar.es/portaArchivo>

- Un modelo de archivo integral: físico, electrónico, preservación digital, difusión y acceso controlado a la documentación.** Esta es, en resumen, la propuesta de Odilo para dar respuesta de manera integral a todas las necesidades del archivo a través de un conjunto de productos (software, soluciones *cloud*) y servicios (consultoría, acompañamiento, preservación...):



2. Más archivo integral en universidades: hacia la conservación a largo plazo de la documentación y los datos científicos

Una vez recordados los conceptos relacionados con la propuesta de archivo integral de Odilo, conviene resaltar algunos elementos claves relacionados con la actividad científica y las universidades:

- Una de las actividades centrales de las universidades es la actividad investigadora. Esta actividad, por lo general, se lleva a cabo de una manera muy independiente, desvinculada de cualquier sistema de gestión documental existente.
- El resultado de la actividad científica se plasma en parte en repositorios institucionales. Un repositorio institucional es un conjunto de servicios de almacenamiento, gestión y difusión de materiales digitales para hacerlos disponibles en abierto o para los miembros de una determinada comunidad científica. Deberíamos añadir que también deben ser accesibles y comprensibles a lo largo del tiempo.
- Es fundamental conservar y difundir este patrimonio intelectual y científico para el progreso de la ciencia y del ser humano. Lo que podríamos denominar una necesidad de conservación a largo plazo.
- La tendencia hacia que el conocimiento científico generado por instituciones financiadas con fondos públicos sea de acceso abierto. Aunque no siempre va a ser así y no todo lo que está relacionado con una investigación va a estar en acceso abierto o público.



De los puntos anteriores resaltaremos uno de especial importancia, en las revistas científicas y otros medios de comunicación científica se publican los **resultados de la investigación** (generalmente no los datos completos que llevan a las conclusiones ni el proceso completo con todo lujo de detalles). Y aunque parezca obvio, no está de más resaltarlo: se publican los resultados de la investigación.

Esto es así porque el procedimiento de publicación científica actual se basa en un sistema de confianza bajo la premisa de que los diferentes elementos de esta cadena de confianza hacen bien su trabajo. Los investigadores trabajan conforme a un método científico y los resultados de las investigaciones se presentan ante la comunidad a través de revistas científicas en las que especialistas anónimos de cada sector (lo pares) evalúan los artículos para poner en cuestión su publicación. Si un artículo no pasa este filtro, no es publicado, pero, si lo supera, se supone que lo planteado es correcto. Por otra parte, existe un sistema complejo de prestigio de las publicaciones que trata de asegurar que no se publica cualquier cosa y que no hay intereses (comerciales, políticos...) de ningún tipo de por medio.



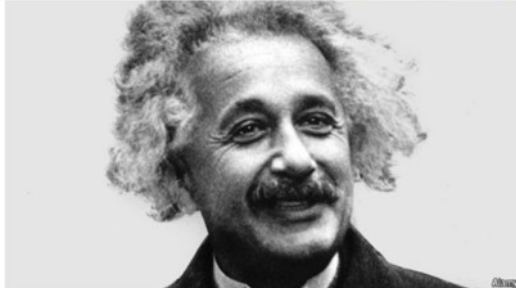
¿Qué ocurre en este sistema aparentemente perfecto? Pues que, como seres humanos que somos, tenemos una cierta tendencia a equivocarnos. Y, sin embargo, a la sociedad solo llega lo más llamativo:

Neutrinos and relativity
Faster than the speed of light

What does an experiment that seems to contradict Einstein's theory of relativity really mean?

Oct 1st 2011

Timekeeper Like 1.8K Tweet



IN 1887 physicists were feeling pretty smug about their subject. They thought they understood reality well, and that the future would just be one of ever more precise measurements. They could not have been more wrong. The next three decades turned physics on its head, with the discovery of electrons, atomic nuclei, radioactivity, quantum

Octubre de 2011

<http://www.economist.com/node/21531006>



Douglas Isbell/Don Savage
Headquarters, Washington, DC
(Phone: 202/358-1547)

Nov. 10, 1999
Embargoed until 2 p.m. EST

RELEASE: 99-134

MARS CLIMATE ORBITER FAILURE BOARD RELEASES REPORT, NUMEROUS NASA ACTIONS UNDERWAY IN RESPONSE

Wide-ranging managerial and technical actions are underway at NASA's Jet Propulsion Laboratory, Pasadena, CA, in response to the loss of the Mars Climate Orbiter and the initial findings of the mission failure investigation board, whose first report was released today.

Focused on the upcoming landing of NASA's Mars Polar Lander, these actions include: a newly assigned senior management leader, freshly reviewed and augmented work plans, detailed fault tree analyses for pending mission events

Noviembre de 1999

<https://mars.jpl.nasa.gov/msp98/news/mco991110.html>

- Que existen partículas que viajan a mayor velocidad que la luz (lo primero que hacemos es sacar la fotografía de Einstein para resarcirnos: «ya, por fin, lo hemos pillado»). Posteriormente se descubrió que se habían producido errores en el proceso de medición.
- Que hemos invertido 125 millones de dólares en una sonda que acabamos de estrellar (Mars Climate Orbiter 1999). Podemos introducir esta noticia en el campo de la investigación porque detrás de cada misión especial lo que hay es un inmenso trabajo de preparación, una labor científica que continúa con los datos que posteriormente recogen las sondas.

Efectivamente, nos equivocamos. Y estos ejemplos son, tan solo, algunos acontecimientos que han saltado a la prensa. Otros, la mayoría, quedan en los debates sectoriales o tienen poca repercusión. Y otros, simplemente no llegan a conocerse. En ocasiones los errores son premeditados. Hay intereses económicos asociados. Algunos saltan a la prensa con mayor o menor celeridad.

AMERICA

50 Years Ago, Sugar Industry Quietly Paid Scientists To Point Blame At Fat

September 13, 2016 - 9:59 AM ET

CARILA DOMONOSKE



A newly discovered cache of internal documents reveals that the sugar industry downplayed the risks of sugar in the 1960s. *Luis Acaribe/Getty Images*

In the 1960s, the sugar industry funded research that downplayed the risks of sugar and highlighted the hazards of fat, according to a newly published article in *JAMA Internal Medicine*.

Septiembre de 2016

<http://www.npr.org/sections/thetwo-way/2016/09/13/493739074/50-years-ago-sugar-industry-quietly-paid-scientists-to-point-blame-at-fat>

EL DIÁ.es
 Canarias La Palma Tenerife Economía Nacional Sociedad Sucesos Deportes Cultura Gastronomía Criterios Más

SUCESOS

Detenidos por vender fármacos contra el cáncer sin homologar

0 veces compartido | Comentarios 0

Efe, Palma 12/abril/17 6:09 AM Etiqueta impresa

Dos catedráticos de la Universidad de las Islas Baleares y tres de sus colaboradores han sido detenidos y acusados de estafa por vender como remedio contra el cáncer un fármaco sin homologar por el que habrían ingresado 600.000 euros a través de una fundación sin ánimo de lucro.

La Policía Nacional informó ayer de que los dos profesores de Biología Celular, socios en la empresa Lipopharma, fueron arrestados la semana pasada, interrogados y puestos en libertad con caros a la espera de declarar ante el

Abril de 2017

<http://eldia.es/sucesos/2017-04-12/8-Detenidos-vender-farmacos-cancer-homologar.htm>

Llegados a este punto ¿basta con que las publicaciones científicas y nuestros repositorios institucionales recojan exclusivamente los resultados de la investigación?



Probablemente no. Ilustrémoslo con un ejemplo:

Las misiones espaciales de la NASA de los años 60 y 70 del siglo XX (que costaron cantidades astronómicas) obtuvieron cantidades ingentes de información y datos que se encontraban almacenados originalmente en soportes magnéticos, pero sin ninguna metodología documental, esto es, muchos de

estos soportes no estaban debidamente descritos; en ocasiones era o es prácticamente imposible determinar la forma en la que fueron estructurados (ausencia o pérdida de información de representación); y en otros casos el hardware necesario para reproducir los soportes no se ha conservado.

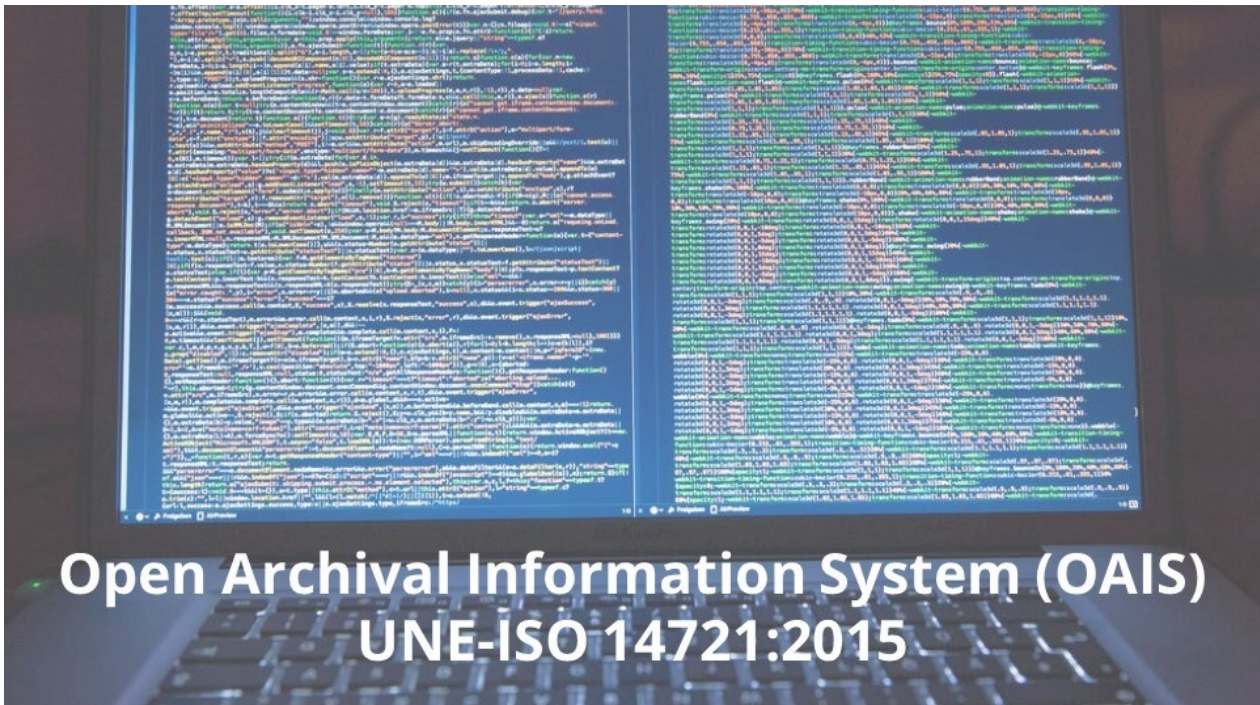
La recuperación de esta información con una ausencia total de medidas de preservación implica un esfuerzo prácticamente desproporcionado en nuestros días. Es decir que en muchos casos los datos están, pero es imposible recuperarlos. ¿Cuál es la solución? ¿Cómo se resuelve este problema?



Desde luego no podemos enviar una misión al espacio cada vez que nos encontramos una situación similar ni podemos confiar ciegamente y sin poner a prueba las investigaciones previas (una ciencia de ver y no tocar). Además, hemos visto que vamos a encontrar otras ocasiones (nos equivocamos) en las que vamos a necesitar acceder a los datos de soporte para verificar el camino utilizado para llegar a los resultados científicos. Pueden concurrir otras circunstancias: desaparece el grupo de investigación o se retoma una investigación realizada años atrás.



¿La solución? Esta problemática asociada a la preservación digital es la que inspiró la creación del estándar OAIS precisamente por la industria aeroespacial (Consultative Committee for Space Data Systems) y su transformación posterior en una norma ISO, cuya versión actual es UNE ISO 14721:2015.



OAIS es el principal estándar de preservación digital pero, ¿qué tiene que ver con la investigación científica. La preservación digital implica una serie de medidas que afectan a los gestores de la información: bibliotecarios, documentalistas, archiveros (quienes se encarguen de la conservación a

largo plazo). Pero que también afecta a los investigadores que deberán trabajar utilizando metodologías de gestión documental o similares. Es el corolario de lo que hemos avanzado ya:

- No basta con conservar los resultados de la investigación, es necesario conservar los datos que dan soporte a dichos resultados por si en el futuro próximo o lejano necesitamos los datos de soporte para realizar nuevas verificaciones.
- Que sean accesibles y comprensibles a largo plazo: tanto los resultados como los datos de soporte.
- Que se distinga entre lo que sea público y lo que no deba serlo.

Y ¿por qué los repositorios institucionales no son las mejores herramientas para alcanzar estos objetivos? Básicamente porque no están diseñados para aplicar un sistema basado en OAI. Expliquémoslo brevemente:

MÁS ARCHIVO INTEGRAL EN UNIVERSIDADES...

ODILO
23

El funcionamiento (simplificado) de las aplicaciones de gestión documental y archivo



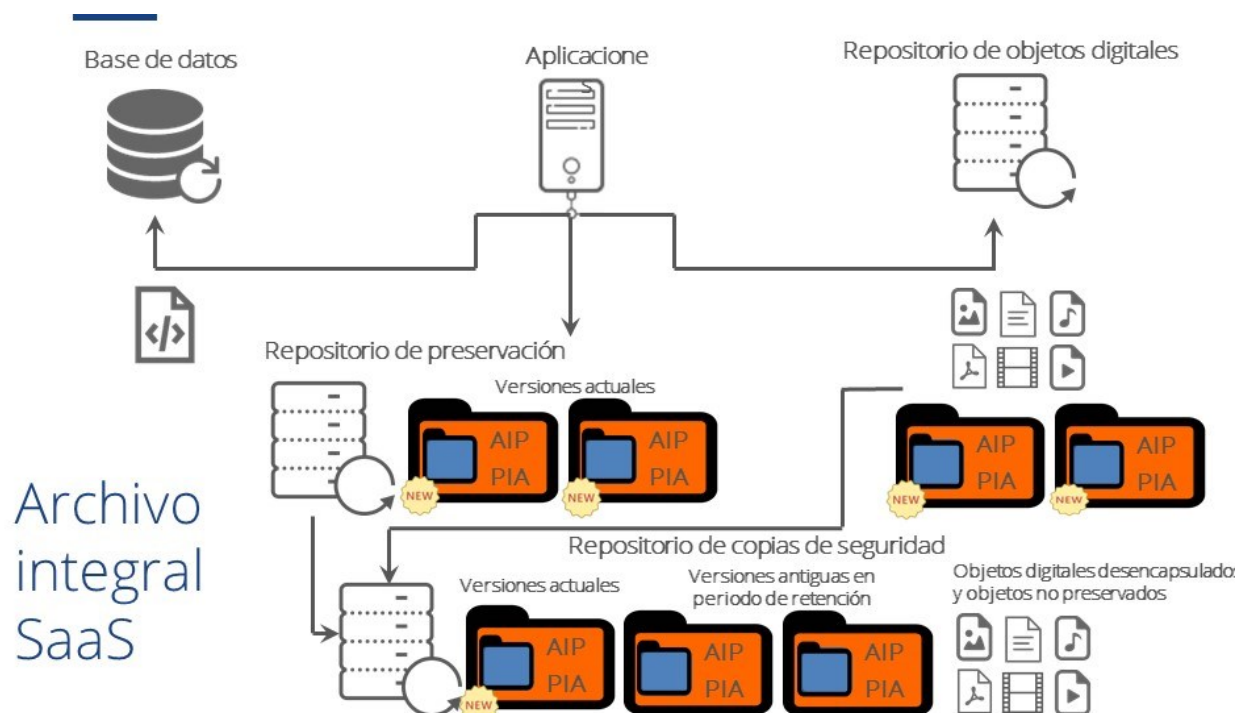
Si lo simplificamos de manera tosca, las aplicaciones de gestión de archivo (como A3W) se ejecutan en un servidor de aplicaciones y almacenan todos los datos necesarios para su ejecución, incluidas las descripciones de la documentación, perfiles de usuarios, movimientos de documentación... en una base de datos (que se encuentra en un servidor de base de datos). Los objetos digitales (documentos electrónicos) se encuentran referenciados en la base de datos y conservados en un repositorio de documentos. Es decir, los elementos que componen el sistema se encuentran en «lugares» (servidores) diferenciados, siendo la aplicación informática el único sistema capaz de interpretar correctamente los datos de la base de datos y las relaciones de la información y con los objetos digitales.

De un modelo de estas características se derivan ciertos riesgos cuando hablamos de conservación a largo plazo:

- Los objetos digitales y sus descripciones se encuentran conservados en ubicaciones diferentes. Normalmente, para asegurar la conservación se realizan copias individuales de la base de datos y de los objetos digitales para tener elementos de respaldo.
- La dependencia de la aplicación para interpretar datos y relaciones entre datos y objetos digitales es total. Si la aplicación deja de estar soportada o es sustituida por otra, es necesario plantear procesos de migración que, en algunos casos y debido a la complejidad de estos, no finalizan de forma completamente exitosa.
- El sistema no está diseñado para soportar eventos catastróficos o que pongan en peligro su integridad: si uno de los elementos que conforman el sistema se pierde total o parcialmente, los otros quedan afectados de forma irremisible.

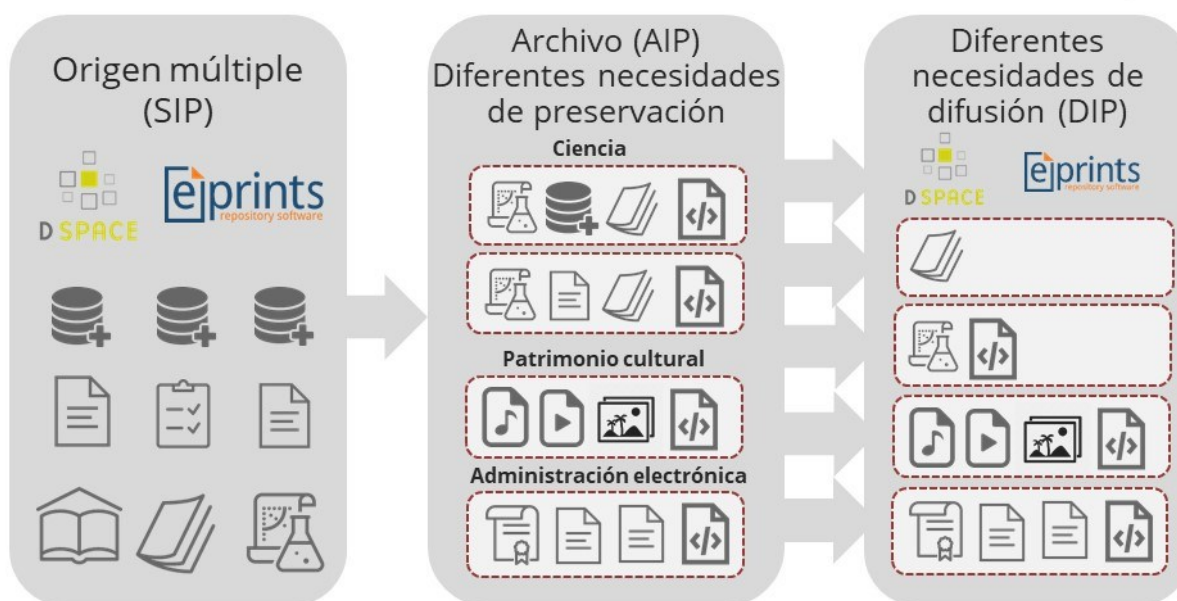
Los repositorios institucionales, en general y las herramientas de gestión científica (cuando existen) no tienen esta visión. Sin embargo, sí la tienen los archivos y por eso los archivos universitarios están llamados a asumir la responsabilidad de conservar la información generada en el contexto de la actividad científica.

Efectivamente, los archivos han comenzado a implementar sistemas OAIS en los que el encapsulado de la información en forma de paquetes de archivo contextualizados asegura la conservación a largo plazo. En la siguiente figura se muestra el ejemplo del modelo de archivo y preservación en la nube (SaaS) implantado por Odilo y sus múltiples instancias de almacenamiento:



Y cómo puede ayudar Odilo a poner en práctica un sistema de estas características?

Solución de Odilo



Odilo proporciona un sistema integral de archivo y preservación que da respuesta a las necesidades tradicionales relacionadas con el archivo, como las de gestión de la documentación administrativa electrónica y en papel, de la documentación patrimonial, etc. y que, al mismo tiempo, permite la gestión de documentación procedente de repositorios institucionales y de otras fuentes de producción científica.

Modelo integrado de archivo físico-digital, preservación, difusión y acceso



Este modelo de archivo integral compatibiliza:

- La capacidad de integración de las soluciones de Odilo con otras aplicaciones de gestión y tramitadores de expedientes.

- El potencial de la solución tecnológica de archivo (A3W-AE). Combina archivo físico, archivo electrónico (tramitación electrónica y contenido digital patrimonial) y difusión:
 - Archivo físico: con funcionalidades completas de gestión, mantenimiento de varios centros de archivo, transferencias, gestión de depósitos, control de las políticas de disposición, etc.
 - Archivo electrónico: adaptado a las exigencias de administración electrónica derivadas de las leyes 39/2015 y 40/2015, el Esquema Nacional de Interoperabilidad y las normas técnicas de interoperabilidad (archivo electrónico único). Se encuentra integrado con la gestión física. Permite la gestión y conservación de colecciones especiales (fotografía, vídeo, etc.).
 - Difusión. Pensado para la difusión de contenido digital patrimonial. Incluye un portal de archivos con el fin de difundir el contenido digital seleccionado o proporcionar un acceso controlado a la documentación. A3W-AE también cuenta con un repositorio OAI-PMH que permite la interconexión con Hispana y Europea.
- Preservación digital (Odilo Preserver) completamente integrada y 100% OAI certificada en el cumplimiento del estándar ISO 14721:



El servicio de preservación (Odilo Preserver) cuenta con cuatro elementos:

1. Odilo Preserver es asesoramiento. Desde el inicio del servicio, nuestro equipo de preservación asesora al archivo en la definición de su plan de preservación. En el ámbito del que venimos hablando, es fundamental que los científicos cambien su forma de trabajar, pero solo en parte, para hacerla más interoperable y compatible con los sistemas de gestión documental existentes.
2. Odilo Preserver es una plataforma tecnológica. Administra el ingreso de paquetes de información y los transforma en elementos de archivo conforme al modelo OAI. El sistema

- gestiona objetos digitales, metadatos, acciones de preservación y accesos asegurando la trazabilidad de los paquetes.
3. Odilo Preserver es almacenamiento. Ofrecemos diferentes modalidades de almacenamiento *cloud* y local.
 4. Odilo Preserver es servicio. El plan de preservación definido al inicio es trasladado a la plataforma de gestión y el personal especializado de Odilo se encarga de identificar riesgos y ejecutar las acciones para asegurar la preservación a largo plazo.

Por último, una de las características más importantes de este servicio es la capacidad de proporcionar un entorno certificado que se hereda por el hecho de usar el servicio:

LA SOLUCIÓN DE ODILO: ARCHIVO INTEGRAL



La visión de Odilo: archivo y preservación como servicio (SaaS)

Nosotros cumplimos con los estándares y exigencias legales y nos encargamos de todo: auditar nuestros sistemas, actualizar las certificaciones...

Nuestros clientes heredan este entorno al utilizar la solución de Odilo y no necesitan realizar certificaciones propias.



F. J. Valentín Ruiz
fjvalentin@odilotid.es

Madrid (ES)

Cartagena (ES)

Denver (US)

New York (US)

Mexico City (MX)

Bogotá (CO)

Buenos Aires (AR)