

PRESERVAR DIGITALMENTE AHORA: SOLUCIONES PRAGMÁTICAS PARA CONTENIDOS PATRIMONIALES

VÍCTOR MACÍAS ALEMÁN*

Fecha recepción: 20 de diciembre de 2021

Fecha de aceptación: 28 de diciembre de 2021

Resumen: El proceso para preservar los contenidos digitales, tanto el que ha sido escaneado desde un original físico como el nacido digitalmente, es un cometido que habitualmente se posterga. Usualmente se suele creer que, con la digitalización y difusión, lo más importante ya está hecho. La preservación digital, erróneamente entendida como un proceso posterior, queda por tanto para más adelante y, en la mayor parte de los casos, aplazada sin fecha. A ello se suma el no asumir su planificación, la complejidad de los conocimientos necesarios, la dificultad técnica de implementación, el alto coste de las soluciones comerciales, la granularidad, escasa o nula integración del software libre disponible y la ausencia de plataformas centralizadas, medios y soporte técnico por parte de la Administración pública para quienes desean llevarla a cabo. Sin embargo, lo primero es tomar conciencia del problema. Luego explicar que se puede asumir gradualmente el trabajo y que ha soluciones asequibles para ello.

Palabras claves: Preservación digital; planificación; implementación.

Abstract: The process to preserve digital content, both that which has been scanned from a physical original and that which was born digitally, is a task that is usually postponed. It is usually believed that, with digitization and dissemination, the most important thing is already done. Digital preservation, erroneously understood as a subsequent process, is therefore left to a later date and, in most cases, postponed without a date. Added to this is not assuming its planning, the complexity of the necessary knowledge, the technical difficulty of implementation, the high cost of commercial solutions, the granularity, little or no integration

* Universidad de Las Palmas de Gran Canaria, Biblioteca Universitaria. Sección de automatización y repositorios digitales. Correo electrónico: victor.macias@ulpgc.es.

of available free software and the absence of centralized platforms, media and technical support from the Public Administration for those who wish to carry it out. However, the first thing is to become aware of the problem. Then explain that you can gradually take on the job and that you have affordable solutions for it.

Key words: Preserve digital; planning; implementation.

I LA VITAL IMPORTANCIA DE LA PRESERVACIÓN DIGITAL PARA NUESTROS SERVICIOS EN LÍNEA, OBJETOS DIGITALES Y METADATOS

El propósito de este artículo va dirigido a exponer la vital importancia de lo que viene al final del proceso documental vinculado a la digitalización o a la generación, captura y gestión de los denominados *born-digital files*¹, la documentación nativa digital. Hablamos de la preservación digital, de las técnicas y procesos necesarios para que lo que hemos generado, o lo que hemos obtenido directamente en formato digital, no se pierda con el paso del tiempo y siga siendo plenamente accesible en el futuro.

Si nos preocupamos, centrándonos en el objeto de las jornadas en que nos encontramos, de digitalizar debidamente, con normalización y últimas técnicas, las imágenes que nos muestran lo que fue nuestro pasado y, acto seguido, las ofrecemos en línea para facilitar su mayor y mejor difusión, almacenándolas informáticamente de la manera que nos parece más adecuada, o quizá de la única forma que está a nuestro alcance, ahí suelen acabar nuestras preocupaciones. Tenemos muchos otros cometidos pendientes y, por tanto, damos carpetazo a ese trabajo, a ese fondo documental, y nos dedicamos al siguiente proyecto.

Pero no es así. Nuestro trabajo no ha terminado porque, ¿cómo sabemos con total seguridad que, dentro de un año, de cinco, de diez, de cincuenta, estos contenidos digitales que hemos generado seguirán siendo accesibles localmente o en «la nube»² y disponibles en línea con todas sus funcionalidades?

1. <https://www.nationalarchives.gov.uk/information-management/manage-information/digital-records-transfer/what-are-born-digital-records/>.

2. https://es.wikipedia.org/wiki/Almacenamiento_en_nube.

Cuando hablamos de preservación digital a largo plazo, queremos decir de forma indefinida en el tiempo. Por tanto, *para siempre*, si eso es posible.

El origen del problema a este respecto es que la mayoría de los proyectos de digitalización documental se conciben sin plantearse siquiera el archivo de los objetos obtenidos, y de sus metadatos vinculados, a largo plazo. No existe conciencia generalizada de que, en realidad, preservar las imágenes que hemos obtenido digitalizando un documento físico es más difícil y, usualmente, más costoso, que preservar el propio documento físico de donde se ha obtenido. A diferencia del original, los objetos digitales obtenidos son dinámicos y pueden reconvertirse en múltiples formatos, generando diversos servicios derivados. Por ello, si no asumimos esta cuestión vital, puede suceder que al final solo nos quede el documento físico de procedencia. Cabe entonces pensar que sucederá con los objetos nacidos digitalmente y que pueden desaparecer o quedar inaccesibles como si nunca hubieran existido, ya que ni siquiera cuentan con soporte material original al que podamos recurrir en última instancia.

Uno de las premisas más importantes es que los recursos tecnológicos: los informáticos, tanto en software como en hardware, pero también los aparatos reproductores de contenidos electrónicos, son sumamente inestables. Esto quiere decir que su periodo de vida es escaso, ya que mutan y evolucionan continuamente y, por tanto, quedan obsoletos en plazos más o menos breves.

A ello debemos añadir el peligro inherente de que los objetos digitales y sus metadatos asociados son mucho más modificables y manipulables que el original físico de donde proceden. Estas modificaciones no tienen por qué ser intencionadas, sino que también pueden producirse por virus informáticos, hongos, calor o humedad en el soporte físico que lo sustenta. Sobre estos ficheros digitales y en sus soportes electrónicos no podemos plantearnos su restauración tal y como la entendemos para un original analógico tradicional, pretendiendo recuperarlo de igual modo que estaba antes del daño que haya sufrido.

Otro planteamiento básico es que la preservación digital cambia según la entidad, el fondo digital generado y su propósito. No hay un planteamiento común para todos los casos. Por ello, diferirán los procedimientos a realizar por quien tiene un solo tipo de documentación digital, en un único repositorio y dirigido a un tipo de usuarios homogéneo con respecto a quienes generan objetos digitales de muy diverso género, gestionados en varios repositorios y dirigidos a propósitos y público variable. Además, y dentro de la temática de estas jornadas, no recibirá el mismo tratamiento, con vistas a su preservación, las digitalizaciones obtenidas desde diapositivas, negativos o papel fotográfico, que aquellos ficheros RAW generados por cámaras fotográficas digitales.

Adquirir o contratar el material necesario para el almacenamiento y, en su caso, puesta en línea de los objetos digitales, sus metadatos y estructura organizativa, no es todo, ni mucho menos, para conseguir la preservación digital. Incluso, a nivel humano, la solución completa va más allá del personal técnico que asume la gestión de estos contenidos.

2 ESTADO DE LA SITUACIÓN

En el año 2016, la empresa Arkivum³ realizó una encuesta entre profesionales de una amplia gama de galerías, museos, archivos, bibliotecas y otras instituciones dedicadas a preservar la memoria en todo el mundo. Cuando se les preguntó sobre la planificación de su estrategia de preservación digital, el 67 % respondió que todavía estaban en la fase de recopilación de información o en sus primarios comienzos. Solo el 30 % tenía un proceso implementado y realizaba activamente la preservación digital. En el caso de España, el sector más desarrollado en este entorno, compuesto por los repositorios institucionales de la red de bibliotecas universi-

3. <https://arkivum.com>.

tarias (REBIUN), manifestaba en 2018 a través de un informe⁴ que tan solo dos repositorios de un total de cincuenta y dos contaban con planes de preservación aprobados y en funcionamiento. Las técnicas de preservación digital se estaban aplicando de forma incipiente y parcial en su mayoría, se detectaban graves problemas en el número de copias de seguridad y su ubicación, escaseaban las comprobaciones de la integridad de ficheros y la capacidad para realizar auditorías en el caso de detectarse incidencias de seguridad y no se estaba aplicando el compromiso público que se reconoce en cuanto a la preservación de los materiales digitales almacenados.

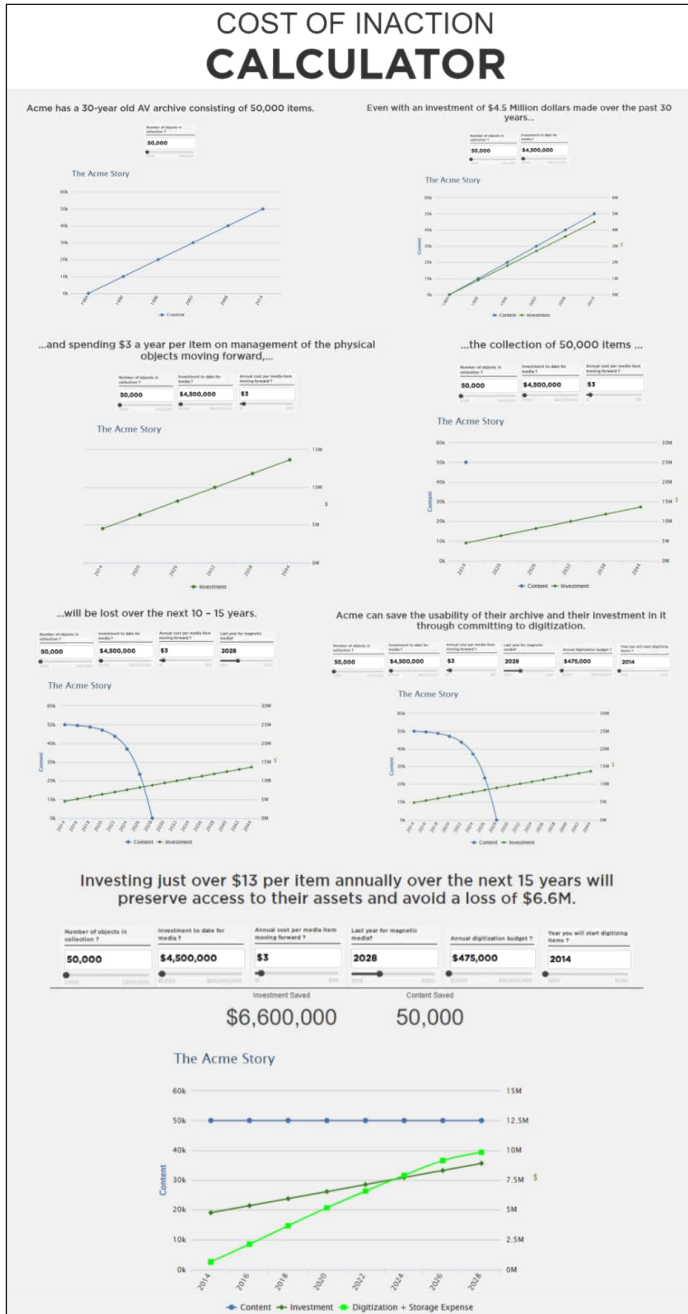
El punto de inicio indica que, al no actuar en el momento en que producimos o capturamos la documentación digital, multiplicamos el problema para el futuro, ya que hay que planificar las acciones necesarias, no cuando ha acabado dicho proceso, sino antes y durante la ejecución del mismo. La falsa percepción de que, al almacenar informáticamente los metadatos y objetos digitales generados, el problema ha quedado resuelto porque todo queda *preservado* por arte de birlibirloque, es generalizada.

Una de las más indeseadas consecuencias es que el no asumir las acciones necesarias para hacer posible la preservación digital puede tener un costo económico considerable que va incluso más allá de la posible pérdida del trabajo realizado o de los *born-digital files* capturados. Para estimarlo, existen herramientas como la calculadora del coste de inacción en materiales audiovisuales⁵ desarrollada por AVPreserve, y que ayuda a las entidades a analizar el riesgo, los diferentes niveles de aplicación deseados, el retorno de la inversión y las pérdidas o ganancias en función de los diferentes escenarios posibles.

En suma, tenemos que tomar conciencia no solo nosotros, sino también el personal de nuestra entidad, y sus órganos directivos (de quienes dependerá la financiación, establecimiento de la

4. Informe sobre la evaluación del estado de la preservación de los repositorios REBIUN 2018. <http://hdl.handle.net/20.500.11967/253>.

5. <https://coi.weareavp.com>.



Ejemplo utilizando la calculadora del coste de inacción desarrollada por AVPreserve.

estrategia y su normalización) para contar con la debida coordinación interna, e integrarnos en redes, consorcios y proyectos cooperativos que nos puedan facilitar toda la ayuda posible.

Llevar a cabo una preservación digital efectiva supone un proceso complejo, indefinido en el tiempo, y caro para el que, por lo general, no se suele contar con personal cualificado y escaso o ningún presupuesto. Además, en España tenemos una pobre cultura de cooperación institucional en estos entornos y tampoco contamos con el apoyo de planes, financiación o infraestructuras nacionales. En estos casos, donde quienes actúan suelen hacerlo individualmente, las entidades pequeñas o medianas no tienen posibilidad alguna para hacerse con la capacidad técnica, económica y organizativa para realizar estos imprescindibles cometidos de forma completa.

Con respecto a los costes, hemos de asumir que no son solo de de naturaleza tecnológica. Es difícil contestar a la pregunta: *¿y esto, cuánto vale?*, porque además de la tecnología hay que contemplar muchas variables, como la formación o la contratación de personal cualificado; la recomendable utilización de arquitecturas y sistemas distribuidos para los contenidos que nos pueden dar más seguridad a menor precio; escoger si optamos por almacenamiento local o en *la nube*; cuantificar los procesos de migración, conversión y verificación; el volumen de datos a tratar; las necesidades de transferencia masiva o el ancho de banda a contratar, además de decidir la frecuencia de las recuperaciones de datos que sean necesarias. Esto puede y debe asumirse por fases, con el fin de hacer menos onerosas, de entrada, las necesidades financieras para llevarlo a cabo y el volumen de trabajo que es necesario para ejecutarla.

3 CÓMO PROCEDER

Para empezar a trabajar, existen propuestas consorciadas⁶ que nos lo indican, como la realizada el pasado año por el Grupo de Tra-

6. Guía para la evaluación de los procesos de preservación en repositorios institucionales de investigación (REBIUN, 2020) <http://hdl.handle.net/20.500.11967/634>.

bajo de Repositorios de la Red de Bibliotecas Universitarias de España (REBIUN) y que propone las siguientes fases:

1. Planificación.
2. Plan de preservación.
3. Integridad de datos.
4. Formatos.
5. Metadatos.
6. Almacenamiento. Copias.
7. Inventarios.
8. Flujos de trabajo.
9. Documentación de procesos.
10. Gestión de riesgos.

Dentro del específico plan de preservación a desarrollar, se encuentran los siguientes apartados:

1. Alcance y propósito.
2. Objetivos.
3. Colecciones y usuarios.
4. Funciones y responsabilidades.
5. Compromisos y políticas institucionales.
6. Acciones de preservación y control de calidad.
7. Sostenibilidad financiera.
8. Sostenibilidad técnica.
9. Plan de contingencia y análisis de riesgos.
10. Formación.
11. Plan de comunicación.
12. Evaluación, seguimiento y revisión del propio plan.

Estructurarlo de esta forma nos ayudará a asumir nuestras posibilidades reales, a contestar a preguntas como: *¿se debe preservar todo o según que formatos?, ¿quién va a usar que contenidos?, ¿cuándo?, ¿cómo?, ¿para qué?, ¿con que frecuencia?, ¿borraremos los ficheros originales una vez migrados a formatos actualizados?*

entre otras, fundamentales para realizar la necesaria autoevaluación y diagnóstico, y acometer el proceso, idóneamente según los niveles de preservación digital de la National Digital Stewardship Alliance (NDSA)⁷.

Hemos de asumir que nadie puede preservarlo todo de forma indefinida, por lo que hay que escoger, desechar y priorizar. Por ejemplo, los editores de periódicos conservan usualmente solo las versiones electrónicas facsímiles, en la mayoría de los casos en formato PDF vectorial, aunque se trate de medios que también se editen en papel. La maquetación en ordenador del original, la versión web y la propia versión impresa suelen desecharse pasado un tiempo. En términos generales, no solo en este sector, el crecimiento imparable de los datos y de los objetos digitales obliga a ello⁸.

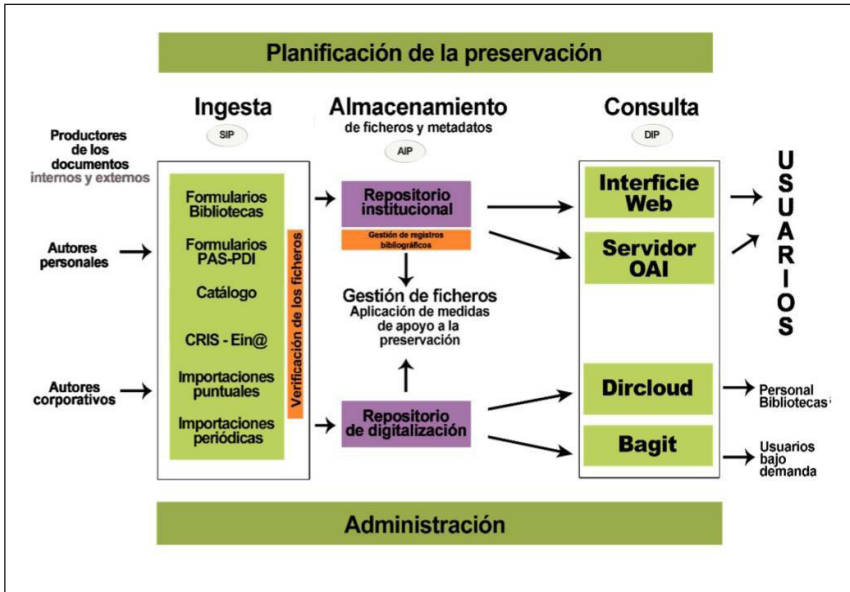
4 INTEROPERABILIDAD DE SISTEMAS

Otra cuestión que no puede obviarse a efectos de preservación digital, es la necesaria vinculación e interoperabilidad entre los diversos sistemas informáticos y de gestión utilizados en nuestra entidad. Si hablamos de una biblioteca o de un archivo, es habitual contar con:

- Catálogo informatizado.
- Repositorio en donde se gestionan, buscan y difunden los resultados de las digitalizaciones.
- Sistema de almacenamiento de datos.
- Sistema de preservación digital local o remoto (una vez establecido).

7. <https://es.surveymonkey.com/r/P9N7FLZ>.

8. Volumen mundial de datos <https://www.dealsunny.com/mobile-usage-statistics/>.



Aplicación del modelo OAIS al sistema de preservación digital del repositorio DDD de la Universidad autónoma de Barcelona (UAB).

que deben operar de la forma más integrada posible a efectos de facilitar este cometido. Nuestros trabajos se simplificarán si utilizamos repositorios basados en web semántica con empleo de datos enlazados y abiertos (LOD) y vinculados a la *nube*, o dificultado si nuestras instalaciones utilizan programas comerciales cerrados o basados en desarrollos informáticos hechos *ad-hoc* que, por lo general, no cumplen con el status de repositorios digitales confiables (ISO 16363:2012)⁹. También el uso de ontologías, esquemas de metadatos específicos, protocolos abiertos y visores interoperables de objetos¹⁰ serán positivos para conseguir nuestro propósito.

Este entorno tecnológico ha de asumir, en su base, lo indicado por el Esquema nacional de seguridad (ENS)¹¹ para el caso de Es-

9. <https://www.iso.org/standard/56510.html>.

10. <https://iiif.io>.

11. https://administracionelectronica.gob.es/pae_Home/pae_Estrategias/pae_Seguridad_Inicio/pae_Esquema_Nacional_de_Seguridad.html#.YVoaJZrMIuE.

paña, con la imprescindible aplicación del modelo OAIS (UNE-ISO 14721:2015)¹² en el sistema de preservación digital escogido y que utilizará la gestión mediante paquetes SIP/AIP/DIP (submission / archival / dissemination information package).

5 BUENAS PRÁCTICAS

Con respecto a los procedimientos correctos a llevar a cabo, en la praxis necesaria para articular la consecución del objetivo perseguido y una vez establecidas las bases tecnológicas adecuadas, podemos señalar la aplicación de los ya mencionados niveles de preservación digital de la NDSA¹³, en forma inevitablemente progresiva; la utilización de formatos estándar estables, abiertos y de uso generalizado¹⁴; aplicar una sintaxis normalizada a los ficheros y directorios generados¹⁵; el uso de identificadores persistentes para la localización continua de los recursos digitales en línea, como PURL¹⁶ o Handle¹⁷; empleo de modelos de empaquetado de ingesta de objetos y metadatos SIP como E-ARK¹⁸ o Bag-it¹⁹ y, por supuesto, la adopción de esquemas de metadatos acordes con la información y los objetos digitales a preservar, entre los que

12. <https://www.iasa-web.org/tc04-es/618-sistema-abierto-de-archivo-de-informaci%C3%B3n-oais>.

13. NDSA Levels of digital preservation 2.0 (26 septiembre 2019). <http://www.apredig.org/npdnds2019/>.

14. Library of Congress recommended formats statement, 2020-2021. <https://www.loc.gov/preservation/resources/rfs/TOC.html>.

15. *Recomendaciones para proyectos de digitalización de patrimonio bibliográfico y fotografía histórica*. Madrid: Ministerio de cultura y deportes, Consejo de Cooperación Bibliotecaria, Grupo de Trabajo de Buenas Prácticas en Digitalización, 2021. Véase: Estructura de contenidos y nombres de archivo, pp. 148-155 <http://hdl.handle.net/10421/9200>.

16. <http://bibpurl.oclc.org>.

17. <https://handle.net>.

18. <https://www.eark-project.com/resources/eark-tools.html>.

19. <https://en.wikipedia.org/wiki/BagIt>.

NDSA%		Niveles de Preservación Digital			
		Nivel			
Área Funcional	Nivel 1 - (Conocer su contenido)	Nivel 2 - (Proteger su contenido)	Nivel 3 - (Controlar su contenido)	Nivel 4 - (Mantener su contenido)	
Almacenamiento	Tener tres copias completas en ubicaciones separadas Documentar todos los medios de almacenamiento donde este almacenado el contenido Proteger el contenido en soportes de almacenamiento estables	Tener tres copias completas con al menos una copia en una ubicación geográfica distinta Documentar el almacenamiento y los recursos y las dependencias que estos requieren para funcionar	Tener al menos una copia en una ubicación geográfica distinta Tener al menos una copia en un medio de almacenamiento de diferente tipo Rastrear la obsolescencia del almacenamiento y los medios	Tener al menos tres copias en ubicaciones geográficas distintas, cada una con una amenaza de desastre diferente Maximizar la diversificación del almacenamiento para evitar puntos únicos de falla Tener un plan y realizar acciones para abordar la obsolescencia del almacenamiento y medios de almacenamiento	
Integridad	Verificar que la información de integridad se ha proporcionado con el contenido Generar información de integridad si el contenido ha sido proporcionada con el contenido Se verifica virus en todo el contenido, se aislado en cuarentena según sea necesario	Verificar la información de integridad al mover o copiar contenido Usar bloqueadores de escritura cuando se trabaja con medios originales Hacer una copia de seguridad de la información de integridad y almacenar en una ubicación separada del contenido	Verificar la información de integridad del contenido en intervalos fijos Documentar los procesos y resultados de verificación de información de integridad Realizar una auditoría de la información de integridad bajo demanda	Verificar la información de integridad en respuesta a eventos o actividades específicas Reemplazar o reparar el contenido dañado según sea necesario	
Control	Se determinan los agentes humanos y de software que deben estar autorizados para escribir, mover y eliminar contenido	Documentar a los agentes humanos y de software autorizados para leer, escribir y eliminar contenido y aplicar estos	Mantener los registros (logs) y se identifican a los agentes humanos y de software para realizar acciones sobre el contenido.	Se realizan revisiones periódicas de acciones / registros (logs) de acceso	
Metadatos	Crear un inventario de contenido, que incluya la descripción de almacenamiento actual de estos Hacer una copia de respaldo del inventario al menos una copia por separado	Almacenar suficientes metadatos para respaldar la información de aspectos administrativos, técnicos, descriptivos, de preservación y estructurales)	Determinar qué estándares de metadatos aplicar Encuentrar y completa los vacíos en sus metadatos para cumplir con esos estándares	Registrar las acciones de metadatos sobre el contenido y cuándo ocurren esas acciones implementa los estándares de metadatos elegidos	
Contenido	Documentar los formatos de archivo y otras características de contenido. Verificar el formato cómo y cuándo fueron identificados	Verificar los formatos de archivo y otras características de contenido esenciales Establecer relaciones con los creadores de contenido para fomentar la elección sostenible de archivos	Monitorear la obsolescencia y los cambios en las tecnologías de las que depende el contenido	Realizar migraciones, normalizaciones, emulación y otros procedimientos que garanticen el acceso al contenido	

Levels of Digital Preservation v2.0 - 26 Septiembre 2019.
Copyright © 2019 by National Digital Stewardship Alliance. This work is licensed under a Creative Commons Attribution 3.0 Unported License. | DOI 10.17865/OSF.IO/GZ398

Niveles de preservación digital de la NDSA versión 2.0.

Hull University Archives – Expanded NDSA Digital Preservation Levels List – 21 February 2017

	Level 1 - Protect Your Data	Level 2 - Know Your Data	Level 3 - Monitor Your Data	Level 4 - Repair Your Data
Storage & Geographic Location	Two complete copies that are not collocated For data on heterogeneous media (optical discs, hard drives etc.) get the content off the medium and into your storage system	Store a backup outside premises At least one copy in a different geographic location Document your storage system(s) and storage media and what you need to use them	At least one of the copies is available with a different storage media	At least one copy is available on a different storage media Have a complete copy on a different storage media for geographic redundancy Consider backup of metadata to maintain all geographic redundancy
File Fidelity & Data Integrity	Check file/folder integrity if it has been processed with its content Create fidelity info if it wasn't provided with the content	Check hashes of all files Use write protectors when working with original media Write-check all content, especially high-risk content	Check hashes of content. If there is a problem, verify or correct corrupt data Write-check all content Write-check high-risk content (high-risk content is content)	Check hashes of all metadata to maintain all geographic redundancy All files to include provenance metadata Check for any errors that occur during the backup process
Information Security	Identify who has read, write, move and delete authorization to approved files Restrict who has those authorizations to individual files	Document access restrictions for content	Monitor for any changes to access restrictions	Set access restrictions to individual metadata Monitor for unauthorized access to metadata Monitor for unauthorized access to metadata
Metadata	Inventory of content and its storage location Ensure backup and non-collocation of inventory	Have a plan to recover metadata Store a backup of metadata and file system	Monitor the backup of metadata	Have a plan to recover metadata
File Formats	When you can give up the location of digital files, archive the use of a limited set of known open format and sources	Inventory of file formats in use	Monitor the format of digital content	Monitor for digital content that is not in an open format Monitor for digital content that is not in an open format
Physical Media	Inventory all physical media in use including hard discs	Evaluate available physical media and recommend suitable archival media Begin to migrate information away from any physical media that is over ten years in age	Plan to replace media if dependent on a format that is over ten years in age Plan to migrate information away from any physical media that is over ten years in age	Plan to replace media if dependent on a format that is over ten years in age Plan to migrate information away from any physical media that is over ten years in age
Access	Dedicated designated community Ability to ensure the security of the material while it is being accessed, including physical security measures Ability to identify and protect personally identifiable information (PII) and other sensitive material	Have publicly available cataloging, finding aids, inventories, or collection descriptions available Create redundant information including PII and personally identifiable information (PII) and other sensitive material	Plan to ensure digital content is available by other means if the original is not available	Plan to ensure digital content is available by other means if the original is not available
PASS	PASS	PARTIAL	PARTIAL	PARTIAL

Ejemplo de aplicación de los niveles de preservación digital NDSA en la Hull University Archives en 2017.

inevitablemente debe figurar METS²⁰ y PREMIS²¹ y que, en su caso, podrán estar incorporados a los propios objetos y generados por los procesos de preservación.

En verde, lo realizado; en amarillo, lo que está en curso y, en rojo, lo pendiente de acometer.

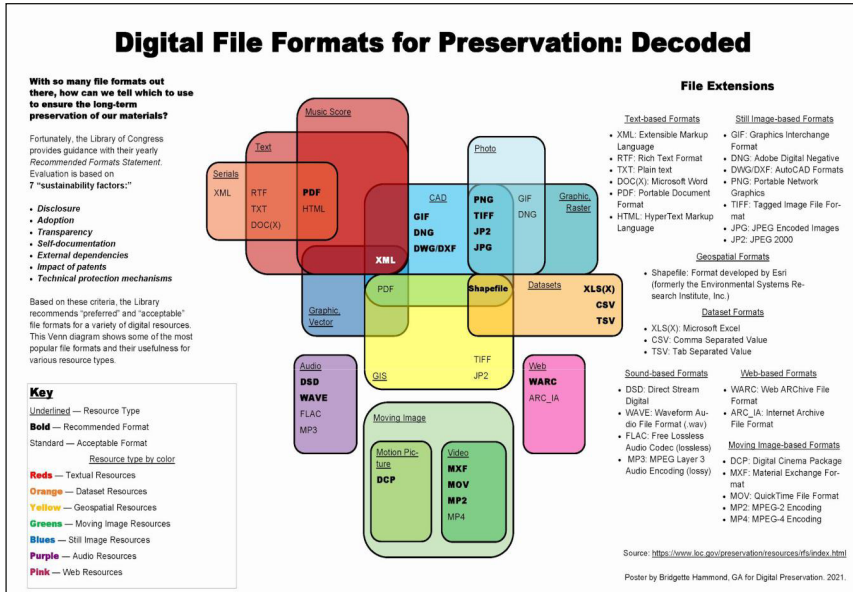
6 FORMATOS

La elección de los formatos a utilizar en los objetos digitales que generamos y gestionamos, al objeto de su preservación, es especialmente delicada. Los formatos tienen un bajo riesgo de obsolescencia cuando:

- Son interoperables, muy utilizados y compatibles en todo el mundo.
- Libres de patentes, preferiblemente de código abierto.
- Bien documentados, identificables y con especificaciones disponibles públicamente, suficientemente detalladas para crear software.
- Estables, lo que comporta que sus nuevas versiones se lancen raramente. Deben ser compatibles con versiones anteriores y posteriores o tener una ruta de migración clara.
- Independientes de plataformas específicas y compatibles con una amplia gama de software o independiente del mismo.
- No usan compresión o la utilizan sin pérdida de datos.
- Soporte técnico disponible fácilmente por proveedores, comunidad internacional o terceros.
- Admiten la incorporación de metadatos.

20. Metadata Encoding and Transmission Standard (METS) https://www.loc.gov/standards/mets/METSOverview_spa.html.

21. Preservation Metadata Implementation Strategies (PREMIS) [https://es.wikipedia.org/wiki/PREMIS_\(PREservation_Metadata_Implementation_Strategies\)](https://es.wikipedia.org/wiki/PREMIS_(PREservation_Metadata_Implementation_Strategies)).



Formatos de objetos digitales y su relación con la preservación digital.

En el entorno que nos ocupa dentro de estas jornadas, hay que hacer hincapié en los formatos utilizados por la generación de fotografía digital original, en la que es usual que cada marca de cámaras profesionales utilice una variante del reconocido formato RAW²². Así, la marca Sony produce ficheros ARW, en las Canon son CR2, Nikon emplea el NEF y Olympus el ORF, por citar algunas. En este caso, es recomendable su conversión a formato abierto DNG²³ (Digital negative) producido en su momento y posteriormente liberado por Adobe. La Library of Congress recomienda convertir los RAW al estándar universal TIFF²⁴.

Como ejemplo la adopción de formatos, la Biblioteca de la Universidad de Las Palmas de Gran Canaria ha adoptado los siguientes basados en los soportes digitalizados y los servicios en línea que presta:

22. [https://es.wikipedia.org/wiki/Raw_\(formato\)](https://es.wikipedia.org/wiki/Raw_(formato)).
 23. <https://es.wikipedia.org/wiki/DNG>.
 24. <https://www.loc.gov/preservation/resources/rfs/stillimg.html#photodigital>.

- RAW.
- TIFF.
- JPEG.
- PDF.
- PDF/A (ISO 19005).
- PDF 1.4 (PDF/A-1).
- PDF 1.7 (PDF/A-2).
- PDF 1.7 (PDF/A-3).
- PDF 2.0 (PDF/A-4).
- TXT.
- EPUB.
- MUSICXML.
- WAV.
- MP4 (H.264/MPEG-4 AVC).
- BWF (broadcast wave format).
- MXF (material exchange format).

Hay que hacer especial mención al formato PDF/A, un estándar de preservación²⁵ gestionado por diverso software y servicios en línea²⁶ que permite que los contenidos en él recogidos sean plenamente accesibles de forma local o remota. Es capaz de incluir texto, imágenes o gráficos, su impresión precisa, la integración de fuentes textuales, elimina lo no preservable y, a partir de la versión 1.7 (A-2), permite encapsular múltiples ficheros en un mismo ítem.

7 OBSOLESCENCIA DE LOS SOPORTES

También hay que considerar la problemática inherente a los soportes de la documentación física a gestionar, donde sus di-

25. PDF/A-1 for long-term preservation (ISO 19005-1:2005) <https://www.loc.gov/preservation/digital/formats/fdd/fdd000125.shtml>.

26. Adobe Acrobat <https://www.adobe.com/es/acrobat.html>; AvePDF <https://avepdf.com/es/convert-to-pdf>; iLovePDF <https://www.ilovepdf.com/es/convert-pdf-a-pdf>; PDF en <https://www.pdfen.com/convert-to-pdf/bulk-conversion-to-pdf>.

ferentes tipos (por ejemplo, bobinas magnetofónicas, cintas de casete, cintas de vídeo, discos de vinilo, bobinas de celuloide, placas de cristal, etc.) comportan numerosos problemas con el paso del tiempo desde el que fueron normalmente utilizados y que, ahora, es difícil o casi imposible acceder a la información que contienen. Entre estas dificultades nos encontramos con: el volumen de los soportes a procesar; su estado de conservación; la dependencia de los mismos respecto a unos aparatos específicos y en muchos casos obsoletos, fuera de fabricación y mantenimiento para poder visualizarlos, escucharlos y reproducirlos; la migración desde estos soportes a otros iguales, similares (o por completo distintos) que conserven hasta cierto punto sus propiedades, contenidos y estructura; la pérdida de información, datos o funcionalidades al realizarla; el coste y tiempo necesario para ello e, incluso, la emulación informática, cuando no es posible, conveniente o rentable el traspaso a otros soportes.

En estos casos, el contar con los denominados laboratorios de «análisis forense»²⁷ es vital para conseguir el acceso a soportes analógicos y digitales y recuperar sus contenidos, interpretarlos, documentarlos y transferirlos, cuando proceda, a medios actuales.

8 LA IMPRESCINDIBLE COOPERACIÓN

Visto de forma somera el panorama, la única vía de éxito, dentro de parámetros y costes razonables, pasa por cooperar. Si nuestra entidad actúa por libre, comportará que la posible solución adoptada sea más cara, difícil de llevar a cabo, contenga mayor número de errores, tenga menor impacto y sea más modesta en el cumplimiento de los objetivos deseados. Las alianzas con entidades con

27. <https://www.dpconline.org/handbook/technical-solutions-and-tools/digital-forensics>; <https://library.stanford.edu/research/digitization-services/labs/born-digital-preservation-lab>.

necesidades similares, el participar en asociaciones y consorcios, siempre que sea posible, es fundamental. Debemos citar, entre otras y el ámbito supranacional, a la Digital Preservation Coalition (DPC)²⁸, el International Internet Preservation Consortium²⁹ o la Asociación Iberoamericana de Preservación Digital (APRENDIG).³⁰ En el entorno nacional, destacan casos como la Asociación para la Preservación del Patrimonio Digital de Chile³¹ que, aunque trata un ámbito más general, se ocupa también de la preservación digital.

Sin embargo, en términos genéricos, prima en nuestra cultura la individualidad, a veces incluso entre organismos dependientes de una misma entidad, lo que hace que las amenazas antedichas fatalmente se cumplan. Además, en el caso de España, no contamos con la asistencia de organismos nacionales de soporte y referencia que ayuden a las entidades a llevar a cabo tan compleja y necesaria tarea. Recordemos propuestas como la del Centro Nacional de Preservación Digital de Archivos que, años después de haber sido planteada, continúa a día de hoy sin concretar³² aun cuando, de existir, solo constituiría un soporte parcial por ocuparse de un ámbito documental sectorial. Sin embargo, sí la hay en comunidades autónomas que han entendido la importancia de esta cuestión y han desarrollado o auspiciado soluciones. En este caso se encuentra COFRE (Conservem per al Futur Recursos Electrònics)³³ de manos de la Biblioteca de Cataluña; *iArxiu*, el archivo seguro de las administraciones públicas catalanas³⁴ y *Padicat*, el archivo de toda la web en lengua catalana³⁵. También el gobierno autóno-

28. <https://www.dpconline.org>.

29. <https://netpreserve.org>.

30. <http://www.apredig.org>.

31. <https://www.aspat.cl>.

32. https://www.congreso.es/docu/pge2021/pge_2021-web/PGE-ROM/doc/1/3/23/3/2/7/N_21_A_R_31_124_1_2_3_1_332A_C_1.PDF.

33. <https://www.bnc.cat/El-Blog-de-la-BC/COFRE-una-nova-eina-per-difondre-fons-digitalitzats>.

34. <http://www.iarxiu.eocat.cat>.

35. <https://www.padicat.cat/es>.

mo vasco ha auspiciado un servicio mixto, público y privado, denominado *Metaposta*³⁶. Dirigido a la sociedad civil, ofrece soporte a particulares y empresas para que puedan preservar su valiosa documentación electrónica y favorecer así oportunidades empresariales y de negocio.

9 ANALIZAR LAS POSIBILIDADES

A la hora de escoger la solución más adecuada para nuestras necesidades, hemos de estudiar pormenorizadamente la oferta existente, idóneamente asesorados, asociados y consorciados.

Los parámetros a evaluar en las soluciones a nuestro alcance deben ser, al menos, los siguientes:

- Funcionalidad.
- Fiabilidad.
- Usabilidad.
- Eficiencia.
- Necesidades de mantenimiento.
- Portabilidad.
- Compatibilidad e integración.
- Computación y almacenamiento local y/o remoto (nube).
- Seguridad.
- Distribución.
- Coste.
- Utilización de sandbox o demostraciones.
- Instalación del sistema, acceso y administración.
- Revisión de funcionalidades durante pruebas de carga, salida y gestión.
- Consultar a otros clientes-usuarios-administradores que la utilicen.

36. <https://www.metaposta.com>.

La oferta se agrupa de la siguiente forma:

- Herramientas para procesos específicos (véase listado en anexo).
- Adaptación de sistemas no específicos para esta finalidad, como los CMS (Content Management Systems).
- Sistemas de Preservación Digital SPD (Sistemas de Preservación Digital). Se trata de plataformas integrales (véase listado en anexo) que incorporan vigilancia tecnológica con alerta de obsolescencia de formatos, previsión de migración, consulta, emulación, sumas de verificación (checksum), comprobación de no alteración de datos o antivirus entre otras, incluyendo la automatización de las operaciones. En estas, como en la elección de cualquier otro software, es fundamental el estudio de la entidad que lo mantiene, sea pública o privada, y la solidez de la comunidad que lo utiliza.
- Plataformas cooperativas internacionales como LOCKSS (Lots of Copies Keep Stuff Safe)³⁷; MetaArchive³⁸, basada en la anterior y sin afán de lucro, lo que no quiere decir gratuita; HathiTrust³⁹ u otras con propósitos sectoriales, como CLOCKSS (Controlled LOCKSS)⁴⁰ para publicaciones y que actúa de común acuerdo con los editores; PKP Preservation Network⁴¹ para salvaguardar las revistas electrónicas hechas con la plataforma libre OJS; o Portico⁴² en el ámbito académico.
- Computación y almacenamiento en la nube (véase listado en anexo), opción creciente que supone, en este cometido, su utilización de forma combinada con las opciones mencionadas para facilitar el cumplimiento de los niveles de preservación pretendidos.

37. <https://www.lockss.org>.

38. <https://metaarchive.org/join-us/>.

39. <https://www.hathitrust.org>.

40. <https://clockss.org>.

41. <https://pkp.sfu.ca/pkp-pn/>.

42. <https://www.portico.org>.

10 CONCLUSIONES

Para solucionar un problema, primero hemos de reconocer que lo tenemos. En este caso, resolver la preservación digital de forma indefinida de los contenidos generados en nuestra entidad u organización: objetos digitales y metadatos, pero también su estructura, gestión y servicios en línea.

Hemos de buscar apoyo, consorciarnos con quienes tengan similares necesidades y cometidos, asociarnos a organismos que nos puedan guiar en este propósito. Esto también favorece la búsqueda de financiación.

Tenemos que formarnos en este ámbito o buscar, pedir ayuda y contratar a quienes tengan experiencia en este sector, cuando con nuestros propios medios humanos no sea posible o recomendable. Hay que optar por lo que sea más procedente y eficaz.

Hay que redactar y aprobar una política de preservación en nuestra entidad y, acto seguido, generar un plan de preservación que ejecute lo dispuesto en aquella. Para que esto pueda hacerse es necesario involucrar a todo el personal de la organización que tenga participación, aún de forma indirecta, con la digitalización documental y los servicios digitales. De esta forma, no solo conseguiremos la aceptación del objetivo perseguido en toda la organización, sino además, con la alta dirección, de quien depende la toma de decisiones estructurales, organizativas y de financiación. Constituir un grupo de trabajo es siempre la opción más recomendable, ya que será quien se encargará de tomar las decisiones, planificar, ejecutar y realizar la vigilancia tecnológica que permitirá la necesaria actualización de las acciones a largo del tiempo.

A la hora de ejecutar nuevos trabajos de digitalización, o bien para continuar con los existentes, debe incluirse en su planificación los conceptos, parámetros, recomendaciones y normalización que hayamos establecido para la preservación digital real de los contenidos generados.

Evalúe con detalle las opciones a su alcance para implementar la preservación digital en su organización. Asumirla por fases, más

numerosas cuanto más escasos sean nuestros recursos humanos, materiales o financieros, le ayudará a avanzar, aunque a pasos cortos. Aplicar en ese proceso los diversos niveles de preservación, gradualmente pero con un cronograma para evitar que el propósito se diluya en el tiempo *sine die*, conllevará que más adelante se sorprenda por lo conseguido en lo que quizá veía imposible llevar a cabo. Y esto, por modestos que sean los resultados, le dará ánimo para continuar esta labor con el fin de que sus valiosos trabajos digitales no desaparezcan en el tiempo. Así, los servicios que oferte podrán seguir siendo utilizados por aquellos que los necesitan y por otros que se sorprenderán al conocerlos.

Recuerde que la peor opción es, siempre, no hacer nada.

REFERENCIAS BIBLIOGRÁFICAS

- CASELLAS, L., BOADAS, J. (dir.). *¿Qué recordarán de nosotros? Cómo preservar documentos personales y familiares en el siglo XXI*. Gerona: Archivo Municipal, 2013. https://www.girona.cat/web/sgdap/docs/dia_2013_def.esp.pdf.
- CRITERIOS *básicos para valorar sistemas de preservación digital*. Área de tecnología del Grupo de preservación digital. Ciudad de México: Universidad Nacional Autónoma de México, Instituto de Investigaciones Bibliográficas, 2020. Serie Instrumenta Bibliographica, 1. <https://www.iib.unam.mx/files/iib/libros-electronicos/Criterios-Basicos-Sistemas-Preservacion-Digital.pdf>.
- CRUZ MUNDET, J. R., DÍEZ CARRERA, C. «El cálculo de costes de la preservación digital: un análisis de modelos». *Anales de documentación*, v. 18, n. 2. (2015). <http://dx.doi.org/10.6018/analesdoc.18.2.228411>.
- GIL GARCÍA, Pilar, VALENTIN RUIZ, Francisco José. «¿Es la preservación digital, estúpido!» *Ruiderae: revista de unidades de información*, n. 16 (2020). <https://revista.uclm.es/index.php/ruiderae/article/view/2542>.
- GIUSTI, Marisa Raquel de. «Las dificultades de la preservación digital: problemas, desafíos y propuestas para los repositorios». En: *VI Conferencia Internacional BIREDIAL-ISTEC (San Luis Potosí, México, 17 al 19 de octubre de 2016)*. <http://sedici.unlp.edu.ar/handle/10915/57424>.

- GUÍA para la evaluación de los procesos de preservación en repositorios institucionales de investigación. REBIUN, Línea 3. Grupo de repositorios, 2020. Colección Estudios e informes. <http://hdl.handle.net/20.500.11967/634>.
- LEIJA ROMÁN, David Alonso, TÉRMENS GRAELLS, Miquel. «Preservación digital distribuida: lecciones de experiencias internacionales». *BiD: textos universitaris de biblioteconomia i documentació*, n. 39 (Diciembre 2017). <https://bid.ub.edu/es/39/leija.htm>.
- MARCOS RECIO, Juan Carlos, SÁNCHEZ VIGIL, Juan Miguel, OLIVERA ZALDUA, María. «La conservación de todos los contenidos digitales no es necesaria: guardemos solo lo imprescindible de los medios». *Ibersid: revista de sistemas de información y documentación*, v. 13, n. 2 (2019), pp. 31-38. <https://dialnet.unirioja.es/servlet/articulo?codigo=7112011>.
- OCHOA GUTIÉRREZ, J., SÁENZ GIRALDO, A., TIRADO TAMAYO, T. «Experiencias de gestión de los procesos de preservación digital a partir del modelo OASIS en repositorios institucionales». *Anales de documentación*, v. 24, n. 1 (2021). <http://dx.doi.org/10.6018/analesdoc.428141>.
- RECOMENDAÇÕES para a produção de planos de preservação digital. 2da. verçao. Portugal: Direçao-geral do livro, dos arquivos e das bibliotecas, 2020. 45 p. https://arquivos.dglab.gov.pt/wp-content/uploads/sites/16/2019/08/Recomendacoes_PPD_v2.pdf.
- ROSENTHALER, Lukas. «Digitization and long-term archiving of photographic material». En: *Imatge i recerca: 11 es Jornades Antoni Vares*. Gerona: Ayuntamiento, Servei de Gestió Documental, Arxius i Publicacions, 2010. https://www.girona.cat/sgdap/docs/6jwps3frosenthaler_english.pdf.
- TÉRMENS, Miquel. *Preservación digital*. Barcelona: Universitat Oberta de Catalunya, 2013.
- TÉRMENS, Miquel, LEIJA, David. «Auditoría de preservación digital con NDSA levels». *El profesional de la información*, v. 26, n. 3 (2017), pp. 447-456. <https://doi.org/10.3145/epi.2017.may.11> <https://fima.ub.edu/pub/termens/docs/EPI-v26n3.pdf>.

ANEXO

POLÍTICAS Y PLANES DE PRESERVACIÓN

- Plan de preservación digital de la Biblioteca Universitaria de Granada. 2018 <http://hdl.handle.net/10481/51469>.
- Protocolo de preservación digital para el sistema de preservación digital del Archivo Español de Media Art (AEMA) 2018. <https://bit.ly/31j7bKX>.
- Política de preservació del Dipòsit Digital de Documents de la UAB (DDD). 2018 <https://ddd.uab.cat/record/189808>.
- Plan de preservación del Repositorio Brújula de la Universidad Loyola Andalucía. 2019. <https://repositorio.ulozola.es/bitstream/handle/20.500.12412/2164/PlanPreservacionBiblioteca.pdf?sequence=1&isAllowed=y>.
- Política preservación digital del O2 Repositorio UOC. 2020 <http://hdl.handle.net/10609/128486>.

HERRAMIENTAS PARA PROCESOS

- 3-Heights document converter <https://www.pdf-tools.com/en/products/document-converter/>.
- BadPeggy <https://www.coderslagoon.com/#/product/badpeggy>
- BWF metaedit <https://sourceforge.net/projects/bwfmetaedit>.
- CranMD. Verificador de formatos y características de ficheros, validador de transferencias, generador de metadatos.
- Droid <https://www.nationalarchives.gov.uk/information-management/manage-information/preserving-digital-records/droid/>.
- Fido <http://fido.openpreservation.org>.
- Fixity <https://www.weareavp.com/es/productos/fixity> <http://www.cobdc.net/programarillure/fixity-para-monitorizar-sistemas-de-archivos-preservacion-digital/>.
- FixityPro <https://www.weareavp.com/es/productos/fixity-pro/>.
- Jhove <http://jhove.openpreservation.org>.
- Quick view plus <https://quick-view-plus.softonic.com>.
- veraPDF <https://openpreservation.org/products/verapdf/>.

REPERTORIOS DE UTILIDADES:

- Community Owned Digital Preservation Tool Registry COPTR https://coptr.digipres.org/index.php/Tools_Grid.
- Preserving Digital Objects With Restricted Resources (POWRR) Tool Grid <https://digitalpowrr.niu.edu/digital-preservation-101/tool-grid/>.

SOLUCIONES INTEGRADAS SPD

- Archivematica: <https://www.archivematica.org/es/>, <https://wiki.archivematica.org/Sandbox>.
- ArchiveSpace <https://archivesspace.org/>; <https://archivesspace.org/application/sandbox>.
- Libsafe / Libdata <https://es.libnova.com/preservacion-digital-libsafe-y-libdata/>.
- Odilo Preserver <https://www.odilo.es/preservacion/>.
- Roda (Repository of Authentic Digital Objects) https://demo.roda-community.org/?locale=es_CL#welcome.
- Rosetta <https://exlibrisgroup.com/products/rosetta-digital-asset-management-and-preservation/>.

COMPUTACIÓN Y ALMACENAMIENTO EN LA NUBE (CLOUD COMPUTING & DATA STORAGE)

- Amazon S3 / Amazon Glacier <https://aws.amazon.com/es/s3/pricing/> Estimación de 50 Terabytes en opción S3 estándar con acceso frecuente: 1, 228 USD/1059€ x mes. Opción S3 Glacier: 230 USD/198€ x mes.
- Archive.org (Wayback machine). <https://archive.org/web/> Cubre solo la capa superficial. Por lo general pierde sistemas relacionales, librerías, sistemas flash, etc.
- Arkivum <https://arkivum.com>.

- Digital Preservation Consortium. Servicios en la nube (2015) <https://www.dpconline.org/handbook/technical-solutions-and-tools/cloud-services>.
- eArchiving (Unión Europea) Apuesta por la interoperabilidad para toda la UE y que sea válida para todos y haga posible la reutilización de datos. Es un «building block» que incluye especificaciones, herramientas tecnológicas, formación, SIP/AIP/DIP, preservación de bases de datos, geodatos, <https://ec.europa.eu/cefdigital/wiki/display/CEFDIGITAL/eArchiving>. <https://www.culturaydeporte.gob.es/dam/jcr:0566f6d4-bb18-420a-b2a8-25d61a2b369c/cef-digital-earchiving-infographic-es.pdf>. Explicación en vídeo en minuto 3:02 <https://www.youtube.com/watch?v=cX-yuHgxwD8>.
- Duracloud <https://duraspace.org/duracloud/> Utiliza Amazon S3, entre 5 y 19 Tb, entre 1235+700USD Tb/año como mínimo 4,735 USD/año en 5Tb y 5,520+625USD Tb/año de coste 17,395USD/año en 19Tb. Para espacios mayores hay que pedir presupuesto basado en <https://wiki.lyrasis.org/download/attachments/90979118/duracloud-large-account-pricing-2020.pdf>.
- Microsoft Sharepoint <https://docs.microsoft.com/es-es/office365/servicedescriptions/sharepoint-online-service-description/sharepoint-online-limits>.
- Perpetua (EBSCO information services & Arkivum) <https://www.ebsco.com/es/productos/perpetua>.
- Preservica (DSpace & Amazon web services) <https://preservica.com/digital-archive-software/products-editions> <https://starter.preservica.com/pricing>.
- Telefónica Virtual Data Center (VDC) <https://www.movistar.es/grandes-empresas/soluciones/fichas/virtual-data-center/#>.