MANUALES Y GUÍAS | 2020

GRUPO DE REPOSITORIOS

M

GUÍA PARA LA EVALUACIÓN DE LOS PROCESOS DE PRESERVACIÓN EN REPOSITORIOS INSTITUCIONALES DE INVESTIGACIÓN

crue Universidades Españolas

Red de Bibliotecas REBIUN

Colección Estudios e Informes, 2020

GUÍA PARA LA EVALUACIÓN DE LOS PROCESOS DE PRESERVACIÓN EN REPOSITORIOS INSTITUCIONALES DE INVESTIGACIÓN

REBIUN Línea 3 (3er. P.E.) Grupo de Repositorios

crue Universidades Españolas Red de Bibliotecas REBIUN

Este informe ha sido elaborado por los miembros de la Acción 5 del Grupo de Repositorios de REBIUN (2020).

Cristina Azorín (ORCID 0000-0002-0063-9643)
Universitat Autònoma de Barcelona. cristina.azorin@uab.cat

José Manuel Barrueco (ORCID 0000-0001-7916-7847)
Universitat de València. jose.barrueco@uv.es

Isabel Bernal (ORCID 0000-0003-2506-9947)
Consejo Superior de Investigaciones Científicas. isabel.bernal@bib.csic.es

Víctor Macías (ORCID 0000-0002-4743-7483)
Universidad de Las Palmas de Gran Canaria. victor.macias@ulpgc.es

Rebeca Marín (ORCID 0000-0002-0010-1544)
Universidad Carlos III de Madrid. rebeca.marin@uc3m.es

Cristal Martínez (ORCID 0000-0003-4562-7571)
Universidade de Santiago de Compostela. cristal.martinez@usc.es

Miquel Térmens (ORCID 0000-0002-7305-3424)
Universitat de Barcelona. termens@ub.edu

REBIUN, Red de Bibliotecas Universitarias Españolas, es una comisión sectorial de la Conferencia de Rectores de las Universidades Españolas (CRUE) desde 1998. El Grupo de Trabajo de Repositorios de REBIUN fue creado en 2011 con la finalidad de potenciar los repositorios de contenidos y datos de investigación y docencia institucionales, y aprovechar las tecnologías e interoperabilidad para impulsar nuevos servicios de valor añadido. Su coordinador es Ciro Llueca (Universitat Oberta de Catalunya).

15 de octubre de 2020

0. Resumen

La presente Guía tiene como finalidad permitir una auditoría interna para establecer posibles acciones de preservación digital en los repositorios institucionales de las universidades y el CSIC. Se engloba dentro de la acción 5 del Grupo de Repositorios de REBIUN para el año 2020.

En 2018 el Grupo realizó una encuesta[1], basada en los niveles establecidos por la National Digital Stewardship Alliance (NDSA), para determinar en qué medida los repositorios están desarrollando actividades de preservación digital. Con una tasa de respuestas del 90% (52 encuestas respondidas de 58 repositorios REBIUN contactados), se pudo concluir que los repositorios no están aplicando las medidas técnicas de preservación al nivel de su compromiso público.

La presente Guía, basada en los mismos criterios NDSA pero ampliada con otras prácticas en preservación digital y con la experiencia de los miembros del grupo de trabajo, pretende servir de referente para plantear a nivel político y técnico los puntos fuertes y las medidas de mejora en cuanto a la preservación de los materiales y metadatos almacenados en los repositorios institucionales.

Palabras clave: repositorios institucionales de investigación; preservación digital; políticas institucionales; evaluación.

[1] Informe sobre la evaluación del estado de la preservación de los repositorios, REBIUN 2018: https://hdl.handle.net/20.500.11967/253

1. La preservación de los documentos digitales

Podríamos caracterizar la ciencia en el siglo XXI como colaborativa; orientada a maximizar los retornos y el impacto en la sociedad; social, por el papel importante que juegan las redes sociales especializadas en la creación e intercambio de conocimiento; y abierta, por la importancia de la disponibilidad en abierto de los resultados y los datos de investigación. En este sentido, estamos experimentando un nuevo paradigma en la forma en que se desarrolla la ciencia, es la Ciencia Abierta.

Los repositorios institucionales de investigación son un elemento esencial en ese nuevo ecosistema de comunicación de la ciencia. Han de garantizar una disponibilidad de calidad a la producción científica en acceso abierto de la institución a la que pertenecen, proporcionando los recursos técnicos y tecnológicos adecuados para asegurar su integridad y su perdurabilidad a largo plazo. Sin embargo, es importante tener en cuenta que la mayoría de programas de gestión de repositorios no preservan por sí solos, aunque sí puedan ayudar en la preservación la ejecución de determinadas funciones, programas o rutinas, como pueden ser las sumas de verificación.

La Digital Preservation Coalition (DPC) define la preservación digital como “el conjunto de actividades gestionadas necesarias para garantizar el acceso continuo a los materiales digitales durante el tiempo que sea necesario…”

En los últimos años se han sucedido las iniciativas internacionales que intentan proporcionar unas directrices a las organizaciones encargadas de la preservación. La primera fue la Carta para la preservación digital de la UNESCO, como documento que sentó las bases en esta materia; el modelo conceptual de referencia OAIS (Open Archival Information System); la metodología de la National Digital Stewarship Alliance (NDSA) para evaluar el nivel de preservación digital en una determinada institución; los metadatos de preservación PREMIS (Preservation Metadata Maintenance Activity); etc.

La presente guía pretende ser un documento de ayuda para el personal de gestión de los repositorios institucionales en los procesos involucrados en la preservación digital, introduciendo iniciativas como las citadas anteriormente, y conceptos clave en la preservación digital.

El planteamiento es eminentemente práctico y se estructura como un documento de autoevaluación, una guía para detectar puntos fuertes y débiles, establecer prioridades y adoptar medidas de mejora. No se plantea como una certificación ni una competición. En muchos casos puede servir para realizar un análisis que permita determinar posibles acciones de preservación, por ello la honestidad en las respuestas revela el compromiso real de la institución.

Se puede realizar una autoevaluación similar en línea a través de la Asociación Iberoamericana de Preservación Digital (APREDIG)[2], aunque utiliza la primera versión de los niveles NDSA.

Se incluye al final un anexo con bibliografía y enlaces a recursos para ampliar información.

2. Concienciación y pedagogía

Como personal de gestión de repositorios, debemos concienciar a los órganos políticos y al personal de la institución sobre la importancia de la preservación, a la vez que debemos hacer pedagogía entre los usuarios para que reconozcan la importancia de seguir nuestras recomendaciones. Las soluciones técnicas no son suficientes si la institución no asume (a nivel financiero, estratégico y normativo) la necesidad de la preservación digital.

La preservación digital no se puede considerar solo desde el punto de vista de la implementación de una determinada tecnología, sino que precisa además de una fuerte coordinación y organización del personal y de la institución (definiendo roles y responsabilidades, por ejemplo). Tampoco puede considerarse como un proyecto aislado (presupuesto excepcional, personal temporal, plazo establecido...) sino como un proceso en el resto de actuaciones rutinarias del repositorio.

La Ciencia Abierta requiere unos sistemas de gestión, acceso y preservación eficientes. Para garantizar la disponibilidad de los contenidos a largo plazo debe ser prioritario establecer una estrategia de control de amenazas, documentación

[2] Herramienta de auto-evaluación: https://es.surveymonkey.com/r/P9N7FLZ

de procesos y gestión de riesgos en paralelo a un plan de seguimiento de acciones concretas que llevar a cabo.

Los repositorios institucionales deberían desarrollar:

- Una política que ponga de manifiesto el compromiso institucional con la preservación de sus colecciones digitales.
- Documentación de procedimientos, calendarios, flujos de trabajo...

A nivel de propiedad intelectual deben respetarse las condiciones de los editores comerciales en cuanto a la versión que es posible almacenar y conservar en el repositorio. Es necesario además registrar en los metadatos las licencias de los contenidos puesto que en algunos casos estas licencias pueden afectar a la preservación digital; por ejemplo, en el caso de que no se permita la migración, ya que durante la migración se modifica el recurso original.

En el caso de publicaciones sin derechos de autor cedidos a una editorial (informes, tesis, datos…) recomendamos el uso de licencias Creative Commons[3] para documentos en acceso abierto y las Rights Statements[4] para informar de derechos de autor y de reutilización de objetos digitales. Ambas tipologías de declaraciones se ponen a disposición como datos enlazados (linked data). Cada declaración de derechos está ubicada en una URI.

3. Autoevaluación, los 10 puntos principales

1. Planificación
2. Plan de preservación
3. Integridad de datos
4. Formatos
5. Metadatos
6. Almacenamiento. Copias

[3] Creative Commons licencias: https://creativecommons.org/licenses/?lang=es_ES

[4] Rights Statements: https://rightsstatements.org/page/1.0/?language=es

7. Inventarios
8. Flujos de trabajo
9. Documentación de procesos
10. Gestión de riesgos

Planificación

Hay diferentes riesgos que pueden afectar a la preservación de datos digitales. Además, el volumen de ítems digitales es mucho mayor que el que nunca hemos tenido en soporte analógico. También sabemos que algunas actuaciones de tipo informático pueden tener un gran coste económico y no se pueden improvisar. Todo ello implica que la conservación de objetos digitales tenga que planificarse aún más que la conservación de objetos analógicos.

Algunas de las acciones que se han de planificar son: la sustitución del hardware obsoleto por otro nuevo, el tiempo de máquina para comprobar la integridad de los ficheros, el tiempo necesario para realizar una gran migración de formatos, la sustitución de soportes de almacenamiento por otros de mayor capacidad (pasar de discos físicos a almacenamiento en la nube, o migración de cintas LTO[5] a una generación más actual, por ejemplo).

La planificación empieza con la redacción y la aprobación de dos documentos: la política de preservación y el plan de preservación. En el primero, la institución u organización declaran su compromiso con la preservación digital de los contenidos del repositorio y hacen explícitas las pautas básicas de su actuación; se trata por tanto de un documento del más alto nivel, muy ligado a los objetivos de la organización y a sus planes estratégicos. En el segundo documento se describen los principios tecnológicos y las actuaciones técnicas en los que se basa su estrategia de preservación digital. Los dos documentos se han de aprobar de manera formal y se han de revisar y actualizar de forma periódica.

La política de preservación digital de una institución ha de ser estable y ha de tener una vigencia equivalente a los planes estratégicos. Se recomienda revisar el documento de política de preservación digital como mínimo cada cinco años.

[5] LTO: Linear Tape-Open, tecnología de cinta magnética de almacenamiento de datos.

Ejemplos de políticas de preservación digital tenemos:

- Política de preservación del Dipòsit Digital de Documents de la UAB (DDD) https://ddd.uab.cat/record/189808
- Yale University Library's Digital Preservation Policy Framework https://web.library.yale.edu/departments/preservation/policies-procedues-guidelines
- Cornell University Library Digital Preservation Policy Framework https://ecommons.cornell.edu/handle/1813/11230

Ejemplo de modelos para redactar una política de preservación digital:

- NEDCC Digital Preservation Policy Template https://www.nedcc.org/assets/media/documents/SoDAExerciseToolkit.pdf
- SCAPE: Scalable Preservation Environments. Catalogue of preservation policy elements https://scape-project.eu/wp-content/uploads/2014/02/SCAPE_D13.2_KB_V1.0.pdf

Acciones imprescindibles para garantizar un primer nivel de preservación en el repositorio
¿La institución ha aprobado formalmente una política de preservación?
Acciones imprescindibles para garantizar niveles avanzados de preservación en el repositorio
¿Está prevista la actualización periódica de la política de preservación?

Plan de preservación

Un plan de preservación es un documento en el que se concretan las principales actuaciones técnicas que ejecuta una organización para asegurar la preservación a largo plazo de sus activos digitales. El plan de preservación es la concreción de los principios genéricos que previamente han sido aprobados en la política de preservación por la misma organización.

Hay diferentes factores y condiciones que pueden influir en la elección de la estrategia de preservación; el plan de preservación debe adaptarse a las necesidades y las estructuras de su entorno, a cuestiones tanto técnicas como organizativas.

La elaboración de un plan de preservación asegura que todas las actuaciones en este campo están coordinadas, siguen unos objetivos determinados y se rigen por normas técnicas y buenas prácticas que son asumidas por todo el personal implicado. A partir del plan de preservación se pueden diseñar flujos de trabajo para la realización de acciones concretas. También tiene la ventaja de fijar por escrito y hacer públicas decisiones que en caso contrario sólo serían conocidas por el personal que las tomó; así se asegura la transparencia de las actuaciones de preservación y que estas puedan pervivir más allá de las personas.

Las buenas prácticas internacionales recomiendan que los planes de preservación tengan en cuenta las orientaciones que proporciona el modelo OAIS (Open Archival Information System). OAIS, como habitualmente se conoce a la norma ISO 14721:2012, establece cuáles han de ser las funciones y conjuntos de acciones que debe desempeñar cualquier sistema de preservación digital. Cada organización debe adaptar este modelo a sus necesidades.

Los planes de preservación han de ser concretos, no de carácter teórico, y han de describir el funcionamiento real de las acciones de preservación en el presente y en el futuro inmediato. Dado que la tecnología cambia continuamente y también lo hacen los mismos datos a conservar, es importante que los planes de preservación sean revisados y actualizados de forma continuada.

Los planes de preservación digital contienen información técnica que puede cambiar a corto plazo, por ello se recomienda que se mantengan actualizados al mismo nivel que se exige a toda la documentación susceptible de ser sometida a auditoría. Aconsejamos revisar el documento de plan de preservación digital cada dos años.

Algunos contenidos que deberían contemplarse en el plan:

a) Alcance y propósito
b) Objetivos
c) Colecciones y usuarios
d) Funciones y responsabilidades
e) Compromisos y políticas institucionales
f) Acciones de preservación y control de calidad
g) Sostenibilidad financiera
h) Sostenibilidad técnica
i) Plan de contingencia y análisis de riesgos
j) Formación
k) Plan de comunicación
l) Evaluación, seguimiento y revisión del propio plan

Ejemplos de plan de preservación digital:

- Protocolo de preservación digital para el sistema de preservación digital del Archivo Español de Media Art (AEMA) https://bit.ly/31j7bKX
- Digital Preservation Plan Wheaton College Library and Archives https://library.wheaton.edu/sites/default/files/Digital_Preservation_Plan.pdf

Ejemplo de modelo para redactar un plan de preservación digital:

- SCAPE: Scalable Preservation Environments. Catalogue of preservation policy elements https://scape-project.eu/wp-content/uploads/2014/02/SCAPE_D13.2_KB_V1.0.pdf

Acciones imprescindibles para garantizar un primer nivel de preservación en el repositorio
¿Existe un plan de preservación en forma de documento aprobado?
Acciones imprescindibles para garantizar niveles avanzados de preservación en el repositorio
¿Está prevista la actualización periódica del plan de preservación? **¿Se dispone de documentos técnicos que desplieguen y detallen el plan de preservación?**

Integridad de los datos

La integridad de los datos supone garantizar que los documentos almacenados permanecen completos, tal como se introdujeron en el sistema, y que no han sufrido corrupción alguna o alteración no autorizada o documentada. La integridad también comprende aspectos como el encriptado o la firma digital.

La integridad es esencial para poder contar con datos confiables en todo momento y, por supuesto, a largo plazo. Asegurar la integridad de los ficheros que contienen los datos de nuestro repositorio comporta su chequeo mediante sumas de verificación (como MD5, SHA1, SHA256, entre otros ejemplos) a lo largo de todo su recorrido en el repositorio. No sólo en su producción, sino en la ingesta inicial en el sistema y, sucesivamente en cada paso posterior, ha de comprobarse el buen estado de los ficheros para verificar que no han sufrido alteraciones en ninguno de los pasos realizados. La verificación incluirá también el análisis en busca de posibles virus y la persistencia de los enlaces a ficheros

y a recursos externos e internos. Las comprobaciones han de quedar reflejadas en los metadatos de preservación correspondientes.

Cuando se produce una migración la organización debe decidir si conserva o elimina los ficheros originales y en ambos casos debe hacer un seguimiento de las diferentes versiones para poder demostrar la autenticidad del material.

Deben establecerse medidas de autenticación para garantizar que el personal no puede hacer cambios en los ficheros almacenados ni, por azar, suprimir algún objeto digital.

La integridad de los datos se ha de comprobar con la siguiente periodicidad:

- En el momento de la ingesta. Las sumas de verificación recibidas deben compararse con las sumas de verificación generadas en la recuperación, se muestra así la posible pérdida de bits durante el transporte. Esta medida debería implementarse en todos los movimientos de ficheros, incluso cuando se trasladen ficheros dentro del propio repositorio.
- Siempre que se produzca una migración de sistema informático.
- Después de que se haya producido un incidente grave que pueda haber comprometido la integridad de los datos.
- De forma periódica, a intervalos regulares, si es posible cada seis meses. Estas revisiones periódicas pueden suponer una gran carga de tiempo de proceso de máquina, por lo que su ejecución deberá programarse según la disponibilidad técnica.

Acciones imprescindibles para garantizar un primer nivel de preservación en el repositorio
¿Todos los ficheros del repositorio tienen información respecto a su integridad? (generación de valor *checksum*). **¿En caso de que sea necesario, se establece un control de virus en la ingesta de los ficheros?**
Acciones imprescindibles para garantizar niveles avanzados de preservación en el repositorio
¿Se verifica la información de integridad al mover o copiar los ficheros? **¿Se verifica la información de integridad con una periodicidad determinada?** **¿Se verifica la información de integridad en situaciones o actividades específicas?** **¿Si se duda de la integridad del fichero se reemplaza o repara su contenido?** **¿Se tiene una copia de la información de integridad en una localización separada de los ficheros que verifica?**

Formatos

La gestión de los formatos de ficheros debe ser parte de una planificación de preservación digital. Es importante tener en cuenta las posibles implicaciones a la hora de determinar los formatos que se aceptarán en el repositorio dado que estos pueden variar significativamente en sus propiedades.

Otra cuestión a tener en cuenta es la proliferación de formatos, lo que puede redundar negativamente en su gestión. Por ello, no es recomendable que un repositorio termine gestionando un número excesivamente alto de formatos y de versiones de formatos ya que suelen evolucionar con bastante frecuencia. En una estrategia de preservación digital es fundamental estar al día del recorrido de cada formato utilizado. En ese sentido, la normalización, unificación en la ingesta a los formatos seleccionados por el repositorio, suele ser una estrategia básica, bastante efectiva y no costosa para asegurar la validez de los ficheros en el futuro. Otras estrategias incluyen la migración o la emulación de formatos.

Los formatos que tienen más opciones de ser accesibles en el futuro son: no propietarios, basados en estándares abiertos, no encriptados, no comprimidos y aceptados por una amplia comunidad de usuarios (por ejemplo: pdf, csv, mp4, jpg...).

Guías y especificaciones de formatos recomendados según tipos de objetos digitales, sostenibilidad y políticas institucionales

- British Library File Formats Assessments

http://wiki.dpconline.org/index.php?title=File_Formats_Assessments

- Harvard University Library Formats Assessment

https://wiki.harvard.edu/confluence/display/digitalpreservation/Format+Assessments

- Library of Congress recommended format specifications

http://www.loc.gov/preservation/resources/rfs/index.html

- Library of Congress sustainability factors

http://www.digitalpreservation.gov/formats/index.shtml

• Wiki File Formats by Extension

http://fileformats.archiveteam.org/wiki/Category:File_formats_by_extension

Se recomienda revisar los formatos de los documentos en el momento de su ingesta. Se ha de comprobar que la extensión del fichero se corresponde con el formato del mismo; también se ha de comprobar que los formatos introducidos están dentro de la política de formatos admitidos por el repositorio.

Acciones imprescindibles para garantizar un primer nivel de preservación en el repositorio
¿Se conocen y se tiene capacidad para gestionar correctamente todos los formatos de ficheros que se encuentran en el repositorio?
Acciones imprescindibles para garantizar niveles avanzados de preservación en el repositorio
¿Se promociona y recomienda el uso de formatos abiertos? **¿Se monitoriza la obsolescencia de los formatos?**

Metadatos

Los metadatos constituyen un pilar fundamental a la hora de montar una estrategia de preservación digital. Los metadatos incluyen cualquier información contextual que permita asegurar un acceso sostenible y continuo a los contenidos. Estos metadatos no son solamente de carácter técnico, sino que deben contener información descriptiva suficiente para la correcta puesta a

disposición, identificación, persistencia, ejecución, comprensión, integridad y autenticidad de los objetos digitales a lo largo del tiempo.

El objetivo principal de los metadatos de preservación es documentar las actividades realizadas sobre los objetos digitales para facilitar su preservación, como pueden ser: análisis antivirus, cambios de nombre y las migraciones de formatos.

Es importante que en la aplicación de los metadatos se sigan estándares reconocidos, por ejemplo PREMIS. Otra recomendación básica es que los metadatos estén en un formato legible por máquinas (XML, por ejemplo) para permitir la automatización de ciertas actividades de preservación y difusión.

Los metadatos deben actualizarse periódicamente en función de los requerimientos que realizan diferentes infraestructuras o plataformas, se expondrán de forma compatible para favorecer una correcta recolección.

Acciones imprescindibles para garantizar un primer nivel de preservación en el repositorio
¿Está determinada y documentada la tipología de metadatos aplicada en el repositorio?
Acciones imprescindibles para garantizar niveles avanzados de preservación en el repositorio
¿Los metadatos usados se revisan de manera periódica y se someten a control a través de listados, incidencias o avisos automáticos?

Almacenamiento. Copias

Disponer de varias copias de los datos (software, metadatos y ficheros) es el principal mecanismo del que disponemos para asegurar su acceso y preservación futuro.

Los datos almacenados han de tener como mínimo dos copias completas en ubicaciones geográficamente separadas, empleando soportes de almacenamiento estables y redundantes. El contenido ha de estar inventariado indicando la localización dentro del almacenamiento y disponer de una copia del mismo en un lugar diferenciado del resto de los datos. Con independencia de si se almacenan los datos empleando sistemas propios o externos, deberemos elaborar un plan de acción frente a la obsolescencia de nuestro software, hardware y medios utilizados para el depósito de los datos.

En los presupuestos a largo plazo deben tenerse en cuenta los costes de un reemplazamiento periódico de los componentes del sistema. Los elementos de hardware normalmente tienen una vida útil en torno a los cinco años.

Acciones imprescindibles para garantizar un primer nivel de preservación en el repositorio

¿Existen dos copias completas de todos los metadatos del repositorio, en localizaciones separadas?

¿Existen dos copias completas de todos los ficheros del repositorio, en localizaciones separadas?

¿Todos los documentos, nacidos digitales o digitalizados, se encuentran localizados en el ecosistema del repositorio? No existen ficheros en soportes o localizaciones sobre los que no se pueda actuar de manera global.

¿Está garantizado el mantenimiento del sistema de almacenamiento en el que se encuentran los ficheros?

Acciones imprescindibles para garantizar niveles avanzados de preservación en el repositorio

¿Existe al menos una copia del software del repositorio (programas, configuraciones, diseño gráfico, etc.)?

¿Existe al menos una copia de los metadatos de todos los registros del repositorio en un formato estándar?

¿Existen tres copias completas de los ficheros del repositorio, y al menos una de ellas en una localización geográficamente separada y con una amenaza de desastre diferente al resto?

Inventario

El inventario de los objetos digitales disponibles en el repositorio, metadatos y ficheros, es una herramienta fundamental para determinar tanto el volumen de datos con que contamos como los formatos utilizados y las versiones de los mismos. Nos permitirá llevar a cabo una monitorización sobre la obsolescencia de los formatos para determinar en qué momento tendremos que llevar a cabo su migración. Además, nos permitirá obtener informes sobre el contenido y el estado de todo el material.

Los elementos mínimos de un inventario deberían ser: nombre, localización física (URI), formato, versión y tamaño de fichero.

Se recomienda actualizar los inventarios:

- Después de una migración de sistema informático.
- Después de una modificación en los medios de almacenamiento usados.
- Después de una migración de formatos.

- Después de que se haya producido un incidente grave que pueda haber comprometido la integridad de los datos.
- Periódicamente para incluir los nuevos documentos ingresados.

Acciones imprescindibles para garantizar un primer nivel de preservación en el repositorio
¿Se dispone de un registro de los formatos de ficheros soportados por el repositorio? **¿Se tienen identificados los formatos de todos los ficheros ingresados?**
Acciones imprescindibles para garantizar niveles avanzados de preservación en el repositorio
¿Se pueden generar estadísticas y listados de los ficheros según sus formatos y versión?

Flujos de trabajo

Una vez definidas las políticas de preservación es necesario documentar los correspondientes flujos de trabajo básicos. Un flujo de trabajo es una secuencia de pasos o actividades necesarios para colocar contenidos digitales bajo control para la preservación. Pueden ser muy variables dependiendo de las políticas institucionales, capacidad o tipología de contenidos del repositorio. Los flujos son dinámicos, cambian con el tiempo, a partir de la experiencia, mejoras en las aplicaciones y otros factores.

Se recomienda revisar la documentación de los flujos de trabajo cada dos años, esta debe mantenerse actualizada al mismo nivel que se exige a toda la documentación susceptible de ser sometida a auditoría.

Acciones imprescindibles para garantizar un primer nivel de preservación en el repositorio
¿Existe un acceso con diferentes niveles de permiso a los ficheros, por parte del personal o por parte de aplicaciones? Permisos delimitados de lectura, escritura, gestión o eliminación. **¿Se establece un sistema de permisos sobre la gestión de los ficheros para evitar su modificación o eliminación por error?**
Acciones imprescindibles para garantizar niveles avanzados de preservación en el repositorio
¿Se tienen identificadas las personas o aplicaciones que tienen algún tipo de permiso de actuación sobre los ficheros? **¿Además de existir esta identificación, se mantiene un registro de los cambios o actuaciones realizados?**

Documentación de procesos

Según la norma ISO 9000[6], un procedimiento es una forma específica de llevar a cabo una actividad o un proceso. En la gestión de repositorios se realizan distintas actividades o procesos habitualmente: comprobación de derechos de autoría, digitalización, ingesta de nuevos ítems, modificaciones para adaptarse a nuevos requerimientos, migraciones... Para la mayoría ya existirá un procedimiento que puede ir cambiando en el tiempo, pero, ¿está documentado? ¿se actualiza la documentación de los procedimientos?

Documentar los procedimientos es importante para la preservación digital ya que ayudará al personal a tomar la decisión correcta cuando se requieran acciones de preservación. Además, esta documentación facilita las transiciones cuando se producen cambios en el personal vinculado al repositorio, ya que sirven de guía para el desempeño de los procesos y actividades.

Se recomienda revisar la documentación de procesos de trabajo cada dos años, debe mantenerse actualizada al mismo nivel que se exige a toda la documentación susceptible de ser sometida a auditoría.

Acciones imprescindibles para garantizar un primer nivel de preservación en el repositorio

¿Están debidamente documentados todos los tipos de soportes de almacenamiento? Información respecto al mantenimiento, obsolescencia…

¿Se documentan los procesos de verificación de la integridad y sus resultados?

[6] ISO 9000:2015(es) Sistemas de gestión de la calidad — Fundamentos y vocabulario: https://www.iso.org/obp/ui/es/#iso:std:iso:9000:ed-4:v1:es

Acciones imprescindibles para garantizar niveles avanzados de preservación en el repositorio

¿Se mantiene un registro de las acciones de preservación asociadas con el contenido y cuándo ocurren esas acciones?

¿El repositorio ofrece la exportación de los registros en diferentes formatos de metadatos?

Gestión de riesgos

El riesgo se define como la combinación de la probabilidad de un evento y sus consecuencias. La gestión de riesgos es un factor clave para garantizar la preservación digital y tiene como objetivo gestionar o tratar de disminuir el riesgo residual. Implica identificar amenazas, evaluar la vulnerabilidad de los activos críticos a amenazas específicas, determinar el riesgo, identificar formas para reducirlo y priorizar las medidas de reducción de riesgos.

Se recomienda realizar un nuevo análisis de riesgos cada dos años. Esta documentación se ha de mantener actualizada al mismo nivel que se exige a toda la documentación susceptible de ser sometida a auditoría.

Algunos de los riesgos identificados para la preservación de los documentos digitales son:

- fallos en los soportes, en el software o en el hardware
- la obsolescencia del: soporte físico, formato, software o hardware
- degradación bit rot
- ataques informáticos deliberados (internos o externos)
- problemas organizacionales, económicos o políticos
- errores en las comunicaciones, en la red, en la prestación de servicios…
- errores humanos

- desastres naturales

Acciones imprescindibles para garantizar un primer nivel de preservación en el repositorio
¿Se realiza un análisis de riesgos del repositorio siguiendo alguna norma?
Acciones imprescindibles para garantizar niveles avanzados de preservación en el repositorio
¿Existen un plan y actuaciones específicas para la renovación tanto del software como del hardware del repositorio? **¿El repositorio está preparado para realizar migraciones, normalizaciones o emulaciones que garanticen el acceso a los contenidos?**

4. Bibliografía

Universidad Nacional Autónoma de México. Área de Tecnología del Grupo de Preservación Digital (2020). *Criterios básicos para valorar sistemas de preservación digital*. Recuperado de https://www.iib.unam.mx/index.php/instituto-de-investigaciones-bibliograficas/publicaciones/libros-electronicos/557-criterios-basicos-para-valorar-sistemas-de-preservacion-digital-2

Barbedo, F. (2019). *Recomendações para a produção de planos de preservação digital. 2ª versão.* Recuperado de http://arquivos.dglab.gov.pt/wp-content/uploads/sites/16/2019/08/Recomendacoes_PPD_v2.pdf

Barrueco Cruz, José Manuel, et al. (2017). *Guía para la evaluación de repositorios institucionales de investigación*. 3ª ed. Recuperado de https://www.recolecta.fecyt.es/node/1199

Beagrie, N., Semple, N., Williams, P. y Wright, R. (2008). *Digital Preservation Policies Study. Part 1: Final Report*. Recuperado de http://www.jisc.ac.uk/media/documents/programmes/preservation/jiscpolicy_p1finalreport.pdf

Bishoff, L. (2010). Digital Preservation Plan. *Information Standards Quarterly,* 22 (2), pp. 20-25. Recuperado de https://www.niso.org/sites/default/files/stories/2019-11/FE_Bishoff_Digital_Preservation_Plan_isqv22no2.pdf

Caplan, P. (2009). *Entender PREMIS*. Recuperado de http://hdl.handle.net/10421/981

Centro Criptológico Nacional. (2011). *Esquema Nacional de seguridad. Política de seguridad de la información. Guía de seguridad (CCN-STIC-805)*. Recuperado de https://www.ccn-cert.cni.es/series-ccn-stic/800-guia-esquema-nacional-de-seguridad/508-ccn-stic-805-politica-de-seguridad-de-la-informacion/file.html

Community Owned digital Preservation Tool Registry (COPTR). (2020). Recuperado de http://coptr.digipres.org/

Consultative Committee for Space Data Systems (CCSDS). (2012). *Reference model for an Open Archival Information System (OAIS). Recommended practice CCDS 650.0-M-2. Magenta Book*. Recuperado de https://public.ccsds.org/pubs/650x0m2.pdf

CoreTrustSeal (2020). Recuperado de https://www.coretrustseal.org/

DigiPres Commons. *Community-owned digital preservation resources*. Recuperado de http://www.digipres.org/

Digital Preservation Coalition. *Digital Preservation Handbook* (2015). 2nd Edition. Recuperado de https://www.dpconline.org/handbook

Erway, R. (2012). *You've Got to Walk Before You Can Run : First Steps for Managing Born Digital Content Received on Physical Media*. Recuperado de http://www.oclc.org/content/dam/research/publications/library/2012/2012-06.pdf

Guia da preservação digital por siglas. (2018). Recuperado de http://preservacaodigital.wixsite.com/guia

Hillmann, D. I., Marker, R., y Brady, C. (2008). Metadata standards and applications. *Serials Librarian*, *54* (1–2), pp. 7–21. Recuperado de https://doi.org/10.1080/03615260801973364

International Conferences on Digital Preservation (iPRES) (2004-). Recuperado de http://www.ipres-conference.org/

International Organization for Standarization. (2012). *Space data and information transfer systems – Open archival information system (OAIS) – Reference model (ISO 14721:2012).* Recuperado de https://www.iso.org/standard/57284.html

International Organization for Standarization. (2012). *Space data and information transfer Systems – Audit and certification of trustworthy digital repositories* (ISO 16363:2012). Recuperado de https://www.iso.org/standard/56510.html

National Digital Stewarship Alliance (NDSA). Levels of Preservation Revisions Working Group. (2019). *Using the Levels of Digital Preservation: an overview for V2.0*. Recuperado de https://osf.io/vnc32

Ravelo Díaz, G., Mena Mugica, M. M. y Del Castillo Guevata, J. (2019). Requisitos para la valoración de riesgos de preservación en repositorios digitales. *Biblios: Journal of Librarianship and Information Science*, (75), pp. 25-34. Recuperado de https://doi.org/10.5195/biblios.2019.484

National Digital Library of India (NDLI), National Library of Australia (Trove), Digital Public Library of America (DPLA), National Heritage Digitization Strategy [Canada] (NHDS), National Library of New Zealand, Europeana Foundation. *Rights statements*. (2016). Recuperado de https://rightsstatements.org/en/

Termens, Miquel (2013). *Preservación digital*. Barcelona: Editorial UOC.

Termens, Miquel y Leija, David (2017). Auditoría de preservación digital con NDSA Levels. *El profesional de la información*, 36 (3), pp. 447-456. Recuperado de https://doi.org/10.3145/epi.2017.may.11

The Library of Congress. *Tools for preservation metadata implementation*. (2010). Recuperado de http://www.loc.gov/standards/premis/tools_for_premis.php

5. Anexo

Directorios de software, herramientas y servicios para la preservación digital

• Community Owned Digital Preservation Tool Registry COPTR

http://coptr.digipres.org/Main_Page

• Digital Curation Centre (DCC) tools and services

http://www.dcc.ac.uk/resources/external/category/preservation-planning

• Inventory of FLOSS (Free/libre open-source software) in the cultural heritage domain

https://docs.google.com/spreadsheets/d/1bOoQiXFjGyR3oEubdLdkfCat7V4TsNLnEXGOJWkJ63c/edit#gid=516255520&vpid=D2

• Preserving Digital Objects With Restricted Resources (POWRR) Tool Grid

http://www.digipres.org/tool

crue Universidades Españolas
Red de Bibliotecas REBIUN