FACULTAD DE HUMANIDADES

DEPARTAMENTO CIENCIA DE LA INFORMACIÓN

LICENCIATURA EN BIBLIOTECOLOGÍA Y DOCUMENTACIÓN

TESINA

Prácticas de preservación digital en los repositorios institucionales de acceso abierto de la Universidad de Buenos Aires

Tutora: Lic. Milagros Pandolfo (CONICET. Universidad Nacional de Mar del Plata. Facultad de Humanidades)

Co-Tutora: Mg. Alicia Beatriz Hernandez (Universidad Nacional de Mar del Plata. Facultad de Humanidades)

Alumna: Bib. Fabiana Laura Salerno

2026

RESUMEN

Este trabajo presenta un análisis de las prácticas de preservación digital en los repositorios institucionales de acceso abierto de la Universidad de Buenos Aires. El estudio investiga el estado actual de los procesos y políticas de preservación digital, evaluando su efectividad y conformidad con los estándares internacionales. Se identifican y analizan las características distintivas de las prácticas existentes, así como las posibles brechas y áreas de mejora, explorando la capacidad de los repositorios para adaptarse a los cambios tecnológicos considerando el contexto institucional, normativo y de los recursos disponibles. El objetivo principal de esta investigación es proponer recomendaciones concretas y viables que fortalezcan las prácticas de preservación digital de los repositorios de la universidad, para asegurar la conservación del patrimonio digital y facilitar el acceso a la información de la comunidad académica.

PALABRAS CLAVE: preservación digital; repositorios institucionales; acceso abierto; gestión de repositorios digitales; políticas de preservación digital

ABSTRACT

This paper presents an analysis of digital preservation practices in the open access institutional repositories of the University of Buenos Aires. The study investigates the current state of digital preservation processes and policies, evaluating their effectiveness and compliance with international standards. The distinctive features of existing practices are identified and analyzed, as well as possible gaps and areas for improvement, exploring the capacity of the repositories to adapt to technological changes considering the institutional, regulatory and resource-available context. The main objective of this research is to propose concrete and viable recommendations that strengthen the digital preservation practices of the university's repositories, to

ensure the conservation of digital heritage and facilitate access to information for the academic community.

KEYWORDS: digital preservation; institutional repositories; open access; digital repository management; digital preservation policies

Agradecimientos

Deseo expresar mi más profundo reconocimiento a la Lic. Milagros Pandolfo y a la Mg. Alicia Beatriz Hernandez, directora y co-directora de este trabajo, por sus valiosos aportes y la generosidad al brindarme las herramientas necesarias para la realización de esta investigación. Asimismo, agradezco a mis colegas, amigos y familia por el apoyo y el acompañamiento durante esta instancia académica.

INDICE

Índice de gráficos, figuras y tablas

Introducción

En la era digital, la información se ha consolidado como un activo invaluable que impulsa el conocimiento, la innovación y el progreso en todas las esferas de la sociedad. Sin embargo, la propia naturaleza de los datos digitales, caracterizada por su volatilidad y dependencia tecnológica, plantea desafíos significativos para su preservación a largo plazo.

A diferencia de los documentos físicos, que pueden perdurar durante siglos con el cuidado apropiado, la información digital es altamente susceptible a la obsolescencia tecnológica (de hardware y software). Este panorama exige la implementación de estrategias sólidas y sistemáticas de preservación digital, especialmente dentro del ámbito académico, donde la producción científica y de investigación constituye un patrimonio intelectual y cultural de valor incalculable.

La presente investigación se fundamenta en una exhaustiva revisión de la literatura especializada en preservación digital a largo plazo. Esta fase inicial se enfocó en el análisis crítico de fuentes académicas y profesionales (libros, artículos científicos, actas de congresos, tesis y documentos técnicos), con un doble objetivo de identificar las problemáticas clave que afectan la sostenibilidad de los repositorios digitales.

Dicha revisión fue fundamental para el desarrollo del marco teórico, permitiendo una profundización sobre principios, técnicas y estrategias de preservación. Se prestó especial atención a la identificación de estándares internacionales, marcos normativos y herramientas tecnológicas esenciales para asegurar la integridad, autenticidad y accesibilidad de los contenidos digitales a largo plazo.

La Universidad de Buenos Aires (UBA), reconocida como una de las instituciones de educación superior más prestigiosas de América Latina,

gestiona un vasto y creciente patrimonio digital a través de sus repositorios institucionales (RI) de acceso abierto.

Estos repositorios custodian la producción científico-académica resultante de las investigaciones, proyectos y creaciones intelectuales de su comunidad, constituyendo un acervo de valor estratégico, tanto para la universidad como para la sociedad. En este contexto, el objetivo principal de la presente investigación consiste en identificar y analizar las prácticas actuales y los desafíos inherentes a la preservación digital. El resultado de este análisis podrá servir como fundamento para la formulación de una estrategia de preservación digital efectiva y sostenible adaptada a las necesidades específicas de la UBA.

Para alcanzar este objetivo, el presente trabajo de investigación adoptará una aproximación empírica centrada en el estudio de las políticas de gestión de la información de los repositorios digitales en las distintas unidades académicas de la UBA. Dicho estudio se llevará a cabo mediante la aplicación de un instrumento de evaluación (una encuesta) diseñado para diagnosticar el estado actual de la preservación y la accesibilidad de los contenidos digitales.

Las instituciones a las que se envió la encuesta fueron aquellas que poseen un repositorio digital propio: Facultad de Agronomía; Facultad de Ciencias Sociales; Facultad de Ciencias Exactas y Naturales; Facultad de Ciencias Económicas; Facultad de Filosofía y Letras; Facultad de Ingeniería; Facultad de Medicina; y Facultad de Odontología; y las instituciones que ofrecen sus contenidos únicamente a través del *Repositorio Digital Institucional de la Universidad de Buenos Aires* (RDI-UBA), desarrollado y gestionado por el Sistema de Bibliotecas y de Información (SISBI): Facultad de Arquitectura, Diseño y Urbanismo; Facultad de Ciencias Veterinarias; Facultad de Derecho; Facultad de Farmacia y Bioquímica; y Facultad de Psicología.

La encuesta ha sido estructurada para abordar las dimensiones críticas de la preservación digital, como protocolos de respaldo de datos, identificación y gestión de formatos obsoletos o en riesgo de obsolescencia, actualización y migración de datos a nuevas plataformas, adopción de esquemas de metadatos estandarizados para facilitar la descripción y recuperación de los objetos digitales, mecanismos para garantizar la autenticidad e integridad los documentos digitales, disponibilidad de infraestructura tecnológica, personal capacitado y recursos financieros para la preservación digital, entre otros.

El análisis de la información recopilada ha permitido elaborar un diagnóstico de las políticas de preservación digital en los repositorios de la UBA, identificar áreas de mejora y proponer recomendaciones específicas para fortalecer la preservación a largo plazo de sus activos digitales. Esta investigación busca contribuir a la construcción de un ecosistema digital académico sólido y sostenible, donde el acceso al conocimiento esté garantizado para las generaciones presentes y futuras.

Definición del problema

Los RI, desde su concepción, manifiestan un compromiso institucional fundamental con la adhesión a los principios del Movimiento de Acceso Abierto y la implementación de sus estándares asociados, facilitando la difusión y el intercambio de la producción científica y académica. No obstante, esta dedicación a la apertura no se ha reflejado en la definición y aplicación de políticas de preservación digital explícitas y sistemáticas. Esta omisión representa un desafío crítico ya que la preservación a largo plazo de los activos digitales es fundamental para garantizar el acceso continuo al conocimiento y la investigación.

En este contexto, se ha identificado un problema de investigación fundamental: la ausencia de un análisis exhaustivo que derive en la definición de políticas de preservación digital. Esta carencia se manifiesta en la falta de una planificación estratégica a largo plazo en los RI de acceso abierto de la UBA. Esta brecha en la gestión de la preservación digital plantea interrogantes sustanciales sobre la sostenibilidad y la accesibilidad futura del patrimonio intelectual de la universidad, lo cual subraya la necesidad de abordar esta cuestión mediante una investigación rigurosa y la subsiguiente formulación de recomendaciones prácticas. Se plantean las siguientes preguntas:

- ¿Qué políticas y prácticas específicas de preservación digital se implementan en los RI de la UBA?
- ¿Cómo se incorporan los esquemas de metadatos, normas y directrices internacionales en las políticas de preservación digital de los repositorios de la UBA?
- ¿Qué áreas de mejora y recomendaciones se han identificado para optimizar las políticas actuales de preservación digital de los RI de la UBA?

Objetivo general

Esta investigación se propone realizar un estudio exhaustivo de las políticas de preservación digital implementadas por los RI de las Facultades de la UBA y por el Sistema de Bibliotecas e Información (SISBI) de la universidad. El objetivo principal es relevar, analizar y comparar las estrategias de preservación digital adoptadas por estas unidades académicas, para identificar las mejores prácticas, las áreas de mejora y los desafíos comunes en la preservación a largo plazo del patrimonio digital de la UBA.

Objetivos específicos

- Realizar un diagnóstico exhaustivo del estado actual de la gestión y preservación digital en los repositorios académico-científicos de acceso abierto de las Bibliotecas Centrales de la UBA.
- Efectuar un relevamiento sistemático de las políticas de preservación digital implementadas en los repositorios de la universidad.
- Analizar y evaluar las prácticas de preservación digital en relación con su adecuación a modelos y estándares internacionales.
- Desarrollar un plan de mejoras integral, si se identifica la necesidad, para optimizar las políticas de preservación digital en los RI de la UBA.

Antecedentes y marco teórico

El interés por la preservación digital comenzó a mediados de la década del noventa con la irrupción de Internet y la consecuente preocupación por la conservación de la información que crecía velozmente en el entorno web. Los primeros esfuerzos en este sentido se llevaban a cabo en forma individual o conformando pequeños grupos de trabajo y se centraban en colecciones muy específicas o en un reducido conjunto de formatos digitales. El resultado de la producción científico-académica, como pilar fundamental de la divulgación de las investigaciones y del reconocimiento de la ciencia y de las instituciones educativas de las cuales emerge, demandó una respuesta más estructurada. En un contexto de crecimiento exponencial de la información, se hizo indispensable establecer procesos de preservación a largo plazo para documentos y datos digitales.

La función tradicional del profesional bibliotecario se circunscribe primariamente a tareas técnicas como la selección, catalogación y clasificación de la información. No obstante, el avance sostenido de las

tecnologías ha provocado una transformación paradigmática en su rol. El bibliotecario actual se posiciona como un agente social y gestor de conocimiento que participa activamente en la construcción de la información (Allendez Sullivan, 2004). Esta evolución ha expandido significativamente sus responsabilidades, permitiéndole la interacción con los usuarios más allá de los límites físicos de la biblioteca y facilitando el establecimiento de redes de conocimiento globales que trascienden las barreras territoriales.

El crecimiento exponencial del volumen de información digital en instituciones de todos los sectores genera un desafío estructural en relación con el desarrollo, la implementación y el cumplimiento de políticas y directrices de preservación (Bodero Poveda *et al*., 2022). Como respuesta a esta coyuntura, los profesionales de la información han intensificado sus esfuerzos, enfocándose en la elaboración y aplicación de nuevas herramientas y procedimientos metodológicos orientados a garantizar la conservación a largo plazo y la accesibilidad continua de los objetos digitales que conforman el patrimonio institucional y cultural.

En la década de 1990, en Estados Unidos, la Comisión de Preservación y Acceso (Commission on Preservation and Access; CPA)[1] junto al Grupo de Bibliotecas de Investigación (Research Libraries Group; RLG)[2] encargaron al Grupo de Trabajo para el Archivo de Información Digital (Task Force on Archiving Digital Information)[3] un estudio que resultaría

[1] La CPA se creó para fomentar y apoyar la colaboración entre bibliotecas de los Estados Unidos, con el fin de garantizar la preservación de los registros documentales en todos los formatos y proporcionar un mejor acceso a la información académica. En 1997 se fusionó con el Consejo de Recursos Bibliotecarios (Council on Library Resources; CLR) para formar el Consejo de Bibliotecas y Recursos de Información (Council on Library and Information Resources; CLIR), que es una organización sin fines de lucro que desarrolla estrategias para mejorar los entornos de investigación, enseñanza y aprendizaje en colaboración con bibliotecas, instituciones culturales y comunidades de educación superior.

[2] El RLG fue un consorcio de bibliotecas norteamericanas dedicado a fomentar el acceso a información de aprendizaje e investigación, fundado en 1974 por cuatro prestigiosas bibliotecas de investigación: la Biblioteca Pública de Nueva York y las bibliotecas universitarias de Columbia, Harvard y Yale, hasta su fusión en 2006 con el consorcio Online Computer Library Center (OCLC).

[3] El Task Force on Archiving Digital Information, creado conjuntamente, en 1994, por la CPA y el RLG en los Estados Unidos, se enfocó en el estudio de materiales en formato digital, identificando la necesidad de preservarlos de la obsolescencia tecnológica.

pionero en el ámbito de la preservación digital. El resultado de este estudio fue el informe de 1996, *Preservación de la información digital: informe final y recomendaciones*[4], que propuso dos recomendaciones clave para la preservación a largo plazo de documentos digitales: primero, que los creadores de contenido aseguren la integridad de los objetos digitales a lo largo de su ciclo de vida y, segundo, que se establezca un sistema de certificación para repositorios confiables, garantizando el acceso a la información a largo plazo para futuras investigaciones.

Como señalan Corda *et al*. (2020), las recomendaciones del informe de 1996 sobre preservación digital obtuvieron un amplio respaldo internacional. Organizaciones clave como la Federación Internacional de Asociaciones e Instituciones de Bibliotecarios (International Federation of Library Associations and Institutions; IFLA), el Consejo Internacional sobre Archivos (Internacional Council on Archive; ICA), la Organización Internacional de Normalización (International Organization for Standardization; ISO), la Organización de las Naciones Unidas para la Educación, la Ciencia y la Cultura (United Nations Educational, Scientific and Cultural Organization; UNESCO) y el Centro de Biblioteca Informática en Línea (Online Computer Library Center; OCLC) han apoyado estas directrices, subrayando su importancia global.

A solicitud de la ISO y del Comité Consultivo de Sistemas de Datos Espaciales (Consultative Committee for Space Data Systems; CCSDS), se desarrolló el Modelo de Referencia del Sistema de Información de Archivo Abierto (Open Archival Information System Reference Model; OAIS). Publicado en 1997 y estandarizado como ISO 14721:2012. OAIS se convirtió en un marco de referencia clave para la preservación a largo plazo de datos digitales.

[4] Preserving Digital Information: Final Report and Recommendations https://www.oclc.org/content/dam/research/activities/digpresstudy/final-report.pdf

A principios del siglo XXI, se intensificaron los esfuerzos globales en la preservación digital. En Estados Unidos, la Biblioteca del Congreso (Library of Congress; LC) lanzó en el año 2000 el Programa Nacional de Preservación e Infraestructura de Información Digital (National Digital Information Infrastructure and Preservation Program; NDIIPP), mientras que en Europa surgieron iniciativas como la Red de Acceso y Preservación de Recursos Electrónicos (Electronic Resource Preservation and Access Network; ERPANET), y la Coalición para la Preservación Digital (Digital Preservation Coalition; DPC), enfocadas en la investigación, concientización y desarrollo de estrategias para el acceso y preservación a largo plazo de contenidos digitales (Baucom, 2019).

En 2003, la UNESCO publicó la *Carta para la Preservación del Patrimonio Digital* y sus *Directrices*, impulsando la conciencia global sobre la fragilidad de los objetos digitales. La Carta reconocía la necesidad inmediata de elaborar políticas y estrategias para materiales del patrimonio digital en peligro, incitando a las universidades y otras instituciones de investigación, públicas y privadas, a velar por la preservación de los documentos digitales. Mientras que en las *Directrices* se define a la preservación digital como un conjunto integral de estrategias, técnicas y procesos que buscan garantizar la longevidad de los objetos digitales. Al alinearse con estándares internacionales, estas prácticas aseguran que los recursos digitales permanezcan accesibles y utilizables a largo plazo, independientemente de los cambios tecnológicos, tanto en el hardware como en el software (UNESCO, 2003).

Entre 2006 y 2010, la Unión Europea financió la Open Preservation Foundation (Open Preservation Foundation; OPF), inicialmente el Proyecto de Preservación y Acceso a Largo Plazo a través de Servicios en Red (Preservation and Long-Term Access Through Networked Services Project; PLANETS)[5] para desarrollar servicios interoperables de preservación digital

[5] Proyecto PLANETS https://www.planets-project.eu/

basados en estándares. Este esfuerzo reunió la experiencia de bibliotecas, archivos nacionales, universidades y empresas tecnológicas de Europa.

En 2013, la UNESCO creó el programa Plataforma para Mejorar la Sostenibilidad de la Sociedad de la Información a Nivel Transglobal (Platform to Enhance the Sustainability of the Information Society Transglobally; PERSIST)[6] cuya misión fue servir de guía para ayudar y fomentar a bibliotecas, archivos y museos a redactar sus propias políticas institucionales sobre preservación digital sostenible a largo plazo.

Tras discusiones en la vigésimo segunda *Conferencia Internacional sobre Preservación Digital* (International Conference on Digital Preservation; iPRES 2015) sobre el almacenamiento para la preservación digital, se elaboraron los Criterios de Almacenamiento para la Preservación Digital (Digital Preservation Storage Criteria), que buscan orientar a las organizaciones en la creación de requisitos de almacenamiento y fomentar la capacitación de profesionales especializados en preservación digital (Schaefer *et al*., 2021).

En el ámbito latinoamericano, la existencia de la Asociación Iberoamericana de Preservación Digital (APREDIG) y el Grupo de Trabajo sobre Gestión y Preservación de Documentos Electrónicos, creado en el marco de la Asociación Latinoamericana de Archivos (ALA), instancias en las que Argentina se involucra con el objetivo de promover políticas de preservación digital a largo plazo, contrasta marcadamente con la limitada implementación efectiva de estas políticas en la mayoría de los países de la región.

Zapata Cárdenas (2023), integrante del Grupo de Trabajo de ALA, subraya la persistente y preocupante disparidad en este aspecto, con Brasil como una excepción destacada. Esta situación implica una vulnerabilidad significativa para la supervivencia a largo plazo del patrimonio digital

[6] UNESCO PERSIST Programme https://unescopersist.org

latinoamericano, originada por la escasa prioridad otorgada por los gobiernos al tema, la consecuente limitación de recursos económicos destinados a la preservación digital, la insuficiencia de marcos regulatorios integrales y las notables carencias en la formación y capacitación especializada de profesionales en preservación digital.

Capítulo 1. El Movimiento de Acceso Abierto y el origen de los repositorios institucionales

El Movimiento de Acceso Abierto surgió formalmente en 2001, durante una reunión en Budapest organizada por el Instituto Sociedad Abierta (Open Society Institute). Académicos, científicos y bibliotecarios se unieron para abordar las restricciones de acceso a los resultados de la investigación científico-académica. El surgimiento y desarrollo del Movimiento de Acceso Abierto se vieron impulsados por la crisis imperante del sistema tradicional de comunicación científica. Algunos de los factores claves de esta crisis fue el monopolio editorial, el aumento de los costos de las suscripciones de revistas electrónicas ("serial crisis"), especialmente perjudicial para países en desarrollo, que limitaban el acceso, la difusión y el impacto de la producción científico-académica, y las crecientes restricciones impuestas por los derechos de autor.

La declaración de la *Iniciativa de Acceso Abierto de Budapest* (Budapest Open Access Initiative; BOAI), de febrero de 2001 durante la Conferencia Mundial sobre la Ciencia (World Conference on Science), al abordar la filosofía del Movimiento de Acceso Abierto, sostiene que el acceso abierto a la literatura científica y académica implica que cualquier persona puede leer, descargar, copiar, distribuir, imprimir, buscar y usar libremente los textos completos de los artículos en Internet, sin restricciones económicas, técnicas o legales; la única limitación es que se debe reconocer la autoría y la integridad del trabajo origina. Además, propone políticas que impulsen la apertura de la comunicación científica y el acceso abierto a los resultados de las investigaciones, partiendo de la base de que el conocimiento científico es un bien público. En este sentido, en consonancia con Hernández Pérez *et al.* (2007), la divulgación y el acceso a los resultados de la investigación son fundamentales para el desarrollo económico, científico y social de las

naciones, dando lugar a una economía del conocimiento con repercusiones globales.

A la Declaración de Budapest le siguieron otras iniciativas: la *Declaración de Bethesda sobre el Acceso Abierto* (Bethesda Statement on Open Access*;* 2003), la *Declaración de Berlín sobre el Acceso Abierto al Conocimiento en las Ciencias y las Humanidades* (Berlin Declaration on Open Access to Knowledge in the Sciences and Humanities*;* 2003), los *Principios de Washington DC para el Libre Acceso a la Ciencia* (Washington DC Principles for Free Access to Science; 2004) y la *Declaración de San Francisco sobre la Evaluación de la Investigación* (San Francisco Declaration on Research Assessment; DORA, 2012). Siguiendo los principios de la BOAI, estas iniciativas proponen un paradigma de difusión del conocimiento que, además de ser sostenible económicamente para las instituciones, maximice la visibilidad y el impacto de su producción científica.

El Movimiento de Acceso Abierto se sustenta principalmente en dos modelos complementarios que garantizan el acceso y la difusión de la información científica: la ruta dorada y la ruta verde. La *Ruta Dorada* implica la publicación de artículos científicos en revistas de acceso abierto, lo que garantiza su disponibilidad gratuita e inmediata en Internet, fomentando así una mayor difusión y reutilización del conocimiento. Por su parte, la *Ruta Verde* consiste en el auto-archivo de artículos científicos en repositorios digitales de acceso abierto, este modelo invita a los autores a depositar sus trabajos en dichos repositorios, garantizando una mayor difusión y preservación de la producción científica, facilitando el acceso a la investigación y promoviendo el avance del conocimiento.

Stevan Harnad, figura prominente en el impulso del Movimiento de Acceso Abierto, junto a Peter Suber en los Estados Unidos, propuso por primera vez la práctica del auto-archivo en su influyente artículo de 1994, *The Subversive Proposal*. En dicho texto, Harnad argumenta que la “ruta verde”

representa la vía más factible hacia el acceso abierto, al no exigir una transformación radical del sistema de edición científica.

En el marco de la ruta verde, los repositorios digitales se clasifican en temáticos e institucionales. Los r*epositorios temáticos* se definen como aquellos que albergan documentos científico-académicos pertenecientes a una misma disciplina o área de conocimiento. Por su parte, los *repositorios institucionales* recopilan la producción científica generada por una o varias instituciones, siendo comúnmente implementados en universidades, donde las bibliotecas desempeñan un papel fundamental en la gestión, acceso y preservación del trabajo académico de la comunidad universitaria.

Si bien la ruta verde y la ruta dorada son las principales expresiones del acceso abierto, también se pueden mencionar: la *ruta negra*, compuesta por redes sociales académicas y sitios de acceso abierto, pero ilegales a las obras, como Sci-Hub; la *ruta de bronce*, que incluye artículos o revistas de lectura gratuita, pero con una licencia identificable; y la *ruta diamante o platino*, que se refiere a las revistas que publican en acceso abierto y que no cobran a los autores por publicar ni a los lectores por leer y que, generalmente, son financiadas por instituciones académicas, gubernamentales o sociedades científicas.

Los repositorios institucionales (RI) se sustentan en la filosofía del Movimiento de Acceso Abierto y tienen como objetivo recopilar, procesar, almacenar y difundir el conocimiento que se genera en el marco de una o más instituciones académicas o de investigación, logrando incrementar la visibilidad de la producción científico-académica de estas instituciones. Se demanda a nivel internacional que los RI no solo sean usables y visibles, sino que también cuenten con políticas institucionales que aseguren su preservación a largo plazo.

Incorporando algunos conceptos de Suber (2006), Alonso Arévalo *et al*. (2008) y De Giusti (2014), los RI se caracterizan por ser una colección de

objetos digitales basada en la Web, que ofrecen el libre acceso a sus contenidos, los cuales pueden estar compuestos por artículos de revistas científicas, tesis y disertaciones, objetos de aprendizaje, archivos de datos, archivos multimedia (audio y video), memorias, libros y capítulos de libros, entre otros tipos de documentos digitales. Su naturaleza institucional se encuentra enmarcada en una organización educativa o de investigación científica y su carácter es acumulativo y perpetuo.

Es importante señalar la Iniciativa de Archivos Abiertos (Open Archives Initiative; OAI)[7] que ha establecido el Protocolo de la Iniciativa de Archivos Abiertos para la Recolección de Metadatos (Open Archives Initiative Protocol for Metadata Harvesting; OAI-PMH) como un estándar fundamental de interoperabilidad entre repositorios digitales. Este protocolo facilita la diseminación de contenidos al permitir el acceso y la recolección de metadatos, lo que a su vez incrementa significativamente la visibilidad y el impacto de la investigación científico-académica.

1.1. Los repositorios institucionales en las universidades argentinas

A principios del siglo XXI, las bibliotecas universitarias, en su misión de apoyo a la docencia y la investigación, iniciaron la implementación de repositorios institucionales como un servicio esencial para sus usuarios. Este desarrollo, impulsado por el Movimiento de Acceso Abierto, permitió a las bibliotecas universitarias reafirmar su rol como centros de conocimiento, extendiendo su alcance a través de la difusión y preservación digital de la producción académica de sus comunidades.

La adopción de los principios del Movimiento de Acceso Abierto al conocimiento científico por parte de la comunidad bibliotecaria académica

[7] Open Archives Initiative Protocol for Metadata Harvesting (OAI-PMH) https://www.openarchives.org/pmh/

argentina impulsó la creación del Sistema Nacional de Repositorios Digitales (SNRD) en 2011, por iniciativa del Ministerio de Ciencia, Tecnología e Innovación (MINCyT). Formalizado mediante la Resolución 469 del mismo año, el SNRD establece una red interoperable de repositorios digitales, fundamentada en políticas, estándares y protocolos comunes. En 2015, el SNRD publicó la última versión de sus *Directrices para proveedores de contenido del Sistema Nacional de Repositorios Digitales*[8], basándose en las Directrices DRIVER para proveedores de contenido[9], que se sustentan en el protocolo de interoperabilidad OAI-PMH para recolectar recursos almacenados en repositorios abiertos. Por otra parte, las directrices del SNRD garantizan la integración con el proyecto Infraestructura de Acceso Abierto para la Investigación en Europa (Open Access Infrastructure for Research in Europe; OpenAIRE)[10] y la Red Latinoamericana para la Ciencia Abierta LA Referencia[11].

Por otra parte, el 13 de noviembre de 2013, se promulgó la Ley Nacional N° 26.899, titulada "Sistema Nacional de Ciencia, Tecnología e Innovación. Repositorios Digitales Institucionales de Acceso Abierto", que estableció la obligatoriedad para los organismos e instituciones públicas que integran el Sistema Nacional de Ciencia, Tecnología e Innovación (SNCTI) y reciben financiamiento estatal, de desarrollar repositorios digitales institucionales de acceso abierto. Estos repositorios debían albergar la producción científico-tecnológica resultante de trabajos, formación y proyectos financiados total o parcialmente con fondos públicos, incluyendo las contribuciones de investigadores, tecnólogos, docentes, becarios de posdoctorado y estudiantes de maestría y doctorado (Ley 26.899, art. 1). Conforme al anexo de la Resolución 753 de noviembre de 2016, los

[8] Directrices SNRD https://repositoriosdigitales.mincyt.gob.ar/files/Directrices_SNRD_2015.pdf
[9] Directrices DRIVER https://upcommons.upc.edu/bitstream/handle/2117/1998/prats_directivasdriver%20.pdf
[10] OpenAIRE Interoperability Guidelines for Literature Repository Managers https://guiasopenaire4.readthedocs.io/_/downloads/es/latest/pdf/
[11] LAReferencia https://www.lareferencia.info/es/

principios rectores del acceso abierto y el ámbito de aplicación de la Ley 26.899 buscan optimizar el uso de los fondos públicos destinados a la ciencia, la tecnología y la innovación. Esta optimización se sustenta en la capacidad de identificar, localizar y acceder a la producción científica existente, y requiere un enfoque colaborativo para impulsar iniciativas que promuevan el avance del conocimiento y su aplicación en beneficio de la sociedad, la academia, la economía y el sector productivo.

Los organismos públicos, además de la creación de repositorios digitales, están obligados a establecer políticas de gestión y preservación digital a largo plazo que garanticen el acceso a los datos primarios de las investigaciones; esto implica definir quiénes pueden acceder a estos datos, bajo qué condiciones y con qué fines. En síntesis, la implementación de un repositorio institucional de acceso abierto se erige como una medida fundamental para democratizar el acceso a la investigación científica, potenciando la difusión y el impacto de los resultados de investigación, y superando las barreras económicas y legales impuestas por el modelo tradicional de comunicación y publicación científica.

La implementación de un RI representa un avance fundamental hacia la consolidación del acceso abierto a la investigación científica financiada con fondos públicos. No obstante, para garantizar la eficacia de esta iniciativa, resulta imprescindible complementarla con políticas institucionales o mandatos de autoarchivo obligatorio. La ausencia de directrices claras que incentiven o exijan a los investigadores el depósito de sus trabajos en el repositorio puede traducirse en una baja participación, lo que limitaría la difusión y el impacto de la investigación, así como su potencial para estimular nuevas ideas y avances científicos. En este contexto, la propuesta de De Giusti (2014), que aboga por la transformación del sistema de evaluación científica para reconocer y valorar positivamente a los autores que autoarchivan, tanto en el ámbito científico como económico, se revela como un elemento crucial para el fortalecimiento del Movimiento de Acceso Abierto.

1.1.1. Los repositorios institucionales de la Universidad de Buenos Aires

Mediante la Resolución Nº 6323/2013 del Consejo Superior de la UBA se crea el *Repositorio Digital Institucional de la Universidad de Buenos Aires* (RDI-UBA)[12]. Este repositorio es desarrollado y gestionado por la Dirección General del Sistema de Bibliotecas y de Información (SISBI), que depende de la Secretaría de Ciencia y Técnica del Rectorado de la Universidad. Además, el SISBI integra el Comité de Expertos en Repositorios Digitales, cuya función es orientar al Consejo Asesor de la Biblioteca Electrónica de Ciencia y Tecnología[13] y al MinCyT en relación con las políticas destinadas a la mejora permanente del SNRD.

En los fundamentos de la creación del RDI-UBA, la Resolución Nº 6323/13 especifica que la Universidad, como institución pública que compone el Sistema Nacional de Ciencia, Tecnología e Innovación (SNCTI)[14] y que recibe financiamiento del Estado Nacional, debe contar con un repositorio digital institucional de acceso abierto que permita reunir, registrar, divulgar, preservar y dar acceso a la producción intelectual y académica. De esta manera, en el artículo 2 de la resolución citada, se designa a la Dirección General del SISBI como la unidad responsable de gestionar, mantener, organizar y dar tratamiento documental a todas las colecciones que la Universidad genere.

El principal objetivo del RDI-UBA es facilitar que los académicos de la universidad compartan los resultados de investigación en acceso abierto y, en particular, centralizar y organizar el conocimiento generado por la UBA y financiado con fondos públicos, garantizar el acceso al mismo de manera libre y abierta a través de Internet e incrementar la visibilidad, tanto de los

[12] RDI-UBA https://repositoriouba.sisbi.uba.ar
[13] Biblioteca Electrónica de Ciencia y Tecnología https://biblioteca.mincyt.gob.ar/
[14] Actual Consejo Interinstitucional de Ciencia y Tecnología (CICYT) https://www.argentina.gob.ar/ciencia/cicyt, creado mediante Ley Nacional 25.467 del año 2001.

investigadores, como de la universidad como institución líder en conocimiento.

Desde sus orígenes, el RDI-UBA tuvo que contemplar en su implementación una diversidad de situaciones para lograr integrar la totalidad de las bibliotecas de la universidad, ya sea mediante la cosecha de metadatos de las bibliotecas que poseen repositorio propio (utilizando el protocolo OAI-PMH) o la incorporación directa de contenidos de las bibliotecas que carecen de repositorio (Elizalde *et al*.; 2013), la que se realiza mediante un sistema en línea desarrollado por el SISBI para la carga de los metadatos y el depósito de los objetos digitales.

Además de reunir la producción científico académica de las trece Facultades que componen la Universidad, el RDI-UBA incluye los recursos digitales de las bibliotecas del Ciclo Básico Común, del Colegio Nacional de Buenos Aires, de la Escuela Superior de Comercio Carlos Pellegrini, del Instituto de Historia Argentina y Americana Dr. Emilio Ravignani, del Rectorado de la UBA y de la Biblioteca del SISBI.

De las trece Facultades que componen la UBA, ocho cuentan con repositorios institucionales propios, estas son: Agronomía; Arquitectura, Diseño y Urbanismo; Ciencias Económicas; Ciencias Exactas y Naturales; Ciencias Sociales; Ciencias Veterinarias; Derecho; Farmacia y Bioquímica; Filosofía y Letras; Ingeniería; Ciencias Médicas; Odontología y Psicología.

Tabla 1. *Facultades de la UBA con repositorios autónomos*

Facultad	***Repositorio***	***URL***
Agronomía	FAUBA Digital	http://ri.agro.uba.ar

Ciencias Económicas	Biblioteca Digital FCE-UBA	http://bibliotecadigital.econ.uba.ar/econ
Ciencias Exactas y Naturales	Biblioteca Digital FCEN-UBA	https://bibliotecadigital.exactas.uba.ar
Ciencias Sociales	Repositorio Digital	https://repositorio.sociales.uba.ar/
Filosofía y Letras	FILO:Digital	http://repositorio.filo.uba.ar/
Ingeniería	Repositorio Institucional Inga. Elisa Bachofen	https://bibliotecadigital.fi.uba.ar/
Medicina	Memoria Institucional	https://www.fmed.uba.ar/node/928
Odontología	FOUBA Digital	https://repositorio.odontologia.uba.ar/

Por otra parte, las cinco unidades académicas que no disponen de un repositorio institucional autogestionado, pero cuyos contenidos son incorporados al RDI-UBA son: la Facultad de Arquitectura, Diseño y Urbanismo; la Facultad de Ciencias Veterinarias; la Facultad de Derecho; la Facultad de Farmacia y Bioquímica; y la Facultad de Psicología. Independientemente del buscador general del repositorio de la UBA, que integra el acceso al conjunto de recursos que gestiona, cada una de estas instituciones posee un sitio web independiente y un buscador personalizado,

pudiéndose acceder además a sus contenidos por colección o tipología documental.

Tabla 2. *Facultades de la UBA sin repositorios autónomos (acceso a sus colecciones en el RDI-UBA)*

Facultad	Portal en el RDI-UBA
Arquitectura, Diseño y Urbanismo	Colección FADU
Ciencias Veterinarias	Colección Ciencias Veterinarias
Derecho	Colección Derecho
Farmacia y Bioquímica	Colección Farmacia y Bioquímica
Psicología	Colección Psicología

Cabe destacar que los únicos repositorios de la UBA que adhieren al SNRD y que son cosechados por el mismo mediante el protocolo OAI-PMH son: el RDI-UBA, la Biblioteca Digital FCEN-UBA, FAUBA Digital y FILO:Digital.

Capítulo 2. Evaluación y certificación de un repositorio institucional

A principios de los años 2000, surgió el concepto de Repositorio Digital Confiable (Trusted Digital Repository; TDR) que establece criterios para que los RI aseguren el acceso a información fiable a largo plazo. Estos criterios impulsaron el desarrollo de modelos de auditoría y certificación para garantizar la confiabilidad y durabilidad de los repositorios, asegurando el acceso y uso continuo de sus contenidos por parte de las comunidades a las que sirven. La credibilidad de un repositorio se basa en evidencias claras y verificables de sus prácticas, permitiendo a los interesados confiar en la integridad, autenticidad y accesibilidad de los datos durante períodos de tiempo prolongados; esta confianza requiere auditorías y certificaciones periódicas (Lin *et al*., 2020).

De acuerdo a Ross y McHugh (2005), la capacidad de un RI para superar las auditorías depende de su eficacia en la gestión de los riesgos que amenazan la integridad y accesibilidad de los datos a largo plazo. Un TDR debe mantener un marco sólido de atributos organizativos, técnicos, de recursos, legales y de seguridad (Research Libraries Group/Online Computer Library Center, 2002) con el objeto de asegurar su confiabilidad a lo largo del tiempo.

En esta línea, La Alianza de Datos de Investigación (Research Data Alliance; RDA) impulsó los *Principios TRUST*[15] que definen las características esenciales de los TDR. Estos principios, centrados en la transparencia, la responsabilidad, el foco en el usuario, la sostenibilidad y la tecnología, buscan asegurar la gestión y el acceso a los datos a largo plazo.

[15] En español, *confianza*.

Tabla 3. *Los Principios TRUST*

Principio	***Guía para el repositorio***
Transparency (transparencia)	Ser transparente sobre los servicios específicos del repositorio y sobre los depósitos de datos, verificables mediante evidencia de acceso público.
Responsibility (responsabilidad)	Ser responsable de garantizar la autenticidad e integridad de los datos almacenados y de la confiabilidad y persistencia de su servicio.
User Focus (foco en el usuario)	Garantizar que se cumplan las normas de gestión de datos y las expectativas de las comunidades de usuarios.
Sustainability (sostenibilidad)	Sostener los servicios y preservar los datos a largo plazo.
Technology (tecnología)	Proporcionar una infraestructura y capacidades para respaldar servicios seguros, persistentes y confiables.

Nota. Traducido de Lin *et al*. (2020, p. 2)

Los *Principios TRUST* exigen que un repositorio digital se comprometa con el modelo OAIS, demuestre viabilidad organizacional y sostenibilidad financiera, establezca criterios claros basados en normas actualizadas, garantice la idoneidad tecnológica y procedimental, implemente medidas de seguridad robustas, desarrolle y mantenga políticas exhaustivas, y establezca procedimientos detallados para asegurar la consistencia y calidad de la preservación digital (Lin *et al*., 2020).

En síntesis, los RI deben ser flexibles y adaptables ante los cambios tecnológicos. Para ello, la preservación digital requiere una gestión proactiva, capaz de prever y mitigar los riesgos que comprometen la integridad de la información. Para garantizar la transparencia y fomentar la mejora continua, es fundamental que los gestores se anticipen a la obsolescencia tecnológica, asegurando la accesibilidad y la conservación de los datos.

2.1. Auditoría externa

En 2003, el RLG y la Administración Nacional de Archivos y Registros de los Estados Unidos (National Archives and Records Administration; NARA) iniciaron la creación de criterios para evaluar y certificar repositorios digitales confiables. Este esfuerzo culminó en la publicación de la *Lista de Verificación para la Auditoría y Certificación de Repositorios Confiables* (Trustworthy Repositories Audit and Certification: Criteria and Checklist; TRAC)[16] que establece los atributos y responsabilidades esenciales para evaluar tanto la autenticidad y la usabilidad de la información, como la fiabilidad de los repositorios que la contienen. La lista de verificación fue creada tanto para auditar repositorios existentes como para guiar la creación o expansión de nuevos repositorios.

La *Recomendación sobre Prácticas para Sistemas de Datos Espaciales: Auditoría y Certificación de Repositorios Digitales de Confianza* (Recommendation for Space Data System Practices: Audit and Certification of Trustworthy Digital Repositories)[17], que fue elaborada por el CCSDS y adoptada como norma ISO 16363:2012, permite evaluar la eficacia de los sistemas de preservación de un RI y, a su vez, ofrece una herramienta de diagnóstico interno para optimizar la gestión a largo plazo de los activos

[16] Trustworthy Repositories Audit & Certification: Criteria and Checklist (TRAC) https://www.crl.edu/sites/default/files/d6/attachments/pages/trac_0.pdf
[17] Recommendation for Space Data System Practices: Audit and Certification of Trustworthy Digital Repositories https://public.ccsds.org/Pubs/652x0m1.pdf

digitales. Puntualmente, audita y certifica la confiabilidad de los repositorios digitales evaluando las políticas y planes de preservación, la infraestructura tecnológica, la capacitación del personal, la garantía de calidad, la seguridad y la gestión de riesgos ante la obsolescencia.

En Alemania, la Red de Expertos en Almacenamiento a Largo Plazo de Recursos Digitales (Network of Expertise in Long-term STORage of Digital Resources; NESTOR), compuesta por expertos de diversas instituciones, desarrolló el *Catálogo NESTOR de Criterios para Repositorios Digitales de Confianza*[18], estableciendo otro estándar para la certificación de repositorios digitales. Las categorías principales de evaluación de esta herramienta son: la infraestructura organizacional, la gestión de los objetos digitales, la infraestructura técnica y la seguridad de un RI.

2.2. Auditoría interna

A diferencia de las auditorías externas, la literatura destaca las iniciativas de autoevaluación como una herramienta alternativa de gran valor para las organizaciones, permitiendo la identificación de buenas prácticas no implementadas y posibles áreas de mejora.

La Red de Bibliotecas Universitarias Españolas (REBIUN) y la Fundación Española para la Ciencia y la Tecnología (FECYT) desarrollaron una *Guía para la evaluación de repositorios institucionales de investigación*, con el fin de fortalecerlos como "infraestructuras tecnológicas esenciales para los procesos de implementación de las políticas de acceso abierto y ciencia abierta" (Barrueco Cruz *et al*., 2021, p. 5). Esta guía, que se alinea con el *Marco Comunitario de Buenas Prácticas en Repositorios* (Community Framework for Good Practices in Repositories; COAR)[19] y con las *Directrices*

[18] NESTOR Catalogue of Criteria for Trusted Digital Repositories
https://files.dnb.de/nestor/materialien/nestor_mat_08_eng.pdf
[19] COAR Community Framework for Good Practices in Repositories
https://coar-repositories.org/coar-community-framework-for-good-practices-in-repositories/

de Interoperabilidad OpenAIRE para la validación de los metadatos asignados a los objetos digitales, unifica criterios de calidad vinculados con variadas temáticas: visibilidad, políticas, aspectos legales, metadatos descriptivos, interoperabilidad, acceso a los contenidos, registro de la actividad del sistema (logs), estadísticas, seguridad, autenticidad e integridad de los datos, y servicios o funcionalidades de valor añadido.

En esta línea, Serrano Vicente *et al*. (2014) propusieron una herramienta de autoevaluación bajo el título *Indicadores para la evaluación de repositorios institucionales de acceso abierto*, que incluye 32 indicadores clasificados en tecnología, procedimientos, contenidos, marketing y personal. Estos indicadores pueden aplicarse tanto a evaluaciones individuales como comparativas entre grupos de repositorios y su propósito es determinar en qué medida cada criterio cumple con los objetivos definidos en la política institucional, con el fin de identificar áreas de mejora que contribuyan al logro de los objetivos estratégicos de la institución.

En 2019, el consorcio Alianza Nacional para la Gestión Digital (National Digital Stewardship Alliance de los Estados Unidos; NDSA) publicó la última versión del instrumento *Niveles de Preservación Digital*[20], que ofrece una metodología para evaluar el nivel de preservación digital de una institución y la elaboración de un plan de mejoras a partir de las carencias detectadas. La metodología NDSA utiliza una tabla simple con preguntas sobre el sistema de almacenamiento y la ubicación geográfica de los archivos, la alteración e integridad de los datos, las medidas de seguridad de la información, los metadatos y los formatos de archivo. De acuerdo con Término y Leija Román (2017), la presentación de la tabla es intencionalmente sencilla y de fácil comprensión para evitar el efecto de rechazo que provocan muchos sistemas tradicionales de auditoría. Tettamanti, De Giusti y Lira (2022) proponen la autoevaluación de los niveles NSDA como una alternativa eficaz para instituciones con recursos limitados,

[20] NSDA Levels of Digital Preservation
https://ndsa.org//publications/levels-of-digital-preservation/

facilitando un diagnóstico rápido de sus políticas de preservación y preparando el camino hacia la certificación.

Capítulo 3. Preservación digital a largo plazo

El concepto de preservación digital ha sido abordado por diversos autores, Leija Román (2017) incorpora los términos “políticas”, “estrategias” y “acciones técnicas” para definirla como una actividad fundamentalmente de *gestión* de los contenidos digitales, cuyo propósito es mantenerlos disponibles y funcionales, sin importar los cambios tecnológicos o el paso del tiempo. Por su parte, Bodero Poveda *et al*. (2022) amplían este concepto, sosteniendo que la preservación digital se debe realizar en base a las prácticas más confiables de almacenamiento durante todo el ciclo de vida del contenido digital, constituyéndose en la rama de la bibliotecología y las ciencias de la información que se ocupa de mantener la *accesibilidad* a los documentos atendiendo a principios de seguridad, longevidad, calidad, fiabilidad e integridad de la información.

La preservación digital enfrenta desafíos críticos, entre los que destacan la naturaleza intrínseca de los contenidos, el mantenimiento de la confianza y sostenibilidad de los formatos, la obsolescencia tecnológica, la vulnerabilidad de los medios de almacenamiento, las complejidades inherentes a los derechos de autor y la propiedad intelectual, y la correcta asignación de metadatos en consonancia con las directrices y normativas vigentes. Por lo tanto, es fundamental priorizar la implementación de políticas y buenas prácticas de preservación en cada fase del ciclo de vida de la información digital, comprendiendo la adquisición, creación, transformación, catalogación, almacenamiento y acceso.

El enfoque de “largo plazo” en la preservación digital debe considerar un horizonte temporal que contemple el impacto de la incesante evolución de las tecnologías. En este contexto, destaca la visión de Justrell (2006, p. 4), que propone concebir la preservación digital a largo plazo como un “ecosistema dinámico”, es decir, en constante transformación. Esta

perspectiva subraya la necesidad de desarrollar una estrategia continua para gestionar los “cambios perpetuos” que permitan “garantizar la accesibilidad y la usabilidad a largo plazo para apoyar la transparencia en los procesos sociales, el patrimonio cultural y la mejora del conocimiento”. Como se ha señalado, la preservación digital se caracteriza por un “esfuerzo persistente” (Cramer, 2023) que implica un desarrollo continuo y cambiante. Lejos de ser una tarea resuelta, exige una revisión y reinvención constante ya que se intensifica su complejidad con el tiempo. Por lo tanto, es fundamental diseñar, evaluar e implementar políticas alineadas con los estándares y directrices vigentes en el ámbito de la preservación digital a largo plazo.

3.1. El modelo de referencia OAIS

El modelo OAIS es un estándar clave para la preservación digital, que guía la gestión de repositorios al definir cómo se deben preparar, archivar, almacenar, conservar y recuperar los objetos digitales. Su enfoque principal radica en asegurar tanto la *preservación* como el *acceso* a la información a largo plazo. Los RI que adoptan el modelo OAIS ofrecen una gestión y preservación integral de la información a lo largo de todo su ciclo de vida (Ochoa-Gutiérrez *et al*., 2021). Esta innovación se convirtió en una práctica recomendada, estableciendo la organización de personas y sistemas necesarios para preservar y acceder a los objetos digitales de forma efectiva y a largo plazo.

En el modelo OAIS se introdujo por primera vez el concepto de 'paquete de información' (information package, IP) que se refiere a los objetos digitales y a su información de referencia, contexto, procedencia, integridad y acceso (metadatos). Es así como la preservación digital se materializa en dichos 'paquetes de información' que aglutinan el 'contenido de la información' (content information; CI) y la 'descripción de la información para preservación' (preservation description information, PDI) que asegura

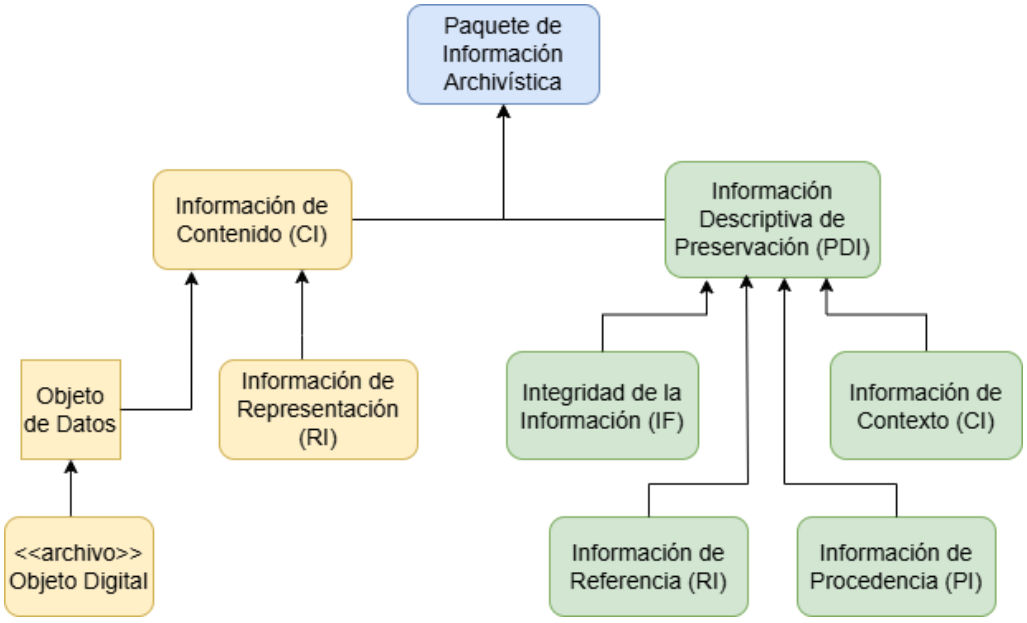
que el contenido esté identificado y se comprenda en el entorno en el que ha sido creado. Ambos, contenido y descripción, se encuentran encapsulados, unidos e identificados por la 'información de empaquetado' (packaging information, PI) y son accesibles a través de la 'información descriptiva' (descriptive information, DI). En resumen, estos paquetes, identificados y accesibles a través de metadatos descriptivos, aseguran que el objeto digital se preserve y sea recuperable para su uso futuro (Cruz Mundet y Díez Carrera, 2016).

El IP, que es el objeto digital en sí mismo, tiene tres variantes según sea su estado. Cuando el objeto digital es transferido por el productor al archivo mediante un 'acuerdo de transferencia' (submission agreement) ingresa al repositorio denominándose 'paquete de transferencia de información' (submission information packages; SIP). Cuando los objetos digitales son almacenados en el repositorio, los SIP se transforman en entidades digitales preservables llamadas 'paquetes de información de archivo' (archival information packages; AIP). OAIS utiliza los términos SIP y AIP para distinguir los objetos digitales que se reciben de los que se preservan, respectivamente. La tercera variante del IP es el 'paquete de diseminación de información' (dissemination information package; DIP) que es el objeto digital proporcionado en respuesta a una solicitud del consumidor, es decir, lo que recupera un usuario al acceder al documento mediante la interfaz de búsqueda de un RI. En síntesis, el modelo define tres estados del IP según su etapa: SIP (al ingresar al repositorio), AIP (al almacenarse) y DIP (al ser recuperado por el usuario). Estos estados distinguen claramente la recepción, preservación y acceso a los objetos digitales

El modelo define metadatos claves para la preservación digital, que incluyen información para representar, documentar, empaquetar y recuperar los objetos archivados. Estos metadatos aseguran la integridad y accesibilidad a largo plazo de los recursos digitales. Como señalan Salvador

Benítez y Ruíz Rodríguez (2005) y la Cornell University Library (2007), la PDI está conformada por cuatro categorías de metadatos de preservación que registran la identidad, las relaciones, el historial y la integridad del objeto digital almacenado en el RI, es decir, la información de referencia, la información de contexto, la información de procedencia y la información de integridad.

Figura 1. *Marco de referencia para la preservación digital de acuerdo al Modelo OAIS*



Nota. Traducido de Rieger (2004, p. 12)

Para garantizar la autenticidad, integridad, accesibilidad y usabilidad de la información digital se deben considerar cuatro tipos esenciales de metadatos. En primer lugar, la 'información de referencia' en OAIS identifica de forma única cada objeto digital, facilitando su gestión y organización

dentro del repositorio. En segundo lugar, la 'información de contexto' describe las relaciones del objeto con su entorno, incluyendo su origen y conexiones con otros objetos, asegurando su comprensión a largo plazo. En tercer lugar, la 'información de procedencia' documenta la historia del objeto digital (sus orígenes, la cadena de custodia y las intervenciones de preservación), el registro de procedencia es esencial para asegurar la trazabilidad y la autenticidad de la información a lo largo del tiempo. Y finalmente, la 'información de integridad' asegura la fiabilidad y el valor de la información digital en el contexto académico y de investigación, para ello se utilizan técnicas, como sumas de comprobación y firmas digitales, que verifican que el objeto digital no haya sido alterado.

El modelo involucra cuatro actores y seis funciones principales para la preservación digital. Los actores (productores, consumidores, gestores y el sistema) ejecutan las funciones de ingesta, almacenamiento, gestión de datos, administración, planificación de la preservación y acceso. La 'ingesta' en OAIS es el proceso de recepción y preparación de los objetos digitales para su preservación en el repositorio, incluyendo controles de procedencia, antivirus y formatos. El 'almacenamiento de archivos' gestiona la preservación física y recuperación de los objetos digitales (AIPs), implementando estrategias para asegurar su acceso a largo plazo. La 'gestión de datos' administra los metadatos generados durante la ingesta y el ciclo de vida del documento, mientras que la 'administración' coordina la operación general del repositorio, integrando funciones, actores y tecnologías intervinientes. Por último, la 'planificación de la preservación' define políticas y estrategias ante los cambios tecnológicos, mientras que el 'acceso' proporciona la interfaz para la consulta del repositorio digital.

3.2. Estrategias para la preservación digital

Una preservación digital efectiva demanda un plan estratégico integral que defina políticas para la accesibilidad y legibilidad a largo plazo. Este plan debe articular una combinación de herramientas y técnicas, tales como migración de formatos, emulación de programas, replicación de datos, encapsulado, preservación tecnológica, estándares, metadatos, arqueología digital y control de autenticidad, para asegurar la permanencia, integridad y fiabilidad de los objetos digitales.

La “migración”, como técnica de preservación digital, asegura la accesibilidad y usabilidad a largo plazo de los objetos digitales mediante la transferencia de formatos a plataformas estables y actualizadas para responder a la evolución tecnológica. La migración enfrenta críticas debido a algunos riesgos inherentes como la potencial pérdida de datos (Granger, 2000) por lo que algunos expertos consideran la emulación como una alternativa más fiable.

La “emulación” consiste en crear programas que reemplacen y reproduzcan el comportamiento de tecnologías de hardware y software antiguas u obsoletas. Para lograrlo, es necesario guardar tanto el programa que imita (emulador), como el sistema operativo y los archivos originales (Bia Platas y Sánchez Quero, 2002). Sin embargo, algunos autores señalan la obsolescencia de los emuladores como un riesgo para la representación fiable de los objetos digitales a largo plazo.

La “replicación” es la creación de copias idénticas de objetos digitales para su restauración en caso de fallos de hardware, software, red, errores humanos, desastres naturales o ciberataques. Para mitigar estos riesgos, se recomienda un sistema de múltiples copias distribuidas en ubicaciones geográficas seguras. Cuando la replicación se realiza debido al deterioro u obsolescencia de los medios de almacenamiento, se denomina

"rejuvenecimiento" (refreshing), implicando la transferencia de los objetos digitales a nuevos soportes. Según Lee *et al*. (2002), la replicación es una estrategia a corto plazo, por lo que se recomienda una replicación frecuente, idealmente diaria, para mitigar riesgos inmediatos.

El “encapsulado” es una estrategia de preservación que agrupa un objeto digital con sus metadatos, especificaciones de formato y toda documentación necesaria para su decodificación futura (El Idrissi, 2019). La inclusión de metadatos de representación, procedencia, fijeza y contexto asegura la autenticidad de los datos.

La “preservación de la tecnología” busca combatir la obsolescencia de los soportes mediante la conservación del entorno tecnológico original (software, hardware y sistema operativo). Esta estrategia, costosa y compleja para instituciones individuales, requiere inversión significativa en equipos y personal especializado, por lo que se sugiere la creación de redes de cooperación como alternativa viable. Formenton y De Souza Gracioso (2020) la consideran una solución temporal y de corto plazo.

La “confianza en los estándares” implica el uso de formatos y software abiertos para mitigar la obsolescencia tecnológica. Se persigue una representación de contenido independiente de software propietario, basada en código abierto y adaptable a la tecnología. Para asegurar la integridad de los objetos digitales, la estrategia también requiere el cumplimiento de estándares estables a largo plazo, como esquemas de metadatos y normas de autenticidad y sostenibilidad de los formatos. La puesta en práctica de la confianza en los estándares es la “normalización”, que como estrategia de preservación implica la elección de un formato para cada grupo de archivos del mismo tipo, debiéndose optar por aquel que ofrezca mejores posibilidades de longevidad y funcionalidad.

La adopción de “metadatos de preservación” es una estrategia crucial para garantizar la longevidad de los objetos digitales. Esto implica asignar

metadatos descriptivos para la localización y acceso, metadatos legales para la gestión de derechos, metadatos administrativos para la integridad y autenticidad, y metadatos técnicos para asegurar la legibilidad y accesibilidad a largo plazo.

La “arqueología digital” recupera datos de software o hardware obsoletos, dañados o en desuso, aplicando técnicas para acceder a información que se ha vuelto ilegible. Sin embargo, la falta de garantías en la interpretación de los datos compromete su integridad y autenticidad. Más que una estrategia de preservación preventiva, es principalmente un recurso de emergencia para la recuperación de datos.

La vulnerabilidad de la información digital exige estrategias de “control de autenticidad e integridad” para garantizar que los documentos no experimenten una alteración que cambie su significado. Esto se logra mediante metadatos de preservación, registro de procedencia y transformaciones, marcas de agua o firmas digitales, sumas de verificación (checksum) e identificadores únicos y persistentes. Estas técnicas, complementarias entre sí, aseguran la inmutabilidad y la localización inequívoca de los objetos digitales.

3.3. Metadatos de preservación digital

Los metadatos son datos que describen otros datos y tienen la capacidad de proveer información necesaria para identificar un recurso de información. Senso y De la Rosa (2003) amplía este concepto indicando que, además de identificar un documento, los metadatos sirven para describirlo, contextualizarlo, autentificarlo, recuperarlo y establecer sus condiciones de uso e interoperabilidad.

De acuerdo a la función que cumplen dentro de un sistema de información, los metadatos se clasifican en cuatro categorías. Los *metadatos*

descriptivos son aquellos que describen los recursos de información para contribuir a su recuperación. Los *metadatos administrativos* facilitan la gestión de recursos digitales mediante datos técnicos sobre su creación, derechos, preservación, acceso y calidad. Los *metadatos estructurales* describen la organización interna de un recurso digital, incluyendo sus componentes y relaciones lógicas con otros recursos. Los *metadatos técnicos* describen las características operativas de los objetos digitales, incluyendo requisitos de software y hardware, formatos y firmas digitales, entre otros.

En cuanto a los *metadatos de preservación*, su función no se limita a un sistema de información, sino que abarca un propósito más amplio, que es garantizar la disponibilidad, identidad, persistencia, capacidad de representación, comprensión y autenticidad de los objetos digitales a largo plazo. Lavoie y Gartner (2013) proponen que la estructura de los metadatos de preservación se fundamenta en tres entidades primordiales: la procedencia, para asegurar la autenticidad e integridad del objeto digital; la gestión de derechos, para detallar las restricciones de uso; y el entorno tecnológico, para definir los requisitos técnicos de acceso.

Respecto al modelo de referencia OAIS, la descripción de la PDI se centra en la información necesaria para gestionar la “perpetuidad” de los objetos digitales, registrando la identidad, las relaciones, el historial y la integridad de los mismos. Sin embargo, si bien OAIS ofrece un marco conceptual fundamental para la preservación digital, no puntualiza qué metadatos se deben recopilar ni cómo se deben implementar para respaldar los objetivos de la preservación (Dappert y Enders, 2010). La falta de una estructura de datos definida en el modelo representa una limitación que traslada la responsabilidad a los gestores de información, quienes deben discernir y adaptar los metadatos de preservación al contexto de su institución, así como establecer los procedimientos de implementación.

Para subsanar esta deficiencia, resulta indispensable complementar el modelo OAIS con estándares y directrices adicionales que proporcionen guías prácticas para la aplicación efectiva de los metadatos de preservación. Entre estos recursos, destacan Dublin Core, los Metadatos de Preservación: Estrategias de Implementación (PREservation Metadata: Implementation Strategies; PREMIS) y el Estándar de Codificación y Transmisión de Metadatos (Metadata Encoding and Transmission Standard; METS).

3.3.1. El Diccionario de Datos PREMIS

Con el objetivo de establecer un marco de referencia común para la gestión y preservación de objetos digitales a largo plazo, el OCLC y el RLG patrocinaron la creación del equipo de trabajo PREMIS, que tenía como objetivo principal establecer un conjunto de estándares que permitieran caracterizar las propiedades de los objetos digitales, registrar las acciones de preservación y mantener un catálogo detallado de los agentes involucrados en este proceso.

Dicho grupo de trabajo, integrado por expertos internacionales en materia de preservación digital, desarrolló el Diccionario de Datos PREMIS para Metadatos de Preservación[21], que es “un esquema de metadatos específicos para preservación que permiten seguir el ciclo de vida de los objetos digitales y garantizar su accesibilidad futura, su correcta interpretación, su autenticidad y su integridad” (De Giusti, 2020, p. 11). En el diccionario, los metadatos de preservación engloban las categorías de metadatos administrativos (incluidos derechos y permisos), técnicos y estructurales, prestando especial atención a la *procedencia* o historia de los objetos digitales y a las relaciones entre los distintos objetos dentro del repositorio.

[21] PREMIS Data Dictionary for Preservation Metadata
https://www.loc.gov/standards/premis/v3/

El Diccionario PREMIS está organizado como un modelo de datos integrado por cinco elementos: objetos, entidades intelectuales, eventos, agentes y derechos. Los “objetos” son unidades de información (entidades intelectuales, archivos, flujos de bits) sujetas a “eventos” (acciones de preservación) en entornos tecnológicos (hardware/software). Los “agentes” son personas u organizaciones y los “derechos” definen los permisos de uso del contenido (Lavoie y Gartner, 2013). Cada una de estos elementos se describen mediante un conjunto de propiedades llamadas “unidades semánticas” que son los atributos que describen las características de los objetos digitales y sus relaciones contextuales.

Un aspecto significativo del modelo de datos PREMIS son las relaciones entre objetos, agentes y eventos, es decir, la “procedencia digital”, que constituyen un registro completo de la historia de un objeto digital o de varios objetos interconectados, desde su creación hasta cualquier transformación posterior, incluyendo la cadena de custodia y cómo se relacionan con otros objetos en el tiempo. En resumen, la procedencia trasciende la mera descripción de objetos digitales aislados, sino que además documenta las interrelaciones entre múltiples objetos.

3.3.2. El estándar de metadatos METS

El estándar de metadatos METS[22] surgió como una iniciativa de la Federación de Bibliotecas Digitales (Digital Library Federation; DLF) que tenía como objeto establecer un formato común para describir y gestionar objetos digitales. En la actualidad, bajo la tutela de la Biblioteca del Congreso de los Estados Unidos, METS se ha consolidado como un estándar internacional que garantiza la preservación a largo plazo y la interoperabilidad de los recursos digitales.

[22] Metadata Encoding and Transmission Standard
https://www.loc.gov/standards/mets/mets-schemadocs.html

METS es una implementación en XML diseñada para actuar como un SIP, un DIP o un AIP del modelo de referencia OAIS, y permite registrar cuatro tipos de metadatos, cada uno dentro de una sección propia de un archivo METS: un inventario de archivos asociados con el objeto digital (como archivos de imagen, archivos de texto, video o audio); una sección para metadatos administrativos, dividida en cuatro subsecciones, información técnica sobre los archivos, información de gestión de derechos, información sobre la fuente sobre la cual se hizo el objeto e información de procedencia digital; una sección para metadatos descriptivos; y un mapa estructural de los contenidos internos del ítem, que indica de manera jerárquica cómo se relacionan entre sí sus diversos componentes (Lavoie y Gartner, 2013). El principal objetivo de METS es proporcionar un mecanismo para registrar las diversas relaciones que existen entre los elementos de contenido y entre el contenido y los metadatos que componen un objeto digital en un repositorio.

Un archivo METS consta de siete secciones principales (Library of Congress, 2022): un *encabezado*, que contiene metadatos con información sobre quién, cuándo y para qué fue creado el objeto digital; una *sección de metadatos descriptivos*, que describen la información representada por el objeto y permiten su descubrimiento; un *mapa estructural*, que indica de manera jerárquica cómo se relacionan entre sí los diversos componentes del objeto y los metadatos relativos al mismo; una *sección de metadatos descriptivos*, que puede contener metadatos descriptivos externos al documento METS o metadatos descriptivos encapsulados dentro del archivo; una *sección de metadatos administrativos*, que proporciona información sobre los derechos de propiedad intelectual e información sobre la procedencia del objeto digital; una *sección de archivos de contenido*, que enumera los archivos cuyo contenido incluye versiones electrónicas del objeto digital, lo que permite subdividir los archivos por versión del objeto; y una *sección de metadatos administrativos*, que contiene información sobre los objetos digitales de la sección de archivos de contenido.

Esta última sección se subdivide, a su vez, en *metadatos técnicos* que especifican información relativa a la creación del archivo, su formato y características de uso; *metadatos de origen* que especifican metadatos descriptivos y administrativos sobre el documento origen a partir del cual se ha generado el objeto digital; *metadatos de procedencia digital*, que especifican los cambios que el archivo ha sufrido desde su origen; y *metadatos de derechos* que detallan las condiciones de acceso legal al objeto digital (copyright e información sobre licencias).

La integración de PREMIS en METS es viable, aunque requiere una consideración cuidadosa de la ubicación de ciertas entidades para asegurar la correcta representación de la información. El Diccionario de Datos PREMIS, en su mayoría, puede integrarse en la sección de metadatos administrativos de un archivo METS.

Así es como las entidades de *evento* y *derechos* de PREMIS se adaptan a las subsecciones de procedencia digital y derechos de METS, respectivamente, mientras que la entidad *objeto* puede ubicarse dentro de los metadatos de procedencia. Por otro lado, la entidad *agente* puede hacer referencia a información tanto de procedencia digital como de derechos, lo que requeriría su ubicación en las subsecciones correspondientes según el contexto.

Capítulo 4. Formatos de información digital

La comprensión del formato de un archivo digital es fundamental para su preservación y para la interpretación de su contenido. Un formato define las reglas sintácticas (estructura y organización de los datos) y semánticas (significado e interpretación de los datos) para el mapeo de un modelo de información a un flujo de bytes y el mapeo inverso de ese flujo de bytes al modelo de información original (Hedstrom y Lee, 2002; Abrams y Flecker, 2005). Según el modelo OAIS, la identificación del formato, es decir, la *información de representación* de un objeto digital, es la clave que convierte datos binarios en información comprensible. La tipificación e interpretación precisa del formato resulta esencial para el uso y el intercambio de la información codificada digitalmente, garantizando que los objetos digitales permanezcan comprensibles y utilizables a largo plazo.

Un formato de archivo digital se define mediante una especificación formal, un documento que detalla su estructura y permite tanto la creación de archivos válidos como el desarrollo de software para su interpretación. La disponibilidad de especificaciones detalladas tiende a ser mayor en formatos ampliamente utilizados. En palabras de Barve (2007, p. 242), “una especificación de formato de archivo indica la subdivisión, codificación, secuencia, disposición, tamaño y relaciones internas adecuadas que identifican de forma única el formato en particular y permiten que se lo interprete y represente correctamente”.

En 2003, bibliotecas y organismos de normalización de Estados Unidos y Europa iniciaron el proyecto Registro Global de Formatos Digitales (Global Digital Format Registry; GDFR) para centralizar el control de especificaciones de formatos digitales. En esencia, el proyecto GDFR fue un esfuerzo colaborativo para crear un sistema centralizado que facilitara la gestión y el acceso a la información sobre formatos digitales, crucial para la

preservación digital. La propuesta dio como resultado un modelo de datos detallado y un modelo de registro basado en la gobernanza compartida, la contribución cooperativa y el alojamiento distribuido de datos (Abrams, 2005). Paralelamente, los Archivos Nacionales del Reino Unido (The National Archive; TNA) desarrollaron PRONOM[23], un registro de formatos digitales de acceso libre, similar al GDFR. Pero ninguno de los dos registros cumplía completamente las necesidades de preservación: PRONOM tenía información técnica, pero faltaba la gobernanza compartida del GDFR. Se intentó combinar ambos en el Registro Unificado de Formatos Digitales (Unified Digital Format Registry; UDFR), pero el proyecto se discontinuó. Como resultado, se priorizó el uso del registro técnico PRONOM.

Otra iniciativa de este tipo es el Entorno de Validación de Objetos JSTOR/Harvard[24] (JSTOR/Harvard Object Validation Environment; JHOVE), que es una herramienta de código abierto desarrollada por la Biblioteca de Harvard que automatiza la identificación, validación y caracterización de archivos digitales. Actualmente mantenida por la OPF, extrae información técnica de formatos de archivo comunes, lo que facilita la preservación digital.

Un registro público de especificaciones de formatos digitales es esencial para la preservación digital. Este registro proporciona información detallada (metadatos descriptivos, administrativos y técnicos) sobre los formatos, lo que facilita la identificación, validación, procesamiento y gestión de riesgos de los objetos digitales a largo plazo.

4.1. Identificación, selección y evaluación de formatos

Para una preservación digital a largo plazo efectiva, es crucial identificar y seleccionar formatos de archivo que garanticen accesibilidad,

[23] PRONOM https://www.nationalarchives.gov.uk/pronom/
[24] JHOVE https://jhove.openpreservation.org/

autenticidad, inteligibilidad, integridad y longevidad. Como indica Moro Cabero (2018), la abundancia y diversidad de formatos constituye uno de los principales desafíos que enfrentan actualmente los gestores de información digital, dificultando su control y comprometiendo su interoperabilidad, portabilidad, replicación, compatibilidad y mantenimiento a largo plazo. Dada la rápida evolución y obsolescencia de los formatos digitales, se requiere la exploración y aplicación de herramientas que faciliten la evaluación y selección de aquellos formatos que aseguren la disponibilidad y reutilización de la información a través del tiempo.

La identificación precisa, la validación y el conocimiento de las normas de los formatos digitales son esenciales para garantizar el acceso y uso a largo plazo de los objetos digitales. Los profesionales de la preservación digital requieren el uso de herramientas y registros técnicos especializados en la identificación y validación de formatos. Algunas herramientas para reconocer formatos digitales son: PRONOM, JHOVE, DROID, Metadata Extraction Tool[25] (de la Biblioteca Nacional de Nueva Zelanda), Format Identification for Digital Objects (FIDO)[26], File Information Tool Set (FITS)[27], Community Owned Digital Preservation Tool Registry (COPTR)[28], NARA File Analyzer and Metadata Harvester[29], FILEINFO: the File Format Database[30], ExifTool[31], FFmpeg[32], entre otras.

Moro Cabero (2018) propone la creación de “perfiles de formatos” detallados para mejorar la identificación y selección de formatos digitales para la preservación. Estos perfiles deben incluir información exhaustiva sobre cada formato, abarcando desde sus especificaciones técnicas y versiones, hasta su naturaleza (abierto o propietario), riesgos, licencias, y su

[25] Metadata Extraction Tool https://meta-extractor.sourceforge.net/
[26] FIDO https://openpreservation.org/tools/fido/
[27] FITS https://github.com/harvard-lts/fits
[28] COPTR https://coptr.digipres.org/index.php
[29] NARA File Analyzer and Metadata Harvester https://github.com/usnationalarchives/File-Analyzer
[30] FILEINFO https://fileinfo.com/
[31] ExifTool https://exiftool.org/
[32] FFmpeg https://ffmpeg.org/

viabilidad a largo plazo. La finalidad es proporcionar una guía completa para la gestión y preservación efectiva de los activos digitales.

En la gestión de activos digitales, es crucial diferenciar entre formatos de preservación y difusión. Para la preservación a largo plazo, se deben seleccionar formatos de alta calidad y sin pérdida de información para los archivos maestros y copias de seguridad. En contraste, para el acceso y difusión al usuario final, se recomiendan formatos de calidad media/baja, con posible pérdida de información y resolución, optimizando así el acceso sin comprometer la integridad de los originales.

La rápida obsolescencia de los formatos digitales representa un desafío para la preservación a largo plazo en los RI. Se recomienda adoptar formatos abiertos, no propietarios, estandarizados y bien documentados para garantizar el acceso continuo a los contenidos. Minimizar la dependencia de software y hardware propietarios, como sugiere Brown (2008), reduce la necesidad de migraciones y prolonga la vida útil de la información digital. Sin embargo, la decisión entre formatos de código abierto y propietarios en la preservación digital no es sencilla. Según la DPC, aunque los formatos propietarios como TIFF son robustos, corren el riesgo de obsolescencia si sus propietarios los descontinúan. Por otro lado, los formatos de código abierto, aunque tecnológicamente neutrales, dependen del apoyo y la estabilidad de sus comunidades de desarrollo, lo que también puede generar incertidumbre en su viabilidad a largo plazo.

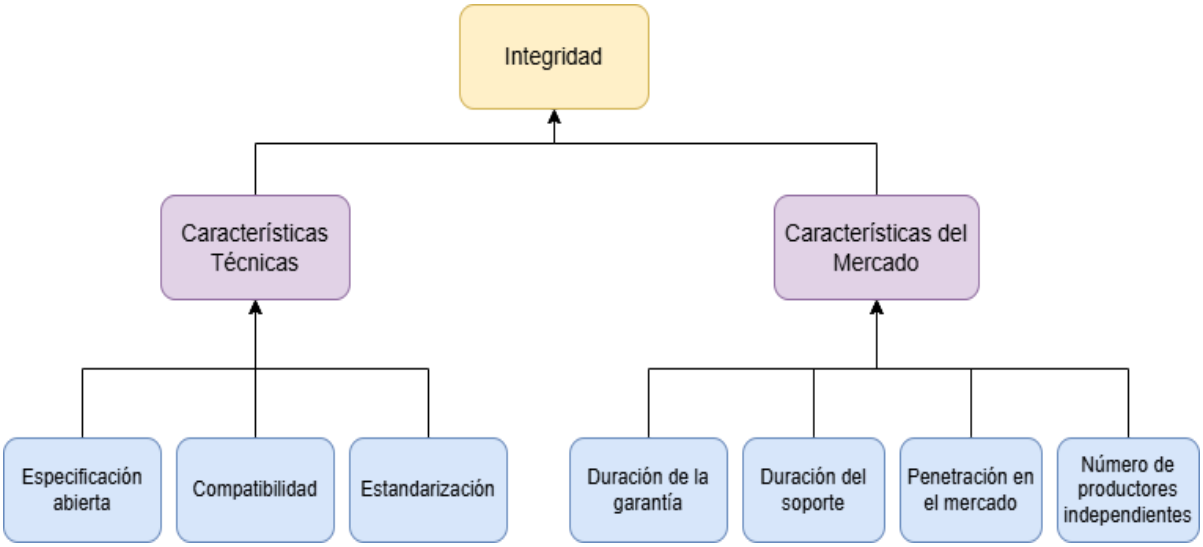
A pesar de las complejidades en la elección de formatos, multiplicidad de autores coinciden en que los formatos de preservación deben ser preferentemente abiertos, sin restricciones de patente o licencia, y estandarizados por organismos reconocidos. La adopción de estándares abiertos es crucial para asegurar la viabilidad a largo plazo de los formatos, ya que garantiza su independencia del medio de almacenamiento, hardware o software propietario. Para evaluar, seleccionar y adoptar un estándar abierto Park y Oh (2012) identifican una serie de criterios que deberían

considerarse: el grado de adopción; la independencia en relación a la aplicación que lo ejecuta; su divulgación y transparencia; su reutilización e interoperabilidad; la robustez, complejidad, viabilidad y estabilidad; junto a la propiedad intelectual y la gestión de derechos.

Otro factor importante a considerar para seleccionar formatos de archivo es su capacidad para embeber metadatos. Estos, ya sean generados automáticamente, ingresados manualmente o una combinación de ambos, enriquecen significativamente los recursos digitales al proporcionar información detallada sobre su origen, interoperabilidad y características técnicas, así como descripciones contextuales. Dentro de las aplicaciones de código abierto para la gestión de metadatos en PDF, se encuentra AutoMetadata, que facilita la edición e incorporación por lote. Otras herramientas disponibles gratuitamente son PDF Shaper y Hexonic PDF Metadata Editor, este último también permite la conversión de imágenes JPEG, TIFF, BMP, PNG y GIF a documentos PDF.

Rauch *et al.* (2007) proponen evaluar la “integridad” a largo plazo de los formatos digitales mediante la consideración de características técnicas y de mercado. Las *características técnicas* incluyen la apertura y disponibilidad de las especificaciones, la compatibilidad, el mantenimiento por empresas de software y la estandarización por organismos reconocidos como ISO. Las *características del mercado* se refieren a la aceptación del formato, la duración del soporte, la penetración en el mercado y la diversidad de productores independientes.

Figura 2. *Criterios para evaluar la integridad de los formatos a largo plazo*



Nota. Traducido de Rauch *et al*. (2007, p. 104)

La gestión de riesgos es fundamental para la preservación de los formatos digitales, dada la amenaza del envejecimiento de la información. La evaluación de riesgos se centra en las características técnicas del formato, los elementos de software asociados y la influencia de los propietarios o proveedores de contenido, permitiendo tomar decisiones informadas para garantizar la accesibilidad a largo plazo. En este sentido, la Universidad de Cornell (Estados Unidos), con el apoyo del Consejo de Bibliotecas y Recursos de Información (Council on Library and Information Resources; CLIR), desarrolló un modelo de evaluación de riesgos para la preservación digital, basado en dos escalas: probabilidad e impacto (Lawrence *et al*., 2000). Estas escalas miden la exposición al riesgo mediante la combinación de ambas variables, utilizando un sistema de cinco puntos con etiquetas, valores y descripciones detalladas. Este modelo permite evaluar de manera estructurada la vulnerabilidad de los objetos digitales en diferentes formatos.

Tabla 4. *Escala de Probabilidad de Riesgos*

Etiqueta	Valor	Descripción
Muy alta	5	Una probabilidad estimada entre 26-99%
Alta	4	Una probabilidad estimada entre 11-25%
Moderada	3	Una probabilidad estimada entre 6-10%
Baja	2	Una probabilidad estimada entre 1-5%
Muy baja	1	Una probabilidad estimada de 1%

Nota. Traducido de Lawrence *et al*., (2000, p. 24)

Tabla 5. *Escala de impacto del riesgo*

Etiqueta	Valor	Descripción
Catastrófico	E	Pérdida total e irreversible de datos. No es posible obtener datos de otras fuentes (impresas o digitales).
Muy serio	D	Pérdida parcial e irreversible de datos. No es posible obtener datos de otras fuentes.
Serio	C	Pérdida total de datos. Los datos pueden reconstruirse completamente a partir de otras fuentes.
Significativo	B	Pérdida parcial de datos. Los datos pueden reconstruirse completamente a partir de otras fuentes.

Menor	A	Pérdida total o parcial de datos. Los datos pueden copiarse de otros archivos de datos.

Nota. Traducido de Lawrence *et al*. (2000, p. 25)

La ponderación de la *probabilidad* se realiza desde valores muy bajos hasta valores muy altos y el *impacto* se mide desde un nivel leve de pérdida de datos menor o insignificante hasta otro nivel catastrófico de pérdida de datos total o irreversible. Lo que finalmente se registra es el valor de probabilidad de riesgo junto al valor de impacto como un valor único:

5E = Probabilidad muy alta con un impacto catastrófico

3D = Probabilidad moderada con un impacto muy grave

2C = Probabilidad baja con un impacto grave

1B = Probabilidad muy baja con un impacto significativo

1A = Probabilidad muy baja con un impacto menor

Por su parte, NARA ha desarrollado una *Matriz de riesgos de preservación digital* para evaluar la conservación de formatos de archivo digitales, tanto actuales como futuros. Esta matriz se basa en 27 preguntas organizadas en ocho categorías clave: divulgación, adopción, transparencia, autodocumentación, dependencias externas, e impacto de patentes y mecanismos de protección. Al reconocer la interdependencia de estos factores, NARA promueve un análisis integral de riesgos para una valoración precisa de los formatos digitales. Una evaluación de riesgos eficaz es esencial para la preservación digital, ya que permite identificar con precisión el riesgo de pérdida a lo largo del tiempo y determinar medidas para mitigarlo. Su utilidad radica en proporcionar estimaciones de riesgo claras y

comprensibles, que sirven como base para la planificación estratégica y la toma de decisiones predictivas.

4.2. Factores de sostenibilidad de los formatos de archivo

La sostenibilidad de los formatos digitales es crucial para asegurar la accesibilidad y legibilidad de la información a largo plazo, evitando la pérdida de datos por obsolescencia tecnológica. La sostenibilidad digital abarca la gestión del ciclo de vida de los objetos digitales y sus desafíos técnicos (Bradley, 2007). La evaluación y selección de formatos debe ponderar cuidadosamente su sostenibilidad, junto con su calidad y funcionalidad, para garantizar la preservación efectiva (Arms y Fleischhauer, 2005).

Diversas entidades líderes en preservación digital han establecido criterios de sostenibilidad para formatos digitales (Library of Congress, British Library, The United Kingdom National Archives, Harvard University Library, National Archives and Records Administration). Estos criterios buscan asegurar la preservación del contenido informativo ante la evolución tecnológica y consideran el contexto institucional para la creación y mantenimiento de objetos digitales, garantizando así su viabilidad a largo plazo.

La sostenibilidad de los formatos digitales se evalúa mediante diversos criterios como: adopción, autodocumentación, complejidad, dependencias externas de software y hardware, divulgación, estabilidad, identificación, patentes, interoperabilidad, protección técnica, transparencia y viabilidad (Library of Congress, 2015; National Archives and Records Administration, 2024; Brown, 2008; Pennock *et al*., 2014; Australasian Digital Recordkeeping Initiative, 2020). Estos criterios son similares a las categorías de la *Matriz de riesgos de preservación digital* de NARA, ya que ambos enfoques vinculan los factores de riesgo con los indicadores de sostenibilidad, buscando garantizar la preservación a largo plazo.

La *adopción y uso* de un formato digital se refiere a su grado de utilización como formato maestro, de difusión y de intercambio. Un formato ampliamente aceptado por otras instituciones es más transparente, reduce el riesgo de obsolescencia y, si es actualizado activamente, garantiza mayor estabilidad y compatibilidad a largo plazo.

Los objetos digitales que se *autodocumentan*, es decir, que incluyen metadatos integrados, facilitan la preservación a largo plazo y son menos vulnerables que aquellos que almacenan los metadatos por separado. La inclusión de metadatos dentro del propio objeto digital asegura que la información contextual necesaria para su comprensión y gestión se mantenga junto con el archivo, reduciendo el riesgo de pérdida o desvinculación. Los formatos que integran metadatos técnicos y permiten incrustar metadatos descriptivos y administrativos facilitan la gestión, control de integridad, autenticidad y usabilidad de los objetos digitales, siendo cruciales para la preservación.

En entornos tecnológicos en constante evolución, la *dependencia de un formato respecto al hardware* representa un riesgo significativo para su preservación a largo plazo. Esta dependencia dificulta la adaptación a nuevas tecnologías y plataformas, y puede llevar a la obsolescencia del formato, impidiendo su migración y preservación. Las *dependencias externas de software* y la falta de interoperabilidad reducen significativamente la probabilidad de acceso a los archivos a largo plazo. Los formatos deben ser independientes de un software específico para evitar la obsolescencia. Además, la compatibilidad con múltiples sistemas operativos es crucial para facilitar la interacción y la transferencia de archivos entre diferentes entornos informáticos.

La *divulgación* de un formato digital se refiere a la disponibilidad pública de sus especificaciones técnicas. Un formato bien divulgado, con documentación accesible y herramientas de verificación de integridad, facilita la creación de software compatible y promueve la interoperabilidad entre

sistemas, siendo crucial para su sostenibilidad a largo plazo. Como se mencionó anteriormente, se recomienda el uso de formatos abiertos, cuyas especificaciones son públicas, debido a su mejor documentación y mayor respaldo en comparación con los formatos propietarios. Estos formatos facilitan la validación y divulgación de archivos, y según Brown (2008), reducen la dependencia de tecnologías específicas, favoreciendo la sostenibilidad a largo plazo.

Un factor crucial en la especificación de un formato es su *estabilidad*, fundamental para mantener la interoperabilidad. La antigüedad de un formato y la frecuencia de sus actualizaciones son buenos indicadores de su nivel de estabilidad, evitar cambios constantes o significativos asegura su confiabilidad.

La *identificación* y *validación* precisas de los formatos son esenciales para su uso continuo y preservación. Un diseño adecuado incorpora información de versión en la estructura interna del archivo, facilitando la identificación. Además, la disponibilidad de herramientas de validación asegura que los archivos cumplan con las especificaciones técnicas del formato, garantizando su integridad y usabilidad.

La *dependencia en relación a patentes* representa un riesgo significativo para la sostenibilidad de los formatos digitales. Los formatos patentados restringen el acceso y la gestión de archivos a aplicaciones y licencias específicas, lo que dificulta la preservación a largo plazo. Por el contrario, los formatos de código abierto, libres de patentes, facilitan la colaboración, reducen riesgos legales y promueven la creación de herramientas de acceso, siendo preferibles para la preservación.

La *interoperabilidad* de un formato, medida por la cantidad de programas que lo soportan, es crucial para su sostenibilidad. Los formatos compatibles con una amplia variedad de software facilitan el intercambio de

registros electrónicos y la migración de un entorno técnico a otro, asegurando la accesibilidad y preservación de los datos a largo plazo.

Para garantizar el acceso futuro, los repositorios deben permitir la replicación y migración de contenido digital, lo que requiere la ausencia de *mecanismos técnicos de protección* que restrinjan su uso y preservación. Aunque algunos formatos permiten mecanismos opcionales para proteger la propiedad intelectual o requerir contraseñas, el cifrado dificulta significativamente el procesamiento y acceso a los archivos.

La *transparencia*, entendida como la capacidad de analizar directamente la representación digital con herramientas básicas, incluyendo la legibilidad humana, favorece la preservación a largo plazo de los formatos digitales. La transparencia reduce las dependencias complejas y facilita la migración, como se evidencia en el formato TIFF sin comprimir, cuya simple codificación permite su recuperación incluso ante la pérdida de especificaciones.

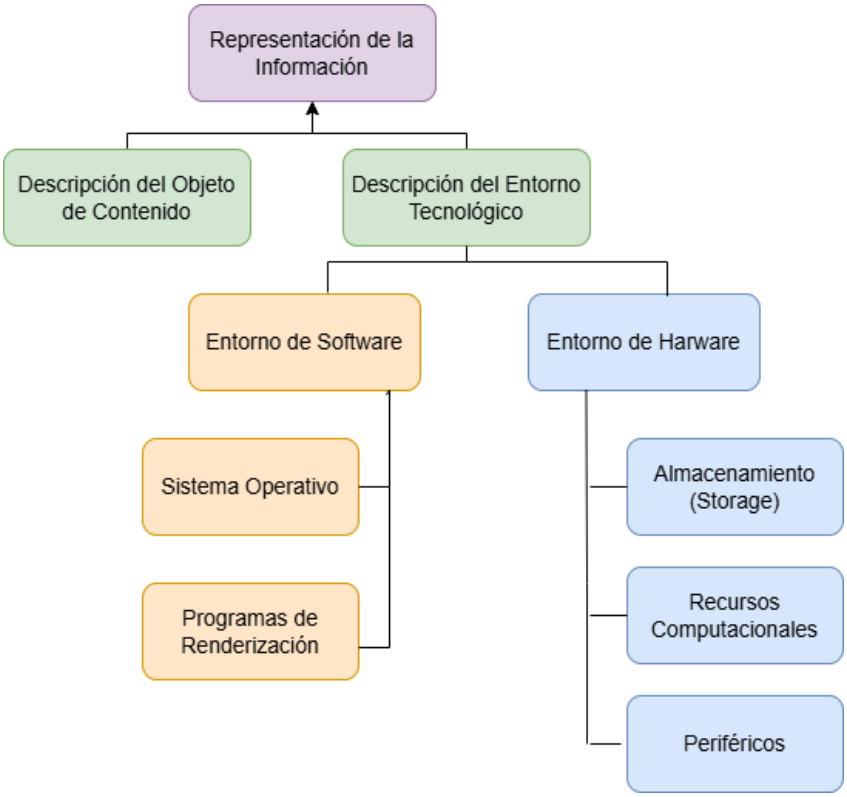
Los formatos digitales que incorporan mecanismos de *detección de errores* son preferibles para la preservación a largo plazo debido a su mayor robustez. Estos mecanismos permiten identificar daños durante la transmisión, facilitando la recuperación de la información y asegurando la integridad de los datos. La presencia de funciones de detección y, en algunos casos, corrección de errores, funciones verificables en la documentación técnica, permiten identificar y reparar daños en los archivos, asegurando la integridad y viabilidad de los datos a lo largo del tiempo.

Capítulo 5. Almacenamiento para la preservación a largo plazo

La tecnología que se utiliza para registrar, conservar, acceder y preservar información digital conforma el denominado “entorno tecnológico de almacenamiento” que, de acuerdo al grupo de trabajo en preservación digital de OCLC/RLG se divide en dos componentes: un entorno de software y un entorno de hardware.

El *entorno de software*, esencial para el acceso al contenido de objetos digitales archivados, se compone de programas de renderización y el sistema operativo. Los programas de renderización permiten visualizar o acceder al contenido, mientras que el sistema operativo proporciona la plataforma necesaria para su funcionamiento. Mientras que el *entorno de hardware* consta de objetos físicos (principalmente equipos informáticos) que son necesarios para el funcionamiento del entorno de software e incorpora tres aspectos: los recursos computacionales, el almacenamiento y los periféricos. Los recursos computacionales aluden a la capacidad lógica para procesar las secuencias de bits del objeto de datos de contenido y su entorno de software.

Figura 3. *Descripción del entorno tecnológico*



Nota. Adaptado y traducido de Online Computer Library Center/Research Libraries Group (2002, p. 25)

El almacenamiento para la preservación digital a largo plazo, enmarcado en el modelo OAIS, se divide en dos entidades funcionales: la ingesta y el almacenamiento de los objetos digitales de un RI. La *entidad funcional de ingreso o "ingesta"* proporciona los servicios para aceptar los SIP de los productores, lleva a cabo el control de calidad y crea los AIP. La *entidad funcional de almacenamiento* de archivo, por su parte, es responsable de recibir los AIP, agregarlos al almacenamiento permanente, actualizar los medios de almacenamiento, realizar verificaciones de errores, garantizar la recuperación ante desastres y entregar los AIP para pedidos de acceso o consulta (Reference Model for an Open Archival Information System, 2012).

Schaefer *et al*. (2021) destacan que, dentro del modelo OAIS, el almacenamiento de preservación digital se apoya en tres funciones clave: la planificación, la administración y la ingesta. La *planificación* de la preservación monitorea la tecnología para el almacenamiento, las normas relevantes y la migración de medios; la *administración* se ocupa de las políticas y normas relacionadas con la gestión del almacenamiento de preservación digital y audita los envíos, desde la recepción hasta el depósito en el almacenamiento; y la *ingesta* crea los paquetes de preservación y entrega los objetos digitales al almacenamiento, asegurando la integridad y accesibilidad a largo plazo de los mismos.

En cuanto a los medios físicos o soportes de almacenamiento de datos, la evolución ha sido constante en las últimas décadas, pero como indica Bhushan (2023), los discos duros se consolidan como la opción más común, tanto en modalidad en línea como externa. Si bien son susceptibles a daños físicos, su capacidad de almacenamiento y bajo costo los convierten en una alternativa viable para grandes volúmenes de datos, ofreciendo una favorable relación costo-beneficio. Por otro lado, las cintas magnéticas, a pesar de su antigüedad, conservan su relevancia para copias de seguridad de grandes volúmenes de datos almacenados fuera de línea, lo que las protege de ataques cibernéticos, y su bajo costo las hace ideales para el almacenamiento a largo plazo. De todas maneras, de acuerdo al principio deseable de múltiples soportes, el almacenamiento exclusivo en un solo soporte constituye un riesgo. En un proceso de preservación digital sistemático es sumamente necesaria la *redundancia* que asegura la durabilidad y recuperación de los datos. Para ello, se utilizan varias copias idénticas almacenadas en diferentes lugares geográficos y en distintos tipos de dispositivos o soportes de almacenamiento (infraestructura de preservación distribuida), minimizando así el riesgo de pérdida parcial o total de la información en caso de fallos del hardware, del software o de la red, alteraciones en la electricidad, fallas humanas, desastres naturales o ataques informáticos como virus y sabotajes intencionales.

En muchos casos se opta por un sistema de almacenamiento en la nube (“cloud computing” o “cloud storage”). El Instituto Nacional de Normas y Tecnología de los Estados Unidos (National Institute of Standards & Technology; NIST) lo define como un modelo que permite a los usuarios acceder a un conjunto compartido de recursos informáticos personalizables (como redes, servidores, aplicaciones y servicios) a través de Internet. Dicho modelo ofrece un acceso ubicuo y bajo demanda a los recursos digitales, permitiendo a los usuarios obtener y liberar almacenamiento con agilidad y con una mínima interacción con el proveedor (Mell y Grance, 2011). Sin embargo, se debe considerar que su implementación exige una gestión rigurosa en cuanto a la seguridad informática y al cumplimiento de requisitos legales. Se pueden mencionar algunos ejemplos de nubes comerciales como Dropbox, Microsoft OneDrive, Apple iCloud Drive, pCloud, Sync.com, IDrive, Box, Backblaze, NordLocker, Koofr, Google Cloud Storage y Amazon S3, entre otros.

Para el ámbito de las bibliotecas, archivos y editoriales, la Universidad de Stanford ha creado el reconocido Programa LOCKSS (Lots of Copies Keep Stuff Safe)[33] que proporciona tecnologías y servicios de almacenamiento a largo plazo de código abierto para una preservación digital segura, mitigando amenazas y favoreciendo la persistencia de los datos.

De acuerdo al modelo de implementación, el almacenamiento en la nube se puede clasificar en cuatro categorías (Mell y Grance, 2011). La nube privada, donde la infraestructura de la nube está destinada para el uso exclusivo de una sola organización. La nube comunitaria, donde el uso es para una comunidad específica de organizaciones que tienen inquietudes compartidas. La nube pública, destinada para el uso abierto del público en general, pudiendo ser propiedad de una organización empresarial, académica o gubernamental, o una combinación de ellas. Y la nube híbrida,

[33] Lots of Copies Keep Stuff Safe https://www.lockss.org/

cuya composición es el resultado de dos o más infraestructuras de nube (privadas, comunitarias o públicas).

En respuesta a la diversidad de enfoques en el almacenamiento para la preservación digital, la conferencia iPRES (2015) generó los *Criterios de Almacenamiento para la Preservación Digital*[34]. Estos criterios, accesibles en línea, proponen que un sistema de almacenamiento de preservación eficaz debe adherirse a prácticas interrelacionadas para asegurar la seguridad, integridad y accesibilidad de los datos a largo plazo. Se estructuran en nueve categorías clave: integridad del contenido, consideraciones de costo, seguridad de la información y del sistema, resiliencia de los medios de almacenamiento, flexibilidad, escalabilidad, rendimiento, soporte y transparencia del sistema.

La integridad del contenido exige mecanismos robustos para asegurar que los datos permanezcan inalterados y recuperables en su forma original a lo largo del tiempo, independientemente de los cambios tecnológicos o de infraestructura. Por otra parte, la sostenibilidad económica del sistema de almacenamiento demanda una planificación cuidadosa de los costos a largo plazo, abarcando almacenamiento, mantenimiento y operación. En lo relativo a la flexibilidad, la capacidad del sistema de almacenamiento debe adaptarse a las evoluciones en términos de capacidad, formatos y tecnologías, además de asegurar la interoperabilidad con otros sistemas y plataformas, facilitando el intercambio de datos y la migración tecnológica.

En cuanto a la seguridad se requiere la protección de los datos contra manipulaciones y eliminaciones no autorizadas, lograda mediante la redundancia de copias en ubicaciones y soportes diversos. Paralelamente, la seguridad se asegura con mecanismos de gestión de acceso y autenticación, previniendo accesos no autorizados y ataques maliciosos. Los medios de almacenamiento deben garantizar, además, una longevidad aceptable para la preservación segura de los datos a largo plazo. Se requiere una alta

[34] Digital Preservation Storage Criteria https://osf.io/ym6ua

durabilidad para asegurar el acceso confiable a los datos, junto a mecanismos efectivos para la detección y corrección de errores.

El sistema de almacenamiento debe garantizar también un rendimiento óptimo en la velocidad de acceso y procesamiento de datos, así como en la eficiencia de entrada/salida. Debe ser escalable, permitiendo ajustar la capacidad de almacenamiento a las necesidades organizacionales sin comprometer el rendimiento ni la disponibilidad de los datos. Y, finalmente, debe ofrecer servicios adicionales como migración de datos, recuperación y asistencia técnica, y garantizar la transparencia mediante auditorías, informes y notificaciones de errores, en este sentido, se recomienda el uso de formatos de archivo abiertos y documentados para facilitar la interoperabilidad y la preservación a largo plazo.

Por su parte, el consorcio NDSA, en sus *Niveles de preservación digital*, considera algunas prácticas estrechamente relacionadas con el almacenamiento de preservación a largo plazo de los contenidos de un RI. En primera instancia, sugiere mantener tres copias completas de los datos y documentar los medios de almacenamiento, indicando los recursos y las dependencias que estos requieren para funcionar. Por otro lado, recomienda verificar la integridad de la información al mover o copiar contenido, utilizando bloqueadores de escritura cuando se trabaja con medios originales y realizando un respaldo de la información, almacenando una copia en una ubicación separada del contenido. Y por último, aconseja registrar y documentar los agentes humanos y de software autorizados para leer, escribir, mover o eliminar contenido, para tal fin se deben implementar medidas de seguridad adecuadas, vinculadas al uso de usuarios y de contraseñas de autenticación.

En cuanto al segundo criterio del NDSA, relacionado con la integridad de los datos a largo plazo, se propone implementar mecanismos de verificación. Las sumas de verificación o *checksum* “son un cálculo cuyo valor sirve para verificar que todos los datos almacenados, transmitidos o

replicados se encuentren libres de error" (International Association of Sound and Audiovisual Archives, 2011) y permiten detectar cualquier alteración de las copias redundantes que se utilizan para reconstruir la información perdida. Idealmente, la verificación de la integridad de los objetos digitales almacenados debe realizarse tanto durante el proceso de ingesta como de forma periódica a través de un mecanismo automatizado.

Dos aspectos, no mencionados hasta aquí, pero cruciales para el éxito del almacenamiento de preservación digital, son: el establecimiento de convenciones de nomenclatura de archivos claras y documentadas, y la asignación de identificadores únicos y persistentes (PIDs) a cada documento almacenado en el RI. Este último procedimiento, que incluye el uso de DOI, Handle, URN, PURL, ARK, entre otros, representa una estrategia de preservación enfocada en garantizar la autenticidad e integridad de los objetos digitales.

Capítulo 6. Políticas de preservación digital

Una política es “un conjunto de principios que guían la toma de decisiones y las acciones con el fin de lograr los resultados deseados para una meta en particular” (InterPARES/ICA, 2012, p. 9). La planificación estratégica de una política, al ser un proceso dinámico y continuo, permite no solo proyectar un futuro deseado y definir los medios para alcanzarlo, sino también identificar problemas potenciales y desarrollar procedimientos para abordarlos de manera preventiva a través del análisis y la gestión de riesgos.

La preservación digital, tal como se ha establecido previamente en su conceptualización, se define como un proceso de gestión documental que requiere de una planificación estratégica sólida que garantice la vida útil de los activos digitales y el acceso a los documentos a lo largo de su ciclo de vida. El diseño de una política de preservación digital depende de su alcance y de los objetivos definidos, lo que permite justificar la importancia de la preservación y establecer la misión y visión que orientan la política.

Una política de preservación digital en el ámbito de un RI tendrá como objetivo principal garantizar la protección de la autenticidad, la integridad y la accesibilidad de los documentos digitales a largo plazo. Para alcanzar este objetivo se procederá al diseño de estrategias de preservación digital adecuadas (migraciones, copias de seguridad, etc.), a la evaluación y selección de formatos para acceso y conservación, a la utilización de modelos y estándares (OAIS, esquemas de metadatos) y a la implementación de auditorías que garanticen la fiabilidad del repositorio. Será fundamental que la planificación de estas políticas se implemente desde la fase inicial de concepción del repositorio, adoptando un enfoque programado, continuo y flexible que permita su adaptación a los avances tecnológicos, asegurando la sostenibilidad de los activos digitales presentes y futuros de la institución.

Dado que la formulación de políticas de preservación digital se enmarca en un contexto institucional y normativo, es imprescindible definir metas y objetivos alineados con la misión y los valores organizacionales, y que se cuente con el respaldo de las autoridades para su aprobación e implementación a largo plazo. Es fundamental que este apoyo se traduzca en acciones concretas, como la asignación de recursos técnicos y financieros destinados a la creación, mantenimiento y actualización de la infraestructura tecnológica, es decir, las necesidades de hardware, software y almacenamiento para el desarrollo de un repositorio digital confiable y eficiente. Asimismo, será importante la asignación de personal con experiencia en preservación digital, incluyendo bibliotecarios y especialistas en tecnología, considerando la capacitación continua de dicho personal, y la previsión para la contratación de expertos o proveedores externos, así como la cooperación con otras instituciones para proyectos específicos, la migración de formatos, la digitalización de impresos a formato digital, el almacenamiento remoto, etc.

A fin de lograr el respaldo de las autoridades, debe someterse a consideración un plan estratégico que ilustre cómo la política de preservación digital se alinea con la misión y los objetivos institucionales. Para ello, este plan deberá especificar los beneficios tangibles de la preservación digital, como el aumento de la accesibilidad y visibilidad de la producción digital institucional, la salvaguarda del patrimonio intelectual para futuras generaciones, el fomento de la investigación y la innovación, y la generación de oportunidades a través de la creación de nuevos productos y servicios digitales.

Adicionalmente, se deberá incluir un análisis exhaustivo de los riesgos inherentes a la ausencia de una política de preservación digital; este análisis deberá abordar la obsolescencia tecnológica, la pérdida de datos y la vulnerabilidad relacionada con la seguridad informática, cuantificando el

impacto potencial de estos riesgos en términos de costos y de su repercusión en el cumplimiento de los objetivos estratégicos institucionales.

La planificación de una política de preservación digital para un RI exige la colaboración de todas las partes interesadas, estableciendo con claridad sus relaciones y responsabilidades dentro de la estructura organizacional. Esto incluye al departamento de tecnologías de la información, la biblioteca responsable del repositorio digital, el departamento responsable de asignar recursos humanos y financieros, y el departamento legal encargado de los aspectos normativos vinculados con la circulación de la información y la aprobación de las políticas del repositorio, considerando la legislación vigente sobre propiedad intelectual y derechos de autor.

En síntesis, los principios rectores de una política de preservación digital, esenciales para la orientación de los agentes encargados de la conservación a largo plazo de los activos digitales, comprenden: la delimitación del alcance y los objetivos de la política, la identificación y evaluación de riesgos potenciales, la articulación de los beneficios y las oportunidades de la preservación, la determinación de los requerimientos legales, tecnológicos y económicos para su implementación, la estimación de costos y la planificación de la sostenibilidad financiera, la definición de roles y responsabilidades, y el diseño de estrategias de evaluación y revisión continua.

Metodología

En el presente estudio, se define como variable independiente el conjunto de políticas y prácticas de preservación digital implementadas por los repositorios institucionales de las unidades académicas de la UBA. Esta variable comprende las diversas estrategias, tecnologías y recursos que cada repositorio adopta con el objetivo de garantizar la preservación a largo plazo de sus contenidos digitales.

Por otro lado, la variable dependiente, el nivel de preservación digital alcanzado, se midió a través de un conjunto de indicadores que reflejan la capacidad de los repositorios para asegurar la perdurabilidad de sus activos digitales. Estos indicadores incluyen la monitorización de la integridad de los archivos, la garantía de accesibilidad a largo plazo, la implementación de estrategias de preservación digital (copias de seguridad, etc.), la adopción de directrices y esquemas de metadatos relevantes, la capacidad de migración de datos ante la obsolescencia tecnológica y el uso de formatos de archivo orientados a la preservación digital, entre otros criterios. En esencia, la investigación se centró en explorar la relación causal entre las políticas y prácticas de preservación digital implementadas (variable independiente) y el nivel de preservación digital resultante en los repositorios analizados (variable dependiente).

El diseño metodológico de esta tesina se caracteriza por ser no experimental y por emplear un enfoque mixto, combinando la recolección y el análisis de datos cuantitativos y cualitativos para examinar las prácticas y políticas de preservación digital en los RI de acceso abierto de las Facultades de la UBA.

Las unidades de análisis son las trece Facultades que componen la UBA: Facultad de Agronomía; Facultad de Arquitectura, Diseño y Urbanismo;

Facultad de Ciencias Económicas; Facultad de Ciencias Exactas y Naturales; Facultad de Ciencias Sociales; Facultad de Ciencias Veterinarias; Facultad de Derecho; Facultad de Farmacia y Bioquímica; Facultad de Filosofía y Letras; Facultad de Ingeniería; Facultad de Ciencias Médicas; Facultad de Odontología y Facultad de Psicología. También se incluye el RDI-UBA, gestionado por el SISBI, que recopila y difunde los contenidos digitales de todas las Facultades de manera integrada. En consecuencia, las fuentes primarias de datos para la presente investigación comprenden, los RI de las unidades académicas de la UBA, el RDI-UBA y los profesionales responsables de la gestión, el mantenimiento y la actualización de dichos repositorios.

La metodología adoptada en esta investigación comprendió una fase inicial de exploración exhaustiva de los sitios web de los RI de cada Facultad perteneciente a la Universidad de Buenos Aires y del RDI-UBA. El objetivo de esta exploración fue verificar la existencia y accesibilidad pública de políticas de preservación digital explícitas para sus respectivos RI. Los datos relevantes identificados durante esta etapa fueron registrados sistemáticamente en una hoja de cálculo de Google Sheets, la cual se estructuró para consignar la siguiente información para cada institución analizada: denominación de la Facultad; nombre del repositorio digital; dirección URL del repositorio; dirección de correo electrónico de contacto del repositorio; dirección URL donde se publican las políticas de preservación digital (en caso de su identificación); dirección URL del sitio web de la biblioteca central de cada Facultad; identificación y datos de contacto del responsable de la biblioteca central; e identificación y datos de contacto del responsable del RI (cuando esta información estaba disponible).

Adicionalmente, se diseñó una encuesta (ver **Anexo 2**) para obtener datos cuantitativos y, en menor medida, cualitativos, sobre las prácticas específicas de preservación digital implementadas en los repositorios de la UBA. Para la administración de la encuesta se empleó un formulario en línea

de Google Forms; este instrumento de recolección de datos consistió en veinticinco preguntas de formato cerrado, incluyendo opciones de respuesta de tipo dicotómico (sí/no) y de selección múltiple (multiple choice). La encuesta incluyó preguntas de formato mixto (cerradas y abiertas) diseñadas para permitir a los participantes profundizar o especificar sus respuestas en relación a algunos de los temas abordados. Asimismo, al finalizar el cuestionario, se ofreció un espacio para consignar cualquier observación adicional que no haya sido consultada.

Los datos recopilados a través de Google Sheets se utilizaron para relevar la existencia de políticas de acceso abierto y de preservación digital explícitas en los repositorios analizados. Paralelamente, se registró la información de contacto de los responsables de cada repositorio con el fin de facilitar el envío posterior de la encuesta. En la misma hoja de cálculo, se consignaron las fechas de envío y recepción de las respuestas, lo que permitió implementar un control eficiente del proceso de recolección de datos, de acuerdo con un cronograma predefinido.

A la par, se implementó una estrategia de seguimiento para incrementar la participación de las instituciones convocadas que consistió en el reenvío de la encuesta a aquellas instituciones que no habían respondido en la primera instancia, seguido de un contacto individualizado por correo electrónico y vía telefónica.

El análisis e interpretación de los datos recopilados se llevó a cabo mediante una hoja de cálculo Microsoft Excel, herramienta que permitió la organización, procesamiento y cálculo de las variables estudiadas. Para la presentación visual de los hallazgos se utilizó el editor de gráficos de Google Sheets, facilitando la creación de las figuras que ilustran los resultados de la investigación.

Resultados

La presente sección se centra en el análisis de los resultados obtenidos a partir de la encuesta diseñada para esta investigación, cuyo objetivo fue relevar las políticas de preservación digital de los RI de la UBA, incluyendo sus trece facultades y el SISBI. También se analizan los datos recopilados durante una primera fase exploratoria, durante la cual se revisaron los sitios web de cada RI con el propósito de identificar y confirmar la existencia de políticas de preservación digital publicadas y accesibles para los usuarios.

El instrumento de recolección de datos fue enviado a un total de catorce instituciones, lo que arrojó una tasa de respuesta del 64,2%, con nueve respuestas completas. Las instituciones que contestaron la encuesta y cuyos datos se analizan en esta sección son: la Facultad de Filosofía y Letras; la Facultad de Ciencias Exactas y Naturales; la Facultad de Ciencias Económicas; la Facultad de Farmacia y Bioquímica; la Facultad de Agronomía; la Facultad de Medicina; la Facultad de Odontología; la Facultad de Psicología y el SISBI.

Es importante destacar que, de las cinco facultades que no participaron en la encuesta, tres carecen de repositorio propio, aunque incorporan sus contenidos en el RDI-UBA (Facultad de Arquitectura, Diseño y Urbanismo; Facultad de Derecho y Facultad de Ciencias Veterinarias) y dos poseen un repositorio autónomo (Facultad de Ciencias Sociales y Facultad de Ingeniería).

Los datos recolectados a través de este instrumento nos permiten analizar el estado actual de la preservación digital en los repositorios institucionales de la universidad, que constituye el objeto de estudio de esta tesina. A continuación, se presenta un análisis de las respuestas, desglosando los hallazgos en áreas clave como las políticas de acceso, la

infraestructura tecnológica, los formatos de archivo, los metadatos y algunas estrategias de preservación digital. Este análisis comparativo busca identificar las áreas de mejora y señalar desafíos comunes en la gestión del patrimonio digital de la UBA.

Políticas y procedimientos de preservación digital

El análisis sobre la existencia e implementación de políticas de preservación digital en los RI de la UBA, objeto de estudio de esta investigación, reveló que cuatro de las nueve instituciones que respondieron la encuesta (44,44%) confirmaron poseerlas o, al menos, contar con procedimientos definidos para este propósito.

Con base en los datos de la encuesta, se identificaron varias modalidades para la formalización y accesibilidad de las políticas de preservación digital. Una institución indicó que estas políticas están detalladas en un documento de carácter interno, tres señalaron que se encuentran formalizadas en una resolución institucional, mientras que dos repositorios optaron por la categoría "Otro", sugiriendo que emplean métodos diferentes para documentar sus políticas de preservación digital o no las tienen. Por otra parte, un análisis exploratorio de la web de los repositorios de las instituciones consultadas en la encuesta identificó que tres de ellas tienen su política de preservación disponible de forma pública (estas son la Facultad de Agronomía[35], la Facultad de Ciencias Exactas y Naturales[36] y el RDI-UBA[37]).

A pesar de la ausencia de políticas de preservación digital formalizadas en algunas instituciones, existe una intención de abordar esta

[35] http://ri.agro.uba.ar/files/download/biblioteca/arcd-2021-106.pdf (pág. 4)
[36] https://bibliotecadigital.exactas.uba.ar/download/documento/Politica-Acceso-Abierto-resolucionCD_2399_23.pdf (pág. 17)
[37] https://repositoriouba.sisbi.uba.ar/gsdl/web/reglamento.pdf (pág. 3)

carencia. Específicamente, las instituciones que actualmente no disponen de estas políticas expresaron en la encuesta su compromiso de redactarlas, publicarlas e implementarlas en el futuro.

La **Tabla 6** compara las políticas de acceso abierto publicadas en los sitios web de los repositorios de las catorce Facultades de la UBA (incluyendo el RDI-UBA) con la existencia o no de políticas de preservación digital, según la información recopilada tanto mediante la encuesta como de fuentes públicas de Internet. Refiere cuáles son las Facultades que no poseen un repositorio autónomo y qué tipo de acceso tienen sus políticas de preservación, cuando estas existen.

Tabla 6. *Políticas de Acceso Abierto y Políticas de Preservación de los RI de la UBA*

Institución	***Política de acceso abierto*** *(pública web RI)*	***Política de preservación digital***
Facultad de Agronomía	Normativa propia	Si *
Facultad de Arquitectura, Diseño y Urbanismo †	Normativa del RDI	Si **
Facultad de Ciencias Económicas	Normativa propia	No
Facultad de Ciencias Exactas y Naturales	Normativa propia	Si *
Facultad de Ciencias Sociales	No posee	s/inf.

Facultad de Ciencias Veterinarias †	Normativa del RDI	Si **
Facultad de Derecho †	Normativa del RDI	Si **
Facultad de Farmacia y Bioquímica †	Normativa del RDI	Si **
Facultad de Filosofía y Letras	Normativa propia	No
Facultad de Ingeniería	No posee	s/inf.
Facultad de Medicina	No posee	No
Facultad de Odontología	Normativa propia	No
Facultad de Psicología †	Normativa del RDI	s/inf.
RDI-UBA	Normativa propia	Si *

Nota. † (no posee repositorio autónomo); * (de acceso público); ** (de acceso público en el RDI-UBA); s/inf. (sin información)

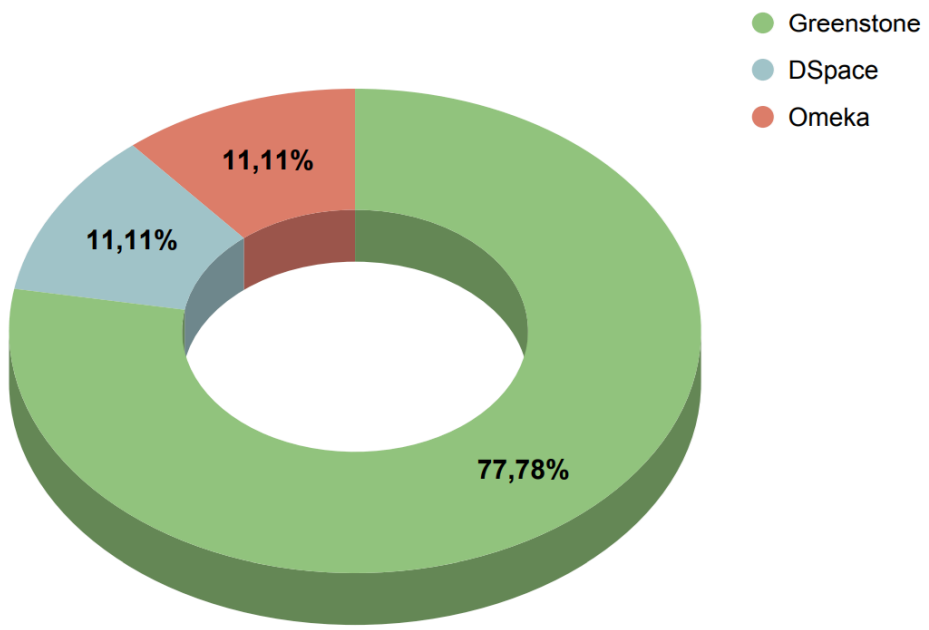
Información general de los repositorios

Los hallazgos de la encuesta evidencian un marcado predominio de Greenstone como software para la gestión de repositorios, concentrando el 55,56% de las elecciones (5 de las 9 respuestas). En contraste, DSpace y Omeka fueron seleccionados solo una vez cada uno, constituyendo cada una

un 11,11% del total. Cabe destacar que las opciones EPrints y "Desarrollo propio" no fueron seleccionadas por ninguno de los encuestados **(Gráfico 1)**.

Es importante señalar que dos de las respuestas (22,22%) clasificadas inicialmente en la categoría “Otro” merecen una consideración especial. Estas fueron proporcionadas por instituciones que, por carecer de infraestructura propia, gestionan sus contenidos a través del RDI-UBA, implementado con Greenstone. Por lo tanto, la clasificación inicial de estas respuestas en “Otro” distorsiona la verdadera prevalencia de Greenstone. Una vez reasignadas, la predominancia de Greenstone ascendería a 7 de las 9 respuestas obtenidas, representando un 77,78% del total.

Gráfico 1. *Software de gestión de los repositorios*

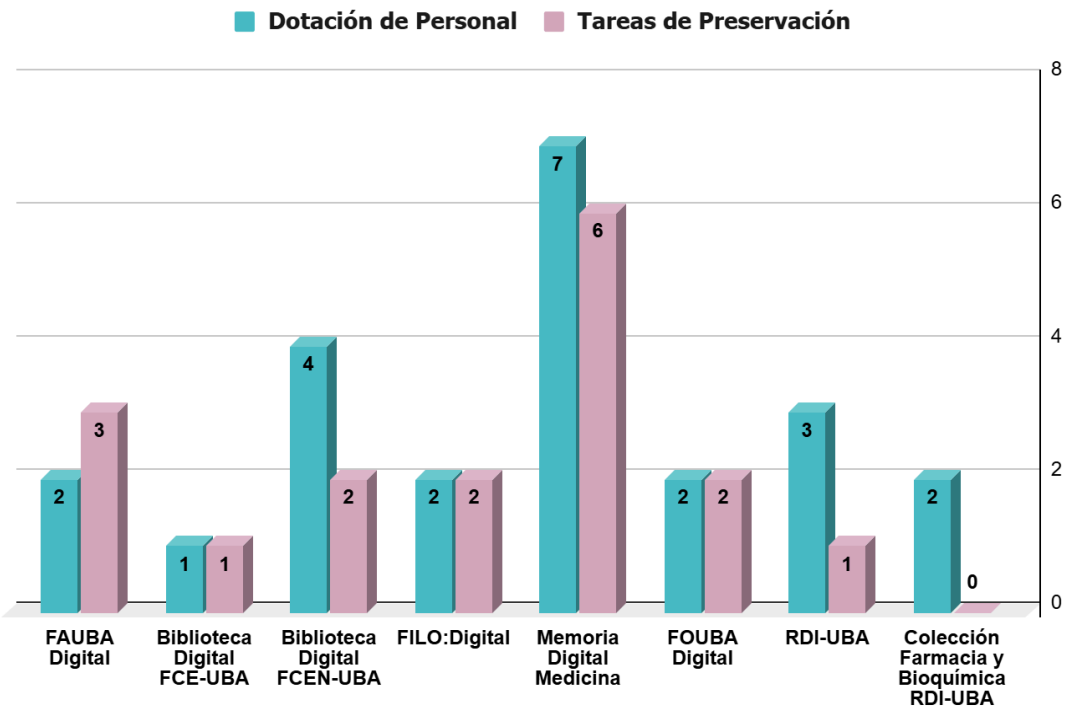


En cuanto a la dotación de personal de los repositorios, de las diez instituciones que respondieron la encuesta, ocho brindaron información sobre el personal asignado a la gestión y mantenimiento de los mismos. De ese conjunto, se observa que todas las instituciones que disponen de un

repositorio autónomo, incluyendo el RDI-UBA, cuentan con al menos una persona designada para su supervisión. Esta función se define generalmente con alguno de los siguientes roles: administrador, responsable, desarrollador, gestor o coordinador.

La dotación de personal de los repositorios, independientemente de si su asignación sea a tiempo parcial o con dedicación exclusiva, se sitúa entre 1 y 7 profesionales. En el **Gráfico 2** se observa la relación entre la dotación de personal de los repositorios y, de esa dotación, cuántas personas se dedican a tareas de preservación digital (en promedio, la mitad o la totalidad del personal). Los valores numéricos situados sobre cada barra indican la cantidad exacta de personal designado para tareas generales del repositorio y para funciones específicas de preservación digital. Considerando el primer conjunto de barras (correspondiente al repositorio FAUBA Digital): el valor "2" (barra azul) indica que dos miembros del personal de la Biblioteca están asignados a las tareas generales del repositorio, mientras que el valor "3" (barra morada) muestra que a esas dos personas se suma un informático adicional, totalizando tres personas que se ocupan de las tareas de preservación digital.

Gráfico 2. *Distribución del personal asignado a tareas de preservación digital*



Se destaca la Facultad de Medicina que se posiciona como la que cuenta con la mayor asignación de recursos humanos para su repositorio, liderando también en el número de individuos específicamente enfocados en las actividades de preservación digital.

Al consultar si los repositorios cuentan con un presupuesto asignado exclusivamente para tareas de preservación digital, todas las respuestas (100%) fueron negativas, indicando que ninguno de los repositorios consultados dispone de un presupuesto específico para este fin.

Almacenamiento y seguridad de la información

Se exploró la existencia de procedimientos formales para la realización de copias de seguridad como estrategia clave para la preservación digital de los objetos digitales de los repositorios. Los

resultados de la encuesta revelaron que ocho de las diez instituciones que respondieron la encuesta cuentan con un procedimiento documentado para este propósito. Las dos instituciones restantes, que no operan repositorios de forma autónoma, no disponen de un procedimiento propio, ya que confían en el servicio de almacenamiento y las copias de seguridad gestionadas por el RDI-UBA, donde depositan sus activos digitales.

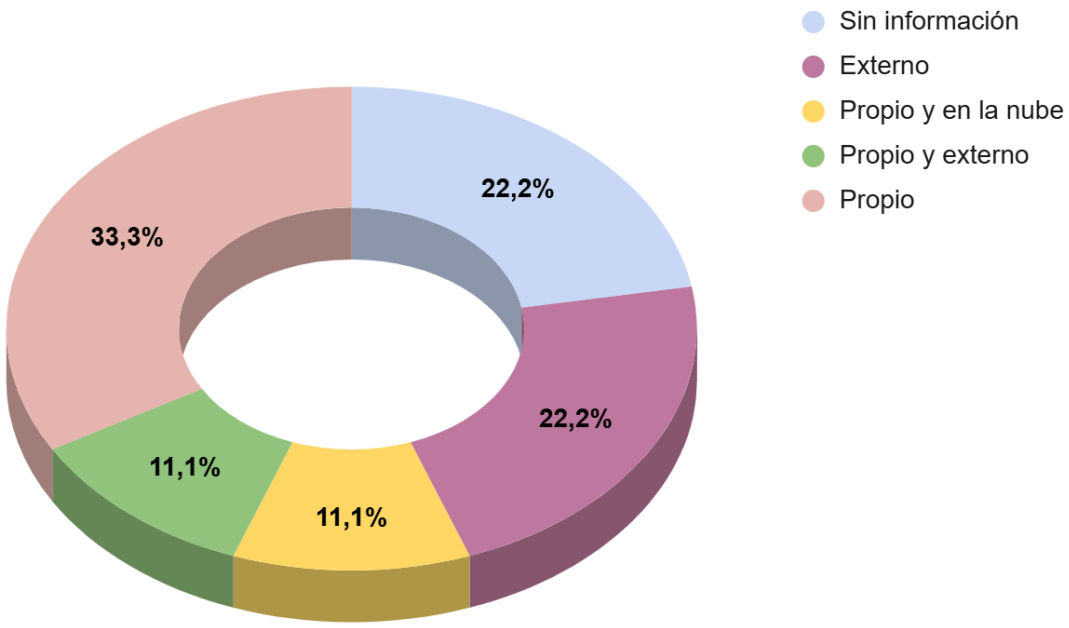
En cuanto a la evaluación de la seguridad de la información, se examinó la implementación de la redundancia de copias, una estrategia fundamental para la preservación digital a largo plazo. Los resultados indican que 5 de los 7 repositorios autónomos de la UBA, incluido el RDI-UBA, cuentan con al menos dos copias de seguridad de sus objetos digitales. Este dato es significativo, ya que demuestra una comprensión generalizada de la importancia de la redundancia para mitigar el riesgo de pérdida de datos.

Los modelos de almacenamiento de los objetos digitales varían, reflejando las capacidades y estrategias particulares de cada institución **(Gráfico 3)**. Se identificó el almacenamiento propio (33,3%), donde la institución gestiona y mantiene sus propias copias; también se encontró el almacenamiento exclusivamente externo (22,2%), que corresponde a las dos instituciones sin repositorio propio y que delegan la responsabilidad en un servicio externo, en este caso el RDI-UBA; y, por último, algunas instituciones utilizan un almacenamiento mixto, combinando el propio con el externo (11,1%) o en la nube (11,1%), una estrategia que garantiza la seguridad al diversificar los riesgos.

Resulta relevante señalar que hubo dos instituciones con repositorio autónomo que no proporcionaron información sobre este aspecto. No obstante, este análisis revela que, si bien la mayoría de los repositorios autónomos aplican la redundancia, el almacenamiento de los objetos digitales en los repositorios evidencia una variedad de modelos que se ajustan a los recursos y estrategias específicas de cada institución, desde la

autogestión hasta la delegación completa en un servicio centralizado como el RDI-UBA.

Gráfico 3. *Tipos de almacenamiento de los objetos digitales*



Se constató que únicamente tres instituciones (33,33%) disponen de un plan de contingencia formal para afrontar fallas en el almacenamiento. Esta situación revela que la mayoría de los repositorios aún no cuentan con protocolos estandarizados para la recuperación de datos en situaciones de emergencia, lo que constituye una vulnerabilidad potencial y sugiere una brecha significativa en sus políticas de preservación digital.

Al explorar las áreas percibidas como críticas para mejorar la gestión y seguridad del almacenamiento de datos, las respuestas de las instituciones revelaron una clara conciencia sobre la necesidad de fortalecer sus estrategias de preservación digital. Los aspectos mayormente señalados fueron:

- La mayoría de las instituciones identificó la necesidad de elaborar planes de contingencia robustos para responder eficazmente a fallas o desastres.

- Se enfatizó la importancia de documentar formalmente los procedimientos y asignar responsabilidades claras para cada etapa de la gestión de los datos.

- Las instituciones sugirieron la necesidad de establecer almacenamientos alternativos y copias espejadas para garantizar la redundancia de los datos.

- Se mencionó la importancia de utilizar una diversidad de formatos de archivo para asegurar la interoperabilidad a largo plazo, así como la necesidad de implementar metadatos de preservación digital para un mejor control y seguimiento de los objetos digitales.

De las nueve instituciones que participaron en la encuesta, el 55,6% (5 de ellas) confirmaron realizar verificaciones periódicas de la integridad de sus objetos digitales. Es relevante destacar que dos de las instituciones que contestaron la encuesta no administran un repositorio propio, lo que explica su ausencia en esta práctica. **(Tabla 7)**

La periodicidad de las verificaciones de integridad varía entre las instituciones, algunas lo hacen de manera irregular, anual o cuando se detectan fallas. La Facultad de Medicina, que actualmente lleva a cabo un control de periodicidad anual, manifestó su intención de implementar controles semestrales.

En contraste, la Facultad de Filosofía y Letras, que implementa su RI con DSpace, realiza la comprobación de integridad de forma permanente, gracias a la función incorporada del *software* que detecta y previene alteraciones o corrupción en los archivos almacenados (*checksums*).

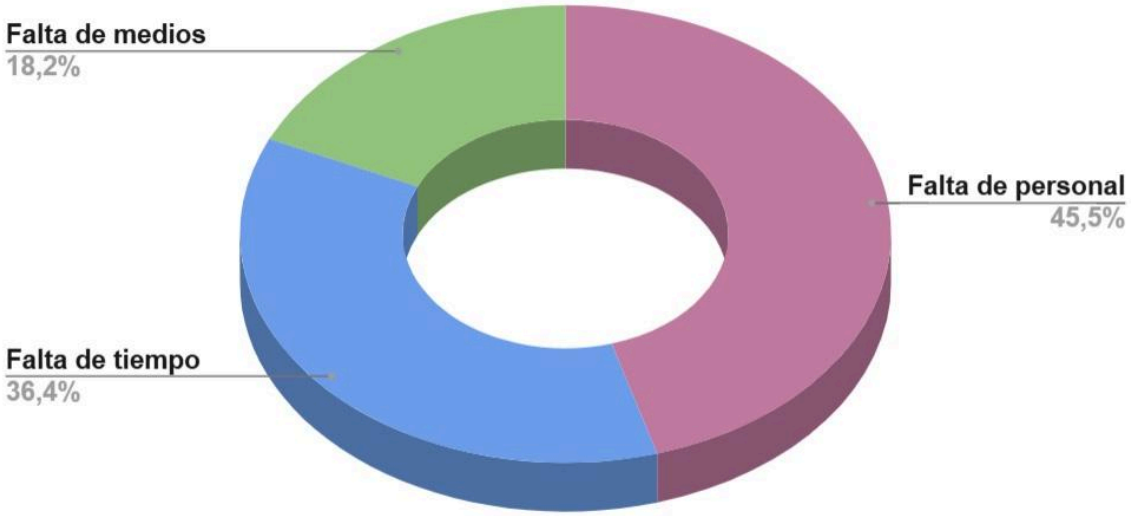
Tabla 7. *Verificación de integridad de los objetos digitales*

Institución	***Repositorio***	***¿Se verifica la integridad?***	***¿Con qué periodicidad?***
Facultad de Agronomía	FAUBA Digital	Si	Irregular
Facultad de Cs. Económicas	Biblioteca Digital FCE-UBA	No	---------
Facultad de Cs. Exactas y Naturales	Biblioteca Digital FCEN-UBA	Si	Anual
Facultad de Filosofía y Letras	FILO:Digital	Si	Permanente *
Facultad de Medicina	Memoria Digital Medicina	Si	Anual
Facultad de Odontología	FOUBA Digital	No	---------
SISBI	RDI-UBA	Si	Irregular
Facultad de Farmacia y Bioquímica †	Colección Farmacia y Bioquímica RDI-UBA	No	---------
Facultad de Psicología †	Colección Psicología RDI-UBA	No	---------

Nota. † (no posee repositorio autónomo); * (implementa el RI con DSpace)

Aunque la mitad o la totalidad del personal asignado para trabajar en los repositorios está dedicada a tareas de preservación digital (como se muestra en el **Gráfico 2**), la principal barrera citada por los entrevistados para no llevar a cabo controles de integridad es la escasez de personal (45,5%). Otros motivos importantes citados incluyen la falta de tiempo (36,4%) y la ausencia de medios técnicos adecuados para implementarla (18,2%). **(Gráfico 4)**

Gráfico 4. *Obstáculos para la verificación de integridad de los objetos digitales*



Al analizar la práctica de verificación de virus al momento de la ingesta de los objetos digitales a los repositorios se revela una disparidad. Cuatro de ellas (44,4%) confirmaron realizar este control de manera sistemática. Por otro lado, tres instituciones (33,3%) no aplican ningún tipo de verificación, mientras que las dos restantes (22,2%) lo hacen de forma ocasional.

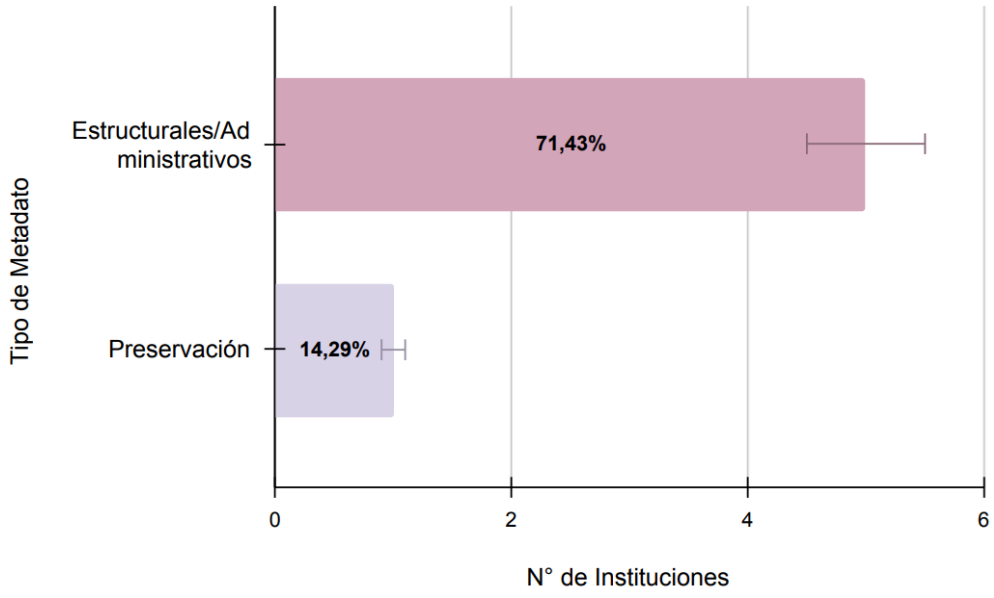
Respecto a la definición de tipos de usuarios que requieren identificación (*login*) para acceder al repositorio, la mayoría de las instituciones que respondieron la encuesta, seis de nueve (66,7%), indicaron

que no han establecido esta estrategia de seguridad de los datos. Por el contrario, tres de las nueve (33,3%) afirman haber definido y configurado usuarios específicos para el acceso. La falta de autenticación de usuarios para acceder a los repositorios tiene implicaciones negativas para la seguridad y el análisis del uso de los recursos digitales.

Metadatos, formatos e identificadores digitales

La mayoría de las instituciones encuestadas, cinco de las siete que respondieron este punto (71,43%), confirmó realizar el registro de metadatos estructurales y administrativos (formatos, relación entre objetos digitales, control de acceso, licencias, etc.). Por el contrario, dos instituciones (28,6%) indicaron que no llevan a cabo esta práctica. A diferencia de lo anterior, el registro de metadatos de preservación presenta un panorama menos favorable, de las siete instituciones, solo una (14,29%) respondió afirmativamente a esta pregunta. **(Gráfico 5)**

Gráfico 5. *Registro de metadatos estructurales-administrativos versus preservación*



Respecto a la práctica de incrustar metadatos descriptivos en los objetos digitales antes de su almacenamiento o ingesta, los resultados indican una baja implementación. Específicamente, solo una institución (11,11%) de las consultadas afirmó llevar a cabo esta práctica, utilizando para ello el software AutoMetadata.

En cuanto al uso de identificadores únicos y persistentes, que garantiza el acceso permanente a los objetos digitales y facilita la verificación de procedencia y autenticidad de los mismos, solo una institución respondió aplicar el sistema HANDLE (11,11%). Las restantes no asignan ningún identificador propio a los objetos, pero utilizan el DOI incorporado en los documentos gestionados en el repositorio.

En cuanto a la evaluación, identificación y selección de formatos de para los distintos tipos de objetos digitales incorporados, los resultados de la encuesta muestran que, de las nueve instituciones que respondieron la encuesta, solo cuatro (44,44%) indicaron contar con criterios definidos para esta práctica de preservación a largo plazo: Facultad de Agronomía; Facultad de Ciencias Económicas; Facultad de Ciencias Exactas y Naturales; y Facultad de Filosofía y Letras.

Finalmente, dos instituciones (22,22%) informaron llevar a cabo un monitoreo periódico orientado a detectar posibles problemas de obsolescencia en los formatos de los objetos digitales almacenados en sus repositorios, lo que garantiza la sostenibilidad y accesibilidad a largo plazo de los mismos: Facultad de Ciencias Exactas y Naturales, y Facultad de Filosofía y Letras.

Conclusiones

El análisis de los resultados obtenidos en la presente investigación permite trazar un panorama general del estado actual de las prácticas de preservación digital en los repositorios institucionales de acceso abierto de la UBA. A partir de los datos relevados mediante la encuesta y del examen exploratorio de los sitios web de los repositorios, se evidencia que la universidad ha alcanzado un grado importante de desarrollo en materia de acceso abierto y visibilidad de su producción científica, pero aún enfrenta desafíos significativos en la consolidación de políticas y estrategias de preservación digital a largo plazo.

En primer lugar, se observa una heterogeneidad considerable en cuanto a la existencia y la formalización de políticas de preservación digital. Solo el 44,4% de las instituciones encuestadas declaró poseer políticas o procedimientos afines, mientras que la mayoría reconoció la urgencia de su futura elaboración e implementación. Este hallazgo subraya una brecha institucional entre el avance en la consolidación del acceso abierto y la efectiva preservación sostenible de los activos digitales.

La falta de documentación sobre procesos clave como la migración de formatos, la procedencia, la autenticidad y el entorno técnico revela una debilidad significativa en las prácticas de preservación a largo plazo. Sin embargo, cabe destacar que la ausencia de normativas formales no refleja necesariamente desinterés, sino que es consecuencia de la limitación de recursos técnicos, profesionales y normativos necesarios para su desarrollo.

Es importante detenerse en el caso de las instituciones que no tienen una política propia y que optan por la delegación de esta responsabilidad al RDI-UBA. Este modelo de trabajo subraya una realidad en muchas instituciones con recursos limitados, al depositar sus activos digitales en un repositorio centralizado, estas instituciones aprovechan la infraestructura y

las políticas de seguridad de una entidad mayor. Por lo tanto, aunque no tengan un protocolo interno, se puede concluir que están aplicando una estrategia de preservación indirecta.

En relación a la infraestructura tecnológica, se debe destacar que Greenstone, si bien ofrece como software la interoperabilidad con otros sistemas mediante el protocolo OAI-PMH, fue concebido inicialmente para la creación de colecciones de bibliotecas digitales, no específicamente para repositorios digitales institucionales. Debido a este enfoque de origen, carece de funcionalidades específicas orientadas a la preservación digital. En contraste, DSpace presenta una ventaja significativa en este aspecto, ya que permite la gestión de la integridad de archivos y *checksums*, funcionalidades cruciales para asegurar la autenticidad y la inalterabilidad de los recursos digitales a lo largo del tiempo.

Otro hallazgo relevante es la ausencia generalizada de presupuestos específicos destinados a tareas de preservación digital. Ninguna de las instituciones encuestadas dispone de financiamiento exclusivo para este fin, lo cual condiciona la capacidad para mantener la infraestructura, realizar auditorías de integridad o adoptar estrategias de preservación. Esta carencia económica se combina con limitaciones de personal, ya que las tareas de preservación recaen en equipos reducidos, muchas veces con responsabilidades múltiples.

Respecto al almacenamiento y la seguridad de la información, se observa que la mayoría de los repositorios ha implementado prácticas básicas de copia de seguridad (backup). Si bien esto indica una alineación con buenas prácticas en la protección de los datos, la carencia de planes de contingencia formales y de procedimientos documentados de recuperación ante desastres revelan una vulnerabilidad ante la necesidad de gestionar eventuales pérdidas de información o fallas de la infraestructura tecnológica.

La asignación de identificadores únicos y persistentes, fundamental para la accesibilidad, la procedencia y la autenticidad de los objetos digitales, evidenció una marcada dependencia de identificadores externos (como el DOI preexistente). En consecuencia, se considera prioritario avanzar en el desarrollo y fortalecimiento de estrategias institucionales orientadas a garantizar la identificación persistente y el control de los activos digitales de la universidad.

Por otra parte, en lo que respecta a la evaluación, identificación y selección de formatos de preservación para los distintos tipos de objetos digitales, los resultados de la encuesta indican que solo una minoría de las instituciones consultadas cuenta con criterios definidos para esta práctica. Asimismo, la situación es más crítica en cuanto al monitoreo periódico de la obsolescencia de formatos, el cual es realizado por un número aún más reducido de instituciones. Esto indica que la sostenibilidad y la accesibilidad a largo plazo de los activos digitales de la universidad es aún rudimentaria.

En cuanto a los metadatos, si bien las instituciones analizadas gestionan los metadatos estructurales y administrativos de forma generalizada, el registro de metadatos de preservación es una práctica casi nula. Este hallazgo destaca la necesidad urgente de fortalecer políticas tendientes al registro de metadatos de preservación que garanticen la longevidad y el acceso a los objetos digitales. La incorporación sistemática de metadatos de preservación, conforme a estándares como PREMIS o METS, constituye una deuda pendiente y una prioridad estratégica para los repositorios de la UBA.

La falta de verificación de integridad y de controles de virus en algunos repositorios también revela un área de riesgo. Aun cuando varias instituciones expresaron su intención de implementar verificaciones periódicas, las limitaciones de tiempo, personal y recursos tecnológicos son obstáculos recurrentes. Esto refuerza la necesidad de consolidar

mecanismos automatizados que reduzcan la dependencia del trabajo manual y que incrementen la fiabilidad de los sistemas.

Los resultados permiten afirmar que los repositorios institucionales de la UBA se encuentran en una etapa de madurez intermedia. Han logrado consolidar políticas de acceso abierto y afianzar la visibilidad de la producción científico-académica, pero requieren profundizar la planificación y gestión de la preservación digital. La heterogeneidad en los niveles de desarrollo, la falta de políticas formales, la escasez de recursos específicos y la limitada adopción de estándares internacionales son factores que obstaculizan la consolidación de una estrategia institucional sostenible.

En síntesis, además de subrayar la urgencia de desarrollar políticas de preservación digital, también se revela una serie de desafíos interconectados relacionados con la falta de recursos, personal, conocimiento y colaboración. Abordar estos problemas de manera integral será fundamental para fortalecer la capacidad de los repositorios para preservar su patrimonio digital y garantizar su acceso a largo plazo. La implementación de políticas sólidas debe ir de la mano con la asignación de recursos adecuados, la capacitación del personal y la exploración de modelos de colaboración más efectivos.

Se destaca el potencial articulador del Sistema de Bibliotecas y de Información (SISBI) como agente central para coordinar, normalizar y fortalecer las prácticas de preservación digital en toda la universidad. La formulación de una política universitaria de preservación digital, común a todas las facultades, acompañada por un plan de capacitación técnica y la asignación de recursos sostenibles, resultan pasos imprescindibles para asegurar la autenticidad, permanencia y accesibilidad del patrimonio digital de la UBA a largo plazo.

Referencias bibliográficas

Abrams, S. L. (2005). Establishing a Global Digital Format Registry. *Library Trends*, *54*(1), 125-143. https://core.ac.uk/download/pdf/4812597.pdf

Allendez Sullivan, P. M. (2004). El impacto de las nuevas tecnologías en la competencia laboral del bibliotecario del siglo XXI. *Biblos. Revista de Bibliotecología y Ciencias de la Información*, *5*(17), 25-35. https://www.redalyc.org/pdf/161/16101701.pdf

Alonso Arévalo, J. A.; Subirats, I. y Martínez Conde, M. L. (2008). *Informe APEI sobre acceso abierto*. Asociación Profesional de Especialistas en Información. https://www.apei.es/wp-content/uploads/2013/11/InformeAPEI-Accesoabierto.pdf

Argentina. Ministerio de Ciencia, Tecnología e Innovación Productiva (2011). *Resolución 469/2011*. https://biblioteca.mincyt.gob.ar/storage/resoluciones/Resoluci%C3%B3n-469-11-17-05-2011-crea-SNRD.pdf

Argentina. Ministerio de Ciencia, Tecnología e Innovación Productiva (2016). *Resolución 753-E/2016*. https://www.boletinoficial.gob.ar/detalleAviso/primera/154125/20161116

Argentina. Sistema Nacional de Ciencia, Tecnología e Innovación (2013). *Ley 26.899. Repositorios digitales institucionales de acceso abierto*. https://www.argentina.gob.ar/normativa/nacional/ley-26899-223459/texto

Arms, C. y Fleischhauer, C. (2005). Digital formats: factors for sustainability, functionality, and quality. *Archiving Conference*, *2*(art00047), 222-227. https://doi.org.10.2352/issn.2168-3204.2005.2.1.art00047

Australasian Digital Recordkeeping Initiative (2020). *Sustainable digital file formats for creating and using records*. https://www.caara.org.au/wp-content/uploads/2023/07/Sustainable-Digital-File-Formats-for-Creating-and-Using-Records-V1.0-April-2020.pdf

Banzato, Guillermo (2012). *El movimiento de acceso abierto al conocimiento científico en la Argentina. Políticas y prácticas en torno a la investigación, las revistas académicas y los repositorios* (Proyecto de investigación PI+D H642). Universidad Nacional de La Plata, Facultad de Humanidades y Ciencias de la Educación, Instituto de Investigaciones en Humanidades y Ciencias Sociales. https://www.memoria.fahce.unlp.edu.ar/proyectos/py.622/py.622.pdf

Barrueco Cruz, J. M. y García Testal, C. (20-22 de mayo de 2009). *Repositorios institucionales universitarios: evolución y perspectivas* [Presentación]. XI Jornadas Españolas de Documentación, FESABID 2009 "Interinformación", Zaragoza, España. https://www.fesabid.org/zaragoza2009/actas-fesabid-2009/99-107.pdf

Barrueco Cruz, J. M.; Rico-Castro, P.; Bonora Eve, L. V.; Azorín, C.; Bernal, I.; Gómez Castaño, J.; Guzmán Pérez, C.; Losada Yáñez, M.; Fabra Marín del Campo, R.; Martínez Galindo, F. J.; Martínez Pousa, C.; Morillo Moreno, J. C. y Prats Prat, J. (2021). *Guía para la evaluación de repositorios institucionales de investigación*. Fundación Española para la Ciencia y la Tecnología; Red de Bibliotecas Universitarias. https://riunet.upv.es/handle/10251/166115

Barve, S. (21-23 de febrero de 2007). *File formats in digital preservation* [Presentación]. International Conference on Semantic Web and Digital Libraries, Bangalore, India. https://core.ac.uk/download/pdf/333967179.pdf

Baucom, E. A. (2019) Brief history of digital preservation. En: *Digital preservation in libraries : preparing for a sustainable future*. American Library Association Editions. (An ALCTS Monograph). https://doi.org/10.1080/24750158.2020.1757577

Beagrie, N.; Charlesworth, A. y Miller. P. (2015). The National Archives guidance on cloud storage and digital preservation. 2nd ed. https://cdn.nationalarchives.gov.uk/documents/CloudStorage-Guidance_March-2015.pdf

Bhushan, B. (2023). Current status and outlook of magnetic data storage devices. *Microsystem Technologies*, *29*(11), 1529-1546. https://doi.org/10.1007/s00542-023-05549-z

Bia Platas, M. y Sánchez Quero, M. (18-19 de noviembre de 2002). *Desarrollo de una política de preservación digital: tecnología, planificación y perseverancia* [Presentación]. III Jornadas de Bibliotecas Digitales. Madrid, España. https://www.cervantesvirtual.com/nd/ark:/59851/bmc5b028

Bodero Poveda, E. M.; De Giusti, M. y Morales Alarcón, C. (2021). La preservación digital a largo plazo y las bases de la planificación estratégica. *3C TIC. Cuadernos de Desarrollo Aplicados a las TIC*, *10*(3), 17-39. https://doi.org/10.17993/3ctic.2021.103.17-39

Bodero Poveda, E. M.; De Giusti, M. y Morales Alarcón, C. (2022). Preservación digital a largo plazo: estándares, auditoría, madurez y planificación estratégica. *Revista Interamericana de Bibliotecología*, *45*(2), e344178. https://doi.org/10.17533/udea.rib.v45n2e344178

Bonal Zazo, J. L. y Lorenzo-Cáceres, M. P. O. (2017). Criterios de certificación y auditoría de repositorios digitales seguros en archivos. En: *Da produção à preservação informacional: desafios e oportunidades.* Nelson Vaquinhas, Marisa Caixas y Helena Vinagre

(dir.), (pp. 529-550). Universidade de Évora, Centro Interdisciplinar de História, Culturas e Sociedades. (Biblioteca. Estudos & Coloquios; Série E-books; 8). https://doi.org.10.4000/books.cidehus.2563

Bradley, K. (2007). Defining digital sustainability. *Library Trends*, *56*(1), 148-163. https://doi.org.10.1353/lib.2007.0044

Brown, A. (2008). *Digital preservation guidance note 1: selecting file formats for long-term preservation*. The National Archives. https://cdn.nationalarchives.gov.uk/documents/selecting-file-formats.pdf

Budapest Open Access Initiative (1-2 de diciembre de 2001). World Conference on Science, Budapest, Hungría. https://www.budapestopenaccessinitiative.org/

Cantara, L. (2005). METS: the Metadata Encoding and Transmission Standard. *Cataloging & Classification Quarterly*, *40*(3-4), 237-253. https://doi.org/10.1300/J104v40n03_11

Caplan, P. y Guenther, R. (2005). Practical preservation: the PREMIS experience. *Library Trends*, *54*(1), 111-124. https://doi.org/10.1353/lib.2006.0002

Clifford Lynch, C. (2000). Authenticity and integrity in the digital environment: an exploratory analysis of the central role of trust. En: *Authenticity in a digital environment* (pp. 32-50). Council on Library and Information Resources. https://www.clir.org/wp-content/uploads/sites/6/pub92.pdf

Consultative Committee for Space Data Systems (2012). *Recomendation for Space Data System Practices. Reference Model for an Open Archival Information System (OAIS)*. Magenta Book. CCSDS. (Recommended Practice; 2) https://dldc.lib.uchicago.edu/dl/OAIS_2012-06.pdf

Corda, M. C.; Viñas, M. y Vallefín, C. (2020). Preservar la producción académica digital para el futuro: políticas diseñadas en los repositorios de Argentina. *Informatio*, *25*(2), 41-61. https://sedici.unlp.edu.ar/handle/10915/119230

Cornell University Library (2007). *Digital preservation management: implementing short-term strategies for long-term problems*. https://dpworkshop.org/dpm-eng/index.html

Cramer, T.; German, C.; Jefferies, N. y Wise, A. (2023) A perpetual motion machine: the preserved digital scholarly record. *Learned Publishing*, *36*(2), 312-318. https://doi.org/10.1002/leap.1494

Criterios básicos para valorar sistemas de preservación digital (2020). Universidad Nacional Autónoma de México, Instituto de Investigaciones Bibliográficas, Área de Tecnología del Grupo de Preservación Digital. (Serie Instrumenta Bibliographica; 1). https://www.iib.unam.mx/files/iib/libros-electronicos/Criterios-Basicos-Sistemas-Preservacion-Digital.pdf

Cruz Mundet, J. R. (2015). Estrategias de preservación digital permanente en los archivos nacionales: un estudio comparativo. *Boletín ANABAD*. *65*(3), 127-148. https://dialnet.unirioja.es/servlet/articulo?codigo=5320603

Cruz Mundet, J. R. y Díez Carrera, C. (2016). Sistema de Información de Archivo Abierto (OAIS): luces y sombras de un modelo de referencia. *Investigación Bibliotecológica: Archivonomía, Bibliotecología e Información*, *30*(70), 221-247. https://doi.org/10.1016/j.ibbai.2016.10.010

Dappert, A. y Enders, M. (2010). Digital preservation metadata standards. *Information Standards Quarterly*, *22*(2), 4-13. https://doi.org.10.3789/isqv22n2.2010.01

De Giusti, M. R.; Lira, A. J.; Villarreal, G. L. y Texier, J. D. (noviembre de 2012). *Las actividades y el planeamiento de la preservación en un repositorio Institucional* [Presentación]. Conferencia Internacional BIREDIAL-ISTEC Acceso Abierto, Comunicación Científica y Preservación Digital, Barranquilla, Colombia. http://sedici.unlp.edu.ar/handle/10915/26045

De Giusti, M. R. (2014). *Una metodología de evaluación de repositorios digitales para asegurar la preservación en el tiempo y el acceso a los contenidos* [Tesis de doctorado, Universidad Nacional de La Plata]. SEDICI Repositorio Institucional de la Universidad Nacional de La Plata. http://sedici.unlp.edu.ar/handle/10915/43157

De Giusti, M. R. (5 de noviembre de 2020) *Preservación digital: normas, prácticas y acciones recomendadas desde un repositorio institucional* [Presentación]. Segundo Encuentro de Preservación Digital, Ciudad de México, México. http://sedici.unlp.edu.ar/handle/10915/108454

Diccionario de datos PREMIS de metadatos de preservación, versión 2.0. (2015). Biblioteca Nacional de España. (Publicaciones Técnicas). https://www.loc.gov/standards/premis/PREMIS_es.pdf

Digital Preservation Coalition (2015). *Digital preservation handbook*. 2nd ed. https://www.dpconline.org/handbook

Electronic Resource Preservation and Access Network (2003). *Digital preservation policy tool*. Information Society Technologies. (ERPA Guidance). http://www.erpanet.org/guidance/docs/ERPANETPolicyTool.pdf

El Idrissi, B. (7-9 de marzo de 2019). *Long-term digital preservation: a preliminary study on software and format obsolescence* [Presentación]. Proceedings of the 6th Annual International Conference on Arab

Women in Computing, Rabat, Marruecos. https://doi.org.10.1145/3333165.3333178

Formenton, D. y De Souza Gracioso, L. (2022). Metadata standards in web archiving technological resources for ensuring the digital preservation of archived websites. *Revista Digital de Biblioteconomia e Ciência da Informação*, *20*, e022001. https://doi.org.10.20396/rdbci.v20i00.8666263/27830

Formenton, D. y De Souza Gracioso, L. (2020). Preservação digital: desafios, requisitos, estratégias e produção científica. *RDBCI: Revista Digital de Biblioteconomia e Ciência da Informação*, *18*, e020012. https://doi.org/10.20396/RDBCI.V18I0.8659259

Fushimi, M.; Banzato, G. (9-10 de diciembre de 2010). *Las políticas de acceso abierto en las universidades estatales argentinas: un análisis a través de la web* [Presentación]. VI Jornadas de Sociología de la UNLP. Debates y perspectivas sobre Argentina y América Latina en el marco del Bicentenario. Reflexiones desde las Ciencias Sociales, La Plata, Argentina. http://www.memoria.fahce.unlp.edu.ar/trab_eventos/ev.931/ev.931.pdf

Garrett, J. y Waters, D. (1996). *Preserving digital information, report of the Task Force on Archiving of Digital Information*. Commission on Preservation and Access; Research Libraries Group. https://www.clir.org/wp-content/uploads/sites/6/pub63watersgarrett.pdf

Granger, S. (2000). Emulation as a digital preservation strategy. *D-Lib Magazine*, *6*(10). https://www.dlib.org/dlib/october00/granger/10granger.html

Elizalde, E.; Ferrando, C. y Vergara Rossi, F. (7-8 de noviembre de 2013). *Repositorio Digital Institucional de la Universidad de Buenos Aires* [Presentación]. XIII Jornada sobre la Biblioteca Digital Universitaria

"Gestión del conocimiento en el entorno digital", Córdoba, Argentina. https://repositoriouba.sisbi.uba.ar/gsdl/collect/event/index/assoc/HWA_806.dir/806.PDF

Hedstrom, M y Lee, C. A. (6-8 de mayo de 2002) *Significant properties of digital objects: definitions, applications, implications* [Presentación]. Proceedings of the DLM-Forum "@ccess and preservation of electronic information: best practices and solutions". Barcelona, España. https://ils.unc.edu/callee/sigprops_dlm2002.pdf

Hernández Peréz, T.; Rodríguez Mateos, D. y Bueno De la Fuente, G. (2008). Open Access: el papel de las bibliotecas en los repositorios institucionales de acceso abierto. *Anales de Documentación*, *10*, 185-204. https://revistas.um.es/analesdoc/article/view/1141

International Association of Sound and Audiovisual Archives (2009). *Directrices para la producción y preservación de objetos digitales de audio*. 2a ed. IASA. (Prácticas y Estrategias Recomendadas; IASA-TC 04). https://www.iasa-web.org/audio-preservation-tc04

InterPARES (2010). *Glosario InterPARES de preservación digital*. International Research on Permanent Authentic Records in Electronic Systems, Team México. http://interpares.org/display_file.cfm?doc=ip3_mx_glosario_interpares_v1-2.pdf

InterPARES/ICA (2012). Developing policy and procedures for digital preservation. En: *Digital records pathways: topics in digital preservation*. International Research on Permanent Authentic Records in Electronic Systems; International Council on Archives. http://www.interpares.org/ip3/display_file.cfm?doc=ip3_canada_gs12_module_2_july-2012_DRAFT.pdf

Justrell, B. (10-14 de septiembre de 2006). *Digital preservation: challenges and opportunities* [Presentación]. CIDOC 06. Gotemburgo, Suecia. https://cidoc.mini.icom.museum/wp-content/uploads/sites/6/2018/12/Justrell_Borje.pdf

Lavoie, B. y Gartner, R. (2013). Preservation metadata. *Digital Preservation Coalition Technology Watch Report*, *13*(13). https://www.dpconline.org/docs/dpc-technology-watch-publications/technology-watch-reports-1/894-dpctw13-03/file

Lawrence, G. W.; Kehoe, W. R.; Rieger, O. Y.; Walters, W. H. y Kenney, A. R. (2000). *Risk management of digital information: a file format investigation*. Council on Library and Information Resources. https://www.clir.org/pubs/reports/pub93/

Lee, K. H.; Slattery, O.; Lu, R.; Tang, X. y McCrary, V. (2002). The state of the art and practice in digital preservation. *Journal of Research of the National Institute of Standards and Technology*, *107*(1), 93-106. https://doi.org.10.6028/jres.107.010

Lechich, R. (2007). *File format identification and validation tools*. Library of Congress, National Digital Information Infrastructure and Preservation Program. https://www.digitalpreservation.gov/education/documents/1_Identify_Module_Preread.pdf

Leija Román, D.A. (2017). *Preservación digital distribuida y la colaboración interinstitucional: modelo de preservación digital para documentos con fines de investigación en universidades de México* [Tesis doctoral, Universidad de Barcelona]. Dipòsit Digital de la Universitat de Barcelona. https://diposit.ub.edu/dspace/bitstream/2445/117368/1/DALR_TESIS.pdf

Leija Román, D. A. (2023). La política de preservación digital: modelos y elementos clave para su redacción. *BiD: Textos Universitaris de Biblioteconomia i Documentació*, *50*. https://doi.org.10.1344/BiD2023.50.12

Leija Román, D. A. y Térmens, M. (2019). *Niveles de Preservación Digital NDSA 2019*. Traducción al español de la versión 2.0. Asociación Iberoamericana de Preservación Digital. https://osf.io/egjk8

Library of Congress (2015). *Sustainability of digital formats: planning for Library of Congress collections*. https://www.loc.gov/preservation/digital/formats/sustain/sustain.shtml

Library of Congress (2022). *METS: introducción y tutorial*. https://www.loc.gov/standards/mets/METSOverview_spa.html

Lin, D.; Crabtree, J.; Dillo, I.; Downs, R. R.; Edmunds, R.; Giaretta, D.; De Giusti, M.; L'Hours, H.; Hugo, W.; Jenkyns, R.; Khodiyar, V.; Martone, M. E.; Mokrane, M.; Navale, V.; Petters, J.; Sierman, B.; Sokolova, D. V.; Stockhause, M. y Westbrook, J. (2020) The TRUST principles for digital repositories. *Scientific Data*, *7*(144). https://doi.org/10.1038/s41597-020-0486-7

Mell, P. y Grance, T. (2011). The NIST Definition of Cloud Computing. Recommendations of the National Institute of Standards and Technology. (Special Publication; 800-145). Department of Commerce, National Institute of Standards and Technology. https://nvlpubs.nist.gov/nistpubs/legacy/sp/nistspecialpublication800-145.pdf

Méndez Saavedra, L. O. (2020) Algunos apuntes sobre la historia de los metadatos de preservación para objetos digitales en el contexto de las bibliotecas. *ACCESO. Revista Puertorriqueña de Bibliotecología y*

Documentación, *1*, Nueva Época. https://revistas.upr.edu/index.php/acceso/article/view/18454

Moro Cabero, M. (2018). Identificación, caracterización y selección de formatos para la preservación del recurso digital. *Métodos de Información*, *9*(16), 5-46. https://doi.org/10.5557/IIMEI9-N16-049090

National Archives and Records Administration (2024). *NARA Digital Preservation Risk Matrix*. https://github.com/usnationalarchives/digital-preservation/blob/master/Digital_Preservation_Risk_Matrix/readme.md

Ochoa-Gutiérrez, J.; Sáenz Giraldo, A. y Tirado Tamayo. T. (2021). Experiencias de gestión de los procesos de preservación digital a partir del modelo OAIS en repositorios institucionales. *Anales de Documentación*, *24*(1). https://doi.org/10.6018/analesdoc.428141

Online Computer Library Center/Research Libraries Group Working Group on Preservation Metadata (2002). *Preservation metadata and the OAIS information model: a metadata framework to support the preservation of digital objects*. Online Computer Library Center. https://www.oclc.org/content/dam/research/activities/pmwg/pm_framework.pdf

Park, E. G. y Oh, S. (2012). Examining attributes of open standard file formats for long-term preservation and open access. *Information Technology and Libraries*, *31*(4), 46-67. https://doi.org/10.6017/ital.v31i4.1946

Pennock, M; Wheatley, P. y May, P. (6-10 de octubre de 2014). *Sustainability assessments at the British Library: formats, frameworks, and findings* [Presentación]. iPRES 2014. Proceedings of the 11th International Conference on Digital Preservation. Melbourne, Australia. https://phaidra.univie.ac.at/detail/o:378110

Rauch, C.; Krottmaier, H. y Tochtermann, K. (junio de 2007). *File-formats for preservation: evaluating the long-term stability of file-formats* [Presentación]. Proceedings ELPUB2007 11th International Conference on Electronic Publishing. Viena, Austria. https://elpub.architexturez.net/system/files/pdf/122_elpub2007.content.pdf

Research Libraries Group/Online Computer Library Center (2002). *Trusted Digital Repositories: attributes and responsibilities: an RLG-OCLC report*. https://www.oclc.org/content/dam/research/activities/trustedrep/repositories.pdf

Rieger, O. Y. (2004). *Metadata standards for managing and discovering image collections: trends* [Objeto de Aprendizaje]. Cornell University Library. https://www.slideserve.com/xaviera-ferguson/metadata-standards-for-managing-and-discovering-image-collections-trends

Rivera Donoso, M. A. (2009). *Directrices para la creación de un programa de preservación digital*. Universidad Tecnológica Metropolitana, Departamento de Gestión de Información. (Serie Bibliotecología y Gestión de Información; no. 43). http://eprints.rclis.org/12989/1/Serie_N%C2%B043_Preservacion_digital.pdf

Ross, S. y McHugh, A. (2005). Audit and certification of digital repositories: creating a mandate for the Digital Curation Centre (DCC). *RLG DigiNews*. https://worldcat.org/arcviewer/1/OCC/2007/08/08/0000070511/viewer/file2396.html

Salvador Benítez, A; Ruíz Rodríguez, A. A. (2005). Metadatos para la preservación de colecciones digitales. *Cuadernos de Documentación Multimedia*, *16*, 21-36. https://revistas.ucm.es/index.php/CDMU/article/view/58930/52988

Schaefer, S. K.; McGovern, N. Y.; Zierau, E. M. O. y Wu, C. C. M. (2021). Deciding how to decide: using the digital preservation storage criteria. *International Federation of Library Associations Journal*, *48*(2), 318-331. https://doi.org/10.1177/03400352211011490

Senso, J. A.; De la Rosa, A. (2003). El concepto de metadato: algo más que descripción de recursos electrónicos. *Ciência da Informação*, *32*(2), 95-106. https://doi:10.1590/S0100-19652003000200011

Serrano Vicente, R.; Melero Melero, R. y Abadal, E. (2014). Indicadores para la evaluación de repositorios institucionales de acceso abierto. *Anales de Documentación*, *17*(2). https://doi.org/10.6018/analesdoc.17.2.190821

Sierman, B.; Jones, C. y Elstrøm, G. (2014). *Catalogue of preservation policy elements*. SCAPE (Scalable Preservation Environments). https://scape-project.eu/wp-content/uploads/2014/02/SCAPE_D13.2_KB_V1.0.pdf

Suber, P. (2006). *Una introducción al acceso abierto*. En: Babini, D.; Fraga, J. Edición electrónica, bibliotecas virtuales y portales para las ciencias sociales en América Latina y El Caribe (pp. 15-33). Consejo Latinoamericano de Ciencias Sociales. https://biblioteca.clacso.edu.ar/clacso/se/20100528031534/2Peter.pdf

Térmens, M.; Leija Román, D. A. (2017). Auditoría de preservación digital con NDSA Levels. *El profesional de la información*, *26*(3), 447-456. https://doi.org/10.3145/epi.2017.may.11

Tettamanti, S.; De Giusti, M. R. y Lira, A. J. (3-7 de octubre de 2022). Evaluación de un repositorio institucional a través de NDSA Levels: Caso CIC Digital [Presentación]. *Actas de la XI Conferencia Internacional de Bibliotecas y Repositorios Digitales (BIREDIAL-ISTEC)* (pp. 304-320). Buenos Aires, Argentina. http://sedici.unlp.edu.ar/handle/10915/148622

Universidad de Buenos Aires. Consejo Superior (2013-03-13). Resolución (CS) Nº 6323/13 - Creación del Repositorio Digital Institucional de la Universidad de Buenos Aires. https://repositoriouba.sisbi.uba.ar/gsdl/collect/normauba/index/assoc/HWA_434.dir/434.PDF

UNESCO (2003). *Carta de la UNESCO para la Preservación del Patrimonio Digital*. 2a ed. Organización de las Naciones Unidas para la Educación, la Ciencia y la Cultura, Consejo Intergubernamental del Programa Información para Todos. https://unesdoc.unesco.org/ark:/48223/pf0000229034_spa

UNESCO (2003). *Directrices para la preservación del patrimonio digital*. Organización de las Naciones Unidas para la Educación, la Ciencia y la Cultura, División de la Sociedad de la Información. https://unesdoc.unesco.org/ark:/48223/pf0000130071_spa

Vermaaten, S.; Lavoie, B. y Caplan, P. (2012). Identifying threats to successful digital preservation: the SPOT Model for Risk Assessment. *D-Lib Magazine*, *18*(9/10). https://doi.org.10.1045/september2012-vermaaten

Zapata Cárdenas, C. A. (2023). La preservación digital de documentos electrónicos de archivo: una materia pendiente en América Latina. *Revista del Archivo Nacional de Costa Rica*, 87, e596. https://www.dgan.go.cr/ran/index.php/RAN/article/view/596/508

Anexo 1. Recomendaciones finales

A partir del análisis realizado y de las conclusiones alcanzadas, se presentan a continuación una serie de recomendaciones orientadas a fortalecer las prácticas de preservación digital en los repositorios institucionales de acceso abierto de la Universidad de Buenos Aires. Estas propuestas buscan ofrecer lineamientos viables, progresivos y sostenidos en el tiempo, capaces de contribuir a la consolidación de una política universitaria común que garantice la preservación, autenticidad y accesibilidad del patrimonio digital de la institución.

Marco normativo y articulación institucional

Se propone la elaboración y aprobación de una política marco de preservación digital, que establezca lineamientos mínimos comunes y asegure la armonización con los criterios institucionales de cada Facultad. Esta política debe ser formalizada mediante una resolución institucional que asegure su cumplimiento y la asignación de recursos.

Las Facultades que aún no lo han hecho, deben documentar y publicar sus propias políticas y procedimientos internos de preservación digital que definan con precisión los objetivos, responsabilidades, estrategias, procedimientos y estándares adoptados. Se recomienda que esta documentación sea de carácter público y se encuentre disponible para su acceso desde el sitio web de cada RI.

Es esencial fomentar la colaboración entre facultades para el intercambio de buenas prácticas que incorporen aspectos centrales como la selección de formatos, las estrategias de migración, la documentación de procedencia, el registro de metadatos de preservación, los mecanismos de control de integridad, etc.

Gestión de recursos

Es crucial establecer y garantizar un presupuesto anual exclusivo para las tareas de preservación digital en los repositorios, superando la carencia identificada en la totalidad de las instituciones encuestadas. El presupuesto debe cubrir tanto la infraestructura tecnológica orientada específicamente a la preservación como a la asignación de recursos humanos con perfiles técnicos especializados en preservación digital. Es importante fomentar el desarrollo de proyectos cooperativos o convenios externos orientados a fortalecer la infraestructura y las capacidades técnicas.

En materia de almacenamiento, copias de seguridad y contingencia, se recomienda establecer procedimientos formales y estandarizados que definan frecuencias de respaldo, responsables y ubicaciones de almacenamiento. Resulta pertinente promover estrategias de redundancia más robustas, apoyadas en modelos mixtos de almacenamiento (local, externo y en la nube) que aseguren el mantenimiento de al menos tres copias de seguridad de los objetos digitales en localizaciones geográficamente separadas. Con el objetivo de explorar soluciones cooperativas, se sugiere analizar la viabilidad de implementar sistemas de almacenamiento compartidos que permitan la optimización de recursos y capacidades técnicas de cada unidad académica.

El fortalecimiento de los recursos humanos constituye otro eje estratégico. Se sugiere asignar personal específicamente dedicado a la preservación digital, evitando que esta tarea recaiga en trabajadores que ya poseen múltiples responsabilidades y que, por ello, no pueden garantizar una atención sostenida a los procesos de preservación. Para potenciar estas iniciativas, sería valioso promover una red de colaboración entre facultades que facilite el intercambio de experiencias, el asesoramiento técnico y el desarrollo conjunto de soluciones.

Plan de contingencia

La elaboración y documentación de un plan de contingencia y recuperación de datos ante desastres se vuelve imperiosa, dado que constituye uno de los puntos más débiles relevados en la mayoría de los repositorios. Estos planes deberían prever tanto la continuidad operativa como el acceso sostenido al patrimonio digital en situaciones críticas.

Complementariamente, se destaca la importancia de evaluar periódicamente la posible obsolescencia de los formatos utilizados y promover procesos de migración cuando sea necesario. La implementación de mecanismos automatizados de monitoreo y detección de corrupción de archivos permitiría reducir la dependencia de procesos manuales y aumentar la fiabilidad del sistema.

Normalización

Resulta indispensable establecer la obligación de registrar sistemáticamente metadatos de preservación, siguiendo estándares internacionales como PREMIS y METS, para documentar la procedencia, autenticidad, entorno técnico e historial de los objetos digitales.

A su vez, es fundamental desarrollar un listado de formatos de archivo preferidos y aceptados para la preservación a largo plazo, y documentar los criterios de evaluación, identificación y selección de formatos para cada tipo de objeto digital. Esto permitirá evaluar periódicamente la obsolescencia de formatos y promover la migración controlada, cuando sea necesario, para garantizar la accesibilidad y sostenibilidad de los activos digitales en el tiempo.

Controles de integridad y seguridad

Se recomienda implementar mecanismos automatizados para la verificación periódica y permanente de la integridad de los objetos digitales (*checksums*), reduciendo la dependencia de la escasez de personal y de las verificaciones manuales.

Asimismo se debe establecer la verificación de virus como un paso obligatorio y sistemático en el flujo de trabajo de ingesta de todos los objetos digitales al repositorio, y la implementación de autenticación de usuarios específicos para el acceso a las funciones administrativas y de preservación de los repositorios, mejorando la seguridad y permitiendo un análisis más preciso del uso de los recursos.

Creación de un Programa de Preservación Digital en el ámbito de la Universidad de Buenos Aires

Se propone desarrollar estándares comunes, herramientas compartidas y procesos de auditoría periódica en la UBA. La creación de un *Programa UBA de Preservación Digital* contribuiría a institucionalizar estas acciones, centralizando la documentación, ofreciendo soporte técnico especializado y promoviendo la mejora continua de los repositorios.

En este sentido, es recomendable la elaboración de criterios institucionales explícitos, respaldados por estándares internacionales y buenas prácticas del campo, lo que favorecerá que la toma de decisiones sea más coherente y que se sostenga en el tiempo, reduciendo la vulnerabilidad del acervo digital ante los cambios tecnológicos.

Por último, se considera necesario consolidar una cultura institucional de seguimiento, monitoreo y mejora continua en materia de preservación digital. Ello incluye la implementación de evaluaciones regulares del estado de los repositorios, la detección temprana de posibles fallas técnicas o

inconsistencias y la generación de espacios de capacitación destinados al personal responsable de la gestión del patrimonio digital. La articulación de estas acciones permitirá garantizar una mayor sostenibilidad de los repositorios en el mediano y largo plazo, fortaleciendo así el rol de la institución como productora, custodia y difusora de conocimiento.

Anexo 2. Encuesta

Prácticas de preservación digital en los repositorios institucionales de acceso abierto de la Universidad de Buenos Aires

Apellido, nombre:

Institución:

Nombre del repositorio:

Cargo o función en el repositorio:

Correo electrónico de contacto:

¿Qué software de gestión para repositorios se implementa?

- DSpace
- Greenstone
- EPrints
- Omeka
- Desarrollo propio
- Otro

Si marcó la opción Otro, especifique:

¿Cuántas personas trabajan en el repositorio con dedicación de tiempo parcial o exclusivo?

¿Cuántas personas realizan actividades relacionadas con la preservación digital con dedicación de tiempo parcial o exclusivo?

¿El repositorio cuenta con un presupuesto asignado únicamente para tareas de preservación digital?

Sí

No

¿El repositorio posee políticas explícitas de preservación digital?

Sí

No

Si el repositorio posee políticas de preservación digital:

Son de acceso público

Están explícitas en un documento interno

Están detalladas en una resolución institucional

Otro

Si marcó la opción Otro, especifique:

Si el repositorio no posee políticas de preservación digital. ¿Se considera su redacción, publicación e implementación en el futuro?

Sí

No

¿Se dispone de un procedimiento para realizar copias de seguridad?

Sí

No

¿Se dispone de, por lo menos, dos copias de todos los objetos digitales que componen el repositorio?

Sí

No

En el caso de que la respuesta anterior haya sido afirmativa. ¿En qué espacio se alojan esas copias?

Almacenamiento propio

Almacenamiento externo

Otro

Si marcó la opción Otro, especifique:

¿Existe un plan de contingencia para una eventual falla en el almacenamiento?

Sí

No

¿Qué aspectos del almacenamiento de datos considera que deberían ser mejorados? (Por ejemplo: documentar procedimientos, establecer planes de contingencia, buscar alternativas de almacenamiento, etc.)

¿Se comprueba regularmente la integridad de los objetos digitales del repositorio?

Sí

No

En el caso de que la respuesta anterior haya sido afirmativa. ¿Cada cuánto tiempo se realiza dicha comprobación?

Si el control de integridad no está siendo realizado. ¿Cuáles son las razones?

Falta de personal

Falta de tiempo

Falta de medios adecuados

Otra/s

Si marcó la opción Otra/s, mencione cuál/es:

¿Se realiza la comprobación de virus al momento de la incorporación de los objetos al repositorio?

Siempre

A veces

Nunca

¿Se definen tipos de usuarios que se deben identificar al momento de acceder al repositorio (login)?

Sí

No

¿Se registran metadatos estructurales y administrativos de los objetos digitales? (procedencia, formato, propiedades técnicas, etc.)

Sí

No

¿Se registran metadatos de preservación? (migración de formatos, controles de antivirus y de integridad, propiedad intelectual, etc.)

Sí

No

¿Al momento de almacenar los objetos digitales se incrustan metadatos descriptivos mediante un software destinado a tal fin? por ejemplo AutoMetadata, ABBYY, Photo Helper, Vidmore, etc.

Sí

No

¿Se asignan identificadores únicos y persistentes para cada objeto digital?

Sí

No

En el caso de que la respuesta anterior haya sido afirmativa, especifique qué tipo de identificador se utiliza (por ejemplo DOI, URI, ARK, HANDLE, etc.)

¿Se definen formatos de preservación a largo plazo para cada tipo de objeto digital que se incorpora en el repositorio? (por ejemplo PDF/A, JPEG2000, TIFF, GZIP, etc.)

Sí

No

Si fuera necesario. ¿Se realiza la migración de los objetos digitales a incorporar al repositorio de su formato original a un formato de preservación a largo plazo?

Sí

No

¿Se realiza un monitoreo periódico sobre posibles problemas de obsolescencia en los formatos de los objetos del repositorio?

Sí

No

¿Desea comentar algún aspecto que no haya sido contemplado en las preguntas anteriores?

Anexo 3. Siglas y abreviaturas utilizadas

AIP	Archival Information Packages
ALA	Asociación Latinoamericana de Archivos
APREDIG	Asociación Iberoamericana de Preservación Digital
ARK	Archival Resource Key
BOAI	Budapest Open Access Initiative
CCSDS	Consultative Committee for Space Data Systems
CI	Content Information
CLIR	Council on Library and Information Resources
COAR	Community Framework for Good Practices in Repositories
COPTR	Community Owned Digital Preservation Tool Registry
CPA	Commission on Preservation and Access
CRL	Center for Research Libraries
DI	Descriptive Information
DIP	Dissemination Information Package
DLF	Digital Library Federation
DPC	Digital Preservation Coalition
DOI	Digital Object Identifier
DORA	Declaration on Research Assessment

EAD	Encoded Archival Description
ERPANET	Electronic Resource Preservation and Access Network
et al.	y otros (latín)
FECYT	Fundación Española para la Ciencia y la Tecnología
FIDO	Format Identification for Digital Objects
FITS	File Information Tool Set
GDFR	Global Digital Format Registry
ICA	International Council on Archive
IFLA	International Federation of Library Associations and Institutions
IP	Information Package
iPRES	International Conference on Digital Preservation
ISO	International Organization for Standardization
JHOVE	JSTOR/Harvard Object Validation Environment
LC	Library of Congress
LOCKSS	Lots Of Copies Keep Stuff Safe
METS	Metadata Encoding and Transmission Standard
MIME	Multipurpose Internet Mail Extensions
MINCyT	Ministerio de Ciencia, Tecnología e Innovación
NARA	National Archives and Records Administration
NDIIPP	National Digital Information Infrastructure and Preservation

	Program
NDSA	National Digital Stewardship Alliance
NESTOR	Network of Expertise in Long-term STORage of Digital Resources
NIST	National Institute of Standards & Technology
OAI	Open Archives Initiative
OAI-PMH	Open Archives Initiative Protocol for Metadata Harvesting
OAIS	Open Archival Information System
OCLC	Online Computer Library Center
OpenAIRE	Open Access Infrastructure for Research in Europe
OPF	Open Preservation Foundation
PDI	Preservation Description Information
PERSIST	Platform to Enhance the Sustainability of the Information Society Transglobally
PI	Packaging Information
PID	Persistent Identifier
PLANETS	Preservation and Long-Term Access Through Networked Services Project
PREMIS	PREservation Metadata: Implementation Strategies
PURL	Persistent Uniform Resource Locator
REBIUN	Red de Bibliotecas Universitarias Españolas

RDA	Research Data Alliance
RDI-UBA	Repositorio Digital Institucional de la Universidad de Buenos Aires
RI	Repositorios Institucionales
RLG	Research Libraries Group
s/inf.	Sin información
SISBI	Sistema de Bibliotecas y de Información
UBA	Universidad de Buenos Aires