

Preservación digital a largo plazo: estándares, auditoría, madurez y planificación estratégica*

Elba Boderó Poveda

Ph. D. (c) en Ciencias Informáticas de la Universidad Nacional de La Plata (Argentina). Magister Scientiae en Gestión para la Creación Intelectual de la Universidad Politécnica Territorial del Estado Mérida (Venezuela). Máster Degree in Distance Education E-Learning de la Caribbean International University (Curacao - Países Bajos). Magister en Tecnología de la Información y Multimedia Educativa de la Universidad Técnica de Ambato (Ecuador). Ingeniera en Sistemas Informáticos de la Escuela Superior Politécnica de Chimborazo (Ecuador). Profesora titular en la Facultad de Ingeniería de la Universidad Nacional de Chimborazo (Ecuador). Investigadora del Grupo de Investigación en Telecomunicaciones, Informática, Industria y Construcción (TEIIC – UNACH - Ecuador). ebodero@unach.edu.ec
<https://orcid.org/0000-0003-3807-5203>

Marisa R. De Giusti

Ph. D. en Ciencias Informáticas, ingeniera en Telecomunicaciones y profesora en Letras todos títulos de la Universidad Nacional de La Plata (Argentina). Investigadora principal de la Comisión de Investigaciones Científicas de la Provincia de Buenos Aires (Argentina). Directora CIC Digital, directora CESGI, directora PREBI-SEDICL, directora Liblink-ISTEC y profesora de posgrado en la Facultad de Informática de la Universidad Nacional de La Plata (Argentina). marisa.degiusti@sedici.unlp.edu.ar
<https://orcid.org/0000-0003-2422-6322>

Cristian Morales

Máster Executive Management Business Administration de la Universitat de Barcelona (España). Magister en Gestión de Sistemas de Información e Inteligencia de Negocios de la Universidad de las Fuerzas Armadas-ESPE (Ecuador). Ingeniero en Sistemas y Computación de la Universidad Nacional de Chimborazo (Ecuador). Investigador externo de la UNACH (Ecuador). Director Ejecutivo de Esprint (Ecuador). cmorales@unach.edu.ec
<https://orcid.org/0000-0002-0197-0581>

Resumen

Este estudio parte de una revisión sistemática de literatura con la finalidad de describir los estándares, modelos de auditoría, madurez y principios de planificación estratégica para la preservación digital. Tiene un enfoque de investigación cualitativo, es de tipo documental, descriptivo y sistemático y se utilizó el protocolo para una revisión sistemática de literatura. Los resultados indican un total de 309 estudios candidatos, 42 trabajos seleccionados y 24 estudios primarios. Se presentan diez estándares, ocho modelos de auditoría y tres modelos de madurez de preservación digital. Además, se analizaron varios estudios en el ámbito de la planificación estratégica, en aspectos como riesgos, normativa, gestión administrativa, archivística, tecnología, organización, niveles, política de estrategia, estructura para la preservación digital, entre otros.

Palabras clave: preservación digital, estándares, auditorías, modelos de madurez, revisión sistemática, planificación.

Cómo citar este artículo: Boderó-Poveda, Elba; De Giusti, Marisa; Morales, Cristian (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>

Recibido: 2020-10-20/ **Aceptado:** 2022-04-03

* Este estudio forma parte de la tesis doctoral “Modelo de madurez para preservación digital a largo plazo aplicando principios de planificación estratégica” del Doctorado en Ciencias Informáticas de la Universidad Nacional de La Plata.

Long-term Digital Preservation: Standards, Auditing, Maturity, and Strategic Planning

Abstract

This study conducted a systematic literature review in order to describe the standards, audit models, maturity, and strategic planning principles for digital preservation. This work has a qualitative research approach, is documentary, descriptive and systematic. It uses the protocol for a systematic literature review. The results indicate a total of 309 candidate studies, 42 selected papers, and 24 primary studies. Ten standards, eight audit models and three digital preservation maturity models are presented. In addition, several studies in the field of strategic planning are analyzed, in aspects such as risks, regulations, administrative management, archival, technological, organizational, levels, strategy policy, structure for digital preservation, among others.

Keywords: Digital preservation, standards, audits, maturity models, systematic review, planning.

1. Introducción

Empresas e instituciones de todos los sectores de la industria a nivel mundial generan una gran cantidad de información. Esta, al ser almacenada de forma digital, puede significar una ventaja competitiva una vez explotada. Además, esta información, una vez preservada, permite afrontar de manera oportuna procesos de auditoría e investigación. Los repositorios digitales proporcionan una estructura para el almacenamiento y permiten gestionar, preservar y difundir esta información. También brindan un adecuado servicio al cliente, por lo que se convierten en un componente fundamental para el desarrollo y funcionamiento de las organizaciones.

La preservación digital se define como las técnicas o procesos que son necesarios para conservar en el tiempo objetos digitales en un sistema de información (Boté, 2012), y que esta información se encuentre disponible y accesible en un futuro. Se realiza con base en las prácticas más confiables de almacenamiento, gestión de datos y ciclo de vida del contenido digital. A medida que la investigación y la práctica identifican mejores métodos

de preservación, las organizaciones ajustan sus políticas y flujo de trabajo (Gallinger, Bailey, Cariani, Owens & Altman, 2017).

Las organizaciones buscan la preservación digital por diferentes razones, que pueden estar relacionadas con ámbitos legales, financieros, archivísticos, de investigación, entre otros, con la finalidad de que la información digital no sufra alteraciones en el tiempo, se preserve de acuerdo con la aplicación de diferentes principios de seguridad y proporcione accesibilidad para diferentes usuarios.

En las bases de datos científicas, el término *preservación digital* devuelve un gran número de trabajos de investigación, los cuales se enfocan en diferentes ámbitos, como casos de aplicación de preservación digital, modelos de auditoría, modelos de conservación de evidencias, plataformas de software e infraestructura de hardware. Hasta marzo de 2020, la base de datos IEEE Xplore devolvió 225 trabajos relacionados específicamente con “digital preservation”; ACM Digital Library, 425 estudios; Science Direct de Elsevier, 736 resultados; y Google Académico, más de un millón de resultados. Por lo tanto, se puede deducir que este término, en el ámbito científico, se encuentra analizado desde múltiples enfoques, lo que permite identificar una necesidad latente a nivel mundial con respecto a la preservación de la información digital.

Existen varios estándares enfocados en la preservación digital, los cuales permiten a bibliotecarios, usuarios de repositorios, administradores de sistemas de gestión de la información y centros de documentación en general contar con una guía de buenas prácticas para preservar la información que se genera en una organización a lo largo del tiempo. Esta guía debe estar protegida de varios factores que pueden afectar su integridad, autenticidad, flexibilidad, normalización y accesibilidad. De igual manera, existen modelos de auditoría y madurez en este ámbito que permiten definir en qué estado se encuentra una organización en lo que a preservación digital se refiere. Por esta razón, es necesario realizar una revisión sistemática de literatura (SLR) orientada a encontrar cuáles estándares y modelos pueden aportar a la creación de un modelo de madurez para la preservación digital a largo plazo con la aplicación de principios de planificación estratégica.

Una SLR identifica vacíos para efectuar una investigación, responde a una o varias preguntas, fenómenos o áreas de interés. La SLR tiene la finalidad de realizar una evaluación justa, mediante la utilización de una metodología fiable, rigurosa y auditable, que resume la evidencia existente de una temática definida previamente y provee un marco de antecedentes (Kitchenham & Charters, 2007; Marshall & Brereton, 2013).

Este estudio responde a las siguientes preguntas de investigación:

RQ1. ¿Cuáles son los estándares que existen en el ámbito de la preservación digital?

RQ2. ¿Cuáles son los principales modelos de auditoría para preservación digital?

RQ3. ¿Qué trabajos de investigación existen relacionados con modelos de madurez de preservación digital?

RQ4. ¿Existen aportes relacionados con modelos de madurez de preservación digital que incluyan los principios de la planificación estratégica?

RQ1 busca identificar los marcos de trabajo referenciales existentes en torno a la preservación digital (PD); RQ2 extrae los modelos de auditoría que permiten evaluar varios factores alrededor de la PD; RQ3 complementa los modelos de auditoría con una valoración del estado actual de madurez para la conservación de objetos digitales; sin embargo, a pesar de que estos estándares y modelos aportan en gran medida a mejorar la preservación digital, estos no indican cómo llegar a una alta madurez a largo plazo a través de la estrategia. Con el antecedente de que una organización que no aplica la planificación estratégica no podrá llegar a cumplir sus objetivos y metas, se plantea la pregunta RQ4, la cual busca identificar qué existe, en materia de planificación estratégica, en la literatura científica y que pueda aportar de manera significativa a la madurez y auditoría de preservación digital.

2. Metodología

En esta sección se presenta la metodología de investigación, la cual tuvo un enfoque cualitativo y sigue un diseño de teoría fundamentada, debido a que se realizan preguntas sobre procesos y relaciones entre

conceptos que conforman el fenómeno. Es documental debido a que se analizan los trabajos científicos en torno a los estándares, modelos de auditoría, madurez de la preservación digital y explica algunos avances acerca de la estrategia de planificación de la preservación digital. Es sistemática porque sigue un orden basado en el protocolo para la realización de una SLR: 1) preguntas de investigación, 2) criterios de inclusión y exclusión, 3) fuentes de información y criterios de búsqueda, 4) búsqueda y depuración de estudios candidatos, 5) selección de estudios primarios y 6) extracción de características. Esta investigación usó un protocolo de revisión basado en la metodología desarrollada por Kitchenham y Charters (2007), complementada con el estudio de Fonseca-Carrera (2014) y con el artículo realizado por Morales, Radicelli, Jaramillo y Boderó (2018). De esta manera se cumple con un procedimiento formal y riguroso para la extracción de conocimiento de los trabajos científicos.

2.1 Criterios de inclusión y exclusión

Se incluyeron trabajos de un periodo de cinco años respecto a la fecha de realización de esta investigación (2015-2020): trabajos en español y en inglés, artículos de revistas, conferencias, revisiones, capítulos de libros, libros y tesis doctorales, que evidenciaron un alto rigor científico (arbitraje, indexación, edición, entre otros). Se excluyeron trabajos que no presentaban aportes conceptuales ni metodológicos. Se realizó la búsqueda en la base de datos Scopus (título, resumen y palabras clave) y en Google Académico; en esta última con la finalidad de complementar la búsqueda sobre todo de tesis doctorales y libros. Los demás trabajos que no se encontraban bajo los parámetros mencionados fueron excluidos de esta SLR.

2.2 Fuentes de información y criterios de búsqueda

En esta investigación se usó la base de datos científica Scopus —pues esta agrupa publicaciones reconocidas por su alto rigor científico como IEEE Xplore, ACM Digital Library, Science Direct de Elsevier, Springer, entre otras— y Google Académico, para la búsqueda de tesis doctorales y libros. A continuación, se muestran las cadenas de búsqueda que llevaron a recopilar los documentos y posteriormente a responder las preguntas de investigación planteadas. Estas cadenas de

búsqueda fueron construidas tomando como base estudios relacionados con las preguntas de investigación, que formaron un grupo de control, y fueron depuradas hasta llegar a las siguientes:

- Cadena de búsqueda para la base de datos científica Scopus: *TITLE-ABS-KEY ((estándar OR standard OR modelo OR model OR metodología OR methodology) AND ("Preservación digital" OR "Digital preservation")) AND (LIMIT-TO (DOCTYPE , "ar") OR LIMIT-TO (DOCTYPE , "cp") OR LIMIT-TO (DOCTYPE , "re") OR LIMIT-TO (DOCTYPE , "ch") OR LIMIT-TO (DOCTYPE , "bk") OR LIMIT-TO (DOCTYPE , "cr") OR LIMIT-TO (DOCTYPE , "ed")) AND (LIMIT-TO (PUBYEAR , 2020) OR LIMIT-TO (PUBYEAR , 2019) OR LIMIT-TO (PUBYEAR , 2018) OR LIMIT-TO (PUBYEAR , 2017) OR LIMIT-TO (PUBYEAR , 2016) OR LIMIT-TO (PUBYEAR , 2015)) AND (LIMIT-TO (LANGUAGE , "English") OR LIMIT-TO (LANGUAGE , "Spanish"))*.
- Cadena de búsqueda para la base de datos científica de Google Académico: *(estándar OR standard OR modelo OR model OR metodología OR methodology) + ("Preservación digital" OR "Digital preservation") + (doctoral thesis or tesis doctoral or book or libro)*.

Para la búsqueda en la base de datos de Google Académico, además, se indicó en la plataforma el “intervalo específico”, que corresponde al periodo comprendido entre 2015 y 2020, ordenado por relevancia. Al no poder especificarse mediante la cadena de búsqueda, el idioma y el tipo de documento se descartaron posteriormente en la revisión individual, para cumplir con los criterios de inclusión y exclusión.

3. Resultados

Una vez ejecutadas las cadenas de búsqueda al 7 de julio de 2020, se presentaron en la base de datos Scopus un total de 249 documentos, que correspondieron a los estudios candidatos. Estos fueron revisados a partir del título, el resumen y las palabras clave para obtener los estudios seleccionados. La búsqueda en Google Académico dio como resultado 60 documentos. Sumados los estudios de las dos bases de datos científicas se obtuvo un total de 309 trabajos de investigación para revisión.

3.1 Búsqueda y depuración de estudios candidatos

Una vez revisado el título, el resumen y las palabras clave, tal y como lo estipula el protocolo de SLR, de los 249 trabajos de Scopus y los 60 estudios de la base de datos Google Académico, se seleccionaron 25 y 17 investigaciones respectivamente, considerando que estos pueden responder al menos una de las preguntas de investigación. Luego se procedió con la recuperación de los documentos, para analizarlos a texto completo y comprobar que cumplieran con los criterios de inclusión y exclusión planteados en esta SLR.

3.2 Selección de estudios primarios

Los siguientes estudios han permitido dar respuesta a las preguntas de investigación en lo que concierne a estándares, metodologías o modelos de auditoría, modelos de madurez de preservación digital y finalmente trabajos de investigación directamente ligados a la planificación estratégica. Los trabajos de investigación antecedidos por el prefijo “GC” pertenecen al grupo de control que permitió conformar las cadenas de búsqueda; los estudios antecedidos por “GA” son los recuperados por la base de datos de Google Académico, y los manuscritos antecedidos por “SC” pertenecen a los estudios primarios recuperados de la base de datos Scopus. Sumados hacen un total de 24 documentos.

3.2.1 Tesis doctorales

- GC-1. *Propuesta de un modelo de preservación digital para pequeñas y medianas instituciones sanitarias (Boté, 2012)*.
- GC-2. *Una metodología de evaluación de repositorios digitales para asegurar la preservación en el tiempo y el acceso a los contenidos (De Giusti, 2014)*.
- GC-3. *Preservación digital distribuida y la colaboración interinstitucional: Modelo de preservación digital para documentos con fines de investigación en universidades de México (Leija, 2017)*.
- GC-4. *Propuesta de un modelo conceptual de preservación digital aplicable a la mejora de la preservación de evidencia digital en instituciones de investigación criminal (Molina, 2017)*.

3.2.2 Artículos científicos

- GC-5. “Modelo para la preservación de documentos digitales” (Castillo & Umaña, 2020).
- GC-6. “Towards a capability maturity model for digital forensic readiness” (Englbrecht, Meier & Pernul, 2019).
- GC-7. “Digital Preservation Capability Maturity Model (DPCMM): Genesis and Practical Uses” (Ashley & Mistic, 2019).
- GC-8. “Estrategias de preservación digital permanente en los archivos nacionales: un estudio comparativo” (Cruz, 2015).
- GA-1. “Continuidad digital en los archivos del sector transporte e infraestructura de Costa Rica: diseño y planificación de un proyecto para la mejora organizacional del Consejo Nacional de Concesiones” (Chacón, 2018).
- GA-2. “Visibilidad e impacto de la literatura gris científica en repositorios institucionales de acceso abierto. Estudio de caso bibliométrico del repositorio Gredos de la Universidad de Salamanca” (Ferrerías, 2016).
- GA-3. “Repositorios institucionales digitales: Análisis comparativo entre SEDICI (Argentina) y Kérwá (Costa Rica)” (Sandí & Cruz, 2017).
- SC-1. “Library eArchiving with ZONTAL Space and the Allotrope Data Format” (Della, Colman, Welker & Rennick, 2020).
- SC-2. “Revisión de distintas implementaciones para preservación digital: hacia una propuesta metodológica para la preservación y la auditoría de confiabilidad de RI” (De Giusti & Villareal, 2018).
- SC-3. “Different Preservation Levels: The Case of Scholarly Digital Editions” (Oltmanns, Hasler, Peters-Kottig & Kuper 2019).
- SC-4. “Open Archival Information System (OAIS): lights and shadows of a reference model” (Cruz & Diez, 2016).
- SC-5. “La importancia de la normalización para el ejercicio profesional del archivista” (Moro & Llanes, 2018).
- SC-6. “Preliminary Study About the Applicability of a Service-Oriented Architecture in the OAIS Model Implementation” (Neto, Borges & Roque, 2017).
- SC-7. “A Survey of Digital Preservation Challenges in Nigerian Libraries: Librarians' Perspectives” (Ejitagha, Sambo & Urhefe 2017).
- SC-8. “Auditoría de preservación digital con NDSA Levels” (Leija & Térmens, 2017).
- SC-9. “Modeling the value of Digital Preservation Activities” (Borbinha, Proença & Tribolet, 2016).
- SC-10. “Cultivate, assess, advocate, implement, and sustain” (Cote & Mannheimer, 2017).
- SC-11. “Digital Preservation Archives. A New Future Architecture for Long-term Interoperability” (Kearney, Giaretta, Garrett & Hughes, 2018).
- SC-12. “Trusting records: is Blockchain technology the answer?” (Lemiux, 2016).
- SC-13. “Bringing Content into the Picture: Proposing a Tri-Partite Model for Digital Preservation” (Corrado & Moulaison, 2017).

3.3 Extracción de características

3.3.1 RQ1. ¿Cuáles son los estándares que existen en el ámbito de la preservación digital?

Una vez analizados los artículos a texto completo, los estándares enfocados en la preservación a largo plazo se representan en la Figura 1. Se puede observar un diagrama de red asociativo, en el cual existen 10 estándares, relacionados con 21 trabajos científicos. Estos últimos representan el 87,5 % de los estudios primarios, lo cual da a conocer la importancia de la aplicación de buenas prácticas en la preservación digital.

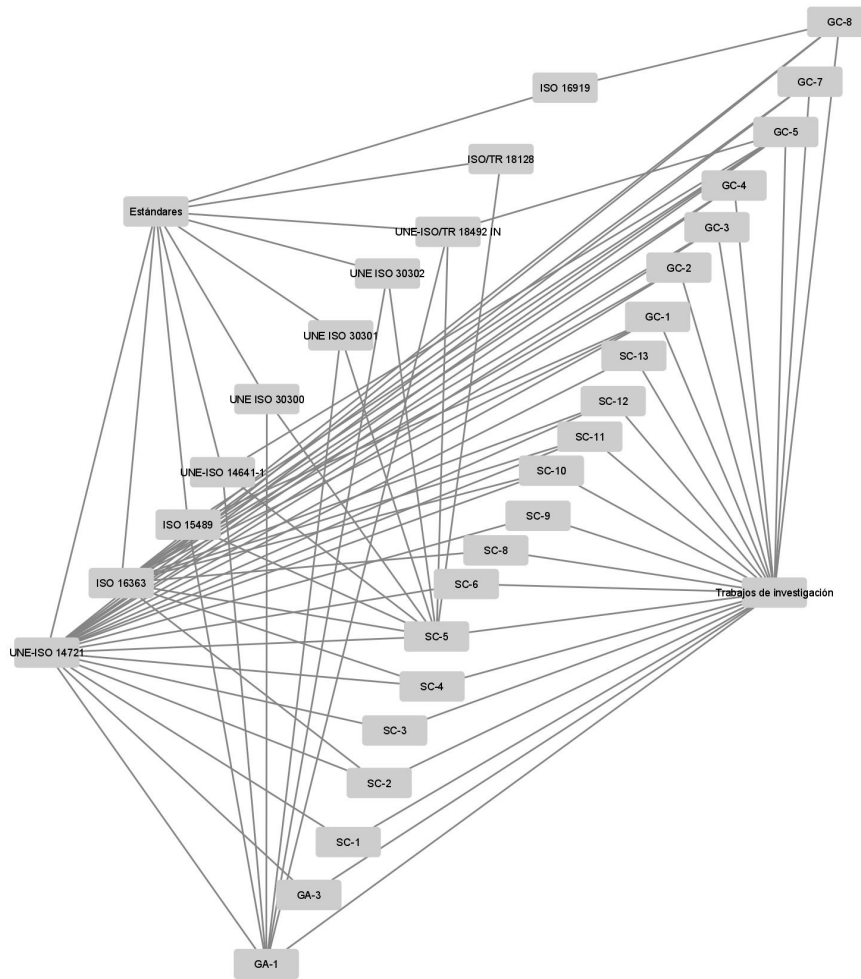


Figura 1. Diagrama asociativo de los estándares en la preservación digital a largo plazo.
Fuente: elaboración propia.

● **UNE-ISO 14721**

“Sistemas de transferencia de datos e información espaciales. Sistema abierto de información de archivo (OAIS). Modelo de referencia”. Es un estándar de alto nivel que proporciona un marco para la definición de una estrategia de conservación universal y definición de modelos propios enfocados en un área en particular propia para cada organización y para todo tipo de soportes digitales. En este se presenta un plan de preservación, contenido asociado a paquetes de información, bases de conocimiento, sistemas de almacenamiento y disponibilidad. Detalla las responsabilidades de cumplimiento

obligatorio. Describe tres modelos: el modelo funcional, el modelo de información y las transformaciones del empaquetado de información.

La entidad funcional Preservation Planning Functional Entity supervisa el entorno del OAIS, esta proporciona recomendaciones y planes de preservación para asegurar que la información almacenada permanezca accesible y comprensible a largo plazo. Las funciones de esta entidad incluyen el seguimiento de la comunidad designada y de la tecnología, el desarrollo de estrategias y normas de preservación, de diseños de empaquetado y planes de migración (Cruz & Diez, 2016).

- **ISO 16363**

Sistemas de transferencia de información y datos espaciales. Auditoría y certificación de repositorios digitales de confianza”. Este estándar es una herramienta para auditar, evaluar, potenciar y certificar repositorios digitales. Proporciona un marco de calidad para el análisis de la consistencia de un repositorio con respecto a la integridad de datos. Además, está íntimamente relacionado con la preservación a largo plazo y su accesibilidad. Los gestores de un repositorio pueden usarlo también como una herramienta de diagnóstico del estado del sistema, así como para planificar acciones que deben ser aplicadas en la gestión. Tres dimensiones son las que están presentes en este estándar: 1) infraestructura organizacional, 2) gestión de objetos digitales y 3) gestión de riesgos de infraestructura y de seguridad. Cada una de estas dimensiones posee diferentes métricas para validar los requisitos que debe cumplir un repositorio digital para su preservación a largo plazo, además de políticas y procedimientos sobre acceso, eliminación y autenticidad de los objetos digitales.

- **ISO 15489**

Información y documentación. Gestión de documentos. Parte 1: Conceptos y principios. Se aplica a los documentos en cualquier formato o soporte, proporciona una metodología de implementación. El documento porta principalmente a este estudio porque describe los principios y conceptos relativos a la gestión de documentos, los sistemas de gestión, el análisis recurrente del contexto de la organización y la identificación de los requisitos. Además, se enfoca en el cumplimiento de un marco legal y reglamentario. Desde el punto de vista estratégico describe las políticas y responsabilidades de los involucrados en el proceso de preservación digital. La norma trata de los beneficios de la gestión de documentos, marco reglamentario, política y responsabilidades, requisitos, diseño e implementación de un sistema de gestión, procesos, controles, supervisión, auditoría y formación.

- **UNE-ISO 14641-1**

Archivo electrónico. Parte 1: Especificaciones para el diseño y funcionamiento de un sistema de información para la preservación de información digital. En esta

norma se encuentra un conjunto de especificaciones técnicas, pero además políticas organizativas para la implementación, el almacenamiento y el acceso a los documentos electrónicos. Está destinado a usuarios como organizaciones que implementan sistemas de información, servicios de tecnología y de archivo de documentos de terceros. Propone en general tres temas: 1) optimizar el sistema para asegurar la preservación, 2) facilitar la búsqueda de información y 3) asegurar la accesibilidad y el uso de documentos electrónicos. En relación con aspectos de planificación, se pueden observar procedimientos, técnicas y un sistema de gestión del cambio y migración.

- **UNE ISO 30300, 30301, 30302**

Información y documentación. Sistemas de gestión para los documentos. Vocabulario (30300), Requisitos (30301), Guía de implantación (30302). Conjunto de elementos que interactúan para llevar a cabo la política de gestión documental; en él se encuentran involucrados personas, roles y responsabilidades; los procesos y controles; y la infraestructura. Además, es un aporte a la consecución de una estrategia, en la cual se encuentran incluidos los fines, la misión y las metas de la organización mediante políticas y planteamiento de objetivos. Con esto se busca una adecuada planificación con respecto a acciones, procesos y mejora continua.

- **UNE-ISO/TR 18492**

Conservación de documentos electrónicos a largo plazo. El informe técnico declara la publicación de un marco amplio para el desarrollo homogéneo de políticas y estrategias de conservación. El objetivo de este estándar radica en preservar la información digital auténtica y asegurar exactitud, fiabilidad e integridad a lo largo del tiempo. Diseña una estrategia organizacional de transferencia de información a un sistema de almacenamiento, soporte estable contra la obsolescencia tecnológica, entre otras características enfocadas a la seguridad de la información. Incluye, además, el diseño de una estrategia en la organización, tomando en cuenta los cambios tecnológicos, compatibilidad en software y hardware para asegurar la preservación a largo plazo y los objetos digitales.

- ISO/TR 18128

Información y documentación. Apreciación del riesgo en procesos y sistemas de gestión documental. Procura ayudar a las organizaciones a evaluar los riesgos para los procesos y sistemas de registros durante el tiempo. Puede ser utilizado por cualquier organización independiente del tamaño. Permite la identificación de riesgos, los cuales han sido propuestos por la Association of Records Managers and Administrators (ARMA), que es la comunidad de profesionales de gestión de registros, gestión de información y gobierno de la información. Además, permite la evaluación de los riesgos administrativos, de control de documentos, legales o normativos y tecnológicos. Incluye el diseño de una estrategia en la organización, tomando en cuenta los cambios tecnológicos, compatibilidad en software y hardware.

- ISO 16919

Sistemas espaciales de transferencia de datos e información. Requisitos para los organismos que realizan auditorías y certificaciones de repositorios digitales

confiables candidatos. Está destinado principalmente a aquellos que crean y administran la organización, que realizan la auditoría y certificación de repositorios digitales. Está diseñado para seguir el proceso de mejora continua, que es indispensable en una correcta planificación estratégica y tiene un mecanismo de verificación probado en cada etapa siguiendo una jerarquía de estándares. Los requisitos que propone la norma son generales, estructurales, de información, del proceso y del sistema de gestión.

3.3.2 RQ2. ¿Cuáles son los principales modelos de auditoría para preservación digital?

Esta pregunta de investigación responde a todos los modelos de auditoría que no se hayan presentado como estándares en RQ1. En la *Figura 2*, se puede observar un diagrama de red asociativo de los trabajos de investigación encontrados en la SLR y los modelos de auditoría de preservación digital. Se presentan ocho modelos de auditoría que están presentes en diez trabajos de investigación, estos últimos corresponden al 41,7 % de los estudios primarios.

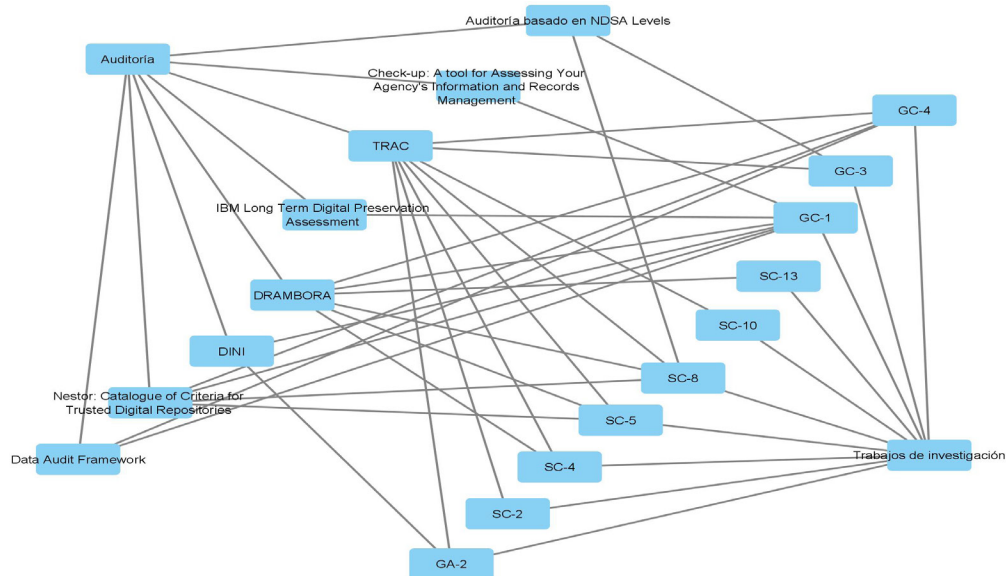


Figura 2. Modelos de auditoría para preservación digital.
Fuente: elaboración propia.

- **Data Audit Framework (DAF)**

Este marco de auditoría fue diseñado para ser aplicado sin personal dedicado o especializado. La metodología DAF sugiere que los datos se clasifiquen para restringir el alcance al pasar del intervalo amplio y poco profundo a la evaluación más detallada de la gestión de datos. Esta metodología consta de cuatro etapas principales que son ejecutadas secuencialmente: 1) planificación de la auditoría, que define el propósito y alcance de la encuesta y la realización de una investigación preliminar; 2) identificación y clasificación de los activos de datos, que consiste en identificar qué activos de datos existen y clasificarlos para determinar en qué enfocar esfuerzos para un análisis más profundo; 3) evaluación de la gestión de los activos de datos, en este punto el ciclo de vida de la información se considera que comprende los flujos de trabajo y se identifican debilidades en prácticas de creación y conservación de datos; e 4) informe de hallazgos y recomendación de cambios, reúne la información recopilada y proporciona recomendaciones para mejorar los datos de administración.

- **Nestor: Catalogue of Criteria for Trusted Digital Repositories**

Está dirigido principalmente a instituciones u organizaciones de memoria documental, como instituciones que almacenan archivos, bibliotecas, museos, entre otros, y sirve como un manual para diseñar, planificar e implementar un repositorio digital confiable a largo plazo. Las etapas generales de este marco de auditoría son 1) concepción, 2) planificación y especificación, 3) realización e implementación y 4) evaluación. Está orientado a la mejora continua y a la gestión de calidad, además, monitorea este proceso de desarrollo. Las medidas están enfocadas a evaluar el marco organizacional, la gestión de objetos digitales, la infraestructura y seguridad del repositorio digital.

- **DINI: Certificate for Document and Publication Repositories**

Este catálogo de criterios se traduce como los requisitos mínimos concretos de los repositorios digitales y servicios de publicación de acceso abierto, basados en la calidad, interoperabilidad, evaluación y certificación de repositorios. Los objetivos a largo plazo son conservar,

almacenar metadatos y documentos en un solo contenedor, mínimo cinco años. Las copias en archivo deben estar libres de medidas de derechos digitales que impidan el uso de estrategias de conservación a largo plazo. Además, se recomienda asegurar la disponibilidad, el uso de formatos de archivo abiertos y la producción de metadatos técnicos, para la preservación a largo plazo, la identificación única del formato de archivo y los metadatos con referencia a registros de formatos de archivos públicos.

- **DRAMBORA: Digital Repository Audit Method Based on Risk Assessment**

Presenta una metodología para la autoevaluación y alienta a las organizaciones a establecer una autoconciencia integral de sus objetivos. Adopta un enfoque ascendente y describe un proceso formalizado que alienta a los repositorios a considerar y documentar la misión, objetivos, limitaciones y actividades. El flujo de trabajo es el siguiente:

1. Definir propósito y alcance de auditoría
2. Formalizar personal y funciones
3. Determinar las clases de funciones
4. Formalizar mandatos
5. Formalizar restricciones
6. Formalizar objetivos
7. Formalizar actividades y activos
8. Identificar riesgos
9. Riesgos activos
10. Gestionar los riesgos

- **IBM Long Term Digital Preservation Assesment (LDTP)**

Permite evaluar a una organización que está en proceso de tener un plan de preservación digital a largo plazo. Para esta evaluación se usa Trustworthy Repositories Audit & Certification: Criteria and Checklist (TRAC) y proporciona resultados gráficos del nivel de riesgo que se encuentra en la organización en el ámbito de preservación.

- **Trusted Repository Archiving Checklist (TRAC)**

Es una metodología que proporciona herramientas para la auditoría, la evaluación y la certificación potencial de repositorios digitales. Es una lista de verificación que intenta reunir las mejores prácticas sobre los requerimientos necesarios para que un repositorio digital sea fiable en cuanto a la conservación de documentos. La lista de verificación TRAC fue reemplazada en 2012 por la ISO 16363, conocida como *Trusted Digital*. Esta evalúa 1) infraestructura organizativa: gobernanza y viabilidad organizacional, estructura organizativa y dotación de personal, sostenibilidad financiera, contratos, licencias y responsabilidades; 2) gestión de objetos digitales: adquisición de contenido, creación del paquete de archivo, planificación de la conservación, gestión de la información, gestión de acceso; y 3) tecnologías, infraestructura técnica y seguridad: infraestructura del sistema, tecnologías apropiadas y seguridad.

- **Check-up: A tool for Assessing Your Agency's Information and Records Management**

Esta guía desglosa procedimientos para el análisis de riesgos vinculados a la gestión de información y archivos. La evaluación del alcance y la calificación del nivel se realiza en términos cuantitativos. Menciona varios aspectos relevantes como el marco, la estrategia y la arquitectura de gestión de la información. Check-up 2.0 es una aplicación web desarrollada para evaluar la capacidad de gestión documental digital de las agencias del Gobierno australiano. En la evaluación de las prácticas de gestión de la información se encuentran los criterios: 1) gobernanza, 2) creación o generación de información, 3) interoperabilidad, 4) almacenamiento y preservación digital de información, 5) eliminación, destrucción y transferencia y 6) operaciones digitales.

- **Auditoría basado en National Digital Stewardship Alliance (NDSA) Levels**

NDSA es un consorcio de instituciones comprometidas en la preservación digital. Presenta un modelo de madurez de preservación digital que se basa en identificar deficiencias en el sistema de almacenamiento, alteración de los ficheros, medidas de seguridad de la información, metadatos y formatos de ficheros en cuatro niveles:

Nivel 1. Proteja sus datos

Nivel 2. Conozca sus datos

Nivel 3. Controle sus datos

Nivel 4. Repare sus datos

3.3.3 RQ3 ¿Qué trabajos de investigación existen relacionados a modelos de madurez de preservación digital?

Es conveniente hablar antes de los modelos de madurez de sistemas de información, que ayudan a las organizaciones a mejorar sus procesos de implementación y desarrollo de software. Entre estos los más difundidos son Capability Maturity Model Integration (CMMI), conjunto de mejores prácticas para mejorar el rendimiento del proceso de desarrollo. Otro de estos modelos de madurez es el que implementa el estándar ISO/IEC 15504 (*Software Process Improvement Capability Determination*). Este estándar permite la evaluación de los procesos de software, mide su capacidad a través de una escala y establece que en una organización la información debe recopilarse, registrarse, almacenarse, ser comprobada, procesada, analizada, recuperada y presentarse por medio de herramientas de software, de forma similar a lo que se espera en un modelo de madurez de preservación digital. La [Figura 3](#) representa un diagrama de red asociativa de los modelos o acercamientos relacionados a la madurez de la preservación digital. Puede observarse que este tema de investigación aún no se encuentra estudiado en gran medida; por esta razón, las implementaciones se tornan complejas al no existir un marco de trabajo que apunte a identificar de forma clara cómo realizar la implementación de un sistema de gestión de preservación digital a largo plazo. Se exponen tres modelos de madurez de preservación digital, presentes en cinco trabajos de investigación. Estos últimos corresponden al 20,83 % de los estudios primarios.

- **Digital Preservation Capability Maturity Model (DPCMM)**

Es un marco utilizado para evaluar la capacidad de una organización y realizar una gama de funciones de gobernanza, operativas y de gestión de datos, asociadas con preservación de archivos digitales a largo plazo (más de diez años) y permanentes. Además, es utilizada para fortalecer el software de gestión de documentos electrónicos, a través de las etapas de los componentes.

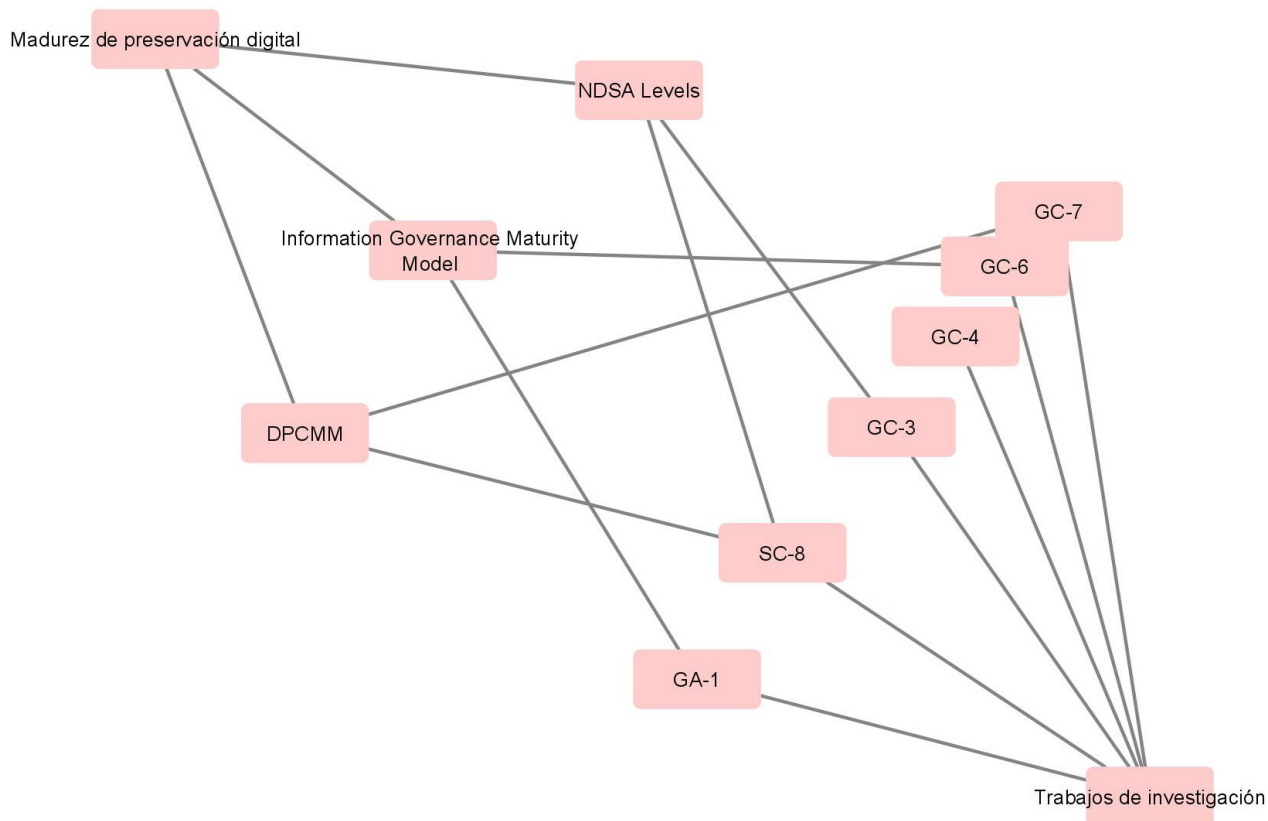


Figura 3. Modelos de madurez de la preservación digital.
Fuente: elaboración propia.

Se basa en las funciones de OAIS (ISO 14721) y los criterios de auditoría de ISO 16363. Busca identificar a un alto nivel las capacidades óptimas de preservación digital, informar brechas en los niveles de capacidad, las métricas de desempeño, educar e involucrar a los asignadores de recursos, establecer prioridades para lograr capacidades basadas en estándares y garantizar el acceso a largo plazo. Presenta cinco etapas o niveles: 1) capacidad nominal de conservación digital; 2) capacidad mínima de conservación digital; 3) capacidad de conservación digital intermedia; 4) capacidad de conservación digital avanzada, y 5) capacidad óptima de conservación digital. Este modelo posee tres características generales de cumplimiento: infraestructura de preservación digital, preservación del repositorio y servicios de preservación digital.

- **Information Governance Maturity Model**

Este marco de trabajo establece autoridades, apoyos, procesos, capacidades, estructuras e infraestructura

de la preservación digital, con la finalidad de permitir que la información sea un activo útil y una responsabilidad reducida. Indica cinco niveles de madurez: 1) deficiente, 2) en desarrollo, 3) esencial, 4) proactivo y 5) transformacional. Se basa en el mantenimiento de registros generalmente desarrollados por ARMA, estos proporcionan pautas de alto nivel de buenas prácticas de preservación digital, incluye los principios de responsabilidad, transparencia, integridad, protección, cumplimiento, disponibilidad, retención y disposición.

- **National Digital Stewardship Alliance (NDSA) Levels**

Anteriormente, en la pregunta RQ2 se había resumido la auditoría basada en NDSA Levels; sin embargo, esta representa una evaluación del modelo que se basa en un conjunto de recomendaciones escalonadas sobre las organizaciones de cómo deberían comenzar a construir o mejorar sus actividades de preservación digital. Se presentan tareas organizadas en cuatro niveles, de los

cuales el nivel 1 es el más bajo y el nivel 4 es el más alto o el cual debe estar como una meta de la organización para la preservación digital. Es considerado un modelo de auditoría, razón por la cual se ubicó en este documento también en la pregunta RQ2.

3.3.4 RQ4. *¿Existen aportes relacionados con modelos de madurez de preservación digital que incluyan los principios de la planificación estratégica?*

No se han encontrado resultados contundentes relacionados con modelos de madurez de preservación que apliquen principios de la planificación estratégica tradicional. Sin embargo, tanto los estándares analizados como las metodologías se orientan a una planificación controlada e incluso gestionan los riesgos de aplicación de una metodología u otra. Además, se encuentran trabajos en esta SLR que dan a conocer ciertos indicios en la planificación estratégica, entre estos se puede mencionar el GC-5, que da pautas organizacionales para la conservación de los documentos generados a través de *Preservation Metadata: Implementation Strategies* (PREMIS). En este existe una evaluación de riesgos por escenarios: normativo, administrativo, archivístico, tecnológico, organizacional; adicionalmente, un plan de conservación, actuación y desempeño de la estrategia. El estudio SC-3 menciona niveles de preservación digital en los cuales se incluye: 1) mantenimiento continuo, 2) aplicación de la conservación, 3) aplicación de la preservación de datos y 4) preservación de Bitstream. El SC-7 habla sobre las obsolescencias de hardware y software, falta de fondos, derecho legal y política de estrategia. El SC-10 indica una estructura de un plan de cinco puntos para construir y mantener la preservación digital en el estado de Montana (USA): 1) base de conocimiento, 2) evaluación del panorama de preservación, 3) defensa del valor de las actividades de preservación e 4) implementación de servicios compartidos, para desarrollar experiencia administrativa de manera colectiva. Por último, el SC-13 aborda estrategias de preservación digital, personal, herramientas, equipos y otros medios de una forma general con un modelo tripartito basado en la administración, tecnología y contenido.

4. Discusión y conclusiones

Las cadenas de búsqueda aplicadas en las bases de datos científicas devolvieron importantes trabajos relaciona-

dos con los estándares, modelos de auditoría y madurez de preservación digital. Se presentó un total de 309 trabajos, de los cuales el 64 % correspondió a la base de datos Scopus y el 36 % restante a Google Académico. Una vez revisados los títulos, resúmenes y palabras clave, 42 trabajos de investigación corresponden a los estudios seleccionados. Posteriormente, se realizó la revisión a texto completo y 24 trabajos aportan significativamente a responder las preguntas de investigación planteadas, estos artículos corresponden a los estudios primarios.

Respecto a los estándares de preservación digital (RQ1), se presentan 21 trabajos en los cuales se han podido recopilar diez estándares de buenas prácticas de preservación digital. Estos estándares describen las mejores prácticas en aspectos de modelo de referencia para la estrategia de preservación digital, auditoría y certificación de repositorios digitales, principios relacionados con la gestión de sistemas documentales —desde el punto de vista estratégico—, especificaciones de funcionamiento de un sistema de información, políticas de gestión documental con roles y responsabilidades de involucrados, la conservación de documentos electrónicos a largo plazo y la apreciación del riesgo en procesos de este ámbito, los cuales están íntimamente relacionados con una visión estratégica de implementación de modelos de preservación digital confiables y que trasciendan en el tiempo.

En relación con los modelos de auditoría de preservación digital (RQ2), se presentan diez trabajos que hacen referencia a ocho modelos y metodologías orientadas a este ámbito, varios de ellos presentan etapas o niveles que incluyen aspectos relacionados con la planificación estratégica y poseen un *checklist* orientado a la autoevaluación, con la finalidad de que una organización aplique las mejores prácticas en cuanto a la preservación de documentos digitales se refiere.

De los modelos de madurez de preservación digital (RQ3), existen seis trabajos relacionados que presentan tres modelos en este ámbito. Estos son muy similares a los presentados por el CMMI y el estándar ISO/IEC 15504. De hecho, el DPCMM se basa en estos para su creación. Estos modelos se basan en fases o niveles de madurez, los cuales buscan que una organización tras-

cienda de una fase inicial a optimizar los procesos y resguardar la información digital en el tiempo.

Cabe señalar que no se han encontrado aportes significativos que implementen los principios de la planificación estratégica (RQ4) tradicional o que permitan el desarrollo e implementación de un proyecto de preservación digital con herramientas de medición y seguimiento, por medio de, por ejemplo, indicadores clave de desempeño (KPIs). Indicadores que permitan tomar decisiones de manera oportuna y establecer procesos de mejora continua de forma clara, con un acercamiento a modelos de planeación estratégica tradicionales que se han implementado por años en miles de organizaciones a nivel mundial para mejorar la calidad de sus negocios. Sin embargo, existen buenos acercamientos a una planificación estratégica para la implementación de la preservación digital. En los trabajos de esta SLR se han podido reconocer aspectos clave en este ámbito, como evaluación de riesgos por escenarios: normativos, administrativos, archivísticos, tecnológicos, organizacionales, planes de conservación y desempeño de la estrategia. Además, niveles de preservación digital, política de estrategia, estructura para construir y mantener la preservación digital, estrategias de preservación digital, personal, herramientas, y un modelo tripartito basado en la administración, tecnología y contenido.

En trabajos futuros se buscará caracterizar las mejores prácticas de cada uno de estos estándares, modelos de auditoría y madurez. Esto con el fin de crear un modelo de preservación digital apto para la implementación en las organizaciones, que disminuya el riesgo a través de la implementación de los principios tradicionales de la planificación estratégica y que han sido probados durante años en las industrias.

5. Referencias

1. Ashley, Lori; Misić, Milovan (2019). Digital preservation capability maturity model (DPCMM). Genesis and practical uses. In *Diverse Applications and Transferability of Maturity Models*, 152-167. <https://doi.org/10.4018/978-1-5225-7080-6.ch006>
2. Borbinha, José; Proença, Diogo; Tribolet, José (2016). Modeling the value of digital preservation

activities. *Procedia Computer Science*, 100, 1050-1055. <https://doi.org/10.1016/j.procs.2016.09.280>

3. Boté, Juan (2012). *Propuesta de un modelo de preservación digital para pequeñas y medianas instituciones sanitarias* (tesis doctoral). Universitat de Barcelona, Barcelona, España. <http://hdl.handle.net/2445/35867>
4. Castillo, María; Umaña, Raquel (2020). Modelo para la preservación de documentos digitales. *Revista Del Archivo Nacional*, 83(1-12), 129-182. <http://www.dgan.go.cr/ran/index.php/RAN/article/view/453>
5. Chacón, Ana (2018). *Continuidad digital en los archivos del Sector Transporte e Infraestructura de Costa Rica: diseño y planificación de un proyecto para la mejora organizacional de Consejo Nacional de Concesiones* (tesis doctoral). Universidad Carlos III de Madrid, Madrid, España. <http://hdl.handle.net/10016/26817>
6. Corrado, Edward; Moulaison, Heather (2017). Bringing content into the picture: Proposing a tri-partite model for digital preservation, *Journal of Library Administration*, 58(1), 1-17. <https://doi.org/10.1080/01930826.2017.1385988>
7. Cote, Conor; Mannheimer, Sara (2017). Cultivate, assess, advocate, implement, and sustain. *Digital Library Perspectives*, 33(2), 100-116. <http://doi.org/10.1108/DLP-07-2016-0023>
8. Cruz, José (2015). Estrategias de preservación digital permanente en los archivos nacionales. *Boletín de la ANABAD*, 65(3), 127-148. <https://dialnet.unirioja.es/servlet/articulo?codigo=5320603>
9. Cruz, José; Díez, Carmen. (2016). Open Archival Information System (OAIS): Lights and shadows of a reference model. *Investigación Bibliotecológica*, 30(70), 221-247. <https://doi.org/10.1016/j.ibbai.2016.10.010>
10. De Giusti, Marisa (2014). *Una metodología de evaluación de repositorios digitales para asegurar la preservación en el tiempo y el acceso a los contenidos* (tesis doctoral). Universidad Nacional de La Plata, La Plata, Argentina. <http://hdl.handle.net/10915/43157>
11. De Giusti, Marisa; Villareal, Gonzalo (2018). Revisión de distintas implementaciones para preservación digital: hacia una propuesta metodológica para la preservación y la auditoría de confiabilidad de RI. *Revista Digital de Biblioteconomía e Ciência da Informação*, 16(2), 273-292. <https://doi.org/10.20396/rdbci.v16i2.8651589>
12. Della, Dennis; Colman, Wolfgang; Welker, Ben; Rennick, Brian (2020). Library eArchiving

[Preservación digital a largo plazo: estándares, auditoría, madurez y planificación estratégica]

- with ZONTAL space and the allotrope data format. *Digital Library Perspectives*, 36(1), 69-77. <https://doi.org/10.1108/DLP-09-2019-0036>
13. Ejitaga, Stella; Sambo, Atanda; Urhefe, Enite (2017). A survey of digital preservation challenges in Nigerian libraries: Librarians' perspectives. *International Journal of Digital Curation*, 12(1), 117-128. <http://dx.doi.org/10.2218/ijdc.v12i1.426>
 14. Englbrecht, Ludwig; Meier, Stefan; Pernul, Günther (2019). Towards a capability maturity model for digital forensic readiness. *Wireless Networks*. <https://doi.org/10.1007/s11276-018-01920-5>
 15. Ferreras, María (2016). *Visibilidad e impacto de la literatura gris científica en repositorios institucionales de acceso abierto* (tesis doctoral). Universidad de Salamanca, Salamanca, España. <https://dialnet.unirioja.es/metricas/documento/TES/155457>
 16. Fonseca-Carrera, Efrain (2014). *Conceptualización e infraestructura para la investigación experimental en ingeniería del software* (tesis doctoral). Universidad Politécnica de Madrid, Madrid, España. <http://oa.upm.es/32845/>
 17. Gallinger, Michelle; Bailey, Jefferson; Cariani, Karen; Owens, Trevor; Altman, Micah. (2017). Trends in digital preservation capacity and practice: Results from the 2nd Bi-Annual National Digital Stewardship Alliance Storage survey. *D-Lib Magazine*, 23(7/8). <http://doi.org/10.1045/july2017-gallinger>
 18. Kearney, Michael; Giarretta, David; Garrett, John; Hughes, Steve (2018). Digital preservation archives—A new future architecture for long-term interoperability. *SpaceOps Conference*. <http://doi.org/10.2514/6.2018-2402>
 19. Kitchenham, Barbara; Charters, Stuart. (2007). *Guidelines for performing systematic literature reviews in software engineering*. Keele: Software Engineering Group, School of Computer Science and Mathematics, Keele University.
 20. Leija, David (2017). *Preservación digital distribuida y la colaboración interinstitucional: modelo de preservación digital para documentos con fines de investigación en universidades de México* (tesis doctoral). Universitat de Barcelona, Barcelona, España. <http://hdl.handle.net/10803/454886>
 21. Leija, David; Térmens, Miquel (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>
 22. Lemiux, Victoria (2016). Trusting records: Is blockchain technology the answer? *Records Management Journal*, 22(6), 110-139. <http://doi.org/10.1108/RMJ-12-2015-0042>
 23. Marshall, Christopher; Brereton, Pearl (2013). Tools to support systematic literature reviews in software engineering: A mapping study. *International Symposium on Empirical Software Engineering and Measurement*, 296-299. <https://doi.org/10.1109/ESEM.2013.32>
 24. Molina, Fernando (2017). *Propuesta de un modelo conceptual de preservación digital aplicable a la mejora de la preservación de evidencia digital en instituciones de investigación criminal* (tesis doctoral). Universidad Nacional Mayor de San Marcos, Lima, Perú. <https://hdl.handle.net/20.500.12672/6850>
 25. Morales, Cristian; Radicelli, Ciro; Jaramillo, María; Boderó, Elba. (2018). Adopción de software de Business Intelligence: una revisión sistemática de literatura aplicando minería de texto. *Revista ESPACIOS*, 39(24). <http://www.revistaespacios.com/a18v39n24/18392429.html>
 26. Moro, Manuela; Llanes Dunia. (2018). La importancia de la normalización para el ejercicio profesional del archivista. *Investigación bibliotecológica*, 32(74), 193-223. <https://doi.org/10.22201/iibi.24488321xe.2018.74.57919>
 27. Neto, Antonio; Borges, María; Roque, Licinio. (2017). Preliminary study about the applicability of a service-oriented architecture in the oasis model implementation. *Association for Computing Machinery*, 1-5. <https://doi.org/10.1145/3144826.3145381>
 28. Oltmanns, Elias; Hasler, Tim; Peters-Kottig, Wolfgang; Kuper, Heinz-Günter (2019). Different preservation levels: The case of scholarly digital editions. *Data Science Journal*, 18(1), 51. <http://doi.org/10.5334/dsj-2019-051>
 29. Sandí, Juan; Cruz, Mainor (2016). Repositorios institucionales digitales: Análisis comparativo entre SEDICI (Argentina) y Kérwá (Costa Rica). *E-Ciencias de la Información*, 7(1), 5-34. <https://doi.org/10.15517/eci.v7i1.25264>