



**VNiVERSiDAD
D SALAMANCA**

CAMPUS DE EXCELENCIA INTERNACIONAL

**MÁSTER EN SISTEMAS DE
INFORMACIÓN DIGITAL**

Softwares para repositorios digitales: Evaluación para la gestión y preservación de los materiales bibliográficos y de los documentos archivísticos de la Biblioteca Nacional de Brasil

Trabajo Fin de Máster

Autor: Wellington da Silva

Tutor: José Antonio Merlo Vega

Salamanca
2019



VNiVERSIDAD
D SALAMANCA

CAMPUS DE EXCELENCIA INTERNACIONAL

MÁSTER EN SISTEMAS DE
INFORMACIÓN DIGITAL

UNIVERSIDAD DE SALAMANCA
FACULTAD DE TRADUCCIÓN Y DOCUMENTACIÓN
MÁSTER EN SISTEMAS DE INFORMACIÓN DIGITAL

Trabajo Fin de Máster

**Software para repositorios digitales:
Evaluación para la gestión y
preservación de los materiales
bibliográficos y de los documentos
archivísticos de la
Biblioteca Nacional de Brasil**

Autor: Wellington da Silva

Tutor: José Antonio Merlo Vega

Salamanca
2019

Asiento catalográfico con formato del repositorio institucional GREDOS.

Título: Software para repositorios digitales: Evaluación para la gestión y preservación de los materiales bibliográficos y de los documentos archivísticos de la Biblioteca Nacional de Brasil

Autor: Da Silva, Wellington

Tutor: Merlo Vega, José Antonio

Palabras clave: Biblioteca Nacional de Brasil, Preservación digital, Repositorios digitales, Evaluación, Materiales bibliográficos y Documentos archivísticos

Keywords: National Library of Brasil, Digital Preservation, Digital Repositories, Evaluation, Bibliographic Materials and Archival Documents

Fecha: 2019

Resumen: Este trabajo se propone a evaluar y analizar las características y funcionalidades de los softwares Archivematica, DSpace y RODA, con el fin de identificar el más adecuado para la gestión y preservación de los materiales bibliográficos digitales que forman la Biblioteca Nacional Digital (BNDigital) y de los documentos archivísticos digitales producidos por el Sistema Eletrônico de Informações (SEI) de la Biblioteca Nacional de Brasil. Considerando que la llegada de la Tecnología de la Información y Comunicación (TIC) afectó a la rutina diaria de la Biblioteca Nacional de Brasil (BNB) y los decretos 8.297/2014 establece en el artículo 2º que la BNB debe adoptar las medidas necesarias para la preservación y protección del patrimonio bibliográfico y digital bajo su custodia, y 8.539/2015 que en el artículo 18 establece que los órganos o entidades deberán establecer políticas, estrategias y acciones que garanticen la preservación, el acceso y el uso continuo a largo plazo de los documentos digitales. Para cumplir las obligaciones establecidas en dichos decretos, la Biblioteca Nacional de Brasil tendrá que adoptar nuevas estrategias de preservación digital, y una de ellas deberá ser la implantación de un repositorio digital, debido a que la longevidad de los objetos digitales depende directamente del medio de almacenamiento utilizado y los repositorios digitales tienen como objetivo garantizar la preservación, el acceso y la autenticidad a largo plazo de los objetos digitales. Sin embargo, este repositorio tiene que ser adecuado a bibliotecas digitales y documentos de archivo. En la actualidad hay una gran variedad de programas para repositorios digitales, con diferentes características, requisitos y funcionalidades y, debido a esta diversidad, es importante realizar una evaluación; para identificar el más adecuado para la BNDigital y para el SEI, lo que constituye el objeto de esta investigación.

Abstract: This study aims to evaluate and analyze the characteristics and functionalities of the Archivematica, DSpace and RODA software, in order to identify the most appropriate for the management and digital preservation of bibliographic materials that form the National Digital Library (BNDigital) and the digital archival documents produced by Sistema Eletrônico de Informações (SEI) of the National Library of Brazil (NLB). This is because the advent of Information Technology and Comunication (ITC) affected the daily routine of the National Library of Brazil, and the legislative decrees 8.297/2014 that in the

artículo 2º establishes that the necessary measures must be adopted for the preservation and protection of the bibliographic and digital heritage under its custody, and 8.539/2015 that in artículo 18 establishes that the organs or entities must establish policies, strategies and actions that guarantee the preservation, access and continuous long-term use of digital documents. To comply with the obligations established in legislative decrees 8.297/2014 and 8.539/2015 the National Library of Brazil will have to adopt new digital preservation strategies, and one of them should be the implementation of a digital repository, because the longevity of digital objects depends directly on the storage used, and digital repositories aim to guarantee the preservation, authenticity and long-term access of digital objects. However, this repository has to be suitable for digital libraries and archival documents. Nowadays there is a great variety of programs for digital repositories, with different characteristics, requirements and functionalities and, due to this diversity, it is important to carry out an evaluation; to identify the most suitable for the BNDigital and for the SEI, which is the aims of this investigation.

Resumo: Este trabalho tem como objetivo avaliar e analisar as características e funcionalidade dos softwares Archivematica, DSpace e RODA, a fim de identificar o mais adequado para a gestão e preservação dos materiais bibliográficos digitais que formam a Biblioteca Digital Nacional (BNDigital) e os documentos arquivísticos digitais produzidos pelo Sistema Eletrônico de Informações (SEI) da Biblioteca Nacional do Brasil (BNB). Considerando que o advento da Tecnologia da Informação e Comunicação (TIC) afetou as rotinas diárias da Biblioteca Nacional do Brasil, e os decretos 8.297/2014 que no artículo 2º estabelece a BNB deve adotar medidas necessárias para a preservação e proteção do patrimônio bibliográfico e digital sob sua custódia, e 8.539/2015, que no artículo 18 estabelece que os órgãos ou entidades devem estabelecer políticas, estratégias e ações que garantam a preservação, o acesso e o uso contínuo de longo prazo dos documentos digitais. Para cumprir as obrigações estabelecidas nos decretos supracitados, a Biblioteca Nacional do Brasil terá de adotar novas estratégias para a preservação digital, e uma delas deve ser a implementação de um repositório digital, porque a longevidade dos objetos digitais depende diretamente do meio de armazenamento utilizado e os repositórios digitais visam garantir a preservação, o acesso e a autenticidade a longo prazo dos objetos digitais. No entanto, este repositório deve ser adequado para bibliotecas digitais e documentos de arquivo. Atualmente, existem uma grande variedade de programas para repositórios digitais, com diferentes características, requisitos e funcionalidades, devido a esta diversidade, é importante avaliar; para identificar o mais adequado para a BNDigital e para o SEI, que são o escopo desta investigação.

A la Biblioteca Nacional de Brasil que
me dio la oportunidad de avanzar en
mi formación académica y profesional

Agradecimientos

A mi madre que siempre me estimuló a estudiar. Mi abuela y mis tías, que ayudaron a mi madre en mi educación. Mi hermana, sobrina y primas.

A la Biblioteca Nacional de Brasil que me permitió realizar este proyecto que aportó y aportará mucho a mi vida profesional y personal.

A mi jefa y amiga Flávia Cezar, que desde el principio apoyó y ayudó a que este deseo se convirtiera en realidad.

A mi familia en Salamanca, las brasileñas Marina, Ana, Tamires, Mariana y Fernanda que me ayudaron a sobrellevar la nostalgia de mi casa.

A mi amigo André que me llamaba todos los días y me hacía reír. A mis amigas de la Biblioteca Nacional de Brasil y ahora de la vida: Gabriela Ayres, Lorrane, Gabriela Jordão y Maria Fernanda, por la ayuda con el TFM, Amanda y Andrea por los mensajes de Whatsapp que hacían mis días más alegres. Larissa, por las conversaciones estimulantes.

A las amigas que el Máster me dio, la chilena Vita con su abrazo latinoamericano y las españolas Laura y Lucía.

A los profesores que tuve en Brasil y aquí en el máster que, con sus enseñanzas y experiencias, han contribuido en mi desarrollo profesional y personal. En especial al profesor y coordinador Carlos que me ayudó desde Brasil con los trámites académicos.

A mi tutor, José Merlo que me ha guiado en esta investigación.

ÍNDICE

PARTE I – INTRODUCCIÓN	12
1. JUSTIFICACIÓN Y OBJETIVOS	12
1.1 Justificación	12
1.2 OBJETIVOS	15
1.2.1 Objetivo general	15
1.2.2 Objetivos específicos	15
1.3 ASPECTOS METODOLÓGICOS	15
PARTE II – CONTEXTO TEÓRICO	17
2. BIBLIOTECA NACIONAL DE BRASIL (BNB)	17
2.1 Presentación	17
2.2 Historia	17
2.3 Misión, función y competencias	18
2.4 Acervo digital	18
2.4.1 Biblioteca Nacional Digital (BNDigital)	19
2.4.1.1 Misión y objetivos	19
2.4.1.2 Política de digitalización	20
2.4.1.3 Colecciones digitales	20
2.4.2 Sistema Eletrônico de Informações (SEI)	21
2.4.2.1 Historia y bases legales	22
2.4.2.2 Organización administrativa/funcional	23
2.4.2.3 Fondo archivístico digital	23
3. PRESERVACIÓN DIGITAL	25
3.1 Concepto y definición	26
3.2 Estrategias de preservación digital	27
3.2.1 Estándares de formatos	28
3.2.1.1 Herramientas para identificación de formatos	32
3.2.2 Metadatos	33
3.2.2.1 Dublin Core	34
3.2.2.2 EAD (Encoded Archival Description)	36
3.2.2.3 MODS (Metadata Object Description Schema)	37
3.2.2.4 PREMIS (PREservation Metadata: Implementation Strategies)	37
3.2.2.5 METS (Metadata Encoding & Transmission Standard)	38
3.2.3 Migración	39
3.2.4 Conversión	39
3.2.5 Emulación	40

3.3	El Modelo OAIS (Open Archival Information System).....	41
4.	REPOSITORIOS DIGITALES.....	45
4.1	Repositorios Institucionales (RI).....	47
4.2	Repositorios de Documentos Administrativos (RDA).....	47
4.3	Interoperabilidad.....	49
4.3.1	Protocolos de interoperabilidad.....	50
4.4	Software para repositorios digitales.....	51
4.4.1	Archivematica.....	52
4.4.2	DSpace.....	53
4.4.3	RODA (Repositorio de Objetos Digitales Auténticos).....	54
	PARTE III – EVALUACIÓN Y RESULTADOS.....	56
5.	METODOLOGÍA.....	56
5.1	Evaluación de los softwares.....	59
5.1.1	Categoría Infraestructura.....	59
5.1.2	Categoría Gestión de contenido.....	60
5.1.3	Categoría Administración del uso.....	60
5.1.4	Categoría Interfaz de usuario.....	61
5.1.5	Categoría Preservación digital.....	61
5.1.6	Categoría: Interoperabilidad e Integración.....	62
6.	RESULTADOS Y ANÁLISIS.....	63
6.1	Resultado por software.....	63
6.1.1	Archivematica.....	63
6.1.2	DSpace.....	63
6.1.3	RODA.....	64
6.2	Resultado general.....	64
6.3	Análisis por categoría.....	65
6.3.1	Infraestructura.....	65
6.3.2	Gestión de contenido.....	67
6.3.3	Administración del uso.....	68
6.3.4	Interfaz de usuario.....	69
6.3.5	Preservación digital.....	70
6.3.6	Interoperabilidad e Integración.....	72
7.	CONCLUSIÓN.....	73
8.	BIBLIOGRAFIA.....	75

ÍNDICE DE TABLAS

Tabla 1. Tipos de estrategias de preservación digital.....	27
Tabla 2. Estándares de formatos.	31
Tabla 3. Funcionalidades relacionadas con la infraestructura	59
Tabla 4. Funcionalidades relacionadas con la Gestión de contenido	60
Tabla 5. Funcionalidades relacionadas con la Administración del uso	60
Tabla 6. Funcionalidades relacionadas con la Interfaz de usuario	61
Tabla 7. Funcionalidades relacionadas con la Preservación digital.....	61
Tabla 8. Funcionalidades relacionadas con la Interoperabilidad e Integración.....	62
Tabla 9. Resultado Archivematica en cada categoría	63
Tabla 10. Resultado DSpace en cada categoría	63
Tabla 11. Resultado RODA en cada categoría	64

ÍNDICE DE FIGURAS

Figura 1. Selección de formatos.	28
Figura 2. Especificaciones técnicas de los componentes.....	30
Figura 3. Elementos Descriptivos del EAD.....	36
Figura 4. PREMIS como subconjunto de todos los metadatos de preservación.....	38
Figura 5. Entidades Funcionales de OAIS	43
Figura 6. Crecimiento de Repositorios de Acceso Abierto en el mundo.	46
Figura 7. Los softwares más instalados en el mundo.....	52
Figura 8. Los softwares más instalados en Brasil.	52

ÍNDICE DE GRÁFICOS

Gráfico 1. Resultado general	65
Gráfico 2. Resultado de los softwares en la categoría Infraestructura.....	66
Gráfico 3. Resultado de los softwares en la categoría Gestión de contenido	68
Gráfico 4. Resultado de los softwares en la categoría Administración del uso.....	69
Gráfico 5. Resultado de los softwares en la categoría Interfaz de usuario	70
Gráfico 6. Resultado de los softwares en la categoría Preservación digital.....	71
Gráfico 7. Resultado de los softwares en la categoría Interoperabilidad e Integración..	72

PARTE I – INTRODUCCIÓN

1. JUSTIFICACIÓN Y OBJETIVOS

1.1 Justificación

Este trabajo es parte de las actividades obligatorias del Máster en Sistemas de Información Digital de la Universidad de Salamanca, curso 2018-2019 y se propone a evaluar y analizar las características y las funcionalidades de los softwares para repositorios digitales Archivemática, DSpace y RODA con el fin de identificar el más adecuado para la gestión y preservación de los materiales bibliográficos y de los documentos archivísticos.

La razón para la elección del tema está relacionada con la primera acción tomada por la Biblioteca Nacional de Brasil para la preservación de su acervo digital, que fue la constitución de la comisión de estudios sobre preservación digital¹, formada por sus funcionarios, con la finalidad de formular, especificar, dimensionar e implementar una política de preservación del acervo digital actual, teniendo en cuenta su crecimiento. Los estudios producidos por la comisión están orientados a la elaboración de proyectos específicos como:

1. Plan de Preservación Digital;
2. Manual de la política de formación de colecciones digitales;
3. Planificación e implantación del Repositorio Digital.

A partir de estos estudios se identificó, entre otras, la necesidad de evaluar y analizar algunos programas para repositorios digitales disponibles en el mercado, para la elaboración del proyecto de Planificación e implantación del Repositorio Digital. El autor de esta investigación es uno de los miembros de la comisión, y la evaluación y el análisis de los softwares quedaron bajo su responsabilidad. Así que, parte de esta investigación será útil para el desarrollo de las actividades que el autor está encargado de desarrollar para los proyectos de la comisión.

Actualmente, la mayoría de las instituciones, utilizan las funcionalidades ofrecidas por la tecnología de la información y comunicación (TIC) para producir, gestionar, almacenar y difundir la información. Desde el escenario globalizado de la Administración Pública, se ha adoptado procesos de modernización, que buscan mejorar la calidad de los servicios públicos y aproximar al ciudadano de la administración (Mateus, 2008). Lo que comprueba la existencia de un proceso de modernización administrativa basado en la desmaterialización de la información. Para Thurston (2012), los gobiernos adoptan sistemas digitales sin tener conciencia de los desafíos que plantea preservar la integridad de los documentos generados por tales sistemas. Estar consciente de eso significa reflexionar sobre las consecuencias del uso de esas nuevas tecnologías, pues estas consecuencias pueden dificultar o imposibilitar el acceso y la comprensión de los objetos

¹ La Comisión fue constituida bajo la Decisión Ejecutiva 025 de la presidencia de la Biblioteca Nacional de Brasil, en 13 de junio de 2017.

digitales en el futuro, debido a la obsolescencia tecnológica (hardware y software), la fragilidad de los soportes y también por problemas para asegurar la autenticidad, la veracidad y la integridad de la información registradas en los soportes y formatos digitales. Según Álvarez Wong (2017), la consecuencia de estos problemas puede ser una pérdida de información con daños irreparables para la gestión y memoria de una institución y de la sociedad.

En la visión del Parliamentary Archives (2008), las estrategias de preservación digital son acciones y actividades capaces de mantener los recursos digitales del Parlamento auténticos y accesibles en el futuro, a pesar de los cambios del entorno tecnológico.

La longevidad de los objetos digitales depende directamente del medio de almacenamiento usado Texier (2013). Por esto uno de los principales enfoques de la preservación digital es el almacenamiento de la información. De acuerdo con Heminger y Robertson (2000), esto se debe a que las nuevas tecnologías de producción y almacenamiento de objetos digitales están evolucionando más rápido que nunca y la rapidez del desarrollo exige la preservación del gran volumen de conocimiento que se ha creado y se sigue creando en el entorno digital. Para que los objetos digitales sean preservados y almacenados con seguridad y control se debe tener en cuenta la implantación de un Repositorio Digital (RD), también conocido como Repositorio Institucional (RI) y Repositorio de Documentos Administrativos (RDA). Polanco-Cortés (2014) define Repositorio Digital, en un sentido amplio, como un medio para gestionar, almacenar, preservar, difundir y facilitar el acceso a los objetos digitales que almacena. Partiendo de la definición del RD, Crow (2002) considera un RI como colecciones digitales que capturan y preservan la producción intelectual de las comunidades universitarias, pudiendo ampliar su campo de implantación. El campo de implantación de un RI se amplió para un RDA y, según Texier (2013), este permite a una institución gestionar los documentos administrativos digitales, así el RDA podrá ofrecer los servicios de recopilación, clasificación, acceso, difusión y preservación de la misma manera que los RI.

La Biblioteca Nacional de Brasil (BNB) también cambió su manera de producir, usar y almacenar información, debido a las TIC. Si antes su acervo estaba compuesto de pergaminos, papiros, libros y documentos administrativos en soporte papel, actualmente se compone además por objetos digitales (nativos y digitalizados) como libros, revistas, periódicos, música, fotografías, y documentos administrativos entre otros; y el volumen de estos objetos digitales están creciendo cada día. Para Lor (1997), una biblioteca nacional es una institución responsable de recoger, registrar, preservar y poner a disposición el patrimonio bibliográfico y documental provenientes o relacionados con su país, es decir, preservar y facilitar el acceso son las principales responsabilidades de una biblioteca nacional. En conformidad con la definición dada por el autor mencionado, el Decreto 8.297/2014 establece que una de las competencias de la BNB es captar, preservar y difundir los registros de la memoria bibliográfica y documental nacional.

Con el objetivo de ampliar dos de sus principales competencias, o sea, preservar la memoria cultural y proporcionar el amplio acceso a las informaciones contenidas en su acervo, en 2000 la BNB empezó con el proyecto de digitalización de sus colecciones bibliográficas, proyecto que originó la Biblioteca Nacional Digital (BNDigital). La BNDigital tiene la misión de ser un vehículo diseminador de la memoria cultural brasileña en todo el país y en el exterior, y aumentar el contenido en portugués disponible en la web, entre

otros. Actualmente la BNDigital cuenta con 2.078.154² documentos de libre acceso, y esta cantidad de documentos sigue creciendo constantemente.

Además de las colecciones bibliográficas el acervo de la BNB también está compuesto por el fondo archivístico, que son documentos producidos en el transcurso de las actividades administrativas, cualquiera que sea el soporte, (Conselho Nacional de Arquivos, 2011). Los documentos archivísticos tienen un alto valor administrativo y legal para la institución, principalmente porque la BNB es una institución pública y estos documentos dan soporte a la democracia, protegen y mantienen derechos de los ciudadanos, y convierten las acciones institucionales en transparentes, constituyendo un importante instrumento contra la corrupción.

Debido al Decreto 8.539/2015, que determina el uso de los medios electrónicos para creación de los procedimientos administrativos, la producción, uso y almacenamiento de los documentos archivísticos digitales en la BNB ha cambiado mucho desde enero de 2018, debido a la implantación del Sistema Eletrônico de Informações (SEI)³, que es una herramienta que permite la producción, edición, firma y trámite de documentos dentro del propio sistema, lo que permite la virtualización de procedimientos, y la actuación simultánea de varias unidades en un mismo procedimiento, aunque estén distantes físicamente, reduciendo el tiempo de realización de las actividades.

La creación de la BNDigital y la implantación del SEI cambiaron principalmente el uso y el almacenamiento de las colecciones bibliográficas digitalizadas y del fondo archivístico digital (nativo). Antes de la creación de la BNDigital los documentos estaban, en su gran mayoría, en soporte papel y almacenados en los almacenes de la institución. Hasta la implantación del SEI, los documentos archivísticos eran producidos en el entorno digital, y aquellos con valor administrativo y legal, eran migrados para soporte papel, es decir, su producción era digital, pero su uso, almacenamiento y la gestión eran realizados por medios analógicos.

Considerando que la llegada de las TIC afectó la rutina diaria de la BNB, permitió ampliar la difusión de sus acervos, que es parte de su misión, mejoró la productividad y el rendimiento de sus actividades, y los Decretos 8.297/2014 que en el artículo 2º establece que la BNB debe adoptar medidas necesarias para la preservación y protección del patrimonio bibliográfico y digital bajo su custodia, y 8.539/2015 que en el artículo 18 establece que los órganos o entidades deberán establecer políticas, estrategias y acciones que garanticen la preservación, el acceso y el uso continuo a largo plazo de los documentos digitales. La Biblioteca Nacional de Brasil tendrá que adoptar nuevas estrategias para la preservación y protección de sus colecciones bibliográficas digitalizadas y del fondo archivístico digital (nativo), para cumplir con las obligaciones establecidas en los dispositivos legales mencionados. Una de las estrategias que la BNB puede adoptar para la preservación digital de sus acervos es la implantación de un repositorio digital, que sea adecuado a la gestión y preservación de los materiales bibliográficos y de los documentos archivísticos. La implantación de un repositorio digital en la BNB es una estrategia recomendable por el Decreto 8.539/2015 que en el artículo

² Cantidad de documentos en julio de 2019.

³ Elaborado a partir de las informaciones disponibles en la página web del Software Público Brasileiro: <https://softwarepublico.gov.br/social/sei/sobre-o-sei>

16.2º establece que los procedimientos administrativos electrónicos cuyos trámites ya se han finalizado y que están a la espera del cumplimiento de los plazos y del destino final, se pueden transferir a un área de almacenamiento específica, bajo el control del órgano o la entidad que los produjo, para: preservación, seguridad y acceso por el tiempo necesario.

Actualmente hay una gran variedad de software para repositorios digitales disponible en el mercado, por eso es necesario evaluar algunos para identificar el más adecuado a las necesidades de la BNB. Basado en esta necesidad, el objetivo de esta investigación es evaluar y analizar las características y las funcionalidades de los softwares Archivematica, DSpace y RODA, con el fin de identificar el más adecuado a la gestión y preservación de los materiales bibliográficos y de los documentos archivísticos. Los objetos digitales (digitalizados) que forman la BNDigital y los documentos archivísticos digitales (nativos) producidos por el Sistema Eletrônico de Informações (SEI) es el alcance de la presente investigación.

1.2 OBJETIVOS

1.2.1 Objetivo general

El objetivo general de la presente investigación es evaluar y analizar los siguientes softwares para repositorios digitales: Archivematica, DSpace y RODA, con vistas a proponer el más adecuado a las necesidades de la Biblioteca Nacional de Brasil, es decir, el que posibilitará a la institución cumplir con su misión de difusión y preservación.

1.2.2 Objetivos específicos

Los objetivos específicos de la investigación son:

1. Presentar aspectos institucionales de la Biblioteca Nacional de Brasil, para justificar la necesidad de implantación de un repositorio digital;
2. Evidenciar que las estrategias de preservación digital ayudan a la Biblioteca Nacional de Brasil a cumplir con su misión y competencias;
3. Evaluar y analizar las características y las funcionalidades de los softwares para repositorios digitales, con el fin de identificar el más adecuado para la gestión y preservación de los materiales bibliográficos y de los documentos archivísticos.

1.3 ASPECTOS METODOLÓGICOS

La metodología utilizada para realizar esta investigación se desarrolló en diferentes etapas: la primera etapa realizó una revisión de la estructura organizacional de la Biblioteca Nacional de Brasil. Este estudio se restringió a los documentos constitutivos de la institución, es decir, los documentos legales donde están definidos su misión, sus

objetivos, competencias y actividades. La finalidad del estudio fue identificar en la legislación las obligaciones legales con respecto a la preservación y difusión, que podrán refrendar los objetivos propuestos en este trabajo.

La segunda etapa de la investigación fue una revisión bibliográfica sobre los temas de preservación digital y sus estrategias (metadatos, estándares de formatos, migración y conversión), repositorios digitales, repositorios institucionales y repositorios de documentos administrativos. Las bases de datos LISA (Library & Information Sciences Abstracts) y LISTA (Library, Information Science & Technology Abstracts), la red social de investigadores ResearchGate⁴, las publicaciones de la Library of Congress y The National Archives of UK fueron las principales fuentes consultadas para poner en marcha esta etapa. Las bases de datos LISA y LISTA fueron elegidas porque son especializadas en la gestión de la información, poseen reconocimiento internacional y actúan en el mercado desde 1969. ResearchGate porque es la red social más grande en esta área, posee más de 130 millones de publicaciones y conecta a más de 15 millones de investigadores; Library of Congress y The National Archives of UK son dos instituciones con reconocimiento internacional en el área de preservación digital. Al ser un tema de constante actualización, la preferencia fue por trabajos publicados a partir del año 2008, pero algunos estudios publicados antes del año 2008, también fueron revisados por ser trabajos muy reconocidos por la comunidad de preservación digital. La revisión de la bibliografía fue realizada en los siguientes idiomas: español, inglés y portugués. Toda información bibliográfica se procesó con el software Mendeley, que además se empleó para la redacción de la bibliografía, que está elaborada según las normas APA.

La tercera etapa de la investigación consistió en elegir los softwares a ser evaluados y analizados. Las características de los softwares para repositorios digitales que han sido consideradas más importantes y relevantes para la Biblioteca Nacional de Brasil fueron el tipo de licencia, el porcentaje de instalaciones registradas en el mundo y los tipos de instituciones que los utilizan. Archivemática, DSpace y RODA fueron los softwares que mejor respondieron a las características definidas. Además de elegir los softwares que debían ser evaluados, en esa etapa también se seleccionaron las funcionalidades más importantes que los softwares deberían poseer, para atender a los objetivos de esta investigación. A partir de estos objetivos se definió que dichas funcionalidades serían seleccionadas a partir de los requisitos de complejidad de instalación, usabilidad, flexibilidad y escalabilidad y agrupadas en seis categorías: Infraestructura, Gestión de contenidos, Administración del uso, Interfaz de usuario, Preservación digital e Interoperabilidad e integración.

En la cuarta etapa se definió cómo se realizaría la evaluación para lograr el resultado. De los modelos de evaluación consultados, el creado por Alvite y Rodríguez (2002), para evaluar las interfaces de OPACs implementadas con UNICORN en las Universidades madrileñas, se reveló el más adecuado para esta investigación. El modelo consiste en atribuir 0 (ausencia), 1 (aplicación incompleta/limitada) y 2 (aplicación completa), para cada funcionalidad. Para atribuir la puntuación propuesta en el modelo fue necesario consultar los manuales técnicos, las páginas web oficiales y las comunidades de soporte de los softwares, con el fin de confirmar si el software tenía las funcionalidades seleccionadas.

⁴ <https://www.researchgate.net/>

PARTE II – CONTEXTO TEÓRICO

2. BIBLIOTECA NACIONAL DE BRASIL (BNB)

2.1 Presentación

Con más de 200 años de historia, la Biblioteca Nacional de Brasil (BNB)⁵ es la más antigua institución cultural brasileña, anterior a la constitución de Brasil como nación independiente. La BNB es la institución que preserva el patrimonio bibliográfico y documental de Brasil, posee un acervo de aproximadamente 10 millones de ítems, lo que la convierte en la mayor biblioteca de Latinoamérica. Su sede está en la ciudad de Río de Janeiro.

2.2 Historia

El origen de la Biblioteca Nacional de Brasil está relacionado con la Biblioteca Real de Portugal. Con la llegada de la familia real a Brasil, la Biblioteca Real de Portugal también fue trasladada. Un Decreto de 29 de octubre de 1810 es el marco oficial de la creación de la Biblioteca Nacional de Brasil. El acervo de 60.000 ítems fue instalado en la Orden Tercera del Carmen, pero la Biblioteca, aunque inaugurada el 13 de mayo de 1811, sólo fue abierta al público en 1814. Antes de eso, el acceso estaba restringido a estudiosos a petición. En los primeros años en Brasil, el acervo tuvo un enriquecimiento considerable, datando de esa época la adquisición de importantes colecciones, como la de Manuel Inácio da Silva Alvarenga, la colección del arquitecto José da Costa e Silva, con dibujos originales, libros, manuscritos, estampas grabadas y camafeos⁶, la colección del conde de la Barca, denominada Colección Araujense, con 2.365 obras en 6.329 volúmenes, muchas obras de artistas italianos entre otras.

Después de la independencia de Brasil, la biblioteca fue comprada por el Imperio del Brasil por 800 contos de réis⁷, cantidad, entonces, considerada exorbitante, como parte de la Convención Adicional al Tratado de Amistad y Alianza firmado entre Brasil y Portugal el 29 de agosto de 1825.

Debido a la necesidad de acomodar mejor un acervo en constante crecimiento, en 1905 comenzó las obras de construcción de una nueva sede para la Biblioteca. Con el proyecto del General Francisco Marcelino de Sousa Aguiar e influencias de los estilos neoclásico y art nouveau, la nueva y actual sede de la Biblioteca Nacional de Brasil fue inaugurada el 29 de octubre de 1910. El edificio tiene salas de lectura y estudio para el público, divisiones y oficinas de trabajo para el personal de la administración y otras dependencias.

⁵ Elaborado a partir de las informaciones disponibles en página web de la Biblioteca Nacional de Brasil: <https://bn.gov.br/sobre-bn/historico>

⁶ El camafeo (del latín *Cammaeus*, que significa piedra esculpida) es una técnica de escultura usada en joyas para formar una figura en relieve (en oposición a la muesca) sacando partido de la existencia de capas superpuestas de colores diferentes. Tradicionalmente esta técnica se hace en broches y colgantes

⁷ Fue la moneda de Portugal desde 1500 hasta 1911.

La Biblioteca Nacional de Brasil estuvo subordinada al antiguo Ministerio del Interior y Justicia, luego al Ministerio de Educación y Salud. Con la creación del Ministerio de Salud, ella pasó a integrar el Ministerio de Educación y Cultura. En 1990 la Biblioteca Nacional, con su biblioteca subordinada, Euclides da Cunha, de Río de Janeiro, el Instituto Nacional del Libro y la Biblioteca Demostrativa, de Brasilia, pasaron a constituir la Fundación Biblioteca Nacional (FBN).

2.3 Misión, función y competencias

Para Takeshy (2000) misión es la propuesta fundamental de la institución, para qué sirve, cuál es su justificación de existencia para la sociedad y la función social que ejerce. Según Drucker (1993) la misión es la base para las prioridades, estrategias, planes y tareas de trabajo. Lor (1997) define Biblioteca Nacional como una institución, principalmente financiada (directa o indirectamente) por el gobierno, que es responsable de recoger, registrar, preservar y poner a disposición el patrimonio bibliográfico y documental provenientes o relacionados a su país. En consonancia con las definiciones de misión y de biblioteca nacional el Decreto 8.297/2014 en el artículo 2º establece que la Fundación Biblioteca Nacional es el órgano responsable de la ejecución de la política gubernamental de captación, guarda, preservación y difusión de la producción intelectual del país. Además establece sus competencias, que entre otras son:

I - captar, preservar y difundir los registros de la memoria bibliográfica y documental nacional;

II - adoptar las medidas necesarias para la preservación y protección del patrimonio bibliográfico y digital bajo su custodia.

A partir de sus competencias es posible concluir que la Biblioteca Nacional de Brasil es una institución gubernamental con la misión de preservar el patrimonio bibliográfico y documental del país y ponerlos accesibles, es decir sus competencias son preservar y dar acceso.

2.4 Acervo digital

A partir del surgimiento de la Internet, el acervo de la Biblioteca Nacional de Brasil sufrió profundos cambios. En vista de estos cambios, la BNB viene buscando utilizar las nuevas tecnologías para cumplir su misión y competencias primordiales de preservar y dar acceso a la memoria bibliográfica/documental que se encuentra bajo su custodia, al mayor número posible de ciudadanos. Esto fue posible a través de las iniciativas de digitalización de las colecciones bibliográficas y la adopción del sistema informatizado para los documentos administrativos. Las dos iniciativas utilizadas para garantizar el amplio acceso y mantener el concepto de democracia, fueron la creación de la BNDigital y la implantación del Sistema Eletrônico de Informações (SEI), que actualmente constituye el acervo digital de la BNB.

2.4.1 Biblioteca Nacional Digital (BNDigital)

Las primeras iniciativas de digitalización en la BNB se realizaron en 1998, con ocasión del lanzamiento de la primera versión de su sitio web, cuando se realizaron algunas demostraciones de asociaciones de multimedia con los registros bibliográficos. El objetivo era probar el sistema con los diferentes tipos de objetos digitales, como textual, visual y sonoro. Estas iniciativas despertaron el interés del lector virtual. (Bettencourt, 2014).

A pesar de los resultados positivos obtenidos en la iniciativa, la falta de recursos financieros necesarios para la digitalización sistemática de las colecciones impidió la ampliación de las iniciativas. Así, la entrada de la Biblioteca Nacional de Brasil en la era digital ocurrió casi dos años después.

Al final del año 2000 el proyecto “Biblioteca Virtual da Cartografia Histórica dos séculos XVI a XVIII”, subvencionado por la Financiadora de Estudos e Projetos (Finep)⁸, permitió la creación del laboratorio de digitalización de la BNB, el cual posibilitó la implantación de la biblioteca digital, es decir este proyecto fue el marco de la creación de la BNDigital⁹.

Oficialmente la BNDigital fue creada en 2006, compuesta por las colecciones digitalizadas desde del año 1998, en el contexto de exposiciones y de proyectos temáticos, en asociación con instituciones nacionales e internacionales.

A partir de 2008 la BNDigital empezó a recibir aporte financiero del Ministerio de Cultura, mediante la inclusión en el “Programa Livro Aberto¹⁰”, cuya finalidad era ampliar y democratizar el acceso de la población a obras que componen el Acervo Memoria Nacional.

La BNDigital está internamente constituida por tres segmentos: captura y almacenamiento de colecciones digitales, tratamiento técnico y publicación de colecciones digitales y programas y proyectos de digitalización y divulgación. Su equipo es interdisciplinario y está compuesto por bibliotecarios, archiveros, historiadores y técnicos en digitalización.

Actualmente la BNDigital recibe cerca de 500 mil accesos mensuales para consulta de los más de 2.078.154 documentos disponibles, que corresponden a más de 11 millones de páginas web.

2.4.1.1 Misión y objetivos

La BNDigital amplía una de las competencias de la BNB que es difundir los registros de la memoria bibliográfica y documental nacional, es decir ampliar el acceso a la información contenida en su acervo. Sus objetivos se basaron en sus competencias, y son ellos:

⁸ Financiadora de Estudos e Projetos (Finep), es una empresa pública brasileña de fomento a la ciencia, tecnología e innovación en empresas, universidades, institutos tecnológicos y otras instituciones públicas o privadas, con sede en Río de Janeiro. La empresa está vinculada al Ministerio de Ciencia y Tecnología e Innovación.

⁹ Elaborado a partir de las informaciones disponibles en la página web de la Biblioteca Nacional Digital: <http://bndigital.bn.gov.br/sobre-a-bndigital/>

¹⁰ El Programa promueve la instalación y modernización de bibliotecas grandes, medianas, pequeñas e itinerantes y tiene como objetivo garantizar y democratizar el acceso al libro.

- Ser vehículo diseminador de la memoria cultural brasileña en todo país y en el exterior;
- Aumentar el contenido en portugués disponible en la web;
- Alcanzar los usuarios, independiente de la localización geográfica;
- Preservar los documentos originales evitando el manejo innecesario;

2.4.1.2 Política de digitalización

La conversión de objetos analógicos a objetos digitales propició nuevas formas de acceso a las colecciones de la BNB y contribuyó para la preservación de las obras originales. La política de digitalización¹¹ establece los procesos y criterios de evaluación y selección de objetos a digitalizar para el desarrollo de las colecciones digitales. La política de selección para digitalización es basada en los criterios del valor histórico, de la importancia y de la rareza de las obras, así como la relevancia de las colecciones, en su totalidad o en parte, que permite la contextualización y la interrelación de las obras que componen la BNDigital. A partir de estos criterios es establecida las prioridades para selección que son:

1. Documentos que reflejaron o reflejan la misión de la BNB;
2. Obras en dominio público, huérfana o cuya reproducción sea autorizada por el titular de los derechos intelectuales y morales;
3. Obras muy demandadas por los usuarios;
4. Obras raras, inéditas o exclusiva;
5. Obras frágiles o dañadas/deterioradas que el acceso y el manejo implican riesgos a su integridad; y
6. Obras con potencial colaborativo, que complemente y/o se complementa por colecciones digitales de otras instituciones.

2.4.1.3 Colecciones digitales

En la BNDigital posee colecciones significativas de la memoria bibliográfica y documental brasileña, estas colecciones son formadas por libros, mapas, fotografías, dibujos, grabados, discos, partituras, revistas, periódicos, manuscritos, y otras publicaciones, un total de 2.078.154 documentos convertidos al formato digital. En la BNDigital están disponibles colecciones certificadas¹² con el Registro de la Memoria del Mundo¹³ de la UNESCO, algunas de ellas son:

- **Lima Barreto** - cartas, originales literarias, recortes de periódicos y documentos personales de Afonso Henriques de Lima Barreto (Río de Janeiro, 13 de Mayo, 1881 - 1 de noviembre de de 1922).

¹¹ Elaborado a partir de las informaciones disponibles en la página web de la Biblioteca Nacional Digital: <http://bndigital.bn.gov.br/sobre-a-bndigital/>

¹² Elaborado a partir de las informaciones disponibles en la página "Memoria do Mundo" de la Biblioteca Nacional de Brasil: <https://www.bn.gov.br/explore/memoria-mundo>

¹³ Es una lista del patrimonio documental que ha sido aprobado por el Comité Consultivo Internacional y ratificado por el Director General de la UNESCO como elemento que cumple los criterios de selección del patrimonio documental considerado de importancia mundial.

- **Artur Ramos** - médico, etnólogo y profesor Arthur Ramos.
- **Guerra del Paraguay** - El conjunto “A Guerra da Tríplice Aliança: representações iconográficas e cartográficas” de nueve instituciones brasileñas y de la Biblioteca Nacional del Uruguay.
- **Cultura y Opulencia de Brasil** - de André João Antonil - pseudónimo del jesuita italiano João Antônio Andreoni.
- **Manuscritos musicales de Ernesto Nazareth** - En el año de 1863, se convirtió en uno de los más creativos compositores brasileños.
- **Atlas y mapa del cartógrafo Miguel Antônio Ciera (1758)** - Primer conjunto iconográfico de la región que, en el siglo XVIII, pasó a recibir el nombre de Pantanal, con cartas geográficas y registros a lápiz y la acuarela de elementos de la fauna, tipos populares y vistas de paisajes.
- **Matrices del grabado de la casa literaria del arco del ciego** - Matrices originales de cobre grabadas, producidas por la tipografía de la Casa Literaria Arco del Ciego, enviadas a la Impresión Régia en Brasil.
- **Viaje Filosófico": Expedición científica de Alexandre Rodrigues Ferreira en las Capitánias del Gran Pará, Río Negro, Mato Grosso y Cuiabá - 1783-1792** - representa en su mayoría la botánica y la fauna de Brasil en el siglo XVIII.
- **Manuscritos musicales de Carlos Gomes** - Que se ha convertido en una de las más antiguas del mundo. Autor de “O Guarani”, se destacó por el estilo romántico.
- **Carta Régia de la apertura de los Puertos Brasileños a las naciones amigas** - En primer Decreto emitido por el rey D. João VI en Brasil, cuatro días después de llegar a Brasil en 1808, se autoriza la apertura de los puertos brasileños al comercio con las naciones amigas de Portugal.
- **La colección del Emperador: Fotografía brasileña y extranjera del siglo XIX** – Colección con fotografías de Brasil y del mundo en el siglo XIX, que fueron donadas por D. Pedro II e integran la Colección Thereza Christina Maria.

La mayor parte de las colecciones disponibles en la BNDigital es parte de la Hemeroteca Digital brasileña, que ofrece acceso a los periódicos nacionales, revistas, anuarios, boletines y publicaciones seriadas de gran importancia para el país.

2.4.2 Sistema Eletrônico de Informações (SEI)

El Sistema Eletrônico de Informações (SEI)¹⁴ es una solución desarrollada por el Tribunal Regional Federal de la 4ª Región (TRF4). Se trata de un sistema de gestión de procedimientos electrónicos, que permite la producción, edición, firma y demás trámites en medio digital.

Las ventajas del uso del SEI incluyen reducción de costes operativos (como gastos en papel, impresión, costes de envío, etc.), optimización de las rutinas de trabajo, reducción

¹⁴ Elaborado a partir de las informaciones disponibles en la página web del Software Público Brasileiro: <https://softwarepublico.gov.br/social/sei>

de tiempo de trámite de los procedimientos, firma de documentos electrónicamente, entre otras.

El SEI es cedido gratuitamente a los órganos, mediante la firma del acuerdo de cooperación técnica con el Ministerio de Planificación, Desarrollo y Gestión.

2.4.2.1 Historia y bases legales

En octubre de 2015 la Presidencia de la República de Brasil, publicó el Decreto 8.539/2015, de acuerdo con ese dispositivo, los órganos y entidades de la administración pública directa, autárquica y fundacional deben adoptar algún sistema informatizado para la gestión y el trámite de procedimientos administrativo, en el plazo de dos años, es decir hasta octubre de 2017.

Esta determinación es parte del Processo Eletrônico Nacional (PEN)¹⁵, que es una iniciativa conjunta de órganos y entidades de diversas esferas de la administración pública de Brasil, con el propósito de construir una infraestructura pública de procedimientos y documentos administrativos digitales, para mejorar el desempeño de los procedimientos de la administración pública, con enfoque en la agilidad, productividad, transparencia, satisfacción de los ciudadanos y reducción de costes.

El Decreto determinaba que los órganos y entidades deberían adoptar algún sistema informatizado, pero no determinaba cual, cada órgano e institución podría elegir cualquier solución de sistema informatizado para la gestión de los procedimientos administrativo. El órgano e institución que ya utilizaban alguno sistema informatizado podrían mantenerlos.

Para facilitar el cumplimiento del Decreto, el gobierno empezó a evaluar y analizar sistemas informatizados en el mercado, con el fin de elegir una solución de procedimiento electrónico para integrar el PEN. Debido a los buenos resultados en la Justicia Federal de la 4º Región (TRF4) y la cesión gratuita, el SEI fue elegido como la solución para procedimiento electrónico del PEN.

El SEI fue cedido gratuitamente a BNB a través del de acuerdo de cooperación técnica firmado electrónicamente con el Ministerio de Planificación, Desarrollo y Gestión el 14 de octubre de 2016, correspondiendo al procedimiento administrativo n ° 05110.00259 / 2016-37.

En virtud de algunas dificultades, el plazo para implantación tuvo que ser ampliado con la anuencia del Ministerio de Planificación, Desarrollo y Gestión. En el mes de septiembre de 2017 el SEI fue instalado y parametrizado de acuerdo con las necesarias de la BNB, sin embargo, fue verificado que, dada la complejidad del sistema, era necesario la contratación de una empresa para el servicio de configuración del sistema. La contratación fue firmada y la configuración realizada en el mes de octubre de 2017.

En los días 18, 19 y 20 de diciembre de 2017 fue realizado el 1º Ciclo de Capacitación en el SEI, fueron capacitados 68 funcionarios. La capacitación fue importante para entrenar los funcionarios en el sistema y también para verificar si la configuración y las parametrizaciones realizadas estaban funcionando correctamente.

¹⁵ Elaborado a partir de las informaciones disponibles en la página web del Software Público Brasileiro: <https://softwarepublico.gov.br/social/sei/pen-projeto-eletronico-nacional>

Como la operación del sistema estaba de acuerdo con el planificado, en 02 de enero de 2018 la BNB empezó la producción y el trámite electrónico de sus procedimientos administrativos a través del SEI.

2.4.2.2 Organización administrativa/funcional

Para atender el Decreto 8.539/2015 de la Presidencia de la Republica de Brasil la Coordinación General de Planificación y Administración (CGPA) de la BNB determinó la creación de una unidad administrativa específica para implantación, gestión, fiscalización y evaluación de la ejecución del SEI.

El 6 de junio de 2016 empezó a funcionar la División de Gestión Documental (DGD), que además de ser responsable por la gestión del SEI, también es responsable por la política de gestión documental de la BNB, con el objetivo de asegurar de forma eficiente la producción, mantenimiento y destinación final de los documentos administrativos, y asegurar que la información gubernamental esté disponible cuando y donde sea necesario a la institución, al gobierno y a los ciudadanos.

2.4.2.3 Fondo archivístico digital

El acervo digital de la BNB está formado por las colecciones bibliográficas y también por el fondo archivístico. Los documentos archivísticos digitales dan soporte al concepto de democracia, para la protección y al mantenimiento de derechos, ellos constituyen pruebas documentales confiables y precisas para toma de decisiones y acciones institucionales, además es un importante instrumento contra la corrupción.

El procedimiento administrativo es un documento archivístico esencial para la administración pública de Brasil, según definición del Arquivo Nacional de Brasil (2005, p. 138,) “el procedimiento es un conjunto de documentos oficialmente reunidos en el curso de una acción administrativa o judicial, que constituye una unidad de archivo”. Es decir un procedimiento es formado por varios documentos archivísticos. Pero la importancia del procedimiento administrativo no es solamente porque es formado por varios documentos archivísticos, también porque de acuerdo con Di Pietro (2016) el procedimiento administrativo es un instrumento indispensable para el ejercicio de las funciones administrativas de la administración pública. Es el instrumento por medio del cual la administración pública materializa sus actos, todas las operaciones administrativas o actos jurídicos, quedan documentados en un procedimiento administrativo.

Debido a su gran importancia para la administración pública el procedimiento administrativo fue elegido para ser gestionado a través del SEI, que es un sistema que gestiona solamente procedimientos, es decir no gestiona documentos sueltos.

El fondo archivístico digital de la BNB, en la actualidad, es formado por los procedimientos administrativos digitales y por los assentamentos funcionais digitais (ADF)¹⁶, pero aunque los ADF pertenecen al fondo archivístico de la BNB su almacenamiento y gestión están

¹⁶ Carpeta que reúne todos los documentos recibidos y producidos a lo largo de la vida profesional de los funcionarios públicos.

bajo la responsabilidad del Ministerio de Planificación, Desarrollo y Gestión, por esta razón estos documentos no están en el alcance de esta investigación.

La gran mayoría de los documentos archivísticos producidos por la BNB, son incluidos en los procedimientos administrativos digitales los demás son migrados para soporte papel para la gestión. Es decir los documentos archivísticos digitales bajo la responsabilidad de la BNB para la gestión son los procedimientos administrativos digitales.

3. PRESERVACIÓN DIGITAL

El tema de la preservación digital aunque parezca reciente, fue tratado por la primera vez en 1964 en el International Congress on Archives realizado en París. Sin embargo, el informe del profesor Robert Henri Bautier para el International Council on Archives - ICA presentado en la Conférence Internationale de la Table Ronde des Archives – CITRA en 1971, en Bonn, Alemania, es considerado el punto de partida de las discusiones sobre este tema.

Según Tinoco (2013), a partir de 1971 han surgido diversos proyectos e iniciativas en el ámbito de la preservación digital, de los cuales se destacan, a nivel internacional, el CEDARS¹⁷ (Consortium of University Research Libraries Exemplars in Digital Archives), InterPARES¹⁸ (International Research on Permanent Authentic Records in Electronic Systems), ERA/NARA¹⁹ (Electronic Records Archives of National Archives and Records Administration), Modelo de Referencia OAIS²⁰ (Open Archival Information System), PANDORA²¹ (Preserving and Accessing Networked Documentary Resources of Australia), PRISM²² (Preservation, Reability Interoperability, Security, Metadata), entre otros.

Estos proyectos e iniciativas de preservación digital tienen el mismo objetivo, que es desarrollar políticas, modelos, estrategias y esquemas de preservación digital para garantizar el uso de la información digital por el tiempo que se necesario. Eso porque la tecnología está cambiando a un ritmo muy rápido, que parece que a cada día se crea un nuevo software, hardware, dispositivo de almacenamiento digital o formato de documento. Este cambio es una amenaza constante de perder el acceso a la información digital, sin tener en cuenta alguna estrategia de preservación digital, la información digital está en riesgo y estos riesgos incluyen la pérdida de datos y la incapacidad de acceso.

¹⁷ Proyecto de preservación digital del Reino Unido financiado inicialmente por JISC a través del Programa de bibliotecas electrónicas (eLib). El proyecto duró cuatro años, desde mediados de 1998 hasta finales de marzo de 2002.

¹⁸ Es una iniciativa académica de investigación en preservación digital, específicamente dirigido a los documentos archivísticos, es coordinado por la Universidad de Columbia Británica, en Canadá, ha desarrollado un conocimiento teórico-metodológico esencial para la preservación a largo plazo de documentos archivísticos digitales auténticos.

¹⁹ Archivos de Registros Electrónicos (ERA) diseñado para conservar y administrar los registros electrónicos de NARA y para administrar el ciclo de vida de los registros y otras participaciones, incluido el soporte para los programas de retención de registros y el proceso de acceso para todos Registros federales.

²⁰ Es un marco conceptual para un sistema de archivo dedicado a preservar y mantener el acceso a la información digital a largo plazo. El propósito del modelo de referencia es aumentar el conocimiento y la comprensión de los conceptos relevantes para archivar objetos digitales.

²¹ El Archivo Web de Australia fue creado por la Biblioteca Nacional en 1996 y es una colección de publicaciones históricas en línea relacionadas con Australia y los australianos. Las publicaciones en línea y los sitios web se seleccionan para su inclusión en la colección con el fin de proporcionarles un acceso persistente y a largo plazo.

²² El Proyecto PRISM de la Universidad de Cornell, se centra en la aplicación de políticas para garantizar la integridad de la información en las áreas de preservación, confiabilidad, interoperabilidad, seguridad y metadatos. PRISM está investigando la supervivencia a largo plazo de la tecnología digital.

3.1 Concepto y definición

Para entender la importancia de la preservación digital para el acceso a la información digital en el futuro, es necesario analizar algunas definiciones sobre el tema, así como sus objetivos:

La preservación digital tiene como objetivo garantizar el acceso a la información digital íntegra y auténtica a largo plazo, de modo que sus funciones pueden ser reproducidas correctamente por una tecnología en el futuro, diferente de la que le originó en el pasado (Conselho Nacional de Arquivos, 2014).

Thurston (2012), define la preservación digital desde el ámbito de acción gubernamental a largo plazo, donde la preservación de documentos digitales depende incondicionalmente de la combinación de software y hardware en constante cambio. Si se desea que los documentos digitales que dependen de las políticas, acciones y transacciones de los gobiernos sean accesibles, auténticos y utilizables a largo plazo, se deben cumplir estándares profesionales internacionales como el de registro de metadatos de preservación para que se defina claramente ¿quién los creó?, ¿cuándo?, ¿dónde?, ¿cómo? y ¿por qué?

Para el Parliamentary Archives (2008), la preservación digital ahora está atrayendo mucha atención del gobierno, del sector público y privado y de los medios de comunicación. Eso porque garantiza que un recurso digital permanezca auténtico y accesible a largo plazo, frente a los cambios y la obsolescencia del hardware y software.

Arellano (2004), presenta tres niveles de preservación digital: física, lógica y conceptual: según el autor la preservación física está centrada en los contenidos almacenados en Medios magnéticos (cintas de audio, cintas VHS y DAT etc.) y discos ópticos (CD-ROM, WORM, discos ópticos regrabables). La preservación lógica busca en la tecnología formatos actualizados para la inserción de los datos (correo electrónico, material de audio y audiovisual, material en red, etc.), nuevos software y hardware que manteniendo vigentes sus bits, preserva su capacidad de lectura. En el caso de la preservación conceptual, el foco son los mecanismos que garantizan la integridad y autenticidad de la información de los documentos digitales.

La UNESCO (2003), a través de su carta para la preservación del patrimonio archivístico digital considera que el objetivo de la preservación de los documentos es garantizar la autenticidad y la integridad de la información. El desafío de la preservación de documentos digitales está en garantizar el acceso a los contenidos digitales, por medio de los recursos tecnológicos disponibles a la época en que ocurrir su uso.

A partir de las definiciones mencionadas, se puede concluir que el enfoque principal de la preservación digital es garantizar el acceso a la información digital, íntegra y auténtica, por medio de los recursos tecnológicos disponibles en el futuro.

3.2 Estrategias de preservación digital

Según Flores y Santos (2014), las estrategias de preservación digital pueden ser entendidas con un conjunto de objetivos y métodos para garantizar el acceso a largo plazo de la información digital, contemplando sus respectivos objetos digitales y su información relacionada. Las estrategias componen un grupo de acciones alternativas, cada una de ellas con el objetivo de preservar la integridad y la autenticidad de los documentos, garantizando un acceso continuo. En este contexto, las estrategias podrán contemplar la preservación del objeto digital en el ámbito de sus tres niveles: físico, lógico y conceptual. El objeto físico es el soporte donde la información está registrada, el lógico es el software/formato el cual el documento fue producido y conceptual es la presentación del documento de manera comprensible por los humanos.

Thibodeau (2002), sugiere que antes de elegir una estrategia de preservación se debe tener en cuenta cuatro criterios:

1. **Viabilidad:** propiedad de los dispositivos de software y hardware de la aplicación con dominio del cumplimiento de la estrategia seleccionada para la preservación.
2. **Sostenibilidad:** la estrategia seleccionada debe ser suficientemente competente para aplicarse indeterminadamente en la perspectiva futura, o debe haber una ruta alternativa si los métodos seleccionados dejan de funcionar.
3. **Practicidad:** la estrategia seleccionada debe ser racional con respecto a la dificultad en la implementación y el retorno de la inversión.
4. **Adecuación:** la estrategia seleccionada debe ser apropiada para que las formas específicas de los objetos digitales sean preservadas de acuerdo en el propósito para que se está preservando.

Las estrategias para la preservación digital están agrupadas en dos tipos: las estructurales y las operativas. Arellano (2004), se considera que las estructurales tratan de las inversiones iniciales por parte de las instituciones que se están preparando para implementar alguna estrategia de preservación digital y las operativas son las medidas concretas aplicadas a los objetos digitales. En la tabla 1, a continuación, están las estrategias de preservación más utilizadas:

Estrategias Estructurales	Estrategias Operativas
Adopción de estándares/patrones	Preservación de software y hardware
Elaboración de normas	Migración de soporte
Metadatos de preservación	Conversión de formatos
Infraestructura	Emulación
Formación de consorcios	Preservación del contenido

Tabla 1. Tipos de estrategias de preservación digital.
Fuente: Arellano (2004)

3.2.1 Estándares de formatos

El debate sobre cual formato de documento se debe utilizar para preservar y proveer acceso a los objetos digitales en el futuro está cada vez más corriente, eso porque según Brown (2008), los formatos de documentos son información codificada que sólo puede ser procesada y comprensible por combinaciones específicas de hardware y software, y su accesibilidad está vulnerable en la actualidad, debido al ambiente tecnológico que está evolucionando rápidamente. Eso preocupa a todos los responsables de la gestión y acceso a los objetos digitales, en una escala de tiempo relativamente corta. Rimkus, Padilla, Popp et al (2014), afirma que los formatos abiertos son más indicados que los formatos propietarios, porque la forma en que codifican el contenido es transparente, pero la adopción de un formato de documento propietario por una amplia comunidad de creadores de contenido, divulgadores y usuarios, puede ser considerado como un indicador confiable de la longevidad del formato. En este sentido, el autor presenta que no existe, de hecho, ninguna fórmula infalible para decisiones de estándares de formatos de documentos.

Según la Kultura 2.0 (2011), en el momento de seleccionar el formato más adecuado para un documento, con el fin de crear estándares, se aconseja tener en cuenta dos categorías básicas de factores: preservación (sostenibilidad) y difusión (calidad y funcionalidad) como se puede observar en la figura 1, a continuación:

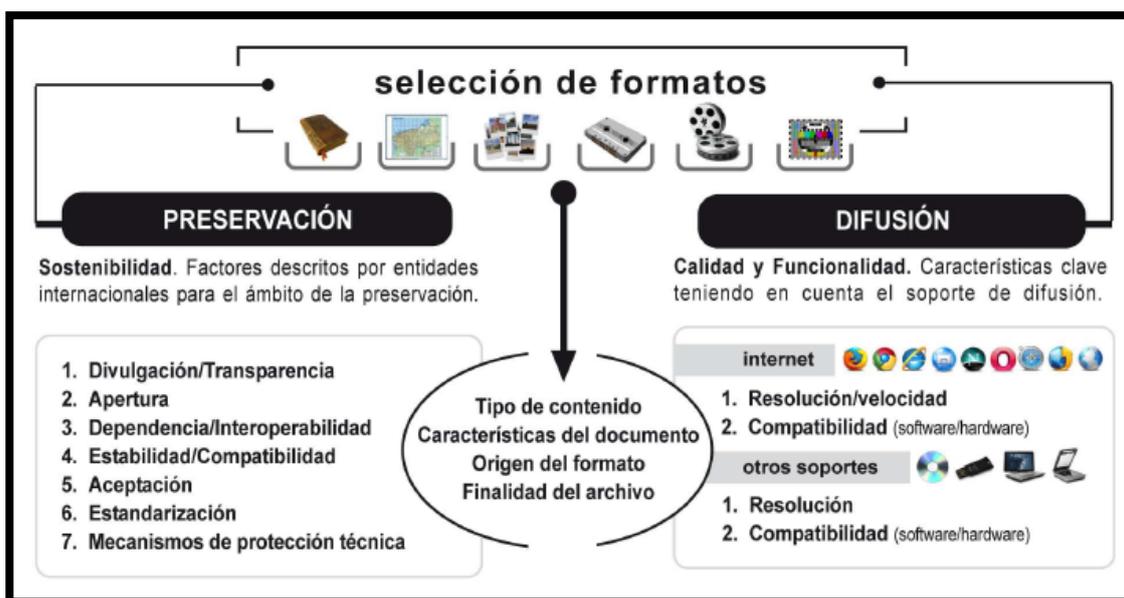


Figura 1. Selección de formatos.
Fuente: Kultura 2.0 (2011)

La Kultura 2.0 (2011, p.12), además enumera los factores de sostenibilidad a tener en cuenta en la selección de los formatos, para creación de estándares, que son:

1. **Divulgación/Transparencia:** tener en cuenta se las especificaciones técnicas de los formatos son de código no propietario.

2. **Apertura:** tener en cuenta si el formato de documento es dependiente de una patente. Instituciones destinadas a la preservación y difusión, en general, utilizan formatos no propietarios.
3. **Dependencia/Interoperabilidad:** tener en cuenta si el formato es dependiente de un hardware, software o de un sistema informático específico.
4. **Estabilidad/Compatibilidad:** tener en cuenta si el formato mantiene su funcionalidad, integridad y accesibilidades con versiones anteriores, actuales o posteriores.
5. **Aceptación:** tener en cuenta el grado de madurez y utilización de los formatos por parte de los usuarios, creadores y distribuidores de los recursos en nivel nacional y/o internacional.
6. **Estandarización:** tener en cuenta si los formatos están adecuados a las especificaciones establecidos por un organismo de normalización (ISO, NISO, W3C...). La normalización de formato asegura la calidad y accesibilidad de los documentos a largo plazo.
7. **Mecanismos de protección técnica:** tener en cuenta que el cifrado (que es un mecanismo de protección técnica) que es utilizado para proteger la información de acceso no autorizado, no deben complicar la recuperación de la información, migración o su adecuación a nuevas necesidades originada de la evolución tecnológica.

Con el fin de crear patrones para garantizar la interoperabilidad, el Gobierno de Brasileiro desarrolló un trabajo en 2007, (actualizado en 2017), llamado e-PING Brasil - Padrões de Interoperabilidade de Governo Eletrônico – que define un conjunto mínimo de premisas, políticas, patrones y especificaciones técnicas que regulan la utilización de la tecnología de información y comunicación (TIC). Entre las especificaciones y patrones establecidos, están los formatos que los órganos del poder ejecutivo federal deben, utilizar. La figura 2, a continuación, presenta los formatos adoptados y recomendados:

Componente	Especificação	SIT	Observações
	A = Adotado R = Recomendado T = Em Transição E = Em Estudo		
Arquivos do tipo documento/publicação	Texto puro (arquivo .txt)	A	
	Open Document ODF 1.2 (.odt) - conforme especificação da OASIS ⁽¹⁰⁾	A	
	EPUB 3.0.1 http://idpf.org/epub/301	R	
	Portable Document Format - PDF ISO 32000-1:2008	R	
	Portable Document Format - PDF/A NBR ISO 19005-1:2009 ⁽¹¹⁾ , quando necessária a preservação digital de documentos	R	
Arquivos do tipo planilha	Open Document ODF 1.2 (.ods) - conforme especificação da OASIS	A	
Arquivos do tipo apresentação	Open Document ODF 1.2 (.odp) – conforme especificação da OASIS	A	
	HTML (.html ou .htm), conforme especificações do W3C.	R	
Intercâmbio de informações gráficas e imagens estáticas	W3C PNG (.png), ISO/IEC 15948:2003 (E) http://www.w3.org/TR/PNG/	A	
	SVG (.svg), gerado conforme especificações do W3C ⁽¹²⁾ .	R	
	JPEG File Interchange Format (.jpeg, .jpg ou .jffif) ⁽¹³⁾ .	R	
Gráficos vetoriais	SVG (.svg), gerado conforme especificações do W3C.	R	
Animação	SVG (.svg), gerado conforme especificações do W3C.	R	
Áudio	Ogg Vorbis (.ogg, .oga) ⁽¹⁴⁾ .	R	
	Ogg FLAC (.ogg, .oga)	R	
	FLAC (.flac)	R	
Video	Ogg Theora (.ogg, .ogv) ⁽¹⁵⁾ .	R	
	Matroska (.mkv)	R	
	Áudio e vídeo MPEG-4, Part 14 (.mp4) ⁽¹⁶⁾ .	T	
Compactação de arquivos	ZIP (.zip).	R	
	GNU ZIP (.gz).	R	
	Pacote TAR (.tar).	R	

Figura 2. Especificaciones técnicas de los componentes.
Fuente: Governo Brasileiro (2017, p. 30-31)

La selección de los formatos adoptados y recomendados por el e-PING Brasil, presentada en la Figura 2, se fundamentó en las especificaciones técnicas, principalmente, en las de patrones abiertos y soporte en el mercado. En la tabla 2, a continuación, está la justificación para selección de los formatos más utilizados en la administración pública de Brasil:

Tipo de Documento	Formato	Estándar ISO	Justificación
Textuales	PDF/A-1, PDF/A-2 y PDF/A-3	ISO (19005-1: 2005), (19005-2: 2011), (19005-3: 2012)	Es un formato mundialmente utilizado, con su código abierto y bien aplicado para textos.

Tipo de Documento	Formato	Estándar ISO	Justificación
Textuales	ODF	ISO (ISO / IEC 26300: 2006)	Es un formato abierto desarrollado por Sun Microsystems, formato de documento basados en XML, independientes de la aplicación e independientes de la plataforma para documentos editables
Fotográfico	JPEG	ISO/IEC 10918 ISO/IEC 14495	JPEG es uno de los formatos más conocidos y utilizados en Internet siendo una extensión de fichero que posee características de pérdida en la calidad de imagen, pero su uso para visualización, acceso y diseminación en la web son recomendables.
Audiovisuales	MPEG-4	ISO/IEC 14496-14:2003	Es un formato de estándar de código abierto que admite la incorporación de metadatos descriptivos para apoyar el acceso en el futuro.
Audiovisuales	Ogg		Es el formato recomendado para audio y video, porque no es propietario, tiene soporte nativo para los formatos de audio y video Ogg. Incluye encabezados internos e identificadores.
Compactación de fichero	ZIP_PK	ISO/IEC 21320-1:2015	Es un estándar industrial de facto, desarrollado, mantenido y documentado abiertamente por PKWARE, combina la compresión de datos, la administración de documento y el cifrado de datos dentro de un formato de documento portátil. Proporciona metadatos ricos en un formato estándar o específico de la aplicación se puede incluir en un archivo ZIP. Esta práctica se ha incorporado en las especificaciones del paquete para EPUB, OOXML, ODF, etc.

Tabla 2. Estándares de formatos.

Fuente: Library of Congress, 2018 y Borba, Siebra et al (2015)

El Decreto 8.539/2015 en el artículo 17 establece que la definición de formatos de los documentos digitales deberá cumplir con las políticas y pautas establecidas en los Padrões de Interoperabilidade de Governo Eletrônico y ofrecer las mejores expectativas de garantía con respecto al acceso y la preservación. En el párrafo único dispone que se deben adoptar formatos interoperables, abiertos, independientes y tecnológicamente ampliamente utilizados.

Los formatos seleccionados como los más adecuados para preservación digital, deben estar bajo a un proceso de monitoreo continuo que puede ser automático o manual. Este monitoreo es necesario debido a rápida actualización tecnológica, incluye también las versiones de los formatos. Siempre que una versión o el formato se muestran obsoletos, un nuevo formato debe ser elegido o su versión actualizada y los objetos digitales almacenados en ellos deberán pasar por el proceso de conversión.

3.2.1.1 Herramientas para identificación de formatos

La comunidad de preservación digital ha creado herramientas especializadas para ayudar en el monitoreo automático de los formatos elegidos como los más adecuados a preservación digital a largo plazo. Actualmente existen herramientas especializadas para identificar, recopilar características y verificar algunos formatos de documentos específicos, la mayoría de ellas son de código abierto y están disponibles gratuitamente con licencia pública. Ni siempre los resultados de dichas herramientas son satisfactorios, por eso es necesario utilizar más de una herramienta y luego combinar los resultados. Según Shala y Shala (2016), las tres herramientas más populares para identificación y verificación de formatos de documentos son JHOVE, DROID y Exiftools. A continuación una breve explicación dada por las autoras para cada herramienta:

- **JHOVE (JSTOR/Harvard Object Validation Environment)**²³ es un marco de software extensible es utilizado para identificación, validación y caracterización de formatos de objetos digitales y también para recopilar metadatos. Fue desarrollada por JSTOR y las bibliotecas universitarias de Harvard, lanzada por primera vez en 2004, en los años siguientes algunos módulos fueron integrados, entre ellos constan los módulos para los formatos de documentos: TIFF, GIF, JPEG, PNG, JPEG-2000, AIFF, WAV, XML, HTML, UTF8, ASCII, PDF y un bytestream genérico. Además de estos, hay otros módulos que son desarrollados por los usuarios para sus necesidades. Los ejemplos típicos de este grupo son los módulos para MP3, ZIP y GZIP. Todos los que quieran agregar un formato de documento adicional a JHOVE puede desarrollar un nuevo módulo para el formato deseado e integrarlo a JHOVE.
- **DROID (Digital Record Object Identificación)**²⁴ es otra herramienta que se utiliza solo para la identificación de formato de archivo, es considerada por la comunidad de investigación de preservación digital como una de las herramientas más precisas para identificación de formato de documento. Fue desarrollada por los Archivos Nacionales de Reino Unido que también desarrollaron el FRF PRONOM²⁵. La comunicación de DROID y PRONOM se realiza a través de un FRF de archivo que se actualiza regularmente de PRONOM a DROID. Este FRF contiene las características necesarias que DROID utiliza para identificar el formato de documento. Por lo general, estas características consisten en extensiones de documentos, magic number²⁶, etc. DROID viene con una interfaz gráfica que permite que los usuarios que no posee ninguna habilidad de programación puedan usarla fácilmente. Además, también tiene la línea de comandos que está diseñada para usuarios avanzados y les permite utilizar técnicas avanzadas para automatizar procesos dentro de DROID. El análisis realizado con DROID, es en el formato XML.

²³ Disponible para descargar en: <http://jhove.openpreservation.org/>

²⁴ Disponible para descarga em: <http://www.nationalarchives.gov.uk/information-management/manage-information/preserving-digital-records/droid/>

²⁵ PRONOM File Registry Format es un registro técnico basado en la web para apoyar los servicios de preservación digital, desarrollado por los Archivos Nacionales del Reino Unido, fue el primero y registro de formato de archivo público operacional en el mundo.

²⁶ Un valor numérico o de texto utilizado para identificar un formato de documento o protocolo.

- **Exiftool (Exchangeable Image File Format Tool)** identifica el formato del documento y también está especializada para recuperar propiedades significativas de los formatos. Esta herramienta está diseñada principalmente para extraer y modificar metadatos del formato de documento EXIF (Exchangeable Image File Format) que es especializado para almacenar metadatos de la cámara digital y la salida del escáner. Además del formato de documento EXIF, Exiftool, también funciona con una gran variedad de formatos de documentos, incluso los más populares, que se utilizan comúnmente para almacenar información. En el contexto de la preservación digital, se utiliza para identificar y extraer propiedades significativas de formatos de documento diferentes. Proporciona numerosas opciones de formato de salida con HTML, XML y JSON delimitados por pestañas.

3.2.2 Metadatos

En general los metadatos pueden ser definidos como datos sobre los datos. La formación del término metadatos se da por el prefijo "Meta"²⁷ que viene del griego y significa "más allá de", "después de", sirve para hacer referencia a un objetivo a cumplir en el futuro, y por el sufijo "Datos", que el Conselho Nacional de Arquivos (2011), define como una representación de todo y cualquier elemento de contenido cognitivo que pueda ser comunicado, procesado e interpretado de forma manual o automática. Es decir, los metadatos facilitan la interpretación y la recuperación de un objeto digital, en el presente y principalmente en el futuro. Por eso que el Decreto 8.539/2015 establece que los elementos descriptivos deben asociarse a los documentos digitales que integren los procedimientos electrónicos, con el fin de apoyar su identificación, indexación, presunción de autenticidad y preservación.

Riley (2017), entiende metadatos como la información estructurada que describe, explica, localiza un recurso informativo y permite que éste sea fácil de usar, administrar y recuperar. Los metadatos deben acompañar todo el ciclo vital del documento, desde la creación hasta su destinación final, ellos admiten diferentes usos y funciones en los sistemas de información. Por eso los metadatos deben ser apropiados a los diferentes objetos digitales, a los diferentes usuarios y a los usos en el presente y en el futuro. Debido a estas diferencias Riley (2017), aborda cuatro tipos de metadatos:

1. **Metadatos descriptivos** – este tipo de metadatos ofrece información sobre el contenido de un recurso que ayuda a encontrarlo o entenderlo, ejemplo: autor, fechas, materia etc.
2. **Metadatos administrativos** – término general que se refiere a la información necesaria para administrar un recurso o que se relaciona con su creación. Entre los metadatos administrativos hay información sobre los documentos digitales necesarios para decodificarlos e interpretarlos, como el tipo de documento.
3. **Metadatos estructurales** – tienen la función de relacionar parte de un objeto digital a otro, que constituyen un mismo recurso informativo.
4. **Metadatos de preservación** – que admiten la gestión a largo plazo y la migración o emulación futura de los documentos digitales.

²⁷ Recuperado de: <http://etimologias.dechile.net/?meta>

Según el IBICT²⁸ (2013), los metadatos se agrupan en esquemas, que los organizan, normalizan y los describen, creando patrones. Estos esquemas permiten el mejor entendimiento de la finalidad de cada metadato y son generalmente mantenidos por institución u organización, permitiendo así el mejor uso y, principalmente, el intercambio de información entre iniciativas que utilizan el mismo esquema de metadatos.

En la actualidad existen varios esquemas de metadatos como: Dublin Core, EAD MODS, PREMIS, METS entre otros, cada cual con sus funciones y finalidades específicas. De acuerdo con las funciones y finalidades los esquemas de metadatos pueden ser clasificados en dos categorías: descriptivos y de preservación. Los descriptivos ayudan a identificar y recuperar los recursos. De manera similar, los de preservación soportan las acciones cuya finalidad es garantizar la utilización a largo plazo de los objetos digitales (Caplan, 2009).

Dublin Core, EAD y MODS son ejemplos de esquemas de metadatos descriptivos ya PREMIS y METS son ejemplos de esquemas de metadatos de preservación. Para Barbedo, Corujo y Sant'Ana (2011) los esquemas de metadatos descriptivos son importantes porque registran algunas informaciones que apoyan a los esquemas de metadatos de preservación, al respecto a la procedencia, autenticidad, integridad, contexto técnico, derechos de autor de los recursos almacenados en un repositorio de preservación digital.

3.2.2.1 Dublin Core

El estándar Dublin Core (DC)²⁹ tiene su origen en Chicago, en la 2ª Conferencia Internacional de la World Wide Web³⁰ en 1994, en una discusión sobre semántica y la Web enfocada en la dificultad del descubrimiento de recursos de información. En el evento, titulado "OCLC/NCSA Metadata Workshop", realizado en la ciudad Dublin, Ohio en marzo de 1995, donde más de 50 profesionales debatieron como un conjunto semántico sería útil para la búsqueda y la recuperación de recursos basados en la Web. Como resultado se definió un conjunto básico de elementos de metadatos llamado "Dublin Core", en razón del lugar del evento.

El estándar Dublin Core incluye dos niveles: Simple y Calificado, el simple consta de quince elementos; y el calificado incluye tres elementos adicionales (*Audience*, *Provenance* y *RightsHolder*), similar a un grupo de refinamientos de elementos (también llamados calificadores) que refinan la semántica de los elementos de manera que pueden ser útiles en la recuperación de los recursos de información. La semántica de Dublin Core ha sido establecida por un grupo internacional e interdisciplinario de profesionales de documentación, informáticos, codificación de texto, la comunidad de museos y otras

²⁸ Instituto Brasileiro de Informação em Ciência e Tecnologia es un órgano brasileño de información, unidad de investigación del Ministerio de Ciencia, Tecnología e Innovación (MCTI), realiza estudios en el campo de la ciencia de la información y temas relacionados.

²⁹ Elaborado a partir de las informaciones disponibles en la página web Dublin Core Metadata Initiative: <http://www.dublincore.org/about/history/>

³⁰ Conjunto total de recursos y servidores accesibles a través de Internet, utilizando el protocolo HTTP, que es una de las diversas maneras en que se puede tener acceso a la información disponible en esta red (el correo electrónico es otra de ellas).

áreas relacionadas. Los quince elementos descriptores de Dublin Core han sido formalmente respaldado en el estándar: ISO Standard 15836-2003.

El Dublin Core está entre los esquemas de metadatos más utilizados por las comunidades internacionales para recuperación de recursos en el dominio Web. Se representa por medio de varias sintaxis como, por ejemplo, codificado en lenguaje de marcado de hipertexto (HTML)³¹ o en el Marco de descripción de la descripción (RDF)³² mediante XML³³. Aunque el Dublin Core tiene una capacidad descriptiva para proporcionar el acceso y ser un exponente de interoperabilidad en la Web, este esquema es muy simple y menos eficaz para la preservación digital y la realización del papel de los metadatos administrativos y estructurales de documentar, por ejemplo, el histórico de estrategias de preservación adoptadas en el recurso digital y los componentes de hardware y software (Campos y Saramago, 2007).

Lamarca (2006), clasifica los elementos de Dublin Core en 3 grupos que señalan la clase o el entorno de la información que contienen:

- Elementos vinculados con el contenido del recurso:
 - **Title** (título)
 - **Subject** (materia)
 - **Description** (descripción)
 - **Source** (fuente)
 - **Lenguaje** (lenguaje)
 - **Relation** (relación)
 - **Coverage** (cobertura)
 -
- Elementos vinculados como la propiedad intelectual:
 - **Creator** (autor)
 - **Publisher** (editor) y, otras colaboraciones
 - **Contributor** (otros autores/colaboradores)
 - **Rights** (derechos)
- Elementos vinculados con la instanciación del recurso:
 - **Date** (fecha)
 - **Type** (tipo de recurso)
 - **Format** (formato)
 - **Identifier** (identificador)

³¹ HTML son las siglas de Hypertext Markup Language Lenguaje empleado para generar páginas de la Malla Mundial, que incluye marcadores para formatear los textos e insertar objetos e hiperenlaces.

³² El Marco de Descripción de Recursos (del inglés Resource Description Framework, RDF) es una familia de especificaciones de la World Wide Web Consortium (W3C) originalmente diseñado como un modelo de datos para metadatos. Ha llegado a ser usado como un método general para la descripción conceptual o modelado de la información que se implementa en los recursos web,1 utilizando una variedad de notaciones de sintaxis y formatos de serialización de datos.

³³ Extensible Markup Language (Lenguaje Extensible de Marcado). Versión simplificada del lenguaje SGML, que se espera que se convierta en una norma ampliamente utilizada para describir estructuras estándar de documentos, de modo que puedan ser comprendidas por la mayoría de los sistemas informáticos.

3.2.2.2 EAD (Encoded Archival Description)

EAD³⁴ es una codificación desarrollada y utilizada para la descripción de metadatos archivísticos basados en el lenguaje de marcado XML. El proyecto, iniciado en la Universidad de California en 1993, tuvo como base el estándar MARC (Machine-Readable Cataloging)³⁵, dando origen a la "EAD.DTD", que fue publicada, en su versión 1.0, en 1998, consolidada y actualizada en agosto de 2015 para la versión EAD3. La versión vigente actualiza e incorpora metadatos relacionados con los patrones de metadatos MARC, ISAD (G) y Dublín Core. EAD permite la descripción, estructuración e interoperabilidad de los metadatos archivísticos que, cuando asociados a XML, posibilitan la decodificación y la presentación de las informaciones referenciales de forma estructurada a los usuarios. EAD, es un estándar no propietario, actualmente se mantiene por el Network Development and MARC Standards Office de la Library of Congress de los Estados Unidos, en asociación con la Society of American Archivists.

EAD3 contiene 165 elementos descriptores, que están organizados por orden alfabético del nombre de cada elemento. En la figura 3 se muestra la información para cada elemento.

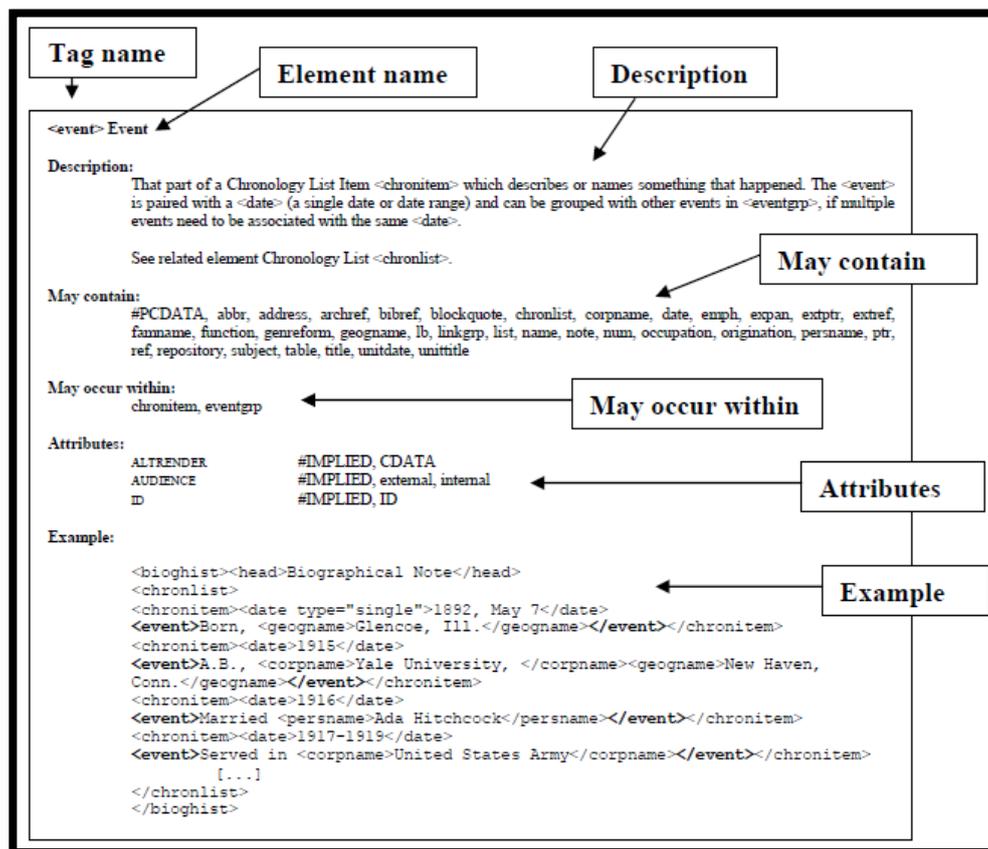


Figura 3. Elementos Descriptivos del EAD.
Fuente: Society of American Archivists. (2002)

³⁴ Elaborado a partir de las informaciones disponibles en la página web Society of American Archivists: <https://www2.archivists.org/standards/encoded-archival-description-ead>

³⁵ MARC 21 (Machine Readable Catalogin) es un conjunto de códigos y asignaciones de contenido definido para codificar registros legibles por máquina. Es un lenguaje estándar internacional para el intercambio de información bibliográfica.

3.2.2.3 MODS (Metadata Object Description Schema)

Metadata Object Description Schema (MODS)³⁶ es un esquema para un conjunto de elementos bibliográficos que se puede usar para una variedad de propósitos y, particularmente, para materiales de biblioteca. El esquema es mantenido por Network Development and MARC Standards Office of the Library of Congress está disponible desde 2002 (Library of Congress, 2018).

Expresado en XML, el MODS contiene un subconjunto de campos del MARC 21 y utiliza elementos textuales, en lugar de códigos numéricos, posibilitando una fácil comprensión. Actualmente en la versión 3.7, el estándar MODS presenta un conjunto de veinte elementos de metadatos descriptivos principales.

Para Formenton, Ferreira, Souza et al (2017), el esquema MODS proporciona información bibliográfica que apoya otros esquemas expresados en XML, como METS y PREMIS. Bajo el enfoque de la preservación digital, tres elementos MODS merecen atención: Información de Origen, Elemento Relacionado y Condición de Acceso. Estos elementos registran informaciones que son útiles para auxiliar a los metadatos de preservación en la corroboración de la procedencia, autenticidad e integridad de los objetos digitales, así como en el reconocimiento de los derechos asociados al recurso que pueden intervenir, sobre todo, en la realización de las acciones de preservación, de acceso y de utilización de sus respectivos contenidos.

3.2.2.4 PREMIS (PREservation Metadata: Implementation Strategies)

PREMIS³⁷ es la sigla de PREservation Metadata: Implementation Strategies. Este nombre tiene origen en el grupo de trabajo internacional de expertos en el área, que elaboró un informe denominado PREMIS Data Dictionary for Preservation Metadata que incluye un diccionario de datos e información sobre los metadatos de preservación. Es mantenido en un esquema de XML.

En verdad PREMIS es un Diccionario de datos que define un conjunto de unidades semánticas fundamentales que los repositorios deben entender para cumplir sus funciones de preservación. De acuerdo con Caplan (2009):

“Si se tienen presentes todos los metadatos que necesita una organización que gestiona un repositorio de preservación, PREMIS puede considerarse el subconjunto que se define en el centro, que no está relacionado con la recuperación y el acceso ni se propone definir los metadatos detallados de un formato específico. Solamente define los metadatos que se necesitan, por lo general, para llevar a cabo las funciones de preservación de todos los materiales” (p. 5).

PREMIS es utilizado principalmente para el diseño de los repositorios y para el intercambio de información almacenada en ellos. Es una guía sobre la información que

³⁶ Elaborado a partir de las informaciones disponibles en la página de la Library of Congress: <http://www.loc.gov/standards/mods/userguide/>

³⁷ Elaborado a partir de las informaciones disponibles en la página de la Library of Congress: https://www.loc.gov/standards/premis/UnderstandingPREMIS_espanol.pdf

los repositorios deben adquirir y registrar para su gestión. Caplan (2009), recomienda que la institución que desea implementar un repositorio digital, debe utilizar PREMIS como una lista de control, eso porque los softwares que soporten PREMIS serán más capaces de preservar los objetos digitales a largo plazo.

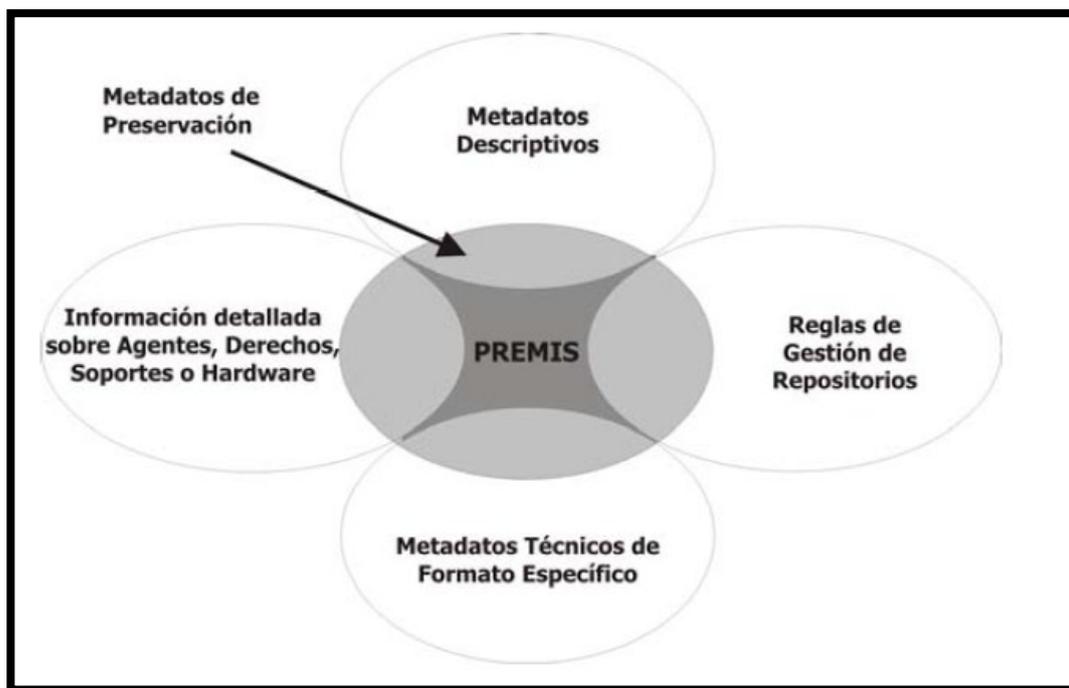


Figura 4. PREMIS como subconjunto de todos los metadatos de preservación.
Fuente: Caplan (2009)

3.2.2.5 METS (Metadata Encoding & Transmission Standard)

METS³⁸ (Metadata Encoding & Transmission Standard) es un esquema en lenguaje de marcado basado en XML que permite la codificación y el intercambio de los metadatos descriptivos, administrativos y estructurales relativos a objetos digitales. Se trata de un patrón de empaquetado que permite organizar en un único archivo comprimido, tanto los datos y los metadatos. Actualmente, el METS es mantenido por la Library of Congress.

Algunas implementaciones del estándar OAIS utilizan METS para estructurar los paquetes SIP, AIP y DIP. Además, algunos repositorios digitales utilizan METS para el intercambio de objetos.

La estructura de un paquete METS es definida por un modelo que detalla los elementos para la estructuración de un objeto digital. Básicamente, un paquete METS posee, obligatoriamente, un encabezado (header) y hasta seis secciones, que abarca los metadatos descriptivos, los metadatos administrativos, la lista de archivos del paquete, sus relaciones y comportamiento. El METS es neutro en relación con los formatos de metadatos encapsulados en el paquete digital. De esta forma, es necesario que se defina

³⁸ Elaborado a partir de las informaciones disponibles en la página de la Library of Congress: <http://www.loc.gov/standards/mets/>

cómo el paquete digital será estructurado, estableciendo los formatos de metadatos que serán utilizados. Esta configuración predeterminada se denomina "perfil de aplicación".

Como buena práctica, se debe, antes de crear un nuevo perfil para un determinado repositorio, investigar si los perfiles existentes atienden a los requisitos deseados (Conselho Nacional de Arquivos, 2014).

3.2.3 Migración

Granger (2000), define de una manera muy sencilla lo que es el proceso de migración, para el autor es la transferencia de datos de un hardware o software que está en peligro de volverse obsoleto a un más actual. La migración no cambia el formato del documento y su objetivo es garantizar el acceso a los objetos digitales frente a los cambios tecnológicos.

Según la norma UNE-EN ISO 13008:2012 (2013), hay algunas razones para conducir la migración digital y son las siguientes:

- a) Por la necesidad de migrar los documentos de una estructura a otra. Por ejemplo, los documentos que existen en varias bases de datos heredadas podrían reestructurarse en una nueva base de datos consolidada, ejemplo desde Oracle a SQL Server.
- b) Si la plataforma en la que los documentos fueron creados ha cambiado y es necesario migrar los documentos a la nueva plataforma. Por ejemplo, documento que tengan que ser trasladados de una plataforma de Microsoft Windows a una plataforma UNIX.
- c) Para introducir un nuevo sistema con mejores funcionalidades, por ejemplo una migración de documentos para apoyar el cambio de una presencia física a otra basada en la web o para migrar documentos desde un disco compartido a una aplicación de gestión de documentos electrónicos (EDRMS).

Durante el proceso de migración todo el contenido del documento y los metadatos asociados a la aplicación de origen deben preservarse hasta que el proceso se termine y la integridad y fiabilidad del formato de destino haya sido controlada y asegurada (UNE-EN ISO 15489-1, 2016, p. 26).

La estrategia de migración, así como la de conversión y emulación, presenta riesgos a la integridad y funcionalidad de los objetos digitales, por lo que debe realizarse de manera planificada y sistemática.

3.2.4 Conversión

En la actualidad la información que las instituciones necesitan para llevar a cabo sus actividades están almacenadas en una gran variedad de formatos: abiertos, propietarios, estables, inestables etc., independiente del tipo de formato, todos están amenazados por la obsolescencia tecnológica. Para evitar la pérdida de información debido a los cambios tecnológicos, es necesario utilizar la estrategia de conversión, que según la norma UNE-EN ISO 13008:2012 (2013), es el proceso de transformación de los documentos de un formato a otro, manteniendo las características del documento.

En general, hay dos razones para convertir los documentos de un formato a otro. La primera es la más común, que es debido a la obsolescencia del formato, es decir documentos que están almacenados en un formato obsoleto, y necesitan ser convertidos a otro más actual, para garantizar el acceso. La segunda está relacionada con la estrategia de Estándares de formato (3.2.1), que es cuando una institución decide utilizar un número limitado de formatos y convertir todos los objetos digitales a los formatos elegidos, para controlar la complejidad y el coste de mantener varios formatos. A partir de la segunda razón es posible concluir que la estrategia de conversión es el complemento de la estrategia estándares de formatos y debe ser ejecutada de manera continua, siempre que en el monitoreo de formatos señalar el riesgo de obsolescencia o inestabilidad de un formato.

El proceso de conversión de formato no es para todo documento producido y recibido en la institución, el calendario de conservación y eliminación es el documento esencial para decidir cuales documentos necesitan ser convertido a otro formato. Si la institución aún no posee este documento debe llevar en cuenta el contexto del negocio y jurídico, así como la importancia del documento para la misma. Eso también se aplica a la cuestión de la eliminación de los documentos originales que fueron convertidos a otros formatos.

Hay una gran cantidad de herramientas³⁹ disponibles para convertir documentos de un formato a otro. Algunas son propietarias, otras son gratuitas y otras son de código abierto. Sin embargo, la cobertura de los formatos puede ser irregular. Para formatos ampliamente utilizados, como procesador de textos, imágenes, puede haber muchas opciones, pero para formatos de nicho o más antiguos, sus opciones pueden ser más limitadas. También hay empresas⁴⁰ que ofrecen servicio de conversión de formato, utilizando sus propias herramientas para realizar la conversión y gestionar del proceso (The National Archives of UK, 2010).

Así como en la migración la estrategia de conversión en sí conlleva riesgos y debe planificarse y ejecutarse cuidadosamente, para garantizar que todo el contenido del documento y los metadatos asociados al mismo serán convertidos al nuevo formato.

3.2.5 Emulación

Diferente de la migración y de la conversión, que se centran en el objeto digital, la emulación se centra en el entorno tecnológico en que se creó el objeto digital. Según la National Library of Australia (2003), la emulación consiste en recrear un software para funcionar con las características de otro. En el contexto de la preservación digital a largo plazo, esta es una manera de posibilitar que las futuras tecnologías se comporten como el entorno original de un objeto digital preservado, de un modo que se presente en su forma original, a partir del flujo de datos originales.

³⁹ Freemake Audio Converter; Doxillion; MediaHuman Audio Converter; Hamster Free Audio Converter; VSDC Free Audio Converter; Pixillion

⁴⁰ FlatWorld Solutions: <https://www.flatworldsolutions.com/data-management/file-format-conversion.php>

OM Data Center India: <https://www.omdataentryindia.com/pdf-conversion-services.html>

Data Entry Outsourced: <https://www.dataentryoutsourced.com/conversion-services/>

Según Van der Hoeven, Lohman y Verdegem (2008), la elección de la emulación como estrategia de preservación es discutible, a pesar de que sus beneficios son reconocidos. No obstante, de acuerdo con estos autores la emulación nunca se ha desarrollado y probado dentro de un entorno operativo de documento digital. La Koninklijke Bibliotheek⁴¹ y el Nationaal Archief⁴² de los Países Bajos creen que la emulación proporciona una buena solución para el acceso a largo plazo a objetos digitales sin afectar su autenticidad e integridad. Estas instituciones están de acuerdo que esta estrategia debe desarrollarse y probarse primero, antes de poder evaluar los potenciales y las limitaciones.

En la actualidad hay algunos emuladores como: MS Virtual PC⁴³, QEMU⁴⁴, Bochs⁴⁵ y las técnicas de virtualización utilizadas en VMware⁴⁶ se han vuelto muy avanzadas y capaces de ejecutar sistemas operativos completos como Microsoft Windows, MacOS de Apple o GNU/Linux, y son compatibles con una amplia gama de dispositivos periféricos. Sin embargo, ninguna de estas soluciones ha sido diseñada específicamente para la preservación digital. El software de emulación o virtualización que funciona hoy no ofrece ninguna garantía de que operará en diferentes condiciones en el futuro. (Van der Hoeven, Lohman y Verdegem, 2008).

La necesidad de un emulador específico para la preservación digital es evidente, pero el desarrollo de una solución de emulación de este tipo es demasiado complejo y costoso.

3.3 El Modelo OAIS (Open Archival Information System)

El modelo de referencia OAIS (Open Archival Information System)⁴⁷ en español Sistema abierto de información de archivo, fue desarrollado en 2002 por el Consultative Committee for Space Data Systems (CCSDS) of the National Aeronautics and Space Administration (NASA), es un modelo de referencia genérico sobre cómo debe ser un sistema de información de archivo para la preservación digital a largo plazo, sin especifica la manera de ponerlo en marcha. En 2003 fue reconocido como norma ISO (14721) y actualizada en 2012. Actualmente es el modelo de referencia más aceptado para la concepción de la preservación digital, es aplicable en la implantación de varios tipos de repositorios digitales, pues define sus componentes esenciales (Ball 2006, Theodoridou, Tzitzikas, Doerr et al 2010)

Cruz-Mundet y Díez (2016), explican que la palabra “abierto” en el nombre del modelo es engañoso, porque se refiere al hecho de que el modelo se desarrolló en foros abiertos, con el fin de incluir información de una amplia gama de expertos, y no que el acceso al archivo es abierto.

El modelo funcional de OAIS recomienda un entorno formado por un productor (quien provee la información, puede ser personas o un sistema), consumidor (quien busca y

⁴¹ Biblioteca Nacional de los Países Bajos.

⁴² Archivo Nacional de los Países Bajos.

⁴³ Microsoft Virtual PC: <https://www.microsoft.com/pt-br/download/details.aspx?id=3702>

⁴⁴ QEMU: <https://www.qemu.org/>

⁴⁵ Bochs: <https://www.qemu.org/>

⁴⁶ VMWare: <https://www.vmware.com/products/workstation-pro.html>

⁴⁷ En Brasil, el modelo OAIS fue traducido por la ABNT y publicado bajo la forma de la norma ABNT NBR 15472: 2007, con el título "Sistema Abierto de Archivo de Información - SAAI"

recupera información) y administración (quien establece el conjunto de políticas del OAIS) y seis entidades funcionales: recepción, almacenamiento, gestión de datos, administración del sistema, planificación de preservación y acceso, Hunter (2006), describe las seis entidades de la siguiente manera:

- **Ingreso (Ingest):** ofrece los servicios y funciones para recibir paquetes de información de envío (SIP) de los productores y preparar los contenidos para su almacenamiento y gestión dentro de los archivos.
- **Almacenamiento de Archivo (Archival Storage):** proporciona los servicios y funciones para el almacenamiento, mantenimiento y recuperación de paquetes de información de archivo (PIA).
- **Gestión de Datos (Data Management):** proporciona los servicios y funciones para ingresar, mantener y acceder, no solo a la información descriptiva sino, a los datos administrativos utilizados para administrar los archivos.
- **Administración (Administration):** proporciona los servicios y funciones para el funcionamiento general del sistema de archivos. También proporciona funciones de ingeniería del sistema para controlar y mejorar las operaciones de los archivos, y para inventariar, elaborar informes y migrar/actualizar el contenido de los archivos.
- **Planificación de la Preservación (Preservation Planning):** supervisa el entorno del OAIS y proporciona recomendaciones para garantizar que la información almacenada en el OAIS siga siendo accesible para la comunidad de usuarios designados a largo plazo, incluso si el entorno informático original se vuelve obsoleto.
- **Acceso (Access):** ayuda a los consumidores a determinar la existencia, descripción, ubicación y disponibilidad de la información almacenada en el OAIS, y les permite solicitar y recibir productos de información.

El modelo OAIS reconoce tres variantes de paquetes de información de acuerdo con Rodríguez Reséndiz (2016), este concepto define la estructura de información que circula dentro y fuera del sistema de archivo. El paquete de información es el objeto digital, enfoque de la preservación, juntamente con los metadatos necesarios para proveer la identificación y preservación a largo plazo, así como facilitar el acceso. De acuerdo con la norma UNE-ES ISO 14271:2012, (2015), los paquetes de información definidos por el modelo son los siguientes:

1. **Submission Information Package (SIP) / Paquete de Información de Transferencia (PIT)** paquete de información que el productor entrega a un OAIS para usarlo en la construcción o actualización de uno o más PIA y/o los metadatos asociados.
2. **Archival Information Package (AIP) / Paquete de Información de Archivo (PIA)** paquete de información, almacenada en un OAIS, y que consta de la información de contenidos y de los metadatos asociados.
3. **Dissemination Information Package (DIP) / Paquete de Información de Consulta (PIC)** es un paquete de información que proviene de uno o más PIA, y enviado al usuario que ha solicitado al OAIS.

En la figura 5 observase la interactuación entre los tres actores y las seis entidades:

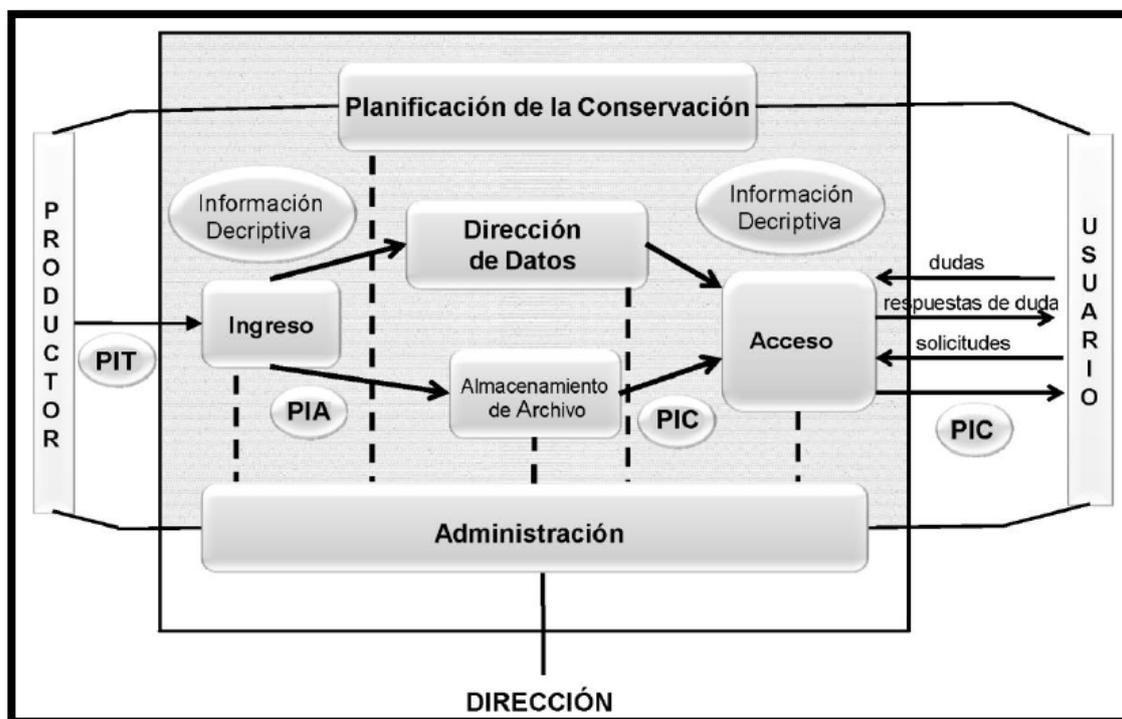


Figura 5. Entidades Funcionales de OAIS
Fuente: UNE-EN ISO 14721:2012 (2015)

El OAIS no prescribe como implementar un sistema de archivo con enfoque en la preservación digital, pero la norma UNE-EN ISO 14721:2012 (2015), establece algunas responsabilidades obligatorias que toda institución debe cumplir para implantar un sistema de archivo que son:

- negociar y aceptar información adecuada de quien la produce;
- obtener un control suficiente de la información para garantizar la preservación a largo plazo;
- determinar qué grupos de consumidores deben considerarse la comunidad designada;
- asegurarse de que la información proporcionada por OAIS sea comprensible por la comunidad sin la necesidad de recurso especiales como la asistencia de expertos o de los productores de la información;
- seguir las políticas y procedimientos documentados;
- poner la información a disposición de la comunidad designada.

A pesar de ser considerado un modelo importante para el tema de la preservación digital y ser un caso de éxito con amplia aceptación, el OAIS sufre críticas de algunos autores como Theodoridou, Tzitzikas, Doerr et al (2010), que expone que el OAIS no recomienda ningún modelo conceptual u ontología específico. Para los autores lo que se llama "Modelo de información OAIS" es muy simplista y no se puede considerar como un modelo conceptual, se parece más a un diagrama de requisitos. Ya la crítica de Cruz-Mundet y Díez (2016), es que OAIS es un modelo elaborado por y pensado para profesionales de base científica, ingenieros, físicos, informáticos etc. los autores explican que las críticas se refiere a la utilización del modelo y no a su valor en sí mismo.

4. REPOSITARIOS DIGITALES

Los repositorios digitales tienen su origen en 2001 en la Budapest Open Access Initiative⁴⁸, una reunión realizada por el Open Society Institute (OSI) con el propósito de acelerar el esfuerzo internacional para lograr el acceso abierto (Open Access⁴⁹) a los artículos de investigación en todos los campos académicos.

La procedencia de la palabra repositorio viene del latín *repositorium*⁵⁰ que significa armario, depósito, el Diccionario de la Lengua Española⁵¹ lo define como “lugar donde se guarda algo”. Uno de los primeros autores que conceptuó repositorios digitales fue Crow (2002), definiéndolos como colecciones digitales que capturan y preservan la producción intelectual de las comunidades universitarias.

Según Córdoba (2011), el arXiv⁵² fue el repositorio digital pionero, creado en el campo de la Física, fue creado en el Laboratorio Nacional de los Álamos (California), en 1991 se transfirió a la Universidad de Cornell. Su éxito en la iniciativa de depositar la producción científica estimuló la ampliación de las asignaturas que cubre actualmente, además se desencadenó una serie de acciones relacionadas que se difundieron alrededor del mundo.

Para Polanco-Cortés (2014) repositorios digitales son medios de almacenamiento para preservar, difundir y facilitar el acceso a los objetos digitales que almenan. Para Tramullas y Garrido (2006) en la actualidad los repositorios digitales comienzan a considerarse como estrategia organizacional. Eso porque según Texier (2013) el acceso a largo plazo a los objetos digitales depende directamente donde éstos están almacenados. El crecimiento de implantación de repositorios digitales en el mundo, de 87 en el año de 2005 para 4.143 en el año de 2019⁵³, comprueba la afirmación de Tramullas y Garrido (2006), que los repositorios ya son considerados como estrategia organizacional:

⁴⁸ Elaborado a partir de las informaciones disponible en la página web Budapest Open Access Initiative: <http://www.budapestopenaccessinitiative.org/>

⁴⁹ El acceso abierto es un amplio movimiento académico internacional que busca acceso en línea gratuito y abierto a información académica, como publicaciones y datos.

⁵⁰ Recuperado de: http://www.colegiosanjose.net/latin/diccionarios/diccionario_latin.pdf

⁵¹ Recuperado de: <https://dle.rae.es/?id=W3mzJyE>

⁵² Repositorio arXiv <https://arxiv.org/>

⁵³ Información recuperada de: https://v2.sherpa.ac.uk/view/repository_visualisations/1.html

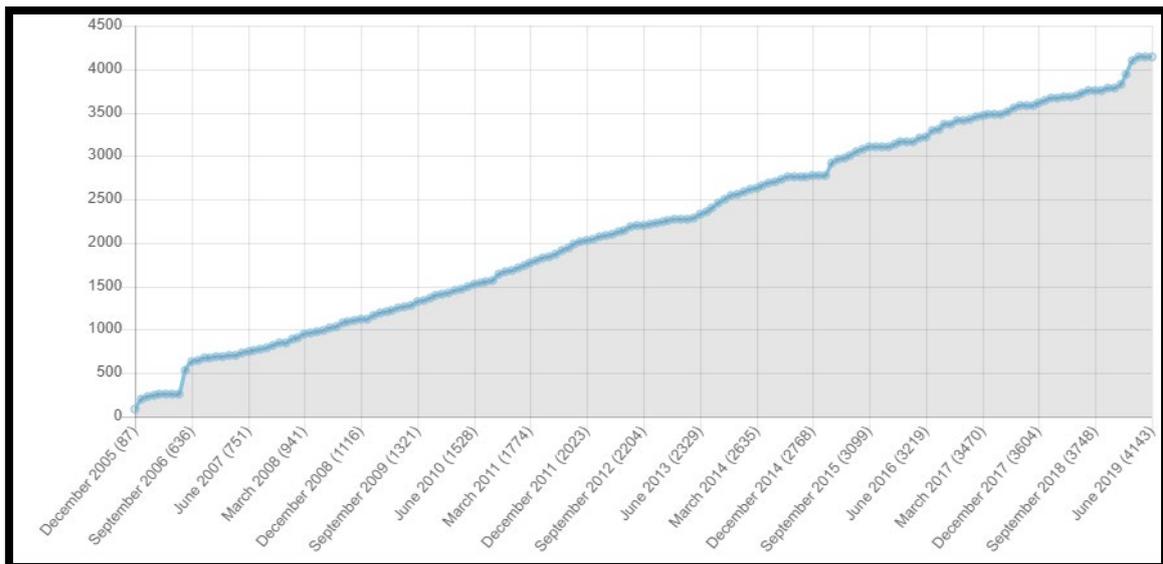


Figura 6. Crecimiento de Repositorios de Acceso Abierto en el mundo.
Fuente: OpenDOAR

Según Thibodeau (2007) existen diferentes contextos y variaciones sustanciales sobre cual tipo de información debe ser preservada en un repositorio, así como lo que eso significa para la institución designada. Para el autor hay categorías principales de repositorios; como bibliotecas y archivos gubernamentales. A continuación la descripción del autor para dichas categorías:

- **Bibliotecas:** las bibliotecas son organizaciones cuyo propósito principal es entregar productos de información relevantes, generalmente libros y publicaciones periódicas, a sus usuarios. También hay, las bibliotecas que posee materiales especializados que abarcan siglos, los cuáles requieren preservación.
- **Archivos gubernamentales:** la democracia es un contexto que diferencia los archivos de un gobierno. La comunidad de usuarios está formada por cualquier persona que tenga interés en los registros del gobierno o la información que contienen, por cualquier motivo. En las naciones democráticas, probablemente el propósito más importante de los archivos de un gobierno es permitir a los ciudadanos determinar lo que el gobierno ha hecho al revisar los registros, que son los instrumentos y subproductos de las actividades del gobierno. Dichas instituciones deben enfatizar las funciones de suministro y de archivo, para garantizar que se adquieren registros adecuados del gobierno y que se conserven de manera auténtica, incluso si las versiones auténticas de los registros no son óptimas para las necesidades de los usuarios.

A partir de las definiciones del autor, es posible concluir que los repositorios digitales de una biblioteca tiene el objetivo de: dar acceso y preservar los objetos digitales que están bajo la custodia de dicha institución y que los repositorios de archivos gubernamentales: gestionan y preservan los documentos administrativos resultantes de sus actividades, para cumplimiento de los requisitos legales. Wheatley (2004) resumió en cuatro los objetivos de los repositorios digitales:

1. La información se mantienen en el repositorio sin sufrir daños, pérdida o alteración maliciosa;
2. La información se pueden encontrar, extraer y entregar a un usuario;
3. La información pueden ser interpretadas y comprendidas por el usuario; y
4. En el objetivo tres se puede lograr a largo plazo.

A partir de lo que fue abordado por los autores es posible inferir que los Repositorios Institucionales (RI), están más relacionados con las bibliotecas digitales y los Repositorios de Documentos Administrativos (RDA) con los archivos.

4.1 Repositorios Institucionales (RI)

Por tener su origen en el campo académico, y por las definiciones de algunos autores, a veces los RI son vistos solamente como un conjunto de servicios para la gestión y el acceso a los materiales de investigación en formato digital producidos por una universidad, es decir una colección digital del producto de la investigación llevada a cabo por esa comunidad. Los repositorios institucionales se están convirtiendo en herramientas fundamentales para la comunicación académica en la era digital (Lynch, 2003).

Vicente y Gozzer (2011) exponen que partir de un punto de vista funcional es posible decir que el concepto de RI se amplió y en la actualidad puede ser definido como una aplicación o sistema de almacenamiento de contenido digital disponible para acceso. En este sentido funcional y también práctico, Leija (2017) define RI como un sistema que almacena objetos digitales como: documentos, audio, videos, fotografías, etc. Este puede albergarse vía un servidor web o de un servidor local. Puede ser de acceso abierto o no, por medio de un software que gestiona los objetos digitales.

De acuerdo con las definiciones de los autores mencionados, es posible definir RI como un medio que amplía y refleja las actividades de una institución, sea una universidad, una biblioteca, un museo etc. Por lo tanto una Biblioteca Digital puede ser considerada un RI, pues amplía la actividad de acceso de las bibliotecas, para la European Commission (2013) bibliotecas digitales son colecciones organizadas y disponible al público que pueden consistir en todo tipo de materiales analógicos que se han digitalizado (libros, audiovisuales, material multimedia, fotografías, documentos de archivo, etc.) y material originalmente producido en formato digital.

4.2 Repositorios de Documentos Administrativos (RDA)

Cuando Tramullas y Garrido (2006) enfatizan que los repositorios digitales están se convirtiendo progresivamente en el medio de almacenamiento en los que varias instituciones gestionan los resultados de sus actividades y el Conselho Nacional de Arquivos (2014) expone que los documentos archivísticos registran y apoyan las actividades de un órgano o institución como evidencia de sus actividades, así como fuente de investigación para garantizar los derechos de los ciudadanos y que dichos documentos, dependen de un buen sistema informatizado que apoye la gestión, el acceso y la preservación, para garantizar el mantenimiento de la integridad y la autenticidad, es

posible nombrar este sistema informatizado de Repositorio de Documentos Administrativo (RDA). El enfoque de este tipo de repositorio es soportar la administración con la gestión de los documentos administrativos en las fases corriente, intermediaria y principalmente en la fase permanente. Este tipo de repositorio también tiene como objetivo el acceso, pero no es el acceso abierto y si acceso público (bajo demanda).

Para Texier (2013) en los RDA se puede encontrar las funcionalidades de los Sistemas de Gestión Informatizada de Documentos (SGID) que son aplicaciones que facilitan la producción, edición y difusión de documentos digitales, que permiten a los usuarios acceder, editar y/o borrar esos documentos, es decir, manejar los documentos en línea. A continuación se enumeran algunas diferencias entre los RDA y los SGID:

- **Tipos de documentos:**

RDA: documentos que ya están producidos.

SGID: documentos en producción y trámite.

- **Modificación de documentos:**

RDA: no está permitido, solo admite incluir la versión final.

SGID: está permitido, es decir, es posible mantener varias versiones de un mismo documento.

- **Versiones de documentos:**

RDA: solo se mantiene la versión final y no existirán versiones de los documentos.

SGID: pueden mantenerse y gestionar diferentes versiones de los documentos.

- **Eliminación de documentos:**

RDA: no está permitido, a menos que se quiera excluir el objeto digital completo del repositorio.

SGID: permitido, pero se conserva en el histórico.

- **Sistema de clasificación y almacenamiento:**

RDA: obligatorio, porque es necesario hacer una clasificación de cada uno de ellos.

SGID: opcional, ya que frecuentemente depende del usuario.

- **Políticas de preservación:**

RDA: los sistemas garantizan una preservación a largo plazo.

SGID: opcional.

- **Objetivo:**

RDA: ofrecer un repositorio adecuado para la preservación y el acceso a los objetos digitales a largo plazo.

SGID: permitir la gestión de documentos en las rutinas de trabajo.

4.3 Interoperabilidad

Uno de los valores de los repositorios reside en su potencial para conectarse entre sí, con el fin de construir una red de repositorios que facilite el acceso unificado, de tal manera que se abran nuevas formas de trabajar con la información (Ferrerías Fernández, 2013). Lo que hace esta conexión posible es la interoperabilidad, que según Governo Brasileiro (2017) se refiere a la capacidad de varios sistemas y organizaciones para trabajar juntos (interactuar) para garantizar que las personas, las organizaciones y los sistemas informáticos interactúen para intercambio de información de manera eficaz y eficiente.

El informe de la COAR⁵⁴ (2011) expone que a nivel del sistema, la interoperabilidad se produce cuando los repositorios están configurados de una manera que permitan a los objetos digitales ser transferidos de uno a otro. Para Bhat (2010) la interoperabilidad asegura que cualquier persona en cualquier parte del mundo puede encontrar documentos en repositorios ubicados en cualquier lugar.

Según la Comisión Europea (2017) en el documento titulado de: Marco Europeo de Interoperabilidad - MEI⁵⁵ hay cuatro niveles de interoperabilidad: jurídica, organizativa, semántica y técnica, a continuación la explicación dada a cada una por la Comisión:

- **Interoperabilidad jurídica:** consiste en asegurar que las organizaciones que operan con acuerdo a diferentes marcos jurídicos, políticas y estrategias pueden trabajar juntas. Ello podría exigir que la legislación no bloquee el establecimiento de servicios públicos europeos dentro de los Estados miembros y entre estos y que existan acuerdos claros sobre cómo plantear las diferencias en la legislación a través de fronteras, incluso la opción de adopción de una nueva legislación.
- **Interoperabilidad organizativa:** se refiere a la manera en que las administraciones públicas adecuan sus procedimientos institucionales, responsabilidades y expectativas para alcanzar los objetivos adoptados de común acuerdo y mutuamente beneficiosas. En la práctica, la interoperabilidad organizativa implica la documentación y la integración o adecuación de los procedimientos institucionales y la información intercambiada. La interoperabilidad organizativa promete satisfacer los requisitos de la comunidad de usuarios permitiendo que los servicios estén disponibles, con facilidad para identificarlos y accederlos y con enfoque en el usuario.
- **Interoperabilidad semántica:** asegura que el formato y el significado exacto de la información intercambiada sea comprensible y conserve en todos los intercambios entre las partes, es decir, entender la información que se transmite. En el contexto del MEI, la interoperabilidad semántica abarca los aspectos semánticos y sintácticos:

⁵⁴ Confederation of Open Access Repositories - COAR es una asociación internacional con más de 140 miembros y socios de todo el mundo que representan bibliotecas, universidades, instituciones de investigación, financiadores gubernamentales y otros. COAR reúne a la comunidad de repositorios y las principales redes de repositorios para construir capacidad, alinear políticas y prácticas, y actuar como una voz global para la comunidad de repositorios.

⁵⁵ El Marco Europeo de Interoperabilidad (EIF por sus siglas en inglés) es un documento que, en la Unión Europea, define una serie de directrices y recomendaciones para los servicios de administración electrónica que garanticen la interoperabilidad de los sistemas.

- **semántico** se refiere al significado de los elementos de datos y la relación entre ellos, incluye la creación de vocabularios controlados y esquemas para describir los intercambios de datos y asegura que todas las partes que se comunican comprendan de la misma manera la información;
- **sintáctico** se refiere a la descripción del formato exacto de la información que se va a intercambiar en términos de gramática y formato.
- **Interoperabilidad técnica:** La interoperabilidad técnica abarca las aplicaciones e infraestructuras que conectan sistemas y servicios. Incluye elementos tales como especificaciones de interfaz, servicios de interconexión e integración de datos, presentación e intercambio de datos y protocolos de comunicación.

Los niveles de interoperabilidad expuesto en el MEI se percibe que los cuatro niveles son aplicables a cualquier tipo de repositorio digital, pero los dos primeros más están relacionados a los RDA.

4.3.1 Protocolos de interoperabilidad

Como explicado en el apartado anterior, existe cuatro tipos de interoperabilidad, la interoperabilidad técnica se produce cuando los repositorios están configurados por aplicaciones o sistemas de una manera que permitan que los objetos digitales puedan ser transferidos de un repositorio a otro. Los protocolos de interoperabilidad son herramientas que posibilitan esta transferencia, en la actualidad hay diversos, que pueden ser clasificados en varias categorías, dependiendo del objetivo, Hakimjavadi y Masrek (2013) describe los más utilizados de la siguiente manera:

- **OAI-PMH (Open Archives Initiative Protocol for Metadata Harvesting)** es un protocolo específico de repositorio, que se basa en el protocolo HTTP y se presenta en XML. Todos los formatos de descripción de metadatos basados en XML son compatibles con el protocolo OAI. Sin embargo, el requisito mínimo para que un repositorio sea compatible con OAI es exponer los metadatos en un formato Dublin Core no calificado.
- **OAI-ORE (Open Archives Initiative Object Reuse and Exchange)** el protocolo OAI-ORE incluye los métodos para exponer e importar el contenido, los metadatos, las propiedades de autenticación, el historial de versiones, etc., como objetos de datos entre repositorios compatibles con OAI-ORE, está representado en RDF, que es el lenguaje estándar para proporcionar datos para la web semántica.
- **ATOM** en el contexto de los protocolos de interoperabilidad se refiere a dos protocolos relacionados: el formato de sindicación ATOM (ASF), que es un protocolo de sindicación basado en XML como RSS, que se utiliza como marco de alimentación web, y el Protocolo de publicación ATOM (APP), que se refiere a un protocolo basado en HTTP diseñado para crear, editar y publicar el contenido de sitios web que admiten este protocolo. ATOM, de hecho, es una tecnología paralela con RSS introducida para abordar algunas insuficiencias de RSS (Sayre, 2005 citado por Hakimjavadi y Masrek, 2013). Diferente del modelo de

extensibilidad RSS, el modelo de extensibilidad de ATOM está diseñado de manera que proporciona interoperabilidad entre las implementaciones extendidas y las otras implementaciones que no reconocen estas extensiones. En términos de interoperabilidad ETD, probablemente la mayor ventaja de ATOM es su conjunto de elementos de metadatos opcionales, como Autor, Colaborador, Generador, Categoría, Derecho, etc., que están predefinidos en el espacio de nombres de ATOM (Ross y James, 2008 citado por Hakimjavadi y Masrek, 2013).

4.4 Software para repositorios digitales

En la actualidad existen varios software para crear repositorios digitales, entre esta variedad, existe software propietario y de código abierto. El software propietario tiene la ventaja de disponer de la empresa propietaria para hacer el sistema funcionar; pero estos sistemas suponen gastos continuos y el peligro de quedar preso en un sistema propietario de lo cual será difícil escapar. Las opciones de software de código abierto tienen a su favor la aceptación a estándares y marcos abiertos que permiten la extracción o migración de contenido en el caso de actualizaciones futuras, es más duradero y más económico, pues está libre de la dependencia de licencias costosas. Su desventaja está en que, a pesar de disponer de la ayuda de las comunidades de soporte, la responsabilidad final del mantenimiento recae en la institución. Sin embargo hay empresas que ofrecen soporte técnico para las soluciones de código abierto.

Para implementar el software, es necesario tener un servidor web donde el repositorio será instalado para gestionar los objetos digitales. Gemetto (2017) orienta que si posible, lo mejor es que el servidor web que hospeda el repositorio esté en el país que reside la institución o el más cercano posible de la mayoría de los usuarios, para que el acceso sea más rápido. Para el autor contar con los servidores en el país favorece también la seguridad jurídica, dado que garantiza que la información almacenada allí quedará sujeta a la legislación nacional.

Según IASA (s.f) la mayoría de los softwares para repositorios permiten las funciones para el acceso, la gestión de objetos digitales y ciertos aspectos de la transferencia de contenidos. Un software de repositorio bien diseñado debe incluir un cierto número de las funciones identificadas en OAIS.

En la página web OpenDoar, es posible verificar los softwares de repositorios digitales más instalados en el mundo, lo que se puede observar en las figuras 7 y 8:

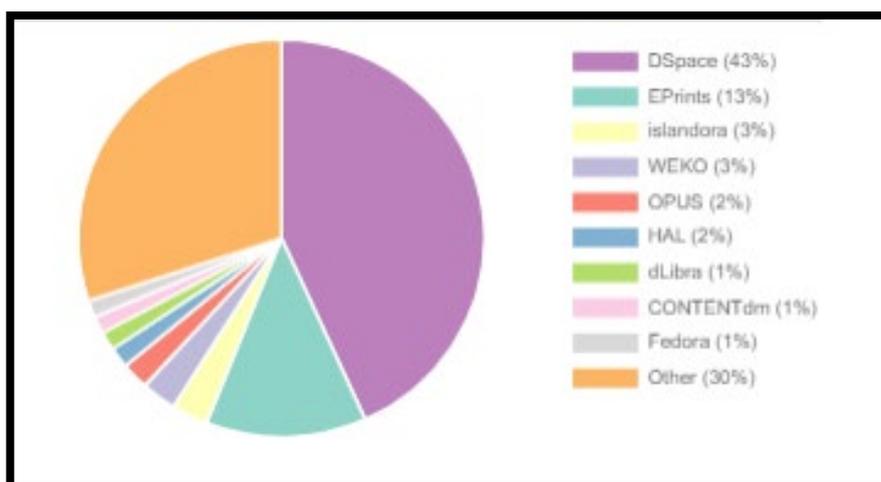


Figura 7. Los softwares más instalados en el mundo.
Fuente OpenDOAR.

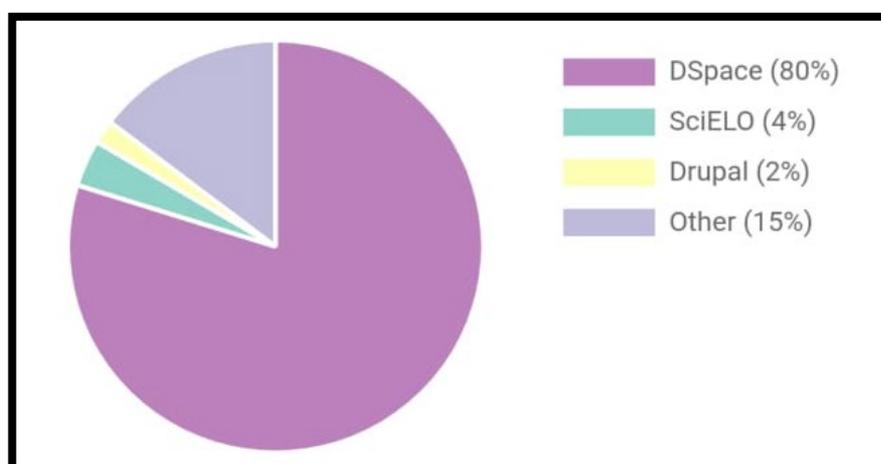


Figura 8. Los softwares más instalados en Brasil.
Fuente OpenDOAR.

4.4.1 Archivematica

Archivematica⁵⁶ es un software de preservación digital gratuito y de código abierto que está diseñado para mantener el acceso a largo plazo a la memoria digital. Archivematica está empaquetado con el sistema de gestión de contenido basado en web AtoM⁵⁷ para acceder a sus objetos digitales. Es gestionado por la empresa

Utiliza un patrón de diseño de micro servicios para proporcionar un conjunto integrado de herramientas de software que permite a los usuarios procesar objetos digitales desde el ingreso hasta el acceso, en conformidad con el modelo funcional ISO-OAIS. Los usuarios pueden monitorear y controlar los micros servicios a través de un panel de control basado en la web. Archivematica usa METS, PREMIS, Dublin Core y otros esquemas de

⁵⁶ Elaborado a partir de las informaciones disponibles en la página web del software Archivematica: https://wiki.archivematica.org/Main_Page

⁵⁷ AtoM es una abreviatura de Access to Memory. Funcionando en ambiente WEB, es una aplicación de código abierto destinada a la descripción normalizada en archivos permanentes permitiendo un acceso multilingüe en una organización con múltiples repositorios integrados.

metadatos, para proporcionar paquetes de archivo (AIP) confiables, auténticos, e interoperables para almacenamiento en repositorio digital.

El código de Archivemática se publica bajo GNU Affero General Public License (AGPL 3.0)⁵⁸, lo que permite estudiarlo, modificarlo, mejorarlo y distribuirlo. La documentación almacenada en la plataforma se publica bajo Creative Commons Share-alike license.

Archivemática implementa Format Policy Registry (FPR) o en español Registro de políticas de formato, esta política es basada en un análisis de las características significativas de los formatos de documento. El FPR ofrece un marco editable y flexible para la identificación de formatos, la extracción de paquetes, la transcripción y la normalización para la preservación y el acceso.

Archivemática ofrece muchos flujos de trabajo de ingreso: importación de metadatos y documentación de envío, admisión de paquete comprimido y descomprimido, procesamiento digital de imágenes, disposición SIP, normalización manual y administración de conjuntos de datos, compatible con cientos de formatos.

Dependiendo de la compatibilidad de las licencias, Archivemática puede integrarse con otras herramientas como: AtoM, DSpace, CONTENTdm, Islandora, LOCKSS, DuraCloud, OpenStack y el Archivists' Toolkit.

4.4.2 DSpace

DSpace⁵⁹ fue desarrollado por MIT Libraries⁶⁰ y Hewlett-Packard (HP) Labs⁶¹, lanzado en 2002, la plataforma es guiada por una comunidad global de responsables, desarrolladores, administradores de repositorios y otras partes interesadas que contribuyen a la gobernanza del proyecto. En 2009 DSpace se convirtió en el proyecto DuraSpace⁶², con la fusión de las organizaciones Fedora Commons y DSpace.

DSpace es un software desarrollado para creación de repositorios digitales con el fin de proporcionar un conjunto de herramientas para la gestión y preservación de contenidos digitales. Es un software de código abierto se distribuye bajo los términos de BSD open

⁵⁸ La GNU Affero General Public License (Licencia Pública General Affero GNU), o simplemente GNU Affero GPL, es una licencia de software libre publicada recientemente por la Free Software Foundation. GNU AGPL tiene el propósito de ser una licencia mínimamente modificada de la GNU GPL y atender las necesidades de proporcionar libertad en software como servicios (SaaS, Software as a Service), es decir, aquellos a los que no se tiene acceso directo al binario / código - objeto.

⁵⁹ Elaborado a partir de las informaciones disponibles en la página web de la empresa DuraSpace y OpenDOAR: <https://duraspace.org/dspace/about/features/> y https://v2.sherpa.ac.uk/view/repository_visualisations/1.html

⁶⁰ El sistema de bibliotecas del Instituto de Tecnología de Massachusetts (MIT Libraries) cubre las cinco escuelas académicas que conforman la universidad.

⁶¹ HP Labs es el grupo de investigación avanzada y exploratoria de HP Inc. La sede de HP Labs se encuentra en Palo Alto, California, y el grupo cuenta con instalaciones de investigación y desarrollo en Bristol, Reino Unido.

⁶² DuraSpace es una organización sin fines de lucro, independiente que proporciona liderazgo e innovación para tecnologías abiertas que promueven el acceso duradero y persistente a datos digitales.

source license⁶³, esto significa que cualquier usuario puede usar, modificar e incluso integrar el código en una aplicación comercial sin pagar ningún permisos de licencia.

En la actualidad DSpace es la plataforma de software de repositorio más instalada en el mundo (eso incluye los software de código abierto y propietario). Según en la página web OpenDoar de los 4.143 repositorios digitales implantados en el mundo, hasta enero de 2019, 43% utilizan DSpace, en Brasil este porcentaje es casi el doble, llega a 80%. Es utilizado por archivos, museos, empresas comerciales para administrar sus activos digitales, pero el uso más común es por parte de las bibliotecas académicas y de investigación, como un repositorio institucional de acceso abierto para administrar y difundir los resultados académicos.

La interfaz de DSpace está basada en la web, puede instalarse en Linux, Mac OSX o Windows. Reconoce y administra una gran cantidad de formatos de documentos comunes como: PDF, Word, JPEG, MPEG, TIFF. Aunque DSpace está configurado solo para reconocer automáticamente los formatos de documentos comunes, cualquier formato puede ser gestionado en la plataforma, también permite registrar cualquier formato no reconocido, para que pueda identificarse en el futuro.

4.4.3 RODA (Repositorio de Objetos Digitales Auténticos)

RODA⁶⁴ es un software de repositorio digital a largo plazo creado por iniciativa de la Direção Geral de Arquivos (DGARQ, anteriormente Instituto dos Arquivos Nacionais/Torre do Tombo), en colaboración con la Universidade do Minho, ambas de Portugal. RODA surge a partir de la propia administración pública lusitana, consciente de la importancia del documento digital y de su incorporación en la gestión archivística, así como de la complejidad del reto que supone garantizar su preservación a largo plazo. El software es el resultado de un ejercicio de análisis y desarrollo de un entorno de gestión basado en el modelo OAIS. Un proyecto acompañado de una importante dotación de recursos tecnológicos y económicos y de un compromiso político para garantizar su desarrollo e implantación.

RODA es una solución de repositorio digital que ofrece funcionalidad para todas las principales unidades del modelo de referencia de OAIS, es capaz de ingerir, administrar y proporcionar acceso a los diversos tipos de objetos digitales producidos por empresas comerciales u órganos públicos. Implementa flujo de trabajo de ingreso que valida los SIPs, y verifica su contenido en busca de virus, realiza la identificación de formatos, extrae metadatos técnicos. Está 100% construido sobre tecnologías de código abierto, toda la infraestructura necesaria para admitir el software es independiente del proveedor, esto

⁶³ La licencia BSD es una licencia de código abierto utilizada inicialmente en los sistemas operativos del tipo Berkeley Software Distribution. A pesar de que se ha creado para los sistemas BSD, actualmente varios otros sistemas se distribuyen bajo esta licencia. Esta licencia impone pocas restricciones cuando se compara las impuestas por otras licencias, como la GNU General Public License o incluso las restricciones estándar determinadas por el copyright, colocándola relativamente cerca del dominio público.

⁶⁴ Elaborado a partir de las informaciones disponibles en la página del software RODA: <https://demo.roda-community.org/#welcome>

significa que puede usar el hardware y las distribuciones de Linux que mejor se adapten a las necesidades de la institución.

Es compatible con estándares existentes tales como el Sistema de información de archivo abierto (OAIS) y los esquemas de metadatos: Dublin Core, EAD, METS y PREMIS. Los componentes de migración admiten la migración de documentos de texto, imágenes, bases de datos relacionales, video y audio a formatos normalizados para la preservación a largo plazo. Un mecanismo de complemento permite a RODA admitir conversiones de formatos adicionales.

Las interfaces del software permiten a los usuarios gestionar fondos, colecciones y definir reglas para acciones de preservación. Todas las interacciones entre los usuarios (humanos y máquinas) y el repositorio son registradas por razones de seguridad y responsabilidad.

PARTE III – EVALUACIÓN Y RESULTADOS

5. METODOLOGÍA

La metodología utilizada para realizar la Parte III del trabajo fue en primer lugar la revisión de la bibliografía especializada. Esta metodología consistió en seleccionar artículos relevantes, centrados en evaluación de las características y funcionalidades de los softwares para repositorios digitales, publicados a partir del año 2008. En el periodo elegido no había muchos trabajos sobre evaluación de las características y funcionalidades de los softwares para repositorios digitales. En general había más trabajos centrados en evaluar los repositorios que ya estaban implantados y no el software en el cual el repositorio fue construido, que es el objetivo de esta investigación. De la bibliografía especializada revisada, los trabajos que se revelaron como más relevantes y adecuados para lograr el objetivo de esta investigación fueron:

- **Evaluation of Three Open Source Software in Terms of Managing Repositories of Electronic Theses and Dissertations: A Comparison Study** (Masrek y Hakimjavadi 2012): evalúa las características de tres software de código abierto (DSpace, EPrints, Fedora).
- **Comparative Analysis of Open Source Digital Library Software: A Case Study** (Verma y Kumar, 2018): un estudio de tres software (Dspace, GSDL & EPrints) de la gestión de biblioteca digital de código abierto que recopila y difunde información para los usuarios.
- **Open Source Software for Digital Preservation Repositories: A Survey** (Rosa, Craveiros y Domingues, 2017): se centra en el estado del arte de once soluciones de software de código abierto para el campo de la preservación digital que se utiliza para asimilar y difundir información.
- **Os repositórios DSpace e archivematica para documentos arquivísticos digitais** (Lampert, 2016): el estudio presenta un análisis sobre aspectos de los programas DSpace y Archivematica, orientados a definir la herramienta más adecuada de repositorio digital para preservación y acceso de documentos archivísticos digitales.
- **Evaluación de software libre para la gestión de archivos administrativos.** Moreiro, Sánchez-Quadrado, Palacios y Barra, 2011): evaluación de tres programas libre: Archon, Archivists' toolkit e ICA-AtoM para la gestión de documentos administrativos y su implantación en el archivo de una universidad. Para la comparación se llevó en cuenta doce requisitos concernientes a la gestión de metadatos, normas de descripción y préstamo, análisis y ciclo de vida del documento.
- **Diretrizes para a Implementação de Repositórios Digitais Confiáveis de documentos arquivísticos.** Conselho Nacional de Arquivos (2014): indica requisitos para repositorios confiables de documentos archivísticos digitales, para garantizar la integridad, la autenticidad, la confidencialidad, la disponibilidad, el acceso y la preservación, teniendo en cuenta la perspectiva de la necesidad de mantener los acervos documentales a largo plazo o, incluso, permanentemente.

- **Institutional Repository Software Comparison** Bankier y Gleason (2014): compara las características de los principales software y está destinada a ayudar a las bibliotecas a centrarse en qué características ayudarán a facilitar el éxito de su repositorio.
- **Filling the Digital Preservation Gap: A Jisc Research Data Spring Project Phase One report.** (Mitcham, Awre, Allinson et al, 2015): introducción para el uso del software Archivemática.

La bibliografía especializada revisada está centrada en softwares para bibliotecas digitales y documentos administrativos. Los trabajos evalúan una variada gama de softwares, que fueron seleccionados a partir de diferentes características y funcionalidades como: estadísticas de ROAR y su idoneidad para una gran biblioteca; los más utilizados y populares; por relevancia para la preservación digital; etc.

Para elegir los softwares a ser evaluados en esta investigación, también fueron consideradas diferentes características y funcionalidades. La primera de ellas fue el tipo de licencia, es decir, que debería ser libre y de código abierto. La segunda fue el porcentaje de instalaciones registradas en el mundo, incluso en Brasil. Estas dos características están fuertemente relacionadas con la preservación digital y el coste para la implantación del repositorio, pues siendo libre no hay gastos con la adquisición ni con la licencia; al ser de código abierto y con una gran cantidad de instituciones utilizándolo, significa más desarrolladores trabajando para mantener el entorno operativo, de lo que resulta una vida más larga del software, y garantiza el acceso a largo plazo. El tercero y último requisito fue definido a partir de los tipos de instituciones que utilizaban el software, debido a que la presente investigación está centrada en materiales bibliográficos y archivísticos, es decir se valoró detenidamente los softwares implantados en bibliotecas y archivos. Los softwares que atendían mejor a los requisitos establecidos fueron: Archivemática, DSpace y RODA. Los tres son libres y de código abierto; DSpace es el software más instalado en el mundo y en Brasil (ver Figura 7. Los softwares más instalados en el mundo. y Figura 8. Los software más instalados en Brasil.), muy utilizados en las bibliotecas; ya el uso de Archivemática y RODA está muy difundido en archivos.

Después de elegir los softwares, fueron seleccionadas las funcionalidades más importantes y relevantes a ser evaluadas. En la bibliografía consultada no hay una metodología en común para la selección de las funcionalidades, es decir, se definía a partir del propósito de la implantación, así como por el tipo de material a ser gestionado por el repositorio.

Teniendo en cuenta el objetivo de la presente investigación y las necesidades de la Biblioteca Nacional de Brasil, las características consideradas más importantes y relevantes fueron: complejidad de instalación, usabilidad, flexibilidad y escalabilidad. A partir de dichas características fueron definidas seis categorías: Infraestructura, Gestión de contenidos, Administración del uso, Interfaz de usuario, Preservación digital e Interoperabilidad e Integración. Dentro de cada categoría fueron elegidas las funcionalidades más utilizadas en la bibliografía consultada, totalizando 40 de ellas. Para esta elección se consideró: los requisitos técnicos para la instalación, la seguridad de la información, estrategias de preservación, interfaz de usuario e integración. A continuación se describe cada categoría, así como las funcionalidades de cada una:

Infraestructura: la categoría de Infraestructura engloba los atributos básicos de cada oferta de repositorio, como: Facilidad para instalar, Facilidad para actualizar, Código fuente modificable, Documentación disponible, Soporta Windows, Soporta Linux, Soporta Java, Plataforma para publicación web, Flexibilidad para personalización, Capacidad de memoria y procesador bajo, Numerosa comunidad para soporte, Soporte externo en Brasil o en portugués y Madurez en el mercado.

Gestión de contenido: este requisito está relacionado con la facilidad con la que se crea, envía, revisa, y organiza el contenido dentro del repositorio. También abarca la provisión para las funciones de búsqueda, la jerárquica y los formatos aceptados. Para esta categoría fueron elegidas las siguientes funcionalidades: Restricción de acceso a información (pública/privada), Flujo de trabajo (*workflow*), Acepta múltiples formatos, Búsqueda avanzada/facetada, Búsqueda por texto integral, Opción de compartir en redes sociales.

Administración del uso: esta categoría es importante porque gestiona cómo los usuarios acceden al contenido en el repositorio, y define perfiles de usuarios con diferentes permisos de acciones dentro del repositorio. Las herramientas de autenticación institucional más comunes entre las comparadas permiten a los usuarios iniciar sesión utilizando las credenciales de inicio de sesión institucionales, sin la necesidad de crear una cuenta de repositorio separada. Otro aspecto importante de la administración de uso son los informes, que proporcionan excelentes estadísticas sobre el uso del repositorio. Esta categoría es importante para la seguridad de la información gestionada por el repositorio, y para el presente trabajo fueron elegidas las siguientes funcionalidades: Autenticación, Diferentes permisos de acceso, Registro de acciones de usuario, Notificaciones y Estadísticas.

Interfaz de usuario: abarca la flexibilidad en la personalización de la interfaz para satisfacer las necesidades de diferentes implementaciones, así como: Personalización, Acceso multilingüe en la interfaz, Administración multilingüe y Adaptación para móvil.

Preservación digital: un software para repositorio digital debe hacer la planificación de la preservación de los documentos bajo su custodia, a fin de frenar los problemas derivados de la obsolescencia tecnológica y la fragilidad del soporte, por eso el software debe incluir: Modelo OAIS, PREMIS, METS, EAD, Dublin Core, Exportar a/Importar desde otros formatos de metadatos; Herramientas/servicios de conversión de formatos, Validación e Identificación de formatos y Respaldo/Copia de Seguridad.

Interoperabilidad e Integración: la interoperabilidad entre los diferentes niveles de los sistemas de información es uno de los factores más básicos considerados por los tomadores de decisiones al seleccionar la plataforma, el software, los estándares y los protocolos de interoperabilidad para los sistemas de información. La flexibilidad que poseen los softwares para integrarse con otros sistemas, también es importante. Para esta categoría se seleccionó las siguientes funcionalidades: OIA-PMH, Integración con otros sistemas e Integración con los softwares evaluados.

Definidos los softwares y las funcionalidades a ser evaluadas, fue necesario definir cuál metodología de evaluación sería utilizada para alcanzar el resultado. Para esta etapa se utilizó el modelo creado por Alvite y Rodríguez (2002), para evaluar las interfaces de OPACs implementadas con UNICORN en las Universidades madrileñas. Para este trabajo se consideró solamente la manera como las autoras puntuaban cada indicador, que consistía en atribuir 0 (ausencia), 1 (aplicación incompleta/limitada) y 2 (aplicación completa). Es decir, cada funcionalidad seleccionada se puntuó con 0,1 o 2, en que 0

significa que el software no posee la funcionalidad, 1 que la posee de manera incompleta o limitada y 2 completa.

La puntuación descrita en el párrafo anterior fue atribuida a partir de la revisión de los manuales técnicos, las páginas web oficiales y las comunidades de soporte de los softwares.

5.1 Evaluación de los softwares

En este apartado se evalúa los softwares Archivemática, DSpace y RODA, según las funcionalidades escogidas, agrupadas en las categorías, que fueron identificadas en el apartado anterior. En este apartado también se destacará las funcionalidades consideradas como más importantes para atender a las necesidades de la BNDigital y del SEI, así como aquellas importantes para ambos. A continuación, la evaluación presentada en tablas:

5.1.1 Categoría Infraestructura

Funcionalidades	Puntuación		
	Archivemática 1.9.1	DSpace 6.x	RODA 3.0
Facilidad para instalar	0	2	2
Facilidad para actualizar	0	1	1
Código fuente modificable	2	2	2
Documentación disponible	2	2	1
Soporta Windows	1	2	2
Soporta Linux	2	2	2
Soporta Java	0	2	2
Plataforma para publicación web	0	2	0
Flexibilidad para personalización	2	2	2
Capacidad de memoria y procesador bajo	0	1	0
Numerosa comunidad para soporte	1	2	0
Soporte externo en Brasil o en portugués	2	2	1
Madurez en el mercado	1	2	0
SUBTOTAL	13/26	24/26	15/26

Tabla 3. Funcionalidades relacionadas con la infraestructura

De las trece funcionalidades, las que más sobresalen teniendo en cuenta la viabilidad de la implantación son la facilidad para instalar, la numerosa comunidad para soporte y madurez en el mercado. Pues al tener estas funcionalidades el software demuestra practicidad en su instalación, lo que se traduce en un coste más bajo para la implantación y también demuestran la satisfacción de los usuarios, por la numerosa comunidad de soporte y madurez en el mercado. La funcionalidad "plataforma para publicación web", es importante para atender a las necesidades de la BNDigital. El hecho de no poseer la

funcionalidad "soporte externo en Brasil", o en portugués, se considera como un punto débil para el mantenimiento y la actualización del software.

5.1.2 Categoría Gestión de contenido

Funcionalidades	Puntuación		
	Archivematica 1.9.1	DSpace 6.x	RODA 3.0
Restricción de acceso a información (pública/privada)	0	2	2
Flujo de trabajo (<i>workflow</i>)	2	2	2
Acepta múltiples formatos	2	2	2
Búsqueda Avanzada/facetada	2	2	2
Búsqueda por texto integral	0	2	2
Opción de compartir en redes sociales	0	2	0
SUBTOTAL	6/12	12/12	10/12

Tabla 4. Funcionalidades relacionadas con la Gestión de contenido

Para gestionar los materiales bibliográficos (acceso abierto) y los documentos archivísticos (acceso bajo demanda) de la Biblioteca Nacional de Brasil es esencial que el software tenga la funcionalidad de restricción de acceso a información (pública/privada). No poseer esta funcionalidad es un punto débil que puede inhabilitar el software para ser el más adecuado a la Biblioteca Nacional de Brasil. Las funcionalidades "búsqueda por texto integral" y la opción "compartir en redes sociales" son importantes para la BNDigital y el flujo de trabajo para el SEI.

5.1.3 Categoría Administración del uso

Funcionalidades	Puntuación		
	Archivematica 1.9.1	DSpace 6.x	RODA 3.0
Autenticación	2	2	2
Diferentes permisos de acceso	0	2	2
Registro de acciones de usuario	0	2	2
Notificaciones	0	2	2
Estadísticas	0	2	2
SUBTOTAL	2/10	10/10	10/10

Tabla 5. Funcionalidades relacionadas con la Administración del uso

Para garantizar la autenticidad e integridad de los documentos archivísticos es esencial que el software tenga la funcionalidad "diferentes permisos de acceso" y "registro de acciones de usuario", pues una define quién puede realizar diferentes acciones en el repositorio y en los documentos y la otra registra qué se ha realizado. No poseer estas funcionalidades por defecto, es decir, sin la necesidad de instalar otro software, es un

punto débil que imposibilita la gestión de los documentos archivísticos en su totalidad. La autenticación es importante principalmente para la gestión del SEI, que no es de acceso abierto.

5.1.4 Categoría Interfaz de usuario

Funcionalidades	Puntuación		
	Archivematica 1.9.1	DSpace 6.x	RODA 3.0
Personalización	1	2	1
Acceso multilingüe en la interfaz	2	2	2
Administración multilingüe	2	2	2
Adaptación para móvil	0	2	0
SUBTOTAL	5/8	8/8	5/8

Tabla 6. Funcionalidades relacionadas con la Interfaz de usuario

De esta categoría se consideran como esenciales dos funcionalidades, no solo para la Biblioteca Nacional de Brasil, sino para cualquier institución pública del país, que son el acceso y la administración multilingüe. Si el software no permite cambiar el idioma al portugués, dificultará o imposibilitará su uso, pues el portugués es la lengua oficial de Brasil y en los órganos públicos no está permitido la utilización de otra lengua que no la oficial así que es imprescindible poseer dichas funcionalidades. La última funcionalidad es importante para la BNDigital, porque el SEI implantado en la Biblioteca Nacional de Brasil no se utiliza mucho en los móviles.

5.1.5 Categoría Preservación digital

Funcionalidades	Puntuación		
	Archivematica 1.9.1	DSpace 6.x	RODA 3.0
Modelo OAIS	2	0	2
PREMIS	2	2	2
METS	2	2	2
EAD	2	0	2
Dublin Core	2	2	2
Exportar a / Importar desde otro formatos de metadatos	2	2	2
Herramientas/servicios de conversión de formatos	2	2	2
Validación e Identificación de formatos	2	2	2
Respaldo/Copia de Seguridad	2	2	2
SUBTOTAL	18/18	14/18	18/18

Tabla 7. Funcionalidades relacionadas con la Preservación digital

De las nueve funcionalidades seleccionadas para esta categoría, dos no son consideradas esenciales, Modelo OAIS y EAD, que ayudan en la preservación de los objetos digitales, pero no son esenciales. Dichas funcionalidades completan y amplían la preservación digital. Lo que no ocurre con las demás funcionalidades, que son esenciales para garantizar la preservación de materiales bibliográficos y documentos archivísticos. Sin ellas, los softwares, no garantizan la preservación digital.

5.1.6 Categoría: Interoperabilidad e Integración

Funcionalidades	Puntuación		
	Archivematica 1.9.1	DSpace 6.x	RODA 3.0
OIA-PMH	2	2	2
Integración con otros sistemas	2	2	2
Integración con los software evaluados	2	2	0
SUBTOTAL	6/6	6/6	6/4

Tabla 8. Funcionalidades relacionadas con la Interoperabilidad e Integración

La primera funcionalidad es esencial para cualquier biblioteca digital, y por lo tanto es la más importante para la BNDigital. En cuanto a la segunda, esta es imprescindible para el SEI, pues el software tendrá que estar integrado en él, ya que los procedimientos administrativos con trámites finalizados en el SEI serán transferidos al software para la gestión, preservación y acceso futuro.

6. RESULTADOS Y ANÁLISIS

En este apartado se presentan los resultados obtenidos por los softwares en cada categoría y en general, originados a partir de la suma de los puntos atribuidos a cada funcionalidad. Se sumó la puntuación obtenida en cada categoría, así como la puntuación máxima que cada software podría obtener. Estas puntuaciones se convirtieron en un porcentaje, con el fin de representar el nivel de asistencia del software en cada categoría y en general.

6.1 Resultado por software

6.1.1 Archivemática

Archivemática 1.9.1			
Categorías	Puntuación obtenida	Puntuación máxima	%
Infraestructura	13	26	50%
Gestión de contenido	6	12	50%
Administración del uso	2	10	20%
Interfaz de usuario	5	8	63%
Preservación digital	18	18	100%
Interoperabilidad e Integración	6	6	100%
TOTAL	50	80	62%

Tabla 9. Resultado Archivemática en cada categoría

A través del resultado obtenido en cada categoría se concluye que la preservación digital, interoperabilidad e integración son los puntos fuertes del software Archivemática y los débiles están en la infraestructura, gestión de contenido y administración del uso. Por los puntos fuertes se observa también que Archivemática no está indicado para ser utilizado solo, sino con otros sistemas. Y por sus puntos débiles se infiere que no atiende a las necesidades de la BNDigital.

6.1.2 DSpace

DSpace 6.x			
Categorías	Puntuación obtenida	Puntuación máxima	%
Infraestructura	24	26	92%
Gestión de contenido	12	12	100%
Administración del uso	10	10	100%
Interfaz de usuario	8	8	100%
Preservación digital	14	18	78%
Interoperabilidad e Integración	6	6	100%
TOTAL	74	80	92%

Tabla 10. Resultado DSpace en cada categoría

De las seis categorías analizadas, DSpace cumple con el cien por cien en cuatro de ellas. A partir del buen resultado en la categoría infraestructura se puede inferir que su instalación es la menos costosa de los tres softwares evaluados y la madurez en el mercado le atribuye un grado mayor de sostenibilidad. Por cumplir con el 100% de las funcionalidades en las categorías de gestión de contenido, administración del uso e interfaz del usuario, se puede concluir que está indicado para la gestión de bibliotecas digitales. Su peor resultado fue en la categoría preservación digital, por no poseer funcionalidades complementarias.

6.1.3 RODA

RODA 3.0			
Categorías	Puntuación obtenida	Puntuación máxima	%
Infraestructura	15	26	58%
Gestión de contenido	10	12	84%
Administración del uso	10	10	100%
Interfaz de usuario	5	8	63%
Preservación digital	18	18	100%
Interoperabilidad e Integración	4	6	67%
TOTAL	62	80	77%

Tabla 11. Resultado RODA en cada categoría

Las categorías Administración del uso y Preservación digital son los puntos fuertes del software RODA, pues posee la totalidad de las funcionalidades. Presentó aun un buen resultado en las categorías Gestión de contenido e Interoperabilidad e integración, y en esta última tan solo no cuenta con una funcionalidad, que sería importante si dos de los tres softwares evaluados fueran considerados como los dos más adecuados.

6.2 Resultado general

En el gráfico 1 a continuación, se presenta el resultado general obtenido por cada software. DSpace fue el que presentó mejor resultado y cumple con el 92% de las funcionalidades seleccionadas para esta evaluación. En segundo lugar está RODA, con un nivel de satisfacción del 77% y en último lugar está Archivematica, con un 62%. En el apartado siguiente se presenta un análisis de la puntuación atribuida a cada software por categoría.

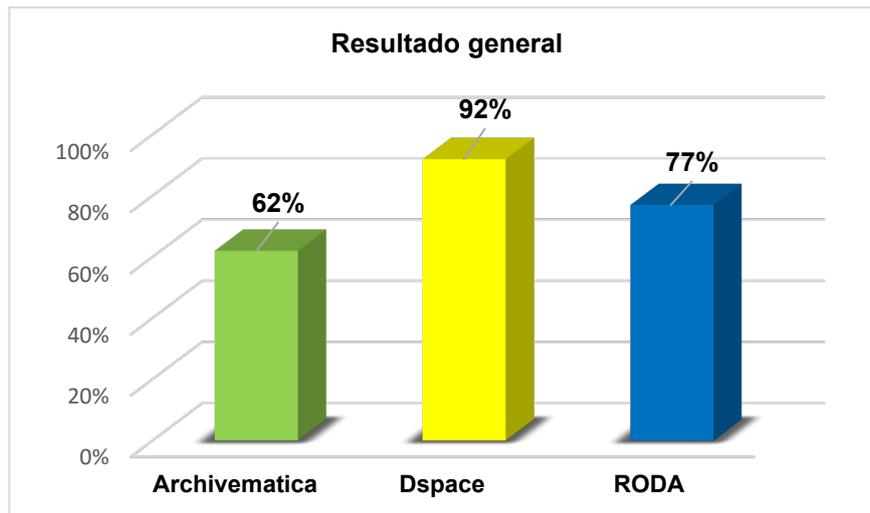


Gráfico 1. Resultado general

6.3 Análisis por categoría

Este apartado se dedica a analizar la puntuación obtenida por los softwares en cada categoría. El análisis presentado aquí se basó en los artículos centrados en el análisis de los softwares evaluados, los manuales, las páginas oficiales y las comunidades de soporte de los softwares. Aquí también se representa en gráficos el porcentaje alcanzado por cada software en cada una de las categorías, además de un comentario sobre las funcionalidades consideradas como las más importantes para atender a las necesidades de la Biblioteca Nacional de Brasil.

6.3.1 Infraestructura

Archivematica: según la información disponible en la página de preguntas frecuentes (FAQ⁶⁵) del software, la plataforma es una pieza compleja de software con muchas dependencias, todas de código abierto. Debido a su naturaleza compleja, no es posible crear un archivo ejecutable de pinchar para la instalación. Además, el software no es ejecutable directamente en Windows y no soporta Java, es decir, que requiere a profesionales expertos en Linux, Ubuntu y Python⁶⁶. Esta complejidad hace que la instalación y la actualización no sean fáciles. Mitcham, Awre, Allinson et al (2015), relatan que Archivematica no permite acceso público a la información y para acceder a la información gestionada por el software, es necesario la instalación e integración con el software AtoM. Los autores mencionados también comentan que Archivematica es flexible y se puede configurar de diferentes maneras para distintas necesidades institucionales y flujos de trabajo. Según Van Garderen, Jordan, Hooten (s.f), presidente de la empresa Artefactual, la comunidad digital está evolucionando y perfeccionándose.

⁶⁵ <https://www.archivematica.org/en/docs/archivematica-1.9/getting-started/troubleshooting/faq/>

⁶⁶ <http://dspace.2283337.n4.nabble.com/Regarding-dspace-and-its-comparison-with-archivematica-td4666304.html>

En Brasil están radicadas las empresas Archivematica Brasil⁶⁷ y META⁶⁸ que ofrecen servicios de instalación y mantenimiento.

DSPACE: de acuerdo con lo indicado en el apartado de preguntas frecuentes (FAQ⁶⁹), de la página web del software, una persona familiarizada con la instalación de Java en bases de datos es capaz de completar un prototipo de instalación de DSpace en unas pocas horas o un día. Rosa, Craveiros y Domingues (2017) explican que los beneficios de contar con una gran comunidad de desarrolladores y colaboradores, que siguen evolucionando y mejorando sus características, hacen de DSpace una de las soluciones más utilizadas en bibliotecas, instituciones educativas, gobiernos, organizaciones sin fines de lucro e incluso comerciales. Para Verma y Kumar (2018) DSpace puede ser personalizado fácilmente para gestionar y preservar el contenido digital y proporcionar a los usuarios el acceso a dichos contenidos. Como es un software de código abierto, una comunidad activa de desarrolladores, investigadores y usuarios de todo el mundo está brindando su experiencia para mejorar la aplicación. Las empresas eScire y NEKI IT ubicadas en Brasil ofrecen servicios de instalación, mantenimiento, actualización, etc, y están certificadas⁷⁰ por DSpace para proveer este tipo de servicios.

RODA: Coruja (2014) expone que una de las características del RODA, como sistema distribuido, es la posibilidad de ser implantado de diferentes maneras, como instalar todos los componentes en un servidor o cada componente en un sistema diferente, y soporta cualquier sistema operativo además de Java. La versatilidad de la instalación hace que esta sea fácil. Según la respuesta dada por Miguel Ferreira⁷¹ a través de la comunidad de soporte, RODA es un sistema de preservación OAIS y no una plataforma de publicación. Esto significa que la manera más común de usarlo es tener un archivo/biblioteca accesible solo para usuarios específicos dentro de una organización. Keep Solution⁷² es la única empresa de consultoría que ofrece servicios de instalación, mantenimiento y actualización del software y está ubicada en Portugal.

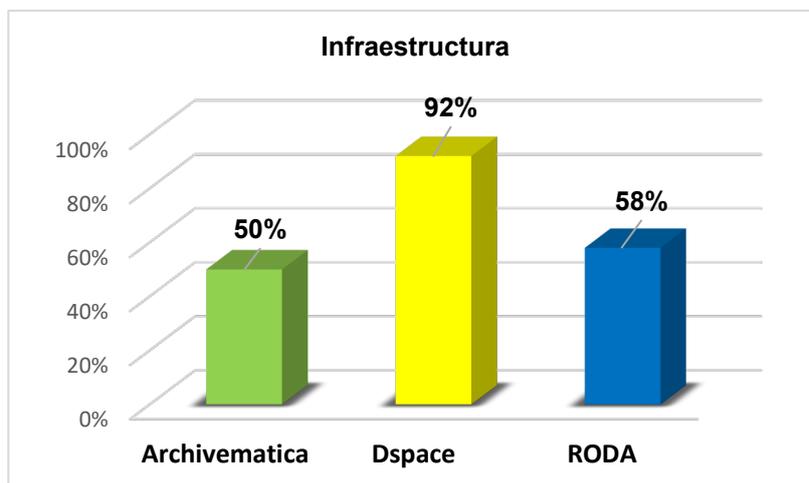


Gráfico 2. Resultado de los softwares en la categoría Infraestructura

⁶⁷ <https://archivematica.com.br/>

⁶⁸ <http://www.metadoc.com.br/>

⁶⁹ <https://wiki.duraspace.org/display/DSPACE/User+FAQ>

⁷⁰ <https://duraspace.org/dspace/resources/service-providers/>

⁷¹ <https://github.com/keeps/roda/issues/1485#issuecomment-506361932>

⁷² <https://www.keep.pt/>

La infraestructura es una categoría esencial para la viabilidad de la implantación del repositorio digital de la Biblioteca Nacional de Brasil. Por ello, elegir un software con más facilidades técnicas, es decir, que su instalación no sea muy compleja y costosa, hace viable la parte financiera de la implantación, que es un punto decisivo para cualquier proyecto en la institución. En esta categoría, el software que demostró más facilidad técnica, resultando así en la viabilidad financiera, fue DSpace, que responde a un 92% de las funcionalidades, además, sus funcionalidades soportan Windows y Java, que son softwares muy utilizados en el mercado. Ya Archivemática no soporta estos softwares. La documentación disponible para ayudar a instalarlo es más completa y las comunidades de soporte son numerosas, de lo que no dispone RODA.

6.3.2 Gestión de contenido

Archivemática: como el software no provee el acceso a la información, no dispone de la funcionalidad de compartir los objetos digitales en redes sociales, y tampoco la restricción de información pública o privada, así que debe ser otro software integrado a Archivemática el que debe proveer esta funcionalidad. Según el manual del usuario⁷³, la búsqueda en Archivemática se realiza a través de los AIP y de los metadatos.

DSpace: para Biswas y Paul (2010), DSpace es una plataforma para construir un repositorio institucional y las colecciones se pueden buscar y recuperar en la web. En el manual consta que los usuarios pueden acceder a páginas para elementos individuales, estas son descripciones de metadatos junto con archivos disponibles para descargar. Además, DSpace puede procesar contenido basado en texto cargado para la búsqueda de texto completo. Esto significa que no solo se podrán buscar los metadatos que proporcione para un archivo dado, sino que también se indexará todo su contenido. Esto permite a los usuarios buscar palabras clave específicas que solo aparecen en el contenido real y no en la descripción proporcionada.

RODA: Miguel Ferreira completa su respuesta en la comunidad diciendo que, no obstante, si se hacen públicos los objetos digitales, se podrá acceder a ellos en la interfaz de usuario sin autenticación, es decir, hay restricción de acceso para información pública y privada. De acuerdo con las preguntas frecuentes (FAQ⁷⁴) hay soporte nativo para búsqueda avanzada y de texto integral.

⁷³ <https://www.archivematica.org/en/docs/archivematica-1.9/user-manual/archival-storage/archival-storage/#search-aip>

⁷⁴ <https://github.com/keeps/roda/blob/master/documentation/FAQ.md>

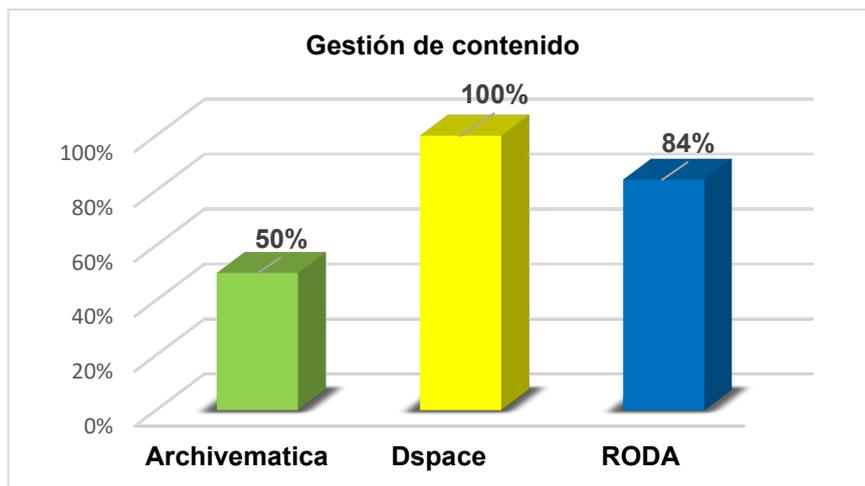


Gráfico 3. Resultado de los softwares en la categoría Gestión de contenido

Esta categoría es muy importante para atender a las necesidades de la BNDigital, principalmente por las opciones de búsquedas y poder compartir en redes sociales. También es importante para gestionar los documentos producidos por el SEI, pues contiene la funcionalidad de restricción de acceso y flujo de trabajo. Igualmente en esta categoría, DSpace es el que mejor atiende a las necesidades de la Biblioteca Nacional de Brasil, ya que posee todas las funcionalidades seleccionadas.

6.3.3 Administración del uso

Archivematica: de la categoría administración del uso Archivematica solo posee uno requisito: autenticación vía usuario y contraseña, LDAP y SHIBBOLETH. En la bibliografía especializada revisada no se encontró información sobre el requisito registro de acciones de usuarios. Como no provee el acceso tampoco hay diferentes permisos de acceso. En la relación de requisitos que Archivematica no ofrece Mitcham, Awre, Allinson et al (2015) destacan la ausencia de proporcionar estadísticas de uso.

DSpace: posee todas las funcionalidades de esta categoría, autenticación por usuario y contraseña, LDAP, SHIBBOLETH y CAS y además cuenta con diferentes permisos de acceso como: Añadir/borrar, leer (solamente visualizar) y editar. Dichos permisos son administrados por comunidad, colección e ítem. Se registra cada acción que el usuario realiza en la plataforma y se le notifica por correo electrónico (el mismo que informó al registrarse en la plataforma). Utiliza la aplicación Apache SOLR para las estadísticas, que pueden ser de dos tipos: de uso y del sistema. En la de uso se pueden recuperar las páginas, los elementos individuales, colecciones y comunidades. Estas páginas de estadísticas de uso muestran: total de visitas a la página; total de visitas al mes; descargas de documentos; países que más visitaron; ciudades que más visitaron. Las estadísticas del sistema permiten varios informes estadísticos sobre los contenidos y el uso del sistema, que pueden ser generados automáticamente por el software. El informe incluye lo siguiente: la cantidad de objetos archivados; de vistas del flujo de bits; de páginas vistas del artículo; páginas vistas de colección; páginas vistas de la comunidad; la cantidad de inicios de sesión del usuario; búsquedas realizadas; rechazos de licencias; solicitudes de OAI; la lista desglosada de visualizaciones de artículos; un desglose

completo de todas las acciones realizadas; inicio de sesión de usuario; las búsquedas más populares; información de nivel de registro (The DSpace Developer Team, 2018).

RODA: según Rosa, Craveiros y Domingues (2017), RODA proporciona el control de acceso y la gestión de permisos con configuración flexible, registro de las acciones del usuario y notificación de eventos por correo electrónico. Los permisos se pueden definir en cuanto al respecto al repositorio, e incluso a objetos de datos individuales. Las estadísticas se organizan por secciones y cada una se centra en un aspecto particular del repositorio, por ejemplo: problemas relacionados con metadatos y datos, estadísticas sobre los procesos de admisión y preservación, datos sobre usuarios y problemas de autenticación, eventos de preservación, gestión de riesgos y notificaciones.

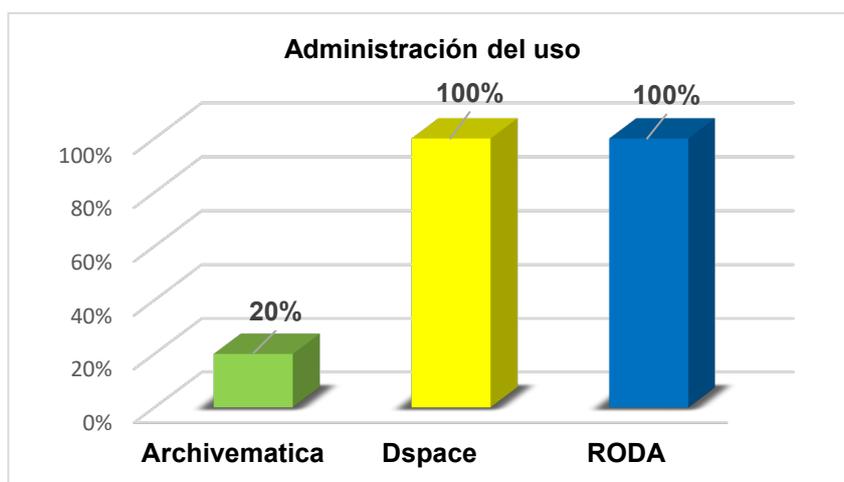


Gráfico 4. Resultado de los softwares en la categoría Administración del uso

Esta categoría es tan importante para la BNDigital como para el SEI, pues ahí están funcionalidades que ayudan a garantizar la autenticidad e integridad de la información, como diferentes permisos de acceso y registro de acciones de los usuarios. Las estadísticas informan sobre el uso de los documentos almacenados en la BNDigital y en el SEI. Este tipo de información es importante para justificar el mantenimiento del repositorio y también para aportar más inversión financiera para mejoras. DSpace y RODA atienden a todas las funcionalidades y Archivematica solo a una.

6.3.4 Interfaz de usuario

Archivematica: el diseño de las interfaces se basó en el modelo OAIS, ni siempre es intuitiva, al usuario que lleva algún tiempo para comprenderla completamente (Mitcham, Awre, Allinson et al, 2015). Por estar basada en el modelo OIAS, la personalización es limitada. Según las instrucciones de su manual, Archivematica se traduce a través de Transifex. Es posible traducir la documentación, la interfaz, el sitio web y el servicio de almacenamiento.

DSpace: para Pyrounakis, Nikolaidou, Hatzopoulos (2014) se puede configurar la interfaz web utilizando la tecnología JavaServer Pages (JSP) o el marco Apache Cocoon (XMLUI). También se pueden personalizar los índices y flujos de trabajo. DSpace admite el uso de

complementos Packager y Crosswalk. Es compatible con múltiples idiomas a través del uso de catálogos de internacionalización, según Cocoon Internationalization Transformer. La adaptación para móvil es posible a través del Mirage 2.

RODA: así como en Archivemática, la interfaz de usuario de RODA está basada en el modelo OAIS, lo que limita su personalización. También se traduce a través de Transifex, por lo que es posible añadir idiomas que no estén en la lista. No contempla la adaptación para el móvil.

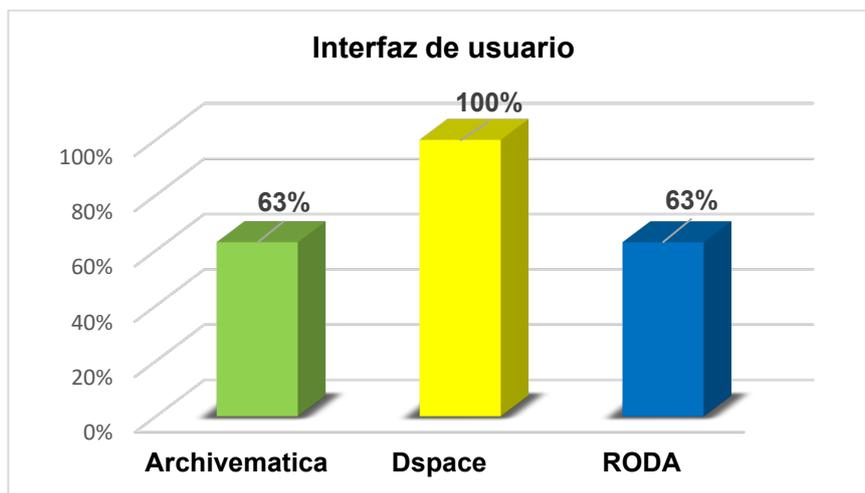


Gráfico 5. Resultado de los softwares en la categoría Interfaz de usuario

Debido a que BNDigital es de acceso abierto, alcanza muchos más usuarios que el SEI, que tiene acceso público, pero bajo demanda. Por la numerosa cantidad de usuarios, la categoría interfaz de usuario tiene más importancia para BNDigital que para el SEI. Archivemática y RODA tienen igual rendimiento y, una vez más, DSpace atiende al 100% de las funcionalidades.

6.3.5 Preservación digital

Archivemática: el sistema Archivemática utiliza un patrón de diseño de micro servicios para proporcionar un conjunto integrado de herramientas de software gratuitas y de código abierto que permite a los usuarios procesar objetos digitales desde el ingreso hasta el acceso de conformidad con el modelo funcional ISO-OAIS. Permite a profesionales de la preservación digital procesar transferencias digitales (objetos digitales, simples y complejos). Archivemática utiliza METS, PREMIS, Dublin Core y otros estándares de metadatos reconocidos. Los planes de preservación se basan en un análisis de las características significativas de los formatos de los documentos. Archivemática admite los planes de preservación de emulación al preservar los flujos de bits originales y respalda los planes de preservación de la conversión al monitorear los formatos de documento en riesgo, proporcionando un proceso para convertirlos en una fecha futura (Garderen, Jordan, Hooten et al, 2011). A partir de la explicación de los autores se puede concluir que Archivemática está centrado en la preservación digital, por eso posee todos los requisitos de esta categoría.

DSpace: Pyrounakis, Nikolaidou, Hatzopoulos (2014) explica que cada flujo de bits está asociado con un formato de flujo de bits, que es una forma única y coherente de referirse a un formato de documento en particular. Se define un nivel de soporte para cada formato bistream, que indica el nivel de preservación para el formato de archivo especificado. El esquema PREMIS se utiliza para representar metadatos técnicos sobre flujos de bits. Admite el uso de checksum y el control de versiones de flujos de bits. DSpace puede realizar copias de seguridad y restaurar todo su contenido como un conjunto de archivos de paquete de información de archivado (AIP). Según la respuesta de un desarrollador en la comunidad de soporte⁷⁵ DSpace proporciona muchas características que pueden considerarse piezas clave de un sistema OIAS, pero no está basada en el modelo OIAS.

RODA: es un sistema de repositorio digital a largo plazo que implementa un flujo de trabajo de ingreso de datos que, no solo valida los SIP, sino que también verifica su contenido en busca de virus, realiza la identificación de formatos, extrae metadatos técnicos y convierte formatos de documento en los más adecuados a la preservación digital. RODA garantiza que los datos introducidos son auténticos al registrar los metadatos PREMIS en todas las acciones realizadas por el repositorio, registra la procedencia en los estándares de metadatos de archivo como EAD o DC y garantiza la integridad y la disponibilidad al monitorear los datos con frecuencia y asegurarse de que no se los hayan manipulado.

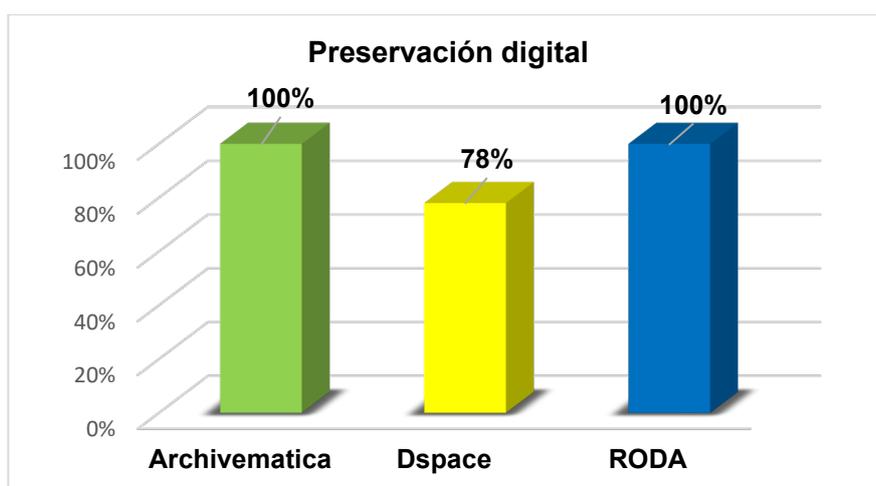


Gráfico 6. Resultado de los softwares en la categoría Preservación digital

Esta fue la única categoría en la que Archivematica y RODA disponen de más funcionalidades que DSpace, debido a que el concepto de los dos softwares está basado en el modelo OIAS y el de DSpace no. Además, DSpace no soporta el metadatos EAD por defecto, que es específico para archivos. Estas dos funcionalidades de las que carece DSpace, lo convierte en el menos completo para el SEI.

⁷⁵ <http://dspace.2283337.n4.nabble.com/OAIS-Reference-Model-td4686621.html>

6.3.6 Interoperabilidad e Integración

Archivematica: según Mitcham, Awre, Allinson et al (2015) Archivematica proporciona varios niveles de integración con DSpace, CONTENTdm, AtoM, Islandora y Archivist's Toolkit para el acceso y Arkivum, DuraCloud y LOCKSS para el almacenamiento. Hay continuas integraciones en curso con ArchivesSpace, Hydra, BitCurator y DataVerse. Además, proporciona una API REST de transferencia que se puede utilizar para iniciar transferencias dentro del software, el primer paso del flujo de trabajo de preservación. El servicio de almacenamiento subyacente de Archivematica también proporciona API REST para facilitar la creación, recuperación y eliminación de AIP.

DSpace: Castagné (2013) explica que DSpace está orientado hacia estándares y protocolos abiertos. Además de ser totalmente compatible con los protocolos OAI-PMH y SWORD, permite el ingreso y exportación de paquetes de información archivística (AIP) según las definiciones del modelo de referencia OAIS. Se integra con el software Archivematica.

RODA: expone toda su funcionalidad a través de una API de servicios REST bien documentada. Las bibliotecas de Java pertinentes están disponibles en GitHub para permitir a los desarrolladores interactuar con RODA a través de sus API principales. Existen varias herramientas para crear y manipular los SIP y enviarlos al flujo de trabajo de inserción de RODA.

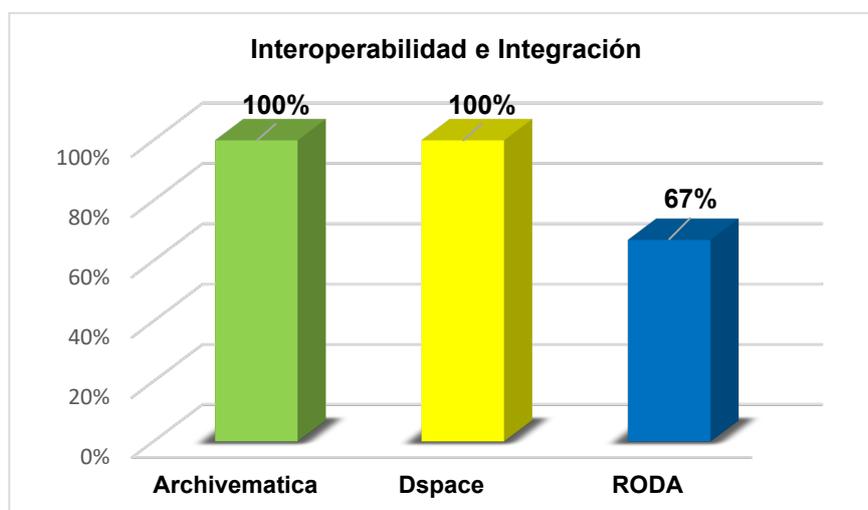


Gráfico 7. Resultado de los softwares en la categoría Interoperabilidad e Integración

El resultado obtenido por los softwares en esta categoría resultó equilibrado, pues los tres demostraron poseer capacidad para la interoperabilidad. Respecto a la parte de la integración, lo único que no tiene desarrollado RODA es la integración con los demás softwares evaluados en esta investigación, pero es posible realizarlo con ayuda externa.

7. CONCLUSIÓN

La presente investigación evaluó los softwares Archivemática, DSpace y RODA, con el fin de identificar el más adecuado para la gestión y preservación de los materiales bibliográficos y de los documentos archivísticos de la Biblioteca Nacional de Brasil. La evaluación demostró que Archivemática, uno de los softwares evaluados, no es un software para repositorio digital, sino un sistema de preservación digital, es decir, una solución tecnológica para un repositorio digital. La investigación llegó a esta conclusión porque, en este trabajo, se consideró repositorio digital el medio de almacenamiento que preserva, difunde y facilita el acceso a los objetos digitales que almacena. De acuerdo con el resultado de la evaluación, Archivemática no provee acceso a los objetos digitales, así que para acceder al contenido preservado por el programa, es necesario la instalación e integración con otro software. El software tampoco posee funcionalidades importantes para la gestión de contenido como: informes estadísticos, gestión de diferentes permisos de acceso, notificaciones de las acciones que los usuarios realizan en el sistema y búsqueda por texto integral. Al no poseer estas funcionalidades no se caracteriza como un software para repositorio digital y no es el más adecuado para la gestión de materiales bibliográficos de una biblioteca digital. Sin embargo, la investigación también constató que Archivemática es el software para preservación digital más reconocido en el mercado, pues fue desarrollado con la colaboración de la UNESCO, a través del Memory of the World's - Subcommittee on Technology.

El software RODA cumple con el 77% de las funcionalidades seleccionadas para esta evaluación. Sin embargo, es un software de preservación OAIS y no una plataforma de publicación. El uso más común de RODA es permitir que un archivo o biblioteca sea accesible solo para los usuarios de dentro de una institución, no obstante es posible hacer públicos los objetos digitales. RODA posee algunas funcionalidades en común con Archivemática, porque ambos están basados en el modelo OAIS. La diferencia es que RODA es un software que contiene algunas funcionalidades ya integradas en la aplicación como el acceso, no habiendo necesidad de instalación e integración con otro software, y también dispone de funcionalidades de gestión. El uso del software está disponible en el mercado desde 2009, pero todavía no está ampliamente difundido, existen pocas empresas que lo utilizan, su comunidad de soporte contiene pocos colaboradores.

La bibliografía revisada ya calificaba DSpace como el software más adecuado para implantación en bibliotecas digitales, es decir, para gestionar materiales bibliográficos y los datos acerca del porcentaje de instalaciones registradas en el mundo y los tipos de instituciones que utilizan el software corroboran esta calificación. Planteaba dudas si DSpace también era adecuado para la gestión de documentos archivísticos y para la preservación digital. Al cumplir con el cien por cien de las categorías gestión de contenido y administración del uso, se comprobó que es el software más adecuado para la gestión de objetos y documentos digitales.

En la categoría preservación digital, el porcentaje de cumplimiento de DSpace fue inferior a la de los otros softwares evaluados, pero de acuerdo con los estudios previos a esta investigación DSpace, así como Archivemática y RODA, tan solo no garantizan la preservación de los objetos digitales, pues los softwares de repositorio son solamente una parte del amplio panorama que constituye la preservación digital. Aunque Archivemática y RODA cumplen a más funcionalidades que DSpace en la categoría preservación digital, ellos no son una solución aislada que garantice el acceso a los

objetos digitales a largo plazo. Gran parte de los atractivos de dichos softwares proviene del hecho de que integran muchas otras herramientas de código abierto, que realizan funciones específicas y complementarias para el complejo y amplio objetivo que es la preservación digital a largo plazo.

DSpace, con la más grande comunidad de desarrolladores y administradores del mundo, está evolucionado desde repositorios institucionales hasta repositorio de documentos administrativos. Esto viene avalado por la creciente cantidad de estudios de casos y publicaciones del uso de DSpace como herramienta para repositorio de documentos administrativos y también su utilización por instituciones como el Departamento Administrativo de Ciencia, Tecnología e Innovación⁷⁶ de Colombia y el Development Bank of Latin America⁷⁷, de Estados Unidos.

Por la explicación en los párrafos anteriores y por el resultado general de la evaluación realizada en esta investigación, DSpace es el software más adecuado para la gestión y preservación de los materiales bibliográficos y de los documentos archivísticos de la Biblioteca Nacional de Brasil, atendiendo al 92% de las funcionalidades consideradas como las más importantes y relevantes para la gestión y preservación de objetos digitales. Aunque algunos autores⁷⁸ consideran que DSpace no contempla la preservación digital y el almacenamiento de documentos archivísticos digitales, en su totalidad el software es el que mejor atiende a las necesidades de la Biblioteca Nacional de Brasil, entre todos los softwares evaluados, ya que el programa más adecuado debe atender a las necesidades de la BNDigital, que tiene como propósito preservar los materiales bibliográficos para proveer el acceso a ellos.

La cantidad de funcionalidades que DSpace posee para preservación de los documentos archivísticos digitales producidos por el SEI es mucho más que las ofrecidas por Archivematica y RODA para la gestión de la BNDigital. La numerosa y amplia comunidad de usuarios, desarrolladores y administradores que DSpace posee es considerada por la comunidad de preservación digital como un indicador confiable de la longevidad del software, y esto es una característica esencial para la preservación a largo plazo.

⁷⁶ <http://repositorio.colciencias.gov.co/>

⁷⁷ <http://www.scioteca.caf.com/>

⁷⁸ Lampert, S. R. (2016)

8. BIBLIOGRAFIA

- Álvarez Wong, B. (2017). Los repositorios digitales para la conservación. Un acercamiento a la preservación digital a largo plazo. *Ciencias de la Información*, 48 (2), 15-22. Recuperado de: <http://www.redalyc.org/articulo.oa?id=181454540003>
- Alvite, M. y Rodríguez, B. (2002). Evaluación de interfaces de OPACs implementadas con UNICORN en las Universidades madrileñas. *III Jornadas españolas de bibliotecas digitales*. Madrid: Universidad Politécnica de Madrid. Recuperado de: <https://buleria.unileon.es/bitstream/handle/10612/3068/UNICORN%20en%20universidades%20madrile%C3%B1as.pdf?sequence=4>
- Arellano, M. A. M. (2004). Preservação de documentos digitais. *Ciência da Informação*, 33 (2), 15-27. Recuperado de: <http://eprints.rclis.org/5800/1/Ci%5B1%5D.Inf-2004-343.pdf>
- Arquivo Nacional de Brasil. (2005). *Dicionário brasileiro de terminologia arquivística*. Rio de Janeiro: Arquivo Nacional. Recuperado de: http://conarq.arquivonacional.gov.br/images/publicacoes_textos/dicionrio_de_terminologia_arquivistica.pdf
- Asociación Española de Normalización y Certificación. (2013). *UNE-EN ISO 13008:2012: Información y documentación – Proceso de migración y conversión de documentos electrónicos*. Madrid: AENOR.
- Asociación Española de Normalización y Certificación. (2015). *UNE-EN ISO 14721:2012: Sistemas de transferencia de datos e información espaciales – Sistema abierto de información de archivo (OAIS). Modelo de referencia*. Madrid: AENOR.
- Asociación Española de Normalización y Certificación. (2016). *UNE-EN ISO 15489-1:2006: Información y documentación – Gestión de documentos. Parte 1: Conceptos y principios*. Madrid: AENOR.
- Ball, A. (2006). *Briefing Paper: the OAIS Reference Model*. Recuperado de: <http://www.ukoln.ac.uk/projects/grand-challenge/papers/oaisBriefing.pdf>
- Bankier J. G. y Gleason K. (2014). *Institutional Repository Software Comparison*. Recuperado de: <https://unesdoc.unesco.org/ark:/48223/pf0000227115>
- Barbedo, F., Corujo, L., Sant'ana, M. (2011). *Recomendações para a produção de planos de preservação digital*. Recuperado de: http://arquivos.dqlab.gov.pt/wp-content/uploads/sites/16/2014/02/Recomend_producao_PPD_V2.1.pdf
- Bettencourt, A. M. (2014). *A representação da informação na Biblioteca Nacional: do documento tradicional ao digital*. (1a. ed). Rio de Janeiro: Fundação Biblioteca Nacional.

- Bhat, M. H. (2010). Interoperability of open access repositories in computer science and IT - an evaluation. *Library Hi Tech*, 28(1), 107–118. <https://doi-org.ezproxy.usal.es/10.1108/07378831011026724>
- Biswas, G. y Paul, D. (2010). An evaluative study on the open source digital library softwares for institutional repository: special reference to DSpace and Greenstone Digital Library. *International Journal of Library and Information Science*, 2(1), 1-10
Recuperado de: <https://academicjournals.org/journal/IJLIS/article-full-text-pdf/5F6BBD9889>
- Borba, V. R., Siebra, S. A., Galindo, M., Machiavelli, J. L., y Gusmão, C. M. G. (2015) Políticas de Formatos de Arquivos para Objetos de Aprendizagem: preservação digital no Saber Tecnologias Educacionais e Sociais. *Informação & Tecnologia (ITEC)*, 2(1), 80-97. Recuperado de: <http://www.brapci.inf.br/index.php/article/download/43789>
- Brown, A. (2008). Selecting File Formats for Long-Term Preservation. *The National Archives*. Recuperado de: <https://www.nationalarchives.gov.uk/documents/selecting-file-formats.pdf>
- Campos, F. M. G.; Saramago, M. L. (2007). Preservação digital de longo prazo em instituições patrimoniais: reutilização e adaptação de metadados. *Actas do Congresso Nacionais de Bibliotecários, Arquivistas e Documentalistas*, 0(9), 1-7. Recuperado de: <https://www.bad.pt/publicacoes/index.php/congressosbad/article/view/540/330>
- Caplan, P. (2009). Understanding PREMIS. *Library of Congress Network Development and MARC Standards Office*. Traducción: María Luisa Martínez-Conde. Recuperado de: https://www.loc.gov/standards/premis/UnderstandingPREMIS_espanol.pdf
- COAR. (2011). *El caso de Interoperabilidad para Repositorios de Acceso Abierto*. Recuperado de: https://www.coar-repositories.org/files/de_la_investigaci%C3%B3n-a-trav%C3%A9s-de-redes-globales-de-Repositorios-de-Acceso-Abierto-final-version.pdf
- Comisión Europea. (2017). *Marco Europeo de interoperabilidad - Estrategia de aplicación*. Recuperado de: https://eur-lex.europa.eu/resource.html?uri=cellar:2c2f2554-0faf-11e7-8a35-01aa75ed71a1.0010.02/DOC_1&format=PDF
- Conselho Nacional de Arquivos. (2011). *e-ARQ Brasil: Modelo de Requisitos para Sistemas Informatizados de Gestão Arquivística de Documentos*. Rio de Janeiro: Arquivo Nacional (Brasil). Recuperado de: http://conarq.arquivonacional.gov.br/images/publicacoes_textos/earqbrasil_model_requisitos_2009.pdf
- Conselho Nacional de Arquivos. (2014). *Diretrizes para a implementação de repositórios digitais confiáveis de documentos arquivísticos*. Rio de Janeiro: Arquivo Nacional (Brasil). Recuperado de: http://conarq.arquivonacional.gov.br/images/publicacoes_textos/diretrizes_rdc_arq.pdf

- Córdoba, S. (2011). Los repositorios institucionales y de acceso abierto en Costa Rica. *E-Colabora: Revista de ciencia, educación, innovación y cultura apoyadas por Redes de Tecnología Avanzada*, 1(2), 86-100. Recuperado de: <http://eprints.rclis.org/18151/>
- Corujo, L. (2014). *Repositórios Digitais e Confiança Um exemplo de repositório de Preservação Digital: o RODA*. (Trabajo de Fin de Máster, Universidade de Lisboa). Lisboa, Portugal. Recuperado de: <http://eprints.rclis.org/30375/>
- Crow, R. (2002). The Case for Institutional Repositories: A SPARC Position Paper. *ARL Bimonthly Report*, 223. Recuperado de: http://www.sparc.arl.org/sites/default/files/media_files/instrepo.pdf
- Cruz-Mundet, J. R., y Díez, C. 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>
- Decreto 8.297/2014, de 15 de agosto de 2014, aprova o Estatuto e o Quadro de Demonstrativo dos Cargos em Comissão da Fundação Biblioteca Nacional. *Diário Oficial da República Federativa do Brasil*.
- Decreto 8.539/2015, de 08 de octubre de 2015, dispõe sobre o uso do meio eletrônico para realização do processo administrativo. *Diário Oficial da República Federativa do Brasil*.
- Di Pietro, Z Maria Sylvia. (2016). *Direito Administrativo* (29a ed.). São Paulo: Atlas.
- Drucker, P. (1993). *Management: tasks, responsibilities and practices* (1a ed.). New York: HarperBusiness.
- European Commission. (2013). Competitiveness and Innovation Framework Programme (CIP). *ICT PSP Work Programme 2013*. Recuperado de: http://ec.europa.eu/cip/files/cip/docs/ict-psp-work-programme-2013_en.pdf
- Ferreras Fernández, T. (2010). *Preservación digital en RIs: Gredos*. (Trabajo de Fin de Máster, Universidad de Salamanca). Salamanca, España. Recuperado: http://eprints.rclis.org/16356/1/TFM_FerrerasT_Preservacion_Digital.pdf
- Ferreras Fernández, T. (2013). La interoperabilidad: el pegamento técnico para unir repositorios. *Blok de Bid*. Recuperado de: <http://www.ub.edu/blokdebid/ca/node/347>
- Flores, D. y Santos, H. M. (2014). Estratégias de preservação digital em arquivos: garantia de autenticidade e acesso contínuo em longo prazo. *VI Congresso Nacional de Arquivologia: Arquivologia, sustentabilidade e inovação: 20 a 23 de outubro de 2014* (pp.674-695). Santa Maria: Associação dos Arquivistas do Rio Grande do Sul. Recuperado de: <http://documentosarquivisticosdigitais.blogspot.com/2015/03/arquivologia-sustentabilidade-e.html>

- Formenton, D., Castro, F., Gracioso, L., Furnival, A., y Simões, M. (2018). Los estándares de metadatos como recursos tecnológicos para la garantía de la preservación digital. *Biblios: Journal of Librarianship and Information Science*, 0(68), 82-95. <https://dx.doi.org/10.5195/biblios.2017.414>
- Gemetto, J. (2017). *Herramientas para crear bibliotecas digitales*. Recuperado de: <https://www.articaonline.com/2017/01/herramientas-para-crear-bibliotecas-digitales/>
- González Ruiz, D., Térmens, M., y Ribera, M. (2012). Modelo de indicadores para evaluar los formatos digitales para la preservación de vídeo. *Revista española de Documentación Científica*, 35(2), 281-297. <http://dx.doi.org/10.3989/redc.2012.2.840>
- Governo Brasileiro. (2017). *e-PING Padrões de Interoperabilidade de Governo Eletrônico*. Comitê Executivo de Governo Eletrônico. Recuperado de: https://www.governodigital.gov.br/documentos-e-arquivos/e-ping/e-PING_v2017_20161221.pdf
- Granger, S. (2000). Emulation as a Digital Preservation Strategy. *D-Lib Magazine*, 6(10). Recuperado de: <http://www.dlib.org/dlib/october00/granger/10granger.html>
- Hakimjavadi, H., y Masrek, M. N. (2013). Evaluation of interoperability protocols in repositories of electronic theses and dissertations. *Program: electronic library and information systems*, 47(1), 34–59. Recuperado de: <https://doi-org.ezproxy.usal.es/10.1108/00330331211296303>
- Heminger, A., y Robertson, S. (2000). The Digital Rosetta Stone: A Model for Maintaining Long-term Access to Static Digital Documents. *Communications of the Association for Information Systems*, 3. <https://doi.org/10.17705/1CAIS.00302>
- Hunter, G. S. (2006). Custody of Digital Records: Options and Implementation Considerations. *The South Carolina Department of Archives and History*. Recuperado de: <https://core.ac.uk/download/pdf/49235161.pdf>
- IASA. (s.f). *Software de repositorio*. Recuperado de: <https://www.iasa-web.org/tc04-es/732-software-de-repositorio>
- Kultura 2.0. (2011). *Formatos de difusión y formatos de preservación de contenidos digitales*. Cuadernos de Formación. Recuperado de: http://www.kultura.ejgv.euskadi.eus/contenidos/informacion/kultura2_0_prestakuntza/es_k20_form/adjuntos/cuaderno-DIG-52.pdf
- Lamarca, María Jesús. (2006). *Hipertexto: El nuevo concepto de documento en la cultura de la imagen*. (Tesis Doctoral, Universidad Complutense de Madrid). Madrid, España. Recuperado de: http://www.hipertexto.info/documentos/dublin_core.htm
- Lampert, S. R. (2016). Os repositórios DSpace e Archivematica para documentos arquivísticos digitais. *Acervo*, 29(2), 143-154. Recuperado de: <http://revista.arquivonacional.gov.br/index.php/revistaacervo/article/view/718/734>

- Leija, D. (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. Recuperado de: https://www.tdx.cat/bitstream/handle/10803/454886/DALR_TESIS.pdf?sequence=1&isAllowed=y
- Library of Congress. (2018). *Sustainability of Digital Formats: Planning for Library of Congress Collections*. Recuperado de: https://www.loc.gov/preservation/digital/formats/fdd/browse_list.shtml
- Lor, P. (1997). Guidelines for legislation for national library services. *General Information Programme and UNISIST*. Paris: UNESCO. Recuperado de: <https://unesdoc.unesco.org/ark:/48223/pf0000109568>
- Lynch, C. (2003). Institutional Repositories: Essential Infrastructure for Scholarship in the Digital Age. *ARL Bimonthly Report* 226, 1-7. Recuperado de: https://www.researchgate.net/publication/215993581_Institutional_Repositories_Essential_Infrastructure_For_Scholarship_In_The_Digital_Age
- Masrek M.N. y Hakimjavadi, H. (2012). Evaluation of Three Open Source Software in Terms of Managing Repositories of Electronic Theses and Dissertations: A Comparison Study. *Journal of Basic and Applied Scientific Research*, 2(11), 10843-10852. Recuperado de : <https://pdfs.semanticscholar.org/4f20/fa7f78e82ed10b57ec0eb6e54dd29e23d0f1.pdf?ga=2.31173943.928927258.1563125175-1286103302.1560078805>
- Mateus, J. C. (2008). O Governo Eletrónico, a sua aposta em Portugal e a importância das Tecnologias de Comunicação para a sua estratégia. *Revista de Estudos Politécnicos*, 4(9), 31-52. Recuperado de: <http://www.scielo.mec.pt/pdf/tek/n9/n9a02.pdf>
- Mitcham, J., Awre C., Allinson J., Green R. y Wilson, S. (2015). Filling the Digital Preservation Gap. A Jisc Research Data Spring project. Phase One report - July 2015. *figshare*. <https://doi.org/10.6084/m9.figshare.1481170.v1>
- Moreiro, J. A., Sánchez-Cuadrado, S., Palacios, V., y Barra, E. (2011). Evaluación de Software Libre Para la Gestión de Archivos Administrativos. *El Profesional de la Información*, 20(2), 206-213. Recuperado de: <https://recyt.fecyt.es/index.php/EPI/article/view/epi.2011.mar.12>
- National Library of Australia. (2003). *Directrices para la preservación del patrimonio digital*. Recuperado de: https://unesdoc.unesco.org/ark:/48223/pf0000130071_spa
- Parliamentary Archives. (2008). *A Digital Preservation Strategy for Parliament*. Recuperado de: <https://www.parliament.uk/documents/upload/digital-preservation-strategy-final-public-version.pdf>

- Polanco-Cortés, J. (2014). Repositorios digitales: Definición y pautas para su creación. Recuperado de: <https://ucrindex.ucr.ac.cr/docs/repositorios-digitales-definicion-y-pautas-para-su-creacion.pdf>
- Pyrounakis, G., Nikolaidou, M. y Hatzopoulos, M. (2014). Building digital collections using open source digital repository software: A comparative study. *International Journal of Digital Library Systems (IJDLs)*, 4(1), 10-24. Recuperado de : <https://pdfs.semanticscholar.org/b439/64f9e128869587fb2a17b9f7fb73a1b3f260.pdf>
- Riley, J. (2017). Understanding Metadata: What is metadata, and what is it for? *National Information Standards Organization (NISO)*. Recuperado de : https://groups.niso.org/apps/group_public/download.php/17443/understanding-metadata
- Rimkus, K. Padilla, T. Popp, T. y Martin, G. (2014). Digital preservation file format policies of ARL member libraries: na analysis. *D-Lib Magazine*, 20(3-4). Recuperado de: <http://www.dlib.org/dlib/march14/rimkus/03rimkus.print.html>
- Rodríguez Reséndiz, P. (2016). El OAIS en la preservación digital de archivos sonoros. *Investigación bibliotecológica*, 30(70), 197-220. <https://dx.doi.org/10.1016/j.ibbai.2016.10.009>
- Rosa, C. A., Craveiro, O., y Domingues, P. Open source software for digital preservation repositories: A survey. *International Journal of Computer Science and Engineering Survey (IJCES)*, 8(3). Recuperado de: <https://arxiv.org/abs/1707.06336>
- Shala, L. y Shala, A. (2016). File Formats-Characterization and Validation. *IFAC-PapersOnLine*, 49(29), 253-258. <https://doi.org/10.1016/j.ifacol.2016.11.062>
- Society of American Archivists. (2002). *Encoded Archival Description Working Group. Encoded Archival Description Tag Library: version 2002*. Recuperado de: https://www2.archivists.org/sites/all/files/EAD_TagLibrary_2002.pdf
- Takeshy, T. (2000). *Estratégia empresarial: tendências e desafios – um enfoque na realidade brasileira* (1a. Ed.). São Paulo: MBooks.
- Texier, J., De Giusti, M. R. Lira, A. J., Oviedo, N., y Villarreal, G. L. (2013). DSpace como herramienta para un repositorio de documentos administrativos en la Universidad Nacional Experimental del Táchira. *Revista Interamericana de Bibliotecología*, 36(2), 109-124. Recuperado de: <https://dialnet.unirioja.es/servlet/articulo?codigo=4612752>
- Texier, J. (2013). *Repositorio de Documentos Administrativos para la Universidad Nacional Experimental del Táchira (REDAUNET)*. Recuperado en: <http://eprints.rclis.org/20813/>
- The National Archives of UK. (2010). *File Format Conversion*. Recuperado de : <http://www.nationalarchives.gov.uk/documents/information-management/format-conversion.pdf>

- The National Archives of UK. (2017). *Migrating information between records management systems*. Recuperado de : <http://www.nationalarchives.gov.uk/documents/information-management/edrms.pdf>
- Theodoridou, M., Tzitzikas, Y., Doerr, M., Marketakis Y. y Melessankis, V. (2010). Modeling and querying provenance by extending CIDOC CRM. *Distributed and Parallel Databases*, 27(2), 169-210. Recuperado de : https://www.researchgate.net/publication/225106892_Modeling_and_querying_provenance_by_extending_CIDOC_CRM
- Thibodeau, K. (2002, octubre). *Overview of Technological Approaches to Digital Preservation and Challenges in Coming Years*. Washington : Council on Library and Information Resources. Recuperado de : http://chnm.gmu.edu/digitalhistory/links/pdf/preserving/8_37e.pdf
- Thibodeau, K. (2007). If you build it, will it fly? Criteria for success in a digital repository. *Journal of Digital Information*, 8(2) Recuperado de : <https://journals.tdl.org/jodi/index.php/jodi/article/view/197/174>
- Thurston, A. (2012). *Preservación digital: preservar el patrimonio y proteger los derechos civiles*. Recuperado de: http://www.unesco.org/new/es/media-services/single-view-tv-release/news/digital_preservation_preserving_heritage_and_protecting_civ/
- Tinoco, A. G. (2013). Políticas de Preservação Digital nos Arquivos Municipais Portugueses. *Cadernos BAD*, 1(2), 74-82. Recuperado de: <https://www.bad.pt/publicacoes/index.php/cadernos/article/view/1031>
- Tramullas, S. J. y Garrido P. P. (2006). Software libre para repositorios institucionales: propuestas para un modelo de evaluación de prestaciones. *El Profesional de la Información*, 15(3), 171–181. Recuperado de: <http://eprints.rclis.org/9345/>
- UNESCO. (2003). *Carta para Preservação do Patrimônio Arquivístico Digital*. Recuperado de: http://conarq.gov.br/images/publicacoes_textos/Carta_preservacao.pdf
- Van der Hoeven, J., Lohman, B. y Verdegem, R. (2008). Emulation for digital preservation in practice: The results. *The International Journal of Digital Curation*, 2(2), 123-132. <https://doi.org/10.2218/ijdc.v2i2.35>
- Van Garderen, P., Jordan, P., Hooten, T., Mumma, C. y McLellan, E. (s.f). The Archivemata Project. *Meeting Digital Continuity's Technical Challenges*. Recuperado de: http://www.unesco.org/new/fileadmin/MULTIMEDIA/HQ/CI/CI/pdf/mow/VC_Van_Garderen_et_al_26_Workshop1.pdf
- Verma, L., y Kumar, N. (2018). Comparative Analysis of Open Source Digital Library Softwares: A Case Study. *DESIDOC Journal of Library & Information Technology*, 38(5), 361-368. <https://doi.org/10.14429/djiit.38.5.12425>

Vicente, A. y Gozzer, S. (2011). Nuevos hábitos de lectura, escritura y publicación en la red. *Abaco: Revista de cultura y ciencias sociales*, (68-69), 82-87. Recuperado de: <https://dialnet.unirioja.es/servlet/articulo?codigo=3782566>

Wheatley, P. (2004). Institutional Repositories in the Context of Digital Preservation. *DPC Technology Watch Series Report 04-02*. Recuperado de: <https://www.dpconline.org/docs/miscellaneous/advice/346-institutional-repositories/file>