



UNIVERSITAT<sup>DE</sup>  
BARCELONA

**Preservación digital distribuida y la colaboración  
interinstitucional: Modelo de preservación digital  
para documentos con fines de investigación  
en universidades de México**

David Alonso Leija Román



Aquesta tesi doctoral està subjecta a la llicència **Reconeixement 3.0. Espanya de Creative Commons.**

Esta tesis doctoral está sujeta a la licencia **Reconocimiento 3.0. España de Creative Commons.**

This doctoral thesis is licensed under the **Creative Commons Attribution 3.0. Spain License.**



UNIVERSITAT DE  
BARCELONA

## **Tesis Doctoral**

Preservación digital distribuida y la colaboración interinstitucional:  
Modelo de preservación digital para documentos con fines de investigación en  
universidades de México.

**David Alonso Leija Román**

Director: Dr. Miquel Térmens Graells

Programa de Doctorado en Información y Documentación en la Sociedad del Conocimiento

Facultad de Biblioteconomía y Documentación

Barcelona 2017



# Índice

<b>Resumen</b>	<b>1</b>
<b>Capítulo 1 - Introducción y definición de estudio</b>	<b>3</b>
1.1 Introducción	5
1.1.1 Almacenar, respaldar, organizar y preservar	8
1.1.2 Preservar intangibles: datos, información y documentos digitales	10
1.2 Planteamiento e interés de la investigación	13
1.2.1 Preservación digital distribuida como modelo y estrategia sostenible	13
1.2.2 Colaboración y cooperación como esencia de la distribución	14
1.3 Definición y delimitación del campo de trabajo	15
1.4 Objetivo y preguntas de investigación	17
1.4.1 Objetivo General	17
1.4.2 Preguntas de investigación	17
1.5 Metodología de estudio e investigación	19
1.5.1 Revista bibliográfica de literatura	19
1.5.2 Estudio, selección y análisis de marcos temáticos	21
1.5.3 Definición, diseño conceptual y propuesta	24
1.6 limitaciones de metodología para estudio e investigación	26
<b>Capítulo 2 - Marco de referencia y contexto general de la preservación digital</b>	<b>29</b>
2.1 La preservación digital y la conservación	31
2.1.1 Definición y concepto de preservación digital	31
2.1.2 Modelos de archivo en preservación digital	39
2.1.3 Estrategias técnicas aplicadas en preservación digital	64
2.1.4 Marco general de directrices y modelos en políticas de preservación digital	78
2.2 Preservación digital distribuida	87
2.2.1 Concepto y definición de preservación digital distribuida (PDD)	89
2.2.2 Terminología de PDD	90
2.2.3 Protocolos en preservación digital distribuida	103
2.2.4 Sistemas de preservación digital distribuida	113
2.2.5 Estructura de organización y gobierno en modelos de PDD	116

2.3 México en la gestión de archivos y preservación digital	122
2.3.1 Contexto actual general	122
2.3.2 Instituciones de educación superior de México y la gestión documental digital	126
2.3.3 Preservación digital en México	132
2.3.4 Marco legal, gestión de archivos y preservación digital en México	134
<b>Capítulo 3 - Análisis y estudio de sistemas, modelos y necesidades en preservación digital distribuida</b>	<b>143</b>
3.1 Definición de análisis y estudio de sistemas, modelos y necesidades PDD	145
3.2 Análisis y estudio de información funcional de sistemas de PDD	146
3.2.1 Listado y análisis de información de sistemas PDD	147
3.2.2 Listado y análisis de funciones generales en sistemas de PDD	153
3.3 Estudio y valoración de coordinación, organización y gobernanza en redes de PDD	163
3.3.1 Antecedentes de redes privadas de preservación digital distribuida específicas	164
a) Antecedentes de MetaArchive Cooperative	164
3.3.2 Indicadores de organización, colaboración y coordinación en redes PDD	171
3.3.3 Aplicación de estudio y valoración OCMC	176
3.3.4 Análisis de respuestas y valoraciones al estudio OCMC	178
3.4 Estudio y análisis de actividades de preservación digital en IES públicas de México	187
3.4.1 Antecedentes y selección de IES-MX para estudio	187
3.4.2 Indicadores de niveles de preservación digital para estudio	189
3.4.3 Aplicación de estudio NDSA-LDP	192
3.4.4 Resultados y análisis de estudio NDSA-LDP en IES-MX	200
<b>Capítulo 4 - Modelo de Preservación Digital Distribuida IES México (PDDIM)</b>	<b>239</b>
4.1 Definición para propuesta de modelo de preservación digital distribuida IES-MX (PDDIM)	241
4.2 Marco general de referencia conceptual para modelo PDDIM	241
4.3 Elementos conceptuales del modelo de referencia, variable y de sistemas PDD para PDDIM.	242
4.3.1 Modelo de referencia OAIS en modelos de PDD	243
4.3.2 Elementos de definición conceptual y funcionales de sistemas PDD para PDDIM	246
4.4 Composición y organización funcional en casos de modelos de redes de PDD	250
4.4.1 Caso de modelo National Danish BitRepository	251
4.4.2 Caso de modelo CARINIANA de Brasil	253
4.4.3 Caso de modelo CSUC en MetaArchive Cooperative	256

4.4.3 Contraste FODA de elementos funcionales en casos de modelos de PDD	260
4.5 Visiones de necesidades, compatibilidad y lineamientos para modelo PDDIM	262
4.5.1 Necesidades de preservación digital en IES-MX ante modelo PDDIM	262
4.5.2 Compatibilidad de elementos y prácticas de modelos internacionales en modelo PDDIM	267
4.5.3 Políticas, lineamientos e instituciones en México ante modelo PDDIM	272
4.6 Integración e implementación de modelo de preservación digital distribuida IES México (PDIIM)	276
4.6.1 La escalera al modelo de preservación digital distribuida de IES México (PDDIM)	276
4.6.2 Ámbitos, estructuras y coordinación de modelo PDDIM	286
4.6.3 Representación integral de prestaciones, roles y relaciones de modelo PDDIM	300
4.6.4 Implementación parcial de modelo PDDIM a IES-MX	311
<b>Capítulo 5 - Conclusiones</b>	<b>323</b>
5.1 Conclusiones Generales	325
5.2 Limitaciones	328
5.3 Líneas futuras de trabajo	329
5.3.1 Estudio A/B de niveles de preservación digital en IES-MX	329
5.3.2 Modelo de organización de gobernanza y financiación	330
5.3.3 Implementación plena de modelo PDDIM	330
<b>6 - Fuentes</b>	<b>331</b>
6.1 Bibliografía y Referencias	333
<b>7 - Anexos</b>	<b>345</b>
7.1 Índice de Tablas	347
7.2 Índice de Figuras	348
7.3 Índice de Gráficas	349



*A Mavis y Alonso,  
por su apoyo incansable,  
y ser ambos el amor de mi vida.*

*A mis padres,  
por su amor, dedicación y legado.*

*A mis hermanos,  
por su cariño, compañía y ayuda.*



*A Miquel Térmens Graells,  
por su tiempo, dirección y aprecio.*

*A Mario Pérez-Montoro,  
por su disposición y ayuda académica.*

*A la Universidad Autónoma de Tamaulipas,  
por su apoyo y paciente soporte.*

*A CONACYT,  
por su confianza y pleno respaldo.*



# Resumen

La preservación digital de archivos utiliza la combinación de políticas, estrategias y acciones para mantener accesibles los contenidos digitales a largo plazo. A nivel funcional la preservación digital puede operarse sobre modelos de archivo centralizados o distribuidos. El modelo distribuido se distingue por hacer uso de la independencia, la coordinación y la replicación geográfica de archivos digitales. Dichos elementos son claves para integrar y sostener de forma descentralizada a grupos cooperativos de organizaciones con diferentes necesidades dentro de un plan colaborativo y distribuido de preservación digital a largo plazo.

Esta tesis aborda el estudio, análisis y definición de un modelo distribuido de preservación digital para la organización funcional de un conjunto de instituciones de educación superior pública de México (IES-MX).

Esta investigación se inicia con el estudio de los contextos y referentes en la gestión de la preservación digital de archivos a nivel internacional, donde se identifican los principales modelos, estrategias, técnicas y sistemas que proveen herramientas de soporte para su labor. Dentro de este marco de contextos se define un panorama general sobre la gestión de archivos digitales en sectores estratégicos de México como influencia para la preservación digital de instituciones de educación superior en México.

Dados los términos de referencia y contextos iniciales se continua con el desarrollo de un capítulo compuesto por un triplete de estudios y análisis integrado en principio por la identificación de los elementos y funciones clave en los sistemas de preservación digital distribuida. En segundo lugar se analizan los principales indicadores de organización, coordinación y gobernanza de redes distribuidas con apoyo de la ratificación participativa de usuarios de redes de preservación digital distribuida (PDD). Finalmente se realiza un tercer análisis enfocado a conocer las actividades y niveles de preservación digital en IES públicas de México pertenecientes al subsistema federal y estatal. El análisis de IES-MX se realizó de acuerdo a una metodología y recomendaciones internacionales referentes al sector de la preservación digital.

Como resultado final de este proyecto y trabajo de investigación se obtuvo una metodología u hoja de ruta con requisitos, recomendaciones y consideraciones generales para preparar y situar la integración paulatina de instituciones de educación superior públicas de México (IES-MX) a un esquema de preservación digital con finalidad distribuida. Otro resultado obtenido fue la definición e integración de un modelo de preservación digital distribuida para IES-MX sobre una implementación parcial de organización, relaciones y participación.



## **Capítulo 1 - Introducción y definición de estudio**



## 1.1 Introducción

¿Por qué preservar nuestra información y patrimonio digital de cara al futuro?. Esta pregunta sugiere poner en perspectiva y a conciencia cultural las diversas posibilidades de uso, valor e importancia futura de nuestra información como producto digital documental tanto a nivel individual como de sociedad y civilización organizada.

En la actualidad los perfiles convencionales y predominantes en la creación de documentos e información digital mantienen ritmos de aceleración impredecible, ya que su producción es incesante, su consumo es inmediato y su conservación sin plazo fijo. La facilidad informática con la que se puede crear, modificar, compartir y almacenar un documento digital adelgaza la importancia panorámica de una gestión digital preventiva con procedimientos formales que ayuden a organizar, preservar y acceder de manera segura a la información en sus distintos formatos digitales a largo plazo. La preservación digital como actividad formal prevé la gestión de archivos ante riesgos digitales ocasionados por errores humanos, obsolescencia tecnológica, desastres naturales y problemas de tipo informático que posibilitan situaciones puntuales de pérdida de datos e información en las estructuras de contenidos con composición binaria digital. Una visión holística sobre la obsolescencia y los errores humanos que pone sobre la mesa la importancia y necesidad de concienciar sobre el impacto que representa la gestión formal y preventiva relacionada con los hábitos humanos y los procesos tecnológicos (incluidos los digitales) se proyecta acertadamente en una frase de Al Gore (2006). La frase hace referencia a que la combinación de viejos hábitos y vieja tecnología dan como resultado consecuencias predecibles. Así cuando los viejos hábitos se combinan con nuevas tecnologías las consecuencias se verán dramáticamente alteradas.

Ante estos escenarios impredecibles donde la soberbia tecnológica de proveedores, usuarios y productores digitales atestiguan una bonanza digital fortuita, existe la necesidad de despertar hacia una noción cultural que procure y haga conciencia sobre el devenir del patrimonio e información digital en sus distintos ámbitos productivos. La posibilidad de entrar en un estado de zozobra digital nos orienta a considerar la visión de la preservación digital como una cultura subyacente que de forma gradual sopesa las necesidades potenciales e impredecibles en el ámbito digital tanto en su almacenamiento, integridad, accesibilidad y durabilidad. La

combinación de los veloces cambios y avances tecnológicos marcan un ritmo incierto sobre la longevidad del patrimonio y memoria digital de nuestra era.

A nivel mundial y principalmente en países desarrollados, organizaciones y sectores especializados en la gestión de conocimiento e información han integrado ya desde hace algunas décadas la conciencia operativa y cultura activa sobre la importancia de establecer protocolos, actividades y colaboraciones formales orientadas a la preservación de sus memorias y patrimonio digital. Institutos, gobiernos y organizaciones mundiales como la NASA, UNESCO, Massachusetts Institute of Technology (MIT), Harvard, Library of Congress, IBM, Stanford University, National Library of Australia o Koninklijke Bibliotheek, han desarrollado investigación, sistemas y modelos de trabajo orientados a mantener el acceso permanente y a largo plazo del patrimonio documental de contenidos nacidos digitalmente y materiales digitalizados.

Son varios los precedentes e hitos en la historia temprana de la preservación digital que han hecho concientizar a comunidades de creadores y gestores de información digital sobre la importancia de su tratamiento y gestión oportuna. Desde la publicación de “Ensuring the longevity of digital documents” (Rothenberg, 1995-99) se visualiza, analiza y prevé técnicamente sobre la accesibilidad futura de soportes y formatos digitales ante la obsolescencia tecnológica. En esta línea de importancia y ante la previsión de necesidades de este tipo la NASA creó la definición del hoy estándar ISO Modelo (OAIS) Open Archival Information System (CCSDS, 2002-12) con el objetivo de integrar un marco formal de trabajo técnico de preservación digital a largo plazo bajo un modelo estándar, flexible y orientativo. Otros hitos importantes como el dictado en conferencia de “La carta sobre la preservación del patrimonio digital” (UNESCO, 2003) y la declaración de Vancouver sobre “La Memoria del Mundo en la era digital: digitalización y preservación” (UNESCO/UBC, 2012), comunican de manera emergente sobre las pérdidas de grandes cantidades de información digital por el desconocimiento de su importancia y gestión. En estas se habla de la urgente necesidad en la creación de programas de preservación digital y colaboración dentro de marcos legales e institucionales que garanticen la conservación del patrimonio mundial digital mediante procesos transparentes de cooperación internacional. Estos programas deben dotar de apoyo y conocimiento que cubra por un lado las necesidades urgentes de capacitación humana especializada en el tema, y por otro de las posibilidades de financiación formal de actividades

de preservación digital con el fin de proteger la memoria digital mundial. Todos estos antecedentes han sucumbido e inspirado hacia una seminal cultura de preservación digital mundial. Algunos expertos aseguran que han sido grandes los avances y adopción de la preservación digital como actividad especializada, sin embargo es vital seguir potenciando su divulgación de importancia tanto en formación, gestión de herramientas, métodos y modelos de trabajo para obtener resultados tangibles y transversales en un futuro de corto, mediano y largo plazo.

Desde hace una década la preservación digital se coloca en torno a diferentes acciones, actividades y desarrollos como un tema clave para la investigación de panoramas y soluciones en la gestión del patrimonio digital de entornos culturales y científicos. Estos esfuerzos tienen como objetivo el crear marcos de trabajo que aseguren el acceso e integridad de información y contenidos digitales a generaciones futuras. Los primeros colectivos en concientizar y activarse ante la gestión documental del patrimonio digital con miras al largo plazo, fueron las comunidades de sectores de la investigación científica, archivos nacionales, bibliotecas y organismos de administración pública. Tal actividad proviene como consecuencia de las necesidades paulatinas presentadas en la transición operativa e inclusión equitativa del tratamiento documental material y digital. El desarrollo de herramientas informáticas de gestión electrónica de documentos (gestores documentales, repositorios, etc.), sistemas de almacenamiento digital y la definición de políticas normativas de gestión documental, nacieron al paralelo como los siguientes pasos a incorporarse en las acciones complementarias del marco de trabajo orientado a la preservación, tratamiento y conservación de documentos de orden digital.

La preservación y tratamiento del patrimonio digital a largo plazo o *long-term preservation*<sup>1</sup>, se encuentran aún en un estado de conciencia y vigilia por su condición de alerta y reconocimiento permanente del problema que representan en sí mismas y para el entorno. Es por ello que existe una mayor actividad de investigación y evaluación de tipo teórico práctico, de organización y esquemas técnicos informáticos que pretenden formalizar estándares y modelos de integración fiables para las actividades de preservación digital en diversos sectores con necesidades de este tipo. El tema de la preservación digital a largo plazo se

---

<sup>1</sup> *Long-term preservation* es el término anglosajón oficializado que refiere a la preservación a largo plazo con un sistema de mantenimiento integral de objetos digitales.

considera a ojos comunes como un tópico de implementación práctico y sistemático no urgente. Uno de los retos que va más allá de aspectos técnicos, se centra en integrar a la preservación digital como una noción cultural de importancia tal como una organización integra protocolos culturales de sostenibilidad ecológica o preservación del patrimonio histórico monumental. La ponderación de una cultura digital preventiva con enclave en la preservación, es el gran reto a cumplir y este debe abordarse principalmente en proyectar los beneficios que la cultura de preservación digital genera en distintos sectores de interés como el proveer acceso permanente, íntegro y auténtico a la investigación y documentación histórica a largo plazo, así como el de exponer los riesgos que supone el no gestionar formal y adecuadamente altas cantidades de datos e información digital que se producen diariamente.

Es necesario procurar una cultura del mantenimiento y gestión preventiva, evitando así la implantación colateral de una cultura correctiva que a la larga es mucho más costosa y difícil de mantener, administrar y operar de cara al futuro. La piedra angular en la cultura de la preservación digital tiene dos caras prioritarias por atender. La primera se representa en concientizar y crear una verdadera cultura de preservación digital de la mano de la divulgación informativa de estrategias, técnicas y modelos de organización disponibles. La segunda cara asiste a la integración de estudios, ámbitos técnicos y de organización en marcos de trabajo flexibles orientados a cubrir las diferentes necesidades de organizaciones y comunidades designadas ante la importancia de mantener íntegro y seguro el patrimonio de activos digitales al largo plazo.

### **1.1.1 Almacenar, respaldar, organizar y preservar**

La preservación digital como disciplina emergente madura día a día su implantación a nivel de tecnología y organización mediante la aplicación de distintos modelos personalizados a las necesidades particulares de usuarios y organizaciones. Como procedimiento planificado, esta contempla una serie de recomendaciones y acciones previas necesarias para iniciarse como un programa formal. Una de las actividades primarias y más comunes para procurar el aseguramiento de la información digital se ve reflejada en los procedimientos de respaldo o backup, los cuales tienen como objetivo el almacenaje para asegurar el uso de materiales digitales a corto plazo. El respaldo o backup digital no integra el uso de técnicas, políticas y organización que le provean de las instrucciones necesarias para asegurar la integridad de los

objetos digitales a largo plazo. El almacenamiento digital que no considera procedimientos instruidos para asegurar su vida a largo plazo, por lo que encuentra más probabilidades de inseguridad y supervivencia a nivel digital. En este sentido, la gestión formal de procedimientos de almacenamiento o *backup* como actividad primaria de aseguramiento digital debe considerar la integración de acciones complementarias bajo un programa o plan de preservación digital a largo plazo. Para la integración de un programa de preservación digital a largo plazo se requiere de la consideración e integración previa de actividades según el tipo de organización. Termens (2013) refiere como ejemplo que en el caso específico de instituciones educativas o bibliotecas, estas deberán dotarse de un buen repositorio digital que integre todos los procedimientos anexos para su correcta operación como lo son el almacenamiento ordenado de información digital (archivo digital, repositorio digital, etc), el respaldo de información (backups físicos, copias redundantes o distribución de copias, etc) y la gestión documental (base de datos, metadatos descriptivos, etc.). En esta línea, es importante no considerar el largo plazo como modelo inicial, ya que debe prestarse atención a todos los requisitos técnicos y de organización previos que aseguran y gestionan la información a corto o mediano plazo, para de esta manera brindar mayor seguridad y fiabilidad a la gestión paulatina de objetos digitales en el presente de cara al futuro.

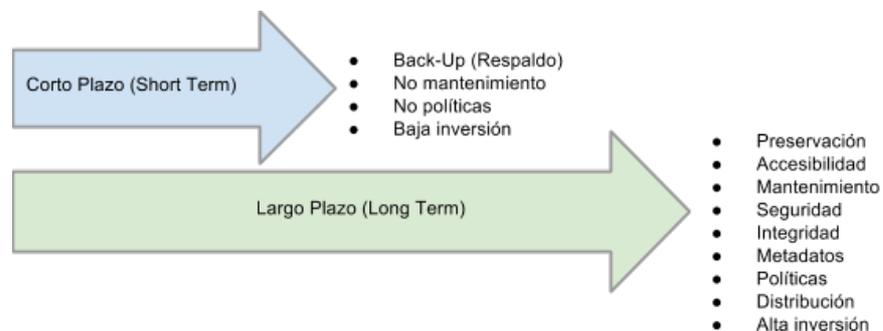


Figura 1. Diferencias de preservación y conservación digital a plazos.

De acuerdo con un estudio realizado por Hitchcock et al. (2007) y mencionado por Ferreras (2010), la mayoría de los repositorios institucionales están más preocupados por el acceso que por la preservación, y solo muy pocos cuentan con políticas de preservación digital a largo plazo. Esta situación plantea una serie complementaria de retos en la gestión y preservación digital a largo plazo, ya que además de los retos culturales sobre la importancia de las

preservación digital se suman los retos de formalización de actividades y políticas, así como los compromisos operativos y económicos que son indispensables para cubrir todas las actividades que integran un programa de preservación digital. Propuestas específicas de modelos de trabajo de preservación digital como la propuesta por la Digital Preservation Coalition (2009), sugieren una serie de actividades esenciales para complementar las actividades cotidianas de los repositorios institucionales con miras a preservar a largo plazo sus materiales almacenados y con ayuda de técnicas básicas de preservación digital, estas son: a) la preservación del software y hardware original con los que fueron creados los materiales digitales, b) preparar software y hardware para emulación los entornos anteriores, c) re codificar o refrescar los materiales en nuevos formatos para su interpretación futura y constante.

Las actividades técnicas sugeridas proveen soluciones centradas en asegurar el acceso digital a nivel de mantenimiento material o formatos lógicos en repositorios digitales, es por ello que a estas se deben sumar soluciones puntuales que complementan su actividad hacia la preservación digital de largo plazo tanto a nivel lógico, como de bits y de organización. Las políticas de organización, los protocolos de almacenamiento descentralizado, el monitoreo de integridad y autenticidad digital son los ejes necesarios para cubrir plenamente las necesidades de preservar, mantener y acceder a la información digital.

### **1.1.2 Preservar intangibles: datos, información y documentos digitales**

Se dice que lo intangible no se puede preservar, pero los documentos digitales si. En este orden, ¿Que son los documentos digitales frente a una constante producción de datos intangibles de origen informático?. En el ámbito académico, científico y de la investigación multidisciplinar cada día a se generan alrededor del mundo millones de documentos y materiales de origen digital en diferentes formatos (texto, audio, vídeo, bases de datos, etc). Schmidt (2010) afirmo que cada dos días del año 2010, se producía la misma cantidad de información digital que la producida desde el inicio de la humanidad hasta el 2003. Ante los diversos pronósticos de cantidades de producción y tipología de información digital, es importante matizar la importancia y diferencia que representan tanto la información, los documentos y los datos digitales. Una de las preguntas más planteadas desde inicios del siglo XXI en diversos sectores y publicaciones especializadas como las Directrices para la

Preservación del Patrimonio Digital (2003) y la Biblioteca Nacional de Australia, fue la de: ¿Qué materiales digitales hay que preservar?, es decir; ¿Qué criterios debemos seguir para decidir el tipo de documentos, formatos o representaciones digitales serán las prioritarias para el resguardo a largo plazo?, ¿Que tiene más valor?, ¿Un dato o un documento?. Aquí cabe hacer un paréntesis en la acepción específica y definición de documentos de investigación desde la perspectiva de dato, información y documento. Pérez-Montoro (2003) sugiere que el dato se define como pequeñas parcelas o trozos de la realidad, las cuales se convierten en unidades de soporte físico de la información. La información es el contenido semántico del dato, derivado de una clave de codificación en un contexto específico que le da sentido. Por tanto, un documento corresponde a la integración de los dos elementos a nivel de todo, es decir, como un soporte de representación formal de la información y datos.

En los documentos digitales con objetivos de integración a un plan de preservación digital existen dos cualidades que deben ser detectadas y diferenciadas dentro de los repositorios institucionales. Las actividades operativas y estructurales de los formatos de documentos digitales según su uso prioritario deben ser definidas y diferenciadas como formatos de consulta y formatos de preservación (Termens, 2013). Ambos formatos contienen prestaciones, elementos y cualidades que no son compartidas por completo entre ambos formatos. Elementos tales como las incrustaciones de metadatos y la calidad de usabilidad estructural del documento ayudan a que los formatos de consulta puedan ser indexados y recuperados con mayor facilidad a nivel de contenidos consultados (más no de preservación). Los formatos de preservación podrían nutrirse de la frontera contraria, con elementos que comprometen a los niveles de accesibilidad para mantenerlos en un régimen de usabilidad y flexibilidad de acceso en su estado de preservación.

Tabla 1. *Diferencias de formatos de consulta y formatos de preservación.*

Formatos de Consulta	Formatos de Preservación
<ul style="list-style-type: none"> <li>•Amplio uso.</li> <li>•Poco peso: transportable, rápido.</li> <li>•Facilidad de uso.</li> <li>•Bajos requerimientos técnicos.</li> </ul>	<ul style="list-style-type: none"> <li>•No es necesario su uso común.</li> <li>•No importa el peso.</li> <li>•Formatos con calidad.</li> <li>•Documentados.</li> <li>•No propietarios estándares</li> <li>•Sin DRM, sin licencias.</li> <li>•Múltiples lectores visualizadores.</li> <li>•Con metadatos incrustados</li> </ul>

La preservación digital como paso consecuente en la gestión de archivos y repositorios digitales ha alcanzado mejoras considerables en implementación e integración de protocolos de tecnología, cooperación institucional, sistemas de almacenamiento y organización. Las necesidades específicas de preservación digital en documentos digitales de investigación científica, busca además de adaptarse a nuevas plataformas de comunicación e incorporar meta información semántica para su recuperación, el consolidar su longevidad digital a modo íntegro y auténtico con todos los elementos y ficheros extras que los componen. La ruta actual en la gestión digital de repositorios institucionales de investigación y su preservación, apunta a incluir en ellos además de los documentos finales de investigación con resultados publicados (paper, artículo, tesis, etc.), los documentos de análisis y evaluación utilizados en la misma (bases de datos, tablas de ecuaciones, encuestas, etc.), como una tendencia paralela de repositorios de datos que favorezca la transparencia de las investigaciones y evitar el fraude científico (Hernández y García, 2013).

La gestión de accesibilidad y preservación digital de documentos de investigación es una de las actividades prioritarias a nivel global en el ámbito científico y académico, esto debido a la importancia que representa el conservar a largo plazo los productos de investigación de los centros e instituciones de educación superior. Es por ello que es muy importante fomentar el desarrollo de modelos, estrategias e investigación que ayude a mejorar los procedimientos de uso, almacenamiento, organización y conservación de la información digital ante agravantes como la obsolescencia tecnológica, desastres naturales y necesidades específicas de consumo de información de investigación para facilitar su comprensión a nivel semántico, sintáctico y de pragmática digital.

## **1.2 Planteamiento e interés de la investigación**

### **1.2.1 Preservación digital distribuida como modelo y estrategia sostenible**

El interés principal de esta investigación se centra en el estudio y definición de los marcos de trabajo necesarios para la creación de un modelo conceptual y funcional de preservación digital distribuida que antele de forma complementaria a la adopción, desarrollo e integración de procedimientos para la gestión electrónica de documentos con fines de investigación en instituciones de educación pública de México (IES-MX). En este sentido, este estudio pretende prever las necesidades de organización funcional para integrar la cultura de preservación digital a largo plazo de la memoria académica de investigación de manera cooperativa y sostenible. El modelo conceptual, funcional y de organización de preservación digital distribuida se estudia y analiza desde el carácter de aplicación interinstitucional de IES-MX del subsistema federal y estatal. El objeto principal es el de sopesar un acercamiento parcial de una estrategia viable, sostenible y alcanzable de preservación digital en modo distribuido que a su vez incentive el flujo, mantenimiento y accesibilidad de la producción de investigadores y docentes universitarios a largo plazo. Desde un punto de vista geográfico, México sugiere una ventaja inherente a la preservación digital distribuida (PDD), ya que la extensión territorial del país está constituida por 1,959,248 kilómetros cuadrados (INEGI, 2016), ya que para algunas directrices y protocolos de PDD la distancia necesaria de lejanía del material preservado de forma distribuida (replicada geográficamente), no puede ser menor de 129 y 200 km (Skinner y Schultz, 2010). Por otro lado, dada la extensa cantidad y tipología de instituciones que conforman el universo de IES-MX en México, una propuesta de modelo de preservación digital distribuida o colaborativa posibilita la creación de un marco de trabajo flexible y equilibrado donde instituciones con experiencia puedan funcionar como transmisores de conocimiento ante instituciones con pocos recursos o experiencia, para que de manera balanceada puedan integrarse paulatinamente a un esquema de preservación digital distribuido siguiendo patrones de independencia, coordinación y replicación, donde la cooperación de economía a escala posibilite desarrollar un esquema de participación dentro de un programa sostenible de preservación digital a largo plazo.

### **1.2.2 Colaboración y cooperación como esencia de la distribución**

En el mundo digital la Información y memoria digital generada por seres humanos solo puede sobrevivir si es organizada, procurada y preservada principalmente por la tarea de seres humanos y con la ayuda complementaria de tecnología adecuada. Así tal, se requiere de una estrategia coordinada entre individuos, instituciones y tecnología a fin de afrontar las necesidades de coexistencia de la memoria digital. Extendiendo esta idea a un mundo real de entidades e instituciones con fines comunes de procurar documentos de investigación, las estrategias de cooperación, colaboración y alianzas son elementos indispensables en la preservación digital como estrategia coordinada de participación múltiple. El compromiso unánime de diferentes instituciones por contribuir en la preservación a largo plazo de su información digital de forma compartida y replicada bajo un mismo sistema de organización y tecnología se relaciona directamente con casos de éxito en el uso de estrategias y sistemas de preservación digital distribuida como es el caso del sistema LOCKSS en MetaArchive Cooperative o CARINIANA, y BitRepository Project en National Danish BitRepository. En estos casos intervienen instituciones, organizaciones y cooperativas que mediante la coordinación, la independencia y la replicación de contenidos digitales integran programas de preservación digital en modo distribuido a largo plazo. A nivel técnico y de acuerdo a Skinner y Schultz (2010), una de las ventajas que ofrecen los sistemas de preservación digital distribuida se manifiesta en que las copias digitales replicadas bajo el sistema LOCKSS dentro de un grupo de instituciones que forman una red cooperativa se pueden regenerar de manera automatizada a nivel digital y mantener la integridad digital de todas las copias de un objeto digital. Para ello es necesario definir en el sistema un plan transversal de acciones reiteradas de revisión de estado de los originales y copias para mantenerlas siempre idénticas al objeto digital original. Esta solución técnica, viene a coste de los varios aprendizajes que han dejado experiencias de pérdida de información digital a nivel mundial como desastres naturales, ataques terroristas (11S, etc.), y accidentes humanos (mala gestión, respaldos olvidados, etc), por lo que a la par de estas soluciones técnicas, se han creado estrategias de organización interinstitucional que de forma convergente ofrecen alternativas viables a corto y largo plazo.

La preservación digital distribuida como estrategia de colaboración y cooperación interinstitucional coordinada, se reitera como un campo de estudio clave para la investigación

de nuevas soluciones y actividades preventivas de organización descentralizada para la gestión de procedimientos de preservación de documentos digitales. La preservación digital distribuida al ser el último paso funcional en una cadena evolutiva de la gestión documental digital compartida, diversifica una estrategia sostenible con la ayuda de la participación de varias instituciones tanto en infraestructura, responsabilidades y compromisos con el objetivo firme para mantener y conservar de manera segura el acceso íntegro a los documentos digitales de investigación en el presente y el futuro.

### **1.3 Definición y delimitación del campo de trabajo**

La definición e investigación para el desarrollo del modelo conceptual, funcional y de organización de preservación digital distribuida para documentos con fines de investigación de instituciones de educación superior públicas de México, se enfoca de manera primaria en conocer, estudiar y evaluar los contextos teóricos de antecedentes de la preservación digital centralizada y de su modelo distribuido. En un siguiente momento se estudiarán los casos prácticos de aplicación y organización de modelos de preservación digital distribuida de redes en Europa, Norteamérica y Latinoamérica. Los modelos distribuidos al ser estrategias emergentes en la práctica e implementación, nos obliga a revisar y estudiar de forma específica sobre el funcionamiento general de sistemas y protocolos tecnológicos de preservación digital distribuida, así como de la organización y colaboración interinstitucional para la puesta en común de infraestructuras tecnológicas de este tipo. Esta revisión nos ayuda a obtener referencias funcionales y marcos de trabajo implementados para correlacionar las necesidades de México ante la preservación digital en modo distribuido. El límite del campo de trabajo queda cedido a temas sectoriales, geográficos y de áreas especializadas enfocadas a dar soluciones estratégicas de preservación digital en ámbitos distribuidos para la gestión digital del sector de la educación superior y de investigación en México.

El sector de la educación superior en México de manera progresiva han integrado estrategias y recursos de tecnologías de la información (TIC) a sus instituciones educativas. Una de las ideas planteadas por la UNESCO (1999) definió que el proceso de integración de las tecnologías de la información en la educación, inicialmente se orientó principalmente a aspectos como el de informatizar la comunicación en los centros educativos a nivel

administrativo y de sus comunidades de estudiantes. A partir del año 2000 en el sector de las instituciones de educación superior de México (IES-MX) se implementaron paulatinamente recursos TIC con mayor peso institucional a distintos ritmos de adopción y mantenimiento. Los repositorios institucionales son un ejemplo del tipo de recursos TIC que han sido implementados en algunas instituciones del sector de IES-MX. Al día de hoy en el ámbito académico mundial esta herramienta representa la mayor opción de custodia digital de contenidos y documentos de investigación, así como la ventana formal institucional para ordenar, depositar y preservar sus documentos de investigación. Una importante proporción de IES y principalmente las pertenecientes al subsistema federal y estatal de México cuentan con repositorios institucionales de acuerdo a datos de REMERI<sup>2</sup>, OpenDOAR<sup>3</sup> y ROAR<sup>4</sup>. Los repositorios institucionales y otras herramientas digitales como bibliotecas o directorios digitales de documentos de investigación han sido implementados de manera asincrónica como parte de sitios web, como infraestructuras independientes o como asociaciones externas a otras instituciones que les dan el servicio. Esta situación sugiere una alta diversidad en los procesos de alcance, adopción y desarrollo formal de herramientas digitales de publicación y gestión de documentos de investigación en todo el mapa de instituciones de educación superior en México. La especialización de los repositorios institucionales como servicio de consulta y preservación de documentos digitales, requiere de protocolos, planes y estrategias puntuales para conservar las crecientes cantidades de información digital que resguardan. Es importante incentivar el reconocimiento de valor implícito que tienen los repositorios institucionales de documentos digitales de investigación, ya que contienen acervos culturales que necesitan ser procurados y gestionados con una visión al futuro ante su accesibilidad e integridad para su activa disposición. La evolución y ruta natural a la que apunta la implementación formal de sistemas de gestión de documentos digitales en repositorios de IES en México, mantiene una orientación de orden y secuencia natural de gestión que requiere definir procedimientos de normalización digital documental que alinee la creación y administración de repositorios, así como en la supervisión de las redes de repositorios bajo la definición formal de planes de preservación digital de los documentos que custodian.

---

<sup>2</sup> IES-MX con repositorios institucionales: <http://www.remeri.org.mx/portal/participantes.html>

<sup>3</sup> Repositorios en abierto de México en OpenDOAR: <http://www.opendoar.org/countrylist.php?cContinent=North%20America#Mexico>

<sup>4</sup> Repositorios en abierto de México en ROAR: <http://roar.eprints.org/view/geoname/geoname=5F2=5FMX.html>

## **1.4 Objetivo y preguntas de investigación**

### **1.4.1 Objetivo General**

Definir un modelo conceptual, funcional y de organización para la preservación digital distribuida de documentos con fines de investigación científica de instituciones de educación superior en México.

Este enfoque nace de la necesidad de sopesar las posibilidades y limitaciones de integrar un modelo de red nacional de preservación digital distribuida, que contribuya a enriquecer la salvaguarda y acceso a largo plazo de documentos de investigación de forma colaborativa y a escala con un grupo de universidades públicas específicas de México.

### **1.4.2 Preguntas de investigación**

Acorde al planteamiento del objetivo de investigación se correlacionan una serie de preguntas temáticas sobre los principales intereses que rigen la dirección de esta investigación.

#### a) Preservación digital

1. ¿Qué es, cómo funciona y qué objetivos persigue la preservación digital como disciplina emergente en el mundo del sector documental y de archivos?
2. ¿Que modelos, estrategias y recomendaciones de buenas prácticas en preservación digital están siendo aplicados a nivel mundial para su gestión?

#### b) Preservación digital distribuida

3. De forma enfática, ¿Como se define, integra y funciona la preservación digital distribuida a nivel de implementación técnica y de organización?
4. ¿Cuales son las principales ventajas, desventajas, similitudes y diferencias operativas en la gestión general de una red de preservación digital distribuida a nivel funcional y organizacional?

c) México en la gestión y preservación de documentos digitales de investigación

5. ¿Cuáles son las acciones, marcos o lineamientos vigentes para la gestión de archivos de documentos digitales en México, así como dentro del subsistema federal y estatal de instituciones de educación superior del país?
6. ¿Qué lugar ocupa y cual es el nivel de preservación digital que existe en la gestión documental de repositorios, BiDis o depósitos digitales de instituciones de educación superior pública que pertenecen al subsistema federal y estatal en México de acuerdo a las recomendaciones internacionales en la materia?

d) Modelo conceptual y funcional de preservación digital distribuida

7. ¿Que tipo de marcos preliminares y esquemas estructurales de organización documental en materia de preservación digital pueden cubrir los distintos cuadros de necesidades y alcances institucionales de tipo individual para la integración equilibrada de colaboración y participación de IES federales y estatales de México dentro de una estrategia de preservación digital distribuida?
8. ¿Que tipo de modelo de preservación digital distribuida puede integrar y articular de forma coordinada a un universo de IES de México de acuerdo a las diferencias que existen en su alcance formativo sobre el tema, su región geográfica o infraestructura institucional?

## **1.5 Metodología de estudio e investigación**

El desarrollo de esta tesis se compone por una investigación de tipo exploratoria y descriptiva integrada por procedimientos metodológicos mixtos de enfoque cualitativo. De manera diferenciada y progresiva en sus etapas de realización se integran técnicas de investigación como la revisión bibliográfica, la observación comparativa, el cuestionario, el test y la entrevista.

De manera general la investigación se concentra en tres procedimientos metodológicos específicos:

1. Revista bibliográfica de literatura para la definición de contextos de estado, cuadro conceptual, modelos de aplicación y marcos de función de la preservación digital, la preservación digital distribuida y la gestión de archivos digitales en México. Este bloque da respuesta a las preguntas de investigación número uno, tres y cinco.
2. Estudio, selección y análisis de marcos temáticos de sistemas, modelos y necesidades institucionales para la preservación digital distribuida mediante la triangulación de métodos y herramientas de recolección de datos diferenciados por tema. Este bloque refiere a un triángulo de análisis y estudios de los temas de interés que responden las preguntas de investigación número dos, cuatro y seis.
3. Definición, diseño y propuesta de modelo funcional teórico de preservación digital distribuida para IES México (PDDIM). Este bloque refiere a dar respuesta a las preguntas de investigación número siete y ocho.

### **1.5.1 Revista bibliográfica de literatura**

En este procedimiento incluye el estudio detallado de los contextos de referencia en los siguientes temas de interés: a) preservación digital, b) preservación digital distribuida, y c) México en la gestión de documentos digitales de investigación. Estos temas se desarrollarán

de forma individual y concatenada para la definición de un capítulo referente al contexto y marco de referencia de la preservación digital en los temas de interés.

- a) Para el tema de la preservación digital se obtuvo el panorama general, histórico y actual sobre la preservación digital, así como de sus distintos modelos, estrategias y técnicas aplicadas a nivel mundial. Para ello se realizaron dos momentos de revisión bibliográfica, donde el primer momento tuvo como objetivo la definición de la preservación digital a nivel conceptual, contexto de antecedentes y modelos de trabajo (referencia, técnicos, etc.). El segundo momento de revisión de literatura tuvo por objetivo definir el contexto actual en función de las estrategias, técnicas, políticas e implementación de la preservación digital. Para ambos se tomó como base el estudio de literatura especializada, artículos e investigaciones realizadas durante el periodo contemporáneo al estudio y complementado este con lo referido a los periodos subsecuentes o anteriores relevantes para el desarrollo de la investigación. El producto principal de la revisión de literatura de la preservación digital fue la descripción de un panorama amplio de elementos y definiciones conceptuales como base de los modelos de preservación digital empleados de manera teórica o técnica en diferentes sectores a nivel internacional y con especial énfasis para conocer su aplicación específica en el entorno de instituciones de educación superior y de investigación.
  
- b) Acorde a la estructura de trabajo planteada para el tema de preservación digital en general, se estableció una estructura similar para el tema de preservación digital distribuida (PDD) como estrategia de la preservación digital para conocer su definición conceptual y funcional, antecedentes, modelos de redes distribuidas, tecnología aplicada y políticas de organización. De igual forma conocer las técnicas y estrategias aplicadas desde el punto de vista de sistemas informáticos, funciones distribuidas, organización de participantes, estándares y marcos normativos. Con la revisión de literatura de la preservación digital general y de la preservación digital distribuida, se listaron los elementos funcionales de modelos de preservación digital detallando sus diferencias a nivel conceptual, técnico funcional y organizacional en los elementos que los integran para de esta forma consolidar un marco argumental de elementos funcionales, tecnología, organización y políticas existentes para la estimación de las herramientas adecuadas para

recolectar datos para el análisis de sistemas funcionales, organización de redes de PDD y de instituciones de educación superior específicas.

- c) Sobre el tema de México en la gestión de documentos digitales de investigación y su relación con la preservación digital, se revisó literatura para sentar un panorama y contexto actual de la gestión digital, instituciones referentes y marcos legislativos que atañen al ámbito de archivos y preservación digital del país. A modo de panorámica general se estudió el uso de herramientas de gestión de colecciones y archivos digitales en entornos de educación superior pública e investigación, prestando atención a bibliotecas digitales (BiDis), repositorios institucionales y tratamientos con fines de preservación digital en México. Para este contexto se consultó la literatura referente a políticas de archivos, gestión digital, educación universitaria, instituciones de influencia en el sector y el papel que juegan los distintos consejos de apoyo a la educación, ciencia y tecnología. Esta revisión nos ayudó a definir los marcos y contextos en el ámbito de los archivos en México, así como a detectar los distintos actores, instituciones y ejes de relevancia en la gestión de herramientas digitales para el manejo, mantenimiento y resguardo de documentos con fines de investigación en IES-MX.

### **1.5.2 Estudio, selección y análisis de marcos temáticos**

Para el desarrollo de este procedimiento metodológico se utilizó la triangulación cualitativa (Denzin y Lincoln, 1998) para tratar la validez y confiabilidad de tres marcos de interés en el estudio. En este se incluyeron diversas acciones y técnicas que tienen como finalidad el análisis de cada uno de los marcos de enfoque temático definidos para nuestra investigación. Los marcos de enfoque temáticos analizados se integraron por: a) información y elementos funcionales de los sistemas de PDD, b) coordinación, organización y gobernanza de casos de redes y modelos PDD, c) necesidades y nivel de preservación digital en IES-MX. Cabe señalar que este procedimiento de estudio y análisis incluyó de manera reiterada y particular la revisión de literatura para redefinir los parámetros adecuados de elementos, conceptos y metodologías funcionales útiles para la creación de indicadores y herramientas de análisis de los temas mencionados.

- a) Para el estudio y análisis de información funcional de los sistemas de PDD se empleó la técnica de revisión de literatura para la obtención de indicadores especializados y la observación comparativa de funciones técnicas. Para la revista de literatura especializada en PDD se recurrió al uso de diferentes recursos informativos como sitios web oficiales de los proveedores de los sistemas PDD, así como de manuales de uso y publicaciones sobre su aplicación. De manera paralela se revisó literatura específica sobre el análisis comparativo de sistemas de PDD para la obtención de indicadores y terminologías especializadas en la comparación de sistemas de preservación digital distribuida. Para la realización de esta técnica de observación comparativa se diseñó una herramienta de fichaje o tabla de indicadores que organizó la información y principales funciones de los sistemas de preservación digital distribuida a nivel práctico. El objetivo primario de esta, fue el describir su perspectiva de uso y prestaciones específicas a modo de contraste de elementos y funciones indispensables en las estrategias distribuidas de preservación desde los puntos de vista de los sistemas estudiados. Para la elección de los sistemas que fueron parte de este estudio y análisis, se seleccionaron de manera indistinta solo el software y protocolos de preservación definido e identificado funcionalmente y de manera específica como sistemas de preservación digital distribuida. De esta forma se dejaron del estudio a otros sistemas que ofrecen preservación digital con opción a integrar *micro servicios* de distribución o control de versiones. Los sistemas fueron analizados en una visión de contraste como elementos funcionales y conceptuales, aclarando que no se realizó una comparación de tipo técnica avanzada (programación de código o instalaciones), ya que esta tipología no estaba planteado en los alcances y posibilidades de esta investigación.
- b) Para el estudio y análisis del tema de coordinación, organización y gobernanza de casos de redes y modelos PDD se hizo efectivo el uso de las técnicas de revisión de bibliografía, el cuestionario test y la entrevista. En el caso específico de la revista de literatura nos interesó recoger e identificar la información especializada en organización y gobernanza de redes de trabajo, roles de trabajo y coordinación en redes de preservación digital distribuida, así como de modelos de redes activas de preservación digital distribuida. La finalidad de la revisión de literatura respondió en un primer momento a la identificación de las principales actividades e indicadores de organización, coordinación y gobernanza de redes distribuidas de trabajo que puedan ser sopesados y valorados de forma voluntaria por un grupo de expertos participantes de redes de PDD. En este sentido la identificación

de tales indicadores nos ayudó a definir un esquema o clasificación de los mismos para la creación de un cuestionario especializado que fue aplicado a modo de entrevista en usuarios clave y voluntarios anónimos participantes de redes activas de PDD. A partir de esta aplicación, se analizaron por un lado las respuestas de validación de indicadores y actividades de organización por segmentos, así mismo las prácticas, opiniones y valoraciones sobre la organización participativa en redes de preservación digital distribuida.

- c) Para el análisis de necesidades y nivel de preservación digital en IES-MX específicas, se realizaron diversas actividades y técnicas dentro de un marco teórico - práctico de acciones recomendadas en preservación digital, las cuales nos permitieron recoger datos sobre la situación y nivel de uso de la preservación digital en universos puntuales de instituciones de educación superior en México. Para tal motivo se llevó a cabo la revisión de literatura específica de recomendaciones y buenas prácticas en preservación digital propuestas por los organismos referentes en el tema a nivel internacional, y se prestó atención puntual al caso de las recomendaciones y niveles de preservación digital propuestos por National Digital Stewardship Alliance (NDSA) y Library of Congress. A partir de esta revisión, se procedió a adecuar (traducir), clasificar y seleccionar los indicadores necesarios para el diseño de una herramienta de test cuestionario que facilitó la aplicación y recolección de datos de las instituciones a estudiar. Sobre la selección de IES-MX que formaron parte del estudio, es importante mencionar que no se realizó ningún procedimiento de definición de muestras proporcionales, ya que el estudio fue aplicado a la población total de IES-MX pertenecientes al subsistema federal y estatal, que de acuerdo a la clasificación de la Secretaría de Educación Pública en México se compone de por 43 IES (9 IES federales y 34 IES estatales). El estudio pretendió obtener una serie de alcances cualitativos globales y sectoriales sobre las recomendaciones propuestas en la metodología de niveles de preservación digital NDSA con la finalidad de generar y situar mediante una perspectiva porcentual a las actividades realizadas por cada institución en grupos de correspondencia de niveles de acuerdo a las recomendaciones de alcance y situación. Para el análisis de resultados del estudio se empleó la técnica de observación comparativa por alcance de niveles con un enfoque cuantitativo sólo para usos de proyección porcentual. Los resultados se proyectaron en matrices de relación porcentual sobre las diferentes actividades, categorías y niveles - acción de preservación

digital alcanzados por cada una de las IES-MX participantes. Posteriormente se realizó de manera cualitativa un contraste diferencial de resultados con diferentes enfoques de interés como los objetivos de recomendaciones de buenas prácticas de niveles de preservación digital (NDSA), PEST, infraestructuras y tipologías de IES por segmentos globales, generales y por subsistema. La finalidad del estudio y análisis fue obtener un panorama general sobre el alcance de aplicación y conocimiento de las prácticas de preservación digital efectuadas para definir un escenario de necesidades y estado ante actividades de preservación digital.

### **1.5.3 Definición, diseño conceptual y propuesta**

Una vez establecido el ámbito argumental de contextos de referencia, funciones prácticas y estimación de necesidades específicas de preservación digital, el siguiente procedimiento integró una serie de actividades basadas en el método deductivo indirecto de tres los marcos de estudio (PD, PDD y nivel de preservación de IES-MX) para la definición y concreción de la propuesta de modelo de preservación digital distribuida de instituciones de educación superior en México (PDDIM). Las actividades fueron las siguientes: a) elementos conceptuales y funcionales para modelo PDDIM, b) composición y organización funcional de casos de redes de PDD para modelo PDDIM, c) necesidades, compatibilidad y lineamientos para modelo PDDIM y d) integración e implementación de modelo PDDIM.

- a) Esta actividad del procedimiento se centró en la técnica de observación para revisar y extraer un esquema de los elementos funcionales y conceptuales esenciales para el modelo PDDIM. Para tal motivo se refirió y esquematizó una propuesta de composición gráfica con los elementos funcionales de la preservación digital distribuida a partir de la revisión estudio del modelo de referencia, modelo variable y los protocolos de sistemas de PDD. El objetivo fue relacionar y simplificar gráficamente en un modelo genérico estructural, los elementos conceptuales y técnicos que participan en la operatividad funcional de preservación digital distribuida.
- b) Con ayuda del modelo genérico estructural de visión conceptual y funcional, en la siguiente actividad del procedimiento se contrastaron las composiciones de organización de casos internacionales de modelos de redes PDDIM. Para ello se llevó a cabo la

revisión de la composición de los esquemas funcionales y diagramas de los modelos de PDD de la red CARINIANA, National Danish BitRepository y CSUSC en MetaArchive Cooperative a nivel de protocolo, sistema y organización funcional. Mediante una matriz comparativa FODA (fortalezas, oportunidades, desventajas y amenazas) se aplicó un análisis comparativo para extraer buenas prácticas de organización y composición funcional con ayuda de la apreciación técnica y observación de acciones, antecedentes y decisiones operativas de las mismas.

- c) Este bloque de actividades concatenadas se inició con la revista de la concentración de resultados del estudio a IES-MX para definir un cuadro descriptivo de necesidades específicas de PD y PDD en IES-MX que sintetizara los escenarios más próximos de necesidades de acuerdo a las prácticas prioritarias por cubrir. Posteriormente y una vez analizadas las visiones de contraste FODA de prácticas funcionales y de organización de casos de modelos de redes PDD, se describió un esquema de compatibilidad en las necesidades de IES-MX. De manera continuada se desarrolló un contraste de los escenarios de lineamientos, políticas y responsabilidades institucionales de principales agentes de influencia en México para el sector de IES-MX ante su cuadro de necesidades y compatibilidades para el modelo PDDIM.
  
- d) El procedimiento de integración e implementación de modelo PDDIM tomó como base los bloques de análisis, contrastes y visiones de necesidades, para describir una propuesta de contextos metodológicos progresivos que integran las necesidades, alcances y compatibilidad de IES-MX para un programa de preservación digital distribuida. Complementario a esto, se creó un esquema de ámbitos de aplicación (geográfico, político e institucional), así como de estrategias de alcance de actividades, estructura de las instituciones integrantes y los esquemas de coordinación y participación en el modelo PDDIM. Se trabajó en una propuesta de representación integral de prestaciones funcionales, roles de participación y estructuras de relaciones clave para concretar una propuesta integral de modelo de preservación digital distribuida IES México (PDDIM) y una proyección parcial de implementación institucional del modelo.

## **1.6 limitaciones de metodología para estudio e investigación**

Dado que nuestra investigación se orienta de forma general al tema la preservación digital y su estrategia distribuida como aplicación dentro de un entorno de instituciones de educación superior, es importante reiterar que aunque esta incluye temáticas especializadas de ámbito técnico o tecnológico como la programación de base de datos, instalaciones de software, estructuras de metadatos, protocolos de redes, etc. Nuestra investigación se limita a tratar estos temas desde un punto de vista general (no técnico) para conocer sus usos y funcionamiento como referencia descriptiva de cara a los usuarios, administradores e instituciones. Las preguntas de investigación de nuestro estudio contemplaron el conocer, estudiar y analizar de forma exploratoria y descriptiva los segmentos temáticos específicos de las funciones de los sistemas de preservación digital distribuida, la organización, gobernanza y coordinación de redes de preservación digital distribuida, así como el nivel de preservación digital aplicado en la gestión documental de instituciones de educación superior en México; es importante mencionar que estos cuentan con algunas limitaciones para su desarrollo y que son descritos a continuación.

El análisis de identificación de funciones de sistemas informáticos de preservación digital solo ha tomado en cuenta los tipos de software o protocolos de preservación digital distribuida que se identificaron de forma literal con esa función de manera clara y publica (sitios web, guías y literatura referente). De esta manera se excluyó del estudio a los sistemas o protocolos que se identificaron como funciones centralizadas de preservación digital. Así tal, este estudio no incluyó análisis comparativos experimentales de tipo técnico sobre funciones avanzadas, programación o pruebas de instalación, ya que algunos de estos sistemas requieren de infraestructuras tecnológicas con costes elevados en su implementación. El límite de este estudio se refirió a la identificación informativa de los sistemas y protocolos de preservación digital distribuida a nivel de usos y funciones principales como herramientas de archivo distribuido.

Para la segunda tipología de análisis referente a de organización, colaboración y coordinación de redes de preservación digital, nuestro estudio se limitó a conocer opiniones y validar el entendimiento de un grupo de indicadores específicos de organización a manera de

cuestionario y entrevista. El estudio se definió para participantes de tres redes de preservación digital distribuida, de las cuales se requirió como mínimo la participación de un gestor, administrador o colaborador de un nodo de la red de preservación digital distribuida. La identificación de procedencia de su respuesta se mantuvo como anónima en el análisis de resultados para evitar de esta manera la exposición específica del participante y de la institución en su aportación, para crear un ámbito general de apreciación y opinión de expertos sobre las mejores prácticas para validar la importancia y uso de los indicadores sugeridos en la organización de las redes de PDD.

En lo que refiere al tercer tipo de análisis de necesidades y nivel de preservación digital en instituciones de educación superior en México, el estudio se enfocó a analizar el universo total de instituciones públicas del subsistema federal y estatal definidas por la Secretaría de Educación Pública en México, por lo que no se tomaron muestras representativas para el mismo. En esta línea debemos mencionar que la matriz de recomendaciones internacionales (NDSA) tiene un carácter puro de auto evaluación, lo que represento una limitación en nuestra aplicación a distancia. Para lograr una aplicación transparente y evitar el sesgo de una auto evaluación conveniente a distancia, se determinó crear un protocolo de contacto reiterado para el envío y control de una herramienta de recolección de datos que utilizara todos los indicadores de las recomendaciones en la matriz a modo de cuestionario de actividades por categorías únicamente, para de esta manera no incidir en la inclinación a los niveles más altos de actividades de preservación digital a propósito de los intereses de los participantes y obtener una mayor transparencia de aplicación de actividades que pudieran ser medidas posteriormente con su alcance de niveles.

Atendiendo a la emergente y escasa posición que mantiene la preservación digital distribuida a nivel conceptual y de literatura, nuestra investigación se ha centrado a la utilización de los términos mayores y de soporte que han sido formalizados en el campo de la preservación digital distribuida por investigadores como Schultz y Zierau (2013). A partir de estos términos se construyó un plano argumental y complementario para atender de manera mas amplia y completa el tema en nuestra investigación.



## **Capítulo 2 - Marco de referencia y contexto general de la preservación digital**



## **2.1 La preservación digital y la conservación**

La decisión de conservar cualquier cosa para poder utilizarla en el futuro plantea diversos caminos. Uno de los más frecuentados es el de usar poco o no usar lo que se conserva para que se mantenga seguro y como nuevo el mayor tiempo posible. Por otro lado, encontramos que se puede conservar algo si constantemente se revisa su estado y se le da el mantenimiento necesario. Un tercer camino es el utilizar las cosas con conciencia, teniendo siempre presente que alguien más lo usara y por tanto, se debe cuidar y procurar para dejarlo como lo encontramos. Un cuarto camino es el de adquirir una copia idéntica de las cosas para usar una y guardar la otra en otro sitio como conserva. En estos caminos encontramos ciertas actividades en coincidencia como lo son: el aseguramiento de protección, el mantenimiento activo y preventivo, la replicación, y la conciencia de procurar la integridad de las cosas de cara a los usuarios futuros. Todas ellas pueden definir una actividad a nivel conceptual: la preservación.

De manera general la preservación es una actividad que se compone de acciones y efectos, ya que en su contexto activo se define como el procurar, proteger, conservar o poner en seguro algo sobre un daño o peligro. Circunstancialmente su definición general se implica con su provenir del latín *Praeservare*, donde el prefijo *Prae-* significa *antes de o delante de* y el verbo *Servare* que significa *guardar o salvar*. La Real Academia Española define la palabra preservación como: “Proteger, resguardar anticipadamente a una persona, animal o cosa de algún daño o peligro”. Por tanto, la preservación aborda su referencia integral como la actividad de salvaguardar de manera anticipada y presente a algo con mira al futuro.

### **2.1.1 Definición y concepto de preservación digital**

En el contexto digital, la preservación integra literalmente elementos técnicos que transponen su actividad como específica en conjunto con distintos conceptos relacionados de origen anglosajón como lo son: digitisation (digitalización), access (acceso), digital records (registros digitales), born digital (origen digital), migration (migración), refreshing (refresco), authenticity (autenticidad), digital publications (publicaciones digitales), metadata (metadatos), etc. Todos estos conceptos técnicos refieren a una serie de actividades que confluyen en un mismo escenario circunstancial digital sobre parámetros de protección, salvaguarda y

mantenimiento organizado de objetos digitales de cara al futuro. Es por ello que el concepto de preservación digital se compone tanto de actividades técnicas como de elementos económicos, legales y de organización que definen un amplio rango de actividades para el mantenimiento de recursos. Todos ellos actúan en un entorno de organización con la influencia de factores externos e internos cambiantes como la obsolescencia tecnológica, los sistemas financieros o marcos legales de operación. Su actuación busca objetivamente el mantener un marco de integridad y acceso permanente a los materiales digitales.

A modo referencial compilamos una complejo de definiciones desde paradigmas especializados en el tema para describir de manera más amplia el campo de acción de la preservación digital.

La *digital preservation coalition (DPC)*<sup>5</sup> refiere a la preservación digital como el manejo de toda una serie de actividades necesarias para asegurar el acceso continuo a materiales digitales por el tiempo que sea necesario. De manera general se refiere a todas las acciones requeridas para mantener el acceso a materiales digitales más allá de los límites de las fallas de medios o de cambios tecnológicos (obsolescencia tecnológica). Los planteamientos de tiempo de resguardo y acceso a materiales digitales son parte fundamental en la línea de vida de la preservación digital.

Tabla 2. Plazos en preservación digital a nivel conceptual DPC.

Preservación a corto plazo (short-term preservation)	Acceso por un periodo definido a los materiales digitales hasta que sean inaccesibles tecnológicamente.
Preservación a mediano plazo (medium-term preservation)	Acceso continuo pero no indefinido, a materiales digitales, más allá de cambios tecnológicos.
Preservación a largo plazo (long-term preservation)	Acceso continuo e indefinido al material digital o al menos a la información que contienen.

De acuerdo con Termens (2013) el término preservación digital debe ser unificado y a su vez diferenciado de los conceptos de conservación y seguridad informática. La conservación pone énfasis en establecer unas condiciones ambientales, de almacenamiento y uso que permitan

<sup>5</sup> Digital preservation coalition es una empresa británica sin fines de lucro que tiene como objeto preservar recursos digitales a nivel internacional para garantizar la memoria digital mundial. [www.dpconline.org/](http://www.dpconline.org/)

conservar un documento en las mejores condiciones posibles de estado físico. Las técnicas de conservación que son empleadas principalmente en documentos analógicos son también de aplicación en los documentos digitales. La seguridad informática se encarga de establecer políticas para análisis, detección y prevención de posibles riesgos de tipo informático que puedan sufrir los datos de origen informático estableciendo procedimientos de copias de seguridad, control y autorización de acceso, así como de análisis y detecciones generales de fiabilidad en el software y hardware utilizado (pg.16,17).

Conway (2010) hace referencia a la necesidad de distinguir de los conceptos de digitalización para preservación y de preservación digital. Donde la digitalización para preservación tiene como objeto crear nuevos productos digitales con valor, y por su parte la preservación digital tiene como objeto proteger el valor de esos productos digitales o colecciones.

Joint Information System Committee (JISC)<sup>6</sup> define a la preservación digital como un conjunto de procesos y actividades para asegurar a largo plazo, el almacenamiento sostenible, el acceso y la interpretación de la información digital. El JISC tiene como objetivo primario el garantizar la disponibilidad de materiales digitales valiosos que están en línea a futuras generaciones de estudiosos, investigadores y usuarios en general.

Autores como Webb (2003) plantean que la preservación digital consiste en procesos dirigidos a asegurar la accesibilidad continua del material digital. Para hacer posible esto es necesario encontrar la forma de representar nuevamente lo que originalmente fue presentado (como objeto digital) a los usuarios con la ayuda de la combinación de herramientas actualizadas de hardware y software. Para ello se deberá manejar y entender los objetos digitales a cuatro niveles: a) Como un fenómeno físico, b) Como codificaciones lógicas, c) Como objetos conceptuales que tienen significado para los humanos, d) Como grupos de elementos esenciales que deben ser preservados para que los usuarios futuros puedan tener o acceder a la esencia de esos objetos digitales.

Por su parte Hunter (2000) divide en tres términos primarios la definición de preservación digital. El primer término es el de preservación, el cual abarca una amplia variedad de

---

<sup>6</sup> JISC es un programa y organismo público no departamental británico, que tiene como objetivo el apoyar a distintos niveles de educación media, superior y de investigación en el uso, aprendizaje, docencia, administración e investigación en tecnologías de la información y comunicaciones.

actividades interrelacionadas que están designadas a prolongar la vida útil de libros, archivos, manuscritos, artefactos. A grandes términos abarca la protección, estabilización y tratamiento de documentos. La preservación es uno de los tres núcleos funcionales de los archivistas, los otros dos son la identificación y el uso. El segundo término es el de digital, que lo describe como estados discretos sin opción de cambio a diferencia del término análogo que mantiene estados variables. En términos de informática lo digital es combinado con binario, donde se emplean dos estados discretos de números. El tercer término es el de información, que se considera que es un punto de progreso entre los datos, documentos y registros para convertirse en archivos. Al dato lo considera como el contenido de hechos que deseamos procesar, la información como datos que se reciben o se comunican, el documento como la información dentro de un contexto, el registro como un documento preservado y finalmente al archivo como el registro preservado por su valor duradero.

En el reporte técnico de RAND Europe (2007) preparado por *Koninklijke Bibliotheek*<sup>7</sup>, se propone dar énfasis en tres áreas estratégicas que definen la labor y análisis de la preservación digital. La primera es la comunicación escolar que incluye todo lo que se produce en entornos escolares y la manera en la que se usa, se accede y se disemina. La segunda son las editoriales y sus modelos de negocio, la composición de su industria, los derechos de autor y el horizonte de temporalidad sobre el uso de sus publicaciones. La tercera es la preservación, que incluye temas legales o regulares tanto en aspectos técnicos, organizacionales y financieros. Estos temas envuelven su actividad y plantean diferentes cuestiones relativas a conocer: ¿cómo se debe hacer la preservación?, ¿cómo afectan las nuevas formas de materiales digitales?, ¿quién debe realizar dicha actividad? y ¿cómo debe pagarse este trabajo en los diversos modelos de aplicación?.

El espectro de acción y definiciones de la preservación digital atañe a varios entornos indispensables para su realización como lo son de tipo técnicos, de seguridad, legales y financiero. Aunado en estos se interpreta como prioridad la concientización cultural sobre la importancia de preservar el patrimonio digital a nivel institucional y personal.

La UNESCO (2003) a través de la carta sobre la preservación del patrimonio digital potencia paulatinamente esta agenda de desarrollo digital como referente e impulsor e las necesidades

---

<sup>7</sup> Biblioteca Nacional de los Países Bajos

tanto de acceso, vigilancia, acción y continuidad; como de estrategias, políticas, protección, promoción, funciones y alianzas articuladas desde los distintos gobiernos y naciones para la preservación del patrimonio digital como compromiso internacional de participación. Mike Ashenfelder (Library of Congress) acentúa que la preservación digital es un tema familiar entre los líderes mundiales de instituciones culturales, la cual tiene más de diez años de trabajo de éxito preservando archivos y colecciones digitales. En el público en general sigue existiendo una inconsciencia sobre qué es la preservación digital o el archivo personal y el por qué debería ser importante (Hawkins et al., 2013). Esta situación nos lleva a considerar un amplio abanico de necesidades en la implementación para fortalecer una cultura general sobre la preservación digital de patrimonios digitales. Las instituciones de gobierno, bibliotecas y archivo en este sentido, llevan la delantera tanto en implementación, desarrollo y transferencia de conocimiento en todos los niveles sobre el quehacer e importancia de una gestión transversal y específica de la preservación digital. Uno de los retos de la preservación digital a nivel institucional es el de construir de manera sólida un compromiso de permanencia de un plan de preservación activo y por ende de establecer un marco de trabajo sostenible económicamente al largo plazo.

Algunos autores consideran a la preservación digital como un cambio de paradigma que conlleva a una serie de definiciones adaptables a su actividad de lo físico a lo digital, ya que se está pasando de una antigua era a una nueva era de gestión de información digital a largo plazo. Harvey (2011) reitera que para entender o discutir sobre la preservación en la era digital hay que poner en consideración dos términos enlazados como lo son: el cambio paradigmático de la preservación y las definiciones de los términos, ya que muchas de las actividades realizadas en épocas de tratamiento de información pre-digital (procurando materiales físicos) siguen permeando en la actualidad. Hoy no solo es necesario contar con una nueva línea de habilidades ante la preservación, si no también es necesario redefinir el campo de la preservación digital en términos de sus usos y así poder describir sus actividades de especialización, ya que el término de preservación digital acoge a diferentes modelos de trabajo según áreas y asuntos específicos como los datos geo espaciales o datos de ciencias sociales.

Conway (2000) menciona años atrás, que es necesaria una nueva definición de la preservación digital en términos de transformación, sugiriendo un consenso sobre el desarrollar un

conjunto de principios fundamentales que deberían guiar el manejo de los recursos disponibles en un maduro programa de preservación ya que estos han sido desarrollados en un mundo análogo y persisten en el mundo digital. Estos términos en esencia definen las prioridades para extender la vida útil de los recursos de información. Los conceptos necesarios son la longevidad, la selección, la integridad y la accesibilidad según las necesidades específicas de los usuarios y ámbitos.

Thurston (2012) define y reitera la importancia de la preservación digital desde el ámbito de acción gubernamental al largo plazo, donde la conservación de documentos digitales depende incondicionalmente de la combinación de software y hardware en constante cambio. Si se desea que los archivos digitales que emplean los gobiernos en la actualidad sean accesibles, auténticos y utilizables a largo plazo, estos deben cumplir a conciencia clara con los 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?.

La gran mayoría de los autores apuntan a una especialización del término de preservación digital orientado a especificar y particularizar su función aplicada a las necesidades de usuarios, es por ello que abogan por una mayor profesionalización de la disciplina a orden de establecer el nuevo paradigma especializado a la preservación de los objetos y materiales digitales esenciales para conservación del conocimiento humano nacido en formato digital. Algunas propuestas de enfoque como las de Zierau (2011), proponen que para entender la preservación digital a nivel activo debe distinguirse en dos segmentos: la preservación de los bits (bits preservation) y la preservación funcional de los bits (funcional preservation). La primera refiere a todas las actividades requeridas para asegurar que el bit-stream (cadena de bits) permanece legible e intacto, y la funcional debe asegurarse que los bits permanecen comprensibles y utilizables de acuerdo con los fines de preservación (pg.10). Acorde a las definiciones de preservación de bits y funcional, podemos integrar de igual forma la definición de preservación lógica, donde las prioridades de cuidado se enfocan a los archivos digitales (files), asegurando su accesibilidad y entendimiento a pesar de la obsolescencia de la tecnología. Con toda esta amalgama de definiciones podemos referir que la preservación digital es una disciplina activa que integra una serie de acciones profesionales concatenadas a nivel tecnológico operativo, de organización humana y factores económicos de sostenibilidad

orientados a mantener el acceso a materiales digitales a nivel de bits y archivo de manera íntegra, segura y representable de manera permanente al largo plazo.

La gran cantidad de desarrollos y proyectos alrededor de la disciplina de preservación digital abordan herramientas especializadas y modelos variables para su implementación. Desde distintos ámbitos se insiste en la necesidad de definir planes específicos y pruebas piloto de preservación digital que marquen un camino a seguir y de manera deseable se compartan las experiencias de los mismos para enriquecer la experiencia a medida de los plazos cumplidos de preservación en sus expectativas. Un ejemplo es el caso de algunas editoriales que mantienen diferentes estrategias de pago y privadas (private networks) a la vez empleando diferentes modelos de preservación digital para atender y prevenir la custodia de sus materiales digitales, donde la tendencia apunta a simplificar procesos y mantener una expectativa alta sobre los planes y estrategias implementados. Las estrategias actuales a niveles no profesionales de preservación digital tal como lo menciona Mike Ashenfelder, en la mayoría de ellas sólo se cubren necesidades de respaldo de información y se trabajan bajo métodos store and ignore (S+I) o almacenar e ignorar, donde se olvida que todos los materiales de origen digital como lo son fotografías, vídeos o emails son completamente dependientes de hardware y software para existir (Hawkins et al., 2013).

Otro enfoque complementario que une actividades de orden técnico con las de definición de políticas y prácticas de preservación digital a corto o largo plazo son las que proponen Hernández y García (2013), las cuales incluyen aplicar políticas de verificación de integridad y copias de seguridad tratando de usar formatos estables de ficheros que aseguren la fiabilidad los medios de almacenamiento. Identificar los datos con valor para el largo plazo con al ayuda de la definición de políticas de depuración, así como de la asignación de identificadores persistentes y la de identificación de posibles datos sensibles (pg. 262).

Para poner en perspectiva el consenso necesario para implementar una cultura de la preservación digital en diferentes niveles y dimensión de gestión, definimos tres niveles de actuación donde la preservación digital tiene participación activa. En primer lugar a nivel institucional, en segundo a nivel empresarial y por último a nivel particular o doméstico. Todos los niveles de acercamiento requieren de diferentes necesidades a nivel organizacional, tecnológica, económica, legal y normativa como se puede apreciar en la tabla 3. Por lo que

una cultura general de preservación digital al día de hoy es una solución de difícil homogeneización, ya que tal como aseguran algunos autores los datos digitales se generan en diferentes formas y tamaños, por lo que a veces son grandes y homogéneos, pero por otro lado pueden ser pequeños y heterogéneos, lo que diversifica su gestión en diferentes dimensiones de su giro de producción digital.

Tabla 3. Dimensiones y necesidades en preservación digital.

<b>Dimensiones y necesidades PD</b>	<i>Datos</i>	<i>Tecnológica</i>	<i>Económica</i>	<i>Legales</i>	<i>Normativas</i>	<i>Organización</i>
<i>Institucional</i>	Grandes cantidades y tipos de datos e información	Altas necesidades de tecnología informática	Alto nivel de inversión y de mantenimiento a largo plazo	Información propiedad de la institución y de externos.	Alta necesidad de normalización de información	Grandes grupos y personal
<i>Empresarial</i>	Medianas y grandes cantidades y tipos de datos e información	Altas necesidades de tecnología informática	Alto nivel de inversión y de mantenimiento a largo plazo	Información propiedad de empresa y de externos	Alta necesidad de normalización de información	Medianos grupos y personal
<i>Particular</i>	Medianas y bajas cantidades de datos e información	Bajas necesidades de tecnología informática	Mediano nivel de inversión y de mantenimiento a largo plazo	Información propiedad particular	Baja necesidad de normalización de información	Pequeños grupos particulares

Sobre una gestión de dimensiones y necesidades tan diversas en formatos, descripciones, contenidos y cantidades, la preservación digital suma una serie de retos a los que debe dar cara para lograr su objetivo. Los retos fundamentales a destacar son el preservar la integridad de los flujos o cadenas de datos, preservar los medios para interpretar los flujos o cadenas de datos, preservar los medios por los cuales se han experimentado los recursos. En la propuesta de Deegan y Tanner (2006) se definen como retos el preservar la integridad de los objetos digitales, del hardware y software. Esta consideración temprana sobre preservar los medios (hardware y software) ha sido superada por otros factores como la obsolescencia tecnológica, lo cual a nivel de hardware o software ha sido modificada por técnicas de refresco, migración y análisis forense digital. La preservación de integridad también ha sido complementada por técnicas para analizar la autenticidad e integridad como es el caso de la suma de comprobación o *checksum*, que suponen incrementar la cobertura de los retos principales de la preservación digital.

Hasta este punto, se ha explorado una serie de definiciones de orden teórico y técnico a nivel conceptual de manera general a la disciplina de la preservación digital. A continuación integramos un panorama práctico de modelos, técnicas y marcos de trabajo bajo ejemplos de implementación en programas de preservación digital que nos otorgan detalles complementarios sobre los elementos legales, económicos e institucionales deben ser tomados en cuenta para describir la línea de actividades que integran su actividad. Las líneas de trabajo y actividades son traducidas en modelos de acciones, políticas y estrategias concretas que otorgan fiabilidad a la sostenibilidad a los objetivos de programas de acceso futuro a los recursos y objetos digitales.

### **2.1.2 Modelos de archivo en preservación digital**

Un modelo puede definirse como un arquetipo, una representación, esquema o figura que funciona como punto de referencia. Un modelo sirve como punto de partida para explorar y desarrollar posibilidades de creación. Cuando hablamos de modelos en preservación digital aludimos principalmente a modelos de archivo digital, ya que estos refieren principalmente a un sistema de almacenaje digital y todas las acciones colaterales que este requiere para mantenerse activo. Por tanto los modelos de preservación digital son meta modelos de archivo que sugieren de técnicas, estrategias y acciones puntuales para preservarlo a largo plazo. Existen diversos ejemplos sobre implementación de modelos de preservación digital, sin embargo hay modelos que son la referencia obligada a nivel de recomendaciones conceptuales y estándares internacionales para integrar un modelo técnico de trabajo en los planes de preservación digital. El punto de referencia es el modelo OAIS, y a partir de él existen una serie de creaciones y adaptaciones del modelo referencia en preservación digital que tienen como objetivo común el mantener accesibles e íntegros los objetos digitales a largo plazo como una necesidad inherente de diversas instituciones y organizaciones que aceptaron la responsabilidad de resguardar el archivo de información digital de diversos campos del conocimiento.

### ***El modelo de referencia OAIS (Open Archival Information System)***

Acorde a la definición simplificada de Giaretta (2011), OAIS es un archivo consistente de una organización (o de una parte de ella) donde personas y sistemas han aceptado la responsabilidad de preservar información y hacerla disponible para una comunidad designada. Este compromiso reconoce una serie de responsabilidades definidas como estándar normativo que permiten a un archivo OAIS distinguirse de otros usos del término archivo.

OAIS (Open Archival Information System) tiene como objetivo general el organizar a personas y sistemas en un mismo entorno de trabajo para el manejo de un sistema de información de archivo con el fin de preservar a largo plazo cualquier tipo de objetos (no solo los digitales) para que estén siempre accesibles a una comunidad designada de usuarios. OAIS integra un marco de trabajo bajo conceptos, recomendaciones, normas y requerimientos en el esquema estándar ISO 14721:2003. Este incluye la integración de acciones humanas y elementos informáticos (en el caso de la preservación digital) como abstracciones de un modelo de trabajo. OAIS fue desarrollado en el año 2002 por Consultative Committee for Space Data Systems (CCSDS) para poderse usar en una amplia gama de organizaciones con necesidades de gestionar un archivo digital y de aquellas que principalmente tienen como objetivo el mantenerlo a largo plazo. El modelo OAIS cumple con dos funciones fundamentales: la primera es la de conservar o preservar información y la segunda es la de hacer posible y garantizar el acceso continuo a la misma.

Este modelo no especifica ningún tipo de diseño o implementación ya que las necesidades de implementación son variables a las necesidades de cada comunidad o institución que aplique dicha referencia. El modelo puede considerarse sólo como un tipo de arquitectura sugerida como base para desarrollar una implementación de software a medida o del flujo de trabajo organizado y coherente de un sistema de conservación digital como los repositorios.

El modelo OAIS ha sido ampliamente adoptado por las comunidades, instituciones y organizaciones interesadas en la preservación digital, es por ello que hoy en día se encuentra como referencia obligada para iniciar un esquema de orden y visualización de un plan de preservación digital. Uno de los recursos más usados en las comunidades de preservación

digital es el de los repositorios digitales, a los cuales se les ha integrado a este esquema para que operen su flujo de trabajo de manera adaptada al modelo OAIS.

### ***Propósitos de modelo OAIS***

Acorde al documento de Reference Model for an Open Archival Information System (CCSDS OAIS, 2002-2012), se norma un modelo de trabajo bajo una serie de responsabilidades de la organización que acuerdan implementar un plan de preservación a largo plazo dentro de un sistema de información de archivo tanto para objetos digitales como físicos. El modelo OAIS establece un conjunto de propósitos, prácticas y recomendaciones como referencia, las cuales incluyen de manera general el proveer los marcos de trabajo y escenarios necesarios que ayuden a entender, analizar e incentivar la actividad de la preservación a largo plazo de información digital. Este proporciona un escenario auxiliar con terminologías y conceptos que describen los procedimientos necesarios para comparar y describir a nivel de arquitectura las estrategias empleadas en la preservación a largo plazo.

El modelo refiere al propósito de entregar un amplio rango de funciones acerca de la gestión de preservar información en un archivo, así mismo reitera un marco de trabajo organizado y coherente acorde a los elementos necesarios para proceder a la preservación de largo plazo. Bel (2012) sugiere que la implementación del modelo OAIS además de atender como propósito primordial el preservar los datos como requisito principal de conservación a largo plazo, debe considerar en igual o mayor importancia que los datos preservados deben ser elegibles para su reutilización después de un período de tiempo no especificado y que normalmente es por más de treinta años. El modelo de referencia tiene como objetivo inherente el proporcionar ejemplos ilustrados sobre buenas prácticas en la materia reflejadas en las responsabilidades, términos y conceptos.

### ***Aplicaciones de modelo OAIS***

En principio el modelo referencia OAIS es aplicable de manera específica a todo tipo de archivo en estado físico o digital de cualquier organización que pretenda y se responsabilice de mantener información disponible a largo plazo. También, es de interés de aplicación en archivos personales que tengan necesidades de adquirir y recuperar información integra a

largo plazo. Es importante aclarar que los conceptos, términos y demás funciones para ayudar a implementar y diseñar esquemas de preservación tomando con base en el modelo OAIS que considera aplicaciones a largo plazo, también considera de manera clara su aplicación en archivos temporales. Esta aplicación temporal surge a causa de los rápido y continuos cambios tanto de comunidades designadas como de tecnología, por lo que el modelo apunta a ser flexible en su adaptación de conceptos y terminología tomando en cuenta necesidades de preservar información en primera instancia de manera temporal y considerando las necesidades futuras de otorgar longevidad indefinida a esa información tratada inicialmente con objetivos de uso temporal.

Tomando en cuenta uno de los propósitos primordiales del modelo que es la comparación de diferentes modelos y las aplicaciones del modelo OAIS, es importante definir los distintos campos de aplicación y modelos de trabajo desarrollados acorde a esta norma para plantear de manera específica los elementos clave en un nivel de comparativo para obtener ejemplos actuales de buenas prácticas en la aplicación y modificaciones del modelo base.

### ***Conceptos Clave y Entidades Funcionales de Modelo OAIS***

Dentro de la norma del modelo OAIS en el apartado 4.2 (CCSDS OAIS, 2002-2012), se describen de manera detallada todos los conceptos funcionales que le integran tanto a nivel de entidades, información, objetos, etc. Para el contexto y necesidades de nuestra definición de conceptos, sintetizamos de manera funcional a modo de grupos a la par del modelo ilustrado para definir una perspectiva práctica de los mismos. Describiremos los conceptos en los cuatro grupos que integran el modelo funcional como: entorno, información, entidades e interacciones.

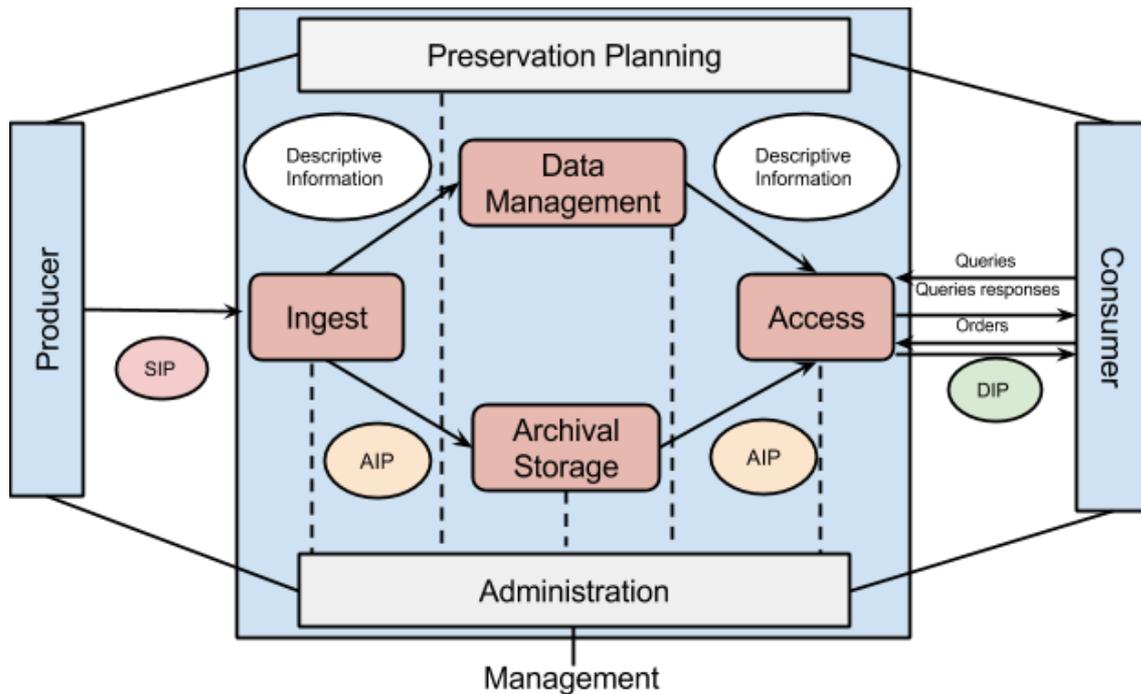


Figura 2. Entidades funcionales del modelo OAIS

### 1. Entorno de Modelo OAIS

Son considerados parte del entorno los elementos productor, consumidor y management ya que existen fuera de OAIS como puntos activos para hacer funcionar el ciclo de funcionamiento.

#### a) Producer (Productor)

En términos objetivos el rol de productor lo ejercen aquellas personas o en su caso sistemas cliente que proveen información a OAIS, es decir que hacen llegar o llevan a cabo la ingesta de información para que esta sea administrada para ser preservada.

#### b) Management (Gestión)

El manejo y gestión es un rol que integra todas las normas o políticas indicadas por una organización o entorno de trabajo implementado en OAIS. Son todas las políticas de gestión y

normas definidas que son una tarea activa de las actividades de administración y que son definidas a modo de tarea previa sin actividades diarias en el proceso.

#### c) Consumer (Consumidor)

Cuando hablamos de consumidor, estos representan personas o sistemas cliente que tienen el rol de usuario y consumidor de información preservada. El consumidor interactúa con el sistema OAIS haciendo diferentes tipos de solicitudes y obteniendo respuestas. En este rol de consumidor tenemos también que definir el término comunidad designada, ya que esta es un tipo de consumidor que tiene como cualidad la de entender la información preservada e incluso puede actuar por momentos con un doble rol tanto de consumidor como productor por el nivel de interacción con el sistema.

## ***2. Información en Modelo OAIS***

La información dentro del sistema OAIS funciona como materia prima, la cual es la principal sustancia que fluye dentro del sistema y por ello que debe ser definida desde sus diferentes funciones a nivel de paquetes de información y descripción de la misma.

#### a) Information Package (IP)

Un paquete de información es un contenedor conceptual de principalmente dos tipos de información: la información de contenido (contenido de datos de objetos) y la información de descripción de preservación (provenir, contexto, referencia, fijación, derechos de acceso). Ambos tipos de información contenidos en el paquete de información integran, describen e identifican a un objeto o representación de información. Estos son entregados, archivados, diseminados y transmitidos entre productores y consumidores. De tal forma un paquete de información obtiene nuevas características en el modelo de acuerdo a su función participativa.

#### b) Submit Information Package (SIP)

Un SIP es el paquete de información presentado por el productor al sistema OAIS como inicio del proceso de ingesta, el cual contiene una personalidad de forma y detalle definida por el productor y que es validada con el sistema. Los SIP incluyen información de contenido e información de descripción de preservación en un nivel básico de identificador.

#### c) Archival Information Package (AIP)

Los AIP pueden considerarse como SIP transformados en una o varias colecciones de AIP, los cuales contienen un paquete completo de información descriptiva de preservación asociada a la información de contenido. Los AIP son más completos a nivel de información que un SIP, ya que contienen información de un objeto con la cantidad y calidad suficiente a nivel de descripción de preservación y de representación de objetos para garantizar su preservación a largo plazo.

#### d) Disseminating Information Package (DIP)

Un DIP es un paquete de información que surge como respuesta de un sistema OAIS a las solicitudes realizadas por un consumidor. El DIP entregado como respuesta puede ser una colección completa de AIPs o solo una parte de ella, y puede tener diversas formas en dependencia de las necesidades de entendimiento del consumidor final del paquete de información.

#### e) Descriptive Information (DI)

Es un grupo de información que consiste en paquetes de información descriptiva que se proveen de las acciones de gestión de datos (data management) para facilitar la búsqueda, recuperación y orden de la información empaquetada y solicitada por un consumidor al sistema OAIS.

### ***3. Entidades funcionales de Modelo OAIS***

El modelo OAIS se compone de seis entidades funcionales, las cuales proporcionan diferentes servicios y funciones mediante una interfaz relacionada de actuación en el modelo.

#### **a) Preservation Planning (Planificación de preservación )**

La entidad planificación de preservación integra actividades prioritarias de monitoreo de los entornos del productor, consumidor y gestión del modelo OAIS mediante distintas funciones y servicios definiendo planes de preservación, estrategias y normativas que garanticen que la información almacenada en OAIS se mantenga accesible y entendible a largo plazo para la comunidad designada y por encima de condiciones de obsolescencia computacional. Dentro de las funciones destacadas de la planificación de preservación se incluye la evaluación de contenidos de archivo, la actualización periódica de la información de archivo, la migración de soportes de archivo, desarrollo de recomendaciones y normativas para el estándar de archivo, reportes y análisis de riesgo de archivo, así como monitorear los cambios en el medio ambiente tecnológico sobre las necesidades de requerimientos de la comunidad designada y el desarrollo de planes de migración o implementación de prototipos de software procurando obtener los resultados objetivos.

#### **b) Administración (Administración)**

Es una entidad que integra una serie de servicios y funciones en las operaciones generales de un sistema de archivo como lo son los acuerdos, negociaciones y envíos de los productores. Realiza actividades de auditoría de los paquetes enviados para que estos correspondan a los estándares necesarios. Otras de las actividades importantes que realiza la administración es la de dar mantenimiento a los sistemas de software y hardware, monitorear, inventariar, actualizar, migrar y reportar el estado del sistema de archivo trabajando en el mantenimiento de las políticas, estándares y necesidades de las peticiones de los usuarios del sistema.

### c) Ingest (Ingesta)

La ingesta es una de las entidades que activan el proceso de un sistema OAIS. Esta entidad incluye funciones y servicios como el de aceptar los SIP de los productores, preparando y asegurando la calidad de los AIP y de que su información descriptiva se establezcan en el sistema OAIS como paso previo a su depósito en el almacenamiento de archivo.

### d) Data Management (Gestión de Datos)

La entidad de gestión de datos, realiza actividades concatenadas de mantenimiento de la información descriptiva y administrativa que asegure el acceso continuo a la población de datos depositados en el sistema. Las funciones que desarrolla son las de administrar las bases de datos (esquemas, definiciones e integridad referencial), actualizaciones de bases de datos y en específico de su información descriptiva o administrativa, así como desarrollar actividades de peticiones regulares para verificar el estado del sistema y generar reportes de estabilidad de los mismos en las respuestas, para así actualizar y gestionar la información necesaria para su recuperación.

### e) Archival Storage (Archivo de Almacenamiento)

La principal función de esta entidad funcional es la de almacenar los paquetes de información, así como mantener y recuperar los paquetes de información en archivo. Las funciones del almacenamiento de archivo más allá de priorizar el almacenamiento se enfoca en jerarquizar con orden los paquetes de información recibidos desde la ingesta y refrescar los medios en los cuales los archivos están almacenados. Lleva a cabo revisiones de rutina sobre errores en los archivos y paquetes de información para acrecentar la capacidad de recuperación de los mismos y mantener su accesibilidad estable.

### f) Access (Acceso)

Esta entidad funcional proporciona el soporte necesario a los clientes sobre la localización, existencia, disponibilidad y descripción de la información almacenada en el sistema. En gran medida lleva a cabo actividades de comunicación constante con los clientes sobre sus

consultas y respuestas a las mismas. El acceso se encarga de generar los paquetes de información diseminados a los clientes finales gestionando los niveles de límite de acceso a la información solicitada y entregando respuestas específicas mediante informes y reportes de las actividades de consulta o solicitud en un ámbito de diseminación.

#### ***4. Interacciones***

##### a) Queries (Peticiones)

Las acciones de peticiones se llevan a cabo en entornos internos y en entornos externos. Dentro del sistema OAIS diferentes entidades como la de gestión de datos, realizan peticiones de información internas. En la entidad funcional de acceso es donde se hacen externas estas acciones de petición, ya que son realizadas por los usuarios.

##### b) Queries responses (Respuestas a Peticiones) y Orders (Ordenes)

Las respuestas a las peticiones son una contra acción realizada en distintos entornos y refieren a entregas de información. Las órdenes son peticiones específicas y objetivas que los usuarios realizan al sistema OAIS de las que obtienen un paquete diseminado de información (DIP).

Después de conocer un panorama general los conceptos, propósitos y entidades funcionales del modelo de referencia OAIS, a continuación definiremos la manera en la que se ha implementado, y adaptado el modelo en diferentes herramientas informáticas y modelos de trabajo en el campo de la preservación digital de archivos.

#### ***Los repositorios digitales y el modelo OAIS***

En este orden de ideas que rodean al modelo OAIS así como a sus adaptaciones en el campo técnico y operativo es necesario poner en perspectiva uno de los sistemas informáticos funcionales dentro de un plan de preservación digital, nos referimos a los repositorios digitales. Gracias a su maduración operativa en la actividad de resguardo y almacenaje digital centralizado, son considerados como una herramienta de colaboración y comunicación de documentos de distribución institucional. Otra visión apunta a que si los repositorios se promocionarán tanto por sus capacidades de preservación y como plataformas de

experimentación de comunicación escolar, nuevas herramientas y estructuras podrían crearse a partir de ellos (Hawkins et al., 2013). Los repositorios son parte fundamental en las estrategias de diseño y adaptación del modelo OAIS a software de preservación digital, ya que funcionan como entes activos en la colaboración técnica y organizacional de modelos y planes de trabajo adaptados a la gestión digital sostenible de instituciones.

La definición de repositorio a nivel de término, proviene del latín *Repositorium*, que significa armario o alacena y donde el DRAE<sup>8</sup> lo define como: “lugar donde se guarda algo”. Por otro lado la palabra repositorio desde un punto de vista funcional puede definirse como lo propone Gozzer (2011), como una aplicación o sistema de almacenaje de contenido susceptible de ser reutilizable.

En este sentido práctico y funcional podríamos decir que repositorio digital es un sistema informático integral que almacena objetos digitales completos (documentos, audio, videos, fotografías, bases de datos, etc.) de manera ordenada y accesible. Este puede alojarse vía un servidor web o de un servidor local. Puede ser de acceso abierto o cerrado mediante un software que gestiona la ingesta, clasificación y mantenimiento de objetos digitales.

En el campo de la preservación digital los repositorios pueden mantener una temática específica o institucional de acuerdo a su contenido, gestión y función. En nuestro estudio nos centraremos principalmente en los repositorios institucionales que gestionan las universidades y sus centros de investigación. De acuerdo a algunos autores la definición de un repositorio institucional puede ser confusa, ya que existen diferentes tipos de repositorios temáticos (especializados en un ámbito científico) que podrían ser considerados como institucionales. Es importante hacer un paréntesis para definir su objetivo diferenciado. Sánchez y Melero (2006) mencionan que un repositorio institucional es definido no tanto por el formato de los documentos constituyentes de su contenido sino por su finalidad. Un repositorio que tiene como finalidad gestionar una colección de documentos con fines de investigación puede ser considerado dentro del rubro de los repositorios institucionales. Secker (2004) defiende la idea de definir a los repositorios como institucionales cuando se refieren al contexto de las publicaciones de investigación que emanan de las instituciones. Un repositorio institucional puede tener diferentes niveles de uso, ya que almacena diversos tipos de materiales de

---

<sup>8</sup> Diccionario de la Real Academia Española <http://www.rae.es/recursos/diccionarios/drae>

investigación como artículos revisados, los e-prints, tesis, informes, ponencias, working papers y otros tipos de documentos (Sánchez y Melero, 2006).

Una definición que concreta acertadamente las visiones anteriores es la de Abadal (2012): “Los repositorios institucionales contienen producción de los miembros de una institución, ya sea una universidad o un centro de investigación. Tienen carácter multidisciplinario, a veces se centran exclusivamente en contenidos científicos (artículos, revistas, tesis, congresos, etc.) pero existen también ejemplos de inclusión de material docente, documentación administrativa, colecciones patrimoniales, etc.”(p. 24).

Como sistema funcional dedicado a la gestión y almacenaje digital, un repositorio digital se constituye de tres partes fundamentales que son un ordenador (hardware) que funciona como servidor y sistema de almacenamiento, de un sistema informático de repositorio (software) que funciona como administrador gestor de las puerta de entrada de objetos digitales, y por último de objetos digitales (ficheros) que pueden ser datos o documentos específicos. En esta línea podemos concretar que un repositorio es en si un sistema centralizado configurado para la gestión de objetos digitales, y por ende un sistema adaptable a actividades de preservación digital desde el punto de vista tecnológico y organizacional bajo la dependencia de actividades planificadas para la prevención de integridad de objetos digitales y el sistema donde se almacenan.

Los repositorios digitales hoy en día implementan diferentes modalidades funcionales. Por un lado existen implementaciones que concentran todo su trabajo en modo centralizado. Otros trabajan en conjunto de una red de recolección distribuida de diversos repositorios, los cuales son concentrados bajo una sola interfaz de repositorio. Estas estructuras de modelos funcionales de repositorios responden a variables en las necesidades de los desarrollos y la experimentación de mejora. En esta línea podemos destacar tres modelos estructurales funcionales de repositorios: modelo centralizado, modelo distribuido y recolector.

El repositorio de modelo centralizado representa a un repositorio con los objetos digitales depositados en un solo servidor central, donde se llevan a cabo todas las actividades primordiales de un repositorio como lo es el almacenamiento de objetos digitales completos, la administración de ficheros y acceso de los mismos mediante una interfaz. Cabe señalar que

el modelo central representa por un lado la integridad institucionalidad, creación específica y la congruencia de trabajo organizado a un mismo fin para facilitar al usuario final una búsqueda en un solo sitio de depósito.

El modelo distribuido de un repositorio es aquel que se define como un modelo dependiente de una red de repositorios enlazados para intercambiar metadatos y objetos digitales de colecciones, para mantener copias extras de objetos digitales en red con la finalidad de disponer de una variedad de colecciones y a su vez de mantenerlos respaldados. Este modelo distribuido en ocasiones se presenta como una estrategia de planificación previa, aunque en ocasiones se debe al crecimiento de una red de unidades o repositorios independientes de una misma organización donde el crecimiento natural es el de la unión inversa. El modelo distribuido de repositorios puede agrupar a diferentes repositorios bajo un buscador único de un sitio web, el cual enlaza a todos los metadatos para presentarlos en una misma capa de acceso. El modelo recolector es una variación del modelo distribuido que funciona como un servicio que recolecta sólo los metadatos mediante un proceso de normalización u optimización para mejorarlos a cara de los proveedores de servicios de repositorios y metadatos.

Dentro del modelo de referencia OAIS un repositorio puede incorporarse en algunas de sus entidades funcionales como lo son el archival storage (almacenamiento de archivo) o el data management (gestión de datos) y así definir una adaptación de un sistema de repositorio con objetivo de preservación digital operativa.

Hay que aclarar que además del término repositorio como archivo digital institucionalizado o formal, existen diferentes acepciones y usos del término que deben ser entendidos en aplicaciones normalizadas o similares. Arreola (2014) las sintetiza con apoyo del ODLIS (Online Dictionary for Library and Information Science, Reitz, s.f.) como biblioteca digital, reservorio digital, repositorio institucional, archivo digital, colección digital y base de datos.

### ***Modelos de preservación digital y adaptaciones de modelo OAIS***

A partir del modelo de archivo referencia y marco de trabajo base OAIS encontramos una serie de implementaciones de modelos de archivo con adaptaciones, variables y tecnología

diferencial en su aplicación. Describiremos las distintas acciones, técnicas, modelos, opciones y arquitectura empleada concretamente en estrategias de preservación digital de modo general con el objetivo de estructurar y visualizar los distintos elementos en un proceso de preservación digital a nivel conceptual.

En preservación digital los modelos de archivo hacen referencia a los sistemas de almacenaje digital en conjunto con todas las acciones y estrategias colaterales que se requieren para mantenerles activos. En todo modelo de archivo y gestión de objetos digitales intervienen dos roles indispensables que son el acceso y la preservación, que como roles divergentes funcionan de forma complementaria con procedimientos de gestión de objetos digitales como la procedencia, la autenticidad e integridad digital que enriquecen su seguimiento, mantenimiento y preservación digital. Estos aspectos pueden ser variables de acuerdo al manejo que una u otra organización realice o en su caso defina como modelo de archivo. (D'Amato, 2012).

Termens (2013) sugiere desde un punto de vista funcional y de enfoque prioritario a cuatro modelos en preservación digital. En primer lugar el modelo centralizado que mantiene su interés primario en las herramientas o aplicaciones informáticas para una gestión de la información y datos a modo central. Por otro lado el modelo de preservación descentralizada que reitera su enfoque en los datos y su distribución como prioridad. Ambos modelos coexisten en dos diferentes modalidades de archivo. El archivo en modo abierto que contempla el acceso y preservación en conjunto como actividades prioritarias, y el modo de archivo cerrado que tiene como prioridad la preservación y es una actividad realizada por separado sobre de los objetos que custodia para dejar el acceso en un segundo plano en modo restringido. Cabe señalar que al hablar de modelos de trabajo en preservación digital no debemos pensar únicamente en resolver problemáticas de aspecto técnico (software, sistemas, aplicaciones, almacenaje, etc.), ya que podemos encontrar barreras de otro tipo como las presentadas a nivel humano, de organización, económico o legal, que pueden ser mayores que las técnicas y podrían dificultar las soluciones de preservación digital y su sustentabilidad al plazo de tiempo necesario.

Una gran mayoría de propuestas iniciales de modelos y estrategias de trabajo con objetivos de preservación digital se enfocan de forma tácita a complementar una estrategia de modelos de

archivo digital o repositorio con modificación e implementación del modelo OAIS en su ecosistema de trabajo. Otras propuestas de modelos de preservación digital presentan características como la flexibilidad y la adaptación de diferentes marcos de trabajo para enriquecer implementaciones prácticas en búsqueda de las mejores soluciones para cada organización. A continuación revisamos una serie de propuestas de opciones estratégicas y modelos de trabajo en preservación digital para ampliar la perspectiva práctica de usos.

Brown (2013) propone ocho opciones de modelos de trabajo para la gestión de un archivo digital con fines de preservación digital como casos prácticos implementados en diferentes organizaciones.

#### *La opción de no hacer nada (To do nothing option)*

Esta opción la define desde la perspectiva de mantener un archivo en una estructura no apropiada, sin inversión de personal y sin un plan de preservación. No hacer nada como opción conlleva a mantener un archivo bajo riesgo de pérdida por la inacción y falta de políticas de preservación a largo plazo.

#### *La opción de repositorio mínimo (The minimal repository option)*

La opción asegura que es posible construir un sistema funcional de preservación digital sin la ayuda de herramientas o sistemas elaborados. Esta sería la opción más realista para una pequeña y mediana organización. La idea es que funcione como un repositorio con objetivos de preservación a largo plazo utilizando eficientemente las herramientas y funciones disponibles en el mismo como lo son la ingesta, metadatos, almacenaje, preservación y acceso. En esta opción se aplicaría una modificación directa del marco de trabajo propuesto en OAIS, donde la potencialidad de las herramientas disponibles se adaptan a un entorno de alcance mínimo necesario.

### *La opción a medida (The bespoke option)*

En definitiva la opción de modelo de trabajo más costosa, ya que refiere a la construcción de un repositorio o sistema de preservación a la medida de los requerimientos de la organización. Esta opción garantiza cubrir todas las necesidades de la misma, pero requiere del trabajo de programadores de software para el desarrollo de un sistema único que cumpla exactamente con lo que una organización necesita. Esta necesidad surge de cuando la oferta de sistemas y software de terceros no cumple con los requerimientos necesarios de una organización.

### *La opción de software de código abierto (The open-source software option)*

Las opciones de herramientas y software de código abierto que existen actualmente en internet son capaces de cubrir las necesidades de creación de repositorios o sistemas de preservación digital de una organización. Muchas de estas opciones de software son gratuitas, sin embargo no por ello significa que la implementación total de un sistema de preservación no acarrea costos implícitos pues en algunas organizaciones será necesario adaptar herramientas o en su caso generar un desarrollo de software adicional que conlleva a un sistema genérico y libre. Una de las ventajas podría ser el uso de los *crowdsourcing* de usuarios del sistema y foros donde se puede obtener apoyo sobre el mantenimiento y adaptaciones del sistema.

### *La opción comercial (The commercial option):*

Esta es una opción segura, cómoda y de bajo riesgo, aunque no por ello libre de un alto costo. La opción de sistemas comerciales por un lado plantean escenarios seguros donde su software es flexible y modificable pues otorgan una alta disposición de soporte funcional. Así tal, las opciones comerciales en ocasiones no cuentan con todas las funciones necesarias para los requerimientos de una organización, lo que significa desarrollos extras en modificaciones y dependencias de propiedad de software.

### *La opción de subcontratada (The outsourced option)*

Es una opción accesible a nivel de implementación por la evasión del costo de inversión en equipos y sistemas para la organización que contrata el servicio. Hoy en día los servicios externos de repositorios seguros y preservación digital representan un mercado latente y en crecimiento pues ofrecen un servicio completo que incluye software propio y técnicas necesarias para mantener archivos replicados, refrescados y accesibles a largo plazo mientras se cumpla una relación responsable con los proveedores del servicio. Esta opción generalmente cuenta con especialistas en la materia quienes evitan que las organizaciones trabajen en planes y protocolos de preservación digital. Sus sistemas de trabajo son de fácil uso e instalación. Como contra se plantea tomar en cuenta los riesgos que acarrea la custodia de la información digital en resguardos externos y en dependencia del tipo de información privada y sensible de la organización.

### *La opción de asociación (The partnership opción)*

La opción de un modelo para colaborar de manera asociada se plantea principalmente cuando varias organizaciones comparten los mismos objetivos y toman la responsabilidad de participar activamente en el desarrollo de servicios compartidos. El mantener un orden entre varias organizaciones para trabajar bajo un mismo objetivo requiere de acuerdos informales, formales o en su caso con entidades legales separadas. Las principales funciones de esta opción de trabajo son las de hospedar y copiar todas o parte de las colecciones de los participantes utilizando distintas tecnologías de colaboración para mantener la información digital compartida segura, auténtica y siempre disponible con distintas réplicas distribuidas. Podríamos decir que mantiene ventajas como las de un club, donde los integrantes comparten buenas prácticas y mejoras continuas.

### *La opción híbrida (The hybrid option)*

En la actualidad esta opción de modelo es la más apropiada por diversos motivos. En primer lugar, por el estado de maduración sobre el cual se encuentran muchos de los servicios, lo que otorga una posibilidad de probar las mejores opciones. En segundo lugar para elevar la seguridad de la información que se resguarda, más cuando las organizaciones (como por

ejemplo editoriales comerciales) necesitan mantener altos rangos de seguridad y evitar margen de error. Hablamos de una opción que mantiene una flexibilidad no solo a la medida de requerimientos, sino más bien considerando situaciones extraordinarias no cubiertas por modelos u opciones.

La propuesta de opciones en modelos de Brown (2013) encierran en gran medida las mejores prácticas que actualmente se desarrollan y se ofrecen tanto a nivel de software como de servicios en el área de la preservación digital, y que de alguna manera en el estado emergente de la preservación digital se podrían considerar como a modo de prueba y previsión de posibles pérdidas al margen de error digital y su evaluación a futuro.

Los autores Dollar y Ashley (2016) definen el estado objetivo al que la gestión de preservación digital debe aspirar y que denominan *Digital Preservation Capability Maturity Model (DPCMM)* o modelo de madurez de preservación digital. Este modelo se soporta en base a 15 componentes que definen diferentes capacidades en la gestión de preservación digital, los cuales se integran por: políticas de preservación digital, estrategia de preservación digital, gobernanza, colaboración, experiencia técnica, tecnología en estándar abierto de formatos neutros, comunidad designada, encuesta de registros electrónicos, ingesta, archivo de almacenamiento, renovación de medios y dispositivos, integridad, seguridad, metadatos de preservación y acceso. Estos componentes definen y sitúan en cinco diferentes etapas evolutivas de trabajo en preservación digital de una organización, mismas que van desde el estado inicial o etapa uno, en el que una organización entra en conciencia de sus prácticas de preservación digital y lleva a cabo actividades cada vez mas comprometidas para poder alcanzar la etapa cinco de maduración en el que desarrolla su capacidades de gestión a un estado sólido y maduro donde se incluyen todas las actividades que van desde capacidades básicas gestión de un archivo digital, hasta las mas optimas de implementación, continuidad y actividad de compromiso en la gestión de sus actividades de preservación digital.

Otra propuesta es la del Modelo de madurez de preservación digital propuesta en el white paper "*digital preservation maturity model*" de Preservica (2014). En este modelo se propone una estructura compuesta por elementos clave a modo de seis niveles, los cuales se integran en tres secciones principales que son: almacenamiento duradero (durable storage), información localizable (findable information) e información utilizable (usable information).

Dentro de la sección de almacenamiento duradero se integran los tres niveles iniciales, en nivel uno corresponde a almacenamiento seguro (safe storage), el nivel dos a la gestión de almacenamiento (storage management) y en nivel tres a la validación de almacenamiento (storage validation). En la sección de información localizable, encontramos al nivel cuatro de organización de información (information organisation) y el nivel cinco de procesos de información (information processes). En la última sección de información utilizable se encuentra el nivel seis que corresponde a la preservación de información (information preservation). En ambos casos de propuestas de modelos de maduración de actividades de preservación digital, encontramos una serie de actividades que deben cumplirse de manera escalonada para poder alcanzar la plenitud de una actividad de preservación digital madura.

Todas estas propuestas de maduración de prácticas provienen de actividades moldeables y adaptables de modelos aplicados a soluciones propias. En algunos casos como el de los grandes grupos editoriales, se ha optado por implementar complejos sistemas informáticos que aseguren la permanencia de su producción digital y además como extra han decidido adherir su producción digital a más servicios externos o de colaboración, ya sea por convicción técnica o comercial, para demostrar a sus clientes su alto compromiso con la seguridad de la preservación de sus datos. Termens (2013) cita como ejemplo la estrategia de la editorial Elsevier que trabaja con cuatro sistemas de preservación digital a la vez. El primero con un sistema propio de preservación digital alojado en un antiguo búnker antiatómico norteamericano. El segundo haciendo uso del servicio e-Depot de la biblioteca de los países bajos. El tercero con la red distribuida y privada de LOCKSS / CLOCKSS. El cuarto usando los servicios privados de Pórtico. Como podemos observar las estrategias y modelos son flexibles a la medida de las necesidades de usuarios, servicios y proveedores de contenidos. Al ser un campo emergente el de la preservación digital se deben considerar de manera necesaria los casos de éxito y buenas prácticas de organizaciones activas en el campo.

Algunos autores como Hitchcock et al. (2007) definieron modelos de trabajo describiendo cuatro modelos de servicios de preservación digital a nivel de esquemas con modificación del modelo OAIS y con la adaptación de la tecnología de software e interfaz de repositorios para dar soporte técnico a los modelos.

### *Modelo de proveedor de servicio (Service provider model)*

Este modelo emplea una adaptación concreta del sistema del sistema OAIS donde las entidades funcionales son modificadas de la siguiente forma. La entidad ingesta (ingest) se modifica a una interfaz de ingesta personalizada al repositorio institucional (IR interface ingest), la gestión de datos (data management) toma personalidad del repositorio institucional (institutional repository) y almacenaje de archivo (archival storage) es modificado a un proveedor de servicio de preservación (preservation service provider). El flujo del modelo referencia continua con las demás entidades con el objetivo de proveer un servicio de preservación.

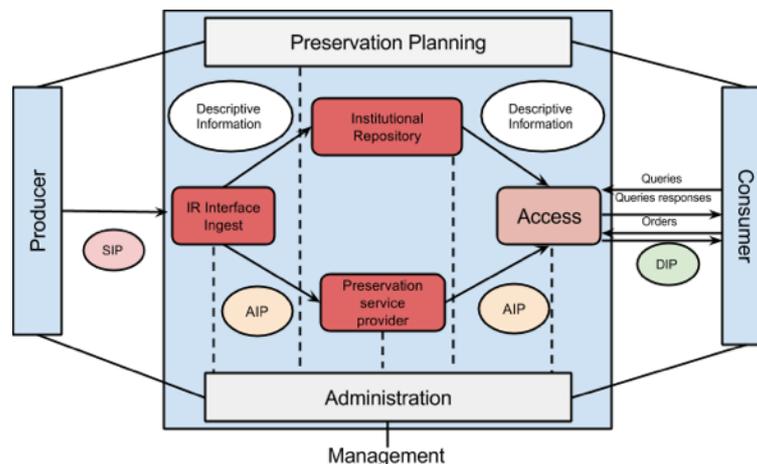


Figura 3. Service provider model (Hitchcock et al., 2007)

### *Modelo institucional (Institutional model)*

La personalización y adaptación del modelo OAIS a una entidad institucional sugiere la integración de una serie de patrones de identificación tanto técnicos como conceptuales en los ejes de este modelo. En terreno de entidades las modificaciones son similares a las del modelo proveedor de servicio pues la ingesta (ingest) se convierte a una ingesta personalizada al repositorio institucional (IR interface ingest), la gestión de datos (data management) se modifica a repositorio institucional (institutional repository) y la tercera entidad modificada es el almacenaje de archivo (archival storage) que se convierte en un repositorio de preservación (preservation repository) que los integra en una unidad de modelo institucional.

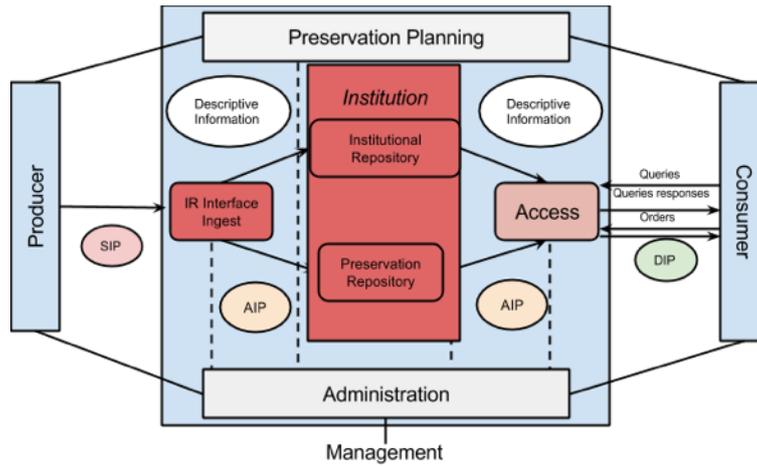


Figura 4. Institutional model (Hitchcock et al; 2007)

*Modelo de software (Software model)*

El modelo de software corresponde a la integración y desarrollo de capacidades de un sistema de software al modelo OAIS. Modificando y agregando capacidades, así como permisiones al modelo de preservación digital.

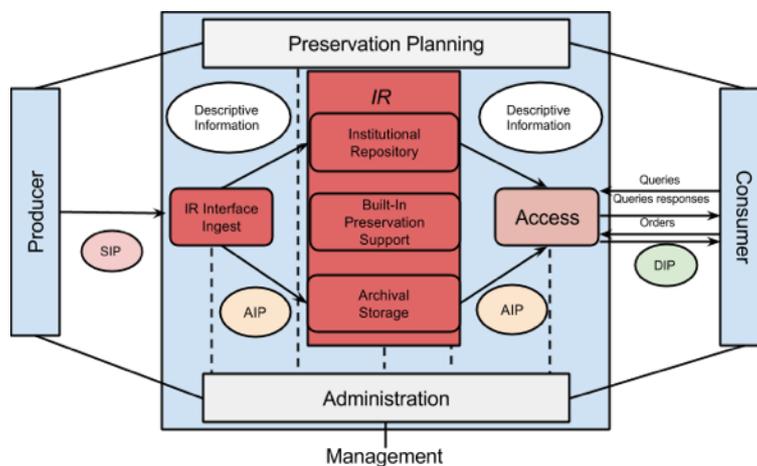


Figura 5. Software model (Hitchcock et al., 2007)

### *Modelo de red y federados (Federated and network models)*

El modelo de red y federado incluye principalmente la colaboración de varios clientes / operadores que comparten contenidos y réplicas similares a niveles de compromiso de socio. Se comprometen a mantener copias idénticas bajo políticas y normativas establecidas en los participantes. Principalmente la modificación del modelo OAIS se aplica al acceso, ingesta y almacenamiento de archivo.

Otro modelo cercano a la propuesta de *Federated and network models* de Hitchcock et al. (2007), y que implementa estrategias de preservación digital en modo colaboración y distribución institucional en red, es la que se propone en la publicación de Educopia Institute denominada “A guide to distributed digital preservation” de Skinner y Schultz (2010).

### *Modelo de preservación digital distribuida (Distributed digital preservation model)*

De acuerdo con “A guide to distributed digital preservation”, mediante la ayuda del modelo de referencia OAIS se definieron los procesos y tecnologías que deberían estar presentes en cada uno de los procesos de envío, diseminación y preservación de objetos digitales que definen el modelo distribuido como *distributed digital preservation* (preservación digital distribuida). Este modelo implementa una logística completa que incluye una plataforma abierta de software, una red de afiliados y un protocolo de organización que formaliza una estrategia de preservación digital a nivel de colaboración y distribución en redes privadas. Este modelo de preservación digital se basa en la metodología de que cualquier sistema responsable de actividades de preservación digital debe mantener y distribuir copias de los ficheros digitales que resguarda en diferentes ubicaciones geográficamente dispersas no solo como respaldo, sino con el objetivo de preservar los ficheros digitales en las distintas locaciones compartidas que previamente fueron seleccionadas para compartir y cumplir con los criterios necesarios de evitar el margen de error de pérdidas de datos o sostén de las actividades de preservación. En la práctica, la distribución de los ficheros se realiza en distintas locaciones que funcionan como servidores o nodos mediante un software y una metodología que periódicamente realiza una actividad de monitoreo, validación y copia de los ficheros para garantizar su integridad digital.

De este modelo de distribución de réplicas como copias idénticas en distintos ordenadores, se han planteado modelos de preservación digital virtual o en la nube complementando la estrategia de preservación digital en modo distributivo a diferentes escalas.

### *Preservación digital en la nube (In the cloud digital preservation)*

La preservación digital virtual o en la nube es un modelo estratégico que en principio pretende suplir un sistema de almacenamiento local físico por un servidor remoto de almacenamiento. En este sentido varias empresas como Google, Jumpbox y Amazon ofrecen servicios de repositorios virtuales en la nube, software de gestión y ordenadores virtuales que proponen integrar una estrategia complementaria a la preservación digital atendiendo a las posibilidades de adherirse en un plan de preservación digital. Rosenthal y Vargas (2013) sugieren en su estudio sobre preservación digital distribuida en la nube que las características de preservar y replicar en la nube aún son lentas y costosas en comparación a el uso de servidores locales físicos geográficamente dispersos. Los servicios y prestaciones que se ofrecen en la nube para integrarse a un plan de preservación van desde un servicio de ordenador virtual, servicio de almacenamiento de objetos digitales simples y servicio de almacenamiento en bloque. El servicio de almacenamiento en bloque es el que más ejecuta cercanía a un modelo de preservación digital pues implementa una tecnología robusta en los recursos de accesibilidad continua a los objetos digitales depositados. Este servicio otorga fiabilidad de disposición de los objetos depositados ante cualquier incidencia tecnológica como reinicio de los ordenadores virtuales, etc. El servicio en la nube de un ordenador virtual puede instalar un software de gestión de repositorios y así gozar de las prestaciones de un repositorio virtual con sus pros o contras en cuanto a la velocidad e incidencias colaterales de las conexiones a internet de terceros. Este tipo de tendencia de servicios en la nube son definidos por Brown (2013) como *Preservation-as-a-Service (PaaS)*, refiriéndose a los servicios ofrecidos por empresas donde estos no solo se limitan a almacenar en la nube, si no que integran en sus sistemas las principales técnicas de preservación digital como migración, auditoría de integridad, reportes, etc.

Hay que mencionar que existen modelos de preservación digital de acceso privado que trabajan sobre una metodología de trabajo lejos del acceso público y bajo redes privadas de archivo conocidas como archivos negros de preservación digital.

### *Preservación digital de archivo negro (Black archive digital preservation)*

Dados los riesgos que propone la custodia a largo plazo de objetos digitales de una organización (ya sea de gobierno, internacional, etc.), algunas organizaciones toman a cuenta un extra de seguridad en las actividades de mantenimiento y gestión de sus memorias digitales aplicando un exigente control de acceso y enfatizando de manera primaria en las actividades de preservación digital (migración, normalización, etc). La preservación es el eje principal de este modelo y se ejerce como prioridad ante una logística que limita el acceso a los objetos custodiados solo de manera autorizada y después de un lapso de tiempo. Un ejemplo de software que permite la gestión de un archivo negro de preservación digital es DAITSS que a diferencia de otros sistemas y aplicaciones trabaja bajo un forzoso estricto control de acceso con prioridad en el aseguramiento de la integridad y autenticidad de los objetos que almacena. Este software aplicación basa su trabajo en el modelo de referencia OAIS y la gestión de un repositorio. Otro ejemplo es el que ofrece el sistema CLOCKSS que mantiene un archivo negro controlado mediante una red federada de instituciones que no pueden acceder a sus contenidos hasta que estos sean liberados según la política definida de acceso.

### *Modelo de sostenibilidad de archivos digitales (Digital archives sustainability model)*

Algunas propuestas y estudios orientados a encontrar un modelo sostenible de preservación digital de archivos, apuntan a potenciar y abrir las posibilidades a fuentes de financiación externas que ayuden a la sostenibilidad de archivos digitales académicos. Más allá del compromiso institucional unilateral de mantener un archivo digital a largo plazo, el modelo de sostenibilidad pone sobre la mesa la participación colaborativa en beneficio de la cultura académica y el impacto social sugiere apoyar a programas de preservación a largo plazo. Otras propuestas puntualizan un modelo sostenible desde ayudas filantrópicas por donaciones, becas, dotaciones, o la generación de ingresos mediante membresía, valores agregados y sistemas de publicidad aplicados a los sistemas de acceso (Maron et al., 2014). En esta línea de la participación, se plantea la mejora de las vías de acceso como un servicio premium en formatos enriquecidos como en licencias de uso. Por otro lado, se presenta otro ámbito de discusión que refiere a los compromisos de algunas instituciones con sus declaraciones de

acceso abierto a sus contenidos, donde estos modelos estudian las opciones de no alterar dichos compromisos y acotar un acuerdo equilibrado.

Estos modelos de generación de ingresos mantienen dos categorías de pago de servicios por uso del sistema de archivo digital en el que proponen el pago de beneficiarios directos por uso de suscripción, pago por uso, pago de contribuyente o tarifa de hospedaje de publicación. El pago de beneficiarios indirectos como el referido de la institución anfitriona del archivo, patrocinadores corporativos, publicistas, fundaciones filantrópicas y licencias de contenido. Aunque este modelo está planteado en mayor proporción a niveles de acceso de contenidos y el pago el mismo, no excluye que para ofrecer servicios de acceso a contenidos digitales se contemple la gestión de preservación digital a largo plazo, ya que todo contenido por el que se paga requiere una disponibilidad permanente. Hay que mencionar que aunque estas propuestas de modelos sostenibles buscan alcanzar un nivel de financiación por medio de la publicidad o un plan de negocios, estos no están muy lejos de lo que algunos autores apoyados en teorías de Laclau y Mouffe (2001) definen como nuevo institucionalismo en las administraciones de bibliotecas que están afianzando factores y agentes clave como la modernidad, leyes de mercado y calidad recursos ínter relacionados. La modernidad es inseparable de desarrollo de tecnologías de la información e infraestructuras, lo cual repercute en la manera clara sobre la eficiencia y racionalidad en la que los usuarios acceden de manera pública a los objetos digitales de consulta (Williams y Golden, 2014).

La orientación de las propuestas de modelos de trabajo en preservación digital contemplan modificaciones técnicas y extensiones funcionales a marcos de trabajo del modelo OAIS, los repositorios y las necesidades específicas de cada organización con la integración de estrategias externas de distribución y seguridad para ampliar los ejes de partida y referencia. En todos los modelos es posible diferenciar entornos de trabajo que incluyen actividades exclusivas de tipo interno y otras de tipo externo. El trabajo de entorno interno se aplica directamente a las funciones del modelo OAIS (ingesta, administración, disseminación, etc.) y a funciones de preservación digital (migración, integridad, etc.), así como a las funciones del repositorio digital (almacenaje, descripción, etc.). El trabajo de entorno externo incluye las colaboraciones instituciones (redes y convenios, políticas, normativas, etc.), contrataciones de servicios mixtos (almacenaje, software, mantenimiento de hardware, etc.). Esta idea de trabajo en dos entornos compartidos en un modelo de preservación digital complementa la

visión de trabajo interno (OAIS, preservación, repositorios) con un esquema o arquitectura de elementos externos (proveedores, patrocinadores, etc.). El trabajo en los modelos y roles que apuntan Termens (2013) y D'Amato (2013) como dualidades esenciales en las actividades de preservación digital definidos en modelos centrales o descentralizados, abiertos o cerrados y roles de acceso o preservación, otorgan una panorámica de todos los elementos necesarios y compartidos en diferentes entornos. La determinación de un sistema de trabajo de relación central o distribuida en la arquitectura logística de un plan de preservación digital ofrece opciones para integrar ambas acepciones tanto como servicios externos como de la estructura interna de OAIS, lo que ayuda a definir un esquema a nivel conceptual de los elementos generales que pueden constituir un modelo de preservación digital flexible y sostenible.

### **2.1.3 Estrategias técnicas aplicadas en preservación digital**

¿Cual es el fin de la preservación digital?. Para contestar esta pregunta debemos situarnos en la esencia de su razón de ser como actividad planificada que integra técnicas y estrategias que tienen por objetivo mantener un acceso continuo, íntegro y seguro de los objetos digitales que custodia a largo plazo. El paréntesis que enmarca a cualquier actividad de orden digital incluye tecnología y procesos especializados que definen un resultado de tipo técnico. Precisamente lo que diferencia a un proceso de preservación digital de un proceso de almacenamiento digital, es el tipo de técnicas informáticas aplicadas a los objetos digitales. La preservación digital integra técnicas específicas en sus procesos de conservación tales como la migración de formatos, el refresco o actualización de soportes, emulación de entornos o sistemas operativos, comprobación de integridad digital, análisis forense o la arqueología digital. Todas las técnicas mencionadas de manera general forman parte de una estrategia de preservación activa y pasiva como se menciona en *Digital Preservation Policies: Guidance for archives* (The National Archives, 2011), donde la preservación activa es definida como toda aquella actividad y técnica pro-activa aplicada a preservar archivos digitales, como por ejemplo la migración de formatos para mantenerlos estables. La preservación pasiva enfatiza en las técnicas de soporte como las actividades de almacenamiento digital y no directamente con los archivos digitales. Es importante mencionar que la agenda mundial sobre la preservación digital avanza en diferentes iniciativas que ayudan a la definición de políticas, software y estándares propios de esta disciplina. Estas recomendaciones y protocolos buscan dinamizar un trabajo coordinado que eficiente por un

lado para la integración de técnicas informáticas de preservación digital en procesos automatizados y por otro lado que equilibre la participación de recursos humanos en los procesos de trabajo ante el crecimiento exponencial de los objetos digitales producidos y con necesidades de normalización para preservar a largo plazo.

### ***Digitalización***

Las técnicas informáticas empleadas en preservación digital se colocan en diferentes segmentos de actuación pues aplican de manera individual en el tratamiento de formatos, soportes de hardware, sistemas operativos, seguridad e integridad de los archivos, entre otros. Debemos mencionar que algunas técnicas como la digitalización no son consideradas como parte del combo principal de estrategias técnicas de preservación digital, sin embargo representan una importante actividad preliminar de preparación para la recepción de los datos, información o documentos a incorporarse a un proceso digital conservación. En este sentido la técnica de la digitalización cumple con el punto de partida en la conversión de datos analógicos a digitales como una técnica de orden primario y fundamental en la discriminación de archivos de origen físico a digital en los procesos previos de las estrategias de preservación digital. La digitalización es definida por la Digital Preservation Coalition (2012) como el proceso de crear archivos digitales mediante el escaneo u otras maneras de conversión de material análogo. Dando como resultado una copia digital o sustituto digital que deberá ser clasificado como material digital y entonces será sujeto de los mismos desafíos involucrados en la preservación de su acceso como los materiales nacidos digitalmente.

### ***Objetos digitales***

Es importante hacer una diferencia de consideración conceptual para centrar y apuntar dos tipos de objetos digitales empleados en preservación digital. En primer lugar los objetos digitales que son denominados como copias a sustitutos digitales deben ser denominados como materiales digitales, ya que se crean y nacen de un proceso de digitalización de origen material, y que a diferencia de los nacidos digitales (born digital) que son y se deben de manera obligada a un origen digital. De manera que tanto los materiales digitales como los nacidos digitales de manera general pueden ser considerados como objetos digitales. La Biblioteca Nacional de España y otras entidades del sector de bibliotecas digitales definen a

un objeto digital como una unidad de información específica en formato digital, el cual puede ser una representación, un fichero, una cadena de bits o una de ficheros. Algunas consideraciones en el campo de los metadatos aplicados a la preservación digital como PREMIS definen a un objeto digital como una combinación de identificador, metadatos y datos. De tales consideraciones y de manera formal para usos en preservación digital, un objeto digital refiere a todo material digital o nacido digital representado en formato de fichero, cadena de bits o cadena de ficheros que es identificado y descrito por medio metadatos. Aclarando tales conceptos, podemos dar paso a la descripción de las técnicas aplicadas a los objetos digitales en preservación digital y su rubro de actuación.

### ***Migración***

El concepto de migración en preservación digital puede entenderse de manera simple como: “La descripción del proceso de copiar el contenido de un formato a un nuevo formato” (JISC Digital Media, 2014). En este sentido, si hablamos de migración de formatos digitales describimos un proceso de transformación y conversión de formatos de lectura y representación informática enfocado a un objetivo concreto que define a nuevos formatos de ficheros bajo necesidades y un demanda específica de transformación. Algunos autores como Waters y Garrett (2006) apuntan el concepto de migración hacia las tareas periódicas de transferencia de materiales digitales y configuración de hardware / software con la finalidad de actualizarse a las siguientes generaciones de tecnología. El concepto de migración se enriquece generacionalmente estableciéndose como el proceso de conversión periódico que haga cara a la obsolescencia tecnológica tanto de formatos como de soportes según sean las necesidades específicas. Podemos entender el concepto de migración como una serie de actividades o tareas de conversión digital realizadas de manera primaria a los ficheros de una versión antigua a una nueva, y a partir de aquí en cada tipo o formato de fichero se encontrarán diferentes intenciones de necesidad de migración digital. De acuerdo a lo anterior, la necesidad de seguridad convoca a todas las acciones programadas y anteladas que mantengan en todo momento el formato más seguro para su acceso, no importando el invertir en nueva tecnología, licencias de formato, etc. Por otro lado, la migración como necesidad obligada será la que de manera emergente concurra a transformar formatos a causa de la obsolescencia tecnológica de formatos presentada como un problema real y presente de acceso a los mismos. Por último, la necesidad convenida de migración digital procura un

estado de conversión de formato de tipo aspiración, el cual ofrezca mejores ventajas y prestaciones de las que se cuenta actualmente con los formatos gestionados.

La migración en preservación digital sugiere trazar un razonamiento propio que diagnostique la necesidad de transformación de la información en formatos digitales. Termens (2013) describe una serie de razones por las que se migra de un formato a otro entre las que se destaca la aparición de versiones modernas de formatos, cambiar los formatos de propietario a libres, usar estándares industriales como preferencia, la obsolescencia de hardware, por requerimiento legal de licencias o unificación de formatos para un uso normalizado.

Con la gran cantidad de fabricantes y desarrolladores de tecnología informática de manera progresiva van alterando los ecosistemas digitales de hardware y software. Algunos casos como el de los dispositivos móviles que cambian su arquitectura de los 32 bits a los 64 bits por ejemplo, generan día a día nuevas necesidades de lectura, adquisidor y aprendizaje de nuevos formatos para su uso cotidiano y profesional. Si esta consideración la trasladamos al campo de la preservación digital nos encontramos con una demanda notable en las necesidades de actualización y definición de políticas de migración de formatos digitales para su resguardo y mantenimiento a largo plazo. Estas necesidades se trasladan a la observación, análisis y diagnóstico que discrimina funcionalmente los formatos aptos o no de migración. Algunos de los problemas más frecuentes presentados en la migración digital que se tratan de corregir con otras técnicas o procesos automatizados son las pérdidas de datos durante las conversiones, las altas cantidades de datos a migrar, las migraciones incorrectas e irreversibles, indefinición de prioridad funcional de ficheros a migrar, etc. Sobre las problemáticas de la técnica de migración, algunos autores como Rothemberg (1998) de manera temprana aseguraban que migración digital no es el método más adecuado para estrategia de preservación digital a largo plazo por los altos costes en mantenimiento, trabajo y tiempo aseguran una alta propensión al error y riesgo de pérdida de datos en el proceso. El autor anteponía la técnica de emulación como mejor opción que la migración. Otros como Granger (2000) reiteran que la técnica de migración es la principal candidata para atender a el tratamiento de formato de altas cantidades de objetos digitales, ya que la emulación solo es usada para casos marginales de tratamiento de datos y que no puede considerarse como la estrategia prioritaria de preservación digital.

La estrategia técnica de migración digital se reitera como una técnica activa que participa de manera primaria en un plan de preservación digital. Para su mejor aprovechamiento debe gestionarse de manera oportuna bajo políticas específicas que establezcan, aseguren y concreten el acceso integral de los objetos digitales y los medios donde se representará su acceso. En la actualidad sistemas y software dedicados a la preservación digital incluyen en su protocolo la migración de manera automatizada, lo que permite a un ecosistema de preservación digital programar la ejecución de este tipo de tareas de manera periódica y bajo los estándares demandados. Es importante reiterar que las acciones automatizadas necesitan ser previamente definidas en un plan de preservación digital que tipifique, clasifique y defina el tipo de estrategia o tratamiento de migración que se va a emplear tanto de manera automatizada como manual.

### ***Refresco de soportes***

Los bits, datos e información digital de manera obligada coexisten dentro de un soporte físico o hardware que los almacena. A partir de esta intrínseca situación, los soportes digitales deben considerarse como una entidad activa que debe ser gestionada de manera periódica. El refresco o actualización de soportes es una técnica que enriquece y se comparte en cierta medida con las actividades de migración digital. Es importante situar el refresco de soportes en un cuadrante separado para optimizar y ordenar las acciones que esta técnica lleva a cabo en las actividades de preservación digital. Tal como la migración sugiere la definición de un plan de acción periódico, el refresco de soportes implica un trabajo más complejo de lo que supone actualizar equipo o software a una nueva generación pues dentro de estas actividades de actualización debe considerarse de igual manera conceptos como la compatibilidad, soporte industrial, estándares de uso, factores externos, obsolescencia tecnológica, etc. Todo soporte físico que almacena datos digitales o bits, mantiene de manera obligada una sensibilidad y fragilidad intangible, la cual solo puede ser sostenida por un respaldo material o de hardware que requiere de cuidados primarios a los materiales físicos que protegen el contenido digital de factores externos que ponen en peligro la preservación digital de estos.

Todos los materiales físicos son afectados por factores externos independientemente de las acciones de migración que tomemos en dado o con suficiente tiempo, ya toda la información física eventualmente se va degradando hasta que toda la información que contiene se pierde

(Del Pozo et al., 2010). En el ámbito de preservación de objetos digitales sucede lo mismo, ya que los medios físicos o discos duros que los contienen también son vulnerables a factores externos independientemente del tipo de hardware portador de la información. Otros puntos de vista acerca del refresco de soportes lo refieren a copiar datos a un nuevo soporte de datos (del mismo tipo o de otro tipo) cuando el soporte falle o se vea amenazado a fallar (Goportis, 2013). Esta estrategia encuentra algunas ventajas preventivas como las de tratar los riesgos de fallo en los soportes y el tratar la disponibilidad de reproductores ya que su implementación es económica y fácil de realizar. Se sugiere como inconveniente de la técnica el hecho de que los problemas de dependencia de software no se ve solucionado, ni el de envejecimiento de formatos y que no es sostenible sin el uso de otras estrategias. En este sentido se debe apuntar que aunque el refresco sugiere su renovación y copia de datos digitales a un nuevo soporte, debemos entender que esta copia sugiere un sentido de identidad única la cual no estará libre de la obsolescencia ya que el formato sería el mismo y por tanto requiere de otras técnicas complementarias.

### ***Emulación***

La emulación es una de las tres estrategias técnicas indispensables en un plan de preservación digital. La emulación nace como una acción de imitar un antiguo sistema operativo o ambiente de operación de hardware y software. Algunos autores como Tzitzikas et al. (2012) enfatizan que la actividad de emulación digital requiere de la creación de emuladores, es decir programas que traducen código e instrucciones del entorno de una computadora para que de esta manera estos puedan ser propiamente ejecutados en otro ordenador. La preservación digital que tiene como objetivo y prioridad el acceso continuo a objetos digitales ante la obsolescencia tecnológica de software y hardware, obtiene apoyo de esta técnica. La emulación ha sido adoptada por distintos desarrolladores de software que emulan sistemas informativos que pueden leer y ejecutar antiguos formatos de ficheros informáticos que solo pueden existir bajo esos ambientes. Podemos nombrar emuladores de diferentes ecosistemas digitales como el de los videojuegos, de los sistemas operativos de ordenadores, móviles y programas informáticos específicos que tienen la capacidad de ejecutarse dentro de sistemas operativos actuales sin necesidad de su antiguo hardware o software.

La historia de esta técnica informática de manera general se remonta a la segunda guerra mundial, donde el gobierno británico mediante el ordenador Colossus pretendía imitar las funciones del código del ordenador nazi enigma (Kaluszka, 2011). Aunque de manera formal las teorías de emulación informática surgen hasta 1962, no es hasta 1965 cuando Larry Moss escribe el primer emulador que ejecutaba entornos IBM. A partir de los años ochenta el desarrollo de emuladores se concreta hasta lo que conocemos hoy en día. La emulación en preservación digital ofrece algunas ventajas significativas a los procesos de preservación digital, ya que disminuye los riesgos de disponibilidad de sistemas operativos o software, del envejecimiento de formatos de archivos y de poder emplear los datos como originalmente se usaban. La emulación como técnica de preservación digital presenta algunos *pros* y *contras* a nivel general en su actividad. Como *pros* encontramos en primer lugar el poder trabajar bajo un auténtico “look and feel” (sentir y ver) del entorno original del sistema informático y formatos. Otra ventaja considerable es que la emulación es aplicable a cualquier tipo de objeto digital. En los *contras* podemos mencionar que requiere de un emulador exacto para cada formato en el software original, se requiere del conocimiento de funciones del antiguo software a emular, se requiere atención a los derechos de autor, patentes de software y en su caso del hardware a emular (Van der Hoeven, 2014).

Por otra parte algunos expertos aseguran que se corre con más desventajas al trabajar con la emulación pues no se puede asegurar que los emuladores puedan seguir el mismo ritmo de funcionamiento en los ordenadores de arquitectura actuales a nivel de compatibilidad. Por otra parte según el tipo de objetos digitales y formatos se puede requerir de una gran cantidad diferencial de emuladores para la gestión de los mismos lo que influye en altos costes de la actividad. Hay que mencionar que aunque existen una gran cantidad de emuladores de código libre para el procedimiento de asegurar de manera formal un archivo digital, lo mas propio según los expertos es utilizar software de emulación confiable que ofrezca la garantía de mantenimiento y compatibilidad asegurada de cara al acceso y emulación futura.

Al implementar una estrategia de emulación en preservación digital se debe considerar anticipadamente ¿Que es lo que se pretende emular y a que nivel de profundidad?, a partir de aquí existen tres opciones concretas de emulación que son el emular aplicaciones, emular sistemas operativos o plataformas de hardware (Granger (2000). Temas como la integridad de los objetos digitales y la manera en la que se debe interactuar con ellos ante aspectos legales y

económicos deben considerarse como aspectos primarios en la técnica. Entre las soluciones que se han desarrollado al respecto se han planteado distintas opciones que incluyen por ejemplo encapsular los datos digitales junto con su software de origen y las descripciones necesarias de metadatos para que se encuentre disponible la característica de emulación del ambiente original. Otras propuestas sugieren el desarrollo de software específico de emulación que pueda ser sostenible al futuro y que automatice los procesos para que en menor escala se realicen actividades de orden manual y así asegurar los sistemas originales según las necesidades y generaciones. Nuevamente se denota que la emulación como estrategia de preservación digital de orden primario se complementa con otras estrategias técnicas como la migración y refresco de soportes dentro de una planificación previa que las integre para optimizar sus objetivos.

La migración, refresco de soportes y emulación son las tres estrategias técnicas que de modo principal se emplean en preservación digital y que sugieren un tratamiento en su mayoría de tipo preventivo bajo una revisión periódica de datos y objetos digitales gestionados. Estas técnicas se se han ido nutriendo y complementando con otras técnicas adoptadas en el área, que trazan un mapa formal de actuación para consolidar un terreno más estable al preventivo sobre el protocolo de preservación a largo plazo. Las técnicas que complementan la actuación preventiva y correctiva se integran por a la de análisis forense, replicación y el análisis de autenticidad e integridad digital.

### ***Análisis forense digital***

A finales de los noventa, diferentes estudios plantearon la necesidad de enfatizar en el campo del análisis digital forense como herramienta metodología para la recuperación de datos en medios y soportes antiguos. Algunos autores también definen esta estrategia con diferentes nombres como arqueología digital o museo de hardware. En todos los casos se define como la recolección de hardware y software antiguo para implementar un sistema integral que permita la ejecución de datos en su ambiente original. El análisis forense digital ofrece a la preservación digital y el mundo de los archivos una serie de patrones de trabajo para la recuperación de archivos y ficheros digitales mediante el uso de nuevas herramientas, metodología y capacidades. Los métodos y resultados de este campo plantean cuestiones legales, éticas y de hermenéutica importante en la naturaleza de los registros culturales, los

límites entre lo público y privado, el conocimiento, los roles y responsabilidades de los donantes, archivistas y el público en una nueva era tecnológica (Kirschenbaum et al., 2010).

El objetivo principal del campo forense digital es la obtención de datos de soportes que necesitan de su hardware original para ser leídos, copiados y así poder migrar a un nuevo ambiente informático. Inicialmente su enfoque estaba dirigido a solventar un problema de soporte de lectura e interpretación de datos. Hoy se presenta como una estrategia más compleja que además de trabajar con el factor obsolescencia de equipos como de formatos de datos en procesos delicados, según los expertos se suma un alto costo de mantenimiento especializado del equipo según la generación original y los componentes o consumibles que en su mayoría están fuera de la serie. Sin embargo, esta estrategia ofrece una alternativa inmediata o de oportunidad para tratar los objetos digitales en su ambiente original y así definir una estrategia de extracción íntegra de los mismos, mediante técnicas cruzadas de sistemas híbridos de hardware y software que ocupan de igual manera a la emulación y migración como puentes de salvación funcional. A nivel de obtención y gestión de datos las herramientas de análisis forense digital permiten crear copias idénticas de datos en los discos, así mismo permiten reflejar el orden original de los materiales, establecer cadenas confiables de custodia, descubrir y (en su caso) exponer la información de contexto asociada e identificar la información sensible que debe ser filtrada, redactada o enmascarada de manera apropiada (Lee, 2014).

Desde un sentido técnico y sólido, el análisis forense digital se enfrenta a dos tipos de actividades para concretar su compatibilidad. Por un lado se encuentran los controladores de dispositivos y sistemas operativos, y por otro lado los equipos de hardware forense especializado con lectores múltiples de soportes. Aunque el trabajo pudiese parecer únicamente técnico, este debe solventarse bajo un esquema de actuación planificada que defina una ruta de trabajo y tratamiento para recuperar datos almacenados a nivel de olvido en múltiples soportes digitales para asegurar a los objetos digitales o datos recuperados una nueva vida funcional para ser tratados y acceder a ellos en un futuro. Actualmente existen en el mercado líneas especializadas de ordenadores y equipos a modo de laboratorios digitales forenses que permiten mediante software y hardware multilector realizar operaciones de arqueología digital para recuperación de datos a modo original y de esta manera ofrecer una segunda vida digital de recuperación, tratamiento y preservación a largo plazo.

## *Autenticidad e integridad digital*

Una de las preocupaciones más frecuentes de la preservación digital a largo plazo es la de preservar y mantener los objetos digitales de manera idéntica a cuando fueron creados, es decir mantener su autenticidad e integridad digital. Según el glosario de The National Digital Stewardship Alliance (NDSA)<sup>9</sup> a nivel digital la autenticidad se define como la característica mecánica de cualquier objeto digital, la cual refleja el grado de confianza en el objeto y en donde los metadatos de apoyo que acompañan a este dejan claro que el objeto poseído es lo que pretende ser. La integridad digital se define y se relaciona con un mecanismo o herramienta para verificar que el objeto digital no ha sido alterado de manera indocumentada. Uno de los principales problemas y retos que representa el implementar distintas técnicas y estrategias de preservación digital es la pérdida o transformación de datos y cadena de bits con problemas de lectura, representación y acceso a los objetos digitales a posteriori. Algunas de las soluciones desarrolladas para atender tales eventos abarcan distintas acciones aplicadas a los objetos digitales. En primer lugar podemos nombrar el uso de registros y tipologías estandarizadas para nombrar formatos digitales únicos como PRONOM<sup>10</sup> o UDFR<sup>11</sup>. Por otro lado existe software especializado para comprobar de manera digital los formatos, integridad y composición de objetos digitales como DROID<sup>12</sup> o JHOVE<sup>13</sup>. En una tercera vía para comprobar la integridad digital o suma correcta (checksums) e integra de dígitos de los objetos se hace uso de aplicaciones como MD5<sup>14</sup>, SHA-1<sup>15</sup> o SHA-256<sup>16</sup>. La mayoría de estas herramientas pueden usarse de manera independiente o en algunos casos integrados dentro de un sistema de repositorio realizando de manera periódica verificaciones sobre las fechas de creación, modificaciones, descripciones e integridad en la cadena de bits de los objetos digitales.

---

<sup>9</sup> The National Digital Stewardship Alliance - <http://www.digitalpreservation.gov/ndsas/>

<sup>10</sup> PRONOM <http://apps.nationalarchives.gov.uk/PRONOM/Default.aspx>

<sup>11</sup> Unified Digital Format Registry (UDFR) - <http://udfr.org/project/reqs.html>

<sup>12</sup> DROID <http://digital-preservation.github.io/droid/>

<sup>13</sup> JHOVE <http://jhove.sourceforge.net/#implementation>

<sup>14</sup> MD5 <http://en.wikipedia.org/wiki/MD5>

<sup>15</sup> SHA-1 <http://en.wikipedia.org/wiki/SHA-1>

<sup>16</sup> SHA-256 <http://en.wikipedia.org/wiki/SHA-2>

En este sentido la integridad digital es considerada como parte de los aspectos de seguridad en la información del estándar de seguridad ISO 27000, el cual según su vocabulario cubre aspectos relacionados con la integridad, disponibilidad y confidencialidad que deben ser tomados en cuenta en los requerimientos de gestión y manejo de repositorios digitales con fines de preservación.

### ***Auditoría y certificación de repositorios de preservación***

Saber que contiene un disco duro o un ordenador de almacenaje digital mediante la observación externa es una tarea imprecisa e imposible. El mundo digital de registros de autenticidad e integridad de objetos digitales se ha planteado una serie de estándares y herramientas que permiten conocer y verificar la manera en la que gestionan la custodia de objetos digitales en repositorios de archivo y sistemas de almacenamiento desde su interior y a detalle. Desde el año 1996 el grupo de investigación Task Force on Archiving of Digital Information creado por la Commission on Preservation and Access and the Research Libraries Group de los Estados Unidos con el objetivo de investigar sobre el continuo e indefinido acceso a futuro de los documentos electrónicos almacenados en forma digital, afirman sobre la necesidad de crear un sistema que certifique de manera externa a los repositorios o archivos digitales con actividades de migración, almacenaje y provisión de acceso. De esta manera se ayudaría a crear un clima de confianza sobre las perspectivas de preservación a largo plazo de manera íntegra y segura.

La tarea de auditar la veracidad y fiabilidad con la que los sistemas de repositorios gestionan objetos digitales ha llevado a la creación de diversos sistemas de auditoría que certifican a repositorios digitales, los cuales se pueden clasificar en tres niveles que van de la certificación básica, a la certificación extendida y la certificación formal. La certificación básica integra una serie de parámetros y metodología de listas de verificación (checklist) sobre los cuales la propia organización evalúa el estado de integridad y confianza en la que los sistemas de archivo integran y custodian archivos. Algunos ejemplos de modelos para certificación básica los podemos encontrar en la metodología inglesa Digital Repository Audit Method Based on Risk Assessment (DRAMBORA)<sup>17</sup> creada por el Digital Curation Centre (DCC) y Digital

---

<sup>17</sup> DRAMBORA <http://www.dcc.ac.uk/resources/repository-audit-and-assessment/drambora>

Preservation Europe (DPE), así como el sistema holandés Data Seal of Approval<sup>18</sup> y el método español Guía de Evaluación Recolecta<sup>19</sup>. Por el lado de la certificación extendida y formal nos referimos a auditorías bajo normas ISO y DIN respectivamente que mantienen un nivel de certificación extendida cuando las actividades de monitoreo y auditoría son realizadas por la misma organización, y cuando estas son realizadas por auditores externos se considera una certificación formal. Como ejemplo de estas metodologías podemos mencionar al sistema de certificación Trustworthy Repositories Audit & Certification (TRAC) creado inicialmente por Research Libraries Group (RLG) y National Archives Records Administration (NARA) de Estados Unidos de Norteamérica y el cual se complementa con la colaboración de distintas organizaciones e instituciones a nivel internacional mediante una serie de criterios y listas de revisión estructuradas que funcionan como una herramienta de auditoría en áreas de políticas y administración, manejo de objetos digitales y tecnología en la gestión de repositorios digitales. Así mismo esta herramienta sirve para evaluar el grado de fiabilidad, compromiso y disposición de la institución ante la preservación segura a largo plazo. En esta misma línea encontramos el estándar de certificación Trusted Digital Repositories (TDR) o ISO 16363:2012<sup>20</sup> que es una revisión de la lista de revisión de TRAC y que fue publicado en el 2011 para ser aprobado como ISO en el 2012. TDR tiene como objetivo guiar a las organizaciones sobre la gobernabilidad de los procesos de auditoría de repositorios y de educar a los auditores de los mismos. Otro ejemplo de certificación formal de origen alemán lo encontramos con la metodología Deutsche Initiative für Netzwerkinformation (DINI) y en Network of Expertise in long-term storage of digital resources (NESTOR). Este último fue creado en el año 2003 e impulsado por el ministerio de educación e investigación de Alemania. El sistema de certificación NESTOR 1 (2006) ha sido actualizado de su versión inicial a la versión NESTOR 2 (2008) bajo la norma DIN 31644:2012. El concepto de auditoría para certificación de repositorios digitales hoy en día mantiene una agenda de interés internacional en el área de archivos digitales y bibliotecas. En lo que respecta a Europa se ha creado el marco de auditoría y certificación de repositorios digitales europeos que tiene como finalidad el estandarizar un sistema de certificación único que cubra todos los niveles de certificación que van del nivel básico, nivel extendido y nivel formal o completo, acorde a las normas ISO/DIN respectivamente.

---

<sup>18</sup> <http://www.datasealofapproval.org/en/>

<sup>19</sup> <http://documenta.sitios.csic.es/alfresco/downloadpublic/direct/workspace/SpacesStore/127f7cc9-651a-4954-aa37-a840726c0d74/recolecta.pdf>

<sup>20</sup> TDR <http://www.crl.edu/archiving-preservation/digital-archives/metrics-assessing-and-certifying/iso16363>

## ***Metadatos y preservación digital***

La custodia de un objeto digital con fines de preservación puede ser gestionada y tratada desde dos niveles de conservación técnica. El primer nivel es como objeto digital puro de representación junto con el medio que fue creado y su descripción. Un segundo nivel es el que algunos autores y proyectos como el CEDARS Project<sup>21</sup> definen como “forma abstracta subyacente”, es decir el conjunto de formas digitales (números) de una cadena de bits-bytes extraída de un objeto digital e independiente del medio en el que fue creado, y que para que este pueda ser interpretado por humanos y sistemas informáticos se acompaña de información de representación e información descriptiva que se denomina metadatos. Los metadatos son información descriptiva y datos adicionales que definen contenidos, procedencia, estructura, autoría, calidad entre otros contextos de un objeto digital. Estos pueden crearse como un archivo separado o incrustado dentro de la cadena de bit-bytes del mismo. Algunos tipos de metadatos como los descriptivos son creados por los autores y otros como los de representación son generados automáticamente.

Algunos autores como Brown (2011) proponen tres categorías amplias sobre las cuales se puede entender el quehacer de los metadatos en un entorno de preservación digital. La primer categoría de metadatos es la de comprensibilidad que refiere a la información de representación que ayuda a identificar, interpretar y entender la información que ha sido digitalmente codificada. La segunda categoría refiere a los orígenes, contexto y restricciones que se define como la información descriptiva de preservación que ayuda a recuperar los contextos originales sobre los que los objetos digitales fueron preservados. La tercera categoría es la de paquete o vinculación de metadatos que define la manera en la cual los datos y metadatos han sido agrupados juntos y así poder identificar los componentes de un paquete de información.

Robin Wendler (2006) propone cinco claves para entender el uso de los metadatos con objetivos de preservación a nivel funcional. El primer nivel es la viabilidad (viability) que supone que todo objeto digital debe mantenerse seguro e intacto. Por ello debe mantenerse íntegro en su secuencia de bits y es donde el registro de metadatos como un identificador o un verificador de *checksum* podrían asegurar el objetivo de mantenerlo viable de uso. El segundo

---

<sup>21</sup> CEDARS Project <http://www.ukoln.ac.uk/services/elib/projects/cedars/>

nivel funcional de metadatos es el de regeneración (renderability) que proporciona la propiedad a los objetos digitales de abrir, ejecutarse, mostrarse o activarse independientemente de su forma o contemporaneidad de software o hardware en el que fueron archivados mediante metadatos estructurales que acompañan a la secuencia de bits a emularse o reinterpretar nuevamente. El tercer nivel es el de comprensibilidad (understandability) que propone la función de los metadatos para ayudar a proporcionar un significado a nivel semántico y entender las estructuras sintácticas de los objetos digitales. El cuarto nivel es el de la autenticidad (authenticity) que mediante metadatos asegura el poder documentar toda acción, movimiento o transformación que el objeto digital tenga en sus procesos para comprender el quien, cómo y cuando le han tratado. El quinto nivel es el de la identificación (identification) que refiere a que aunque a un objeto digital goce de distintas propiedades como la integridad y usabilidad de preservación, estas necesitan estar identificadas con handles (identificadores de manejo) o URNs (nombres de recursos uniformes) que ayuden su descubrimiento y recuperación (Deegan y Tanner, 2006).

Por su parte Gladney (2007) recomienda que para asegurar la veracidad e integridad de objetos digitales preservados más allá de realizar pruebas informáticas de veracidad digital (las cual pueden ser inestables e incorrectas), todo objeto digital debe contener de manera inherente información propia o metadatos que evidencien su integridad y provenir. La consolidación de un sistema de gestión de metadatos con información histórica integrada a los objetos digitales sobre su origen y manejo incrementa la veracidad y autenticidad de un objeto digital en preservación digital. Todo objeto digital que pretenda ser recuperado o transformado en otro formato de interpretación o representación con fines de preservación a largo plazo debe ser descrito en diferentes metadatos a nivel de información de representación, información descriptiva de preservación y adicionalmente si es necesario de información vinculante de paquetes de datos. Los metadatos son el principal recurso de ayuda para su representación, descripción y gestión de recuperación. En preservación digital los esquemas de metadatos o “datos sobre datos” que se emplean por excelencia funcional son METS (Metadata Encoding & Transmission Standard)<sup>22</sup> y PREMIS (Preservation Metadata: Implementation Strategies)<sup>23</sup> que funcionan a nivel de representación y descripción (Termens, 2013).

---

<sup>22</sup> METS - <http://www.loc.gov/standards/mets/>

<sup>23</sup> PREMIS - <http://www.loc.gov/standards/premis/>

### **2.1.4 Marco general de directrices y modelos en políticas de preservación digital**

La preservación digital como política emergente a nivel mundial mantiene un estado de integración sobre dos regímenes prioritarios: el régimen económico financiero y el régimen de patrimonio cultural de la organización. Estos criterios regentes se basan en el entendimiento de que un plan de preservación digital es necesario y benéfico para una organización si es sostenible en ambas direcciones. Para mantener un sostén equilibrado a largo plazo se deben priorizar y discriminar de manera funcional el tipo de objetos digitales y colecciones que se deben conservar ante las altas cantidades de información y objetos digitales producidos cada año tanto de manera heterogénea u homogénea. Para tal acción, distintas organizaciones han propuesto, creado y personalizado políticas de preservación de los objetos digitales alojados y gestionados en repositorios depósitos digitales. Sierman (2012) define a las políticas de preservación como una declaración escrita autorizada para la gestión de un repositorio que describe el enfoque o camino a tomar por un repositorio ante la preservación de objetos digitales que gestiona y accede. Las intenciones de una política de preservación deben describir las intenciones de gestión de una organización y la manera detallada en como serán realizadas.

En *Digital Preservation Policies: Guidance for archives* (The National Archives, 2011) se menciona que las políticas de preservación digital creadas por una organización tienen dos propósitos fundamentales. El primero describe el porqué una organización está haciendo preservación digital acentuando en su propósito, manejo y pretensiones de logro con la preservación de archivos digitales. Este propósito incluye de igual manera la explicación de los beneficios de coparticipación con otras entidades externas, los beneficios que tiene el acceso y re-uso a largo plazo de archivos digitales, y los riesgos económicos, culturales y de reputación que contrae el no atender un programa de preservación digital. El segundo propósito es la actuación de las políticas como autoridad para aquellos que llevan a cabo un programa de preservación digital, ya que articula los términos de roles y responsabilidades de la organización y actores externos definiendo el conjunto de criterios de éxito y medición, la cobertura de las actividades de preservación digital, el cómo y cuándo se debe evaluar un archivo y sus registros para archivarlos. Además, este propósito pretende identificar la presencia y dueños de una estrategia de preservación digital e indicar cuáles estándares serán adheridos a la estrategia.

Las políticas de preservación representan una oportunidad coherente de integrar de manera formal y tipificada una nueva actividad como prioritaria en una organización, la cual debe considerar dentro de su gestión la implementación de programas de capacitación, educación y promoción de un nuevo orden de trabajo para preservar objetos y digitales como una herencia cultural y funcional. Distintas organizaciones y centros de estudios tecnológicos realizan esfuerzos por estandarizar normas y directrices que sirvan de base para la cohesión de una sólida práctica que colabore a la gestión de actividades de preservación digital.

Una de las primeras aportaciones para formalizar marcos de trabajo a nivel organizativo en preservación digital y que no refiere con exactitud a un documento de políticas específico fue el propuesto por la UNESCO en su división de la sociedad de la información en conjunto con la Biblioteca Nacional de Australia en el año 2003. Estas definieron y estructuraron un documento donde se abordan las principales directrices para la preservación del patrimonio cultural, científico y de información digital creado en el mundo. El documento nace a partir del proyecto de carta de la UNESCO para la preservación del patrimonio digital. En dicha carta inicial se contempla de manera primaria las necesidades de promoción y políticas públicas para atender el tema de la preservación del patrimonio digital. La carta se compone de 12 artículos que establecen de manera general la esencia, descripción y significado del patrimonio digital, la forma en la que se accede a él, así como del peligro latente que implica la pérdida del mismo y la necesidad de ser vigilado de manera formal. Los artículos establecen las necesidades de pasar a la acción continua bajo la integración de medidas necesarias que incluyan la elaboración de estrategias y políticas específicas que determinen una discriminación funcional de los elementos u objetos digitales a preservar y proteger a nivel de patrimonio digital. Los artículos proponen la promoción al entendimiento de la diversidad cultural y patrimonial que existe alrededor del mundo, y de su importancia como un tema que requiere por tanto, la atribución de responsabilidades, funciones y alianzas estratégicas necesarias que emancipan un cooperación mundial ante el tema. La UNESCO termina el articular de su carta definiendo sus funciones ante el tema y por tanto, sus compromisos de colaboración activa. El proyecto de carta conforma una antesala al trabajo consolidado a nivel de políticas públicas y de promoción que complementa a su vez un argumento robusto ante la preservación del patrimonio digital y que son consolidadas en las directrices técnicas y prácticas elaboradas por la Biblioteca Nacional de Australia. En las

directrices para la preservación del patrimonio digital se definen ampliamente terminología, principios, perspectivas de gestión organizativa, técnica y prácticas recomendadas sobre la preservación y accesibilidad digital.

Podemos decir que estos documentos considerados en el ámbito de preservación digital como detonantes de toda una cultura de importancia y relevancia de la misma se plantea el bastión de consideración más importante de la preservación digital, pues construyen el argumento holístico de su gestión incluyendo la totalidad de aspectos que van desde el nivel técnico y práctico, al nivel de organización, promoción y de políticas públicas. En ella se establece la base de los programas de preservación digital en todo el mundo y especifica la importancia de normalizar, gestionar políticas y definir estrategias técnicas que ayuden a salvaguardar de cara a las generaciones futuras todo el patrimonio de conocimiento colectivo mundial digital mediante la creación de protocolos de trabajo, propuestas y directrices que puedan crear conciencia activa sobre su gestión prioritaria.

Organizaciones alrededor del mundo de ámbitos de tecnología informática, innovación científica y producción académica ante las necesidades imparables de atender su patrimonio digital y conscientes de la necesidad de normalizar esta actividad emergente, han desarrollado estudios e investigaciones para establecer protocolos estandarizados que puedan plantear políticas de actuación sobre los objetos digitales y su preservación a largo plazo a niveles de manejo y tratamiento legal, económico y de organización. Según el estudio de análisis de políticas de preservación digital en bibliotecas, museos y archivos realizado por Sheldon (2013) se muestra que del año 2008 al 2013 se localizaron 33 políticas de estrategias de preservación digital disponibles alrededor del mundo, las cuales en su mayoría equilibrada pertenecen a bibliotecas (45%) y archivos (48%), así como de manera proporcional a museos (7%). Las políticas de preservación digital están distribuidas principalmente en Norteamérica, norte de Europa, Australia y Nueva Zelanda. El estudio concluye con datos de tendencias y una tipología de las principales instituciones que publican políticas o estrategias de preservación digital. En el caso de Norteamérica las instituciones académicas y gobiernos estatales son la mayor fuente de apoyo en publicación de políticas, a diferencia del norte de Europa, Australia y Nueva Zelanda donde la mayor parte de apoyos y publicación de políticas o estrategias de preservación digital proviene de provincias o cuerpos de gobiernos

nacionales. Australia es la única nación que tiene publicadas políticas para su archivo, librería y museo.

Es evidente que las políticas de preservación digital disponibles en sitios web a nivel mundial aún son limitadas en cantidad a comparación a la cantidad de organizaciones e instituciones (universidades, librerías, archivos, museos, etc.) que en la actualidad inician o consolidan programas de preservación digital en todo el mundo y que trabajan día a día con toma de decisiones sobre sistemas de preservación necesarios, marcos de trabajo para adquirir e ingresar objetos digitales a sus repositorios, capacidad de personal para dichas labores, planeación de presupuestos, etc. La creación de políticas otorga consistencia a la labor de un programa de preservación digital ya que activa formalmente un sistema de interacción organizacional, define un marco de trabajo transparente sobre roles de trabajo y rendición de cuentas del mismo y permite un escaparate para generar intercambio de conocimiento sobre el mismo.

Para la creación y diseño una política de preservación digital existen varios modelos o guía de políticas de instituciones que sirven para la definición de un planteamiento conceptual o de metodología estructural de funciones sobre el proceso de un programa de preservación digital. Un modelo otorga de manera intrínseca la inspiración para la definición de nuevos modelos, siempre y cuando se considere ante su concepción la personalización que define todos los niveles de requerimientos y necesidades en diferentes ámbitos como el legal, institucional, económico, tecnológico y humano que atañen a una organización específica. El marco de trabajo de los modelos de políticas o estrategias en preservación digital son la referencia de los planteamientos pormenorizados que otras organizaciones requirieron en sus necesidades específicas. Es por ello que algunas de estas políticas, estrategias y estudios han servido de guía e inspiración para crear marcos de trabajo en distintas organizaciones que pretenden aplicar un programa de preservación digital.

Sobre modelos de políticas podemos mencionar el Digital Preservation Policies Study que fue realizado por la consultoría Charles Beagrie en colaboración con JISC en el año 2008. El estudio plantea y propone un modelo de políticas de preservación digital institucional que sirven como guía para crear y personalizar según las necesidades de cada institución una estrategia propia de acción. El modelo propuesto se compone por dos secciones de trabajo que

incluyen por un costado la definición de políticas y por otro la guía de implementación a nivel práctico. En cada sección de este modelo se incluye una serie de cláusulas a personalizar. La sección de políticas se inicia por la definición de una declaración de principios que estipule la manera en la que servirá y beneficiará a las necesidades de la organización las políticas de preservación digital. En las siguientes secciones se propone definir el enlace de contexto sobre el cual las políticas de preservación digital se relacionan con otras políticas de alto nivel en la organización, la definición de los objetivos de preservación digital, identificación de contenido y sus alcances ante la creación de contenidos propios, definición de cómo se dará la rendición de cuentas del procedimiento, definición de la orientación e implementación del programa de preservación en diferentes documentos tanto para el equipo de trabajo como para la organización, un glosario y la definición de un control de versiones del documento de manera histórica y temporal. De manera continua en la sección de implementación se inicia por describir las responsabilidades del equipo de trabajo y la definición de la sustentabilidad financiera de la política de preservación para integrarse a el plan financiero de la organización, describir como la propiedad intelectual y derechos de autor serán recogidos en los planes institucionales, definición de cláusulas de servicios distribuidos con terceros o externos sobre algunas o todas las actividades de preservación digital, listado de la manera en la que se llevará a cabo el cumplimiento de estándares, como se realizará la revisión y certificación de la política de preservación en tiempo y forma, que procedimientos de auditoría y evaluación de riesgos serán necesarios, identificar los grupos de interés implicados en la implementación del plan y por último la definición de las estrategias de preservación que han sido adoptadas en las políticas.

Este modelo facilita un abanico de necesidades genéricas que atender a nivel inicial el ámbito formal de dirección de programas de preservación digital. La creación de políticas es un elemento fundamental en esta tipología de objetivos pues atiende a formalizar un modelo de operación así como de deberes y quehaceres bajo responsabilidades puntuales de preservación digital.

### ***Aspectos legales en preservación digital***

En una disciplina de continua integración a ámbitos institucionales, organizacionales y de cooperación internacional como lo es la preservación digital, los temas y aspectos legales se

mantienen en un continuo cabildeo y gestión política que aborda los cumplimientos panorámicos y de conveniencia que envuelven a los documentos, datos e información que forman parte de un programa de preservación a largo plazo. Tal como lo sintetiza Fernández (2010) en Norteamérica, Europa y Oceanía distintos esfuerzos en políticas, acuerdos y estudios han entramado una serie de posibilidades legales que recogen de una u otra forma los aspectos mas importantes a legislar sobre problemáticas que atienden por un lado a las necesidades técnicas y de organización, y por otro a las instituciones, autores y geografía política de actuación en la gestión de objetos digitales. El orden legal se aplica principalmente sobre la reproducción o transformación, acceso y distribución de objetos digitales depositados en repositorios de preservación digital. Como objetivo se pretende atender a aspectos de fondo (contenidos) y forma (formatos).

Un tema a destacar en este sentido es la definición de las principales figuras implicadas en la cadena de procedimientos y afectaciones legales alrededor de los objetos digitales con fines de preservación. Por un lado encontramos a los productores, por otro a los gestores o administradores y por último a los consumidores o usuarios de los objetos digitales. La figura de productor en preservación digital debe distinguirse en dos tipos. El primer tipo es el productor intelectual que debe ser definido como la entidad humana que concibe, define y crea contenidos a nivel intelectual. Por otro lado encontramos al productor digital, el cual es definido como una entidad informática que recibe, procesa y genera una representación informática de los contenidos intelectuales integrándose como un objeto digital. Desde este punto de vista podemos apreciar que la figura de productor se compone de la dualidad de entidades que en preservación digital mantienen una cooperación necesaria a nivel intelectual e informático, y que es traducida en la producción de un objeto digital. A nivel legal esta situación sugiere una responsabilidad doble, por un lado de autoría intelectual y por otro lado la autoría informática.

En este sentido la interpretación de escalones de quienes crean los objetos digitales en los dos niveles (intelectual e informático) refieren a distintos derechos, ya que por una lado se avala la propiedad intelectual del autor humano con derechos morales y de explotación, por otro lado la patente o licencia de creación de formatos digitales informáticos puede encontrar los derechos de propietario industrial en software y medidas técnicas de digital right management (DRM) (Charlesworth, 2012). En las entidades fuera del productor los límites sobre usos,

manipulación o actualización de los objetos digitales con fines de conservación a largo plazo se agudizan. Los gestores o administradores de planes de preservación digital se encuentran con situaciones por convenir a nivel de derechos de propiedad intelectual (propiedad industrial y derechos de autor) que deben acordar con los productores (intelectuales e informáticos) para la aplicación atemporal o planificada de técnicas como la migración o transferencia y conversión del los objeto digitales de su formato de origen a uno mas nuevo asegurando que los objetos digitales se mantendrán íntegros en el proceso de transformación y manejo con posibilidades de pérdida de información en la conversión. Los usuarios de objetos digitales encuentran su frontera legal más cercana a esta situación con la posibilidad de copia o reproducción de los contenidos, así como de la distribución de contenidos de los mismos. Todas estas posibilidades de gestión de propiedad intelectual en actividades de preservación digital se enfocan de manera primaria en los derechos de autor, los cuales (según sea su adaptación internacional) se integran por los derechos patrimoniales (o de explotación) y los derechos morales. Dentro de los derechos patrimoniales básicos se encuentran el de reproducción, el de distribución, el de comunicación pública y transformación. Por el lado de los derechos morales encontramos el de paternidad e integridad (Fernández, 2010).

A partir de los derechos de autor se definen límites de actuación para situaciones específicas en las que esté en peligro la integridad de los objetos digitales, ya que con la implementación de técnicas como la migración o refresco digital de formatos y medios de reproducción como de representación se recae directamente en un posible conflicto de derecho moral de autoría intelectual que aborda al autor como el único autorizado para realizar modificaciones o transformaciones a la obra. Así mismo podría en un momento dado contemplarse de manera estricta sobre el mismo derecho de autoría la permisión otorgada por la empresa o productor digital que creó el programa para generar formato de objeto digital de representación como el único autorizado para transformar o migrar su formato a otro según el mismo derecho, y requerirá de una nueva perspectiva que controle la autoría original y seguimiento digital de transformación.

Otro conflicto relacionado a los derechos de autor y en específico a tanto el derecho patrimonial como el moral es el que se encuentra a nivel del derecho de reproducción y transformación, ya que será necesario el consentimiento o permiso del autor para realizar cualquier copia, reproducción o modificación actualizada de la integridad su obra original que

para labores de preservación digital afecta en concreto en actividades de migración y refresco soportes. Otros niveles de gestión de objetos digitales a largo plazo son la accesibilidad y distribución al que se pueden anteponer posibles conflictos de derechos de distribución y comunicación pública al publicar obras del autor a nivel público en internet.

Algunas de las soluciones propuestas para atender las situaciones legales de propiedad intelectual en preservación digital de información y publicaciones digitales apuntan a la gestión de acuerdos de mesa que clarifiquen los objetivos de los gestores o repositorios de objetos digitales, con creadores o productores para que estos cedan los derechos de patrimoniales para ejercer un tipo de explotación limitada y transformación necesaria de sus trabajos con fines de mantenimiento a largo plazo a nivel digital. Mossink y Estelle (2010) listan una serie de soluciones acordes a la gestión de propiedad intelectual y usos de material digital en organizaciones e instituciones que gestionan publicaciones editoriales, de investigación o académicos. La primer propuesta es la de la creación de un método de acuerdos y licencias entre autores y depósitos de material digital que defina cómo será usado el material digital y qué restricciones tiene. Para tal cometido se necesita un consenso entre diferentes sectores como el académico, el de la industria editorial y el de los legisladores. Una segunda propuesta aunque más cerrada es la de mejorar el método de DRM o derechos que gestión digital que mediante herramientas tecnológicas se regulan el uso y acceso de los datos digitales. El DRM fue creado principalmente para prohibir el infringir el copyright y prevenir la piratería a nivel técnico. Este método es costoso a nivel de tecnología, por lo que puede optarse por desarrollar tecnología propia o pagar licencia de desarrollo de terceros. El tercer método propuesto, es el potenciar el conocimiento del *fair dealing* o acuerdo limpio, el cual de manera implícita define la permisión y uso de objetos digitales con fines académicos o de investigación. Un repositorio de preservación gestionado a nivel académico o institucional por una universidad podría encajar en esta definición del acuerdo limpio, sin embargo la interpretación de derechos que actualmente define esta solución podría anteponerse a los derechos de autor por las posibilidades de fines económicos que limitan los alcances de su aplicación.

Dado a que las diversas interpretaciones de propiedad intelectual aplican a distintos niveles de interés multidisciplinario y por momentos dejan a un lado a los autores morales por anteponer más la valía de sus derechos patrimoniales, Labastida (2008) apunta que distintos modelos

alternativos han tenido cabida al respecto: “Afortunadamente, la respuesta ha sido que el debate sobre la propiedad intelectual se extiende a toda la sociedad y que muchos autores no se sientan identificados con estas campañas. Es por eso que han aparecido varios movimientos a favor de hacer más accesible la cultura y el conocimiento en general. Sobre todo, se trata de movimientos que reconocen los derechos de los autores y apuestan por ejercerlos de una manera más flexible, permiten un intercambio de conocimiento y la posibilidad de aumentar la creación. En estos movimientos podemos incluir desde los defensores del software libre hasta los del acceso abierto al conocimiento científico, pasando por todas las variantes del llamado movimiento copyleft” (p.10). Con las tendencias de movimientos flexibles de derechos de autor como las licencias Creative Commons y visiones como la del acceso abierto a publicaciones de investigación, la gestión de objetos digitales, las reglas de propiedad intelectual y derechos de uso se están dotando de nuevos entendimientos sobre los aspectos que la hacen divergente del mundo análogo.

Los intereses de editoriales, universidades y usuarios entretejen nuevas posibilidades híbridas en el manejo de derechos de propiedad intelectual de publicaciones de acceso libre o privado, aunque debemos entender que no es lo mismo una publicación libre que una publicación gratuita y por tanto, debe entenderse que toda publicación digital u objeto digital sea cual sea el nivel de acceso libre que goce, contiene de manera intrínseca derechos que deben ser manifestados, publicados y reiterados como licencias o acuerdos de manera clara tanto en repositorios como en sitios web donde se albergan y de manera visible dentro de la publicación a la que se accede para que los derechos convenidos queden claros y puedan ser gestionados de manera eficiente tanto para usos a corto plazo como a largo plazo con fines de preservación. Los acuerdos de derechos de propiedad intelectual son necesarios, heterogéneos y al borde de intereses de adecuaciones convenidas. La importancia de clarificar desde la planificación de un programa de preservación digital un marco legal particular ayudará a un planteamiento de preservación eficiente sobre los aspectos legales en altas las cantidades y tipos de información a gestionar con este fin mediante políticas, mandatos o normativas legales de preservación que se anticipen al estado de operación de los sistemas de preservación que están pensados para crecer sin fin sobre las gestiones del mundo digital como el almacenar, compartir, reproducir, reutilizar y distribuir con facilidad, rapidez y economía instantánea los objetos digitales. Si se combinan las necesidades específicas sobre una logística que asegure que tanto autores, académicos y usuarios puedan trabajar en un

marco de legalidad, libertad y transparencia propiedad intelectual de orden digital, será más probable su gestión permisiva de cara a ser preservada a largo plazo.

Asentando un panorama y contexto a nivel general de la preservación digital como estrategia de archivo a largo plazo a continuación nos enfocaremos a una de las estrategias de preservación digital que refiere a el modo distribuido de preservación digital con la finalidad de profundizar en su definición, organización, normativas, arquitectura y tecnología.

## **2.2 Preservación digital distribuida**

Todas las estrategias técnicas y de organización para implementar un programa de preservación digital coinciden en que no existe un traje único o a la medida para que la gestión de archivos digitales sea sostenible a largo plazo. Cada estrategia responde y soluciona necesidades específicas de tipo técnico, de infraestructura económica y de organización. Las necesidades a nivel de acceso, integridad y replicación de objetos digitales archivados son traducidas en la construcción personalizada de planes y desarrollos de preservación que cubran el orden técnico de implementación, de organización y colaboración institucional. Las tendencias preliminares de los programas de preservación digital mantenían como premisa prioritaria la independencia centralizada como estándar de confidencialidad y seguridad de los documentos u objetos digitales custodiados. Los archivos digitales preservados debían ser tratados como un archivo personal cerrado o negro con personalidad de propiedad institucional. Estas consideraciones justifican que los primeros desarrollos de modelos y técnicas como protocolos de preservación digital apostaran por una tendencia centralizada como modelo de conservación digital. Las experiencias de pérdidas de información depositada en un solo sitio dio paso a la necesidad de mejorar la seguridad de la copia única y multiplicar las copias preservadas como estrategia de preservación en modo distribuido en diferentes locaciones e incluso instituciones en modo de red. Rosenthal (2007) menciona como ejemplo histórico de distribución de copias el caso de la Biblioteca de Alejandría, quien mostró que la mejor forma de asegurar que los contenidos sobrevivan es implementando una estrategia de distribución de copias a través de una serie de repositorios independientes. Esta fue la forma en que el papel impreso funcionó y sobrevivió durante cientos de años (Rosenthal, 2007). En el ámbito digital la distribución sí sola no asegura la

preservación a largo plazo. Para ello es necesario fundar colaboraciones de compromiso entre los puntos de distribución y establecer estrategias efectivas que aseguren la integridad de los contenidos digitales en conjunto con su distribución y acceso.

Las tendencias de colaboración y distribución como conciencia de cambio entre las relaciones individuales e institucionales viene como resultado de la preocupación de preservar con objetivos de beneficio a corto plazo como lo son: el desarrollo de estrategias sostenibles de colaboración a nivel económico y organizacional, incrementar el conocimiento en el área de la preservación digital de colaboración entre instituciones, desarrollar tecnologías de sistemas redundantes y distribuidos en posiciones geográficamente dispersas, mejorar el aseguramiento de la gestión digital de archivo a largo plazo con alianzas y la definición de políticas de trabajo de compromiso compartido (Hawkins, 2013).

Para el ámbito bibliotecario el concepto de cooperación y trabajo por consorcios no es una idea nueva. En los años setenta ya se hablaba de catalogaciones compartidas entre diferentes instituciones, así como de diferentes formas de cooperación y consorcios que se centran en diferentes agrupaciones basadas en tipo de institución, factores regionales, o sobre factores como el tipo de material a compartir (Lindlar et al., 2013). Algunas de las razones que afincan la tendencia de colaboración institucional bajo premisas de distribuir copias de objetos digitales con fines de preservación digital responde a diferentes motivos, razones y aprendizajes al respecto. Algunos investigadores sobre el tema mencionan que después de atestiguar situaciones sobre las que se sufrieron grandes pérdidas de colecciones digitales como lo fueron los desastres naturales a gran escala sufridos en algunas regiones de Norteamérica en el año 2003, las cuales provocaron interrupciones de la red de energía por un lado y eventos a nivel local como fallos en medios, errores humanos, las actividades de hackers, inundaciones e incendios; se aprendió sobre la vulnerabilidad de las colecciones digitales y de la necesidad urgente de prácticas sustentantes de colaboración entre instituciones para abordar este tipo de escenarios (Skinner y Schultz, 2010). La colaboración interinstitucional con estos fines también puede encontrarse en los principios abordados en la carta proyecto de la UNESCO sobre la preservación del patrimonio digital, que en su artículo once refiere a las alianzas y cooperación necesaria a nivel de gobiernos, editoriales, creadores, industriales del sector e instituciones que mantienen patrimonios digitales para reforzar la cooperación y solidaridad internacional con énfasis en la mejora de la brecha digital del tema

mediante la formación, educación y acuerdos compartidos de recursos. El fin es democratizar el conocimiento de las técnicas de preservación de objetos digitales (UNESCO, 2003).

### **2.2.1 Concepto y definición de preservación digital distribuida (PDD)**

Los temas de cooperación entre instituciones dentro de una agenda regional o internacional para la gestión de patrimonio digital da cabida al desarrollo conjunto de estrategias viables para afrontar la distribución, replicación y aseguramiento del patrimonio digital en común. Durante la última década las iniciativas de colaboración entre instituciones, desarrolladores informáticos y profesionales en biblioteconomía han definido sistemas y marcos de trabajo que prometen asegurar la gestión de integridad y accesibilidad de objetos digitales de manera distribuida. Sørensen (2014) refiere sobre la creación de un consorcio representado por un grupo llamado "Distributed Digital Preservation", el cual está formado por diferentes instituciones como Library of Congress, Stanford LOCKSS program, Internet Archive, DuraCloud, entre otros. Los integrantes definieron el término de preservación digital distribuida (PDD) como el uso de la replicación, la independencia, y la coordinación para asegurar que el contenido digital permanezca accesible al abordar las amenazas conocidas a través del tiempo.

Brown (2013) define a la preservación digital distribuida como preservación digital por asociación, ya que se trabaja bajo un enfoque de colaboración entre un número de organizaciones que comparten un conjunto de requisitos comunes y establecen una asociación para desarrollar y compartir servicios.

D'Amato y Giordano (2012) refieren que para entender más la perspectiva de la preservación digital distribuida (PDD) debemos hablar también de la preservación digital colaborativa (PDC), donde la PDD refiere a tener una copia de contenido digital en diferentes locaciones geográficamente distribuidas y la PDC refiere a cuando diferentes organizaciones cooperan una con otra en orden de alcanzar el objetivo de preservación. De tal manera que para el ejercicio de buenas prácticas en preservación digital con fines de colaboración y distribución se debe tomar en cuenta el uso coordinado de los distintos elementos de cada una, ya que ambas actividades deben estar vinculadas de manera clara. En este sentido, podemos referir a la PDD no solo como una técnica, sino como una estrategia integral de preservación que

coordina organizaciones, sistemas y técnicas de replicación independientes con la finalidad de conservar objetos digitales de manera íntegra, segura, accesible y atemporal.

### **2.2.2 Terminología de PDD**

Schultz y Zierau (2013) afirman que no hay una definición de término comúnmente aceptada para la preservación digital distribuida, ya que la distribución en el campo de la preservación ha sido muy usada a medida de que esta ha ido madurando. El término de PDD se ha ido aplicando a diversas iniciativas que incluyen en sus modelos de preservación infraestructuras distribuidas con diversos fines tanto técnicos como organizativos. Los autores influyen en la necesidad de crear un marco de trabajo (framework) que defina una terminología clara y formal de dos tipos representados por el grupo de los términos mayores (major terms) que incluyen la definición extendida de los elementos del concepto de preservación digital distribuida y del nivel de actuación de la replicación, independencia y coordinación. En el mismo renglón sugieren que se definan por igual los términos de apoyo o soporte (supporting terms) que explican el significado propio de elementos como la replicación, independencia, coordinación, y de términos como los de unidad de almacenaje (storage unit), nodo de almacenamiento (storage node), entorno de almacenamiento (storage environment), cache/pilar (cache/pillar), así como extensiones y definiciones más amplias de términos OAIS como la replicación (replication) y recuperación ante desastres (disaster recovery). La estrategia de preservación digital distribuida mantiene características particulares y diferenciales que pueden ser definidas tanto a un nivel técnico como organizacional. La preservación digital distribuida y centralizada hacen uso de las tres técnicas esenciales de la PD como lo son la migración, refresco y emulación.

La principal diferencia entre la PDD y la PD es que la distribuida se especializa y prioriza en la técnica de replicación como base para diseñar una arquitectura de red colaborativa con nodos coordinados como unidades de almacenamiento geográficamente dispersos e independientes. La PDD potencia y desarrolla el uso de técnicas como la suma de comprobación de integridad (checksum) para el control y monitoreo de calidad de las cadena de bits de los objetos digitales altamente replicados y distribuidos en la red. Para concretar una perspectiva más amplia del trabajo de PDD definiremos a nivel conceptual (y no técnico)

los términos de apoyo (supporting terms) propuestos por Schultz y Zierau (2013) para de formalizar una visión condensada de tipo funcional de esta estrategia de preservación digital.

### ***Replicación (replication)***

Para entender el término de replicación debe ser explicado desde la visión de unidad y acción. Como unidad la réplica en su forma simple se define como copia idéntica. Como acción la replicación (Del lat. replicatĭo, -ōnis) se define como la acción de replicar, repetición o reiteración (DRAE). Como término simple podemos entenderlo como la actividad de realizar y reiterar copias idénticas. En el caso de la terminología de preservación digital empleada en OAIS (CCSDS OAIS, 2002-2012) la replicación se define como una migración digital (de objetos de información) donde no hay cambios en la información empaquetada, la información de contenido, y en la información descriptiva de preservación (PDI). Los bits utilizados para representar estos objetos de información se conservan (tal y como son) en la transferencia a nuevos medios de almacenamiento. Se realizan copias idénticas de los objetos digitales de origen a otros medios.

Por otro lado a nivel informático la replicación ha sido definida desde diferentes puntos de vista como la replicación de bases de datos o la replicación de disco de almacenamiento. La replicación a nivel funcional se puede relacionar con la definición del concepto referida por ORACLE (2014)<sup>24</sup> sobre el terreno reiterado de las bases de datos como el proceso de copiar y mantener objetos de bases de datos en múltiples bases de datos que componen un sistema de base de datos distribuido. A esta escala podemos decir que en la replicación se realizan actividades de copias y actualizaciones de bases de datos que están almacenados en diferentes nodos (servidores) mediante la acción principal de un nodo maestro (master) que actualiza varios nodos esclavos (slaves). La acción fundamental de la replicación es la de generar copias idénticas de objetos digitales en distintos medios o nodos y mantenerlos actualizados.

Acorde a la visión de Zierau (2011) la técnica de réplica es una de las tareas fáciles para la gestión del material digital ya que los objetos digitales pueden ser replicados en cualquier momento a diferentes medios de comunicación sin cambiar el propio objeto digital. La técnica de réplica en su nivel básico de informática aborda una tarea básica de realización de

---

<sup>24</sup> [https://docs.oracle.com/cd/A58617\\_01/server.804/a58227/ch\\_repli.htm](https://docs.oracle.com/cd/A58617_01/server.804/a58227/ch_repli.htm)

copias (back-ups) de objetos digitales en distintos soportes y locaciones. La réplica de un objeto digital a causa del envejecimiento del medio que lo contiene podría considerarse ya una actividad de migración de medios según OAIS, lo cual podría representar un método indirecto de preservación digital.

La replicación va más allá de solo realizar un respaldo técnico pues su actividad puede interpretarse en dos niveles de actuación. El primer nivel se realizan actividades de respaldo a modo de réplica íntegra en soportes externos y periféricos como discos duros portátiles u ordenadores en un mismo sitio geográfico. En este nivel se definen una serie de actividades que en su mayoría refieren a la preservación a corto plazo sobre el acceso a copias idénticas en distintos soportes distribuidos en un mismo radio o en su caso cercano. La réplica de este nivel puede o no contener una estrategia que vigile la integridad o autenticidad de los mismos. El segundo nivel define un estado de réplica que además de considerar soportes externos y periféricos en un radio cercano, considera el factor de dispersión geográfica como estrategia extra para mantener una réplica íntegra de los objetos digitales en diferentes puntos o locaciones distantes. Diversas iniciativas y proyectos precursores como Internet Archive<sup>25</sup>, Digital preservación Network (DPN)<sup>26</sup>, MetaArchive Cooperative<sup>27</sup>, HathiTrust<sup>28</sup> o la Biblioteca Nacional de Francia con su sistema SPAR (Système de Préservation et d'Archivage Réparti)<sup>29</sup> embanderan el uso de réplicas federadas geográficamente dispersas con fines de preservación a largo plazo. La mayoría de esta iniciativas tienen como objetivo el replicar múltiples copias de diversas colecciones provenientes de repositorios o archivos de diferentes instituciones de gobierno o académicas a modo de archivo cerrado en diversos nodos. De esta manera se pretende proteger dichas colecciones ante riesgos de pérdidas catastróficas a distintos niveles tanto naturales, tecnológicos u organizacionales.

Chen y Curtmola (2013) afirman que la estrategia de réplica dispersa representa una ventaja a nivel de cantidad de copias idénticas distribuidas, ya que al combinar la replicación basada en

---

<sup>25</sup> <https://archive.org>

<sup>26</sup> <http://www.dpn.org>

<sup>27</sup> <http://www.metaarchive.org>

<sup>28</sup> <http://www.hathitrust.org>

<sup>29</sup> [http://www.bnf.fr/fr/professionnels/spar\\_systeme\\_preservation\\_numerique/a\\_spar\\_presentation.html](http://www.bnf.fr/fr/professionnels/spar_systeme_preservation_numerique/a_spar_presentation.html)

sistemas de almacenamiento distribuido con revisiones periódicas de integridad se proporciona un mecanismo eficiente para asegurar la fiabilidad de los datos a largo plazo

La distribución de réplicas representa un reto adicional para los planes de preservación digital por el aumento del margen de error en la pérdida de datos durante el transporte o realización de réplicas automatizadas entre los distintos nodos o sistemas de almacenamiento. Para minimizar el peligro de pérdida o daño de datos, se utilizan estrategias adicionales de revisión de integridad que de manera coordinada y automatizada a las estructuras de almacenamiento de réplicas verifican la autenticidad e integridad digital de copias que de manera automática que en su caso reintegran y reemplazan las réplicas dañadas con una copia del archivo original de manera periódica bajo las políticas acordadas por cada institución que suelen ser de 3 a 4 meses (como media variable) para el mantenimiento preventivo o correctivo de los objetos digitales. Dichas actividades quedan registradas y documentadas a manera de histórico de información sobre su gestión.

Halbert y Skinner (2010) subrayan que la replicación y distribución digital mantienen la promesa de preservación indefinida de materiales (digitales) sin degradación. Pero para ello es necesario el establecimiento de procedimientos eficaces (técnicos y organizativos). En el caso de algunos sistemas se recomienda que los sistemas de preservación al menos tengan tres a cuatro réplicas como mínimo durante la gestión de un archivo digital con fines de preservación a largo plazo en un modelo de almacenamiento distribuido geográficamente (Skinner y Schultz, 2010).

### ***Independencia (independence)***

El término de independencia puede ser tratado desde distintos puntos de vista bajo sinónimos como la autosuficiencia, individualización, emancipación, autonomía, etc. Los sinónimos la definen a modo general. Para la línea que atiende a la preservación digital distribuida nos enfocaremos en dos puntos de vista específicos del término que son referidos como independencia institucional e informática. En el entorno institucional de preservación digital la independencia refiere a la gobernanza de una organización independiente y es un concepto clave para cualquier federación de preservación digital. En una federación con este tipo de modelo cooperativo entre organizaciones de confianza, hay instituciones independientes

(agentes independientes) y personas que forman parte de ellas, las cuales se comunican, interactúan y toman decisiones (Walters y McDonald, 2008). La independencia desde el sentido institucional puede ser el acercamiento más claro de la referencia de independencia en un sistema PDD por el grado de compromiso y responsabilidad unilateral de mantener contenidos replicados.

La independencia a nivel técnico informático debe considerarse para la descripción del término y su entendimiento a nivel funcional sobre la réplicas de datos. Técnicamente la independencia sugiere describir la manera en la que un sistema, datos o aplicaciones tienen la posibilidad de funcionar activamente con y sin necesidad de dependencias. A este nivel podemos atender dos conceptos clave como lo son independencia de datos (data independence) y el de independencia de dispositivos (device independence). En un sistema de repositorio o archivo digital los datos y los dispositivos entretienen una relación funcional de contenedor y contenido de réplicas. Stonebraker (1974) en “Functional View of Data Independence” anticipadamente resumió dos definiciones generales sobre la independencia de datos que sirven de antecedente para ponerla en perspectiva. El primero lo define como el aislamiento de un programa (informático) de las consideraciones de los datos que procesa. Por otro lado la segunda como la capacidad de un programa de aplicaciones para ejecutarse correctamente e independientemente del almacenamiento real de sus datos.

Daintith (2004) define a la independencia de datos como la facilidad para modificar un esquema de base de datos sin necesidad consiguiente de modificar vistas de usuario o programas que interactúan con la base de datos, ni ninguna necesidad de volver a cargar los datos. Acorde a estas definiciones de la independencia de datos podemos decir que estos tienen la capacidad de mantenerse almacenados y procesarse sin dependencias de programas, de tal suerte que estos pueden actualizarse, modificarse o gestionarse de forma independiente a los programas o dispositivos que los almacenan. La independencia de datos mantiene de manera intrínseca relación con el concepto de independencia de dispositivos pues estos tienen como función principal el brindar acceso al almacenamiento de los mismos ya sea para lectura y escritura de los datos (réplicas de objetos digitales). En lo que refiere a la independencia de réplicas se requiere que estas estén geográficamente distribuidas, coordinadas y organizadas a través de la comprobación de la integridad oportuna y eficaz (Schultz y Zierau, 2013). El

concepto de independencia de dispositivo acorde a la guía de W3C España<sup>30</sup> refiere a que esta se aplica principalmente en la idea de que independientemente del dispositivo o dispositivos usados para acceder a la información, ésta (la información) va a estar siempre disponible y accesible para el usuario (W3C, 2014). Hay que aclarar que la independencia de dispositivos a nivel conceptual define su funcionamiento activo de independencia ante los procesos de acceso a la información, datos o aplicaciones de programas sea cual sea el dispositivo.

Esta visión de universalidad de acceso y proceso de la independencia de dispositivos a modo general puede ser contradictoria aun en la actualidad, ya que aún existen acuerdos exclusivos de algunos sistemas operativos y aplicaciones informáticas que mantienen sus procesos para dispositivos específicos. En el caso de los sistemas de preservación digital distribuida los actores implicados tanto informáticos como de organización deben cumplir con una serie de prestaciones y requerimientos compatibles en una red distribuida de colaboración como la capa de programas y aplicaciones, acuerdos, políticas y normativas, pero no de contenidos. En el ámbito de archivos digitales y preservación distribuida la definición de independencia sugiere principalmente que los contenidos de información y datos mantienen un grado de independencia como colecciones o cajas de réplicas en su gestión a nivel institucional e informático en una red de colaboración compartida o distribuida, así mismo los dispositivos, objetos digitales y las aplicaciones de programas forman parte de una unidad o nodo independiente que es gestionado por cada institución.

### ***Coordinación (coordination)***

La coordinación puede ser definida como una actividad de organización que combina de manera armónica el acoplamiento de distintas partes en una unidad funcional. En preservación digital la coordinación surge como una necesidad de integrar a las partes implicadas en un proceso de índole distribuido para que tanto contenidos digitales, organizaciones o instituciones, tecnología y autores puedan organizarse en una red federada de colaboración. El término de coordinación ante las necesidades de orden distribuido precisan la necesidad de creación de modelos y estructuras que ayuden a desarrollar estrategias de colaboración con fines de preservación digital en ámbitos centralizados y en esquemas distribuidos como extensión de una red. Para poner un ejemplo de los beneficios de

---

<sup>30</sup> <http://www.w3c.es/Divulgacion/GuiasBreves/IndependenciaDispositivo>

la coordinación integrada a un modelo de orden distribuido (como una red de bibliotecas) Björnshauge (2003) menciona que esta en esencia tiene como objetivo la creación de una estructura unificadora en la que todas las sub estructuras participen como socios iguales. La estructura central sirve como un punto de encuentro virtual donde las ideas, los métodos de trabajo, planes, estrategias y misiones son acordados por todos los socios. A través de su función de coordinación permite la redistribución de métodos y conocimiento de una manera eficaz.

Para que la actividad de coordinación pueda alcanzar un nivel y modelo distribuido de participación en repositorios digitales institucionales se deben concebir buenas prácticas de comunicación entre los integrantes coordinados, ya que una estrategia de preservación digital distribuida tal como lo afirman Schultz y Robbins (2010) exige por definición la comunicación y la colaboración a través de múltiples ubicaciones y numeroso personal. El éxito para el establecimiento y mantenimiento de una red privada, requiere de una coordinación oportuna entre las instituciones miembros y el personal central que ayuda a monitorear y mantener la red (Skinner y Schultz, 2010).

La coordinación es un elemento clave para la colaboración de varias partes en un modelo distribuido. Para que esta colaboración coordinada sea efectiva se han realizado una serie de recomendaciones como las propuestas en las Directrices para la Preservación del Patrimonio Digital (UNESCO, 2003) donde se resalta que para establecer una buena colaboración coordinada según la experiencia de las organizaciones se deben tener en cuenta factores como comprender lo que se desea alcanzar en conjunto, contar con socios adecuados que contribuyan compartiendo intereses y compromisos con comunicación frecuente y eficaz respecto a la operaciones, mantener objetivos realistas y evaluaciones periódicas de los acuerdos de colaboración. La coordinación como actividad dentro de la preservación digital distribuida tiene por objetivo la organización de las colaboraciones institucionales y la de equilibrar las operaciones técnicas de replicación de repositorios digitales que forman parte de una red distribuida procurando la función armónica y formal de todas ellas bajo un fin común de preservación a largo plazo.

### ***Unidad de almacenamiento (storage unit)***

La definición del término unidad de almacenamiento en modo básico desde la composición de las dos palabras que lo integran se puede definir de la siguiente manera. Unidad es considerada como una propiedad (física o conceptual) de calidad de único e indivisible. Por su parte el almacenamiento digital se define como el lugar o tipo de hardware (dispositivo, disco duro, etc.) que de manera unitaria realiza actividades de mantener en almacén el contenido digital. Una definición práctica y de función informática que sugiere Rose (2005)<sup>31</sup> sugiere que el almacenamiento es considerado con dos usos principales. El primer uso es el que se refiere a los dispositivos y datos conectados a un ordenador a través de operaciones de entrada y salida que incluyen a discos duros, sistemas de cinta y otros sistemas de almacenamiento que no incluyen a sistemas de memoria y otros medios de almacenamiento dentro del ordenador. El segundo uso es a nivel formal en donde el almacenamiento de datos se divide en un almacenamiento primario que contiene los datos de la memoria RAM y de memoria caché del procesador, y el almacenamiento secundario que contiene los datos de discos duros, cintas y otros dispositivos con funciones de entrada y salida.

A nivel conceptual la unidad de almacenamiento refiere a un dispositivo único que realiza funciones de contener datos funcionales acorde a las necesidades informáticas requeridas. El glosario de Gartner (2014)<sup>32</sup> lo define como un tipo de dispositivo informático que proporciona datos o gestiona los datos de otros dispositivos informáticos conectados a la red. A diferencia de los aparatos de servidores los dispositivos de almacenamiento ofrecen o gestionan los datos sin un contexto de aplicación. Aun cuando el almacenamiento digital ha sido una pieza fundamental para el desarrollo y maduración de la era informática desde su creación formal en la década de los años cincuenta del siglo XX, en la actualidad las unidades de almacenamiento siguen consolidando su tecnología bajo distintos desarrollos, mejoras y optimización que lo mantenga estable para los fines que sean requeridos. Como tecnología base o básica en el terreno informático y en el área de preservación digital, el reporte "Storage Solutions for Digital Preservation" refiere al almacenamiento como un componente central e importante por la responsabilidad técnica que este conlleva de mantener bajo lectura y escritura distintas cantidades contenido digital a largo plazo. Es por ello que el

---

<sup>31</sup> <http://searchstorage.techtarget.com/definition/storage>

<sup>32</sup> <http://www.gartner.com/it-glossary/storage-appliance/>

almacenamiento requiere de una funcionalidad especial a fin de abordar adecuadamente las necesidades de un sistema de conservación que pueden ser sustancialmente variables entre los usuarios o socios, por la capacidad requerida, el número de objetos, el tamaño, ubicaciones geográficas, etc. Así mismo en el reporte se abordan las dependencias variables que deben tomarse en cuenta del almacenamiento las cuales recaen en la naturaleza de los datos, los patrones de uso y el rendimiento de las unidades de almacenamiento (APARSEN, 2014).

La unidad de almacenamiento mantiene una alta relación con su entorno (entorno de almacenamiento) pues las necesidades de cada programa de preservación digital es único y por tanto las tipologías de contenido y necesidades de cada institución deben referir el uso de una o varias unidades de almacenamiento específicas. El tema del almacenamiento en preservación digital es uno de los mas abordados a nivel de sostenibilidad, pues el mercado tecnológico responde a diferentes cambios económicos (fabricación, oferta y demanda, etc.) y a cambios sustanciales en la cantidad de producción de objetos nacidos digitalmente (crecimiento de almacenamiento). Rosenthal (2012) menciona que las unidades de almacenamiento cumplían hasta antes de las inundaciones de Tailandia del 2011 la ley de Kryder<sup>33</sup> donde cada año las unidades de almacenamiento bajan en un 40% de su precio, lo que sugería una época de bonanza en almacenamiento digital a corto y largo plazo más barato por la reducción gradual del costo de los discos de almacenamiento. La ley de Kryder a largo plazo ya no cumple con los niveles de reducción de porcentajes a causa de la influencia de tres datos variables de crecimiento anual como lo son la demanda de almacenamiento (60%), el crecimiento de densidad o capacidad de los discos de almacenamiento (20%) y los presupuestos para las áreas de informática (0% y 2%). Estos porcentajes variables vistos a diez años dejan a los presupuestos de las áreas de informática (principal soporte a actividades de preservación digital) muy por encima e incluso fuera del presupuesto actual del 2%, elevando un 100% mas allá de las predicciones de Kryder y la gestión actual de preservación de archivos digitales a largo plazo. Rosenthal sugiere prestar atención a soluciones sostenibles más allá de rapidez de acceso a datos y plantear una estrategia más viable que mantenga al almacenaje estable bajo modelos de costo apropiados para el almacenamiento y necesidades de preservación a largo plazo (Rosenthal, 2012)<sup>34</sup>.

---

<sup>33</sup> Ley Kryder - <http://www.nature.com/scientificamerican/journal/v293/n2/full/scientificamerican0805-32.html>

<sup>34</sup> "Designing Storage Architectures" - <http://blog.dshr.org/2012/09/talk-at-designing-storage-architectures.html>

Hay que señalar que las dos tecnologías de unidades de almacenamiento más implementadas actualmente en el campo de la preservación digital son la de HDD (hard disk drive) y SSD (solid state drive). El HDD sigue siendo la tecnología más implementada acorde a la norma para fines de preservación digital por su calidad probada de almacenamiento que asegura al día de hoy solidez y estabilidad. Los vendedores de tecnologías de almacenamiento prueban y buscan mejoras en técnicas de almacenamiento como HARM (heat-assisted magnetic recording) y BPM (bit-patterned media) para mejorar e incrementar la densidad y velocidad de grabación en disco, así como de la capacidad de almacenaje del disco magnético giratorio HDD y la tecnología flash del SSD (Brown, 2013).

Unidad de almacenamiento en PDD refiere a un dispositivo único con la capacidad principal de almacén digital y como secundaria de conexión a otros dispositivos para transferir los datos resguardados. La unidad de almacenamiento en este sentido mantiene una actividad pasiva como ente único que depende de otros dispositivos para generar actividad e interacción.

### ***Nodo de almacenamiento (storage node)***

Un nodo a modo general puede ser definido como un punto de encuentro para el almacenamiento. En términos de informática se define como un dispositivo u ordenador que puede gestionar su dirección activa y funcional dentro de una red de dispositivos similares. Acorde a la definición técnica de ORACLE<sup>35</sup> un nodo de almacenamiento es típicamente una máquina física con su propio almacenamiento persistente local, ya sea disco duro (HDD) o de estado sólido (SSD), una CPU con uno o más núcleos, memoria y una dirección IP. A modo práctico el término nodo de almacenamiento en preservación digital distribuida juega un papel importante junto con la replicación redundante, ya que hablamos de una red de nodos de almacenamiento (servidores) y nodos de replicación (clientes) conectados entre sí por un sistema común (peer-to-peer) que constantemente replica copias de objetos digitales en modo redundante. Esta acción de redundancia de replicas ayuda a predecir y enfrentar eventos externos que puedan afectar a la disponibilidad de los recursos (objetos digitales) así mismo a restaurar y monitorear la integridad de los mismos en base a un protocolo. En el glosario de “A Guide to Distributed Digital Preservation” un nodo es definido como un “caché”

---

<sup>35</sup> Storage Node - <http://www.oracle.com/technetwork/database/nosql/overview/nosql-storage-497226.html>

individual de LOCKSS que participa en una larga red de LOCKSS. Así tal “cache” es definido como un servidor que corre software LOCKSS y el cual almacena contenido recolectado en su disco local (Skinner y Schultz, 2010). En esta terminología el nodo de almacenamiento además recibe el nombre de cache LOCKSS o caja LOCKSS (box). En la oferta de software y sistemas que realizar actividades de preservación digital distribuida se proponen distintas terminologías a efectos similares como es el caso de Bitrepository Project donde los nodos de almacenamiento son llamados pilares (pillars) o sistema de almacenaje de datos (data storage system).

Acorde a estas visiones del almacenamiento implementadas en estrategias de preservación digital distribuida debemos mencionar que en la actualidad se cuenta con diferentes opciones de unidades de almacenamiento que podemos distinguir en local (dedicado), virtual (compartido) y en la nube (dedicado / compartido). Las opciones de almacenamiento refieren a distintas prestaciones, necesidades y costos. La arquitectura de software para preservación digital distribuida en su mayoría es flexible y permiten el uso de diferentes tipos de nodos de almacenamiento. Sobre este tema Sorensen (2014) menciona que el sistema PDD Bitrepository Project cuenta con la flexibilidad de integrar diferentes opciones de nodos de almacenamiento como los de la nube los cuales denominan *cloud pillar*. Acorde a estos contextos un nodo de almacenamiento es un ordenador o dispositivo con capacidad de almacenaje, interconexión y proceso de datos que trabaja bajo condiciones activas de lectura, escritura y comprobación de integridad en conveniencia de una red de ordenadores distribuidos.

### ***Entorno de almacenamiento (storage environment)***

El entorno de almacenamiento debe ser visto en dos tipos: el físico y digital. El término de tipo físico de entorno de almacenamiento refiere a las condiciones y factores externos tales como la calidad del aire, la humedad relativa, y el control de temperatura sobre los cuales se mantiene la preservación de objetos físicos. Las condiciones externas afectan principalmente a la degradación y daño de los objetos analógicos que repercuten en su funcionamiento o uso adecuado. El hardware (ordenadores, dispositivos de almacenamiento, etc.) como objeto físico contempla dichos factores como parte de un entorno que puede dañar su integridad

material y funcional. En específico condiciones básicas como la humedad relativa y la temperatura.

En entornos de almacenamiento digital con fines de preservación digital se mantiene una paridad de condiciones tanto del orden analógico (hardware) como del orden digital (datos, software, etc.) ya que toda información digital se conserva en contenedores de información digital a nivel físico. El término de entorno de almacenamiento también se relaciona con el término de red de área de almacenamiento (storage area network) donde esta categoría de dispositivos de computación incluye el almacenamiento conectado en red (NAS) y red de área de almacenamiento (SAN). Estos dispositivos son los que proporcionan la conectividad y distribución de comandos entre los nodos que forman parte de una red para ofrecer servicios adicionales con la ayuda de software de aplicación. El entorno de almacenamiento por tanto considera a todas aquellas condiciones externas (temperatura, humedad, aire, etc.) e internas (capacidad, velocidad, redundancia, etc.) que afectan o modifican el funcionamiento de las unidades, nodos o sistemas de almacenamiento digital.

The National Archives<sup>36</sup>, dentro de sus recomendaciones para gestionar un archivo de preservación digital define como entorno de almacenamiento seguro a un ordenador servidor o un nodo que provee de servicio a otros nodos clientes. En este caso el entorno de almacenamiento incluye también una tipología de entorno de almacenamiento de modo local, es decir que se realiza de manera física en un ordenador ínter conectado a una red de ordenadores similares o virtuales. Por otro lado el modo en la nube (cloud) que se realiza de manera remota en un servidor de almacenamiento virtual mediante tecnología de internet.

La consideración de un entorno de almacenamiento en PDD se refiere a un ordenador o grupo de ellos con aplicación de software que actúa como servidor, el cual cuenta con la potencia necesaria para mantener en un entorno funcional y seguro el almacenamiento de datos o archivos digitales bajo y con las prestaciones de distribución a otros ordenadores que forman parte de una red.

---

<sup>36</sup> <http://www.nationalarchives.gov.uk/archives-sector/digital-collections.htm>

### ***Cache/Pilar (cache/pillar)***

Cache/pilar es un término que contiene una doble función como elemento de acuerdo a las anteriores definiciones donde el caché es un componente que mantiene funciones de nodo, servidor o almacén de disco/memoria de datos. Así tal puede definirse como una colección de datos similares que están almacenados en un lugar escondido o inaccesible en el ordenador para potenciar la velocidad del uso de los mismos. En esta línea un pilar puede ser definido como un elemento estructural a modo sintáctico que forma parte de una arquitectura de preservación digital distribuida. Acorde a The BitRepository Project<sup>37</sup> un pilar es un medio de comunicación e infraestructura con características específicas previamente definidas en un acuerdo de nivel de servicio. Es una representación de una copia de los datos que se puede ver y se analiza como una unidad individual a nivel abstracto. Podemos referir de manera integral que un caché/pilar es nodo o espacio activo que tiene la funcionalidad de servidor de almacenamiento con prestaciones de cosecha (harvesting), distribución y restauración de contenidos digitales de otros nodos dentro de una red de preservación digital distribuida. La función de un caché/pilar es la de mantener actualizados y preparados en memoria los datos necesarios para ser llamados por los servidores y distintos pilares, cajas o nodos para mantener potencialidad de respuesta ante la interacción de una red distribuida.

### ***Recuperación ante desastres (disaster recovery)***

Ante las perspectivas de prevención de desastres y necesidades de recuperación de información digital, la recuperación ante desastres puede definirse como una planificación sistemática (técnica y de organización) que ayuda a la prevención y recuperación ante desastres de orden informático. Digital Preservation Coalition<sup>38</sup> refiere a una tipología de desastres informáticos propuesta por The Data Archive y que pueden ocurrir principalmente a los archivos digitales. En primer lugar el archivo corrupto o dañado (corrupt file) que contiene información dañada que no puede ser detectada mediante el procesamiento de datos. En segundo lugar tenemos el tipo de archivo ilegible (unreadable file) que es un solo archivo que no puede ser leído por el medio debido a un mal bloque en la cinta. En tercer lugar tenemos el medio corrupto o ilegible (corrupt media) que es cuando la cinta completa está dañada o no

---

<sup>37</sup> <https://sbforge.org/display/BITMAG/Pillar>

<sup>38</sup> <http://www.dpconline.org/advice/preservationhandbook/organisational-activities/storage-and-preservation>

puede leerse con fiabilidad. Un cuarto tipo es el área de sombra corrupta o ilegible como el área principal (corrupt shadow area as well as main area) donde todas las áreas no pueden ser leídas, ni ninguna de las áreas pueden ser renovadas o refrescadas. En quinto lugar la pérdida completa de datos (complete loss data) que refiere a cuando todos los datos son ilegibles y todos los sistemas están dañados sin posibilidad de reparación, es decir un desastre mayor. Para todos los casos y guardando las diferencias necesarias entre toda la tipología la DPC recomienda como posibles soluciones la realización de una copias de seguridad, la migración, el refresco o en su caso de solicitar al proveedor de datos una copia del sistema de preservación para reemplazar los archivos dañados. La recuperación ante desastres deberá ser un plan preventivo que provea de seguridad anticipada a situaciones conocidas, desconocidas y posibles en la que los archivos digitales puedan estar expuestos en su integridad y funcionamiento.

A modo de recapitulación podemos decir que la terminología que da soporte para la preservación digital distribuida puede brindarnos un paraguas de enfoque y dirección sobre las actividades necesarias involucradas en las acciones de un plan de PDD. Los términos de soporte dejan ver las necesidades de organización (plan, políticas, etc.) y tecnología (software, hardware, etc) que son requeridas para la correcta función de los modelos distribuidos. Para definir el ámbito de organización y tecnológico de la preservación digital distribuida debemos estudiar los protocolos existentes de organización y colaboración para situar las funciones de los sistemas, arquitectura de redes y coordinación para atender a los programas de preservación digital distribuida.

### **2.2.3 Protocolos en preservación digital distribuida**

La preservación digital distribuida como sistema de gestión integral se compone tanto de elementos técnicos como de organización que funcionan de manera específica y en coordinación de las estrategias de preservación digital que sugieren a modo de principios normativos las acciones necesarias a nivel de buenas prácticas para intentar reducir los intentos de posibles fallos en una red de preservación digital distribuida. Investigadores y desarrolladores de sistemas distribuidos de preservación han definido en sus publicaciones propuestas de guías y recomendaciones sobre los principios y prestaciones necesarias para una óptima función técnica y de organización en una estrategia de preservación digital

distribuida. A continuación sintetizamos de manera panorámica en dos grupos a nivel técnico y de organización las principales funciones protocolarias que proponen autores como Skinner y Schultz (2010) en “A Guide to Distributed Digital Preservation I MetaArchive”, Trehub y Halbert (2012) en “Safety in Numbers: Distributed Digital Preservation Networks” y Sørensen (2014) en “From stand-alone preservation to cross-institutional collaboration”.

### ***Protocolos de organización PDD***

Los protocolos de organización se componen de la guías y principios que son necesarios para realizar una estrategia de preservación distribuida. Algunas de estas son adaptadas a las necesidades y prestaciones que les otorguen sus infraestructuras tanto a nivel de recursos humanos como económicos. De acuerdo a las recomendaciones y buenas prácticas propuestas para la organización y funcionamiento de las redes privadas distribuidas en un sistema de preservación distribuida podemos puntualizar los siguientes protocolos de organización.

#### *a) Políticas de trabajo y compromiso institucional*

1. Un mínimo seis instituciones para iniciar una red distribuida de preservación. Esta recomendación se plantea como escenario de existencia mínima de instituciones u organizaciones para asegurar que la red sea robusta ante la posibilidad de fallo de uno o dos nodos y que el resto de nodos operativos pueda llevar a cabo la recuperación de desastres.

2. Los acuerdos, políticas y compromiso de trabajo entre varias instituciones debe definirse en un documento de políticas de trabajo y gobernanza que clarifique las necesidades y requisitos de los participantes para atender la preservación a largo plazo de manera múltiple y cooperativa.

#### *b) Definición de nodos y dispersión geográfica*

1. La lejanía geográfica entre los sitios que preservan el mismo contenido debe estar de 75 a 125 millas de lejanía. Esta recomendación sugiere que la dispersión crea una ventaja al diversificar las replicas en un mínimo de dos sitios diferentes como protección de contenidos similares compartidos.

2. Los sitios geográficos de preservación deben distribuirse de manera variada y en menor cabida en sitios de riesgo natural. Si un sitio de preservación se encuentra en una zona de alto riesgo de huracanes o temblores, se debe considerar que los demás sitios replicados de preservación estén en zonas fuera de peligro para evitar el margen de error tanto a nivel natural, analógico y digital.

3. Los sitios de preservación deberán ser distribuidos a través de diferentes redes de energía y no colocar a todas las infraestructuras de preservación digital en una misma área red eléctrica, pues será más seguro ante cualquier incidencia de este tipo que los demás sitios de preservación se alineen de otras fuentes independientes o diversas para evitar pérdidas conjuntas.

*c) Acuerdo de nivel de administración y comunicación*

1. Los sitios de preservación deben estar bajo el control de distintos administradores para asegurar que todas las actividades de gestión, monitoreo y control se puedan realizar de manera específica. Se recomienda ampliamente que cada sitio cuente con su propio administrador para asegurar una saludable red de preservación distribuida en modo colaborativo.

2. El contenido preservado en sitios dispares debe mantenerse en comunicación online y en constante revisión dada la naturaleza y fragilidad digital. Para evitar una posible degradación a nivel digital, los archivos deben ser constantemente evaluados y revisados con independencia de los factores externos identificables.

*d) Definición de políticas y seguridad*

1. En las políticas de distribución y replicación se debe considerar que el contenido tiene que ser replicado por lo menos tres veces para evitar cualquier desastre a nivel de sitio de preservación. Se recomienda conservar copias idénticas en un archivo compartido y autorizado en la red para que este pueda ser recuperable ante los procedimientos de reemplazo de las copias deterioradas.

2. Debe emplearse un sistema de seguridad para los contenidos replicados y distribuidos. Este debe contar tanto con sistemas que anticipen entradas de intrusos a una red privada o que los contenidos distribuidos puedan ser interceptados por terceros en una red pública.

Encontramos que los protocolos de organización responden a la planificación previa de la estrategia de organización que requiere de la simpatía, compromiso y presupuesto de las organizaciones implicadas y que pueden ser previstos a la par de definir una línea sostenible de operaciones.

### ***Protocolos técnicos PDD***

A nivel técnico los protocolos incluyen las acciones y estrategias necesarias sobre las cuales y de manera específica un sistema de preservación digital distribuido funciona. Los protocolos técnicos se integran por la ingesta, la recolección, el control - monitoreo y la recuperación de contenido para su gestión a largo plazo y de manera múltiple en una red privada de preservación digital de orden distribuido. Estas recomendaciones técnicas son necesarias y especialmente implementadas en sistemas distribuidos como complemento de las actividades principales de preservación como lo son la migración, refresco de medios, emulación, etc.

#### *a) Ingesta y recolección (cosecha) de contenido*

Esta actividades definen la manera en la que el contenido se va integrar a una red de preservación distribuida. La ingesta cuenta con dos opciones principales las cuales son vía web o directamente alimentando al servidor de almacenamiento con un dispositivo de almacenamiento externo. La vía web se sugiere como la vía más usada por tratarse de sistemas que están geográficamente dispersos. El servidor web alimenta a los demás nodos de manera que estos recolectan el contenido ingresado. Este en principio sugiere la acción necesaria de un sistema distribuido de *peer-to-peer* (de igual a igual) que realizan cada nodo sobre los contenidos

*b) Verificación de ingresos y recolección*

Una de las recomendaciones es que el sistema de preservación distribuida cuente con la manera técnica de verificar la coincidencia de que cada copia que se ha ingresado / recolectado coincida con el contenido original y las demás copias de la red.

*c) Control y monitoreo de contenido*

Los servidores de preservación deben de cumplir una función regular de control y monitoreo de los contenidos que protegen. Estas actividades se llevan a cabo mediante comprobaciones regulares de los *checksums* o sumas correctas de verificación de de los documentos mediante sistemas de votación de red que definan el estado de los mismos o de otros sistemas de monitoreo. De manera deseable se recomienda que esta actividad genere documentos de monitoreo que reflejen el estado de la red y de los contenidos en ella.

*d) Recuperación y reparación de contenido*

Una de las virtudes y objetivos de una red de preservación geográficamente distribuida en múltiples servidores es la de garantizar que los contenidos preservados en distintos sitios siempre mantengan copias sanas e idénticas para poder reparar una copia. Para ello es necesario contar con un sistema de recuperación y reparación de objetos digitales.

*e) Usar diferentes tipos de hardware, software y medios*

Se recomienda la diversidad tecnológica como protocolo para evitar fallos sistemáticos o deficiencias de particulares tipos de tecnología en un sistema de preservación digital. Esta recomendación principalmente se aborda para que los nodos de almacenamiento no estén casados con un solo tipo de oferta tecnológica y probar de manera transversal los soportes de software, hardware y medios.

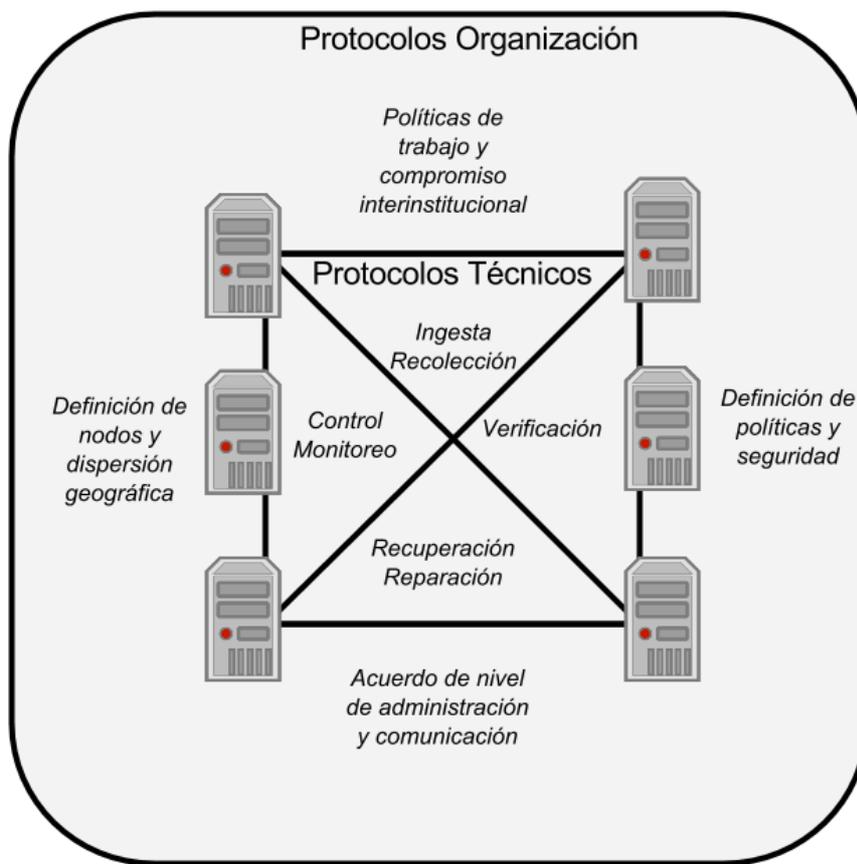


Figura 6. Diagrama de arquitectura de protocolos de preservación digital distribuida

### ***Estructura funcional de PDD***

Para definir la estructura a nivel funcional de la preservación digital distribuida nos orientaremos en los protocolos técnicos y de organización para referir las funciones, entidades y relaciones que la integran a nivel conceptual. Las redes de preservación digital distribuida al basarse en los modelos peer-to-peer o de igual a igual, definen a modo técnico una serie de ordenadores servidores (nodo, pilar o caja) fijos y a instituciones específicas (nodo institucional) que funcionan como colaboradores e iguales entre si integrando una red de intercambio de información, formatos y funciones. La arquitectura de un sistema de preservación distribuida tiene varias opciones de arquitectura para su funcionamiento. La flexibilidad de su composición acorde a las propuestas ejemplificadas en función de las redes privadas del sistema LOCKSS sugeridas por Rosenthal (2014) muestran las diferentes posibilidades y opciones de arquitectura estructural en un sistema distribuido de preservación

digital. Estas propuestas definen un escenario de necesidades específicas de organizaciones en la planificación y composición funcional de una red privada distribuida.

*a) Abierto o Cerrado (Open vs close)*

Esta opción define la capacidad de integración a una red distribuida a partir de una arquitectura abierta y libre donde cualquier participante con el solo hecho de instalar software de PDD en un ordenador y convertirlo en nodo, puede conectarse a una red de preservación distribuida de manera abierta. Por lo contrario y para defender la vulnerabilidad, sostenibilidad económica y control de una red PDD se definió el modelo funcional cerrado. La estructura cerrada trabaja por medio de un protocolo técnico distingue el acceso de los interesados o integrantes de una red de preservación digital distribuida en base a el pago de mantenimiento y soporte continuo de la red como modelo viable para su sostenimiento económico y seguro a largo plazo.

*b) Negro o Claro (Dark vs light)*

La arquitectura oscura o clara deja entrevista la definición de acceso al sistema de preservación. La arquitectura oscura se refiere a que el sistema de preservación no admite acceso alguno al contenido preservado, a diferencia de la arquitectura clara que define la posibilidad de acceso y preservación. En esta arquitectura existe la flexibilidad de mediar sobre las necesidades del sistema de preservación ya que algunas veces será indispensable tener un alto acceso y mantener la preservación por igual. Mediar entre los grises que puede suponer las necesidades particulares de cada estrategia de preservación y contenidos.

*c) Local o Huésped (Local vs hosted)*

En principio la mayoría de la mayoría de los nodos, cajas o pilares son locales pues concentran sus contenidos en local principalmente. En el caso de las GLN (Global LOCKSS Network) por condiciones y requisitos se requiere que los contenidos estén en local. En otros casos existe la posibilidad de que alguna de los nodos este de manera externa hospedada en otro sitio de terceros (p.e.: en la nube), lo que refiere a un tipo de arquitectura mixta de almacenamiento de contenidos.

*d) Colecciones Homogéneas o Heterogéneas (Homogeneous vs heterogeneous collections)*

Hay redes que preservan el mismo contenido en todos sus nodos y se definen como colecciones homogéneas. Existen redes donde cada nodo mantiene diferentes contenidos o material. Para un sistema de preservación distribuida es más fácil trabajar con colecciones homogéneas ya que en el caso de las heterogéneas el proceso de comparación de contenidos en busca de votación de idénticos en cada nodo exige más recursos y tiempo en su monitoreo ejecutivo. Por su parte las arquitectura homogéneas requieren de mayor exigencia de tipo de almacenamiento pues se hacen copias idénticas de contenidos de otros nodos en otros nodos.

*e) Reparadores Confiables o no Confiables (Untrusted or trusted repairers)*

Dentro de una red con colecciones heterogéneas suele suceder que se requiere de un nodo especial que valide el contenido diferente de los nodos, es decir que guarde conocimiento de que un nodo está autorizado a copiar cierto tipo de contenidos y donde debe inclinarse el proceso de votación para decidir la comparación de contenidos para reparar posibles daños. Los reparadores no confiables solicitan la autorización de contenido de nodos, ya que esta propiedad refiere a que no debe escaparse contenido que no pertenece a otros nodos. En el caso de los reparadores confiables no hay necesidad de contar con un nodo que conserve la informaciones antecedentes de confianza de nodos ya que como todos tienen el mismo contenido homogéneo en la misma red se desarrolla con base a un protocolo de cuadro de seguridad que identifica los nodos que son parte de la red para así poder entrar las reparaciones directamente en una red.

Las prestaciones que ofrecen distintos sistemas de preservación digital en modo distribuido entretejen una serie de capacidades que son moldeadas en las arquitectura según las necesidades de acceso, disposición o alojamiento que cada organización requiera. Lo que es importante al definir una arquitectura de un modelo conceptual o abstracto de funcionamiento de una sistema de trabajo combinado (informático y humano) es el referir a los agentes y elementos que forman parte del mismo con funciones específicas dentro del sistema.

### *Elementos de referencia en modelos PDD*

A nivel general los agentes de referencia aceptados para definir un sistema de preservación digital distribuida son la replicación, la independencia y la coordinación. Todos ellos combinados con estrategias, acciones y políticas que tengan por objetivo el asegurar la accesibilidad y disposición de los contenidos digitales durante el paso del tiempo. De esta referencia debemos considerar a nivel conceptual la función de los elementos que participan en un modelo de preservación digital distribuida. Los alcances técnicos y de organización para interactuar mediante elementos de referencia y realizar las funciones de preservar de manera distribuida.

### *Sistema PDD*

Un sistema de preservación distribuida podemos definirlo tanto a nivel de software como de protocolo. Como software o sistema informático realiza las funciones de preservación técnicas para interconectar a distintos nodos de almacenamiento y otorgarles facultades de ingresar, intercambiar, monitorear, controlar, reparar y preservar contenidos digitales. En el caso de un protocolo de preservación digital distribuida (el cual no es considerado como un sistema pleno PDD) es un conjunto de reglas informáticas y de comunicación sobre las cuales diferentes instituciones pueden preservar y replicar copias de sus datos utilizando sistemas propios de almacenamiento como de otras instituciones geográficamente dispersas que se integren en el protocolo bajo las consideraciones necesarias de seguridad, integridad, accesos y comunicación en un sistema único. Los ejemplos más especializados al día de hoy en sistemas y protocolos de preservación distribuida refieren a el sistema pionero LOCKSS y en el caso de protocolos a The BitRepository Project. Ambos integran en su arquitectura técnica las funciones de un sistema integral de preservación digital distribuida.

### *Nodos*

De manera general y para evitar cualquier inclinación o preferencia sobre un sistema hablaremos de los nodos como los ordenadores físicos, virtuales o servidores que realizan acciones de almacenamiento, enlace y cache de contenidos digitales con la instalación de un

sistema PDD en ellos. En dependencia del tipo de sistema pueden ser llamados caja (LOCKSS Box), Pilar (BitRepository Pillar), etc.

### *Políticas y acuerdos*

las políticas de un sistema de preservación digital distribuida son un elemento fundamental que dará organización a la interacción técnica y de recursos humanos. Debemos decir que las políticas definen los acuerdos de participación y trabajo dentro de la red distribuida que devienen en tres tipos principales que son las políticas y acuerdos de gobernanza de la red (centralizada, descentralizada, etc), las políticas de gestión técnica de los contenidos digitales (generales e individuales) y las políticas de tipo administrativo (generales e individuales). Además de estos tres tipos de políticas se puede considerar un sistema de acuerdos o nivel de colaboración que refiere de igual manera las condiciones específicas de cada institución en dependencia de sus necesidades o composición heterogénea.

### *Replicación geográfica*

La replicación geográfica es un elemento que mantiene en principio un sentido organizacional y por ende técnico. La replicación de contenidos en diferentes lugares geográficamente dispersos es un elemento fundamental para trazar un esquema de preservación distribuida pues con este factor como protocolo a nivel tanto técnico como de organización, se refiere al desarrollo y crecimiento de los participantes de una red PDD sobre la cual de deben crear soluciones que mantengan los diferentes nodos distribuidos en operación.

### *Independencia y colaboración institucional*

La diferencia de la preservación digital distribuida es su estado descentralizado. La integración de diferentes nodos que mantienen un sentido independiente de sus sistemas, contenidos y de políticas individuales integran a su vez un plano de colaboración interinstitucional que relaciona a todos los nodos de forma armónica con un mismo fin y objetivos de gestión de objetos digitales.

## 2.2.4 Sistemas de preservación digital distribuida

Como se ha definido y mencionado anticipadamente, todo programa de preservación digital debe de manera obligada atender factores de organización y planificación en conjunto de soluciones tecnológicas de sistemas informáticos. Una solución tecnológica por si sola no puede gestionar y definir en su totalidad las actividades de un archivo digital, ya que para ello serán necesarias diferentes acciones como la programación de mantenimiento de objetos digitales, definición de permisos de acceso y solicitudes específicas de los objetos de archivo, son actividades de orden y decisión humana ante necesidades específicas de una institución u organización. Los sistemas y protocolos informáticos que ofrecen soluciones para gestionar y dar soporte tecnológico a la preservación digital distribuida refieren a diversos tipos de software de código libre, servicio comercial o de asociación que otorgan diferentes prestaciones y capacidades de gestión de contenidos digitales distribuidos bajo el uso de la redundancia, replicación, independencia y coordinación. Cabe hacer mención que existe también una gran lista de software y servicios enfocados al almacenaje distribuido (no confundir con preservación digital distribuida) o solo preservación que a modo de integración flexible pueden adecuar sus sistemas a servicios distribuidos con prestaciones de preservación a largo plazo. Para el objetivo de esta investigación nos centraremos en el software, protocolo y servicios especializados con arquitectura integral de preservación digital distribuida.

LOCKSS se considera el sistema pionero en la especialidad de sistemas de preservación digital distribuida. De acuerdo con Dobson (2003) los antecedentes del programa LOCKSS de la Universidad de Stanford, inician en el año 1999 bajo las ideas y esquemas de Vicky Reich y David Rosenthal así como del apoyo de la Universidad de Stanford, Library of Congress, Sun Microsystems, entre otros. El objetivo de este programa es el dar apoyo a la comunidad de librerías de la Universidad de Stanford para proteger y controlar los crecientes activos digitales de la misma en un sistema replicado. El lanzamiento del sistema LOCKSS fue en el año 2004<sup>39</sup> y en la actualidad el sistema LOCKSS corre por debajo de una gran cantidad de servicios y redes de preservación digital redundante o de distribución geográfica de réplicas. El sistema LOCKSS desde el punto de vista técnico recrea su funcionamiento distribuido convirtiendo mediante su software a un ordenador en lo que se denomina “LOCKSS Box

---

<sup>39</sup> <http://www.lockss.org/about/history/>

(LB)” o “Caja LOCKSS (CL)”, la cual tendrá la capacidad de recolectar, preservar y proveer acceso en línea a contenidos digitales. Las característica principal del sistema LOCKSS se enfoca en la preservación digital distribuida y para que esta se lleve a cabo se requiere de tres acciones principales que inician con la autorización y permiso del editor de preservar contenidos específicos. Con el permiso la biblioteca pueda poner en línea una CL que ha sido autorizada para el acceso a sus contenidos y de esta manera la CL puede registrarse con un número de serie en la red de la alianza LOCKSS. Una CL realiza cinco funciones principales que son: la ingesta de contenido de sitios web específicos mediante rastreadores web, la preservación de los contenidos mediante acciones de comparación con otras CL para reparar cualquier diferencia encontrada en los mismos, la entrega de contenido autorizado cuando este no esté disponible mediante tecnología segura, el mantenimiento mediante una interfaz web para seleccionar, supervisar y controlar el acceso a los contenidos preservados o a preservar y por último la migración dinámica de contenidos a nuevos formatos convenientes para su acceso. Como pionero el sistema LOCKSS ha servido de experiencia e inspiración de otros sistemas.

The BitRepository Project de los Archivos Nacionales de Dinamarca nace por la necesidad de realizar un protocolo para un sistema de red compartida interinstitucional de preservación digital para tres instituciones principalmente: Los Archivos Nacionales, la Biblioteca Real y la Biblioteca del Estado y la Universidad. Explorando diversas opciones de preservación compartida como la tecnología peer-to-peer de LOCKSS el proyecto BitRepository necesitaba adecuaciones personalizadas para las diferentes necesidades de las instituciones participantes del proyecto de preservación. Tal como lo describe Sørensen (2014) las soluciones de preservación bits debían incluir el soporte de seguridad de bits así como confidencialidad y disponibilidad. Estos requisitos debían adaptarse a cada colección en cuestión. Por tanto la arquitectura de la Bitrepository debía apoyar todo tipo de combinaciones con el fin de asegurar que los diversos requisitos podrían cumplirse para cada colección en este ambiente de “una talla no sirve para todos”. Para ello se necesito crear un sistema protocolo con arquitectura flexible a medida de las necesidades de las instituciones. En los términos del protocolo BitRepository se denominan “Pillars” o “Pilares” a las diversos ordenadores o sistemas de almacenaje que se conforman (por políticas de los creadores y del sistema) de diferentes tipos de tecnología de almacenamiento que van desde DVD, cinta, SAN, DAS y con posibilidad de usos de almacenamiento en la nube. En este modelo cada institución tiene

un cliente que le permite intercambiar información con los diversos pilares mediante una capa media (medium-layer) o capa de coordinación BitRepository.

Los sistemas especializados en preservación digital distribuida son relativamente minoría ante los sistemas y servicios generales de preservación digital centralizados. Debemos señalar que existe una serie de sistemas de preservación digital que se complementan de diferentes tipos de sistemas, aplicaciones y microservicios que no están especificadas objetivamente para preservación digital pero ofrecen funciones correlativas ya sea sobre distribución, gestión o control de datos, empaquetar datos etc. Podemos mencionar algunos ejemplos como es el caso de iRODS<sup>40</sup> que ofrece un sistema gestión de datos distribuida y que funciona en redes de preservación digital distribuida como Chronopolis. Otro ejemplo es Coda<sup>41</sup> que es un sistema especializado en la gestión de los paquetes de información de archivo (AIP) y que ha sido implementado junto con varios microservicios en la Universidad del Norte de Texas (UNT) con el sistema Aubrey<sup>42</sup> que es un sistema de interfaz *front-end* que provee de acceso, búsquedas, navegación y visualización de los recursos a los usuarios finales del repositorio de preservación de la universidad.

Hay que hacer mención de los sistemas que ofrecen control de versiones de datos a con posibilidades de distribución como es el caso de Bazaar, Git, Mercurial entre otros con posibilidades de integrarse a un sistema completo de preservación digital a modo distribuido. En el mercado de la preservación digital como servicio existen diferentes opciones que se comercializan como PDD y que poco a poco normalizan el tipo de estrategia de distribución como funcionalidad añadida en algunos de los servicios de preservación digital centralizada. Como podemos observar a modo de ejemplo ilustrativo en la siguiente tabla existen algunas organizaciones que ofrecen soluciones de preservación digital distribuida, redundante en la nube o geográficamente distribuida ya sea a modo comercial, por asociación, por membresía o sin fines de lucro. La preservación digital sigue explorando sus alcances como herramienta tecnológica mediante diferentes consorcios, herramientas tecnológicas y servicios añadidos para dar solución a la estrategia distribuida.

---

<sup>40</sup> <http://irods.org>

<sup>41</sup> <http://digital.library.unt.edu/ark:/67531/metadc282640/>

<sup>42</sup> <http://vphill.com/journal/post/4084>

Tabla 4. Servicios y redes especializadas en preservación digital distribuida.

Nombre	Tipo	Licencia	Responsable	Prestación
DuraCloud	Servicio / Software	Comercial	DuraSpace	Sistema de almacenamiento y preservación digital distribuida basado en la nube
MetaArchive Cooperative	Servicio / Red	Sin fines de lucro / Membresía	Educopia Institute	Cooperativa de red interinstitucional de preservación digital distribuida bajo el software LOCKSS
Chronopolis	Servicio / Red	Comercial	UC San Diego	Infraestructura (herramientas, software y manejo) en una red de preservación digital geográficamente distribuida. Trabaja con software iRODS.
Data-PASS	Asociación	Voluntaria	University of Michigan	Alianza para la preservación digital de datos de ciencias sociales con uso de réplicas distribuidas bajo el software LOCKSS y SafeArchive
DPN	Servicio Red	Comercial	Digital Preservación Network	Red de preservación digital federada y distribuida geográficamente
CLOCKSS	Asociación / Servicio	Sin fines de lucro / Membresía	CLOCKSS	Archivo controlado o negro de preservación digital distribuida con software LOCKSS
OCLC Digital Archive	Servicio / Sistema	Comercial	OCLC	Servicio de sistema de preservación digital con distribución geográfica

## 2.2.5 Estructura de organización y gobierno en modelos de PDD

Para entender las prestaciones tecnológicas existentes ofrecidos en los diferentes sistemas y servicios de preservación digital distribuida debemos conocer los modelos de organización de socios, participantes y arquitectura de las redes distribuidas. Existen diferentes estructuras aplicadas que designan responsabilidades y roles de trabajo tanto a nivel económico como de dirección y coordinación. Tal como se sugiere en las directrices para la preservación del patrimonio digital de la UNESCO existen modelos estructurales de preservación digital a modo colaborativo que revisan el ámbito distribuido y que se pueden englobar en cuatro categorías. En primer lugar encontramos a los *modelos distribuidos centralizados* que tienen

por característica principal que la dirección general de un sistema de preservación distribuido (políticas, infraestructura, asesoramiento, etc.) la dirige un solo socio mientras que los demás participan solo cumpliendo funciones específicas. La siguiente estructura es la del *modelo distribuido más parejamente* en la cual los socios tienen responsabilidades idénticas en la toma de decisiones, definición de políticas, infraestructura, etc. Otra estructura es la de las *colaboraciones altamente distribuidas* en la cual cada uno de los socios tiene una responsabilidad específica de preservación. Finalmente la estructura de *modelos autónomos* que se puede considerar dentro de los modelos de colaboración distribuida ya que se trabaja inicialmente solo en espera de colaborar con otras entidades más adelante (UNESCO, 2003).

En este sentido todas las estructuras de modelos distribuidos representan una válida organización de miembros para definir un punto de partida de colaboración conjunta y distribuida con un parámetro de roles de los participantes dentro de un proyecto de preservación digital distribuida. A continuación sintetizamos los pros y contras de las estructuras de organización de modelos de preservación digital distribuida acordes a estas categorías propuestas en el documento de la UNESCO bajo factores de responsabilidad, control y colaboración.

Tabla 5. Estructuras de organización y gobierno en modelos distribuidos.

Tipo	Pros	Contras
Estructura Distribuida Centralizada	Alto control y objetividad unilateral en toma de decisiones	Única responsabilidad y limitación de colaboración y recursos de socios
Estructura Distribuida Parejamente	Equilibrada colaboración, responsabilidad y uso de recursos compartidos por socios	Diversidad en el control y objetividad en la toma de decisiones
Estructura Altamente Distribuida	Limitada colaboración, responsabilidad y recursos compartidos por socios	Diversidad en el control y objetividad en la toma de decisiones
Estructura Autónoma	Alta responsabilidad y control unilateral	Nula colaboración y limitación de recursos

En esta tabla podemos observar que las diferentes estructuras de organización en un sistema de preservación distribuida pueden mantener actuaciones limitadas con base a mayor control y mayor colaboración sobre menor control. Se necesita definir una política de actuación que estime las acciones de responsabilidad, control y colaboración de forma clara para que la estrategia de preservación digital distribuida pueda gestionar eslabones de socios productivos,

transparentes y con una personalidad real de organización que respete la independencia y la coordinación de manera objetiva. La organización estructural dentro del flujo de trabajo de un equipo de preservación digital distribuida se deben mantener objetivos claros de lo que se pretende alcanzar con un plan colaborativo y del tipo de socios que integrarán la colaboración debe mantener un perfil adecuado con respecto su contribución (ideas, responsabilidad y compromiso). Se debe tener claro que se requiere por cada una de las partes para contar con recursos necesarios y políticas de comunicación claras para que se logren los objetivos de manera realista.

McGovern y Skinner (2012) proponen seis enmiendas para lograr que los esfuerzos de organización en una estrategia de preservación digital sean efectivos o exitosos. Estas recomendaciones aplicadas a una estrategia de preservación distribuida representan un clave de alto interés a los niveles de colaboración integrada por varios socios en red. Dentro de las enmiendas recomendadas en primer lugar se menciona que la preservación digital requiere de un compromiso de largo plazo como una solicitud referida de manera general para cualquier estrategia de preservación digital sea distribuida o centralizada. Cuando se dirige a un grupo de integrantes de estrategias distribuidas o socios esta enmienda será la clave para la selección y formación de grupos. La segunda enmienda es la de considerar que los costos de una estrategia de preservación digital son más efectivos cuando se trabajan a escala, es decir que los recursos económicos destinados a sostener una estrategia a largo plazo de manera compartida garantizan su alcance y sostenibilidad junto con el compromiso a largo plazo. La tercera enmienda recomendada define que se requiere de una interacción efectiva entre los puntos implicados como lo son usuarios, productores y archivos digitales. Esta interacción y cercanía de intercambio puede mantener una mayor coherencia y control en lo que se refiere a la identificación de material, decisión de materiales de importancia, mayor organización de los materiales y servicios. La cuarta enmienda define que son más valiosos los beneficios que se pueden recibir de una estrategia de preservación digital compartida, que de la explotación del enfoque de beneficios único pues la unión hace la fuerza y la visión de trabajo de esfuerzo común y colectivo enriquece más que la singularidad. La quinta enmienda refiere a que las iniciativas de preservación digital deben generar un alto impacto público sobre la importancia de la preservación digital en el presente para lograr la suficiente atención y compromiso de la sociedad y fundaciones para que éstas apoyen a la sostenibilidad de la preservación a largo plazo. Una estrategia distribuida es una iniciativa incluyente de diferentes organizaciones e

involucrados que debe referir a que los esfuerzos en común que se logren el presente con conciencia será fruto para las generaciones futuras. La sexta enmienda que proponen es que la preservación digital requiere de una reconocida división de trabajo (profesional) desde la perspectiva digital donde se necesita de especialistas que trasladen y realicen la labor de actuación de manera práctica a nivel organizacional, colaborativo / distribuido en diferentes ámbitos desde el punto de vista de preservación digital para poder obtener una visión plural a escala y ampliar el abanico de buenas prácticas.

En estas enmiendas realizadas por especialistas de distintos grupos y ámbitos especializados en preservación digital reflejan el ímpetu sobre las necesidades de organización y de colaboración necesarias para seguir abordando la conciencia de importancia sobre la preservación digital, donde los modelos distribuidos deben afrontar de tal manera retos claros, comprometidos y de divulgación para lograr avanzar sobre en diferentes fronteras una colaboración distribuida y sostenida. Autores como Rosenthal (2013) en términos concretos hace referencia a la estructura del modelo distribuido como preservación a escala (Preservación at scale) que se promueve como una ventaja a nivel económico cuando varios participantes actúan para relativizar los costos de operación en preservación digital. El autor afirma que este término de escala debe ser visto con alta ética y responsabilidad tanto por las organizaciones que llevan a cabo los planes de preservación a modo distribuido, como y primordialmente de las empresas que ofrecen servicios de almacenaje distribuido para evitar el monopolio y realmente aplicar el término de preservación a escala en beneficio colectivo y no particular en el juego de precios para que la definición pueda realmente operar en la escala que se necesita.

El cuadro general que engloba a los diferentes modelos de preservación digital combinan políticas, acciones y estrategias con el fin de asegurar el acceso continuo de materiales digitales en el tiempo donde la preservación digital central y los modelos o estándares como la referencia OAIS definen a grandes rasgos los elementos, entidades y funciones que se realizan para que un archivo digital sea preservado. Los modelos de orden distribuido, federado o descentralizados con base en la replicación múltiple y dispersión geográfica integran una red de servicios de preservación digital externa compuesta por varios participantes que requieren de la coordinación de los diferentes repositorios de contenidos preservados tanto a nivel técnico como administrativo.

A continuación describiremos de manera específica su composición desde un punto de vista conceptual con enfoque semántico, sintáctico y pragmático de los elementos que integran el modelo de preservación digital distribuida como sistema de información de composición referida de datos, entidades, protocolos y correlaciones. Trataremos de definir de manera general los elementos integrales a nivel conceptual y de manera necesaria elemental en aspectos técnicos que describen la composición del modelo distribuido basándonos principalmente en los elementos para su función distribuida y las actividades principales de preservación digital. En una estrategia de preservación digital distribuida los temas de políticas de gobernanza interinstitucional, políticas normativas de uso de material y guías, refieren a un tipo *sui generis* ya que la gran mayoría de las políticas y normativas existentes en esta estrategia distribuida refieren a la personalización de necesidades particulares de cada organización o grupo de ellas como en su estrategia raíz de preservación digital centralizada. Como mencionamos anteriormente y acorde al estudio de análisis de políticas de preservación digital de Sheldon (2013) en la actualidad una gran cantidad de organizaciones no cuentan con políticas de preservación digital o estas son limitadas, sin embargo y de manera contradictoria se realizan acciones de preservación digital a la par del desarrollo de las políticas de preservación. En cuanto a guías, debemos mencionar a "A Guide to Distributed Digital Preservation" de Educopia Institute por miembros de MetaArchive Cooperative (Skinner y Schultz, 2010) representa al día de hoy el único documento guía especializada sobre el tema de preservación digital distribuida. En este documento mediante una colección de ensayos especializados se puede obtener una visión general de la preservación digital distribuida de tipo teórico y práctico para implementar o unirse a una red de preservación digital distribuida según las experiencias de redes bajo el sistema LOCKSS.

Uno de los temas complejos a tratar por la importancia conlleva a nivel de organización y definición de actividades de rol de responsabilidades, coordinación y cooperación de diferentes miembros de una red de trabajo dentro de una estrategia de preservación digital distribuida es la gobernanza. Para definirla necesitamos referirnos a el término de gobierno que dentro del ámbito de una estructura de estado se puede definir como el órgano o autoridad que tiene por funciones la de administrar, dar dirección y controlar a sus instituciones. En la actividad del gobierno denominada gobernanza las acciones del gobierno son ejercidas mediante la definición de una estructura de trabajo (en el caso de un modelo de estado) de

varias personas con funciones definidas como las del presidente o primer ministro, y distintos secretarios o ministros sea el caso. En el ámbito empresarial la gobernanza corresponde en algunos casos a las actividades desarrolladas por una estructura interna como la junta de directores que lleva a cabo la toma de decisiones sobre los intereses de los involucrados en la empresa.

En lo que corresponde a la preservación digital distribuida la estructura de gobernanza Walters (2010) menciona que puede ser aplicada en redes distribuidas por algunas razones que principalmente pueden ser para crear y establecer un órgano de toma de decisiones, para proporcionar de un medio de comunicación que coadyuve a fomentar el compromiso de colaboración entre los contribuyentes de contenido en redes distribuidas, y por último para asegurar la supervisión de una adecuada gestión de tipo estratégico, fiscal y operacional. Acorde a estas razones, el autor cita y adecua el trabajo de (Provan y Kenis, 2007) quienes refieren a tres modelos de gobernanza en redes de trabajo. El modelo gobernado por participantes (the participant-governed model) el cual es la forma más simple de gobernanza pues es ejercida por cada uno de los miembros de red sin necesidad de una entidad separada de gobernanza, lo que involucra a todos los participantes de una manera descentralizada ya sea formal o informal acorde al interés comunal. Por otro lado tenemos el modelo administrativo de red (the network administrative model) que es un modelo centralizado que tiene por fin la creación de un órgano de administración con la finalidad de dar gobernanza a la red y sus actividades. Y por último tenemos a el modelo de organización principal (the lead organization model) que se lleva a cabo para dar orden a los modelos descentralizados y donde es necesario establecer un orden principal mediante un miembro participante o organización líder donde las principales actividades y toma de decisiones se llevan a cabo a través de él (Skinner y Schultz, 2010).

En la práctica existen varios ejemplos de redes de trabajo enfocadas a preservación digital distribuida como es el caso de ADPNet<sup>43</sup> que ejerce una gobernanza basada en el modelo administrativo que está definido por un comité directivo que está integrado por miembros de la red que se encarga de definir y llevar a cabo las políticas. Por otro lado cuenta con un comité técnico que está enfocado a los temas de software y hardware. Smallwood (2014) por su parte propone el término de gobernanza de información (information governance) que

---

<sup>43</sup> <http://www.adpn.org>

sugiere que la información digital que se gestiona en una organización privada o pública debe tratarse bajo este concepto de gobierno de la misma y que debe considerarse en conjunto con las políticas de la organización como lo son el portafolio de tecnología, privacidad y protocolos de seguridad. Podemos observar que el entretreído de la gobernanza y la ejecución de políticas de una red distribuida puede personalizarse acorde a las necesidades propias de una o varias organizaciones, y por tanto se deben estudiar y analizar los modelos de casos activos de gobernanza distribuida para sopesar las necesidades propias de un proyecto de preservación digital.

## **2.3 México en la gestión de archivos y preservación digital**

### **2.3.1 Contexto actual general**

México es un país que en los últimos años normaliza la brecha digital en el sector de archivos de diferentes instituciones a nivel nacional. La gestión de archivos digitales de ámbito gubernamental desde el año 2002 mantiene un desarrollo constante ante las necesidades de integración de prácticas y estándares internacionales en ámbitos de transparencia digital. A partir del proceso normativo propuesto en La Ley Federal de Transparencia y Acceso a la Información Pública Gubernamental<sup>44</sup> y otras leyes con modificaciones posteriores a la Constitución de los Estados Unidos Mexicanos sobre el tema de gestión de información pública y archivos, se activó el detonador para la actualización y organización de archivos físicos institucionales de ámbito público. Esta iniciativa representó el apoyo y empuje a nivel federal para el inicio formal de la normalización digital y organización de archivos institucionales en México. Acorde a este marco legal el Archivo General de la Nación (AGN) de México, definió los lineamientos de apoyo para la organización y conservación de archivos de la administración pública federal, las dependencias y entidades de gobierno federal mediante la implementación de cuadros de clasificación por funciones, catálogos de vigencia documental, sistemas de gestión documental, así como de un esquema de organización interna acordes a la época y necesidades de cada institución (Barnard, 2009).

A partir del año 2012 en México se publica la Ley Federal de Archivos<sup>45</sup> que define la

---

<sup>44</sup> Ley Federal de Transparencia y Acceso a la Información [http://www.diputados.gob.mx/LeyesBiblio/pdf/244\\_140714.pdf](http://www.diputados.gob.mx/LeyesBiblio/pdf/244_140714.pdf)

<sup>45</sup> Ley Federal de Archivos <http://www.agn.gob.mx/pt/LFA.pdf>

obligación formal para la administración pública federal de atender, organizar y conservar sus archivos (análogos y digitales) en posesión de distintos órganos federales fomentando su resguardo, difusión y acceso a todo archivo privado que sea de relevancia histórica, social, técnica, científica o cultural para el país. La Ley Federal de Archivos se establece como la única política existente y vigente desde su publicación para abordar la gestión de documentos digitales y en específico lo que refiere a su organización, conservación, acceso y preservación. El ámbito de aplicación federal de esta ley actúa y aplica de manera limitada sólo a grupos de dependencias federales, dejando al margen de posibilidades a archivos y bibliotecas de orden municipal y estatal. Estas iniciativas y leyes forman parte de un avance sustancial en materia de políticas y normativa archivística digital de México, ya que aborda al enfoque de gestión, conservación, acceso y preservación digital con influencia de aplicación en distintas capas, organismos e instituciones de orden público para su actuación en busca de mejorar la preservación del patrimonio a nivel nacional.

Acorde al enfoque vigente los archivos de tipo histórico de gobierno y administración pública federal en México son gestionados concretamente por el Archivo General de la Nación (AGN) que tiene la responsabilidad de custodiar, conservar y preservar el patrimonio documental histórico del país integrado principalmente por los archivos y documentos de valor histórico antiguos y nuevos documentos que se generan en la administración pública y alrededor del país. En la última década el AGN ha estado inmerso en proyectos de digitalización del archivo histórico de México y de la gestión del material de origen digital para la creación formal del repositorio digital del AGN con la finalidad de organizar de manera segura el acceso y la preservación del mismo. De manera paralela a estas actividades y proyectos el AGN mantiene una actividad de divulgación y cabildeo a nivel nacional con el apoyo de diversas instituciones de educación superior mediante el ejercicio de foros y reuniones con representantes de archivos, bibliotecas e instituciones de orden federal y estatal sobre la importancia de una gestión archivística homologada a nivel general (nacional) en México resaltando las virtudes que esto conlleva evitando la pérdida del patrimonio general de archivos y enriquecer los indicadores del país mediante la preservación del patrimonio documental. Este tenor normativo podría colocar en una situación susceptible al patrimonio nacional documental de diferentes instituciones (mas allá del ámbito de administración pública gubernamental) que de manera ulterior e indirecta les podría sugerir la normalización generalizada de sus archivos y depósitos digitales. Las instituciones de educación superior

pública representadas por universidades, consejos y centros de investigación alrededor del país que disponen bajo el índice de recursos públicos de documentos de interés al patrimonio nacional tales como disertaciones, tesis o artículos científicos que cumplen anticipadamente con el reclamo de documentos de patrimonio documental del país con necesidades de preservación y donde las figuras de autonomía educativa con actuación en los IES públicos podrían referir especificaciones excepcionales para la gestión de sus activos digitales.

Instituciones de educación superior (IES) y diversos centros de investigación en México han realizado esfuerzos a vanguardia mediante la implementación y gestión de BiDis (bibliotecas digitales), repositorios digitales (RD), repositorios institucionales (RI) o repositorios de tesis digitales (RTD) como respuesta a las necesidades de colaboración de investigación e interoperabilidad de documentos de investigación (artículos, tesis, preprints, etc) y metadatos de interés para la transparencia nacional. Ante esta situación se han ido priorizando previo a los marcos legales referentes, a llevar a cabo una gestión y organización de archivos digitales mediante esfuerzos independientes para articular la conservación y acceso de sus archivos de documentos de investigación. Tales esfuerzos en algunos casos se pueden colocar como una respuesta espontánea a las necesidades propias de cada institución de organizar, acceder, colaborar y conservar la producción de su conocimiento a manos de investigadores, profesorado y alumnado de diversas IES en México. En el caso específico de la investigación hasta el mes de Mayo del 2014 México contaba en su registro del sistema nacional de investigadores (SNI) con una plantilla de 43 mil investigadores en todo el país, los cuales publican 12 mil artículos por año (Villa, 2014). Tales índices muestran de manera prioritaria la necesidad de formalizar los canales de depósito, difusión y soporte a la investigación científica en instituciones de educación superior y centros de investigación bajo infraestructuras de gestión digital que alojan y gestionan documentos de investigación en su acceso, resguardo y preservación.

Cabe señalar que aunque México se encuentra dentro de los países que ha aumentado el número de sus publicaciones científicas de acuerdo a datos del Consejo Nacional de Ciencia y Tecnología (CONACYT), estos indicadores sólo representan cifras cuantitativas y no datos cualitativos sobre la forma de acceso, gestión y preservación digital de las mismas en sus archivos o repositorios digitales. Esta situación representa una área de oportunidad y razón tangible para realizar un análisis del estado, aplicación de procedimientos y modelos

necesarios para aplicar al ámbito la preservación digital de publicaciones y documentos digitales del sector de IES en México como patrimonio cultural con prioridad de gestión formal de su uso, acceso y preservación digital a largo plazo.

Voutssás (2012) menciona que en Latinoamérica la mayoría de los países están trabajando activamente en las bibliotecas y archivos digitales, pero con poca o ninguna consideración por el futuro digital a largo plazo. Abordo de esto y aunque las consideraciones del estado actual de la preservación digital en México en el sector de educación superior e investigación plantean que en la última década con esfuerzos considerables a nivel de infraestructuras tecnológicas y redes mediante la implementación y uso de repositorios o bibliotecas digitales, estas iniciativas nuevamente no es un indicador de que existan estrategias formales de preservación digital sólidas basadas en estándares internacionales y desarrollo de políticas específicas para este fin. La importancia de consolidar el marco legal de apoyo hacia el tema de la gestión de archivos digitales con énfasis en el ámbito digital de valor patrimonial nacional servirá de inspiración para desarrollar de igual manera acciones que ayuden al formalizar la cultura de aseguramiento y preservación del conocimiento generado en el presente para generaciones futuras, así como para consolidar el desarrollo de políticas institucionales integradas a estrategias de preservación digital específicas para cada institución y el desarrollo de aspectos metodológicos que ayuden a consolidar esta agenda digital. En este sentido es necesario trabajar en procesos de implementación de la cultura de modelos de preservación digital en instituciones de educación superior de México, así como el crear un modelo de trabajo que corresponda a la tendencia actual enfocada al desarrollo y gestión de repositorios de documentos de investigación con contenido científico (tesis, artículos, informes, documentos de investigación, etc) en el ámbito universitario de México en función de filosofías de trabajo de colaboración, economía de escala y distribución de esfuerzos en la gestión de preservación digital.

Es importante abordar la definición y tipología de instituciones de educación superior pública ante el uso de tecnologías e infraestructuras en gestión de archivos digitales, su preservación a largo plazo, los marcos legales y los parámetros formales de seguridad para su custodia digital. A partir de este estudio y análisis de situación se podrá colocar un balance analítico sobre los estándares internacionales de normalización para el tratamiento de objetos digitales, así como de las posibilidades de colaboración y distribución del almacenamiento de archivos

digitales con fines de preservación para definir un modelo sólido de preservación digital de archivos a largo plazo para el sector.

### **2.3.2 Instituciones de educación superior de México y la gestión documental digital**

El panorama genérico sobre la estructura general de los niveles de educación en México nos ayuda a integrar un balance cuantitativo de instituciones por sistemas y alcances de influencia para referir los esfuerzos específicos de la gestión de documentos digitales de investigación en instituciones de educación superior del subsistema estatal y federal. La Secretaría de Educación Pública en México (SEP) es la Secretaría de Estado que tiene por objetivo crear las condiciones necesarias para asegurar una educación de calidad en México en sus diferentes sistemas de oferta educativa. La SEP en México además tiene competencia directa en la manutención de la cultura y sus centros de exhibición, así como del control de registros de derechos de autor y registro de marcas en el país. La estructura general del sistema de educación pública vigente en México se divide en el sistema de educación básica que incluye al nivel preescolar con 3 grados para edades de 3 a 5 años, nivel primaria con 6 grados para edades de 6 a 11 años y el nivel secundaria con 3 grados de 12 a 15 años. El sistema de educación media superior también llamada preparatoria o bachillerato se integra por 2 y 3 grados en edades de 16 a 18 años. Por último el sistema de educación superior o sistema de universidades públicas que incluye estudios de grado y posgrado universitario.

La oferta del sistema de educación superior en México está compuesta de 847 Instituciones de educación superior (IES) que se subdividen en 11 tipos de subsistemas (tabla 6) de acuerdo con datos de la Secretaría de Educación Pública (SEP, 2017). Aunque en México se cuenta con un alto y variado número de instituciones públicas y ofertas educativas en el ámbito de educación superior que en número pueden superar incluso a diversos países de América Latina y el resto del mundo, la cantidad de instituciones no es un factor de relevancia con relación a los diversos indicadores de calidad que ayudan a la clasificación académica de universidades a nivel mundial. Las clasificaciones de relevancia mundial en la educación superior responden a metodologías de objetivos bibliométricos y no bibliométricos. Los objetivos bibliométricos incluyen a publicaciones indexadas, investigaciones académicas, citas de trabajos, galardones, entre otros. Mientras que los objetivos no bibliométricos consideran el número de alumnos inscritos, egresados, académicos con doctorado, cursos de

calidad y mas. Estas situaciones nos dejan ver que es importante considerar nuevamente la gestión oportuna de sistemas que permitan dinamizar la producción y la gestión de documentos de investigación en el ámbito de la educación superior para atender las tendencias cualitativas de reconocimiento mundial.

Tabla 6. IES de México por subsistema.

Tipo	Descripción	Centros
Universidades Públicas Federales	Las Universidades Públicas Federales son las que mantienen su gestión a nivel de organismo público federal con operaciones centrales.	9
Universidades Públicas Estatales	Este subsistema universitario de Universidades Públicas Estatales está integrado por instituciones universitarias de educación superior creadas bajo decretos de congresos locales o estatales, y regidos bajo figuras jurídicas como organismos públicos descentralizados	34
Institutos Tecnológicos Federales	Forman parte del Sistema Nacional de Institutos Tecnológicos (SNIT) son coordinados por la Dirección General de Educación Superior Tecnológica, los Institutos Tecnológicos Federales mantienen soporte del gobierno federal central.	132
Institutos Tecnológicos Descentralizados	Forman parte del Sistema Nacional de Institutos Tecnológicos (SNIT) son coordinados por la Dirección General de Educación Superior Tecnológica, los Institutos Tecnológicos Descentralizados mantienen soporte de gobierno estatales o locales.	130
Universidades Públicas Estatales con Apoyo Solidario	Estas universidades reciben aportaciones del programa de presupuesto de los gobiernos federales, así mismo recibe apoyo del Gobierno Federal.	23
Universidades Tecnológicas	Este tipo de universidades ofrece formación intensiva técnica de 2 años, con el objetivo de incorporarse al campo laboral o en su caso continuar con estudios de licenciatura.	104
Universidades Politécnicas	Se crearon en el año 2001 para ofrecer estudios de ingeniería y licenciatura en modo de especialidad con orientación hacia la investigación aplicada al desarrollo tecnológico de colaboración con el sector público y social del país.	50
Universidades Interculturales	Tienen por objetivo ofrecer formación académica superior vinculada a revalorización de los conocimientos de pueblos indígenas de México, así como de sintetizar al conocimiento científico, la preservación de sus costumbres, lenguas y tradiciones.	12
Centros Públicos de Investigación	Tienen por objetivo el desarrollo de investigación científica y tecnológica con la finalidad de divulgar, vincular y crear mecanismos de aplicación en la sociedad y sector productivo.	6
Educación Normal Superior	Este tipo de centro de educación superior tiene como objetivo la formación del profesorado para los niveles de preescolar, primaria y secundaria.	261
Otras Instituciones Públicas	Son instituciones que no pueden ser clasificadas en algún tipo, ya que tienen características particulares y de educación diversa.	86

La gestión de documentos digitales con fines de investigación de algunas IES en México tiene al día de hoy avances importantes y también algunas lagunas con respecto a su gestión independiente. Como avances podemos citar sus esfuerzos en la implementación de depósitos digitales de orden universitario como bibliotecas digitales (BiDi), repositorios institucionales (RI), repositorios de tesis digitales (RTD) y depósitos de colecciones digitales (DCD). Ante los avances de infraestructuras de depósito digital aún queda pendiente la normalización general del conocimiento, disposición e importancia de los productos académicos en las universidades mexicanas tal como lo menciona Arreola (2012): “Aún cuando las IES mexicanas realizan esfuerzos constantes para distribuir entre sus usuarios información relevante de I&D en línea, las experiencias han sido parciales y contrastantes debido a múltiples dificultades (técnicas, operativas, presupuestarias y colaborativas) que inciden en la selección, calidad, accesibilidad, procesos de adquisición, organización, mantenimiento y preservación de la información electrónica/digital; especialmente la generada por las propias IES. Pese a ello, la mayoría cuenta ya con la infraestructura de redes y la tecnología de información y comunicación (TIC) que hace posible el acceso de sus comunidades de usuarios a sistemas digitales de información en línea, bases de datos, suscripción de revistas especializadas y recursos propios desarrollados por sus integrantes.” (p.2).

Aunque los esfuerzos y avances de las IES por potenciar la visibilidad, gestión y difusión de la investigación académica en México siguen un rumbo prioritario, el producto de la actividad no se ve reflejado de manera específica en la proyección internacional de este campo. Muñoz (2011) afirma que en México hace falta una política nueva que estimule la participación la universidades a escala internacional en el terreno de la investigación con calidad, coordinación y capacidad de atracción para enlazar a las grandes universidades Mexicanas con las del extranjero y por ende fortalecer una red de conexión con el todo el sistema universitario en México (p.32). Datos como estos sugieren que al igual de las infraestructuras tecnológicas y difusión, existe trabajo por hacer en lo que se refiere a la proyección y prestigio de investigación e interoperabilidad internacional de los centros de educación superior en México.

En lo que toca al ámbito de gestión de documentos digitales con fines de investigación un claro indicador de iniciativas de mejora en lo que se refiere a la visibilidad y proyección de la investigación en México es el depósito especializado y normalizado de los productos de

investigación en repositorios digitales de investigación institucional. Lluca (2012) menciona que cuando una institución alberga en un repositorio su producción de investigación, consigue más visibilidad, difusión y prestigio de sus autores; así como de su producción documental y científica; y de sus centros de investigación y divulgación a un entorno internacional (p.4). Es importante mencionar que algunas instituciones de educación e investigación en México por iniciativa propia y con miras de atraer visibilidad de sus productos de investigación han implementado estándares de acceso abierto en sus repositorios de documentos digitales que al contar con los estándares de normalización, calidad e implementación han podido integrarse en acceso abierto a la red mundial de conocimiento e investigación como avance particular ante una nueva perspectiva de visualización internacional.

Según datos de OpenDOAR<sup>46</sup> (2017) México cuenta con 30 repositorios en acceso abierto, de los cuales 19 pertenecen a instituciones públicas y 9 a privadas. Estos repositorios en acceso abierto con investigación académica de centros, institutos y universidades representan un 3,3% del total de 900 instituciones de educación superior públicos y privados que podrían disponer de este tipo de recursos de gestión y depósito de objetos digitales. El crecimiento en el número de repositorios abiertos en México se acentúa en los años 2011 y 2012 con un 40%. Hay que mencionar que a partir de la iniciativa y publicación los lineamientos generales y técnicos para la creación del repositorio nacional e institucional así como de los programas de apoyo para su creación (CONACYT, 2014) durante el año 2015 y 2016 se inicia una nueva etapa de desarrollo formalizado en el ámbito de repositorios en México por lo que la tendencia de creación de repositorios y acceso abierto en universidades e institutos de México está en pleno desarrollo y oportunidad. En países como España, el 49,2% de sus universidades cuentan con repositorios en acceso abierto y forman parte de los principales países en la gestión de este tipo de iniciativas según datos del 2017 de Webometrics<sup>47</sup> y ROAR<sup>48</sup>.

Un dato importante con respecto al acceso abierto en la investigación universitaria depositada en repositorios mexicanos se presentó en el año 2013 por la senadora Ana Lilia Herrera Anzaldo<sup>49</sup> en colaboración y apoyo de la Universidad del Estado de México. La iniciativa y

---

<sup>46</sup> Directorio de repositorios en acceso abierto <http://www.opendoar.org>

<sup>47</sup> Ranking web de repositorios <http://www.webometrics.info/es>

<sup>48</sup> Registro de repositorios con acceso abierto <http://roar.eprints.org>

<sup>49</sup> <http://www.senado.gob.mx/?ver=sp&mn=2&sm=2&id=39774>

proyecto de decreto por el que se reforman diversos artículos de la Ley General de Ciencia y Tecnología y la Ley General de Educación pretende en específico establecer que toda investigación realizada en instituciones públicas, con recursos públicos o con el uso de alguna infraestructura financiada por recursos públicos, estén disponibles en acceso abierto a través de plataformas en línea (repositorios). Dicha iniciativa responde por igual a un estudio realizado por la Universidad Nacional Autónoma de México (UNAM) donde se muestra que 1 de cada 10 investigaciones científicas realizadas en universidades de México, 9 provienen de instituciones universitarias públicas. Dicha propuesta proyecta la importancia de legislar más apoyos para las universidades públicas mexicanas y para que se pueda disponer a nivel de ley el acceso abierto de todas las investigaciones (tesis, artículos científicos, ponencias, etc.) de las mismas para de esta manera incentivar a mejorar la cantidad y la calidad de las investigaciones científicas fomentando el intercambio de conocimiento científico en México y a nivel internacional.

El estado actual de los repositorios institucionales, abiertos o especializados en México mantiene un centralismo de origen principalmente en la capital del país el Distrito Federal y ciudades importantes del interior de la república como Estado de México, Puebla, Guadalajara y Nuevo León. Podemos mencionar que del total de 30 repositorios en acceso abierto existentes en México el 63% pertenece a universidades e instituciones públicas y el resto al sector privado universitario. Esto nos reduce a un número de 19 repositorios de universidades públicas de las cuales 6 repositorios pertenecen a la UNAM y 4 a universidades que se encuentran en el centro del país como lo son el Instituto Politécnico Nacional (IPN), Universidad Autónoma Metropolitana (UAM) y La Universidad Autónoma del Estado de México (UAEM). Los 9 repositorios en abierto restantes pertenecen a instituciones que se ubican en el interior de la República Mexicana. Para tener una mejor perspectiva sobre los repositorios en México debemos mencionar que estos cuentan con diferentes tipos de modelos de repositorios y gestión de documentos digitales de investigación. Como ejemplo podemos mencionar el modelo distribuido y centralizado del repositorio institucional de la UNAM que está integrado por un grupo de 10 repositorios (6 en abierto y 4 internos) denominado Red de Acervos Digitales UNAM<sup>50</sup>. También de la UNAM surgió la iniciativa de un modelo de

---

<sup>50</sup> RAD UNAM <http://www.rad.unam.mx>

catálogo y directorio interinstitucional a nivel nacional e internacional LATINDEX<sup>51</sup>. Otro modelo a mencionar es el del repositorio interinstitucional centralizado REDALYC<sup>52</sup> de la Universidad Autónoma del Estado de México, así como la red de repositorios distribuidos de metadatos REMERI<sup>53</sup>. Existen además otros modelos de repositorios institucionales de universidades, centros de investigación y dependencias públicas mexicanas que están dados de alta como proveedores de datos en servicios con interoperabilidad de distintos servicios internacionales de metadatos en abierto y otros en privado.

De tal manera el estado actual de la gestión de documentos digitales de algunas universidades y centros de investigación en México de ámbito público a primera vista y principalmente en las de ámbito federal, se percibe que trabajan ya con algunas estrategias y elementos básicos de gestión digital como la organización, depósito y acceso abierto de documentos digitales. Estas actividades los prospectan a considerar las necesidades a priori de cara al futuro de la planificación de una estrategia formal de normalización y preservación digital. Cabe señalar que la mayoría de instituciones de educación superior en México del subsistema de IES estatales no cuentan a la fecha con repositorios dedicados de manera abierta. Definir una estrategia institucional sobre la gestión de documentos digitales mediante modelos de servicios de archivo o repositorios institucionales es la antesala obligada para quien considera el plantear la dirección hacia el tratamiento y mantenimiento de los archivos en una estrategia de preservación digital. Esta coyuntura de disparidad en la implementación tecnológica y políticas sobre la gestión de documentos digitales hace una brecha de desconocimiento que debe tratarse bajo un estudio analítico de necesidades, prioridades, nivel de gestión, etc. Estos temas pueden redefinir el porqué algunas universidades en México tienen más valoración y prestigio ante temas de visibilidad de trabajo en el área de investigación, así como de la implementación de un orden de publicación o depósitos institucionales propios o centrales de sus productos de la actividad investigadora o la dispersión de los mismos en repositorios de terceros. Si México pretende integrarse en el mercado internacional del conocimiento, se requieren en principio tres cosas: cambiar la visión del Estado hacia la educación superior, la ciencia y la tecnología. Para así conjugar en una sola política a la ciencia y a la educación superior con miras de largo plazo (Muñoz, 2011).

---

<sup>51</sup> Sistema regional de información en línea para revistas científicas de América Latina, el Caribe, España y Portugal. Directorio, catálogo e índice. <http://www.latindex.unam.mx>

<sup>52</sup> Red de Revistas Científicas de América Latina y el Caribe, España y Portugal <http://www.redalyc.org/home.oa>

<sup>53</sup> Red federada de Repositorios de acceso abierto de las Instituciones Mexicanas de Educación Superior <http://www.remeri.org.mx/portal/index.html>

### **2.3.3 Preservación digital en México**

La preservación digital en México está en un proceso de abstracción a nivel cultural, de valor y conocimiento disciplinar que hace que su implementación formal sea inexistente. Voutsáss (2009) afirma con respecto a la preservación del patrimonio documental digital: “En México, no contamos con ningún proyecto, programa, plan, iniciativa ni organización que actualmente trabaje en este rubro a gran escala. Todos los que existen son proyectos puntuales a nivel de ciertas instituciones puntuales. Reitero, nada estratégico ni de alcance nacional.” (p.189). A la par de las políticas y leyes de gobierno en México sobre transparencia, acceso a la información y archivos que sugieren la obligación de implementar estrategias formales para la gestión de archivos digitales de orden público se trabaja desde el rubro de la investigación superior nacional en divulgar sobre la necesidad de contar con una cultura de preservación del patrimonio digital general que cubra más allá del nivel del sector de administración pública federal o el Archivo General de la Nación, y que reconozca el esfuerzo individual que realizan actualmente algunas instituciones de educación superior e investigación con las acciones de preservación digital de sus acervos digitales y desarrollo de políticas individuales al respecto. Serra (2014) afirma sintéticamente que si el archivo de una organización pretende lograr una preservación digital exitosa, lo primero que deberá hacer es desarrollar una exhaustiva política de preservación digital basada en modelos y estándares confiables que incluya el compromiso institucional en el proyecto de manera permanente. En este sentido las políticas propias de preservación digital en sectores como el de la educación superior, la investigación científica o disciplinas artísticas deberían contar con un respaldo nacional que los ayude y apoye a considerar, capacitar e implementar políticas de preservación digital de manera formal desde procesos iniciales de construcción de sus archivos digitales con la finalidad de poner en contexto su creación y procedencia para pasar al siguiente nivel o necesidad que refiere una estrategia formal e integral de preservación digital a largo plazo del conocimiento producido en el país.

Para que en México se considere la existencia de una cultura de preservación digital lo primero que debe ocurrir es actividad puntual de gestión digital de acuerdo a distintos momentos de necesidades institucionales. Por un lado algunos investigadores mencionan que la estrategia de preservación digital debe ser considerada desde el inicio de la creación de un archivo, aunque otras consideraciones definen que una estrategia de preservación digital inicia

con la implementación de un sistema de gestión documental. Lo que si está claro es que la preservación digital tiene una cualidad activa de principio y fin, por tanto, es flexible a su aplicación siempre y cuando se consideren los diferentes factores que repercuten en la implementación de una estrategia de este tipo como lo son los culturales, tecnológicos, metodológicos, económicos, legales y sociales (Voutsáss, 2012). Considerar la integración adecuada de factores de influencia, necesidades y soluciones tecnológicas en un mismo sentido ayudará a consolidar a diferentes escalas la adopción de la cultura y estrategia de preservación digital.

En México existen diferentes proyectos nacionales de datos abiertos o estrategias digitales en las que se omiten políticas, metas u objetivos relacionados con la preservación digital aun cuando esta se encuentra como una consideración de asunto estratégico en universidades a nivel internacional y en donde México no debería hacer la excepción como lo menciona Arreola (2012): “La preservación y disponibilidad de los productos científicos y tecnológicos que las IES producen han sido señalados como asuntos urgentes de atender en México según indicadores y análisis comparativos en rubros sobre desarrollo humano, educativo, económico, tecnológico, sustentable, entre otros.” (pg.17). La UNESCO refiere a que todas las actividades de preservación y conservación forman parte de una prioridad de desarrollo, por lo que es imprescindible la elaboración de leyes y derechos claros sobre la gestión del material digital mediante el desarrollo de políticas públicas que orienten a la creación de estrategias de gobernanza de archivo de manera transparente mediante garantías legales que le den soporte. Los materiales digitales son activos de las organizaciones, y como tal, deberían estar sujetos al escrutinio de la ciudadanía y de los órganos de control. Este tipo de iniciativas también deben hacerse extensibles a profesionales para que inviertan en este rubro y consolidar un ejercicio hacia la preservación digital de alianzas para optimizar el compromiso tanto en el ámbito económico como en el ámbito social y cultural de preservación digital como disciplina de apoyo con miras al largo plazo. En este sentido, se puede decir que México atraviesa aún por la etapa donde digitaliza para conservar y no para preservar, y por tanto con la inclusión de todas las actividades paralelas necesarias como políticas, normativas y desarrollo de tecnología podría aspirar a cambiar esa visión a mediano y largo plazo. Voutsáss (2014) afirma que para el nivel en el que se encuentra hoy en día la gestión digital con miras a la preservación se requiere de un 90% de método y control, y solo un 10% de aspectos técnicos.

### **2.3.4 Marco legal, gestión de archivos y preservación digital en México**

Aunque existe un alto antecedente sobre el marco legal que implica obligaciones y gestiones sobre las publicaciones, documentos y autores de material documental físico en México como algunos autores como Voutáss (2009) han publicado de manera excepcional y completa sobre el tema de depósito legal, derecho de copia o copyright en México, debemos enfocarnos a grandes rasgos en la historia contemporánea que afecta en principio y a modo de leyes la gestión documental digital para mantener una línea puntual y actualizada a los cambios digitales. Dicho lo anterior y tal como se mencionó antes la Ley Federal de Archivos es hasta ahora la única política que dispone México acerca de la preservación de archivos digitales aunque en otras disposiciones jurídicas se habla sobre la gestión de los materiales digitales en ninguna se menciona de manera puntual o específica acerca de los controles, procesos o políticas necesarias para su producción, manejo y preservación. Así mismo como acotamos anteriormente la normalización sobre la elaboración, clasificación, catalogación, capacitación y gestión de archivos digitales (a nivel de gobierno federal) en México trasciende a partir a dos momentos legislativos como lo son la Ley Federal de Transparencia y Acceso a la Información, y posteriormente con la Ley Federal de Archivos.

Además del esfuerzo e influencia federal que representan las leyes federales tanto de archivo como de transparencia, 28 entidades federativas de México cuentan con leyes propias para la gestión de archivos. En la ocupación de la gestión de archivos aún se presentan irregularidades en lo que respecta a sus normativas a nivel de municipios en la labor archivística. La formalización y expedición de una Ley General de Archivos pretende integrar, homogeneizar y estandarizar los procesos de archivo a nivel nacional. De manera general definiremos las composiciones, definiciones y pertinencias objetivas de las mismas para el ámbito documental digital en México. Es preciso mencionar lo que apunta Arreola (2012) con lo que respecta a los derechos de los autores y las universidades en cuanto a la visualización y consulta de contenidos digitales en instituciones de educación superior en México (IES): “Cada IES opera en función de una legislación propia que la mayoría de las veces no contempla lo relativo a contenidos digitales, seguridad y protección de datos, porque no hay legislación nacional —ni estatal— en la materia”. (p.34).

## ***La Ley Federal de Transparencia y Acceso a la Información Pública Gubernamental***

Para atender el acceso y la gestión de información de archivos públicos del gobierno federal en México se creó la Ley Federal de Transparencia y Acceso a la Información Pública Gubernamental (LFTAIPG)<sup>54</sup> que fue publicada en el año 2002 y reformada en mes de Septiembre del año 2014. De manera general esta ley refiere a las obligaciones de acceso de información pública de orden federal y concibe también obligaciones necesarias para ofrecer el acceso a la información de normalización, organización y creación de archivos públicos. Esta ley define obligaciones del órgano e institutos para la elaboración de metodología de trabajo concernientes a la definición de sistemas de gestión documental. En las disposiciones generales de su artículo primero refiere su finalidad:

*Artículo 1. “La presente Ley es de orden público. Tiene como finalidad proveer lo necesario para garantizar el acceso de toda persona a la información en posesión de los Poderes de la Unión, los órganos constitucionales autónomos o con autonomía legal, y cualquier otra entidad federal.”*

En lo que refiere a la gestión, clasificación y conservación de archivos encontramos menciones y definiciones en el Artículo 29 apartado 5 que escribe:

*Artículo 29 - V: “Establecer y supervisar la aplicación de los criterios específicos para la dependencia o entidad, en materia de clasificación y conservación de los documentos administrativos, así como la organización de archivos, de conformidad con los lineamientos expedidos por el Instituto y el Archivo General de la Nación, según corresponda”*

Sobre las responsabilidades del órgano responsable para la elaboración de criterios de elaboración para actividades de catalogación, clasificación y conservación de documentos se recogen en el Artículo 32:

*Artículo 32. “Corresponderá al Archivo General de la Nación elaborar, en coordinación con el Instituto, los criterios para la catalogación, clasificación y conservación de los documentos administrativos, así como la organización de archivos de las dependencias y entidades. Dichos criterios tomarán en cuenta los estándares y mejores prácticas internacionales en la materia. Los titulares de las*

---

<sup>54</sup> [http://www.diputados.gob.mx/LeyesBiblio/pdf/244\\_140714.pdf](http://www.diputados.gob.mx/LeyesBiblio/pdf/244_140714.pdf)

*dependencias y entidades, de conformidad con las disposiciones aplicables, deberán asegurar el adecuado funcionamiento de los archivos. Asimismo, deberán elaborar y poner a disposición del público una guía simple de sus sistemas de clasificación y catalogación, así como de la organización del archivo.”.*

Sin entrar en interpretaciones mayores de orden legislativo y jurisprudencia referimos los elementos que tienen incidencia en las estructuras de archivos y necesidades de gestión documental ya sea como obligación y por ende de la integración de un trabajo de especialización ante dicha ley. También se recoge en esta ley la obligación de que cada entidad implicada en estas tareas defina un reglamento específico de transparencia, acceso a la información pública y protección de datos personales como mandato y referido para la generación de un marco de trabajo y entrega de reportes de transparencia de las dependencias, institutos u órganos implicados. Entendamos que esta ley además de recoger obligaciones sugiere la actividad de profesionalización en el área archivística (catalogación, clasificación, organización, etc.) a nivel dependencias y organismos federales en México, lo que representa un avance en esta materia y debe ser reconocido literalmente como ello ante la heterogeneidad que representa la estructura federal.

### ***Ley Federal de Archivos México***

En el año 2012 se crea la Ley Federal de Archivos que especializa la obligación de poderes, organismos e instituciones federales de México a la gestión obligada de sus archivos en poder para que puedan ser manejados con fines de acceso y preservación. Analizaremos la ley acorde a los temas y artículos relevantes para el ámbito de la preservación digital en México. La ley en sus disposiciones generales refiere que su finalidad es el fomento del resguardo, difusión y acceso de archivos privados en su artículo primero:

*Artículo 1. “El objeto de esta Ley es establecer las disposiciones que permitan la organización y conservación de los archivos en posesión de los Poderes de la Unión, los organismos constitucionales autónomos y los organismos con autonomía legal, así como establecer los mecanismos de coordinación y de concertación entre la Federación, las entidades federativas, el Distrito Federal y los municipios para la conservación del patrimonio documental de la Nación, así como para fomentar el resguardo, difusión y acceso de archivos privados de relevancia histórica, social, técnica, científica o cultural.”*

Un apartado importante de abordar es el referente a los efectos y ámbitos de aplicación de la ley en lo que refiere a los archivos y su preservación, ya que habla de los la creación de una área coordinadora de archivos así como de los sujetos obligados y de la composición de un archivo de documentos. En este sentido el área coordinadora de archivos en conjunto con el comité de información de cada sujeto obligado deben junto con el área de tecnologías definir además de los diferentes procesos necesarios de gestión documental de archivo, la preservación de archivos electrónicos o digitales. Hay que mencionar que los sujetos obligados en este caso es todo el gobierno federal, así como toda institución con presupuesto federal, secretarías centralizadas y descentralizadas, congreso y órganos autónomos. Cada uno de los sujetos obligados bajo ciertas características sin importar el formato, ni el medio por las que producen documentos las entidades gubernamentales son consideradas documentos de archivo como se define en el Artículo 4 de su apartado segundo referente al archivo:

*Artículo 4. Para efectos de la presente Ley y su ámbito de aplicación se entenderá por:*

...

*II. Archivo: Conjunto orgánico de documentos en cualquier soporte, que son producidos o recibidos por los sujetos obligados o los particulares en el ejercicio de sus atribuciones o en el desarrollo de sus actividades;*

...

*VIII. Área coordinadora de archivos: La creada para desarrollar criterios en materia de organización, administración y conservación de archivos; elaborar en coordinación con las unidades administrativas los instrumentos de control archivístico; coordinar los procedimientos de valoración y destino final de la documentación; establecer un programa de capacitación y asesoría archivísticos; coadyuvar con el Comité de Información en materia de archivos, y coordinar con el área de tecnologías de la información la formalización informática de las actividades arriba señaladas para la creación, manejo, uso, preservación y gestión de archivos electrónicos, así como la automatización de los archivos;*

...

En el Artículo 5 se definen de los principios a los cuales están regidos los sujetos obligados de la Ley de Archivos integrados por la conservación, procedencia, integridad y disponibilidad de los archivos:

*Artículo 5. Los sujetos obligados que refiere esta Ley se regirán por los siguientes principios:*

*I. Conservación: Adoptar las medidas de índole técnica, administrativa, ambiental y tecnológica para la adecuada preservación de los archivos;*

*II. Procedencia: Conservar el orden original de cada fondo documental producido por los sujetos obligados en el desarrollo de su actividad institucional, para distinguirlo de otros fondos semejantes;*

*III. Integridad: Garantizar que los documentos de archivo sean completos y veraces para reflejar con exactitud la información contenida; y*

*IV. Disponibilidad: Adoptar medidas pertinentes para la localización de los documentos de archivo.*

En esta línea el Artículo 20 refiere a la obligación de instrumentación de sistemas automatizados para la gestión documental así como que para la preservación de documentos electrónicos. Los sujetos obligados deberán contar con un sistema de preservación a largo plazo:

*Artículo 20.*

...

*En la preservación de archivos electrónicos en el largo plazo, sea por necesidades del sujeto obligado o por el valor secundario de los documentos, se deberá contar con la funcionalidad de un sistema de preservación en el largo plazo, el cual deberá cumplir las especificaciones que para ello se emitan.*

...

En el Artículo 21 se define la obligación del Archivo General de la Nación en coordinación del instituto y la secretaría de función pública de coordinar los lineamientos de creación y uso de sistemas automatizados para la gestión y control de documentos físicos y electrónicos que deben considerar las actividades y estrategias de preservación digital.

### ***Ley General de Archivos***

En la búsqueda de contar con un Sistema Nacional de Archivos en México que integre una misma metodología y plano de obligaciones mínimas para los diferentes órdenes de gobierno y dependencias nacionales a manera de regularizar la gestión archivística, a partir del año 2014 y acorde a datos publicados en el portal del Archivo General de la Nación en México (AGN, 2014) se ha iniciado la gestión para la construcción de la Ley General de Archivos<sup>55</sup>. Esta acción refiere al decreto que reforma y adiciona diversas disposiciones de la Constitución Política de los Estados Unidos Mexicanos en materia de transparencia publicadas en el Diario Oficial de la Federación el 7 de febrero del 2014 que facultan al Congreso de la Unión para expedir la Ley General de Archivos. Esta ley deberá establecer la organización y administración homogénea de los archivos en los órdenes federal, estatal, del Distrito Federal (hoy Ciudad de México), municipal y del Sistema Nacional de Archivos. La Ley General de Archivos garantiza la preservación del patrimonio documental de México y será pieza fundamental en la construcción de un sistema nacional articulado la rendición de cuentas y mecanismos de coordinación entre los tres niveles de gobierno en materia archivística con respeto pleno a la soberanía de las entidades federativas y autonomía de los municipios. Una Ley General de Archivos representa el encontrar a la actividad archivística de México con ámbitos estandarizados contemporáneos que ayuden a la sistematización, gestión y desarrollo de procedimientos en favor del orden nacional orientados a mejorar la articulación del sistema nacional de rendición de cuentas, el acceso a la información y transparencia del acervo documental del país.

### ***Ley de Ciencia y Tecnología***

El 5 de Junio del año 2002 se publica en el Diario Oficial del a Federación la Ley de Ciencia y Tecnología que tiene por objeto de manera general el regular los apoyos del gobierno federal destinados a la investigación científica, desarrollo tecnológico e innovación. Determinar los instrumentos de obligación del gobierno federal ante la ciencia y la tecnología, así como establecer mecanismos e instancias de coordinación de gobierno, sectores educativos y productivos bajo sistemas de apoyo y fortalecimiento de actividades que regulen y fomenten el desarrollo tecnológico y de innovación en México. A partir de ello se crea el Consejo

---

<sup>55</sup> Ley General de Archivos <http://www.agn.gob.mx/LeyArchivos/>

Nacional de Ciencia y Tecnología (CONACYT) de México como organismo responsable de establecer las políticas, lineamientos y acciones en cumplimiento del fortalecimiento de la ciencia y tecnología del país.

Con lo que respecta a la relación a esta ley en la gestión de archivos y documentos digitales se le adicionaron diversas disposiciones a la Ley de Ciencia y Tecnología que fueron publicadas en el Diario oficial de la Federación el 20 de Noviembre del 2014, en las cuales se instruyo al CONACYT para expedir los lineamientos y disposiciones correspondientes para el funcionamiento del Repositorio Nacional en México. El Consejo Nacional de ciencia y Tecnología presentó los lineamientos generales para el repositorio nacional y los repositorios institucionales, y durante el año 2015 se trabajó en los lineamientos secundarios o técnicos que consolidan el documento para su aplicación formal. Tales lineamientos refieren en su artículo 16 de manera específica a la preservación digital, mencionando lo siguiente: “Los Repositorios Institucionales deberán establecer mecanismos para la preservación a perpetuidad de los recursos de información contenidos; asimismo, deberán contar con los mecanismos para cosechar automáticamente la información de otros repositorios (Nacional, Institucionales e Internacionales), debiendo permitir que su información sea recolectada automáticamente por parte de éstos. Los administradores y operadores de los Repositorios Institucionales deberán asegurar la interoperabilidad entre sí y con el Repositorio Nacional. Igualmente, deberán contribuir a la función de proveer el acceso público gratuito a sus contenidos, adecuándose para la interoperabilidad con otros repositorios y sistemas de información de investigación con base en los estándares internacionales, y las directrices para la descripción de los metadatos que el CONACYT establezca a través los Lineamientos Técnicos que emita.” (CONACYT, 2014). Aunque la iniciativa de lineamientos abarca de manera integra una herramienta formal de depósito digital para las instituciones de educación superior, podemos decir que la preservación digital en México continúa una ruta verde de integración y consideración como un rubro de futura consolidación para formalizarse de manera obligada a consecuencia de las prioridades actuales de integración de una red de repositorios institucionales.

El marco legal que hasta aquí se define se resalta la importancia de la integración y normalización de la actividad profesional de documentación y sistemas de gestión de archivo, así como las nuevas obligaciones que apuntan a la automatización de los sistemas

archivísticos para atender su preservación digital incluyendo el mantenimiento y preservación de metadatos. Establecer procedimientos de actualización de los documentos electrónicos mediante la migración, refresco y respaldo para documentar los cambios que puedan influir en el contenido y contexto de los documentos, como asegurar su acceso futuro bajo las necesidades de integridad, autenticidad, disponibilidad y seguridad. Estas normalizaciones federales representan un gran avance en las iniciativas a nivel de gobierno para atender a las actividades de preservación digital de manera formal con una inclinación técnica mayor que a el desarrollo de políticas de actuación administrativa que complementen las herramientas automatizadas para su gestión. No debemos olvidar que tanto las Leyes de Archivo como la Ley de Transparencia y Acceso a la Información Gubernamental tienen entre sus diferentes objetivos mejorar la rendición de cuentas, es decir de transparencia y acceso a la información de las dependencias y poderes de gobierno mediante el orden sistematizado de gestión archivística. Para conseguir ese fin se necesitan homogeneizar los procesos de archivo a un orden contemporáneo que permita construir una arquitectura de información de archivos apropiada para la sociedad democrática que representa un país en constante desarrollo como México. En este sentido las reformas y adiciones a la Ley de Ciencia y Tecnología y la expedición de los lineamientos del repositorio nacional y repositorios institucionales, hasta ahora amplían el círculo de posibilidades de normalización, gestión y preservación de los diversos patrimonios en México. Cabe mencionar que la importancia de crear conciencia ante la preservación de los archivos y documentos de herencia cultural como registro de una sociedad, será cada día más valorada y necesaria tal como lo apuntó Waters (1997) en su frase: “Cultura, cualquier cultura depende de la calidad en la que registra su conocimiento”.



**Capítulo 3 - Análisis y estudio de sistemas, modelos y  
necesidades en preservación digital distribuida**



### 3.1 Definición de análisis y estudio de sistemas, modelos y necesidades PDD

El siguiente capítulo tiene como objetivo analizar a la preservación digital distribuida desde tres enfoques prioritarios. El primer enfoque se centra en contrastar la información de identificación y funciones básicas de los principales sistemas de preservación digital distribuida para definir los elementos conceptuales que componen su arquitectura como sistema. El segundo enfoque analiza y define las buenas prácticas de organización, coordinación, colaboración y gobernanza en casos de integrantes de redes de preservación distribuida. El tercer enfoque corresponde a analizar el estado y nivel de prácticas de preservación digital en instituciones de educación superior de México (IES-MX) para definir un esquema de necesidades y requerimientos de estas ante una estrategia formal de preservación digital en modo distribuido. La finalidad de realizar el análisis desde tres enfoques refiere a la necesidad de cubrir en una panorámica desde la perspectiva funcional, de organización y niveles de estado de prácticas de preservación digital para apoyar la definición de un modelo con elementos funcionales, conceptuales y de buenas prácticas que sirvan de argumentos clave para integrar una propuesta de estrategia de coordinación y gestión interinstitucional en un modelo de preservación digital distribuida.

Cabe reiterar que este análisis se pretende definir a nivel conceptual, funcional y de buenas prácticas refiriendo por ello a abstracciones de las figuras clave de conceptos y funciones técnicas generales de los sistemas de preservación digital distribuida. El alcance de análisis técnico se limitará a un nivel elemental (no a nivel técnico avanzado de programación y códigos) ya que el interés primordial es que los resultados sirvan de modo pragmático y conceptual para entender los niveles funcionales y operativos de los sistemas de preservación digital distribuida en sus posibilidades y requerimientos de integración complementados con las prácticas de coordinación y necesidades de buenas prácticas.

Tabla 7. Marcos de análisis PDD.

Marcos		
1) Información y funciones de sistemas PDD	2) Coordinación, organización y gobernanza interinstitucional en redes de PDD	3) Estado y nivel de preservación digital en IES-MX

### 3.2 Análisis y estudio de información funcional de sistemas de PDD

El objetivo de este estudio es el definir un marco práctico de contraste e información sobre las funciones, elementos y principales prestaciones de los sistemas de preservación digital distribuida. Para la definición de este estudio de sistemas PDD se elaboró una estructura de indicadores con el objetivo de organizar las funciones de los sistemas de preservación digital distribuida a nivel práctico e informativo, para describir su perspectiva de uso y prestaciones específicas a modo de contraste de elementos y funciones indispensables en las estrategias distribuidas de preservación desde los puntos de vista de los sistemas a estudiar. Para el análisis se emplearon técnicas de tipo exploratorio descriptivo sobre diferentes recursos disponibles para la obtención de información general de los sistemas tales como sitios web oficiales de los proveedores de los sistemas PDD, así como de recursos de información y literatura disponible especializada en el uso y configuración de los mismos.

Para el estudio y análisis se utilizaron algunos indicadores y terminologías especializadas en la comparación de sistemas de preservación digital distribuida propuestos por Educopia Institute en *Comparative Analysis of Distributed Digital Preservation (DDP) Systems* (Schultz y Skinner, 2014). Así mismo se definieron indicadores complementarios basados en las terminologías mayores de elementos de preservación digital distribuida propuestos por Schultz y Zierau (2013) como el de replicación e independencia. De igual manera se emplearon algunos indicadores del sistema de evaluación propuesto en *From Theory to Action: "Good Enough" Digital Preservation Solutions for Under-Resourced Cultural Heritage Institutions* (Schumacher, Thomas y VandeCreek, 2014).

Este análisis incluye de manera indistinta como muestra a tres sistemas de preservación digital distribuida, LOCKSS (Lots of copies keep stuff safe), el sistema BRP (BitRepository Project) y IRODS (Integrated Rule-Oriented Data System). Estos son los sistemas que cumplen con los requerimientos de identificación de software o sistema informático con funciones de preservación digital distribuida. Los indicadores a emplear en este análisis refieren principalmente a dos rubros. El primero es el de la información general del sistema, es decir sobre temas como identificación del sistema y usos del mismo. En una segunda instancia se definen las funciones que distinguen a un sistema de preservación digital

distribuida que se integran principalmente por las funciones generales en modelos de referencia de preservación digital y los modelos empleados como estrategia de replicación distribuida. A continuación se muestran los indicadores por grupo.

#### Información General

1. *Identificación del sistema*
2. *Usos del sistema*

#### Funciones Básicas

1. *Ingesta y contenido*
2. *Almacenamiento y réplicas*
3. *Mantenimiento y monitoreo*
4. *Recuperación y acceso*

### **3.2.1 Listado y análisis de información de sistemas PDD**

A nivel de posicionamiento, identificación y disponibilidad de información general sobre sistemas de preservación digital distribuida, podemos decir que LOCKSS es el sistema más conocido y sobre el cual se encuentra más información de primera mano como sistema especializado. A nivel de identificación por búsqueda de nombre y dominio web, los tres sistemas cuentan con direcciones web bajo nombres de dominio concretos, sólidos e identificables. La información identificada en el sitio web del sistema LOCKSS y de IRODS se muestra bajo una arquitectura de información institucional, comercial y técnica. La información del sitio web del sistema BRP está orientada casi en su totalidad a información técnica ya que su dirección web solo funciona como enlace con dirección al sitio web de mantenimiento del sistema y código abierto en GitHub.

Los sistemas están definidos en su información con objetivos de sistemas de preservación digital distribuida en el caso de LOCKSS y BRP. El sistema IRODS no especifica de manera literal su función para preservación digital distribuida, aunque si lo expresa hacia la colaboración distribuida de flujos de trabajo de datos y contenidos digitales.

Tabla 8. Identificación de sistema PDD.

<b>Informacion General</b>	<b>LOCKSS</b>	<b>BRP</b>	<b>IRODS</b>
<b>Identificación del sistema</b>			
<i>¿Quien desarrolló el sistema?</i>	Stanford University Libraries	National Archives, the Royal Library, The State and University Library of Denmark	San Diego Supercomputer Center y mantenido por un consorcio por RENCI
<i>¿En que año se creo el sistema?</i>	2004	2010	2006
<i>¿Cual es la dirección web oficial con información del sistema?</i>	<a href="http://www.lockss.org">http://www.lockss.org</a>	<a href="http://www.bitrepository.org">http://www.bitrepository.org</a>	<a href="http://irods.org">http://irods.org</a>
<i>¿Como se define el sistema en su web oficial?</i>	The LOCKSS Program is an open-source, library-led digital preservation system built on the principle that “lots of copies keep stuff safe.”	The purpose of the Bitrepository system is to enable longterm preservation of data in a distributed, highly redundant architecture. The data integrity is ensured by using multiple, independently developed data storage systems across different organizations, together with functionality for maintaining the Integrity of the data over time.	The Integrated Rule-Oriented Data System (iRODS) is open source data management software used by research organizations and government agencies worldwide. iRODS is released as a production-level distribution aimed at deployment in mission critical environments.
<i>¿Cual es el principal objetivo del sistema?</i>	Preservación digital en modo distribuido	Preservación digital en modo distribuido	Control de datos en distribuidos y colaborativos
<i>¿El sitio web oficial del sistema cuenta con un apartado que informe sobre la descarga directa del software?</i>	No directamente, pero en el apartado de soporte técnico hay un enlace a sobre la licencia open source, la cual refiere a el sitio donde se mantiene el programa del sistema el cual es: <a href="http://sourceforge.net/projects/lockss/">http://sourceforge.net/projects/lockss/</a>	No directamente; el código se encuentra en GitHub: <a href="https://github.com/bitrepository">https://github.com/bitrepository</a>	No directamente: el código se encuentra en GitHub: <a href="https://github.com/irods/irods">https://github.com/irods/irods</a>
<i>¿El el sitio web del sistema ofrece información sobre soporte, configuración y uso del mismo?</i>	Si, el sitio web ofrece información detallada del sistema, su configuración y uso, así como de actualizaciones y nuevas integraciones.	Si, el sitio web ofrece información detallada del sistema, su configuración y uso, así como de actualizaciones y nuevas integraciones.	Si, el sitio ofrece información de soporte, entrenamiento y certificación.
<i>¿Existe información de terceros (artículos, ponencias, manuales, etc.) sobre el sistema?</i>	Si existe mucha información principalmente de artículos, reseñas y guías.	Si existe informacion, aunque poca literatura de terceros.	Si, existe mucha información de artículos, presentaciones, sumarios y descripciones.

En los tres casos de sistemas la información técnica de acceso a la descarga o adquisición del sistema se enlaza a sitios externos de mantenimiento. LOCKSS y BRP requieren de una búsqueda en diferentes enlaces para dar con los sitios que alojan los códigos del sistema alojados en SourceForge (LOCKSS), en GitHub (BRP y IRODS) respectivamente. Esta situación de antemano sugiere la necesidad de conocimientos de programación y código para participar en el mantenimiento del sistema o en su caso para instalar los diferentes modelos del mismo. Los sistemas en sus sitios web ofrecen información sobre las instalaciones así como de los procesos de su configuración a nivel general.

Llama la atención el caso del sistema LOCKSS, pues en su sitio web se ofrecen diferentes modalidades de servicio para formar parte de la comunidad del sistema con soporte directo de Stanford University Libraries y que se denomina LOCKSS Alliance. Esta se divide en dos grandes grupos que son Global LOCKSS Network que está orientada hacia editores (sin costo pero con sitios limitados y bajo estudio del prospecto) y a bibliotecas e instituciones (costo de soporte) con cuotas de participación según el tamaño de la institución. Otra modalidad es la de Private LOCKSS Network para redes de grupos específicos institucionales a los cuales también se les sugiere contactar directamente con LOCKSS para más detalles de su participación y soporte. El sistema IRODS mantiene la opción de integrarse al IRODS consortium que da acceso a diferentes finalidades de mantenimiento, desarrollo, soporte y entrenamiento en el uso del sistema. En lo que refiere a BRP es notable que ha sido un sistema desarrollado a medida para un proyecto específico, lo que lo mantiene como un sistema en desarrollo e implementación con usos privados en código abierto pero disponible para colaboraciones futuras tanto de desarrolladores como de participantes. La literatura e investigación alrededor de los sistemas está escrita principalmente por terceros o usuarios del sistema. Es notable una gran diferencia entre la información disponible de los sistemas, ya que 3 de cada 5 artículos, reseñas y o guías están especializados en el sistema LOCKSS. En segundo lugar el sistema IRODS con guías, descripciones, artículos, etc. En el caso de The Bit Repository Project encontramos en principio literatura desarrollada por miembros del National Danish Bit Repository, una tesis y algunas reseñas. Cabe señalar que toda la información de los sistemas está escrita en su totalidad en inglés y algunos apartados en danés (BRP).

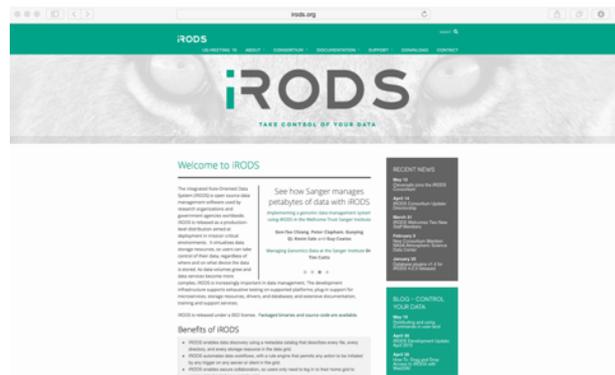
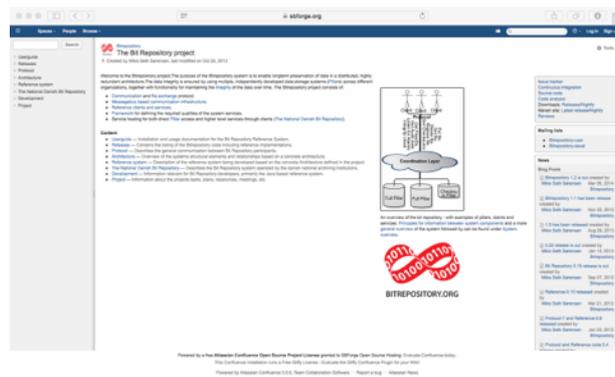
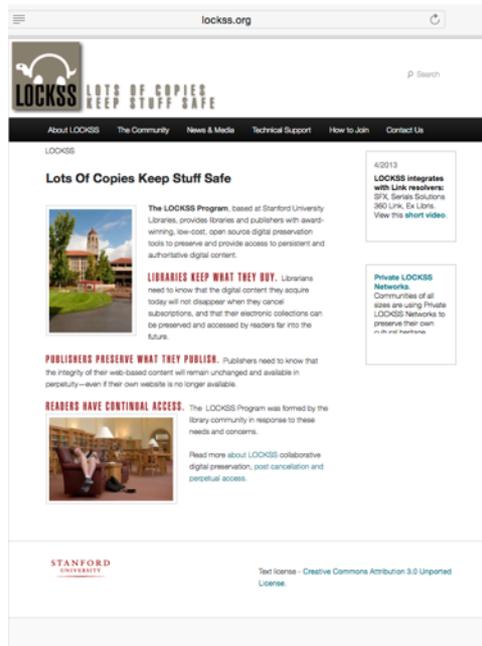


Figura 7. Home Page de sistema LOCKSS, The BitRepository Project y IRODS

Tabla 9. Uso de sistemas PDD.

<b>Informacion General</b>	<b>LOCKSS</b>	<b>BRP</b>	<b>IRODS</b>
<b><i>Usos del sistema</i></b>			
<i>¿Que tipo de licencia de software utiliza el sistema?</i>	Open Source / Código abierto	Open Source / Código abierto	Open Source / Código abierto
<i>¿Que tipo de instituciones u organizaciones utilizan el sistema?</i>	Universidades, Bibliotecas, Centros de Investigación, Archivos de gobierno.	Universidades, Bibliotecas, Centros de Investigación, Archivos de gobierno.	Organizaciones de investigación y agencias gubernamentales
<i>¿Existen redes privadas que ofrecen el servicio externo de preservación digital distribuida bajo el uso del sistema? ¿Cuales son?</i>	Si, como servicio: MetaArchive Cooperative, SAFE Archiving Preservation, Digital Commons, Data-PASS, CLOCKSS y ADPN. como redes privadas: PNAS, PKP, LuKII, CARINIANA IBIC, DFDLP, Synergies, COPPUL.	No hay servicios externos que usen el sistema; actualmente el sistema solo es utilizado por The National Danish BitRepository por las instituciones: National Archives, the Royal Library and the State and University Library of Denmark	Si, como servicio: Chronopolis. Como redes privadas y usuarios del sistema Taiwan National Archive, NASA Langley Atmospheric Sciences Center, Carolina Digital Repository, French National Library, Texas Digital Libraries, National Optical Astronomy Observatory, etc.
<i>¿Que grado de configuración e instalación se requiere para usar el software del sistema?</i>	Configuración e instalación avanzada	Configuración e instalación avanzada	Instalación empaquetada y configuración avanzada
<i>¿Que lenguaje de programación y sistema operativo utiliza el sistema?</i>	CentOS 6 / Linux / Java	Java / Linux	Python / Linux
<i>¿Una vez instalado el software del sistema PDD en un servidor, cuales son los principales usos que otorga?</i>	Se crea un nodo (cache, LOCKSS Box) para integrarse a una red de preservación digital distribuida, el cual tendrá la posibilidad de ingesta, cosecha, acceso, monitoreo y recuperación de contenido digital.	Se crea un enlace en una capa de coordinación entre servidores, la cual permite ingesta, acceso, monitoreo y recuperación de contenido digital.	Se crea una cuadrícula o capa de datos (data grind) la cual permite un flujo de trabajo colaborativo de catalogación, visualización, almacenaje distribuido de datos y contenido digital.
<i>¿El sistema de PDD puede usarse con un sistema de repositorio o archivo digital existente?</i>	Si vía plugins	Si vía plugins	Si vía plugins

Sobre el uso de los sistemas de preservación digital distribuida podemos distinguir algunos detalles que es importante mencionar. En primer lugar el tema de sistemas de código abierto como sistemas de preservación digital, representan de manera intrínseca la apertura a la colaboración de su código para su mejora y optimización. En esta línea todos los sistemas cumplen y llevan a cabo bajo diferentes protocolos como wikis, foros, alianzas, comunidades, etc. De manera puntual hay diferenciar nuevamente del término abierto y libre, los cuales conllevan diferentes niveles de actuación y por tanto del uso abierto que corresponde a la

licencia de apertura para su estudio, modificación o mejora del código abierto. En el caso del término libre se identifica más con un sentido filosófico de libertad de uso. Los sistemas de PDD estudiados se describen como empleados en diferentes organizaciones e instituciones a nivel mundial haciendo uso de la replicación geográfica y definiendo un nivel de colaboración coordinado que integra responsabilidades y obligaciones sobre el uso del sistema.

Llama la atención que el uso del sistema LOCKSS está altamente implementado en diferentes tipos de redes de preservación digital distribuida de universidades, organizaciones e instituciones gubernamentales donde algunas de las cuales (principalmente organizaciones privadas) ofrecen de servicios de preservación bajo membresía y administradas por representaciones de organizaciones sin ánimos de lucro en pro de la preservación del patrimonio cultural, de investigación, etc. Como se mencionó antes este modelo de trabajo colaborativo en diferentes niveles de gestión (membresía, grupo, etc.) refiere a colaborar económicamente para mantener productiva a una red de preservación en la prevención y formalidad la seguridad de los contenidos preservados, replicados y con redundancia geográfica. El objetivo de privatizar una red es para mantenerla libre de ataques y bajo un control de usuarios de los sistemas que otorguen un uso prioritario, auténtico e íntegro del mismo. El sistema IRODS también cuenta con una alta implementación que a diferencia de LOCKSS y BRP se destaca en el ámbito de instituciones de investigación y gestores de datos especializados en el área de astrofísica, ciencias atmosféricas, clima u oceanografía, y de manera menor en el ámbito de repositorios universitarios o archivos. Actualmente los principales usuarios de los sistemas de preservación distribuida se concentran en instituciones de educación, gobierno, bibliotecas, centros de investigación y editoriales. Los segmentos empresariales y particulares podrían ser el siguiente eslabón.

Los sistemas LOCKSS y BRP no cuentan con paquetes de instalación *easy to use*, solo IRODS cuenta con instalación empaquetada. LOCKSS y BRP requieren de configuraciones avanzadas con necesidad de conocimientos en programación de código para su mantenimiento y de sistemas operativos Linux, así como de lenguaje Java o Python para su implementación. Los sistemas cuentan para su uso e instalación con guías en sus sitios web los cuales en su mayoría resuelven su implementación general. En el caso de LOCKSS también se puede lograr por medio del apoyo de los desarrolladores del sistema bajo previo contacto, diagnóstico y consulta. En el caso de BRP se pueden realizar consultas vía correo

electrónico directamente con los desarrolladores del sistema. El caso de IRODS bajo contacto previo y en todos los momentos se debe otorgar créditos de uso que cada sistema provee en sus avisos legales.

A nivel de usos generales del los sistemas otorgan usos diferenciales entre sí. El sistema LOCKSS convierte a un ordenador en un sistema de almacenamiento y servidor *cache* que permite conectarse a una red de otros ordenadores que corren el mismo sistema con las prestaciones de preservación (ingesta, cosecha, mantenimiento, monitoreo, etc.). En el caso de los sistemas BRP y IRODS estos no convierten a ordenadores individuales en sistemas de almacenamiento para preservación, más bien crean una capa de coordinación o punto de unión entre varios servidores (sistemas de almacenamiento individuales) para gestionar diferentes actividades como poner, obtener, reemplazar, borrar archivos, obtener identificadores, auditar colecciones, etc. Todos los sistemas mantienen un uso diferencial pero general como sistemas de preservación digital distribuida y se pueden resumir en sistemas preservación digital bajo independencia, coordinación y conexión.

### **3.2.2 Listado y análisis de funciones generales en sistemas de PDD**

Para definir cuales son las funciones generales de un sistema de preservación digital distribuida, en primer lugar nos enfocamos en las funciones que otorga un sistema de preservación digital centralizado como lo son la ingesta de contenidos, mantenimiento, almacenamiento, recuperación, etc. A partir de ahí complementamos con las funciones que distinguen a la estrategia distribuida como lo son las funciones de replicación geográfica, control de versiones redundantes, independencia coordinada, etc. Los indicadores de funciones generales se definieron mediante un listado sobre las principales funciones de un sistema de preservación digital distribuido entre los tres sistemas de preservación estudiados y con la finalidad de obtener contrastes de usos en los elementos compartidos.

Tabla 10. Ingesta de contenido.

<b>Funciones Generales</b>	<b>LOCKSS</b>	<b>BRP</b>	<b>IRODS</b>
<i>Ingesta de contenido</i>			
<i>Métodos de ingesta de contenido del sistema</i>	Web UI, Servidores HTTS, Dispositivos de almacenamiento directo.	Web UI, Servidores HTTS, Dispositivos de almacenamiento directo.	Web UI, Servidores HTTS, Dispositivos de almacenamiento directo.
<i>Formatos digitales de contenido aceptados</i>	Multiformatos	Multiformatos	Multiformatos
<i>Dosis de contenido aceptado por ingesta</i>	Archivo individual, paquetes de datos y colecciones	Archivo individual, paquetes de datos y colecciones	Archivo individual, paquetes de datos y colecciones
<i>Métodos de ingesta de metadatos</i>	Manual y automatico	Manual y automatico	Manual y automatico
<i>Tipo de metadatos generados en ingesta</i>	Identificadores, estructurales, descriptivos y administrativos	Identificadores, estructurales, descriptivos y administrativos	Identificadores, estructurales, descriptivos y administrativos
<i>Cosecha (harvesting) de contenidos y metadatos</i>	Si	Si	Si
<i>Creación de checksum de contenido en ingesta</i>	Si	Si	Si

Sobre la ingesta del contenido los sistemas PDD estudiados son agnósticos en el ingreso de formatos digitales, es decir aceptan todo tipo de formatos en su ingreso por la interoperabilidad de los mismos en diferentes sistemas operativos. Debemos mencionar que en la actualidad la demanda de uso de los sistemas de preservación digital distribuida en su mayoría atañe a la línea de multiformatos a diferencia de lo que se ofrecía en sus inicios sobre formatos html, bases de datos o formatos exclusivo de documentos, extendiéndose a formatos multimedia y especializados. En el caso de los métodos de ingesta en los sistemas existe la flexibilidad de ingreso ya sea mediante una interfaz de usuario web o directamente desde dispositivos de entrada en servidores u ordenadores; estos métodos ofrecen una ventaja para el rol de ingesta de contenido al sistema PDD. En el caso del sistema BRP ofrece una interfaz de usuario WebGUI a modo de dashboard que ofrece datos, gráficos y estadísticas de las colecciones con información sobre alarmas, auditoría, integridad y estatus de la ingesta. Por su parte la interfaz de ingesta del sistema LOCKSS además de la función de ingesta o agregado de contenidos, ofrece funciones de configuración manual unidades de archivo (AU), control de acceso a contenido, usuarios y conexiones de la caja. En lo que se refiere al soporte

de ingesta de grupos de datos o paquetes, el sistema LOCKSS soporta diferentes tecnologías de como BagIt<sup>56</sup>, BagSplit Utility para interoperabilidad, directorios de datos sin comprimir, y archivos comprimidos en ZIP y TAR con solicitud de checksum de archivos individuales (Schultz y Skinner, 2014). En el caso de IRODS su ingesta se lleva a cabo también mediante una interfaz gráfica web que conecta con la capa media que ofrece el sistema donde se puede realizar toda las funciones necesarias de ingesta, réplica, etc. Todo sistema de preservación enfrenta en su proceso de ingesta el primer problema de la preservación digital, la integridad. Los sistemas necesitan de un método de control con el cual todo ingreso quede registrado bajo la suma de comprobación (checksum) de secuencia de datos digitales. Todos los sistemas cuentan con la función de registro de checksums de los contenidos ingresados al sistema con algunas diferencias de aplicación y almacenamiento de los mismos para verificar cualquier discrepancia.

Los sistemas PDD coinciden en las funciones de ingesta como en el uso de identificadores de ingreso, así como para el ingreso de nuevas versiones de contenidos en el sistema de manera manual. A nivel de registro de metadatos y tratamiento de los contenidos, los sistemas comparten prestaciones similares en la creación de metadatos de modo automático y manual tanto para la descripción de contextos con metadatos de preservación, descriptivos, representación y estructurales en los contenidos ingresados con la flexibilidad de integrar diferentes esquemas de estándares de metadatos de preservación como PREMIS, Dublin Core, entre otros. IRODS se enfoca en los datos descriptivos para recuperar y descubrir datos, los cuales están almacenados en lo que se denomina IRODS metadata catalog (ICAT), la cual es una base de datos relacional que funciona como almacén de información de los datos, usuarios y zonas que facilita el manejo de compartir con los servidores de IRODS los datos en una red distribuida a modo colaborativo de datos. Los sistemas PDD cuentan con la función de cosecha de contenidos y metadatos ingresados al sistema siempre y cuando sean parte de una red o protocolo de coordinación autorizado, así como de que se cuente con un manifiesto de permiso de cosecha de contenidos o quede estipulado en un acuerdo de nivel de servicio de las unidades de almacenamiento pertenecientes a una red.

---

<sup>56</sup> <http://en.wikipedia.org/wiki/BagIt>

Tabla 11. Almacenamiento y réplicas.

<b>Funciones Generales</b>	<b>LOCKSS</b>	<b>BRP</b>	<b>IRODS</b>
<i>Almacenamiento y réplicas</i>			
<i>Denominación de unidad de almacenamiento en el sistema</i>	LOCKSS Box (LB)	Pillar (Full pillar, copy pillar and CheckSum Pillar)	Storage device / IRODS Zone
<i>Tipo de tecnología y medios de almacenamiento aceptados</i>	Diferentes tipos de tecnología y medios de almacenamiento.	Diferentes tipos de tecnología y medios de almacenamiento.	Diferentes tipos de tecnología y medios de almacenamiento.
<i>Creación automática de archivo de preservación</i>	Si	Si	Si
<i>Formato de archivo de preservación almacenado</i>	Lógicos y cadena de bits	Lógicos y cadena de bits	Lógicos y cadena de bits
<i>Replicación de contenido en red geográfica dispersa</i>	Si	Si	Si
<i>Independencia de unidades del almacenamiento replicadas</i>	Si	Si	Si
<i>Coordinación de unidades de almacenamiento replicadas</i>	Si	Si	Si

La gestión de almacenamiento de los sistemas de preservación digital distribuida sobresale como uno de los elementos más importantes en la ejecución de los sistemas y estrategias de preservación digital a nivel distribuido debido a factores como los del tipo de almacenamiento o tecnología a emplear, las dependencias de necesidades de la red a la que se pretenda integrarse, así como de las capacidades ofrecidas por los miembros de la misma. Estas situaciones de almacenamiento podrían ser supeditadas según previos acuerdos de la red pues algunos nodos ocuparan mayor capacidad que otros y los convenios de niveles de servicio y tamaño de los participantes de la red distribuida será crucial en esta definición.

Los sistemas contrastados de PDD dentro de sus protocolos de implementación y arquitectura definen a sus unidades de almacenamiento con diferentes terminologías e incluso funciones. En el caso del sistema LOCKSS las unidades de almacenamiento utilizan el nombre de LOCKSS Box (LB) que de manera integral cumplen diferentes funciones como las de servidor, almacén y cache. El hardware que compone a un LB es un ordenador con capacidad de servidor tal como se sugiere en el manual de instalación de LOCKSS<sup>57</sup> donde para

<sup>57</sup> Manual de instalación LOCKSS en Linux. <http://www.lockss.org/docs/LOCKSS-Linux6-Install.pdf>

implementar una LOCKSS Box se recomienda un hardware con arquitectura no menor de 64-bit y un procesador x86 Intel compatible con un CPU dual-core o de preferencia un CPU quad-core con 8GB de memoria como mínimo y lector CD o DVD. La capacidad y tipo de disco duro dependerá de las necesidades de la red a la que se unirá cada LB por lo que de manera básica sugieren que una típica caja utiliza 4 discos duros, aunque algunos servidores aceptan de 8 a 12 discos duros por igual la capacidad de cada uno dependerá de necesidades de la red a unirse. De manera básica el sistema PDD en su instalación utiliza para su operación entre 16 y 20 GB para funcionar como punto de partida, y de manera independiente deberán tomarse en cuenta las capacidades extras necesarias para almacenar contenido propio y de los demás integrantes de una red. Preservación digital no solo es almacenamiento en backup pues solo es una parte importante de todas las partes que incluyen la solución y estrategia.

Es importante mencionar que el sistema LOCKSS denomina a las colecciones independiente de contenido hospedadas dentro de un LOCKSS Box o caché como Archival Unit (AU). El caso de BRP las unidades de almacenamiento se denominan *pillars* o *pilares*, y tal como los definen en su vocabulario de BRP<sup>58</sup>. Los *pilares* de manera independiente pueden cumplir diferentes funciones aparte de ser un medio e infraestructura específica (disco magnético, óptico, cinta, etc.), tienen características de representar una copia de datos que puede ser empleado acorde a un previo acuerdo de nivel de servicio (SLA) del mismo (pilar) para ser visto y analizado de manera individual como una unidad a nivel abstracto y que puede ser representada como un pilar de copia completa, pilar de checksum (suma de comprobación) o pilar Bitmagasin que corresponde a una entidad o institución específica. En el caso de IRODS este funciona bajo una capa que define a las unidades de almacenamiento como storage device y no especializando con un nombre a los dispositivos pero si a donde se encuentran. IRODS define a la IRODS Zone como la zona donde junto los servidores y sistemas de almacenamiento confluyen en el sistema. Cabe señalar que al igual que BRP, IRODS permite la definición de una capa unificada de sistemas de almacenamiento como una sola para su usabilidad.

Para la valoración de los tipos de tecnología y medios de almacenamiento que los sistemas de preservación distribuida pueden utilizar debemos enfatizar que los tres cuentan con la

---

<sup>58</sup> Pilar BRP <https://sbforge.org/display/BITMAG/Pillar>

flexibilidad de integrar medios y tecnología de almacenamiento tanto físicos en red o virtuales. Esta situación anticipa una ventaja para las posibilidades de escalar o actualizar los medios de acuerdo al crecimiento de espacio digital. Una característica importante en el ámbito del almacenamiento para preservación digital a largo plazo y que es que en la principal referencia para implementarla radica en las políticas de almacenamiento. El caso de la implementación del sistema BRP en The National Danish Bit Repository de manera particular define en su política de almacenamiento que cada BRP Pillar de almacenamiento debe utilizar diferente tecnología que las demás pilares, es decir un pilar usa unidades ópticas, otro cintas, otro disco magnético, etc. Esta implementación además de optar por la diversidad en medios de almacenamiento también sugiere el uso de sistemas operativos diferentes para alojar a cada BRP Pillar su sistema de almacenamiento, todo con la finalidad de brindar mayor independencia (de tecnología y seguridad) a los sistemas de almacenaje digital. En lo que se refiere a una de las características especiales de los sistemas de preservación digital distribuida como el almacenamiento distribuido con redundancia geográfica o réplicas dispersas, en específico las unidades de almacenamiento LOCKS Box o Pillar de ambos sistemas cuentan con las funciones de obtener, copiar o replicar contenidos de manera coordinada con otros nodos caches dispersos geográficamente mediante la conexión de internet y de manera independiente al resto de los nodos. El sistema IRODS por su parte dentro de su zona IRODS contiene una serie de protocolos que deben ser autorizados por su sistema de metadatos ICAT y solicitados en su servidor IRODS para después llevar a cabo la distribución dispersa bajo la unificación de los recursos dentro de la zona.

Tabla 12. Monitoreo y mantenimiento.

<b>Funciones Generales</b>	<b>LOCKSS</b>	<b>BRP</b>	<b>IRODS</b>
<b><i>Monitoreo y mantenimiento</i></b>			
<i>Monitoreo de integridad de réplicas</i>	Si	Si	Si
<i>Migración de contenido de preservación</i>	Si	Si	Si
<i>Reparación de contenido replicado</i>	Si	Si	?
<i>Programación automática y manual de monitoreo y mantenimiento</i>	Si	Si	Si

El monitoreo y mantenimiento en los sistemas de PDD son funciones clave para llevar a cabo la comprobación y monitoreo de los contenidos digitales preservados a largo plazo para mantener la integridad y originalidad de los archivos preservados. En los tres sistemas bajo algunas diferencias de tecnología (microservicios) y terminología se llevan a cabo actividades de monitoreo remoto sobre protocolos de redes privadas en internet. El caso del sistema LOCKSS integra una tecnología de monitoreo que funciona como un sistema de comparación de contenidos preservados en las distintas LOCKSS Box de una red para verificar que sea la versión autorizada el mismo contenido que está almacenado en todos los nodos. Si existiera alguna diferencia con la versión autorizada en alguna de las LB, mediante un sistema de votación de versiones de contenidos y haciendo uso de comparaciones de checksums el sistema puede decidir organizar la reparación automáticamente desde otro nodo siempre y cuando se decida que algún contenido está dañado, corrupto o perdido. El estado y monitoreo del contenido preservado puede ser consultado desde la interfaz web de administrador del sistema. En el caso del sistema BRP se realiza un protocolo similar el cual denomina servicio de integridad que lista archivos, recupera checksums y repara archivos dañados. IRODS en este sentido lleva tiene la capacidad de llevar a cabo actividades que denomina en su sistema como *digital data management*, los cuales integran actividades de monitoreo y mantenimiento con verificación de integridad bajo la revisión de checksums. Es obligatorio que un sistema con una arquitectura altamente distribuida de preservación digital cuente con un servicio de monitoreo automatizado de integridad de contenidos. Otra de las funciones importantes a nombrar es la de la migración digital del contenido preservado en distintos nodos distribuidos pues algunos sistemas de preservación digital centralizados como el de la tecnología Piql<sup>59</sup> ofrecen almacenaje digital con fines de preservación a largo plazo en tecnología de película (film) que promete una duración de 500 años libre de migración digital de los contenidos alojados bajo ciertos parámetros de durabilidad del soporte, lectura, encapsulado de los datos y metadatos preservados (representación, estructurales y descriptivos) a largo plazo. La libertad de no migrar a nuevos formatos como ventaja pasiva que podría acercarse al tradicional backup y a la ecuación de S+I (store and ignore) pues elimina una de las estrategias básicas de la preservación digital como lo es la migración por demanda pues esta proviene de una acción preventiva de monitoreo.

---

<sup>59</sup><http://www.piql.com>

En el caso de sistemas distribuidos dinámicos la migración de contenido se presenta sobre los ejes de acceso a los contenidos y en especial en la preservación de tipo lógica de formatos de representación. A nivel de preservación de bit la migración es mínimamente usada ya que se trata de preservar tal y como originalmente se recibió el contenido digital. IRODS permite migraciones a cinta u otro sistema de almacenamiento. A nivel de reparación de contenidos los sistemas LOCKSS y BRP comparten la posibilidad de reparar contenidos dañados o perdidos en una red replicada. La diferencia principal radica en que el sistema LOCKSS las cajas son las que llevan a cabo el proceso de monitoreo y cabildeo con otras cajas para reparar el contenido; en el caso de BRP se define la reparación bajo el esquema de la capa de coordinación sobre mensajes de alarma y partir de ahí se ordena la ejecución del servicio bajo los acuerdos y permiso de servicios en la red. En el caso de IRODS no queda muy clara esta prestación que debe ser funcional integrando plugins al sistema.

Tabla 13. Recuperación y acceso.

<b>Funciones Generales</b>	<b>LOCKSS</b>	<b>BRP</b>	<b>IRODS</b>
<b><i>Recuperación y acceso</i></b>			
<i>Recuperación de contenido replicado</i>	Si	Si	Si
<i>Migración de contenido a formato de recuperación</i>	Si	Si	Si
<i>Interfase privada de acceso a contenido preservado</i>	Si	Si	Si
<i>Gestión de acceso a contenido preservado</i>	Si	Si	Si

Otra de las cualidades de los sistemas distribuidos analizados es que disponen de herramientas o servicios que permiten la recuperación de copias de versiones autorizadas en otros nodos cuando uno de estos no está disponible ya que integran una alta disponibilidad redundante acorde al número de nodos que forman parte de la red de preservación distribuida. Esta función de recuperación nace del sistema LOCKSS ya que forma parte de sus principales funciones y objetivos iniciales en el desarrollo del sistema acorde a las necesidades que presentaban con las suscripciones a revistas y publicaciones en línea que en algunos momentos sufrían bajas de servicios o en su caso de problemas técnicos que impedían acceder a la suscripción, por lo que el objetivo del sistema LOCKSS fue el integrar una copia de cada suscripción en una red de suscriptores de la PLN (Private LOCKSS Network) para que al

presentarse un problema como los mencionados se pudiese disponer y recuperar de una de las copias de las suscripciones de archivo, para que independientemente de la accesibilidad de los mismos en los servidores de las editoriales, siempre se pueda acceder a ellas o recuperarlos. Las funciones de recuperación son compartidas por igual en el sistema BRP bajo demanda explícita de recuperación según los acuerdos y permisos de acceso definidos entre los pilares. Así mismo el sistema IRODS corresponde a su capacidad de distribución de copias y backups bajo su capa intermedia que le permite al usuario recuperar copias replicadas en otros sistema de almacenamiento con la previa autorización del servidor ICAT de metadatos que gestiona la solicitud de recuperación.

Para la recuperación y acceso se pueden realizar actividades de migración según las necesidades de obsolescencia del contenido a recuperar como en el caso del sistema LOCKSS que utiliza un método de migración que denomina “*migration on access*” que funciona únicamente cuando se necesita acceder a un contenido en formato lógico de preservación que está obsoleto. El caso de BRP igualmente solo se realiza una migración a formatos de representación cuando es necesario y bajo una operación específica. En el caso de recuperación simple de contenido solicitado por los usuarios el sistema crea un enlace con el contenido. Nuevamente reiteramos que esta operación se realiza en la capa de coordinación bajo las acciones de obtención de archivo en sus protocolos de comunicación y en específico de su “*file exchange protocol*”.

Los sistemas LOCKSS y BRP mantienen sus contenidos de preservación a modo de archivos negros, sin embargo acorde a necesidades de recuperación y bajo la definición y permisos de acceso se puede acceder a ellos solo con la finalidad de reemplazar archivos bajo solicitud. IRODS es un archivo transparente que enfoca sus esfuerzos al acceso, aunque es un sistema extensible que cuenta con una gran cantidad de microservicios que pueden integrarse para gestionar las actividades de acceso más restringido y así poder mantener archivos negros o de preservación, pero no de manera prioritaria. Los sistemas cuentan con diversos sistemas de control y gestión de acceso mediante interfaces web con las que se puede acceder al contenido. En el caso de LOCKSS este gestiona los accesos desde su interfaz en las opciones “*admin content options*” y “*content admin options*” donde se puede gestionar quien accede al contenido preservado, quien lo puede cosechar o copiar, acceder y exportar. El caso de BRP se puede acceder desde la capa de coordinación y seguir bajo el WebGUI que ofrece a modo

de dashboard. El sistema IRODS de manera prioritaria mantiene la función de acceso y usabilidad a los datos bajo validación y unificación visual en su interfaz de usuario lo que mejora el acceso a ellos.

A modo general hemos observado que los sistemas de preservación digital distribuida contrastados mantienen comportamientos diferentes en sus funciones de coordinación distribuida, ya que los sistemas funcionan de manera predominante a modo de dos entornos de red distintos que podemos denominar como de ínter coordinación y coordinación integral. En este sentido podemos observar que el sistema LOCKSS crea una red coordinada entre sus nodos que son convertidos en cajas LOCKSS que ejecutan las funciones de red desde ellas mismas en una capa de sistema independiente. BRP y IRODS funcionan como una red coordinada que es lo único que une o pone en común a los sistemas de almacenamiento o pilares pues ambos mantiene un protocolo de interacción que es gestionado por una capa de coordinación que mantiene todos los servicios de referencia para el mantenimiento con fines de preservación a largo plazo de la arquitectura distribuida de los mismos a modo de nodos o pilares independientes uno del otro pero dependientes de una capa de coordinación. Este acercamiento nos ofrece una variante de los modelos de preservación digital distribuida que otorga potencialidad a un estado de coordinación agregando a nuevas posibilidades de gestión la gobernabilidad de una red con varios participantes bajo mayor flexibilidad y por otro lado de limitando la dependencia para negociar el nivel de participación en la misma.

### **3.3 Estudio y valoración de coordinación, organización y gobernanza en redes de PDD**

El objetivo de este estudio es el analizar a modo general las buenas prácticas referidas a la participación, organización y coordinación de instituciones específicas que forman parte de redes de preservación digital distribuida. Las redes de preservación digital distribuida privadas a estudiar con la valoración voluntaria de usuarios expertos de algunas de las instituciones participantes que se componen por MetaArchive Cooperative, National Danish Bit Repository y la red CARINIANA de Brasil. Como primera fase de este estudio se obtiene un panorama general de antecedentes y necesidades de integración como red de preservación, su organización y los elementos comunes de formación y coordinación. La primera fase de este estudio se integra por tres actividades específicas.

La primer actividad consiste en la definición de antecedentes de organización y creación redes de PDD. En la segunda actividad hace una revisión de literatura, manuales, políticas y acuerdos que refiere a terminología referente a la organización, coordinación y gobernanza de redes de trabajo organizacional e interinstitucional. La tercera actividad consiste en la integración de términos comunes de antecedentes de organización de redes PDD, así como de revisión de literatura especializada de organización y coordinación de redes de trabajo para definir un esquema de indicadores de organización, gobernanza y coordinación institucional que nos ayude a obtener una valoración y opinión sobre las actividades y conceptos por parte de los miembros de instituciones participantes de redes de preservación digital distribuida (MetaArchive Cooperative, National Danish Bit Repository, CARINIANA de Brasil) con ayuda de un instrumento de preguntas cuestionario agrupado por temas de acuerdo a las terminologías, elementos y actividades que reflejan los principales ventajas y desventajas de organización en el mantenimiento de intereses y colaboración entre los diversos participantes como de las relaciones y alianzas en una red distribuida de preservación digital.

### **3.3.1 Antecedentes de redes privadas de preservación digital distribuida específicas**

La organización y coordinación de diferentes instituciones, requiere esfuerzos y trabajo previo de cabildeo, gestión y relaciones. En este contexto la idea de crear una red cooperativa o coordinada de trabajo con fines de preservación digital solo puede convertirse en realidad una vez que es destrozada en elementos organizados y procesables (Belsky, 2012). Para la realización de este análisis será importante conocer de manera previa los antecedentes clave sobre el surgimiento de iniciativas de organización y de coordinación de instituciones participantes de una red privada de preservación digital que nos ayuda a definir un campo más claro sobre indicadores y elementos necesarios de construcción de un protocolo de colaboración y organización de una red de preservación digital distribuida. Definiremos los antecedentes y puntos clave de trabajo de las redes seleccionadas para el estudio basándonos en documentación propia de las redes, sitios web y artículos donde se abordan antecedentes de las mismas.

#### **a) Antecedentes de MetaArchive Cooperative**

De acuerdo con los datos de su sitio web<sup>60</sup>, MetaArchive Cooperative es la primera red de preservación digital distribuida privada fundada en el año 2004 gracias a la iniciativa de colaboración de seis bibliotecas del sureste de los Estados Unidos las cuales son Auburn University, Florida State University, Emory University, the Georgia Tech Library, University of Louisville, and Virginia Tech, las cuales buscaban desarrollar una solución de preservación digital para sus materiales y colecciones especiales. A partir del consenso de las bibliotecas miembros se crea MetaArchive como una iniciativa de red de preservación liderada a modo de propiedad comunitaria y compuesta por bibliotecas, archivos y otras instituciones con interés de preservar su memoria digital.

La iniciativa MetaArchive también corresponde a el trabajo de cooperación con Library of Congress a través de su programa NDIIPP (National Digital Information Infrastructure and Preservation Program) con lo que lograron consolidar una infraestructura propia a modo de repositorio seguro y rentable con el que se ofreciera el cuidado a largo plazo de materiales digitales evitando subcontratación externa y participando activamente en la preservación de

---

<sup>60</sup> <http://www.metaarchive.org/the-cooperative>

contenidos propios de la red privada. Esta proyecto e iniciativa de las seis instituciones se traduce en la creación de un modelo organizacional de trabajo en red con fines de preservación distribuida, el cual más tarde y bajo necesidades de convertir el proyecto en un programa sostenible de preservación digital tendría que ser ofrecido por una institución que no fuese miembro de la red, es por ello que en el año 2006 Educopia Institute se encarga de la gestión del programa sostenible. Actualmente MetaArchive Cooperative es una red internacional que cual atiende a más de 50 instituciones en 13 estados de la unión americana y dos países como lo son España y Brasil.

MetaArchive Cooperative (MAC) a nivel de información es transparente con respecto a su metodología de trabajo y organización que se define principalmente por el uso e implementación del sistema de LOCKSS en sus redes privadas. Cada red privada MAC se integra por siete nodos cache o servidores que replican la ingesta que se realiza en alguno de los servidores en siete veces guardando una copia cada uno para preservación. En la red privada de MAC los archivos y ficheros crean una red segura de acceso cerrado de servidores creados por los miembros participantes de la red MAC, lo que significa que no hay acceso a los archivos ya que su finalidad es la preservación a largo plazo. Los servidores son implementados por cada institución y unidos a una red segura donde cada servidor está almacenado en diferentes locaciones geográficas y por ende es mantenido por diferentes administradores para incrementar la seguridad e independencia. Los servidores continuamente revisan las fuentes de contenido o ingesta para tomar cualquier contenido que ha sido cambiado o agregado. La red tiene la función de almacenar, cambiar e integrar contenido nuevo junto con el contenido original que denominan *versioning* y que refiere a poder recuperar cualquier versión del contenido. El sistema LOCKSS que funciona en la red permite que cada servidor realice la función de monitoreo entre sí con los demás servidores para detectar cualquier archivo corrupto o dañado y bajo la votación de versiones entre los servidores la red repara el archivo dañado colocando una copia junto con el original.

La gobernanza es un tema primordial en MAC pues a partir de integrar a un grupo de miembros la coordinación de estos corre bajo el el tipo de membresía que adopten en el programa, lo cual les permite integrarse a los diferentes comités de la red. Esta cuenta con tres tipos de membresía que en orden de importancia inician con la membresía *sustaining members* (*miembros sostenedores*) que al ser los miembros que más colaboran

económicamente, también son los que más reciben a cambio a nivel de un alto control y liderazgo en la red tanto en las pruebas, desarrollos y mantenimiento de hardware y software, así como en el establecimiento de las normas de conectividad y transmisión de la red. De este tipo de miembros se compone el comité directivo de la red. De manera seguida está la membresía *preservación members* (*miembros de preservación*) en la que sus miembros son participantes y beneficiarios en lugar de líderes de la red, por lo que no están involucrados con las operaciones de MAC como colaboradores. Los miembros de este tipo de membresía tienen la posibilidad de nombrar colectivamente a un representante que tiene la oportunidad de estar presente en las juntas del comité directivo pero sin derecho a voto. Por último está la membresía *collaborative member* (*miembros colaboradores*) que de manera específica se conforma por un grupo de instituciones que actúan como un ente unificado que usa un servidor central, lo que simplifica la sostenibilidad de participación económica de un solo miembro a una fracción de uso por institución. En esta membresía también nombran un representante sin derecho a voto.

En lo que respecta a la gobernanza MAC trabaja con el modelo administrativo de red en el cual una institución sin ánimo de lucro administra a los integrantes de la red, en este caso el administrador es Educopia Institute. De acuerdo con su estatuto de cooperación Cooperative Charter (2015) la estructura de gobernanza de MetaArchive Cooperative consiste en un comité directivo, un comité de contenido, un comité de preservación, comité de alcance y un comité técnico. Todos los miembros de la red pueden nominar a un representante dentro de los comités de contenido, técnico, alcance y preservación. En el caso del comité directivo, este solo es votado y nominado por los miembros sostenedores (*sustaining members*). Educopia Institute y MetaArchive Cooperative realizan una serie de documentos como lo son el estatuto de cooperación (*Cooperative charter*), el cual define los principios de la organización, la misión, niveles de membresía y las responsabilidades de colaboración, así como documentos de acuerdos, de especificaciones técnicas, de auditorías TDR y los documentos sobre plan de gestión de operaciones.

### ***b) Antecedentes de National Danish Bit Repository***

La red distribuida National Danish Bit Repository (NDBR) surge como un proyecto conjunto entre tres instituciones del patrimonio cultural de Dinamarca, el cual se integran por Los Archivos Nacionales (The National Archives), La Biblioteca Real (The Royal Library) y La Biblioteca del Estado y la Universidad (State and University Library). NDBR tiene por objetivo la preservación de bits de sus acervos culturales de manera coordinada y de acuerdo a una arquitectura de redundancia distribuida. El antecedente más claro de organización y colaboración al proyecto de NDBR de acuerdo con Sørensen (2014) de The Danish National Archives, se define bajo la experiencia previa de trabajo interinstitucional y realización actividades de preservación digital *outsourcing* entre los Archivos Nacionales y la Biblioteca Estatal y la Universidad en Aarhus. La Biblioteca del Estado y la Universidad mantenía una copia de los archivos digitales de los Archivos Nacionales para gestionar su preservación digital. Las actividades de envío y resguardo de material se hacían transportando copias en soportes físicos de forma manual, por lo que con el tiempo surgieron ideas de como mejorar los costos, seguridad y preservación digital interinstitucional bajo el modelo de economía a escala entre las instituciones de interés.

Es importante mencionar que entre las instituciones de patrimonio cultural de Dinamarca mencionadas siempre había existido una buena relación y conocimiento de sus labores de cuidado y gestión de archivos en donde la principal referencia sopesa a la heterogeneidad de sus colecciones y las diferencias de operación de cada una de las instituciones, ya que de manera independiente cada una realizaba actividades de almacenaje digital bajo tecnologías de desarrollo propio y en algunos casos subcontratando el servicio. Estas situaciones los lleva a plantear la idea de crear un proyecto conjunto de preservación digital entre las tres instituciones que resguardan patrimonio cultural danés en el que se pudieran fundir una cooperación interinstitucional en el campo de la preservación digital que les permitiera explorar las necesidades y requerimientos de las mismas bajo una arquitectura compartida de preservación a largo plazo y bajo un control total e independiente de sus colecciones, acceso, derechos de propiedad intelectual y preservación lógica de sus archivos.

Durante el desarrollo del proyecto denominado BitRepository Project se exploraron diferentes posibilidades de desarrollos existentes en el mercado de la preservación digital compartida y

uno de los desarrollos que les inspiró en su búsqueda fue el sistema *peer-to-peer* LOCKSS. El caso específico de desarrollo de BitRepository Project las necesidades especiales sobre las que dependían las misiones y las colecciones de cada una de las instituciones contenían una variedad de diferentes requerimientos que se establecen como prioridades funcionales la seguridad de bit, accesibilidad y confidencialidad; es por ello que la necesidad de contar con una plataforma que encaja en los requerimientos compartidos se planteó la demanda de crear un desarrollo propio y a la medida que cumpliera con el objetivo de las instituciones involucradas y por ello que se desarrolló el sistema y protocolo de BitRepository Project que actualmente funciona en la red de preservación digital distribuida de NDBR (Zierau y Kejser, 2010).

La operación en la gestión de la seguridad y cuidado de archivos con fines de preservación de bits dentro de una red compartida junto en la definición del tipo de colecciones que gestiona cada una de las instituciones, definieron el patrón principal de operaciones y método de trabajo colaborativo donde por ejemplo algunas instituciones dentro de sus colecciones gestionan una mayor cantidad de material audiovisual que otras, mientras las demás dentro de sus objetos digitales prioritarios gestionan por igual archivos o documentos en papel ya sea nacidos digitalmente o digitalizados. Todas estas situaciones en conjunto con otras diferencias técnicas y de organización de las colecciones como los tamaños y peso digital, restricciones de acceso o procedencia, sugerían la definición de un sistema de trabajo bajo niveles de servicio y seguridad entre las diferentes colecciones de los socios de la red. Aunado a esto los requerimientos de confidencialidad y disponibilidad debían contar con la posibilidad de adaptarse a cada una de las colecciones, por lo que el sistema de operación y organización de trabajo necesitaba de una arquitectura que soporta diferentes combinaciones de acuerdo a las necesidades particulares. La premisa y diferenciación principal de la red de preservación distribuida NDBR radica en una arquitectura atípica en la que cada institución gestiona uno o diferentes pilares (pillars) que deben ser completamente independientes y basados en diferentes tipos de tecnología y medios. Cada institución debe tener el control total e independiente de sus pilares, la tecnología y medios que deben variar entre cintas, DVD, NAS, DAS, nube. Esto mismo aplicado a el uso de diferentes sistemas operativos. La paranoia de usar diferentes tipos de tecnología en cada pilar responde a evitar los más conocidos problemas de pérdida de bits mediante la creación de pilares independientes que son operados por diferentes organizaciones y en diferentes localizaciones. Las medidas

también sugieren el evitar la dependencia de un solo proveedor de tecnologías y medios en ayuda a evitar cualquier simple fallo de tecnología única, desastre natural y humano. Es por todo esto que la arquitectura de la red los datos pueden ser replicados en diferentes pilares tal y tanto como sea necesario.

La metodología de trabajo de NDBR se basa en una capa media del sistema la cual permite que cada cliente o institución pueda intercambiar información con el resto de los pilares y la seguridad e integridad de los datos en los pilares es verificada por medio del resguardo de los *checksums*, así mismo de una revisión de *checksums* de las copias entre pilares que es realizada constantemente entre instituciones para verificar cualquier alteración y corrección de los mismos. El objetivo principal de la red NDBR es el proveer de preservación de bits, por lo que es responsabilidad de cada institución el crear, generar y mantener sus metadatos de representación, estructurales y descriptivos para la recuperación y migración futura de sus archivos.

El software de la red NDBR BitRepository se definió que fuese en código abierto para mantenerlo y actualizarlo en comunidad de desarrolladores. La organización de trabajo así como el sistema de la red NDBR podemos decir que es a la medida y donde todos puedan caber de manera flexible bajo un previo y extenso conocimiento de las necesidades de los participantes en un sentido específico y coordinado. Al ser un proyecto de interés de orden cultural nacional el proyecto fue apoyado y financiado por el ministerio de cultura de Dinamarca aportando aproximadamente 525,000 € para el desarrollo del sistema y el equipo humano involucrado en el proyecto (Sørensen, 2014). Los gastos posteriores de implementación, hardware y almacenaje fueron responsabilidad del presupuesto de cada una de las instituciones participantes de la red. La red NDBR se integra por tres instituciones miembros los cuales mantienen una participación especial en cada una de las actividades del repositorio de preservación de bits. La premisa principal de la red NDBR es la arquitectura flexible y la heterogeneidad de las colecciones de los integrantes de la misma, por lo que las relaciones de operación se desarrollan bajo un documento que se ocupa de la organización en la que los miembros a nivel de servicios que necesitan u ofrecen cada uno de los integrantes de la red que se denomina acuerdo de nivel de servicio o service level agreement (SLA).

### ***c) Antecedentes de Red CARINIANA***

Brasil ha mostrado una productividad constante en el área de las bibliotecas digitales, gestión de visibilidad de publicaciones científicas y programas de redes de bibliotecas digitales para países en desarrollo. Un ejemplo clave es el desarrollo del proyecto Scientific Electronic Library Online (SCIELO) que opera desde 1998 y tiene como objetivo el brindar una herramienta de publicación de revistas científicas electrónicas que pueda gozar de una herramienta de visibilidad, distribución y diseminación de la comunicación científica del trabajo de investigación de países de América Latina y el Caribe. La Biblioteca Digital Brasileña de Tesis y Disertaciones (BDTD) integra un sistema de gestión documental de referencia el cual tiene como objetivo el integrar en un solo portal todas las referencias y texto completos de tesis y disertaciones para su acceso a nivel nacional e internacional.

El crecimiento de sus sistemas de gestión digital de publicaciones y documentos bajo un esquema de orden lógico de desarrollo en la gestión de material digital de investigación, genera la necesidad de procurar el acceso continuo y la preservación a largo plazo de los mismos. En el año 2012 nace la red CARINIANA de preservación digital distribuida de Brasil que es la primera y única red de América latina que funciona con el programa LOCKSS de la Universidad de Stanford desde el año 2013 bajo la gestión del Instituto Brasileño de Información en Ciencia y Tecnología (IBIC) con apoyo de la financiadora de estudios y proyectos (FINEP). CARINIANA tiene como objetivo conservar a largo plazo la información científica de Brasil, en específico de los periódicos nacionales que emplean la plataforma OJS, libros electrónicos, tesis y las disertaciones que están almacenados en la herramienta DSpace en una red de arquitectura distribuida nacional e internacional (Márdero, 2013).

De acuerdo con la información del portal de la red CARINIANA<sup>61</sup> la integración de los sistemas de gestión documental y repositorios de instituciones brasileñas al programa LOCKSS requirió de diferentes etapas de implementación como de esfuerzos a nivel técnico y organización como lo son un programa de definición de personal, capacitación, gestión e invitación de instituciones miembros, pruebas de instalación del software LOCKSS, pruebas de envío, etc. En la primera etapa de prueba o *beta* se integró por la instalación e invitación de cinco instituciones de educación de Brasil (USP, UNICAMP, UFPB, UFSM y UEMA) en las

---

<sup>61</sup> <http://cariniana.ibict.br>

cuales los representantes técnicos realizaron la instalación del software LOCKSS en servidores locales y posteriormente definieron 16 títulos de revistas electrónicas como primera prueba de envío a el resto de las cajas LOCKSS. Posteriormente a las pruebas pilotos realizadas y analizando los resultados de las mismas se definió el oficializar el programa de preservación digital realizando reuniones técnicas con los equipos de Brasil y la Universidad de Stanford que sirvieron de manera oficial para establecer los acuerdos de coordinación técnica y políticas necesarias de la red de preservación. En las siguientes etapas de implementación se definieron los políticas de integración de contenido de revistas y documentos electrónicos. El objetivo de las siguientes etapas es incrementar la inclusión de material digital que debe incluir a todas las revistas brasileñas bajo el sistema OJS / SEER registradas en IBIC, así como de libros electrónicos Portal Libro Abierto y del portal de Tesis y Disertaciones de BDTD. El proyecto ha sido desarrollado bajo una estructura descentralizada y distribuida que funciona bajo la participación de instituciones que producen y custodian este tipo de documentos bajo un entorno e infraestructura normalizada de seguridad que puedan garantizar el acceso y monitoreo ininterrumpido del material digital. Actualmente el programa de la Red CARINIANA incluye a seis universidades brasileñas las cuales son: Universidad de São Paulo, Universidad Federal de Bahía, Universidad Federal de Santa Maria, Universidad Estatal de Campinas, Universidad Federal de Goiás, Universidad Federal de Paraíba, Universidad Federal de Rio Grande del Norte.

### **3.3.2 Indicadores de organización, colaboración y coordinación en redes PDD**

El objetivo de la definición y estructuración de un grupo de indicadores relacionados con la organización, colaboración y coordinación de las redes de preservación distribuida y de sus grupos de trabajo activos tiene como objetivo el poder conocer las estructuras, buenas prácticas y posibilidades de mejora en el sistema de organización de redes distribuidas. A partir de la identificación de los principales elementos que definen estos indicadores de trabajo y organización de redes de preservación digital distribuida así como de los antecedentes, documentación e información referente a la organización, gobernanza y coordinación sugerida por diferentes autores como Walters (2010), Provan y Kenis (2007), Huppé, Creech y Knoblauch (2012), así como de instituciones y redes de preservación digital distribuida activas estudiadas como MetaArchive Cooperative, CARINIANA y National Danish BitRepository; hemos integrado y definido un grupo de indicadores específicos sobre

los conceptos funcionales de organización en sistemas de redes distribuidas puntualizando en sus sistemas de trabajo y organización. Como consecuencia de la identificación de indicadores específicos hemos definido cuatro escalas de grupos que integran a la totalidad de los indicadores bajo los rubros de organización y gobernanza, colaboración y preservación distribuida, marco legal y políticas de trabajo, así como coordinación y comunicación (OCMC) que a continuación definimos.

### ***Organización, gobernanza y grupos de trabajo***

La organización y gobernanza integran a los indicadores generales y primarios de las redes de preservación digital distribuida pues a partir de sistemas de organización o esquemas de trabajo se pueden definir responsabilidades específicas o figuras de función humana de autoridad en un grupo de trabajo con necesidades de coordinación como lo son los roles de trabajo, las estructuras de gobierno y la organización de trabajo. Un tema importante que menciona Smallwood (2014) es el de la gobernanza de información que sugiere como punto de unión del desarrollo de políticas y ejes de organización que define políticas de actuación. La organización y gobernanza dentro de la figura de función de roles de trabajo incluye temas específicos como lo son el rol de dirección, colaborador, desarrollador o intermediarios en actuaciones principales.

La figura funcional de estructuras de gobierno integra a indicadores específicos principalmente sugeridos por Provan y Kenis (2007) quienes en su propuesta de modos de gobernanza en red: estructuras, manejo y efectividad; integran las estructuras de gobierno principal, administrativa y participativa que dejan opción a un sistema híbrido de las mismas. Por último mencionaremos la figura de la organización de trabajo como un indicador preponderante que integra una serie de actividades que se ven enriquecidas por la cantidad de personas con las que se puede contar en una red de preservación distribuida como lo son la organización y definición de comités específicos, equipos de trabajo, grupos técnicos o de mantenimiento, gestores de contenidos, grupos que supervisen la parte legal, etc. Todos estos indicadores definen de manera amplia la actuación de la organización y gobernanza de redes de trabajo institucionales.

### ***Colaboración y preservación distribuida***

Uno de los principales tópicos de las redes de trabajo compartidas y en específico de preservación digital distribuida es la colaboración. Esta define principalmente el nivel de compromiso en acciones sobre el trabajo compartido y recursos necesarios para establecer una red. Dentro de un programa, servicio o red de preservación digital debemos recordar que existen actividades específicas de preservación digital que deben ser cubiertas a nivel técnico y organizacional pues suponen valores agregados a la colaboración interinstitucional distribuida tanto en costes de infraestructura, mantenimiento y contratación de servicios externos de preservación digital. Es importante considerar dentro de estas actividades a las figuras funcionales a nivel técnico y colaborativo en servicios de preservación como lo afirma Zierau (2011) ya que cuando se pretende realizar preservación de bits bajo un modelo como el de IR-BR model o modelo de repositorio institucional bit repository esta será en su mayoría sobre sistemas distribuidos con la finalidad de evitar riesgos de la pérdida de bits. En muchos de los casos este sistema distribuido cubrirá por igual la colaboración al colocar las réplicas del material preservado de manera distribuida en diferentes organizaciones en conjunto con acciones de ingesta de contenido, almacenaje, monitoreo y mantenimiento, recuperación y acceso, etc. Se debe recordar que habrá una serie de servicios que otorguen mayores lazos y responsabilidades de colaboración como lo son la interoperabilidad con posibilidades de acceso abierto, normalización de objetos digitales, cosecha de contenidos, manifiestos, dependencias técnicas y sistemas, etc. Tales indicadores en conjunto con la integración a modo colaborativo en infraestructura de recursos técnicos, recursos humanos, pública o privada, de desarrollo a la medida o filosofía de crecimiento sostenibles refieren a este escalón como uno de los más importantes en lo referido a las relaciones de integrantes de red a nivel de acciones, actividades e infraestructura.

### ***Marco legal y políticas de trabajo***

La colaboración, infraestructuras y otros componentes que unen a diferentes actores que integran una red de preservación digital distribuida requieren de una serie de acuerdos, marcos legales y políticas de trabajo que sitien los derechos y responsabilidades de cada uno de los integrantes en un panorama específico de acción ya sea en su figura de miembro como coordinador o integrantes comunes, adhiriéndose garantías e incluso alguna tipología de

infracciones al incumplimiento de los compromisos de trabajo. En *Guidance on Cloud Storage and Digital Preservation* se sitúa que dentro de buenas prácticas de marcos legales y políticas de trabajo de preservación distribuida que se debe asegurar principalmente la satisfacción de necesidades individuales que en los contratos de servicio (SC) o acuerdos de niveles de servicio (SLA) soporta en elementos jurídicos de relación entre archivos y proveedores (servicio o colaboradores), así también del compromiso de proveedores o colaboradores y la garantía del nivel de tiempo de funcionamiento, integridad de datos y continuidad de colaboración. También se deben tomar en cuenta indicadores sobre los sentidos de pertenencia y federación que una red distribuida de preservación digital puede dictar como privada o mediante un servicio como lo son las necesidades de membresía y cuotas, las tipologías de grupos especiales de trabajo, la forma en cómo se usan y se manejan los contenidos en ellas, así como si los integrantes pertenecen a una red heterogénea u homogénea de una misma institución (The National Archives, 2014). Los marcos de trabajo en un grupo de integrantes y usuarios de una red compartida con fines de preservación digital como los acuerdos, niveles de trabajo y servicios deben ser construidos bajo documentos de contratos o acuerdos, derechos de autor, niveles de servicios, etc.

### ***Coordinación y comunicación***

La coordinación de grupos heterogéneos de usuarios en una red de preservación digital en combinación de sus necesidades y comunicación deben considerarse como un indicador para conocer diferentes niveles requeridos de comunicación institucional, interacción de los miembros, así como su cultura y visibilidad. La actividad o pasividad de los integrantes en conjunto como la percepción de comunicación entre ellos como obligación o necesidad, la identidad institucional y la constancia o periodicidad de comunicación; definen aún más las fortalezas de la coordinación. Los miembros deben interactuar en diferentes niveles ya sean de tipo técnico o de conocimiento y entendimiento institucional bajo ciertas estructuras protocolarias como reuniones, eventos, formaciones o asesoría que requieren de herramientas puntales de comunicación para medir y facilitar su interacción. Costa (2009) afirma que la comunicación eficaz antes de ser un “instrumento para...” es, necesariamente, una estrategia. La cultura y visibilidad de las instituciones e integrantes de una red debe reflejar claramente su compromiso institucional así como difundir el proyecto de preservación para lograr un impacto social que pueda de una u otra manera tocar más fibras de la sociedad para sumar

esfuerzos en las redes y compromisos institucionales de preservar patrimonios culturales, económicos o de salud, que faciliten la visibilidad de resultados públicos de los miembros de una red o de un grupo de instituciones comprometidas con ello.

A modo de resumen podemos colocar a los indicadores para valorar la coordinación, colaboración y organización de la redes de preservación digital distribuida en una tabla propuesta que nos ayuda a definir una herramienta de conocimiento de buenas prácticas en un entorno de valoración de actividades reales. La siguiente tabla de indicadores denominada OCMC incluye a la organización de gobernanza, la colaboración de preservación digital distribuida, los marcos legales de trabajo y la coordinación de comunicación en la misma.

Tabla 14. *Indicadores generales OCMC.*

<b>Indicadores generales OCMC</b>			
<b>Organización y gobernanza</b>	<i>Rol de trabajo</i>	<i>Estructura de gobierno</i>	<i>Organización de trabajo</i>
<b>Colaboración y preservación distribuida</b>	<i>Servicio de preservación</i>	<i>Infraestructura</i>	<i>Interoperabilidad</i>
<b>Marco Legal y Políticas de trabajo</b>	<i>Derechos y responsabilidades</i>	<i>Pertenencia y federación</i>	<i>Acuerdos y niveles de trabajo o servicio</i>
<b>Coordinación y comunicación</b>	<i>Comunicación institucional</i>	<i>Interacción de miembros</i>	<i>Cultura y visibilidad</i>

En la tabla de indicadores podemos correlacionar los principales elementos para darle una mayor profundidad de posibilidades a cada campo e indicadores que requiera un tratamiento práctico y de implementación para trasladarlo a las posibilidades de buenas prácticas.

Tabla 15. Indicadores generales y específicos OCMC.

<b>Indicadores generales y específicos OCMC</b>			
<b>Organización y gobernanza</b>	<b>Rol de trabajo</b> <ul style="list-style-type: none"> <li>- Directivo</li> <li>- Colaborador - socio</li> <li>- Desarrollador</li> <li>- Intermediario</li> </ul>	<b>Estructura de gobierno</b> <ul style="list-style-type: none"> <li>- Principal</li> <li>- Administrativa</li> <li>- Participativa</li> <li>- Híbrida</li> </ul>	<b>Organización de trabajo</b> <ul style="list-style-type: none"> <li>- Comites de personal</li> <li>- Equipos de trabajo</li> <li>- Grupos técnicos y m.</li> <li>- Gestores de contenido</li> <li>- Legal y supervision</li> </ul>
<b>Colaboración y preservación distribuida</b>	<b>Servicio de preservación</b> <ul style="list-style-type: none"> <li>- Ingesta contenido</li> <li>- Almacenaje replica</li> <li>- Monitoreo y Mantenimiento</li> <li>- Recuperación y acceso</li> </ul>	<b>Infraestructura</b> <ul style="list-style-type: none"> <li>- Recursos Técnicos</li> <li>- Recursos Humanos</li> <li>- Pública / Privada</li> <li>- Desarrollo a medida</li> <li>- Crecimiento Sostenibilidad</li> </ul>	<b>Interoperabilidad</b> <ul style="list-style-type: none"> <li>- Acceso abierto</li> <li>- Normalización</li> <li>- Estándares</li> <li>- Cosecha y manifiestos</li> <li>- Niveles de operación</li> <li>- Dependencia</li> </ul>
<b>Marco Legal y Políticas de trabajo</b>	<b>Derechos y responsabilidades</b> <ul style="list-style-type: none"> <li>- De miembro</li> <li>- De coordinador</li> <li>- Comunes</li> <li>- Garantías e infracciones</li> </ul>	<b>Pertenencia y federación</b> <ul style="list-style-type: none"> <li>- Membresía y cuotas</li> <li>- Grupos especiales</li> <li>- Uso y manejo de contenidos</li> <li>- Grupo de instituciones</li> </ul>	<b>Acuerdos y niveles de trabajo o servicio</b> <ul style="list-style-type: none"> <li>- Contratos y acuerdos</li> <li>- Documentos</li> <li>- Copyright</li> <li>- Nivel de servicio</li> </ul>
<b>Coordinación y comunicación</b>	<b>Comunicación institucional</b> <ul style="list-style-type: none"> <li>- Pasiva / activa</li> <li>- Obligada / necesaria</li> <li>- Identidad institucional</li> <li>- Constante / periodica</li> </ul>	<b>Interacción de miembros</b> <ul style="list-style-type: none"> <li>- Conocimiento y entendimiento institucional</li> <li>- Reuniones y eventos</li> <li>- Capacitación y asesoría</li> <li>- Herramientas de interacción</li> </ul>	<b>Cultura y visibilidad</b> <ul style="list-style-type: none"> <li>- Compromiso institucional</li> <li>- Difusión de proyecto</li> <li>- Impacto social</li> <li>- Resultados públicos de miembros o grupo</li> </ul>

### 3.3.3 Aplicación de estudio y valoración OCMC

Para la aplicación del estudio, recopilación y valoración de buenas prácticas referidas en los indicadores OCMC llevamos a cabo una metodología de estudio y opinión bajo una selección de muestra no probabilística. Los participantes se seleccionaron en función de los contactos públicos de los sitios web de las redes PDD y los responsables de nodos o instituciones

participantes en la red de preservación digital distribuida, así como de contactos referidos por expertos en la materia para el estudio. La invitación a participar concurre de manera voluntaria y flexible por lo que definimos de un universo total de responsables o contactos de instituciones pertinentes como muestra de alcance de mínimo de un participante y un máximo de tres participantes por cada una de las tres redes de preservación digital distribuida (MetaArchive Cooperative, la red National Danish Bit Repository y la red CARINIANA de Brasil). Los resultados y respuestas se analizaron posteriormente a modo cualitativo de valoración y opinión como un grupo unificado de usuarios de redes de preservación digital distribuida para conocer generalidades de uso, relación y práctica con las terminologías OCMC entre los participantes de redes. Hay que dejar claro que este estudio no pretende obtener resultados porcentuales o proporcionales, sino no como lo hemos dicho antes pretende obtener opiniones, referencias y validaciones de buenas prácticas alrededor de los indicadores OCMC para reiterar y evaluar un marco de referencia de elementos de organización que nos ayuden a la definición de una ruta de modelo de trabajo y conceptos más aplicados alrededor de los indicadores OCMC.

Para la aplicación del estudio se elaboró un instrumento de recolección integrado en cuestionario con los indicadores OCMC. El cuestionario de manera primaria se realizó en castellano para traducir al portugués y al inglés para su aplicación en los usuarios de interés de Red CARINIANA, National Bit Repository y MetaArchive Cooperative. Posteriormente con la ayuda de la herramienta de formularios de Google<sup>62</sup> se crearon tres formularios en sus respectivos idiomas para aplicarlos en línea para facilitar su envío y acceso a los usuarios de redes. Para la puesta en marcha y envío del estudio se llevaron a cabo acciones de contacto vía correo electrónico, redes sociales como LinkedIn<sup>63</sup> y entrevistas personales con la finalidad de hacer llegar los formularios de la manera más transparente posible y explicar el objetivo del mismo. Es importante mencionar que en cada uno de los contactos e incluso en el formulario se dejó claro la finalidad del estudio y por otro lado que los resultados serían anónimos con respecto a su fuente para evitar especificación de prácticas y por ende señalización personalizada.

---

<sup>62</sup> Formularios de Google <https://www.google.es/intl/es/forms/about/>

<sup>63</sup> LinkedIn <https://www.linkedin.com>

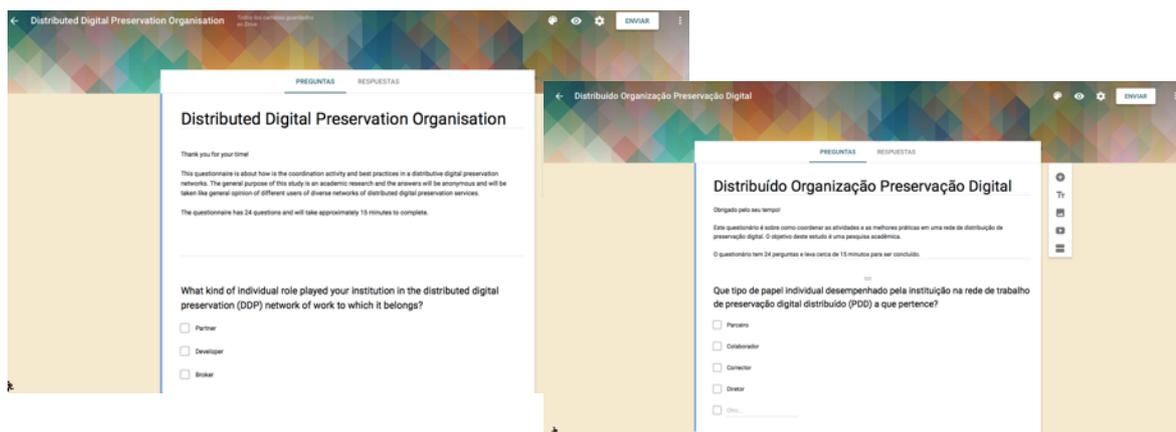


Figura 8. Formularios de estudio de buenas prácticas de organización

### 3.3.4 Análisis de respuestas y valoraciones al estudio OCMC

Las respuestas y valoraciones obtenidas del estudio corresponden a la participación de cuatro usuarios que voluntariamente accedieron a contestar y valorar los indicadores a quien agradecemos plenamente su disposición y participación. De los cuatro usuarios participantes debemos decir que dos de ellos forman parte de una red específica y los otros dos participantes forman parte de las otras dos redes respectivamente, por lo que el estudio concreta la participación deseada de mínimo un usuario experto por red PDD. Como lo hemos explicado antes para el análisis de las respuestas de opinión y valoración al cuestionario de indicadores OCMCs organizamos las respuestas de los participantes individuales de manera global y descriptiva de acuerdo a los cuatro grupos en los que se organizaron los indicadores en el cuestionario para de esta forma situar las valoraciones específicas dentro de las respuestas a las preguntas, y así poder desarrollar sus valoraciones con validaciones de los conceptos para entretejer un caso de estudio general y posibilitar un rumbo coherente de sugerencias sobre sus prácticas mencionadas en las respuestas. Cabe señalar que algunos usuarios manifestaron que algunas preguntas y conceptos resultan redundantes por lo que algunos plantearon como respuestas extendidas en el campo de respuesta “otros” y complementaron de manera enriquecida esquemas adicionales en el campo de la organización y retos de la preservación digital distribuida.

## ***Organización, gobernanza y grupos de trabajo***

En las respuestas obtenidas sobre este grupo de indicadores de referencia a la organización y roles la mayoría de los usuarios al referirnos sobre el papel individual o rol de trabajo que desempeñan en la red de preservación digital distribuida a la que pertenecen, afirman que mantienen un *rol de trabajo* de *socio* de manera integral e individual. El caso de uno de ellos refiere a su papel individual de trabajo como *broker o intermediario* de un grupo de instituciones a las que se les otorga el servicio de preservación distribuida y que funcionan como miembro representante de una red ante la red de servicio de preservación. Otro usuario se define como *desarrollador y socio*.

Hay que apuntar que estas opiniones nos arrojan tres roles de organización de trabajo interesantes con respecto a el tratamiento de servicios externos o en su caso de una red de preservación privada, ya que en una organización de trabajo de este tipo intervienen diferentes capas como roles, y en donde pueden existir una, dos o hasta tres *capas de organización* de trabajo. Estas capas ayudan a gestionar a uno o varios grupos de tipo colaborativo y distribuido que podemos denominar bajo los indicadores de organización OCMC en sus diversos roles como la *capa de socios colaboradores*, *capa de intermediarios* y la *capa de directivos o administradores de servicio*. No existe un modelo rígido de organización colaborativo de roles de trabajo que cierre las opciones de participación o en su caso decida un orden específico, por lo que mantiene una flexibilidad de integración. En cuanto a las opiniones de los usuarios sobre ventajas y desventajas del rol de trabajo que desempeñan, reflejan que hay más *ventajas* que desventajas. Todos coinciden en que encuentran como gran ventaja en su posición dentro de la red el asegurar *tener preservados de manera distribuida y con copias extras sus acervos institucionales, académicos y de investigación*. Un usuario menciona que las ventajas en realidad pueden ser expresadas meramente como *necesidades por cubrir*, ya que en sí la preservación digital distribuida es una necesidad al realizar una estrategia de preservación digital, así como las auditorías, la colaboración y la creación de un sistema flexible son necesidades que pueden ser ventajas aunque para que sean ello se requiere de trabajo duro y constancia plena. Como desventajas dentro del rol de trabajo en una red de preservación digital distribuida uno de los usuarios manifiesta que estas *desventajas* podrían ser las *dificultades para involucrar a las personas y la disponibilidad de los recursos humanos* para el desempeño de las funciones. Otro usuario comenta que las desventajas

principalmente son las *complejidades y dificultades que se encuentran en el proceso y mantenimiento* pleno de un sistema de preservación digital distribuida.

En lo que respecta al conocimiento del tipo de *estructura de gobernanza o gobierno* de la red a la que pertenecen la mayoría de los participantes afirma que la red a la que pertenecen mantiene un gobierno *administrativo, participativo y de organización principal*; solo la mitad afirma que es *centralizado*, aunque en esta línea no se concreta un solo tipo de gobernanza predominante, lo que se percibe son *redes flexibles y colaborativas*. En solo un caso se afirma que se desconoce a grandes rasgos el tipo de gobierno por terminología, sin embargo se afirma que se mantiene una gran colaboración en la misma por lo que se sobreentiende que la gobernanza es participativa. Sobre ventajas del tipo de gobierno de la red a la que se pertenecen un participante menciona como ventaja a una gobernanza distribuida ante una centralizada. Los demás participantes no confirman otras ventajas al respecto. Como desventaja uno de los participantes menciona que en el tipo de gobierno que tiene su red se requiere de mayor participación de los técnicos en la misma. Otro menciona que no encuentra ninguna desventaja ya que los objetivos que persiguen con respecto a su participación como el tener una copia extra preservada y distribuida geográficamente antepone todo, incluso el no tener voto en una estructura de gobierno a nivel de la función general de la red.

El indicador de tipo de gobierno se muestra como un tópico de conocimiento difuso a su objetivo y como indicador de organización de un grupo de instituciones. La gobernanza debe ser referida de manera clara para todos los miembros de las redes a fin de contener la estructura y como permean las acciones individuales, de grupo y de red por escrito bajo un documento que defina la estructura de organización de red.

Con respecto a la *organización o grupos de trabajo* la mayoría afirma que forman parte de comités técnicos, comités de contenidos y comités especiales. Un participante del estudio afirma que no pertenece a ningún grupo o comité específico pero sí en algunas ocasiones ha colaborado en parte técnica, pero no es parte de sus funciones o de un grupo específico. En el caso de las ventajas y desventajas algunos participantes afirman que una de las ventajas de formar parte de comités o grupos de trabajo es el intercambio de experiencias y desarrollo profesional en el trabajo de una red distribuida. Como desventajas un participante afirma que no encuentra ninguna ya que nivel de trabajo (como intermediario), su labor técnica y servicio

ofrecido a el grupo de instituciones que representa no supone problemas mayores de esfuerzo técnico pues se tiene plenamente coordinado a nivel técnico y humano la atención de su nodo de preservación. Por otro lado, algunos usuarios anteponen como desventajas el lugar donde encuentra la coordinación de la red (otra ciudad), así mismo el tiempo que lleva la actualización de datos. Otro participante refleja como desventaja que además de su participación en la red distribuida tienen que realizar otras actividades paralelas, por lo que no les deja tiempo para dedicarse de manera plena a la red.

En este subgrupo de indicadores y opiniones encontramos que mientras en el modo de organización o capa de intermediario las labores técnicas o de organización mantienen un mayor control ante actividades de trabajo técnico y colaboración como nodo de una red. Esta situación puede ser a causa de que esta capa superior otorga un grado extra de coordinación que atiende de manera personalizada las necesidades de un grupo de instituciones. En el caso de los participantes individuales que se encuentran en una capa inferior a la de coordinación vemos que se acentúan de algunos inconvenientes con respecto a la atención, la centralización y el tiempo de atención a las actividades integrales de su participación como nodo de preservación distribuida. Estas valoraciones nos dejan entrever los diversos ejercicios y donde pueden colocarse las situaciones que dejan a la organización, la gobernanza y los grupos de trabajo en posibilidades de mejora de su colaboración en red.

### ***Colaboración y preservación distribuida***

En lo que respecta a los indicadores del grupo de colaboración y preservación distribuida las consideraciones sobre las que estudiamos los indicadores OCMC corresponden a conocer en primer lugar cómo colaboran y que tipo de conocimiento existe sobre los *servicios de preservación digital* que ofrecen las redes distribuidas a las que pertenecen los participantes. Los usuarios participantes manifestaron de manera mayoritaria que el tipo de servicios ofrecidos por la red de pertenencia se centra en el almacenamiento replicado y en la reparación de contenido replicado. Otro usuario agrega además de estos, los servicios a el monitoreo de checksums, integridad y seguridad de preservación de datos de forma distribuida. Estas valoraciones dejan ver que los participantes tienen claros los servicios y su colaboración como nodo en la red distribuida de preservación, así mismo que mantienen las premisas de replicación, seguridad e integridad como servicios fundamentales ofrecidos y en

los que a su vez colaboran activamente. Sobre las *ventajas* de los servicios ofrecidos por las redes distribuidas se valoró de manera amplia la seguridad de los contenidos preservados. Sobre las *desventajas* un usuario afirma que en general la principal desventaja se ve reflejada en *el acceso y procesamiento en masa (de contenidos digitales) ante la seguridad de bits y la confidencialidad*, lo cual representa un reto para todas las redes PDD. Sobre la *infraestructura colaborativa* que ofrece la red a la que se pertenece se confirma que está de manera mayoritaria es infraestructura técnica y de recursos de acceso permanente. En esta misma línea otro usuario complementa que además de estas infraestructuras técnicas que ofrece el software o sistema por los que corren los servicios de la red una infraestructura importante que les brinda la red a la que pertenece es la de contar con socios colaboradores para intercambio de copias replicadas y distribuidas geográficamente. Otro participantes complementan que además de estas se ofrecen *infraestructuras humanas y planes de sustentabilidad* para la red de PDD. Sobre las *ventajas* de los servicios e infraestructuras ofrecidas por las redes de preservación a los usuarios se destaca la del mantenimiento y actualización de software, así como de equipamiento. Otro usuario destaca como ventaja que *se evita el uso infraestructuras técnicas centralizadas* que son vulnerables y costosas en su operación y existencia. Como *desventaja* un usuario apunta que en una infraestructura técnica la *asociación depende de la colaboración*.

Otros de los servicios destacados por los usuarios de las redes es el de la *interoperabilidad* de la cual destacan principalmente los servicios de harvesting o cosecha de contenidos. De igual manera, destacan el acceso abierto al material de otros nodos (para réplica y no consulta), el acceso permanente, así como de la integración de manifiestos y servicios técnicos (plugins) para dinamizar la interoperación de los servicios entre los socios o nodos. A nivel de interoperabilidad algunos usuarios refieren que la red a la que pertenecen les ofrece una capa extra de gestión o interfaz de usuario que denota la orquestación sobre la cual se gestiona el orden sobre el cual actuarán los nodos en el flujo, distribución y reparto equilibrado de nuevos nodos en la red, así como de la recepción o envío de material a los mismos. Un usuario expresó como desventaja que el sistema de interoperabilidad que utilizan es un sistema muy frágil para la seguridad.

Podemos observar que para los usuarios y colaboradores de una red de preservación digital distribuida las prestaciones de colaboración, servicios, infraestructura e interoperabilidad de

acciones de preservación digital distribuida es fundamental el soporte técnico, asesoría e infraestructura tecnológica que la red a la pertenece les ofrece pues ello garantiza la capacidad de transparencia y dinámica de colaboración entre los distintos participantes de la red para procurar la seguridad de sus contenidos y los objetivos de formar parte de una estrategia distribuida de preservar digital.

### ***Marco legal y políticas de trabajo***

En lo que respecta al marco jurídico, legal o políticas de trabajo colaborativo los usuarios de redes distribuidas mantienen una situación similar en el conocimiento de su posición sobre los *derechos y obligaciones* que les comportan dentro de la red. Algunos consideran que son *usuarios normales* con derechos y obligaciones similares a otros miembros de la red, por otro lado también se definen como *miembros especiales* que cuentan con el mismo marco de acción además de gestionar a un grupo de instituciones ante el servicio de conexión a la red. Un usuario manifiesto sus posiciones como *coordinador y desarrollador* de su red PDD. En esto debemos reiterar que aunque existan capas de gestión grupales, los derechos y obligaciones permean de manera similar en toda la sombrilla de usuarios, por lo que la denominación que caracteriza a un usuario dentro de una red de preservación digital distribuida debe puntualizar de acuerdo en una etiqueta de miembro que defina sus límites y coberturas de manera clara de todos los usuarios. Un participante afirma que es necesario tener o *contar con responsabilidades distribuidas para lograr la independencia entre réplicas*, aunque afirma que esto es duro de conseguirlo. En este indicador de derechos y responsabilidades tiene una actuación directa el tipo de *federación o pertenencia de servicio* que como miembro tenga cada nodo en la red, y esto depende directamente de la aportación económica (mayoritariamente) o de desarrollo técnico de la red para fundir el termómetro de responsabilidades y obligaciones.

En el caso del *sentido de pertenencia* de los usuarios a una red se refleja en que los mismos tienen por entendido que son parte de una grupo de instituciones o un grupo especial de las mismas dentro de una programa de preservación digital distribuida. Otros se manifiestan como los representantes de un grupo usuarios. Esta situación podría expresarse de forma más específica si se aplica directamente a los diferentes nodos que conforman una red si el estudio solo es planteado a instituciones directas estas representan puramente nodos y prácticas

independientes. En el caso de consultar a un nodo que gestiona a una sub red de usuarios, este determinara que la pertenencia es también a un grupo especial que tiene su razón de ser en la cooperación y gestión efectiva. Algunos participantes manifiestan que pueden existir algunas desventajas sobre el mantener una pertenencia o federación a una red, ya que si se quisiera formar parte de otra red se complicaría pues el sistema está preso o pegado al sistema actual de preservación. Esta opinión de desventaja nos otorga una huella de la importancia de plantear la pertenencia a una red privada y cuáles podrían ser sus ventajas a niveles de exclusividad, seguridad y gestión más allá del aislamiento o imposibilidad de añadirse a otras redes. Para que la comunión colaborativa de diferentes usuarios de una red con fines similares funcione debe manifestarse bajo un *acuerdo inicial o contrato de servicio* que soporte la colaboración, cooperación o en su caso trabajo en red.

La definición jurídica o legal sobre la cual cada miembro mantiene una actuación directa en la red, algunos usuarios manifiestan que ante dicha necesidad se funge como una prioridad de conocimiento de pertenencia y federación a una red, así como el documento SLA (Acuerdo de nivel de servicio). Otros manifiestan que los acuerdos de trabajo deben quedar manifestados en contratos y documentación o contratos de grupo que establezcan las reglas y obligaciones de los mismos. Una opinión interesante al respecto es la de que *estos documentos deben dejar claro por ejemplo los tipos de derechos específicos que goza un nodo (individual o colaborativo)* para que por ejemplo *en caso de desastre o pérdidas de un nodo por razones inesperadas, el resto de los nodos colaboren en la restauración del material perdido y por otro lado aporten ayuda ya sea en desarrollo o económico*. Esta valoración viene a cuenta de que en el caso de algunas redes parte del presupuesto de las aportaciones que hacen los miembros de una red se guarda a modo de reserva de desastres para apoyar a un miembro que los necesite como tema de recuperación a nivel cooperativo y compromiso de la pertenencia a una red colaborativa o cooperativa, el cual queda por escrito en un acuerdo de nivel de servicio o contrato de trabajo. Un participante afirma que existen muchos acuerdos en diferentes niveles por lo que es difícil situarlos a todos en uno solo. Un usuario refiere *desventajas* en su participación en la red y de los acuerdos ya que estos le obligan a mantener actualizados las infraestructuras y la compatibilidad económica puede complicarse. Esta valoración de nuevo muestra la importancia de necesidades y requisitos de perfil para integrarse en una red de preservación digital distribuida pues si no se cuenta con estabilidad y sostenibilidad económica que asegure actualizaciones, interacción, mantenimiento y

operación de los sistemas se podría comprometer de manera extraordinaria a los demás usuarios y sus porcentajes de participación en la red.

Como podemos observar las opiniones vertidas en este sentido otorgan un alto ímpetu de formalidad que debe ser entendida más a ella de un compromiso institucional como un compromiso integral visto por igual a largo plazo, donde se denota que algunos servicios cuentan con un sentido de integración más maduro y controlado que otras redes donde la inconsistencia y claridad del marco legal que les envuelve puede hacer difusa la interpretación de responsabilidades y obligaciones.

### ***Coordinación y comunicación***

Ante la gestión de grupos de organizaciones o instituciones que forman parte de una red de preservación digital distribuida los planteamientos de comunicación deben centrar sus ejes de manera sólida para conocer no solo en qué terreno se está actuando, si no también quien les acompaña. Al incidir en el indicador de cómo es la *comunicación de las instituciones dentro de la red* a la que pertenecen algunos usuarios manifestaron de manera concreta que es *periódica e institucional*, por su parte otro usuario manifiesta que su comunicación con los demás nodos a los que pertenece su red es *nula* ya que su comunicación institucional se realiza con la fuente del servicio de la red, es decir la capa de organización superior ya que los nodos replicados incluso pueden ser cambiantes según la estrategia de distribución y equilibrio de la red. Un participante define su comunicación en un espectro más amplio, dejando claro que es *activa, constante, periódica e institucional*, y refiere que hay comunicación en muchos niveles. La comunicación institucional con las capas de administración es más formal, la comunicación interinstitucional se plantea a nivel informal.

Hay que mencionar que algunas redes realizan eventos especiales, grupos de trabajo o comisiones de seguimiento donde reúnen a sus colaboradores o nodos y es donde se produce una interacción que podría considerarse institucional. Las capas intermedias o administrativas formales realizan actividades para dinamizar a las instituciones en su comunicación cooperativa o de colaboración. Un participante pondera y complementa que es muy importante *conocer y entender el trabajo que realizan los demás miembros de la red* y el intercambio de conocimiento en situaciones específicas ya que pueden existir muchos y

diferentes tipos de interacciones entre los miembros además de las formales. Algunos participantes de las redes sugieren que podría ser una *desventaja la poca participación que hay entre ellas* en su comunicación pues esta podría enriquecer más allá de los eventos a experiencias y mejoras de prácticas.

Un tema que encierra de cierto modo el objetivo de una red distribuida es el hecho de contar con un reflejo de la labor y las posibilidades de apoyos futuros. En la consulta de los participantes sobre la *visibilidad de la cultura de preservación* que llevan a cabo y cómo se comunica, algunos de ellos manifestaron que el *nivel de compromiso de preservación es un indicador importante que debe ser manifestado* y por ende revisado continuamente para su valoración y mejora. Por otro lado también se considera que la comunicación de los resultados de la estrategia operativa de preservación refiere a un impacto social importante que debe ser comunicado. Algunos participantes definen que existe una necesidad de buenos resultados. Por último, un participante reitera que el alto nivel de compromiso y entendimiento de la importancia del riesgo de los datos es una necesidad para la preservación a largo plazo. Estas opiniones reflejan un tema importante en una estrategia de tal índole, ya que nos sugiere la necesidad de contar no solo con un sistema de comunicación que otorgue fama a la actividad, si no que de manera real logre aterrizar los beneficios de inversión y sostenimiento de una estrategia cooperativa y de colaboración institucional a nivel de comunicación interna como externa, por lo que no debe dejarse a un lado el ejercicio constante de interacción, conocimiento y valoración de todos los pasajeros del viaje para así mantener un objetivo en el mismo horizonte de equilibrio y beneficios. Estos indicadores de opinión y buenas prácticas de organización, colaboración, acuerdos y comunicación de manera integral nos ayudan a tener una panorámica cercana de las necesidades por cubrir en el entretejido de nuestro sentido de trabajo en conjunto. Los planteamientos y ADN colaborativo que debe existir en los planteamientos para fundar una red de preservación digital distribuida más allá de sugerir las características necesarias de tecnología e infraestructuras, se centran en el ejercicio humano por establecer un rasgo de importancia por la memoria colectiva de grupos que pueden ser el modelo de otros grupos y subgrupos para concretar un sentido global como lo plantea la UNESCO para preservar el patrimonio digital la humanidad, donde el factor de organización en la preservación digital tiene un peso mayoritario en la organización humana a diferencia de la implementación de tecnología.

### **3.4 Estudio y análisis de actividades de preservación digital en IES públicas de México**

El objetivo de este estudio es medir y analizar el estado actual en la gestión de contenidos digitales y almacenamiento con fines de preservación digital a largo plazo en instituciones de educación superior superior de México (IES-MX) con ayuda de métricas de buenas prácticas relacionadas a los niveles deseables o de prácticas recomendadas en la preservación digital según los estándares de uso en métricas internacionales. Para tal motivo se creó una metodología de estudio y análisis específico que nos ayudó a definir y seleccionar una muestra conveniente del universo institucional de educación superior en México que con ayuda de una herramienta y protocolo de recolección de datos se obtuvo un acercamiento parcial de actividades y acciones de preservación digital en las áreas de IES que gestionan recursos digitales como archivos, repositorios, depósitos digitales o bibliotecas digitales y conocer sus niveles y categorías de implementación.

#### **3.4.1 Antecedentes y selección de IES-MX para estudio**

Desde hace más de una década una gran cantidad de IES en México realizan grandes esfuerzos por llevar y entregar a usuarios de las mismas información relevante bajo y con la ayuda de herramientas en línea como bibliotecas digitales, repositorios institucionales, depósitos digitales, boletines digitales de nuevas adquisiciones, etc. Las experiencias aún siguen siendo contrastantes y parciales ya que se siguen presentando distintas dificultades tanto a nivel técnico, operativo, presupuestal y de colaboración que inciden directamente en diferentes momentos de organización, operación y mantenimiento con fines de preservación de la información electrónica/digital (Arreola, 2012). En noviembre del año 2014 el Consejo Nacional de Ciencia y Tecnología (CONACYT) publicó los lineamientos generales para la creación del repositorio nacional y repositorios institucionales<sup>64</sup> de IES y otras dependencias de interés. Los lineamientos persiguen concretamente los siguientes objetivos: “I. ampliar, consolidar y facilitar el acceso a la información derivada de las actividades académicas, científicas, tecnológicas y de innovación nacional e internacional a texto completo, en

---

<sup>64</sup> <http://www.conacyt.mx/sicyt/images/Lineamientos-Repositorio-Nacional.pdf>

formatos digitales. II. Fijar los requisitos para acopiar, integrar, estandarizar e interoperar. III. Establecer las políticas que regulen la seguridad, almacenamiento, sostenibilidad, así como la gestión y preservación de la información, y IV. Definir criterios y estándares de calidad en las políticas de comunicación pública de ciencia, tecnología e innovación”, (CONACYT, 2014). Esta iniciativa anticipa un gran empuje a la canalización, normalización e integración del universo de IES-MX a la creación linear formalizada de repositorios institucionales y de la integración de comunicación por disposición de actividades de interoperabilidad de datos activos que agrupe una representación a un repositorio nacional en modo recolector. En este renglón es importante mencionar lo que se estipula en el objetivo número tres de los lineamientos que va sobre el establecer políticas que regulen la seguridad, almacenamiento, sostenibilidad, gestión y preservación de la información. Esta referencia reitera la importancia y responsabilidad de mantener los objetos digitales de los repositorios institucionales bajo los parámetros y lineamientos que aseguren su acceso a largo plazo.

La preservación digital como otras actividades dentro de la gestión de recursos de información en las IES mexicanas aún encuentra contrastes por la adaptación irregular o no formalizada, lo cual refiere a un marco abierto de situaciones para entender el estado de la gestión de sus prácticas pues el universo de IES-MX está integrado por tres niveles que son el de técnico superior o profesional asociado de universidades tecnológicas y otras, la licenciatura normal o carrera docente universitaria y tecnológica, y el posgrado de especialidad, maestría y doctorado. En estos niveles pertenecientes a un mismo subsistema de educación encuentra diferencias en su producción, almacenamiento y organización de objetos digitales de acuerdo a distintos factores como los de infraestructura, especialización en el tema y necesidades de gestión.

### ***Selección de IES-MX para estudio***

El número de IES-MX que integran todo el segmento de instituciones de educación superior públicas se compone por 847 instituciones y se subdivide en 11 poblaciones o universos tipificados de IES-MX (tabla 6). La necesidad objetiva de conocer el nivel de comportamiento y necesidades de IES-MX estatales y federales ante la gestión de preservación digital corresponde a una representación finita de dos poblaciones a nivel pragmático. De acuerdo la revisión formal de literatura referida en Arreola (2014), Rodriguez (2012), Espinosa (2014),

CONACYT (2014), así como de herramientas y registros digitales de directorios como Registry of Open Access Repositories (ROAR), The Directory of Open Access Repositories (DOAR) y la Red Mexicana de Repositorios Institucionales (REMERI<sup>65</sup>); se certifica que los subsistemas federal y estatal contienen a la mayoría de instituciones públicas que de manera activa cuentan con herramientas formales de gestión de archivos digitales como repositorios, depósitos o almacenes digitales que posibilitan a nuestro estudio la representación necesaria específica. El segmento o universo seleccionado de IES mexicanas a estudiar se centra en los compuestos por 9 IES federales y de 34 IES estatales. Cabe aclarar para mantener la objetividad de las prácticas referidas en las IES del estudio, los resultados publicados no referirán de manera específica a universidades o centros de educación superior. Es por ello que se asignan nomenclaturas que nos facilitaran la categoría de segmentos IES para así evitar confusiones y agilizar el procesamiento de los resultados del estudio. Las nomenclaturas asignadas a las IES quedarán definidas del rango de IES-1 a IES-43 y estas se registraron de manera en la que las IES respondan al estudio y así, se completará en una lista el total de participantes. Bajo la definición anterior el universo total a medir se integra por 43 IES de la suma de 34 IES estatales y 9 IES federales de acuerdo a la clasificación de la Subsecretaría de Educación Pública en México.

### **3.4.2 Indicadores de niveles de preservación digital para estudio**

Dada la necesidad de conocer y medir el nivel de prácticas, seguimiento, implicación y uso de protocolos de preservación digital en IES-MX, en el estudio se empleó de modo amplio la metodología y guía de buenas prácticas denominada “Levels of Digital Preservation (LDP)”<sup>66</sup> desarrollada por National Digital Stewardship Alliance (NDSA). Esta guía define el estado de actividades aplicadas de manera escalonada sobre los niveles recomendados de preservación digital en las organizaciones. NDSA - LDP se integra por buenas prácticas y guías estructuradas en 36 actividades de preservación digital englobadas en cinco categorías y ponderada en cuatro niveles - acciones específicas de preservación digital.

Como antecedente y de acuerdo al documento “The NDSA Levels of Digital Preservation: An Explanation and Uses” (Phillips, Bailey, Goethals y Owens, 2013) la NDSA es un grupo

---

<sup>65</sup> <http://www.remeri.org.mx/portal/index.html>

<sup>66</sup> <http://www.digitalpreservation.gov/ndsactivities/levels.html>

diverso formado por más de 140 organizaciones que tienen como objetivo el establecer, mantener y anticipar la capacidad para preservar los recursos digitales nacionales (EE.UU) para el beneficio de las generaciones presentes y futuras, es por ello que crearon los niveles de preservación digital. NDSA-LDP tiene por objetivo principal el conocer en contraste con sus buenas prácticas el nivel de alcance en donde se encuentra una organización ante las actividades de preservación digital. Los autores de la guía definen por igual cinco objetivos de uso que de la misma como lo son: 1) identificar dónde hay y donde no hay un consenso general en la comunidad de preservación, 2) educar y desarrollar directrices para los creadores de contenidos y colaboradores, 3) validar la orientación de preservación dada localmente, 4) desarrollar requisitos para los proveedores de servicios de preservación de terceros, 5) evaluar el cumplimiento de las mejores prácticas de preservación e identificar áreas clave para mejorar.

La guía refiere a un sentido de holística formal de posibilidades enfocadas a actividades de preservación digital referidas en un nivel de implementación real en organizaciones que puede servir de referente primario para diagnosticar las mismas. Dada la obtusa situación del estado de la cuestión práctica de las IES (federales y estatales) de México ante actividades de preservación digital, este estudio y análisis nos ayudará a plantear bajo un alcance descriptivo e inferencial de acciones al acercamiento sobre la situación de gestión de contenidos digitales, archivos y bibliotecas digitales ante las actividades formalizadas de preservación digital de acuerdo a NDSA-LDP.

Tabla 16. Niveles de Preservación Digital (NDSA-LDP Ver. 1).

	<b>Nivel 1</b> <b>(Proteja sus datos)</b>	<b>Nivel 2</b> <b>(Conozca sus datos)</b>	<b>Nivel 3</b> <b>(Controle sus datos)</b>	<b>Nivel 4</b> <b>(Repere sus datos)</b>
<b>Almacenamiento y localización geográfica</b>	Dos copias completas que no estén unidas.	Como mínimo tres copias completas.	Como mínimo una copia en una localización geográfica con una amenaza de desastres distinta.	Como mínimo tres copias en localizaciones geográficas con amenazas de desastres distintas.
	Para datos en soportes heterogéneos (discos ópticos, discos duros, etc.) mover el contenido a otro soporte dentro de su sistema de almacenamiento.	Como mínimo una copia en una localización geográfica distinta.	Controlar el proceso de obsolescencia de su(s) sistema(s) de almacenamiento y de sus soportes.	Disponga de un plan integral preparado para mantener los ficheros y los metadatos accesibles en los actuales soportes o sistemas.
		Documentar el(los) sistema(s) de almacenamiento y soportes de almacenamiento y lo que usted necesite para usarlos.		
<b>No alteración de ficheros e integridad de los datos</b>	Comprobar la integridad de los ficheros en el momento de la ingesta si sus valores han sido proporcionados junto con el contenido.	Comprobar la integridad de todas las ingestas.	Comprobar la integridad del contenido a intervalos regulares.	Comprobar la integridad de todo el contenido en respuesta a situaciones o actividades específicas.
	Crear la información de integridad si no fue proporcionada junto con el contenido.	Usar dispositivos con escritura bloqueada cuando se trabaje con soportes originales.	Mantener registros de la información de integridad; realizar auditoría bajo demanda.	Capacidad para reemplazar o reparar datos corrompidos.
		Comprobar virus en contenido de alto riesgo.	Capacidad para detectar datos corrompidos.	Asegúrese de que ninguna persona tiene acceso de escritura a todas las copias.
			Comprobar virus en todo el contenido.	
<b>Seguridad de la información</b>	Identificar quien ha leído, escrito, movido o eliminado la autorización a ficheros concretos.	Documentar las restricciones de acceso de los contenidos.	Mantener registros de quien ha realizado que acciones con los ficheros, incluyendo las acciones de borrado y preservación.	Realizar auditorías de los registros.
	Restringir quien tiene este tipo de autorizaciones a ficheros concretos.			
<b>Metadatos</b>	Inventario del contenido y de su localización en el almacenamiento.	Almacenar metadatos administrativos.	Almacenar metadatos estándar técnicos y descriptivos	Almacenar metadatos estándar de preservación.
	Asegurar la copia de seguridad separada del inventario.	Almacenar metadatos de las transformaciones y registrar las incidencias.		
<b>Formatos de ficheros</b>	Cuando usted puede participar en la creación de archivos digitales fomenta el uso de un conjunto limitado de formatos abiertos y conocidos de ficheros y de códecs	Disponer de un inventario de los formatos de ficheros usados.	Monitorear los problemas de obsolescencia de los formatos de ficheros.	Realizar migraciones de formatos, emulaciones o actividades similares si es necesario.

### 3.4.3 Aplicación de estudio NDSA-LDP

La aplicación del estudio se llevó a cabo siguiendo diferentes etapas que se pueden resumir en las siguientes: a) Proceso de indicadores y diseño de instrumento de recolección, b) definición y sistema de contactos clave IES-MX, c) contacto, envío y control de respuesta.

#### a) *Proceso de indicadores y diseño de instrumento de recolección*

Para la realización del estudio era necesario contar con un instrumento de recolección de datos que cubriera principalmente dos necesidades específicas. En primer lugar el ser capaz de integrar todos los indicadores de la metodología NDSA-LDP en formato de sistema de valoración múltiple de cada una de las actividades de categorías de preservación digital, ya que la metodología NDSA-LDP se compone de actividades, categorías y niveles - acciones como indicadores, por lo que para obtener una visión más completa a nivel cuantitativo y cualitativo de los indicadores a procesar en el instrumento fue necesario ordenarlos, contabilizarlos y definir sus componentes.

Tabla 17. *Componentes y cantidad de categorías NDSA-LDP.*

<b>Componente de NDSA-LDP</b>	<b>Cantidad de Categorías</b>
Categoría	5

Las *categorías* NDSA-LPD a razón de su organización las hemos concentrado bajo un solo término a modo práctico de referencia y se definen en almacenamiento, integridad, seguridad, metadatos y formatos. A los términos de las categorías les corresponden las siguientes cantidades de actividades:

Tabla 18. *Categorías y cantidad de actividades NDSA-LDP.*

<b>Nombre de Categoría</b>	<b>Cantidad de Actividades - Categoría</b>
Almacenamiento	9
Integridad	12
Seguridad	5
Metadatos	6
Formatos	4

Los *niveles - acciones* NDSA-LPD se integran por el nivel 1 - proteja sus datos, nivel 2 - conozca sus datos, nivel 3 - controle sus datos y nivel 4 - repare sus datos. A las cuales les corresponden las siguientes cantidades de niveles de acción y actividades:

Tabla 19. *Componentes y cantidad de Nivel - Acción NDSA-LDP.*

<b>Componente NDSA-LDP</b>	<b>Cantidad de Nivel - Acción</b>
Nivel - Acción	4

Los niveles - acciones definen las diferentes fases de actuación sobre actividades y categorías específicas de preservación digital. Estos son indicadores de gran importancia pues por un lado definen a partir de la concentración de actividades la fase específica de actuación de una institución. Por otro lado los niveles acciones pueden ponderar un alcance de contexto sobre la totalidad de actividades pues los niveles 4 y 3 dependen de los alcances de los niveles 1 y 2, por lo que puede realizarse una apreciación de alcance de actividades concatenadas y definir cuándo un nivel se ha implementado de forma total o parcial, por lo tanto afectando al resto de la cadena. De acuerdo a los autores los niveles son progresivos y los primeros niveles son prerrequisito de los últimos.

Tabla 20. Nivel - Acción y cantidad de actividades NDSA-LDP.

Nombre de Nivel - Acción	Cantidad de Actividades Nivel - Acción
Nivel 1 - Proteja sus datos (N1 PSD)	9
Nivel 2 - Conozca sus datos (N2 CSD)	10
Nivel 3 - Controle sus datos (N3 CSD)	9
Nivel 4 - Repare sus datos (N4 RSD)	8

Las *actividades* NDSA-LPD como componente activo y principal indicador a utilizar dentro del instrumento de estudio se contabilizaron y definieron de la siguiente manera:

Tabla 21. Actividades y cantidades totales NDSA-LDP.

Componente NDSA-LDP	Cantidad de Actividades
Actividades	36

En lo que refiere a actividades por nivel - acción y categorías de almacenamiento, integridad, seguridad, metadatos y formatos se distribuyen de la siguiente forma.

Tabla 22. Nivel - Acción, categorías, actividades y cantidades totales NDSA-LDP.

Nivel Acción	Actividades de Categoría Almacenamiento
N1 PSD	Dos copias completas que no estén unidas.
N2 CSD	Como mínimo tres copias completas.
N3 CSD	Como mínimo una copia en una localización geográfica con una amenaza de desastres distinta.
N4 RSD	Como mínimo tres copias en localizaciones geográficas con amenazas de desastres distintas.
N1 PSD	Para datos en soportes heterogéneos (discos ópticos, discos duros, etc.) mover el contenido a otro soporte dentro de su sistema de almacenamiento.
N2 CSD	Como mínimo una copia en una localización geográfica distinta.
N3 CSD	Monitorear (controlar) el proceso de obsolescencia de su(s) sistema(s) de almacenamiento y de sus soportes.

N4 RSD	Disponga de un plan integral preparado para mantener los ficheros y los metadatos accesibles en los actuales soportes o sistemas.
N2 CSD	Documentar el(los) sistema(s) de almacenamiento y soportes de almacenamiento y lo que usted necesite para usarlos.

<b>Nivel Acción</b>	<b>Actividades de Categoría Integridad</b>
N1 PSD	Comprobar la integridad de los ficheros en el momento de la ingesta si sus valores han sido proporcionados junto con el contenido.
N2 CSD	Comprobar la integridad de todas las ingestas.
N3 CSD	Comprobar la integridad del contenido a intervalos regulares.
N4 RSD	Comprobar la integridad de todo el contenido en respuesta a situaciones o actividades específicas.
N1 PSD	Crear la información de integridad si no fue proporcionada junto con el contenido.
N2 CSD	Usar dispositivos con escritura bloqueada cuando se trabaje con soportes originales.
N3 CSD	Mantener registros de la información de integridad; realizar auditoría bajo demanda.
N4 RSD	Capacidad para reemplazar o reparar datos corrompidos.
N2 CSD	Comprobar virus en contenido de alto riesgo.
N3 CSD	Capacidad para detectar datos corrompidos.
N4 RSD	Asegúrese de que ninguna persona tiene acceso de escritura a todas las copias.
N3 CSD	Comprobar virus en todo el contenido.

<b>Nivel Acción</b>	<b>Actividades de Categoría Seguridad</b>
N1 PSD	Identificar quién ha leído, escrito, movido o eliminado la autorización a ficheros concretos.
N2 CSD	Documentar las restricciones de acceso de los contenidos.
N3 CSD	Mantener registros de quien ha realizado que acciones con los ficheros, incluyendo las acciones de borrado y preservación.
N4 RSD	Realizar auditorías de los registros.
N1 PSD	Restringir quien tiene este tipo de autorizaciones a ficheros concretos.

<b>Nivel Acción</b>	<b>Actividades de Categoría Metadatos</b>
N1 PSD	Inventario del contenido y de su localización en el almacenamiento.
N2 CSD	Almacenar metadatos administrativos.
N3 CSD	Almacenar metadatos estándar técnicos y descriptivos
N4 RSD	Almacenar metadatos estándar de preservación.
N1 PSD	Asegurar la copia de seguridad separada del inventario.
N2 CSD	Almacenar metadatos de las transformaciones y registrar las incidencias.

<b>Nivel Acción</b>	<b>Actividades de Categoría Formatos</b>
N1 PSD	Cuando usted puede participar en la creación de archivos digitales fomente el uso de un conjunto limitado de formatos abiertos y conocidos de ficheros y de códecs
N2 CSD	Disponer de un inventario de los formatos de ficheros usados.
N3 CSD	Monitorear los problemas de obsolescencia de los formatos de ficheros.
N4 RSD	Realizar migraciones de formatos, emulaciones o actividades similares si es necesario.

Una vez contabilizados, definidos y organizados los indicadores NDSA-LDP el siguiente objetivo para la integración del instrumento de estudio fue la definición e integración de categorías, niveles - acciones y actividades. Para obtener un sistema de opción múltiple de actividades y que se pudiese sugerir la categoría de actividades de los diferentes niveles - acciones a modo general temático decidimos crear de una pregunta global que incluyera la categoría (almacenamiento, integridad, seguridad, metadatos, formatos) y todos los niveles de acción sin definir las descripciones del nivel - acción. Esta decisión se tomó con el fin de concretar la construcción de un instrumento de evaluación objetivo que no sugiriera la interpretación literal de que la elección de un mayor nivel es mejor opción que los niveles menores. Con este diseño se evita el sesgo de situar las respuestas directamente en una acción o nivel específico y de esta forma se puede contar con una valoración más limpia y centrada en actividades individuales por tema que describan mejor las prácticas individuales y que posteriormente se corresponda en el análisis de resultados a integrarse a una acción o nivel correspondiente. Esta decisión y variación corresponde también porque la metodología NDSA-LDP por sí misma representa un instrumento de autoevaluación.

Para tal motivo de búsqueda de objetividad en las respuestas se definieron las preguntas por categoría siguiendo la definición de la metodología e integrando la descripción completa de cada categoría. Debemos puntualizar que se definió utilizar el término de “archivo digital” de acuerdo con el Glosario de Preservación Archivista Digital (Voutsass y Barnard, 2014) como sinónimo de objeto de información dentro de una computadora, conjunto documental o colección (fondo documental). En México el concepto “archivo digital” mantiene las consideraciones tanto en singular como en plural. A partir de esta concepción se integraron las siguientes preguntas: 1. ¿Cómo se gestiona el almacenamiento y localización geográfica de sus archivos digitales?, 2. ¿Cómo se gestiona la no alteración de archivos e integridad de los datos en su archivo digital?, 3. ¿Cómo se gestiona la seguridad de la información en sus

archivos digitales?, 4. ¿Cómo se gestionan los metadatos en sus archivos digitales?, 5. ¿Cómo se gestionan los formatos de archivos en su archivo digital?. Complementario y previo a la pregunta temática se definieron 3 módulos generales para situar a los usuarios del estudio en un contexto de identificación, usos y herramientas del ámbito digital los cuales son: 1. Nombre de universidad (opcional), 2. ¿La IES cuenta con un repositorio digital, depósito digital, biblioteca digital o similar?, 3. ¿Que tipo de archivos digitales mantienen en su repositorio, depósito, biblioteca digital o similar?.

**1. Identificación de Universidad**  
Nombre de Universidad (opcional)

**2. ¿Su IES cuenta con un repositorio digital, depósito digital, biblioteca digital o similar?**  
 Sí  
 No  
 Otro (especifique)

**3. ¿Que tipo de archivos digitales mantienen en su repositorio, depósito, biblioteca digital o similar?**  
 Tesis académicas  
 Artículos de investigación  
 Libros digitales  
 Material educativo  
 Base de datos  
 Videos  
 Audios  
 Fotografías  
 Otro (especifique)

Figura 9. Diseño de preguntas generales de identificación, herramientas y usos.

Para el instrumento de evaluación se definió el uso de un sistema o herramienta informática que cumpliera con las necesidades de aplicación y recopilación de respuestas del estudio, es por ello que se decidió utilizar el sistema y herramienta de encuestas en línea Survey Monkey<sup>67</sup> con el objetivo de elevar la usabilidad del instrumento de estudio y poder emplear las prestaciones que la herramienta de creación de bloques de preguntas temáticas con sus respectivas actividades, además de contar con la facilidad de administración de formularios y de relacionar envíos personalizados por correo electrónico a los usuarios y contactos de interés. Finalmente una vez integrados los indicadores en el instrumento de estudio el sistema nos creo un enlace web para acceder al formulario y poder realizar el envío del mismo a los contactos clave de las IES.

<sup>67</sup> [www.surveymonkey.com](http://www.surveymonkey.com)

### ***b) Definición y sistema de contactos clave IES-MX***

Para la definición de los usuarios y contactos clave para el envío del estudio se integró un esquema o base de datos de contactos de IES basado en dos fuentes principales. En primer lugar se empleó el directorio de IES de la Secretaría de Educación Pública (SEP) que integra los sitios web de las mismas. En segundo lugar y complementario a esto se enriqueció el formato de contactos con el directorio proporcionado y publicado en REMERI<sup>68</sup> acentuando las especificaciones de IES de interés con los números telefónicos, nombres y emails de contactos clave. Creamos un directorio de IES y contactos clave como guía para realizar el envío del formulario de estudio en línea. Cabe señalar que en esta etapa también fue necesario definir perfiles o roles profesionales que nos ayudarán a mejorar la objetividad de contacto y de respuestas, por lo que se definieron dos perfiles de posibles contactos clave que situamos como perfil primario y perfil secundario de interés. Como perfil primario asignamos al rol profesional del responsable u otro cargo superior de biblioteca del IES. Como perfil secundario asignamos al rol profesional de encargado de sistemas informáticos o responsable de los sistemas de almacenamiento de IES. Esta decisión fue importante pues consideramos que sobre la base de hacer contacto telemático con los usuarios clave para nuestro estudio debíamos enfatizar en que las áreas correspondientes según las prácticas internacionales de gestión de colecciones digitales en repositorios, depósitos o bibliotecas digitales se realiza en las áreas principalmente de biblioteconomía con soporte de las áreas informáticas de las IES.

### ***c) Contacto, envío y control de respuesta.***

Como la finalidad del estudio es obtener indicadores del estado objetivo de IES-MX en prácticas de preservación digital decidimos elaborar un esquema que nos facilitara la obtención de respuestas objetivas del estudio. En relación al orden y definición de contactos clave definimos un esquema de contacto formal para realizar el envío del formulario del estudio por correo electrónico que consistió en los siguientes pasos:

1. Se realizaron llamadas telefónicas a los contactos clave IES con la finalidad de presentarnos personalmente, así como perfilar al contacto más adecuado para hacerle

---

<sup>68</sup> <http://www.remeri.org.mx/repositorios/>

llegar el formulario vía correo electrónico y obtener la respuesta más objetiva acorde la finalidad de estudio.

2. Se enviaron correos electrónicos personalizados con el nombre de los contactos clave y un mensaje breve que explica nuevamente por escrito la finalidad del estudio.
3. Se llevó a cabo una segunda tanda de llamadas para completar u orientar sobre dudas recibidas por correo electrónico acordes al formulario.
4. Una vez obtenidas las respuestas de los contactos clave se les hizo llegar un email de agradecimiento por su participación en el estudio.

Consideramos que al realizar este procedimiento se pudo disminuir el riesgo de subjetividad en el contacto, recepción de correos electrónicos y finalidad del estudio. Algunos contactos por sus ocupaciones o disposiciones de tomar las llamadas telefónicas nos enlazaron con otros perfiles que fueron los encargados de transmitir nuestra intención de contacto, por lo que recomendamos que al llevar a cabo cualquier tipo de estudios a distancia debemos enfatizar en la formalización bajo un esquema de acción que garantice la el conocimiento de causa, envió, recepción y respuestas.

El control del estudio se llevó a cabo bajo el siguiente lapso de tiempo que considera el contacto telefónico, envió y última recepción de respuestas.

Tabla 23. Línea de tiempo de aplicación de estudio.

Inicio de llamadas	Inicio de envío de formulario	Inicio de recepción respuestas	Fin y cierre de respuestas
04/05/2015	18/05/2015	26/05/2015	08/07/2015

### 3.4.4 Resultados y análisis de estudio NDSA-LDP en IES-MX

El objetivo de este estudio fue (como lo hemos dicho antes) obtener un nivel de alcance, estado de la cuestión y una primera aproximación a las prácticas y acciones que se llevan a cabo en las IES-MX con respecto a actividades de preservación digital con ayuda de la metodología NDSA-DPL. Para el análisis de los resultados del estudio definimos tres categorías de proyección del alcance de resultados: a nivel general, por subsistema y global del total de instituciones participantes. Estas categorías tienen la finalidad de mantener un límite objetivo que nos ayude a lograr una visión amplia a nivel de alcance descriptivo - inferencial, y obtener las prácticas en una panorámica general de situación y un nivel de preservación digital de acción que posteriormente se contraste con visiones de interés y objetivos de la metodología NDSA-DPL sobre su finalidad de uso y aplicación.

Para el orden de presentación de resultados se definieron tres segmentos que se componen en primer lugar por el segmento de identificación IES que incluye los resultados generales de IES por participación de subsistema, tipo de herramienta digital con la que cuenta (repositorio, depósito, BIDI, etc), más el tipo de material digital que gestiona en la herramienta digital (documentos, BD, videos, etc). El segundo segmento es el de nivel - acción de preservación digital que incluye las categorías y las actividades de preservación digital. Por último el segmento de estado de la cuestión de IES que incluye el análisis de uso de categorías, actividades en nivel - acción de preservación digital en contraste con visiones de interés y objetivos de las metodología NDSA-LDP.

#### a) *Identificación:*

#### IES por Subsistema y participación

IES Subsistema	No. IES	No. IES. Participantes	Porcentaje Participación
Federales	9	8	89 %
Estatales	34	28	82 %
IES Totales	43	36	86 %

Tabla 24. Cantidad de IES participantes del estudio.

En este sentido podemos afirmar que la participación de IES en el estudio se logró por encima de la media del universo total de IES-MX, lo que nos comunica una participación notable para el estudio del 86%. A partir del tamaño del universo de 43 IES y de la participación real de 36 IES en el estudio, se puede definir que este cuenta con un margen de error del  $\pm 7\%$  sobre un nivel de confianza del 95%.

### ¿La IES cuenta con un repositorio digital, depósito digital, biblioteca digital o similar?

IES General	No. Respuestas	Porcentaje
Si	35	97 %
No	0	0
Otro	1	3 %

IES Federales	No. Respuestas	Porcentaje
Si	7	87 %
No	0	0
Otro	1	13 %

IES Estatales	No. Respuestas	Porcentaje
Si	28	100 %
No	0	0
Otro	0	0

Tabla 25. Disposición de herramientas digitales en IES participantes del estudio.

En estos resultados se definió la confirmación de disposición de herramientas digital de gestión como repositorios, depósitos digitales o colecciones digitales donde el porcentaje mayoritario del universo cuenta con algún tipo de repositorio o BIDI. Matizando estos resultados la omisión de respuesta de un usuario y con la respuesta de una IES como “*otro*” al denominar a diferencia de un si o un no como respuesta y especificando que se debía aclarar el término repositorio.

**¿Que tipo de archivos digitales mantienen en su repositorio digital, depósito digital, biblioteca digital o similar?**

<b>IES Material</b>	<b>No. Respuestas</b>	<b>Porcentaje</b>
<b>Tesis académicas</b>	27	77 %
<b>Artículos de investigación</b>	21	60 %
<b>Libros digitales</b>	25	71 %
<b>Material educativo</b>	16	46 %
<b>Base de datos</b>	17	49 %
<b>Videos</b>	12	34 %
<b>Audios</b>	7	20 %
<b>Fotografías</b>	9	26 %
<b>Otro*</b>	5	14 %

Tabla 26. Tipo de herramientas digitales en IES participantes de estudio.

En los resultados encontramos el predominio de archivos digitales a tipología de documentos de tesis académicas, seguidas de los libros digitales y los artículos de investigación. Estos resultados dejan entrever las prioridades documentales de tipo académico y científico como documentos con fines de investigación como predominantes. Cabe señalar que a estos le siguen las bases de datos ya que se pudo comprobar según a una segunda ronda de llamadas telefónicas que una gran cantidad de IES cuentan con el recurso de bibliotecas digitales que alojan únicamente base de datos externas como servicio a su comunidad. El caso de los materiales audiovisuales refieren una minoría considerable en el estudio junto con materiales específicos como mapas, manuscritos, dibujos, entre otros.

A nivel de subsistema de IES los resultados refieren lo siguiente:

<b>Estatales</b>	<b>No. Respuestas</b>	<b>Porcentaje</b>
<b>Tesis académicas</b>	21	20 %
<b>Artículos de investigación</b>	15	15 %
<b>Libros digitales</b>	20	19 %
<b>Material educativo</b>	11	11 %
<b>Base de datos</b>	13	13 %
<b>Videos</b>	8	8 %
<b>Audios</b>	4	4 %
<b>Fotografías</b>	6	6 %
<b>Otro*</b>	5	5 %

<b>Federales</b>	<b>No. Respuestas</b>	<b>Porcentaje</b>
<b>Tesis académicas</b>	6	17 %
<b>Artículos de investigación</b>	6	17 %
<b>Libros digitales</b>	5	14 %
<b>Material educativo</b>	5	14 %
<b>Base de datos</b>	4	11 %
<b>Videos</b>	4	11 %
<b>Audios</b>	3	8 %
<b>Fotografías</b>	3	8 %
<b>Otro*</b>		

\*Otro: Patrimonio documental, mapas, revistas electrónicas, publicaciones periódicas, Manuscritos, dibujos.

Tabla 27. Tipo de herramientas digitales en IES por subsistema.

En este resultado encontramos como dato interesante que en las IES federales el material educativo mantiene una similitud con los materiales predominantes, así mismo los videos como recurso gestionado. En general por subsistema mantienen un resultado común al general con respecto al predominio de resguardo y gestión de materiales como tesis académicas, artículos de investigación y tesis digitales.

## ***b) Acciones - Nivel de Preservación Digital***

En este segmento analizamos los resultados de actividades por categoría seleccionadas por los usuarios acorde a la metodología NDSA-LDP donde se destacan resultados generales de participación global de IES en conjunto y posteriormente el tratamiento específico por subsistema de IES-MX.

### **1. Almacenamiento y localización geográfica**

<b>Clave NDSA LDP</b>	<b>Actividad</b>	<b>No. Respuestas</b>	<b>Porcentaje general</b>
<b>N1 PSD</b>	Dos copias completas que no estén unidas.	9	29 %
<b>N2 CSD</b>	Como mínimo tres copias completas.	0	0 %
<b>N3 CSD</b>	Como mínimo una copia en una localización geográfica con una amenaza de desastres distinta.	3	10 %
<b>N4 RSD</b>	Como mínimo tres copias en localizaciones geográficas con amenazas de desastres distintas.	1	3 %
<b>N1 PSD</b>	Para datos en soportes heterogéneos (discos ópticos, discos duros, etc.) mover el contenido a otro soporte dentro de su sistema de almacenamiento.	11	35 %
<b>N2 CSD</b>	Como mínimo una copia en una localización geográfica distinta.	12	39 %
<b>N3 CSD</b>	Monitorear (controlar) el proceso de obsolescencia de su(s) sistema(s) de almacenamiento y de sus soportes.	3	10 %
<b>N4 RSD</b>	Disponga de un plan integral preparado para mantener los ficheros y los metadatos accesibles en los actuales soportes o sistemas.	10	32 %
<b>N2 CSD</b>	Documentar el(los) sistema(s) de almacenamiento y soportes de almacenamiento y lo que usted necesite para usarlos.	9	29 %

Tabla 28. Resultados brutos de almacenamiento y localización geográfica en IES.

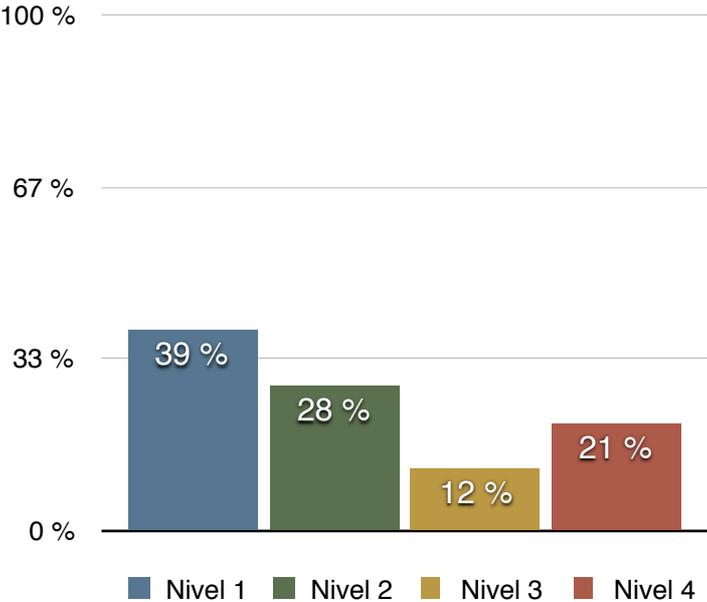
Llevando a cabo el cálculo en función de las respuestas obtenidas (restando las omitidas) del total en la categoría de *almacenamiento* los porcentajes extraídos corresponden a la selección concreta de actividades por participantes, lo que nos define un porcentaje real global del tipo de actividad y de la interpretación de la práctica en 31 respuestas. Lo que coloca un margen de omisión del 13,8%.

Al procesar los resultados de modo en el que las actividades individuales de la categoría de *almacenamiento* se contabilizaron por respuesta individual y se colocaron de modo proporcional a la suma representada en cada acción o nivel de preservación se pudo obtener un dato global del nivel - acción de preservación digital representado en la generalidad de la práctica de la siguiente forma:

Almacenamiento y localización geográfica	Porcentaje absoluto
Nivel 1 - Proteja sus datos	32 %
Nivel 2 - Conozca sus Datos	23 %
Nivel 3 - Controle sus datos	10 %
Nivel 4 - Repare sus datos	17 %

Tabla 29. Porcentajes brutos de almacenamiento y localización geográfica en IES.

Posterior a procesar los porcentajes brutos de la suma nominal de participación en cada actividad y situarlos en un nivel de acción, generamos una regla general de porcentajes totales por nivel para generar un porcentaje real global que sitio las actividades de las IES-MX en la categoría de *almacenamiento* en los diferentes niveles de acciones de preservación:



Gráfica 1. Niveles de preservación digital de IES-MX en categoría de almacenamiento y localización geográfica.

### ***Análisis de niveles de preservación digital de IES-MX en almacenamiento y localización geográfica***

Acorde a los *porcentajes globales* de las prácticas de *almacenamiento y localización geográfica en IES-MX*, estas se sitúan en su mayoría en el nivel 1 de proteja sus datos (39%), seguidas del nivel 2 de conozca sus datos (28%). La interpretación acorde a la metodología NDSA - LDP refiere que las actividades elegidas como prioritarias y en práctica de almacenamiento apuntan a un cumplimiento de nivel básico - medio sobre los estándares y prácticas de almacenamiento a largo plazo, ya que la mayoría de las IES-MX consultadas sólo disponen de una *copia de su archivo digital en una localización geográfica distinta o solo mueven el contenidos de un soporte de almacenamiento a otro*. Las buenas prácticas de la metodología recomiendan que cuando las actividades de organizaciones se encuentren en un nivel básico - medio el primer paso que deberá realizarse es asegurar el acceso en un futuro a los materiales creando una segunda copia (no unida y en una localización geográfica distinta) como paso inicial al almacenamiento de largo plazo. De acuerdo a los resultados, cuando se cuenta con una *copia en una locación geográfica distinta (nivel 2)* esta deberá de integrar un sistema de monitoreo que le permita obtener información sobre el estado de las copias idénticas en las otras localizaciones para asegurar su estado de integridad, autenticidad y acceso. En esta línea de prioridades resultantes encontramos que las actividades de *mover contenidos de archivos a otros soportes de almacenamiento (nivel 1)* para proteger sus datos juega un papel importante en la práctica general y debe tomarse en cuenta con alta prioridad para refrescar soportes de almacenamiento. Otros de los resultados por su parte apuntan a que un porcentaje de las IES-MX llevan a cabo actividades de *documentación de los sistemas de almacenamiento (nivel 2)* para conocer su ubicación en usos futuros. Estos porcentajes refieren a una práctica de nivel medio de implementación, ya que sugiere que el orden de documentar los temas de almacenamiento y ubicaciones además incrementa las posibilidades de integrar en un futuro nuevos sistemas de almacenamiento o ubicaciones geográficamente dispersas ordenadas. Hay que destacar que los resultados muestran que un bajo porcentaje de las IES dispone de un *plan integral (nivel 4)* para mantener ficheros y metadatos disponibles en los sistemas o soportes como medida de reparación de datos. Este rubro tendría que ser estimado más a fondo y de manera específica con cada una de las IES-MX firmantes del estudio pues esta actividad supondría un ejercicio de alto nivel el cual cubre casi la totalidad o formaliza las buenas prácticas de almacenamiento dado a que refiere a un documento formal

de compromiso y prácticas de almacenamiento que soportan estas actividades a modo planificado. Estos resultados colocan a las actividades generales de almacenamiento y localización geográfica en IES MX en el estado de cobertura media y básica de implementación formal de protocolos y buenas prácticas de actividades de almacenamiento y localización geográfica, lo que significa un estadio de posibilidad de revisar de manera formal las mismas y mejorar sus prácticas más allá del nivel básico y medio situadas en el estudio ante el almacenamiento con fines de preservación digital a largo plazo pues las actividades de almacenamiento y localización geográfica con más altas recomendaciones de buenas prácticas están situadas en su mayoría en el nivel 4 con una cobertura de tres copias en tres distintas localizaciones geográficas y de distintos riesgos de amenazas soportadas plenamente en un plan formal de preservación digital.

## 2. No alteración de archivos e integridad de los datos

Clave NDSA LDP	Actividad	No. Respuestas	Porcentaje
N1 PSD	Comprobar la integridad de los ficheros en el momento de la ingesta si sus valores han sido proporcionados junto con el contenido.	14	45 %
N2 CSD	Comprobar la integridad de todas las ingestas.	11	35 %
N3 CSD	Comprobar la integridad del contenido a intervalos regulares.	7	23 %
N4 RSD	Comprobar la integridad de todo el contenido en respuesta a situaciones o actividades específicas.	7	23 %
N1 PSD	Crear la información de integridad si no fue proporcionada junto con el contenido.	1	3 %
N2 CSD	Usar dispositivos con escritura bloqueada cuando se trabaje con soportes originales.	6	19 %
N3 CSD	Mantener registros de la información de integridad; realizar auditoría bajo demanda.	3	10 %
N4 RSD	Capacidad para reemplazar o reparar datos corrompidos.	12	39 %
N2 CSD	Comprobar virus en contenido de alto riesgo.	10	32 %
N3 CSD	Capacidad para detectar datos corrompidos.	7	23 %
N4 RSD	Asegúrese de que ninguna persona tiene acceso de escritura a todas las copias.	18	58 %
N3 CSD	Comprobar virus en todo el contenido.	12	39 %

Tabla 30. Resultados brutos de no alteración de archivos e integridad de datos en IES.

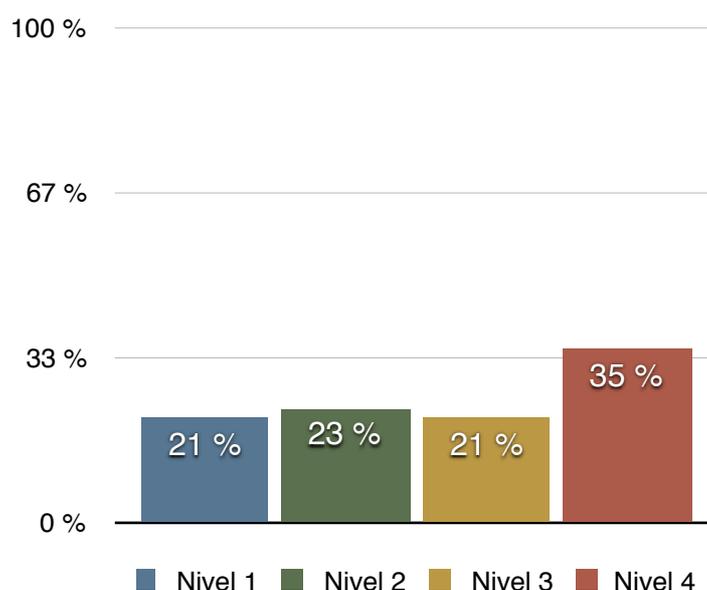
Llevando a cabo el cálculo según las respuestas obtenidas (restando las omitidas) del total en la categoría de *integridad* los porcentajes extraídos corresponden a la selección concreta de actividades por participantes. Esto nos define un porcentaje real global del tipo de actividad y de la interpretación de la práctica en 31 respuestas que coloca un margen de omisión del 13,8%.

Al procesar los resultados de modo en el que las actividades individuales de la categoría de *integridad* se contabilizaron por respuesta individual y se colocaron de modo proporcional a la suma representada en cada acción o nivel de preservación, de esta forma se pudo obtener un dato global del nivel - acción de preservación digital representado en la generalidad de la práctica de la siguiente forma:

No alteración e integridad de los datos	Porcentaje absoluto
Nivel 1 - Proteja sus datos	24 %
Nivel 2 - Conozca sus Datos	27 %
Nivel 3 - Controle sus datos	24 %
Nivel 4 - Repare sus datos	40 %

Tabla 31. Porcentajes brutos de no alteración de archivos e integridad de datos en IES.

Posterior a procesar los porcentajes brutos de la suma nominal de participación en cada actividad y situarlos en un nivel de acción generamos una regla general de porcentajes totales por nivel para generar un porcentaje real global que sitio las actividades de las IES-MX en la categoría de *integridad* en los diferentes niveles de acciones de preservación:



Gráfica 2. Niveles de preservación digital de IES-MX en categoría de no alteración de archivos e integridad de los datos

### ***Análisis de niveles de preservación digital de IES-MX en no alteración de archivos e integridad de los datos***

En lo que respecta a este rubro de *no alteración e integridad* existe una alta participación seleccionada por las IES-MX que les sitúa de manera media en los cuatro niveles específicos, los cuales van del nivel 1 de proteja sus datos (21%), nivel 2 de conozca sus datos (23%), nivel 3 de controle sus datos (21%) y del nivel 4 repare sus datos (35%). Estos resultados si los contrastamos con las prácticas de interpretación de la metodología NDSA-LDP, encontramos que el objetivo de esta categoría comprende un reto para muchas organizaciones pues integra una serie de 12 pasos que deberán cubrir las organizaciones para lograr la no alteración (fijación) e integridad de sus datos. El primer rubro a destacar en los resultados es el de *asegúrese de que ninguna persona tiene acceso de escritura a todas las copias (nivel 4)*, aunque en la traducción literal de la actividad sugiere un nivel de seguridad con base en evitar el acceso por peligro de alteración de copias no debe entenderse esta como una actividad intransigente o impositiva de protección bajo no dar acceso, ya que este no comprendería el fin de la metodología. Los resultados de las IES-MX otorgan a este rubro la más alta selección, es por ello que cabe hacer la reflexión de que si en realidad esta selección aplica como modo de imposición, bloqueo y control de acceso a usuarios del modo escritura, o si

realmente integra una robusta implicación de procesos comprendidos y sistematizados de confianza en la calidad y desempeño íntegro de los sistemas de almacenamiento pues esta actividad afirmará que se cuenta con un alto nivel seguridad y capacidad confianza de que los contenidos no se pueden alterar y se mantienen fijos e íntegros.

El siguiente rubro de alta selección es el de *comprobar la integridad de los ficheros en el momento de la ingesta si sus valores han sido proporcionados junto con el contenido (nivel 1)* que es el primer paso para comprobar que los contenidos que se pretenden preservar es lo que realmente se quiere preservar. La comprobación de integridad bajo instrumentos de encriptado como por ejemplo MD5 o SHA-1 que definan bajo un grado de seguridad la composición numérica de integridad de los contenidos ingestados al sistema de preservación digital para su acceso o seguridad. Además, si no han sido proporcionados se deberán proporcionar complementariamente. De acuerdo a los resultados solo una de las IES marco como actividad en práctica, lo que nos puede dejar ver que esta ya sea por la integración en sus sistemas de repositorios (dspace, eprints, etc.) u otra actividad, está cubierta en su nivel básico. Esta actividad podría ser complementada por herramientas como BagIt o Bagger según las recomendaciones.

El siguiente rubro a destacar en tercer lugar de selección general en IES-MX es el de *capacidad para reemplazar o reparar datos corrompidos (nivel 4)*, donde nuevamente debemos aclarar que existen actividades de orden individual o manual con las cuales se puede reparar o reemplazar datos y que no significan lo mismo a lo representado en la metodología y las prácticas internacionales, ya que a esta actividad debemos situar en el funcionamiento de un sistema completo de preservación digital que abarca una herramienta que supone por igual la mas alta práctica del rubro de no alteración e integridad de datos pues refiere a sistemas automatizados que en el caso de detectar anomalías bajo un procedimiento de monitoreo aplican la actuación necesaria. Nuevamente, esta actividad debe aclararse y orientarse bajo una *regla sistemática y no manual ocasional*.

El último rubro por tratar como práctica prioritarios en el estudio por las IES-MX fue el de *comprobar virus en contenido de alto riesgo (nivel 2)* y *comprobar virus en todo el contenido (nivel 3)*, este rubro debe ser entendido por igual a la regencia de herramientas de comprobación de virus y por tanto de la revisión de fijación o no alteración en todas las

ingestas. Habría que comprobar bajo sistemas específicos como se comprueba dicha actividad y definir si se trata de una actividad aislada de un sistema de antivirus tradicional previo al paso de la ingesta o de un sistema que incluye en la ingesta el procedimiento.

Hasta aquí podemos decir que la práctica de las IES-MX en el rubro de no alteración de archivos e integridad de datos, se ubica de manera general en un nivel medio superior al básico, siempre y cuando se compruebe la parte técnica *versus* la sistemática, lo cual podría otorgarnos un claro ejemplo de improvisación o actividad formal.

### 3. Seguridad de la información

Clave NDSA LDP	Actividad	No. Respuestas	Porcentaje
N1 PSD	Identificar quién ha leído, escrito, movido o eliminado la autorización a ficheros concretos.	16	53 %
N2 CSD	Documentar las restricciones de acceso de los contenidos.	17	57 %
N3 CSD	Mantener registros de quien ha realizado qué acciones con los ficheros, incluyendo las acciones de borrado y preservación.	13	43 %
N4 RSD	Realizar auditorías de los registros.	5	17 %
N1 PSD	Restringir quien tiene este tipo de autorizaciones a ficheros concretos.	19	63 %

Tabla 32. Resultados brutos de no alteración de archivos e integridad de datos en IES.

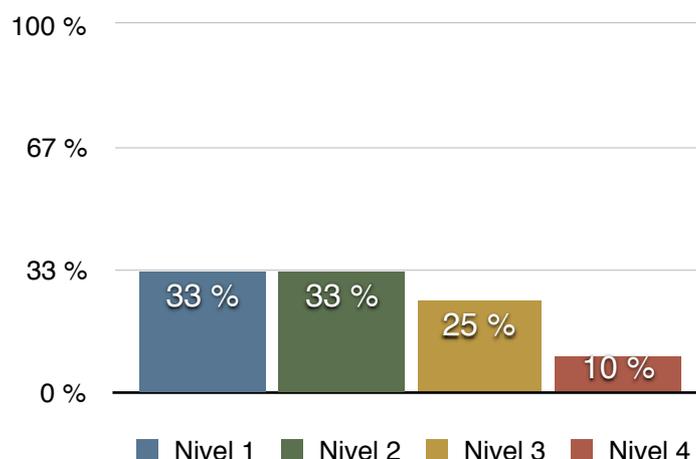
Llevando a cabo el cálculo basándose en las respuestas obtenidas (restando las omitidas) del total en la categoría de *seguridad* los porcentajes extraídos corresponden a la selección concreta de actividades por participantes, lo que nos define un porcentaje real global del tipo de actividad y de la interpretación de la práctica en 30 respuestas, lo que deja un margen de omisión del 17%.

Al procesar los resultados de modo en el que las actividades individuales de la categoría de *seguridad* se contabilizaron por respuesta individual y se colocaron de modo proporcional a la suma representada en cada acción o nivel de preservación se pudo obtener un dato global del nivel - acción de preservación digital representado en la generalidad de la práctica de la siguiente forma:

Seguridad de la información	Porcentaje absoluto
Nivel 1 - Proteja sus datos	58 %
Nivel 2 - Conozca sus Datos	57 %
Nivel 3 - Controle sus datos	43 %
Nivel 4 - Repare sus datos	17 %

Tabla 33. Porcentajes brutos de no alteración de archivos e integridad de datos en IES.

Posterior a procesar los porcentajes brutos de la suma nominal de participación en cada actividad y situarlos en un nivel de acción, generamos una regla general de porcentajes totales por nivel para generar un porcentaje real global que sitió las actividades de las IES-MX en la categoría de *seguridad* en los diferentes niveles de acciones de preservación:



Gráfica 3. Niveles de preservación digital de IES-MX en categoría de seguridad de la información

### ***Análisis de niveles de preservación digital de IES-MX en seguridad de información***

En la categoría de *seguridad* hay que destacar que acorde a los resultados finales procesados existe un predominio absoluto de la práctica referida a el nivel 1 de proteja sus datos (33%), seguida del nivel 2 de conozca sus datos (33%), dejando en niveles menores de práctica al nivel 3 de controle sus datos (25%) y nivel 4 de repare sus datos (10%) de manera continua.

Esto a nivel de análisis significa que las prácticas más destacadas de las IES-MX se reiteran en las actividades de *restringir quien tiene este tipo de autorizaciones a ficheros concretos (nivel 1)* que a orden de las recomendaciones de la metodología las IES-MX cumplen de manera rigurosa con el paso básico de identificar quién tiene acceso al contenido y que puede hacer este al mismo. En este sentido se habla de controlar accesos y nuevamente hay que anteponer que dicha actividad tiene su línea de imposición sobre limitar el acceso a un grupo específico, y por otro lado si en realidad existe un procedimiento escrito y en poder de las IES-MX que dictamine dichas acciones de restricción de acceso pues a este nivel básico será esencial contar con un procedimiento concreto a nivel de seguridad que garantice que los contenidos digitales en resguardo estén fuera del riesgo de ser borrados sin rastro alguno.

De manera continuada la segunda prioridad definida en los resultados fue la actividad de *documentar las restricciones de acceso de los contenidos (nivel 2)* que incrementa el grado de restricciones llevándolas de manera progresiva a un nivel documental que se complementa con las actividades del nivel 1. Un cuarto de las IES-MX, es decir 7 instituciones afirman que cuentan con este procedimiento en sus actividades de gestión de archivos digitales.

La tercera prioridad identificada en la mayoría de las IES-MX refiere a *identificar quién ha leído, escrito, movido o eliminado la autorización a ficheros concretos (nivel 1)*. Esta actividad forma parte por igual de las actividades básicas de las recomendaciones de seguridad, por lo que se correlaciona con las dos actividades anteriores ponderadas a modo íntegro donde la necesidad de contar con procedimientos previos que estipulan tanto niveles de acceso, tipologías y movimientos de usuarios a el nivel más básico. Podemos decir que las IES-MX del estudio se encuentran en un nivel de seguridad básico con fines de preservación digital, y en este sentido general de los resultados prioritarios los niveles de control y reparación de datos quedan muy por debajo de la práctica global de actividades de registros (logs) y auditorías de registros que dan la mejor orientación y referencia de buenas prácticas a nivel archivístico a las organizaciones según las recomendaciones de la metodología, por lo que son esenciales para brindar seguridad en la información y datos. Se puede decir que estas recomendaciones se definieron tomando en cuenta necesidades básicas por cubrir para ir adhiriendo nuevas capas de actividades y requisitos más avanzados que minimicen el riesgo inicial.

## 4. Metadatos

Clave NDSA LDP	Actividad	No. Respuestas	Porcentaje
N1 PSD	Inventario del contenido y de su localización en el almacenamiento.	11	34 %
N2 CSD	Almacenar metadatos administrativos.	12	38 %
N3 CSD	Almacenar metadatos estándar técnicos y descriptivos	22	69 %
N4 RSD	Almacenar metadatos estándar de preservación.	13	41 %
N1 PSD	Asegurar la copia de seguridad separada del inventario.	11	34 %
N2 CSD	Almacenar metadatos de las transformaciones y registrar las incidencias.	2	6 %

Tabla 34. Resultados brutos de no alteración de archivos e integridad de datos en IES.

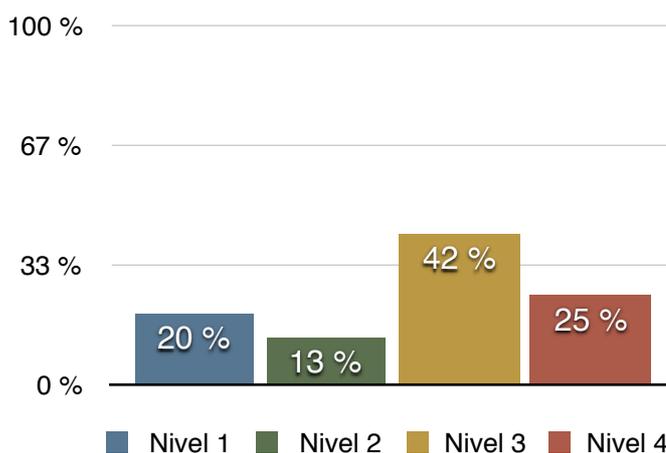
Llevando a cabo el cálculo a partir de las respuestas obtenidas (restando las omitidas) del total en la categoría de *metadatos*, los porcentajes extraídos corresponden a la selección concreta de actividades por participantes que define un porcentaje real global del tipo de actividad y de la interpretación de la práctica en 31 respuestas y que coloca un margen de omisión del 13,8%.

Al procesar los resultados de modo en el que las actividades individuales de la categoría de *metadatos* se contabilizaron por respuesta individual y se colocaron de modo proporcional a la suma representada en cada acción o nivel de preservación, de esta forma se pudo obtener un dato global del nivel - acción de preservación digital representado en la generalidad de la práctica de la siguiente forma:

Metadatos	Porcentaje absoluto
Nivel 1 - Proteja sus datos	34 %
Nivel 2 - Conozca sus Datos	22 %
Nivel 3 - Controle sus datos	69 %
Nivel 4 - Repare sus datos	41 %

Tabla 35. Porcentajes brutos de no alteración de archivos e integridad de datos en IES.

Posterior a procesar los porcentajes brutos de la suma nominal de participación en cada actividad y situarlos en un nivel de acción generamos una regla general de porcentajes totales por nivel para generar un porcentaje real global que sitio las actividades de las IES-MX en la categoría de *metadatos* en los diferentes niveles de acciones de preservación:



Gráfica 4. Niveles de preservación digital de IES-MX en categoría de metadatos

### ***Análisis de niveles de preservación digital de IES-MX en categoría de metadatos***

En los resultados las actividades de metadatos definidas como prioritarias por las IES-MX en el estudio dejan una posición liderada por las actividades del nivel 3 de controle sus datos ( 42%), seguidas del nivel 4 de repare sus datos (25%), nivel 1 proteja sus datos (20%) y del nivel 2 de conozca sus datos (13%). Las IES-MX de acuerdo con los resultados definen a la actividad de *almacenar metadatos estándar técnicos y descriptivos (nivel 3)* como una práctica constante en la que acorde con las recomendaciones de la metodología se debe de tomar en cuenta que las IES-MX firmantes de la práctica cuenten con un sistema automatizado computarizado y no manual para almacenar los metadatos estándar técnicos pues los metadatos descriptivos necesariamente deben introducirse manualmente. Con lo que respecta a la segunda elección como práctica de IES-MX se eligió a la actividad de *almacenar metadatos estándar de preservación (nivel 4)*. Cabe señalar que el nivel 4 y 3 solo incluye una actividad que corona esta medición de niveles de preservación digital en la categoría de

metadatos. En este caso hay que poner en contexto que aunque las recomendaciones de los niveles básicos obligan a la identificación del inventario y localización en los sistemas de almacenamiento del contenido, al ampliarse la forma y especialidad o fin de usos de metadatos en los contenidos digitales las recomendaciones se elevan al grado de complementarse en todos los niveles. Dicho esto, los resultados que apuntan a la práctica de almacenar metadatos estándar de preservación en IES-MX está sugiriendo de manera anticipada una práctica que mira a la preservación de contenidos pues esta práctica debe contar con un procedimiento que avale la finalidad de tratamiento de información (a nivel de metadatos) con fines de preservación real y no solo como una práctica aislada.

La siguiente actividad más valorada en el estudio es la de *almacenar metadatos administrativos (nivel2)* que definen el cuando y como ha sido creada la información, así como quien tiene acceso a ella que de manera relativa está relacionada de modo básico con la identificación de inventario y su acceso. Podemos decir que las prácticas de IES-MX con respecto a la categoría de metadatos reafirma un nivel balanceado que cumple los niveles básicos esenciales y que mantiene una participación continuada a los siguientes niveles mas por margen de especialización que de necesidad, lo que deja entrever un posible horizonte de futuras implementaciones evolutivas de metadatos administrativos, estructurales, técnicos, descriptivos y de preservación.

## 5. Formatos

Clave NDSA LDP	Actividad	No. Respuestas	Porcentaje
N1 PSD	Cuando usted puede participar en la creación de archivos digitales fomenta el uso de un conjunto limitado de formatos abiertos y conocidos de ficheros y de códecs	20	65 %
N2 CSD	Disponer de un inventario de los formatos de ficheros usados.	15	48 %
N3 CSD	Monitorear los problemas de obsolescencia de los formatos de ficheros.	10	32 %
N4 RSD	Realizar migraciones de formatos, emulaciones o actividades similares si es necesario.	12	39 %

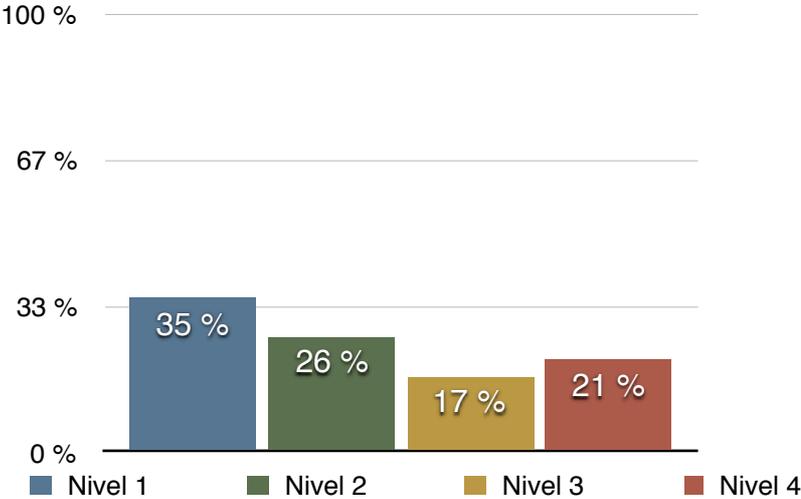
Tabla 36. Resultados brutos de no alteración de archivos e integridad de datos en IES.

Llevando a cabo el cálculo en función de las respuestas obtenidas (restando las omitidas) del total en la categoría de *formatos*, los porcentajes extraídos corresponden a la selección concreta de actividades por participantes que define un porcentaje real global del tipo de actividad y de la interpretación de la práctica en 32 respuestas que coloca un margen de omisión del 11%. Al procesar los resultados de modo en el que las actividades individuales de la categoría de *formatos* se contabilizaron por respuesta individual y se colocaron de modo proporcional a la suma representada en cada acción o nivel de preservación para obtener un dato global del nivel - acción de preservación digital representado en la generalidad de la práctica de la siguiente forma:

NDSA - Levels of Digital Preservation	Porcentaje absoluto
Nivel 1 - Proteja sus datos	65 %
Nivel 2 - Conozca sus Datos	48 %
Nivel 3 - Controle sus datos	32 %
Nivel 4 - Repare sus datos	39 %

Tabla 37. Porcentajes brutos de no alteración de archivos e integridad de datos en IES.

Posterior a procesar los porcentajes brutos de la suma nominal de participación en cada actividad y situarlos en un nivel de acción generamos una regla general de porcentajes totales por nivel para generar un porcentaje real global que sitio las actividades de las IES-MX en la categoría de *formatos* en los diferentes niveles de acciones de preservación:



Gráfica 5. Niveles de preservación digital de IES-MX en categoría de formatos

### ***Análisis de niveles de preservación digital de IES-MX en categoría de formatos***

En lo que refiere al rubro de la categoría de formatos los resultados de las IES-MX sitúan a las actividades de la categoría de manera mayoritaria en el nivel 1 de proteja sus datos (35%), seguido del nivel 2 de conozca sus datos (26%), continuando con el nivel 4 de repare sus datos (21%) y por último el nivel 3 de controle sus datos (17%). Las actividades más valoradas como prácticas se encabezan por la actividad de *cuando usted puede participar en la creación de archivos digitales fomentando el uso de un conjunto limitado de formatos abiertos y conocidos de ficheros y de códecs (nivel 1)*. Dicha actividad como nivel básico sugiere que las prácticas más recurrentes de las IES-MX se encaminan a la utilización de formatos de archivo conocidos y de ser posible formatos de tipo abierto (no propietario) bajo usos de series limitadas de los mismos acordadas como procedimiento. Este procedimiento es extensible de materiales nacidos digitalmente como a formatos que provienen de la digitalización, ya que acuerdos objetivos como este marcan una ruta productiva y viable como inicio de actividades de preservación digital. De acuerdo con la metodología los formatos que son recogidos de fuentes externas que cuentan ya con un formato establecido (como el caso de materiales en la web) no será recomendable como requisito básico el forzar el cambio a otros formatos.

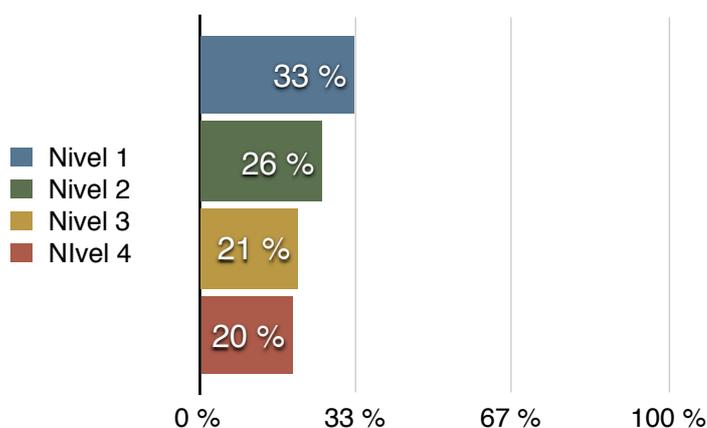
La siguiente práctica que resultó más valorada a usar es la de *disponer de un inventario de los formatos de ficheros usados (nivel 2)* la cual de manera sucesiva define la práctica de las IES-MX de documentar todos los formatos usados en su organización de los mismos, así como de mantener un sistema de inventario que anticipe cualquier actividad fuera de la normalización asignada. Otra práctica que llama la atención y que debemos colocar en contexto es la de *realizar migraciones de formatos, emulaciones o actividades similares si es necesario (nivel 4)* que define de manera estratégica a actividades de mantenimiento y preservación digital. Aquí debemos hacer distinciones de actividades manuales como efecto de mantenimiento correctivo y de actividades sistematizadas de mantenimiento preventivo que un sistema completo de preservación digital realiza según las prioridades y monitoreo de estado de obsolescencia y posibles anomalías en los archivos (nivel 3). Aunque la valoración de las IES-MX deja ver como práctica realizada la realización de migraciones en formatos, emulaciones o actividades similares debemos tener en cuenta los contextos reales de aplicación bajo el

tratamiento correctivo y preventivo sistemático. De modo general podemos decir que la categoría de formatos cumple con los niveles básicos y demuestra que se va por una ruta similar a la de la gestión de metadatos, y que para fines de preservación digital a largo plazo demuestra la validez de que es mejor hacer algo básico como principio de procedimientos a no hacer nada al respecto de la gestión de los archivos.

### ***c) Estado de cuestión global de IES-MX en niveles NDSA-LDP y visiones de contraste***

Es importante apuntar que este estudio como primera aproximación sobre la situación de niveles de preservación digital acordes a la metodología NDSA-LDP muestra una ponderación global porcentual de niveles mediante la suma de todos los resultados por categoría como una concentración del estado general de situación de IES-MX en la práctica de preservación digital.

### ***Análisis de posición global de IES-MX en niveles de preservación digital NDSA-LDP***



Gráfica 6. Porcentaje global IES-MX sobre situación y niveles de preservación digital

De manera global porcentual el nivel donde se concentran las acciones de las IES-MX ante actividades y recomendaciones de buenas prácticas en preservación digital de la metodología NDSA-DPL es en el *nivel 1 de proteja sus datos o protección de datos (33%)*. Esta posición global acentúa que las actividades y procedimientos de preservación digital de categorías como almacenamiento, integridad, seguridad, metadatos y formatos están orientadas como

práctica regular en el nivel elemental o básico de buenas prácticas acorde al estudio. La prioridad es mantener inventarios de protección de datos e información digital con énfasis en actividades mínimas de mantenimiento de soportes de almacenamiento, integración de información de integridad, identificación y restricciones de seguridad, inventarios de contenidos y su locación, así como de normalización y acuerdos de usos de formatos. De tal manera que mantiene una diferencia del 7% sobre las actividades del *nivel 2 de conozca sus datos o conocimiento de datos (26%)*, lo que significa que las actividades de conocimiento de los datos es el nivel que le sigue como práctica más referida y que prioriza sobre actividades de documentar sistemas de almacenamiento, ubicaciones, accesos restringidos, inventario de formatos, etc. Así tal se aleja por un 12% de la práctica del *nivel 3 controle sus datos (21%)* y del *nivel 3 de repare sus datos (20%)*.

Todas estas concentraciones y resultados globales de prácticas de situación sobre niveles resultantes en porcentajes de concentración y valoración práctica cuantitativa del estudio sobre niveles de preservación digital en el que se encuentran las IES-MX a modo global, nos otorga además una aproximación global cualitativa positiva orientada al conocimiento, desconocimiento o nivel de entendimiento sobre el manejo y dominio de las terminologías técnicas y conceptos teóricos sobre la gestión documental y preservación digital por parte del personal clave que respondió el estudio, como fue el caso de directores de bibliotecas, coordinadores técnicos, personal de soporte informático de bibliotecas y recursos en línea. A nivel cualitativo refiere que el trabajo realizado e importancia de los gestores y personal humano en las IES-MX a cargo de estas actividades específicas como piezas fundamentales que entienden y ejercen acciones sobre la tendencia de las buenas prácticas referidas en la metodología. Para ampliar y tener una panorámica más completa de estas referencias globales y poder construir una situación más específica y objetiva sobre el estado de cuestión de las IES-MX ante actividades de preservación digital, analizamos y contrastamos los resultados del estudio en bloques por subsistema IES-MX por objetivos de uso de la metodología NDSA-LDP, por nivel PEST y por último por infraestructura y tipología de IES-MX.

### ***Contraste de resultados globales de IES-MX y resultados por subsistema***

Para centrarnos en un objetivo ejercicio de contraste sobre prácticas resultantes en el estudio por subsistema IES-MX nos enfocaremos en una pregunta concreta: ¿Qué subsistema de IES-

MX de manera porcentual concentra mejores niveles de prácticas en preservación digital y cuáles son estos niveles destacados?. Para responder esta pregunta debemos situarnos de manera objetiva en el universo de IES federales y estatales que participaron en el estudio. Cabe señalar que en estos dos sub grupos existen diferencias básicas sobre las que hay que asentar el contraste para validar su aplicación y entendimiento. La primera diferencia se centra en el número de IES que las conforman pues debemos recordar que el número de IES federales corresponde a un número de ocho instituciones, mientras que las estatales corresponden a un total de veintiocho. En este punto debemos decir que un comparativo porcentual de actividades y situación de las mismas en un nivel específico corresponderá a una regla proporcional que nos arroje dicha cercanía para entender qué subsistema de IES-MX realiza una mayor cantidad de actividades por categoría y nivel de preservación digital acorde a la metodología.

Para definir la orientación de los esfuerzos por subsistema en los resultados del estudio nos apoyamos en la siguiente fórmula de proporcionalidad y concentración:

$$f\acute{o}rmula = \frac{nN^{100} \div nP}{nA}$$

*nN*: Número total de respuestas de actividades por nivel

*nP*: Número total de IES participantes por subsistema

*nA*: Número de actividades por nivel de categoría

La idea de integrar una fórmula de proporcionalidad de concentración se define para que en función de conocer el número total de respuestas sobre las actividades seleccionadas como prácticas activas de cada una de las IES y del subsistema que pertenecen, y proporcionado a la división del número total de IES con participación real del estudio se estima un porcentaje de participación real a su proporcionalidad sobre el número total de actividades por categoría. Esto nos ayudará a tener un porcentaje de alcance y relación comparativa entre subsistema y categoría de la metodología.

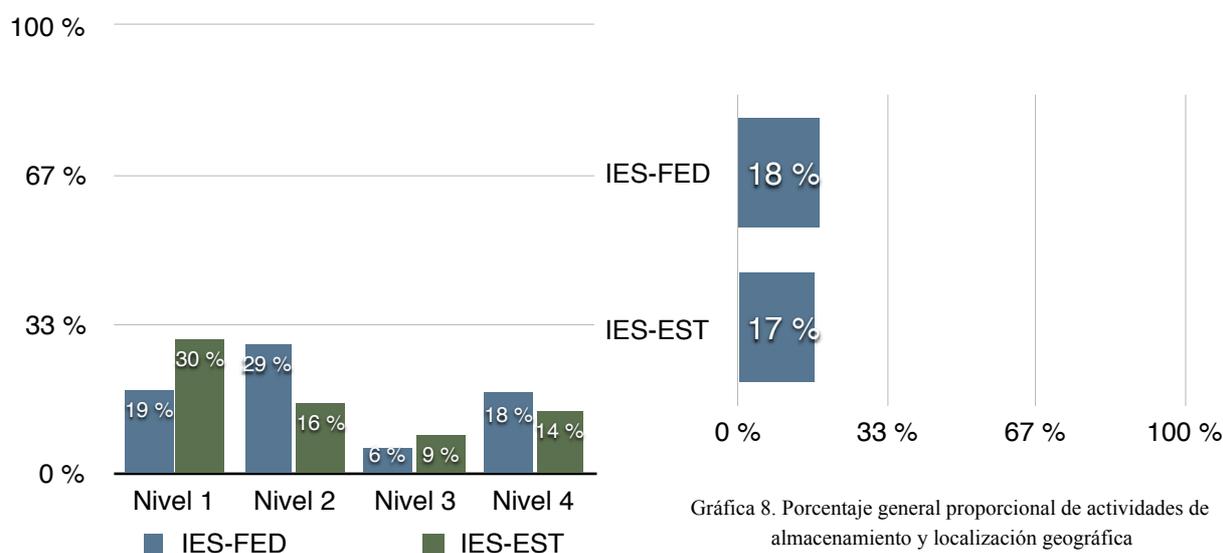
## 1. Almacenamiento y localización geográfica

Clave NDSA	Actividad	No.R. FED	No.R. EST
N1 PSD	Dos copias completas que no estén unidas.	1	8
N2 CSD	Como mínimo tres copias completas.	0	0
N3 CSD	Como mínimo una copia en una localización geográfica con una amenaza de desastres distinta.	1	2
N4 RSD	Como mínimo tres copias en localizaciones geográficas con amenazas de desastres distintas.	1	0
N1 PSD	Para datos en soportes heterogéneos (discos ópticos, discos duros, etc.) mover el contenido a otro soporte dentro de su sistema de almacenamiento.	2	9
N2 CSD	Como mínimo una copia en una localización geográfica distinta.	3	9
N3 CSD	Monitorear (controlar) el proceso de obsolescencia de su(s) sistema(s) de almacenamiento y de sus soportes.	0	3
N4 RSD	Disponga de un plan integral preparado para mantener los ficheros y los metadatos accesibles en los actuales soportes o sistemas.	2	8
N2 CSD	Documentar el(los) sistema(s) de almacenamiento y soportes de almacenamiento y lo que usted necesite para usarlos.	4	5

Tabla 38. Resultados globales por subsistema en almacenamiento y localización geográfica.

NDSA - Levels of Digital Preservation	No. Actividades	IES Federales	% / Absoluto	IES Estatales	% / Absoluto
Nivel 1 - Proteja sus datos	2	3	19 %	17	30 %
Nivel 2 - Conozca sus Datos	3	7	29 %	14	16 %
Nivel 3 - Controle sus datos	2	1	6 %	5	9 %
Nivel 4 - Repare sus datos	2	3	18 %	8	14 %

Tabla 39. Participación proporcionada por subsistema en almacenamiento y localización geográfica.



Gráfica 7. Porcentaje absoluto proporcional por niveles NDSA-DPL

Gráfica 8. Porcentaje general proporcional de actividades de almacenamiento y localización geográfica

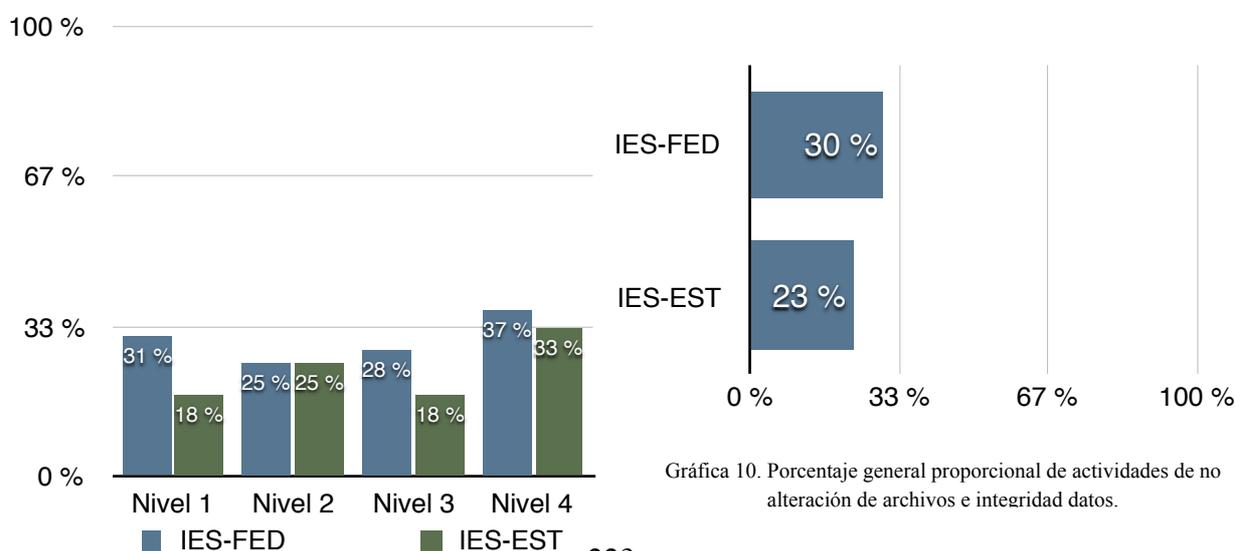
## 2. No alteración de archivos e integridad de los datos

Clave NDSA	Actividad	No.R. FED	No.R. EST
N1 PSD	Comprobar la integridad de los ficheros en el momento de la ingesta si sus valores han sido proporcionados junto con el contenido.	5	9
N2 CSD	Comprobar la integridad de todas las ingestas.	2	9
N3 CSD	Comprobar la integridad del contenido a intervalos regulares.	1	6
N4 RSD	Comprobar la integridad de todo el contenido en respuesta a situaciones o actividades específicas.	1	6
N1 PSD	Crear la información de integridad si no fue proporcionada junto con el contenido.	0	1
N2 CSD	Usar dispositivos con escritura bloqueada cuando se trabaje con soportes originales.	1	5
N3 CSD	Mantener registros de la información de integridad; realizar auditoría bajo demanda.	2	1
N4 RSD	Capacidad para reemplazar o reparar datos corrompidos.	5	7
N2 CSD	Comprobar virus en contenido de alto riesgo.	3	7
N3 CSD	Capacidad para detectar datos corrompidos.	2	5
N4 RSD	Asegúrese de que ninguna persona tiene acceso de escritura a todas las copias.	3	15
N3 CSD	Comprobar virus en todo el contenido.	4	8

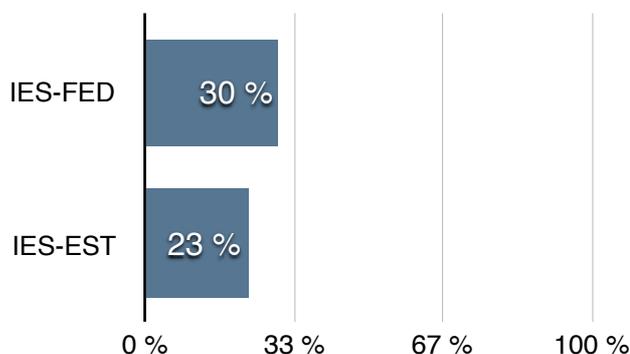
Tabla 40. Resultados globales por subsistema en no alteración de archivos e integridad de los datos.

NDSA - Levels of Digital Preservation	No Actividades	IES Federales	% / Absoluto	IES Estatales	% / Absoluto
Nivel 1 - Proteja sus datos	2	5	31 %	10	18 %
Nivel 2 - Conozca sus Datos	3	6	25 %	21	25 %
Nivel 3 - Controle sus datos	4	9	28 %	20	18 %
Nivel 4 - Repare sus datos	3	9	37 %	28	33 %

Tabla 41. Participación proporcionada por subsistema en no alteración de archivos e integridad de los datos.



Gráfica 9. Porcentaje absoluto proporcional por niveles NDSA.



Gráfica 10. Porcentaje general proporcional de actividades de no alteración de archivos e integridad de los datos.

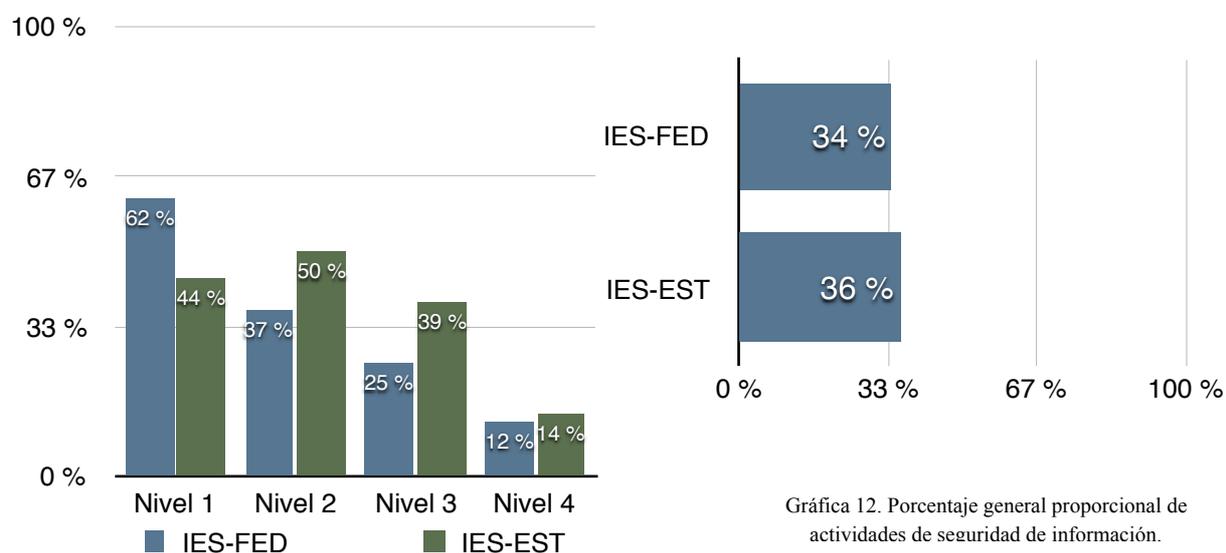
### 3. Seguridad de la información

Clave NDSA LDP	Actividad	No.R. FED	No.R. EST
N1 PSD	Identificar quién ha leído, escrito, movido o eliminado la autorización a ficheros concretos.	5	11
N2 CSD	Documentar las restricciones de acceso de los contenidos.	3	14
N3 CSD	Mantener registros de quien ha realizado qué acciones con los ficheros, incluyendo las acciones de borrado y preservación.	2	11
N4 RSD	Realizar auditorías de los registros.	1	4
N1 PSD	Restringir quien tiene este tipo de autorizaciones a ficheros concretos.	5	14

Tabla 42. Resultados globales por subsistema en seguridad de la información.

NDSA - Levels of Digital Preservation	No Actividades	IES Federales	% / Absoluto	IES Estatales	% / Absoluto
Nivel 1 - Proteja sus datos	2	10	62 %	25	44 %
Nivel 2 - Conozca sus Datos	1	3	37 %	14	50 %
Nivel 3 - Controle sus datos	1	2	25 %	11	39 %
Nivel 4 - Repare sus datos	1	1	12 %	4	14 %

Tabla 43. Participación proporcionada por subsistema en seguridad de la información.



Gráfica 11. Porcentaje absoluto proporcional por niveles NDSA-DPL.

Gráfica 12. Porcentaje general proporcional de actividades de seguridad de la información.

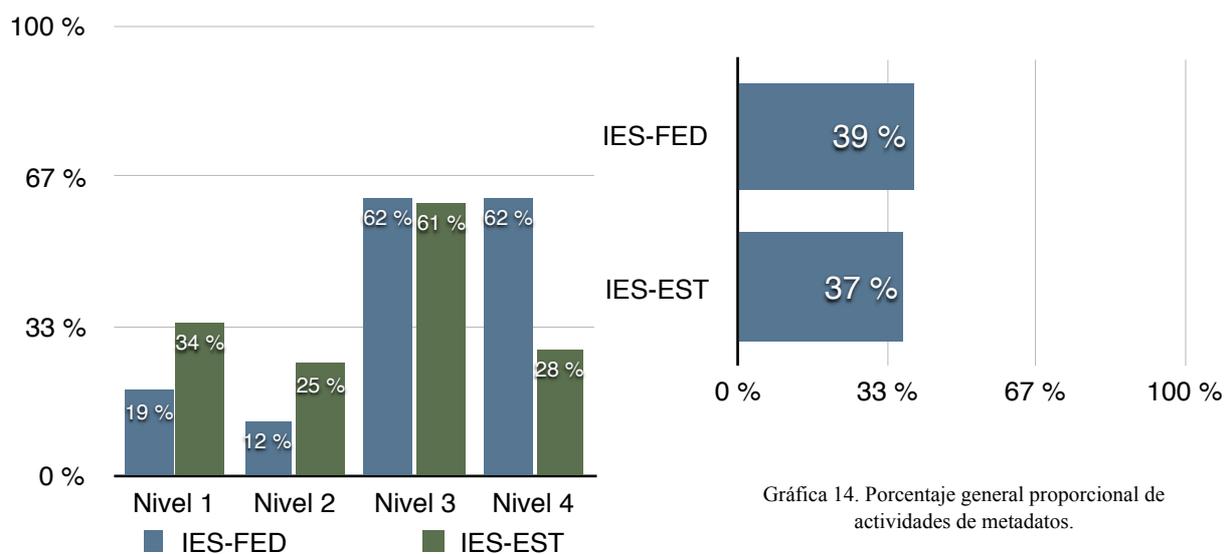
## 4. Metadatos

Clave NDSA LDP	Actividad	No.R. FED	No.R. EST
N1 PSD	Inventario del contenido y de su localización en el almacenamiento.	1	10
N2 CSD	Almacenar metadatos administrativos.	2	12
N3 CSD	Almacenar metadatos estándar técnicos y descriptivos	5	17
N4 RSD	Almacenar metadatos estándar de preservación.	5	8
N1 PSD	Asegurar la copia de seguridad separada del inventario.	2	9
N2 CSD	Almacenar metadatos de las transformaciones y registrar las incidencias.	0	2

Tabla 44. Resultados globales por subsistema en metadatos.

NDSA - Levels of Digital Preservation	No Actividades	IES Federales	% / Absoluto	IES Estatales	% / Absoluto
Nivel 1 - Proteja sus datos	2	3	19 %	19	34 %
Nivel 2 - Conozca sus Datos	2	2	12 %	14	25 %
Nivel 3 - Controle sus datos	1	5	62 %	17	61 %
Nivel 4 - Repare sus datos	1	5	62 %	8	28 %

Tabla 45. Participación proporcionada por subsistema en metadatos.



Gráfica 13. Porcentaje absoluto proporcional por niveles NDSA-DPL.

Gráfica 14. Porcentaje general proporcional de actividades de metadatos.

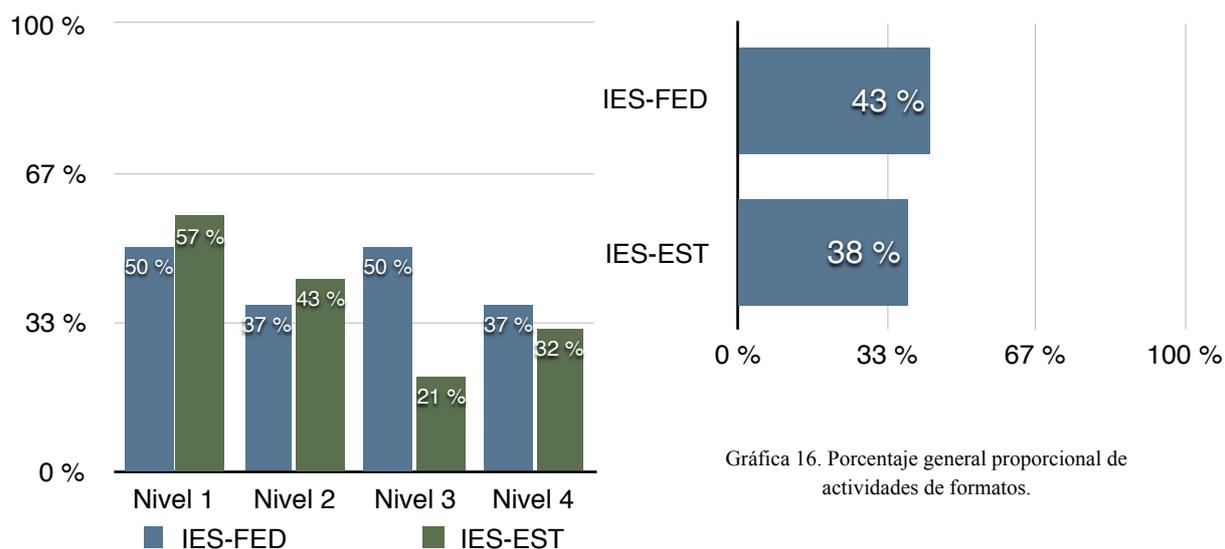
## 5. Formatos

Clave NDSA LDP	Actividad	No.R. FED	No.R. EST
<b>N1 PSD</b>	Cuando usted puede participar en la creación de archivos digitales fomente el uso de un conjunto limitado de formatos abiertos y conocidos de ficheros y de códecs	4	16
<b>N2 CSD</b>	Disponer de un inventario de los formatos de ficheros usados.	3	12
<b>N3 CSD</b>	Monitorear los problemas de obsolescencia de los formatos de ficheros.	4	6
<b>N4 RSD</b>	Realizar migraciones de formatos, emulaciones o actividades similares si es necesario.	3	9

Tabla 46. Resultados globales por subsistema en formatos.

NDSA - Levels of Digital Preservation	No Actividades	IES Federales	% / Absoluto	IES Estatales	% / Absoluto
Nivel 1 - Proteja sus datos	1	4	50 %	16	57 %
Nivel 2 - Conozca sus Datos	1	3	37 %	12	43 %
Nivel 3 - Controle sus datos	1	4	50 %	6	21 %
Nivel 4 - Repare sus datos	1	3	37 %	9	32 %

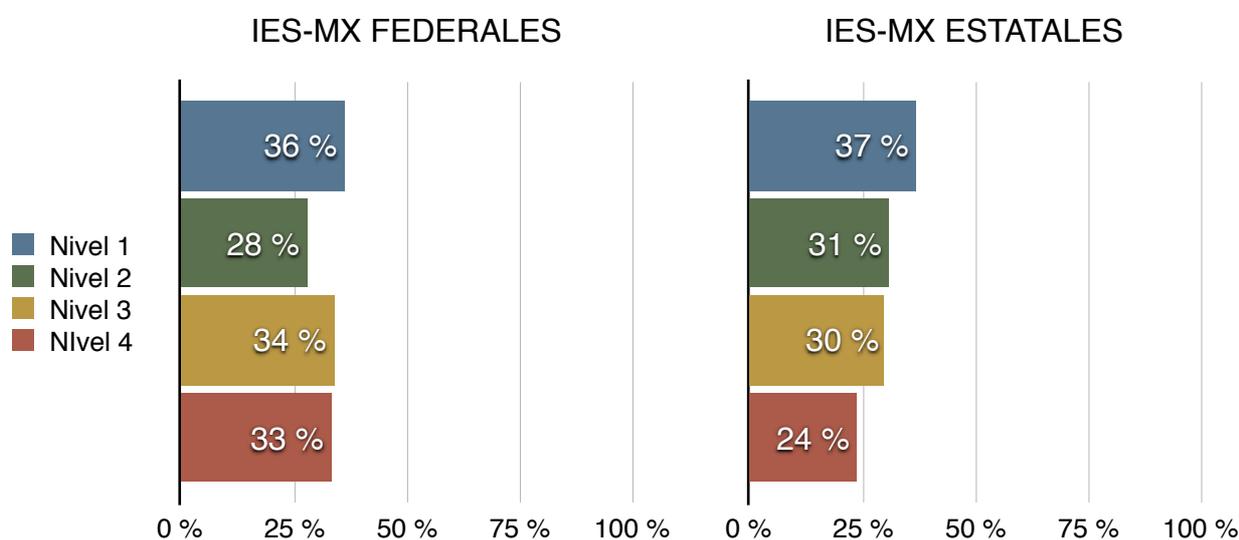
Tabla 47. Participación proporcionada por subsistema en metadatos.



Gráfica 15. Porcentaje absoluto proporcional por niveles NDSA-DPL.

Gráfica 16. Porcentaje general proporcional de actividades de formatos.

## Resultados globales y análisis de niveles de preservación digital por subsistema IES-MX



Gráfica 17. Porcentaje global por subsistema de IES-MX sobre situación ante niveles de preservación digital

De acuerdo con los resultados generales y globales, se puede observar un contraste comparativo de prácticas de concentración por subsistema de IES-MX estudiados podemos afirmar que existe una paridad y una diferencia notable entre las prácticas referidas por ambos grupos de IES-MX. Por un lado, las IES Federales al igual que las IES Estatales ponderan en primer lugar las actividades del nivel 1 de protección de datos que identifican en la metodología el nivel básico de preservación digital. De manera diferencial las IES Federales concentran sus actividades del nivel 1 en un 36%, muy cerca porcentualmente de las actividades del nivel 3 de control de datos con 34%, y de las actividades del nivel 4 de reparación de datos con un 33% que indica que además de actividades básicas de protección de datos la segunda prioridad es la de controlar los datos y la tercera en la reparación de los mismos. En comparación y acorde a la metodología esto significa que hay una aplicación asimétrica de buenas prácticas de actividades de preservación digital en las IES Federales pues se deja en última consideración a las actividades de conocimiento de datos del nivel 2 con un 28% de la concentración total y en la práctica significan en gran parte las actividades de documentar, inventariar y registrar movimientos que son necesarios para alcanzar los siguientes niveles de preservación digital de manera plena.

Por su parte las IES Estatales concentran su actividad de manera mayoritaria a los demás niveles en el nivel 1 de protección de datos con un 37%, lo que nos da una referencia de concentración notable en los niveles básicos de prácticas que a vista de la metodología otorga un sentido natural de orden y cumplimiento de los primeros niveles para continuar con los siguientes, ya que de manera relacionada las actividades de los niveles 2 y 3 comparten un 30% y 31% de concentración que muestran un equilibrio de las mismas. Hay que decir que este equilibrio no significa que mantienen de manera completa o en balance las actividades de la metodología y por consiguiente debemos observar que a nivel de porcentajes absolutos de concentración por categorías y número de actividades totales por subsistema. Tanto IES Federales como IES Estatales no cubren más del 50% de realización de actividades totales propuestas en la metodología, solo las IES Federales en la categoría de formatos alcanzan un 43% general de cumplimiento de las actividades propuestas como buenas prácticas. De acuerdo a la metodología NDSA-LDP se puede calificar a ambas IES-MX como “incompletas” con respecto a las actividades y prácticas de preservación digital en la gestión de contenidos y archivos digitales. Esto, otorga una visión primaria y aproximada de estado actual global de implementación o ejecución no formalizada de la preservación digital como un actividad de conocimiento y buenas prácticas. Estos resultados suponen un estadio informal por cubrir, y por ende de posibilidades de incorporación para formalizar el estado pleno de buenas prácticas en preservación digital para ambos grupos de IES-MX.

### **Contraste de resultados de IES-MX y los objetivos de NDSA-LDP**

Dentro de la metodología NDSA-LDP encontramos una serie de objetivos que se persiguen al hacer uso de la misma, y los cuales contrastamos con los resultados del estudio de niveles de preservación en IES-MX. El primer objetivo que se plantea es la *definición e identificación de consensos* existentes en la comunidad gestora de los recursos digitales con fines de preservación. En este objetivo las IES-MX al mantener una situación de concentración predominante de niveles básicos y medios de prácticas de preservación digital y con predominio en actividades de protección de datos y actividades de conocimiento de datos, nos refiere y afirma que los *consensos* específicos con fines de preservación por los que se podrían regir los resultados del estudio de IES-MX no existen a nivel interinstitucional pues no se muestra una paridad compartida por instituciones en actuaciones y prácticas en consenso. Se

ha podido identificar que una minoría de instituciones mantiene acciones pujantes sobre actividades individuales y categorías específicas pertenecientes a niveles superiores (3 y 4) que sugiere la necesidad y posibilidades de futuros planteamientos de consensos interinstitucionales bajo foros generales, que permitan informar, especializar, normalizar y delinear requerimientos mínimos, máximos y comunitarios de procedimientos aplicables en archivos, bibliotecas, fondos o repositorios de contenidos digitales y de su tratamiento de preservación a largo plazo.

Otro de los usos objetivos de la metodología es el de *educar y desarrollar directrices* para los creadores de contenidos y colaboradores con visión de preservación a largo plazo, ya que acorde a las buenas prácticas existe una relación directa productiva entre los esfuerzos y actividades realizados por los creadores y gestores de contenidos para fortalecer los vínculos e importancia de preservar a largo plazo el material creado y gestionado. El uso de formatos abiertos, codecs y el grado de descripción de contenido están relacionados con el nivel de servicio de preservación que se pretenda otorgar pues estas prácticas bien planteadas y aplicadas pueden convertirse en guías o directrices para otros usuarios o instituciones, lo que en situaciones y actividades de creación de marcos normativos (reformas a ley de bibliotecas, ley nacional de archivos, etc.) y lineamientos institucionales (lineamientos de creación de repositorios CONACYT) dan paso a integrar directrices seminales de preservación digital sólidas que complementan las acciones actuales en la gestión de contenidos digitales en México, y donde el papel de IES-MX pueda de manera progresiva integrar directrices específicas (técnicas y de organización) que garanticen un sistema útil de educación, cultura y dirección de preservación a largo plazo a ritmo de los crecimientos e integraciones de las actuales políticas nacionales en el ámbito digital institucional y educativo.

Un contraste interesante que refiere a nuestro estudio con el objetivo de uso de la metodología es para *validar la orientación de preservación dada localmente*. El estudio refleja que las acciones de una gran cantidad de IES-MX dejan entre ver en los resultados las prácticas puntuales de actividades específicas que llevan a cabo y las que no. Por otro lado también se acentúa el conocimiento o desconocimiento de procedimientos y posibilidades de las mismas ante acciones de preservación digital como posibilidades débiles y fuertes de las actividades desarrolladas para planear a nivel local procedimientos complementarios a las recomendaciones de buenas prácticas que otras organizaciones llevan a cabo de manera local,

y que en general representan las prácticas, manejo y forma de pensar de la comunidad activa internacional en el sector de la preservación digital pues estas han sido las creadoras y responsables de delinear las recomendaciones por niveles de práctica.

Otro de los usos de la metodología en contraste a los resultados de las IES-MX plantea el objetivo de *desarrollar requisitos para los proveedores de servicios de preservación de terceros*. Esto se traduce en los alcances de las prácticas desarrolladas e implementadas por las IES-MX ante las actividades, categorías y niveles reflejados en el estudio, donde la situación de prácticas se refleja en su mayoría en niveles básicos y medios que abren la posibilidad de evaluar, considerar y reflexionar sobre los servicios y prestaciones especializadas disponibles en el ámbito de la preservación digital que son ofrecidas por agentes externos, y que puedan ayudar a crear, implementar o consolidar sus actuales prácticas a posiciones de niveles de mejora y así posibilitar el alcance las recomendaciones para asegurar un completo esquema de buenas prácticas.

Orientando la metodología a un último objetivo de uso, esta se propone el *evaluar el cumplimiento de las mejores prácticas de preservación e identificar áreas clave para mejorar*. El estudio realizado bajo una orientación general del estado de las IES-MX ante actividades de preservación digital nos otorga diferentes marcos a nivel de categorías, actividades y niveles de actuación los cuales por lo tanto ayudan a evaluar qué prácticas se desarrollan mejor que otras. En nuestro estudio las prácticas de los niveles 3 y 4 se realizan en menor escala que las básicas y medias referidas en niveles 1 y 2, por lo que la metodología recoge y orienta estas acciones a grupos de mejores prácticas que deberán evaluarse de modo general, específico y por subsistema de IES-MX para así poder identificar áreas clave de las prácticas y su posibilidad de integración plena.

### **Contraste de resultados de IES-MX y enfoque PEST de México**

Ante el contexto y comportamiento político, económico y sociocultural por el que atraviesa México actualmente las actividades de preservación digital mantienen una agenda inocua, por lo que no podemos afirmar que las prácticas de preservación digital no son una prioridad reconocida o en su caso una práctica prioritaria no formalizada. Para tratar de reflejar el escaño sobre el que se sitúan las actividades en la agenda digital de la educación científica y

tecnológica en IES de México creamos un enfoque PEST (político, económico, socio cultural y tecnológico) que con ocupación sintética, nos refiera a una ventana puntal y nos posibilite la visión de situar las actividades de preservación digital de las IES-MX en un marco de afectación o influencia externa.

A nivel político y legal a partir del año 2012 México se ha envuelto en paquete de reformas estructurales como resultado del acuerdo político nacional denominado “*Pacto por México*” que tienen por objetivo el fortalecimiento del Estado Mexicano, la democratización de la economía, política y derechos sociales, así como la participación de la ciudadanía como eje rector en el diseño, ejecución y evaluación de las políticas públicas. Dentro del paquete de reformas de gran alcance y aplicación en su sector podemos listar a la reforma laboral, la reforma de amparos, el nuevo código de procedimientos legales, la reforma educativa, la reforma fiscal, la reforma de competencia económica, la reforma política, y las reformas al sector financiero, telecomunicaciones y energético (OECD, 2015). Aunque estas iniciativas han planteado diversas opiniones según sus efectos, implementación e interpretación de la mismas, a groso modo las reformas suponen una serie de avances estructurales que dejan a México como un país inmerso en posibilidades de cambios e innovación en áreas estratégicas para su implementación progresiva. A nivel de educación, ciencia y tecnología del sector de las IES-MX y en función de la combinación de reformas que tienen influencia y afectación directa como lo son la reforma educativa y de telecomunicaciones, nos confirma que las actividades de preservación digital a largo plazo podrían enfrentar una inclusión considerable puesto que el margen que abarcan las reformas estructurales dan continuidad indiscutible a el apoyo de la innovación, investigación y tecnología que bajo la agenda digital supone la potencialización y soporte de utilización de herramientas de tecnologías de la información, y que a su vez anticipa la necesidad de considerar políticas de la conservación y preservación a largo plazo del patrimonio, memoria y producción digital de las IES-MX y centros de investigación como principales actores del sector productivo de investigación y conocimiento en México.

En el escalón económico la inmersión de reformas en políticas estructurales específicas en diferentes sectores y en específico del sector financiero (banca) y fiscal (sistema tributario, recaudación y responsabilidad fiscal) dejan entrever un panorama de recuperación y reactivación macroeconómica de México (OECD, 2015). México mantiene en su agenda de

prioridad a la investigación, tecnología e innovación, lo cual se deja ver en el Plan de Desarrollo Nacional de México que plantea como una de sus prioridades la educación de calidad y su enfoque al desarrollo científico, tecnológico e innovación. De acuerdo a datos del portal de Transparencia Presupuestaria<sup>69</sup> y de la Secretaría de Hacienda y Crédito Público<sup>70</sup>, durante el año 2015 se aprobó un presupuesto económico para actividades de investigación científica y desarrollo de \$13,535.000.000 (trece mil quinientos treinta y cinco millones de pesos mexicanos m/n). Esto refiere a un indicador denominado en gasto en investigación científica y desarrollo experimental (GIDE) ejecutado por la Instituciones de Educación Superior (IES) respecto al producto interno bruto (PIB), donde el destino de recursos del presupuesto nacional ha sido enfocado a programas sectoriales de educación que tienen por objetivo impulsar la educación científica y tecnológica como elemento indispensable para la transformación de México en una sociedad del conocimiento. Este sector de manera nítida acoge a las IES-MX en actividades de apoyo a la investigación, desarrollo de plataformas tecnológicas de gestión documental, repositorios y preservación digital. El punto de nivel básico y medio en el que se sitúan las IES-MX ante actividades de preservación digital en nuestro estudio responde principalmente a que las actividades ponderadas van cubriendo las prioridades de atender la agenda actual, como por ejemplo la implementación de programas nacionales de integración y alineamiento de repositorios nacionales de IES-MX (CONACYT, 2014) así como la continua producción, especialización y formación de cuerpos de investigadores e investigación de orden de desarrollo experimental que mantienen los primeros sitios de atención.

En lo sociocultural, México actualmente se encuentra inmerso en diferentes visiones y problemáticas sociales (reputación política, inseguridad, combate a la pobreza, etc.) que dejan en afectación directa a actividades de investigación de IES-MX así como del desarrollo de proyectos de integración tecnológica, organización y preservación de contenidos digitales. A modo pesimista, en el ámbito sociocultural la visión responsable de salvaguardar la producción digital de los diversos factores de riesgo como lo son fenómenos naturales, efectos informáticos y acciones humanas debe ser analizada y considerada a detalle pues se puede encontrar que en ámbitos de inseguridad social y civil a causa de situaciones como guerras civiles o narcotráfico, los efectos de riesgo potencial de acciones humanas irresponsables

---

<sup>69</sup> <http://www.transparenciapresupuestaria.gob.mx>

<sup>70</sup> <https://www.sistemas.hacienda.gob.mx/ptpsed/datosProgramalLave.do?id=11E021>

podrían ocasionar junto con los desastres naturales un efecto informático de orden de riesgo de información prioritaria en diversos niveles y sectores, lo que de manera preponderante afecta a la actividad de conocimiento, educación e historia patrimonial cultural del país. Los pormenores deben ser también propinados para establecer una cultura preventiva de preservación digital. Debemos mencionar que ante la amplitud de instituciones federales y estatales de ámbitos culturales como el Archivo General de la Nación, Centro Nacional de las Artes, Instituto Nacional de las Bellas Artes, Consejo Nacional para la Cultura y las Artes, Fondo de Cultura Económica, Institutos de Cultura de los Estados y Universidades; las actividades de producción cultural entretajan y mantienen una relación constante sobre la preocupación de gestión y protección de su actividad reflejada como patrimonio digital, donde se mantienen ejes de importancia ante actividades de preservación de acervos culturales y memorias históricas. Como lo hemos mencionado el Archivo General de la Nación es uno de los principales propulsores de la Ley Nacional de Archivos, lo que a nivel México representa una pieza clave para la extensión a diversos sectores como el de la educación en sus diversos sistemas para de esta manera integrar la fundamentación y definición de protocolos de archivo y preservación de materiales físicos como digitales, lo que permea directamente en mejorar las prácticas referidas en el campo y posibilita a la apertura de consensos entre IES-MX, instituciones culturales y sociales para integrar directrices especializadas en el ámbito de la preservación a largo plazo bajo el seguimiento de buenas prácticas y recomendaciones necesarias.

A nivel tecnológico, encontramos a México en contraste con los resultados del estudio de IES-MX y su nivel de preservación digital que mantiene una influencia directa bajo la *reforma en telecomunicaciones (2013)*. Esta de manera general establece a cargo del ejecutivo federal la política de inclusión digital, donde se citan temas de infraestructura y conectividad, tecnologías de la información y la comunicación, habilidades digitales, programas de gobierno digital, gobierno y datos abiertos, sistemas y aplicaciones de contenidos digitales. Suman a ello el fomento a la inversión en telesalud, telemedicina, y en el expediente clínico electrónico, así como la garantía de la integración y libre acceso de la población a la sociedad de la información y conocimiento.

Todo esto significa que las iniciativas de las IES-MX ante actividades de preservación digital van por buen camino cubriendo los niveles básicos de implementación e investigación que se

traduce en posibilidades de adaptar y posicionar en un terreno prioritario a la preservación digital sobre el que deben entretorse las prioridades de ejecución necesarias que adviertan de la necesidad evolutiva natural y futura de la ejecución de las actuales políticas nacionales en el ámbito digital y telecomunicaciones, las cuales requieren por ende de especialización y gestión de actividades orientadas a la preservación digital, así como de la integración cultural de la responsabilidad de tratamientos efectivos de la producción digital como patrimonio y activo nacional.

### ***Contraste de resultados globales de IES-MX y tipología e infraestructura de subsistema***

De acuerdo con el análisis de resultados del estudio de prácticas de preservación digital de las IES-MX, cabe hacer una pregunta para poner en contexto los resultados a modo comparativo por subsistema: ¿Que coherencia podemos encontrar de manera general sobre las prácticas de preservación digital en IES-MX de acuerdo a los contenidos y materiales digitales que gestiona cada subsistema, su estructura de profesorado, investigadores, publicaciones, etc.? Las respuestas a estas interrogantes mantienen un calado específico sobre las actividades de investigación y biblioteconomía desarrolladas en el sector de las instituciones de educación superior y los sistemas de pertenencia de las mismas, así tal como en los apoyos económicos o presupuestos destinados a áreas de especialización en materia de gestión documental, repositorios digitales y preservación digital. Los presupuestos federales anuales otorgados a IES-MX se aplican en uso a diversas actividades generales distribuidas a nivel operativo, administrativo, académico y de investigación. En esta línea nuestro enfoque recae en las actividades de investigación, innovación académica, plataformas tecnológicas, sistemas de gestión y productos resultantes de trabajo de investigación de IES-MX, ya que estas representan de manera primaria las principales orientaciones para contrastar coherencias con el estudio. Para tener una panorámica general a nivel económico de los presupuestos otorgados a IES-MX por subsistema y acorde a nuestro universo de estudio, hemos elaborado con datos del documento “Descripción del proyecto de presupuesto de egresos de la federación 2015 para la educación superior” (ANUIES, 2014), un cuadro comparativo de los presupuestos aprobados para el periodo 2014 entre el subsistema de IES Federales y Estatales:

Tabla 48. Comparativo de presupuesto en IES federales y estatales de México.

Programa	Periodo	Presupuesto	Moneda	Numero de IES
IES Federales	2014	57,563.4	Millones de Pesos Mexicanos	9*
IES Estatales	2014	47,062.8	Millones de Pesos Mexicanos	57**

\*Extra: Centro de enseñanza técnico industrial

\*\* Universidades Públicas Estatales y Universidades Públicas Estatales con Apoyo Solidario

De manera cuantitativa si comparamos los presupuestos otorgados a cada subsistema de IES-MX y lo dividimos en el número de IES que le pertenecen a cada subsistema de manera equitativa (lo cual en práctica real no es aplicable), tendríamos como resultado hipotético una alta ventaja operativa sobre recursos económicos generales por parte de IES Federales con un presupuesto individual del 6,11% o \$6,395 millones de pesos mexicanos ante el porcentaje de presupuesto individual de las IES Estatales de un 1,7% o \$825 millones de pesos mexicanos por IES. Esta situación específica no podría atenderse ni sopesar únicamente de manera cuantitativa, ya que para comparar la coherencia de resultados de nuestro estudio con la de presupuestos debe considerarse el número total de alumnado, profesorado, investigadores, publicaciones, patentes o infraestructuras por IES. Como ejemplo, algunas IES federales como la UNAM<sup>71</sup> atienden aproximadamente a más de 300 mil alumnos y una plantilla de 30 mil profesores, esto en comparación con algunas IES federales como el COLMEX<sup>72</sup> que atiende a cerca de 400 alumnos y a 200 profesores. A razón de comparativa nos dejaría una situación de probabilidad y justificación a nivel del presupuesto otorgado, operado y destinado a acciones específicas como las de gestión digital en su caso. En porcentajes globales de nuestro estudio y contrastando ante un nivel de ocupación y atención de prácticas de actividades de gestión y preservación digital de repositorios, archivos o BIDs en IES-MX, las actividades relacionadas con almacenamiento, integridad, seguridad, metadatos y formatos de su actividad de gestión digital de manera porcentual entre instituciones, no muestra una coherencia general comparada por subsistema de IES-MX, pues en los resultados de nuestro estudio las IES-Federales sólo aventajan con diferencias mínimas de manera porcentual

<sup>71</sup> Profesorado y Alumnos de UNAM: [https://es.wikipedia.org/wiki/Universidad\\_Nacional\\_Autónoma\\_de\\_México](https://es.wikipedia.org/wiki/Universidad_Nacional_Autónoma_de_México)

<sup>72</sup> Profesorado y Alumnos de COLMEX: [https://es.wikipedia.org/wiki/El\\_Colegio\\_de\\_México](https://es.wikipedia.org/wiki/El_Colegio_de_México)

general sobre la realización de actividades en cuatro categorías, y mostrando una desventaja en la categoría de *seguridad de información* ante las IES-Estatales que muestran una ventaja de realización y concentración de las mismas.

Otro dato interesante que hay que tomar en cuenta para tener una aproximación de realización y prioridad, es el de que algunas IES-MX no cuentan en su composición estructural y orgánica un referente que afirme que dentro de sus actividades prioritarias tiene la gestión y actividades de centros de investigación. Llama a la atención la definición de la misión general que tienen por subsistema en la web de la Secretaría de Educación Pública<sup>73</sup>, donde a las IES Federales se les describe como: “*Las instituciones que conforman este subsistema realizan, además de las funciones de docencia, un amplio espectro de programas y proyectos de investigación (generación y aplicación innovadora del conocimiento), y de extensión y difusión de la cultura.*”; por otro lado, las IES Estatales son definidas como: “*Las Universidades Públicas Estatales, son instituciones de Educación Superior creadas por decreto de los congresos locales, bajo la figura jurídica de organismos públicos descentralizados. Estas instituciones estatales desarrollan las funciones de docencia, generación y aplicación innovadora del conocimiento, así como de extensión y difusión de la cultura.*”. Si nos centramos de manera literal en su definición general, dejamos entrevisto que las IES Federales son las únicas que tienen una misión y estructura institucional que además de la docencia se ocupa de manera prioritaria a la creación y aplicación de proyectos de investigación. Por su parte en las IES Estatales según la definición, las prioridades y estructuras están más enfocadas a la docencia que a la investigación. Por tanto, cabe reflexionar si las IES Estatales se ciñen a la letra de su definición sobre prioridad a la docencia y en menor escala a la investigación y dejan como asignatura pendiente a todas las actividades que con ella circulan como las plataformas y sistemas de gestión digital. De acuerdo con los resultados de nuestro estudio NDSA-LDP IES-MX, encontramos una coherencia a la atención derivada de sus prácticas de gestión digital de IES estatales, las cuales se encuentran como prioridad los niveles básicos (nivel 1) al igual que las IES federales. Por el contrario, las federales deberían estar muy por encima en la cobertura de recursos prioritarios de gestión de prácticas activas y la disposición de los mismos para actividades relativas a la investigación que les permita depositarla, difundirla y preservarla si fuese el caso literal de su misión. Las IES Federales tienen a nivel de presupuesto federal, prioridades y organización sistemática, la

---

<sup>73</sup> <http://www.ses.sep.gob.mx/instituciones-de-educacion-superior>

misión de mantener el liderazgo en el sector de actividades de investigación y subyacentes. Bajo un este índice debemos mencionar nuevamente que en el sector de IES Federales y Estatales existen diferentes esfuerzos no especializados en preservación digital, pero si en el ámbito de gestión de contenidos digitales de IES y su producción de investigación científica y académica que combinan la labor de IES del sector privado y público. El caso de REMERI (Red Mexicana de Repositorios Institucionales), LATINDEX o REDALYC, que deben considerarse como un ejemplo activo sobre el esfuerzo de integración, consenso y definición de directrices que apuntan a las IES-MX sobre un panorama de movimiento natural y especialización en la agenda de gestión digital que dictamina un panorama a escala de las prioridades y sus posibilidades de la composición nacional de un sistema de preservación digital.



**Capítulo 4 - Modelo de Preservación Digital Distribuida  
IES México (PDDIM)**



## **4.1 Definición para propuesta de modelo de preservación digital distribuida IES-MX (PDDIM)**

Al haber concretado el marco referencial de la preservación digital en su estrategia distribuida, el análisis funcional, organizacional y del nivel actual de IES-MX ante actividades de preservación digital; contamos con los enfoques, conceptos y argumentos necesarios para la integración de la propuesta del modelo de preservación digital distribuida para IES-MX.

Para la integración holística del modelo PDDIM y de acuerdo a los parámetros de nuestra investigación en los capítulos antecedentes, realizamos una serie de abstracciones concentradas con la finalidad de obtener conceptos funcionales y prácticas clave que argumenten la definición de nuestro modelo PDDIM a partir de los siguientes aspectos. a) Marco general de referencia conceptual de la preservación digital y su estrategia distribuida, b) Elementos conceptuales del modelo de referencia, modelo variable y sistemas de preservación digital distribuida, c) Composición y organización funcional en casos de modelos de preservación digital distribuida, y d) Visión de necesidades en IES-MX, compatibilidad de buenas prácticas de casos de modelos PDD, y lineamientos institucionales en México con relación a las IES-MX ante la estrategia de PDD.

## **4.2 Marco general de referencia conceptual para modelo PDDIM**

La preservación digital distribuida (PDD) más allá de los marcos teóricos y prácticos puede ser interpretada a nivel funcional desde varios enfoques de pertinencia. Por un lado es considerada como una necesidad complementaria de la estrategia de preservación digital centralizada, donde la distribuida prioriza en el almacenamiento con dispersión geográfica y coordinada de réplicas de contenidos digitales en diversos nodos de almacenamiento independientes pero enlazados entre sí por una figura intermedia de coordinación. Otra acepción la interpreta a un nivel de estrategia complementaria de organización que asume los límites de la preservación digital centralizada para potenciar y asegurar mediante la distribución de réplicas dispersas geográficamente una mayor cantidad de copias idénticas de

archivos digitales en diferentes puntos o nodos de almacenamiento. Para ambas, es necesaria la ayuda de la coordinación, la colaboración interinstitucional y una capa intermedia de organización que administre los compromisos de responsabilidad, independencia y relaciones de cada nodo institucional. Una vez establecido lo anterior, la implementación de herramientas informáticas para la gestión técnica de preservación digital distribuida (monitoreo de integridad, replicación y reparación de archivos) podrán funcionar de forma armónica y organizada entre las instituciones para alcanzar de manera plena y sostenida el acceso a largo plazo de los contenidos preservados. La situación sobre la que se anticipan ambas premisas es la de asegurar de manera organizada y colaborativa la disposición de copias idénticas e íntegras en diversos sitios con parámetros continuos de monitoreo y reparación, por lo que aunque se pudiera referir a la preservación digital distribuida como una estrategia complementaria de organización a las medidas de precaución técnicas, debemos tomar en cuenta que los modelos de preservación digital distribuida integran elementos que los diferencian de estrategias puramente técnicas, y estos radican principalmente en el proceso de coordinación y los acuerdos de colaboración interinstitucional que definen la participación humana dentro de procesos y servicios técnicos en los que tienen cabida roles, infraestructuras, grupos de interés, ámbitos de alcance geográfico, etc. Para integrar una visión práctica y funcional que nos ayude a construir un modelo PDDIM, sintetizamos a modo de esquemas y definiciones los conceptos de referencia funcional y sistemas, modelos de organización y prácticas referentes a la preservación digital distribuida.

### **4.3 Elementos conceptuales del modelo de referencia, variable y de sistemas PDD para PDDIM.**

Para la concepción y construcción de un modelo de preservación digital distribuida es necesario identificar de forma objetiva a los elementos técnicos, estrategias y entidades funcionales específicas que participan en la organización y coordinación de sus actuaciones dentro en un ecosistema cerrado con fines comunes de preservación digital en modo distribuido. Partiendo de esta idea, sintetizamos a modo conceptual y gráfico las bases elementales de los modelos de referencia, estrategias y herramientas de preservación digital que definen la agenda temática y marcos de acción. También utilizamos la abstracción de los

análisis de contraste funcional de elementos técnicos, herramientas y sistemas especializados en preservación digital distribuida, y de la ascendencia de buenas prácticas organizacionales para entretener de manera sintetizada los elementos y actores comunes.

#### **4.3.1 Modelo de referencia OAIS en modelos de PDD**

El modelo de referencia OAIS o ISO 14721:2012 para procedimientos de preservación digital, además de identificar los actores indispensables en un proceso de ingestión, creación de unidades de archivo, gestión y administración, así como de diseminación o acceso a contenidos custodiados; encuentra en su referencia el ejercicio de la integración simplificada y concreta de actores indispensables para la correcta gestión de contenidos digitales a largo plazo. Debemos entender que el modelo de referencia sirve de modelo estándar para identificar las entidades que ejecutan acciones, y que estas deberán integrar un procedimiento técnico que identifique las funciones necesarias para realizarlas. Investigadores especializados en la preservación digital distribuida como Zierau y Schultz (2011), en *“Creating a Framework for Applying OAIS to Distributed Digital Preservation”* anticiparon que a partir de la definición de un marco de trabajo de componentes que incluye la definición de conceptos y terminología adecuada que describa formalmente el ejercicio de preservación distribuida sería posible la integración de roles de organización interinstitucional y relaciones al modelo de referencia OAIS en un ambiente distribuido de repositorios digitales. Se debe pensar y definir cuales son las entidades y relaciones que necesitamos integrar en un modelo de trabajo de PDD, y que capacidades funcionales, prestaciones y necesidades deberá cubrir.

Hoy en día el modelo OAIS como regla referencial está integrado en la mayoría de los modelos de trabajo de sistemas de preservación digital centralizada, los cuales anticipan la integración completa del modelo de referencia OAIS en sus operaciones. A nivel de preservación distribuida el modelo de referencia se sigue equilibrando e integrando a las diversas necesidades de adaptar una figura de coordinación intermedia como base para realizar la conexión de diferentes nodos, entidades o repositorios que realizan actividades basadas en el modelo de referencia, adaptándolo a lo que el grupo de expertos define en su terminología y la definición de la preservación digital distribuida como aquella que hace uso de la coordinación, la independencia y la replicación.

### **4.3.2 Modelo de distribución variable en preservación digital**

Para poner a modo visual y conceptual el amplio marco de posibilidades de actuación que ofrece la estrategia de preservación digital distribuida haciendo uso de diversas variables de integración y organización operativa dentro de una red, revisamos el modelo de distribución variable para preservación digital propuesto por Zierau (2011) con la finalidad de entender las diferentes posibilidades de roles, funciones y servicios en un entorno distribuido y asimilar las combinaciones de entidades y elementos que le integran para connotar las necesidades de organización como parte de las relaciones de distintos marcos interinstitucional u organizacional.

En el modelo de distribución para preservación digital de Eld Zierau, los elementos que le conforman se explican con las siguientes referencias. Las figuras de casas representan a las organizaciones (instituciones), los rectángulos podrían representar unidades de almacenaje y los círculos a servicios y procesos de preservación digital. Los colores de los rectángulos y los círculos mantienen una variación dependiendo de su color, en el caso de los colores blancos representan a una solución distribuida compartida, el caso de los de color oscuro representan a soluciones plenamente internas de las organizaciones. Los rectángulos como unidades de almacenaje pueden contener cada una, una copia de datos. La autora sugiere que en el caso de la casa “A”, esta es representada como sólo consumidora de la solución de preservación sin tener copias de datos en sí misma, es decir mantiene un servicio externo. El caso de la casa “D”, esta solo mantiene un rol de proveedor de unidad de almacenamiento compartido para mantener una copia de los datos. En el caso de la casa “F”, esta funge con el rol de proveedor de un servicio de procesamiento coordinado con una solución distribuida (Zierau, 2011).

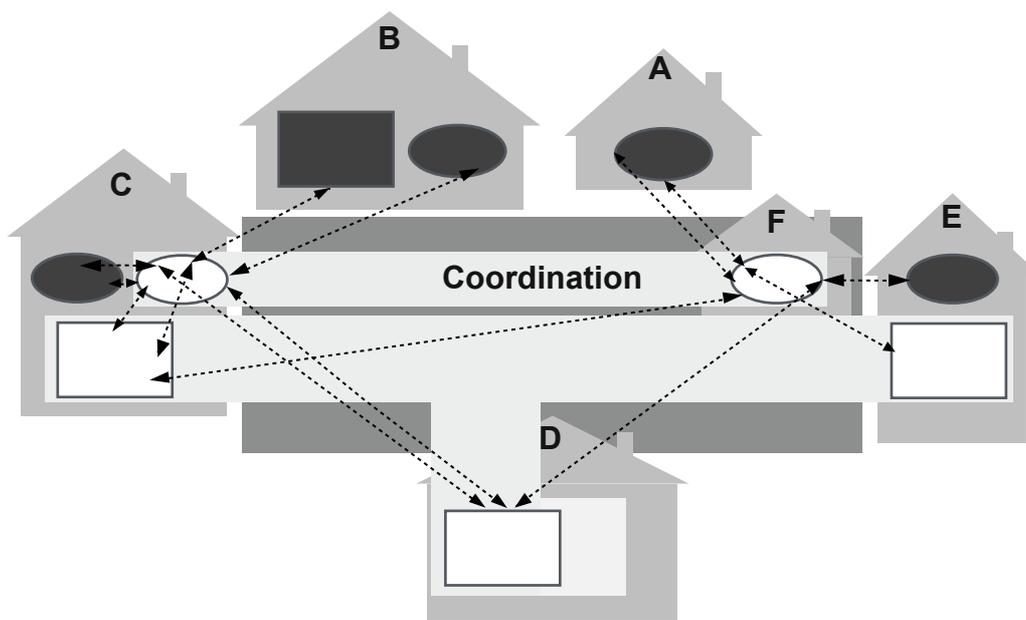


Figura 10. Modelo de distribución para preservación digital de Zierau (2011)

Este modelo nos anticipa las complejas composiciones y posibilidades de tratamiento de redes coordinadas de preservación digital distribuida bajo distintos niveles de integración, y que en el caso de la propuesta del modelo PDDIM, nos ayudará a definir las distintas actuaciones de las IES-MX de acuerdo al estudio realizado sobre los niveles de preservación digital (NDSA-LDP) en IES-MX según los factores clave de sus prácticas de preservación digital y estimando su participación con respecto a su tipología de IES, necesidades, situaciones, alcances, infraestructuras, etc. En esta línea algunas organizaciones gozarán de mayor participación en entornos compartidos y otras en entornos privados o internos, o solo como clientes de servicios de preservación dependiendo de las necesidades de cada una.

En el modelo variable distribuido la existencia de una capa de coordinación intermedia de servicios y procesos compartidos entre las diversas organizaciones que integran al modelo en sus operaciones es de alta importancia, ya que por ejemplo esta capa de coordinación gestiona directamente las actividades de replicación de las unidades de almacenamiento de datos de los repositorios en la red, donde la coordinación de servicios y procesos de preservación replican a los demás repositorios de manera directa e independiente. La capa coordina que los nodos de la red distribuida compartan contenido bajo una regla de operación y acuerdos de niveles de servicio (SLA), procurando así que cada unidad de almacenamiento del repositorio se comuniquen directamente con otra entre ellas, y de esta forma evitar un caos de organización o

vulnerar la seguridad o privacidad de la red y las organizaciones. A modo simplificado podemos decir que la capa de coordinación mantiene dependientes y operativos los nodos en el sistema, ya que cuando cada uno de ellos al realizar la ingesta un nuevo paquete de datos a su sistema de preservación, este por igual es entregado al sistema de ingesta de la capa de coordinación, la cual a su vez llevará a cabo la acción de replicarlos en los demás repositorios. Aunque parezca un procedimiento sencillo, este integra una serie de procesos complejos que deben distinguirse como estratégicos y funcionales de acuerdo a las necesidades de cada uno de los nodos que en conjunto completan la actividad de los modelos distribuidos.

#### **4.3.2 Elementos de definición conceptual y funcionales de sistemas PDD para PDDIM**

A nivel conceptual son tres los elementos concretos que definen y diferencian la actividad de la preservación digital distribuida. Estos son la independencia, la coordinación y la replicación, estos realizan actividades y otorgan roles puntuales en una red de arquitectura distribuida con fines de preservación digital a largo plazo.

En esta línea pragmática, la *independencia* se refiere a las actividades autónomas e internas de cada institución ante sus acervos digitales, ya que cada una de ellas es responsable de crear, gestionar y sostener a largo plazo su archivo digital interno. Al participar una institución en una red distribuida, deberá proporcionar una infraestructura compartida para poder procesar, servir y custodiar de manera responsable e independiente las copias de los archivos de los otros nodos de la red.

La *replicación* sugiere la acción de multiplicar las copias en un número definido por la cantidad de nodos de la red o en su caso específico de lo que se defina en los acuerdos de niveles de servicio, ya que podría replicarse sólo dos o tres ocasiones si el acuerdo así se define, aunque las buenas prácticas sugieren como mínimo tres copias o en otros casos siete.

La *coordinación* además de gestionar la intersección, acuerdos y servicios de los nodos participantes de la red, sobrepone el valor de independencia a los nodos para darles un punto de unión y por ende respetar y coordinar la independencia de cada uno de ellos. Todo antecedente de coordinación y colaboración de tipo interinstitucional y alianzas puntuales de colaboración coordinada como los casos puntuales de la Red Mexicana de Repositorios

Institucionales (REMERI) y la implementación de los lineamientos de repositorios institucionales del Consejo Nacional de Ciencia y Tecnología (CONACYT 2014-15), representan un una gran ventaja ante las posibilidades de concretar un indicio y buen camino en la coordinación colaborativa con fines de preservación digital distribuida bajo una integración sólida y paulatina. Algunos expertos en el área de coordinación y gestión técnica de instituciones en programas de preservación digital distribuida como Ricard de la Vega del Consorci de Serveis Universitaris de Catalunya (CSUC) refieren que si la cooperación no está en el ADN de trabajo de cada organización, será difícil integrar una estrategia distribuida con participación múltiple e institucional de preservación digital, por lo que es vital su entendimiento desde el punto de vista organizacional y técnico para hacer ejecutar la independencia institucional, así como posibilitar la replicación de datos o unidades de archivo de manera transparente y ordenada. Atendiendo las propuestas de modelos variables e integraciones de nodos de almacenaje a los modelos de trabajo distribuido con énfasis en los elementos conceptuales, podemos proyectar los elementos conceptuales a modo simplificado y en vista de diagrama general de la siguiente forma:

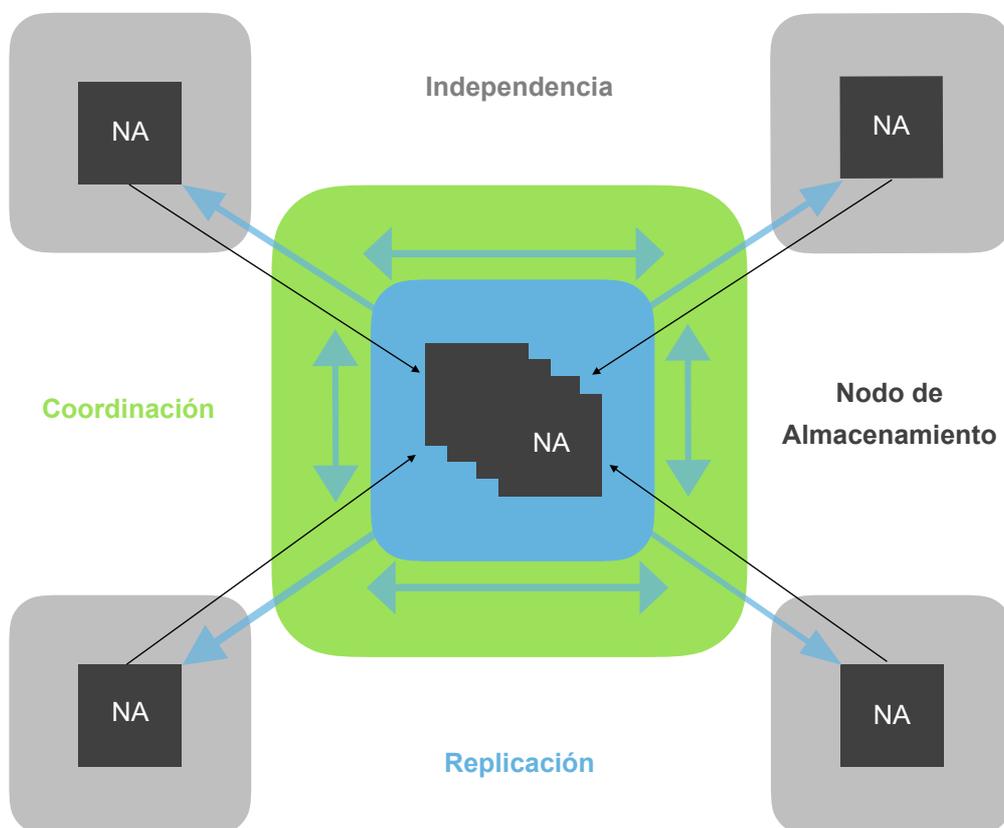


Figura 11. Elementos conceptuales de la preservación digital distribuida

A modo literal en la figura, el concepto de *independencia* (cuadrados grises) pertenece a organizaciones, las cuales mantienen su independencia de operaciones como tal, es decir la gestión de sus contenidos digitales, así como la integración de un repositorio digital con prestaciones de *nodo de almacenaje* (cuadro oscuro) la cual comparte su contenido con el área o capa de coordinación (cuadro verde) que puede ser la organización coordinadora y proveedora de servicios y procesos como el de replicación (cuadro azul) a los nodos de almacenamiento generales (o compartidas) pertinentes. Los tres conceptos que definen la actividad de la preservación digital distribuida mantienen sus roles prioritarios distintivos alrededor de otra serie de operaciones y funciones técnicas específicas de sistemas informáticos de preservación digital distribuida que veremos a continuación.

Los elementos funcionales que integran a los sistemas de preservación digital distribuida tienen su punto de partida del modelo funcional de referencia ISO 14721:2012 bajo el orden de funciones principales de preservación de los elementos de referencia como lo son el entorno (productor, gestión, consumidor), la información (paquete de información, paquete de información enviado, paquete de información archivado, paquete de información diseminado), entidades funcionales (ingesta, planificación de preservación, administración, gestión de datos, archivo de almacenamiento, acceso) e interacciones (peticiones, respuestas a peticiones y órdenes).

Los sistemas especializados en preservación digital distribuida requieren bajo la necesidad específica de almacenar copias de forma distribuida y sincronizada en múltiples sitios, el integrar una serie de elementos funcionales complementarios que proporcionan servicios y procesos para hacer servir las actividades de preservación en una arquitectura distribuida. Estos elementos funcionales de las estrategias distribuidas buscan cumplir una serie de objetivos tales como el de procurar la integridad y seguridad de los contenidos al hacer uso de la función técnica de replicación mediante el *monitoreo* para validar y en su caso *reparar* la estabilidad de los contenidos replicados. Desde el momento de la ingesta de paquetes de información, estos son puestos en contexto y descritos con la información pertinente (metadatos, identificadores, etc.), así mismo el sistema crea las sumas de verificación o checksums necesarios para correlacionar la ingesta con la integridad de paquete de información único (y de los elementos que le integran). A partir de aquí, en la mayoría de los

sistemas de PDD los siguientes procesos son automatizados de acuerdo con las políticas y acuerdos de nivel de servicio (SLA), ya que el paquete de información pasa a convertirse en un paquete de información de archivo que va directo al nodo de almacenamiento (NA) de una institución específica, la cual bajo la previsión de gestión, administración y planificación de preservación de los contenidos almacenados y con la finalidad de incrementar los procedimientos de seguridad e integridad de los mismos, realiza copia idénticas o redundantes del contenidos almacenado con ayuda de un sistema o protocolo informático de preservación digital distribuida, el cual está preparado a nivel de infraestructura, procesos y servicios (harvesting / peer to peer) para replicar copias idénticas del mismo a el nodo de almacenamiento y coordinación central (NAC), para que este replique a los nodos independientes, distribuidos y conectados a su red central de gestión, coordinación y almacenamiento, para así lograr la replicación de copias a nodos geográficamente dispersos.

La replicación como actividad de distribución requiere de otras prestaciones técnicas que aseguren el correcto envío y recepción de los paquetes de información, es donde la función de *monitoreo* de entrega y cosecha de copias realiza una comprobación de integridad (checksums y bits), y posteriormente a modo periódico puede realizar diferentes procesos y servicios de monitoreo de comprobación o en su caso de *reparación* de contenidos dañados. Como ejemplo, el caso del sistema LOCKSS que bajo un sistema de votación que realiza entre todas las copias distribuidas en otros nodos, compara las versiones íntegras tanto para reemplazar o reparar copias dañadas en función a las versiones de integridad. A continuación podemos de manera gráfica las principales funciones y adaptaciones de integración en el modelo distribuido.

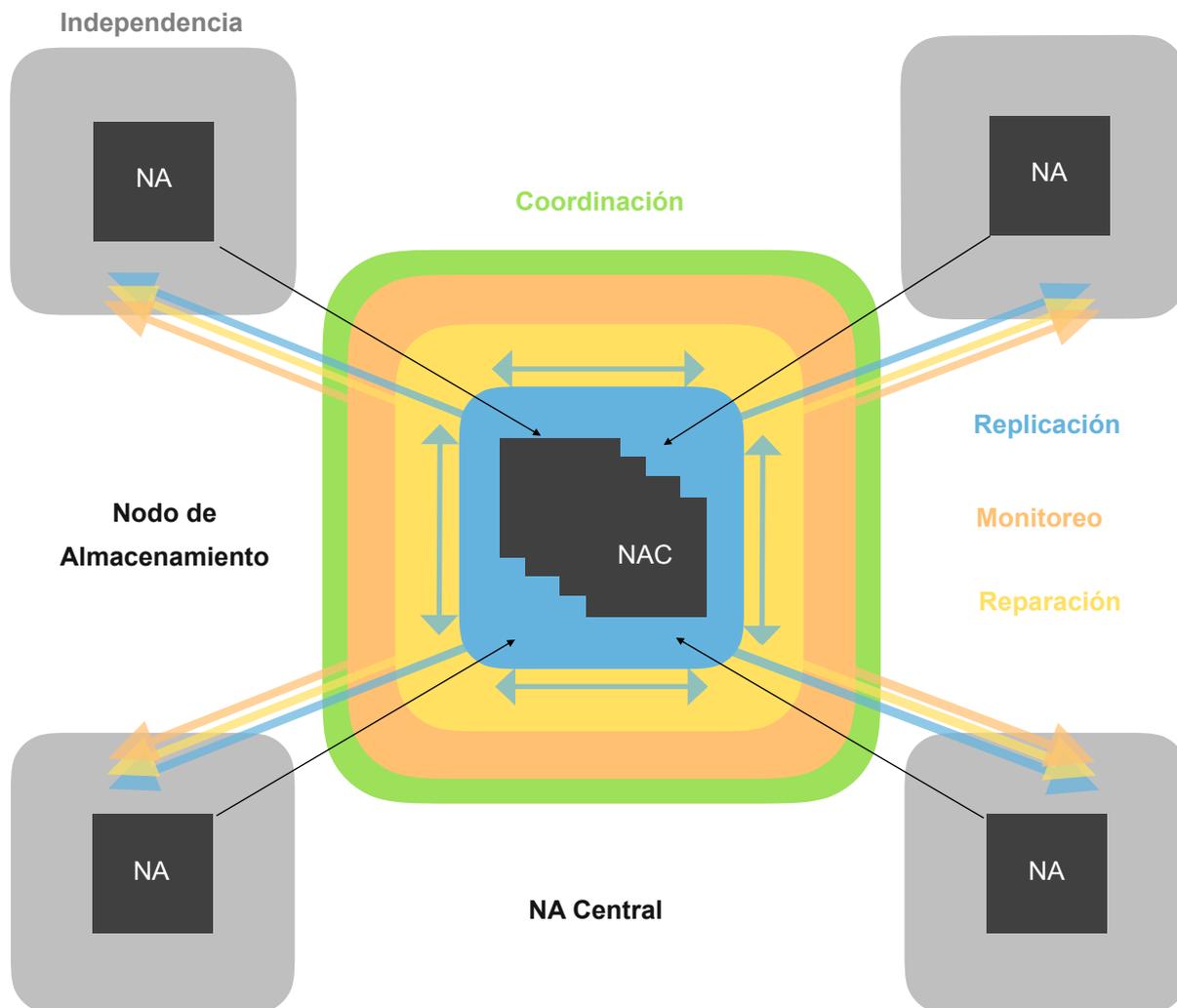


Figura 12. Modelo de elementos de definición conceptual y funcionales en preservación digital distribuida

#### 4.4 Composición y organización funcional en casos de modelos de redes de PDD

A partir de la abstracción de elementos y actores de que componen el ecosistema de la preservación digital distribuida, nos enfocaremos a describir los modelos de The National Danish Bit Repository, Red CARINIANA de Brasil y el Modelo empleado por el Consorci de Serveis Universitaris de Catalunya (CSUC) para que a modo pragmático obtengamos una

panorámica de la organización funcional de los mismos con base en sus objetivos, elementos, servicios y tamaño que nos ayuden a plantear los pros y contras de su implementación como base para un modelo PDDIM.

#### **4.4.1 Caso de modelo National Danish BitRepository**

El objetivo funcional de la red National Danish Bit Repository es el de proveer un protocolo y sistema de preservación digital con arquitectura distribuida en función de la coordinación de tres instituciones participantes (The National Archives, The Royal Library, State and University Library) utilizando una capa de coordinación que gestiona las diferentes acciones y a los pilares (unidades de almacenamiento) de cada una de las instituciones participantes actuando bajo un protocolo de desarrollo propio que las coordina bajo específicos acuerdos de nivel de servicio (SLA) de cada una de ellas y sus pilares que son evaluados y gestionados por la capa de coordinación que mantiene la información sobre el estado de los contenidos digitales almacenados en las instituciones participantes, así como sobre las diferentes necesidades técnicas que van surgiendo en las operaciones e integraciones.

##### ***Elementos conceptuales y funcionales***

De acuerdo con el modelo de NDBR<sup>74</sup> y su visión en conjunto a nivel conceptual y funcional, este hace uso de los conceptos elementos de *cliente*, *protocolo*, *capa de coordinación* y *pilares*. En este sentido las funciones técnicas están definidos en el elemento de *protocolo* bajo diferentes acciones y prestaciones para los *clientes* (nodos de almacenamiento) que representan a los repositorios o software de las organizaciones y la forma en la que pueden acceder a ellas, las cuales son coordinadas por el elemento de *capa de coordinación* el cual realiza las gestiones oportunas de distribución en los diferentes *pilares* (nodo o unidad de almacenamiento independiente) distribuidos en las tres instituciones participantes, que como en el capítulo 2 hemos explicado con más detalle. Estas integran diferentes tipos de medios de almacenamiento como propuesta alternativa a la dependencia de tecnologías únicas y dando un enfoque más amplio al integrar desde unidades de disco magnético, unidades de cinta, unidades estado sólido y posibilidad de unidades en la nube.

---

<sup>74</sup> Visión de conjunto conceptual de bitrepository project en NDBR : <https://sbforge.org/display/BITMAG/The+Bit+Repository+project>

Incorporando y describiendo los conceptos de la visión del modelo de definición conceptual y funcional de PDD al conjunto de conceptos del modelo NDBR, *el cliente* a modo general forma parte de los servicios del nodo de almacenamiento a nivel de software de la institución que solicita servicios y procesos al *protocolo* dentro de la capa de coordinación en la cual se desarrollan las actividades de coordinación de diferentes procesos y servicios (incluida la replicación), las diversas acciones, procesos y servicios refieren por igual a los pilares o nodos de almacenamiento compartidos.

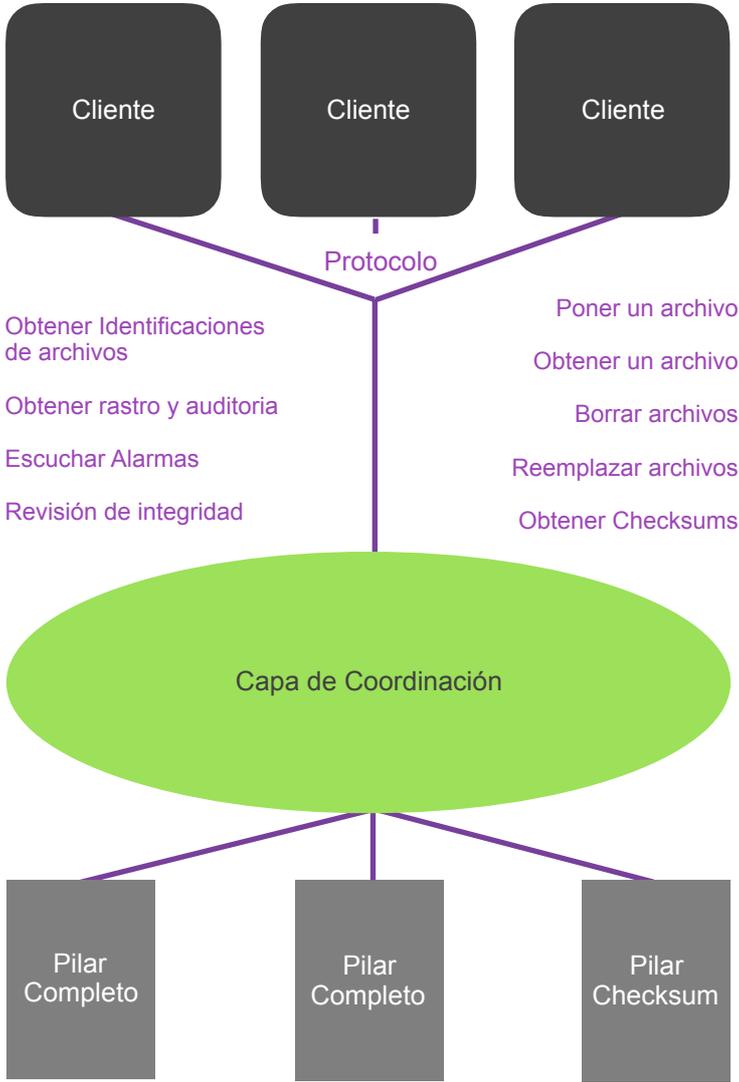


Figura 13. Modelo visión conceptual y funcional NBRP

La propuesta del modelo de funciones de NDBR destaca particularmente en que cada una de las instituciones puede gestionar diferentes pilares de manera independiente o unidades compartidas, y dentro de la caridad de ellas se utiliza un pilar de manera exclusiva para recoger los checksums (por ejemplo), lo que da mayor autonomía a este tipo de registros y replica su existencia por seguridad.

Hay que destacar que en el modelo conceptual y funcional de NDBR los factores de seguridad, confidencialidad y necesidades diversas de las tres instituciones participantes los llevaron a crear un sistema propio, esta situación es un factor muy importante ya que el tipo de contenidos y material a custodiar tiene una alta sensibilidad a nivel nación (archivos nacionales) y de instituciones culturales (Biblioteca Real y Biblioteca del Estado y Universitaria), por lo que la visión de mantener su independencia operativa y funcional en sus actividades de preservación sin hacer uso de servicios de preservación digital externos o cooperativas internacionales, debe entenderse como parte del nivel de necesidades propias de instituciones, gobierno y su administración, dejando claro que todas las partes mantienen un acuerdo, apoyo y entendimiento a la importancia de la preservación digital, lo cual nivel práctico no es una tarea fácil, por lo que es un modelo a seguir.

#### **4.4.2 Caso de modelo CARINIANA de Brasil**

El modelo de la red privada de preservación digital CARINIANA de Brasil está integrado por la propuesta de una metodología y solución tecnológica propia que adopta un dispositivo tecnológico (sistema de preservación digital distribuida LOCKSS) con el objetivo de crear una red de servicios de colaboración de preservación digital distribuida (Márdero, 2013). Aunque el modelo inicial consistía en la colaboración de cinco instituciones socios en la red, una capa de colaboración (IBIC) y una capa de soporte del sistema (LOCKSS); actualmente está conformado por cajas de instituciones socios y dos capas de colaboración que gozan de diferentes prestaciones y necesidades en la red. Cabe señalar que este modelo de red privada de preservación tiene la visión y objetivo de ampliar diferentes subredes temáticas en una red integral como parte de su plan integral de colaboración entre instituciones de educación superior, el instituto Brasileño de ciencia y tecnología y la alianza LOCKSS de Stanford.

El modelo de organización de la red CARINIANA distingue a tres grupos importantes de colaboración. En primer lugar los colaboradores integrales que la integran instituciones públicas o privadas que comparten infraestructura de almacenamiento y que poseen documentos digitales en sistemas electrónicos de gestión. Otro grupo es el de las instituciones colaboradoras las cuales participan en algún proyecto de preservación digital y que poseen documentos digitales compatibles con los sistemas de preservación digital. Por último los colaboradores individuales que son personas que colaboran en proyectos e investigación en preservación digital.

### ***Elementos conceptuales y funcionales***

Dentro del modelo de red CARINIANA podemos observar a tres actores principales. Por un lado encontramos a la integración de una red y subredes privadas de instituciones de educación superior y centros de investigación, por otro a una institución coordinadora de la red (IBIC) y por último la adopción en modo de alianza internacional de la tecnología de un sistema de preservación digital en modo distribuido (LOCKSS). Esta integración de actores hace uso del modelo de elementos conceptuales y funcionales del protocolo de trabajo de la tecnología LOCKSS para entretejer un red de preservación digital distribuida privada que cuenta con la gestión de soporte tecnológico de una capa externa de apoyo en su conexión e implementación de la red.

A nivel específico el modelo de PDD CARINIANA está formado por cuatro elementos de *subred* que se integran por seis *cajas* (LOCKSS Box) que mantienen una integración compartida. En este modelo y visión general podemos observar que cada *subred* mantiene dentro de sus *nodos* o cajas a una caja de IBIC que funciona como el nodo que conecta con la *capa de coordinación IBIC* de subredes y de la red general. Encontramos también la integración de una capa de soporte de sistema o Alianza LOCKSS que mantiene relación directa con la capa de colaboración IBIC.

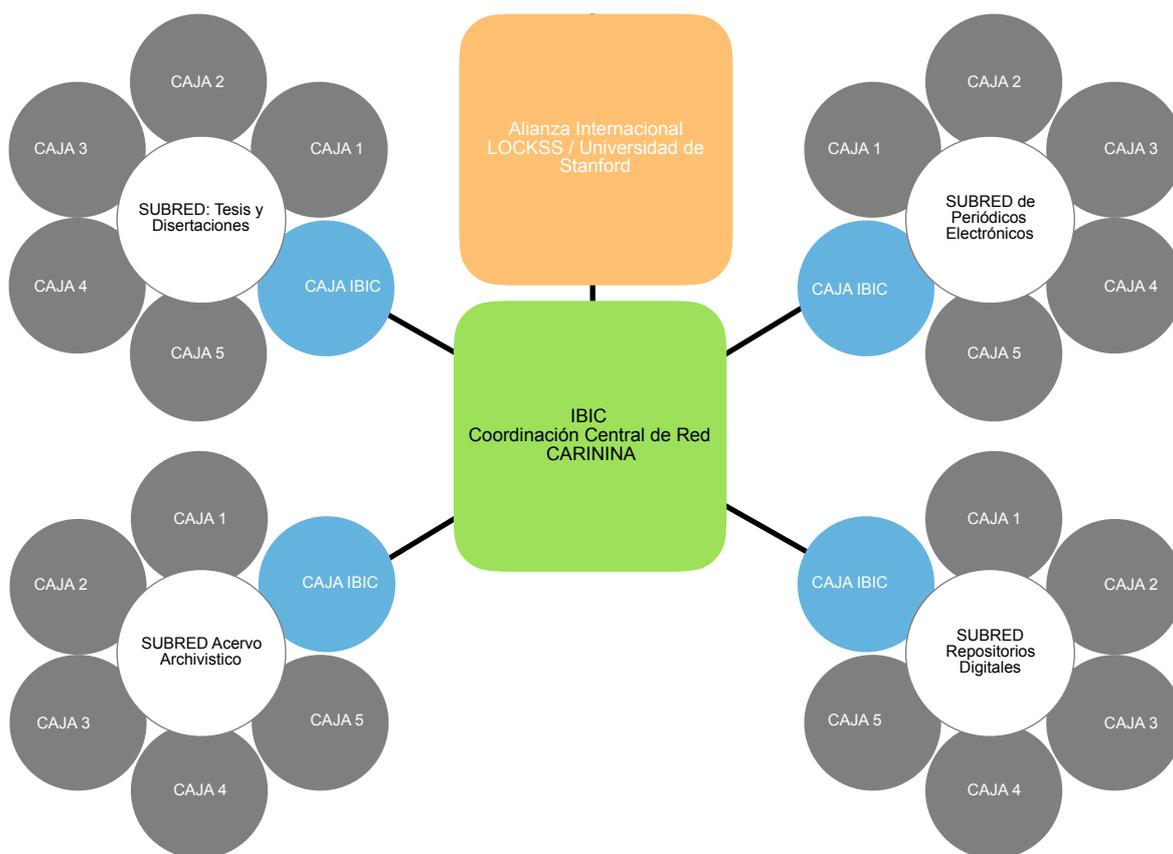


Figura 14. Modelo de red privada CARINIANA y misión de integración de subredes.

Basándonos en la definición del modelo conceptual y funcional PDD, la red CARINIANA en sus actividades de coordinación y soporte con apoyo de la alianza externa LOCKSS, se integran de manera relacional en un solo núcleo de operaciones centralizadas por la capa de coordinación (IBIC) donde cada una de las cajas de las subredes mediante el uso de un cliente (software) puede realizar las actividades de ingreso de contenidos y así como su caja está supeditada y conectada a los servicios ofrecidos por el sistema LOCKSS para replicar, monitorear, reparar y validar sus contenidos almacenados y de las réplicas de las demás cajas de la red bajo su protocolo Library Cache Auditing Protocol (LCAP).

En este modelo de servicios de preservación digital distribuida la alianza internacional con el sistema LOCKSS además de brindar apoyo en soporte, integra a la estructura de la red un servidor web con sede en la Universidad de Stanford que es responsable del depósito de la información que se transmite entre las instituciones asociadas. El servidor (que desempeña el

papel de nodo central o servidor propietario) no interfiere con la administración de la red (Duarte y Márdero, 2015). Este modelo destaca a nivel de colaboración y coordinación ya que integra una alianza estratégica colaborativa con el sistema tecnológico pionero en el ámbito de la preservación digital distribuida desarrollado por la Universidad de Stanford que aporta la seguridad, soporte y garantía de contar con un sistema sólido y robusto de PDD. Por otro lado, la figura de actuación de una institución coordinadora como el IBIC aporta la vinculación, infraestructura y metodología de trabajo colaborativo para integrar las necesidades de varias instituciones socias que tienen en diferentes opciones a nivel de colaboración, lo que ofrece flexibilidad de integración.

La composición de factores tecnológicos y de organización cumplen sus roles y responsabilidades para la implementación bajo dos actores representantes (IBIC y LOCKSS), lo que crea una capa de coordinación que une a los demás nodos (instituciones) sobre un mismo protocolo tecnológico y metodología de organización que cumple con su cometido integral. Hay que destacar que a nivel de propiedad de servidores y a diferencia de la red NDBR que dispone de protocolos y servidores de desarrollo propio, en el modelo de la red CARINIANA la inclusión de un servidor propietario externo (con funciones no administrativas y solo de verificación de metadatos) en conexión con la red privada, representa un factor importante a considerar según sea el tipo de contenidos a preservar y replicar, ya que podría ser un tema sensible ante la colaboración e incorporación de los socios de una red privada por temas de privacidad de la red o en su caso de la seguridad de la información de registros que se intercambia entre cada una de las cajas y que a nivel de metadatos recoge y verifica el servidor propietario. Este punto aunque pueda quedar en una paranoia de seguridad y privacidad, debe ser considerada y discutida plenamente para esclarecer cualquier duda al respecto y definir un modelo de conveniencia de red de alianza colaborativa interinstitucional a nivel internacional.

#### **4.4.3 Caso de modelo CSUC en MetaArchive Cooperative**

El modelo de red de preservación digital distribuida del que participa el Consorci de Serveis Universitaris de Catalunya (CSUC) es un modelo interesante de analizar por la integración de diferentes capas de coordinación de tipo colaborativo y de cooperación. CSUC emplea a la preservación digital a nivel de servicio externo en MetaArchive Cooperative donde el reto de

coordinar implementación tecnológica y organización interinstitucional, destaca en función de consorcios e iniciativas previas y sólidas de colaboración entre dos instituciones que datan desde principios de los años noventa y consolidadas a inicios del año 2000 como lo son el Centre de Supercomputació de Catalunya (CESCA) en el área TIC y del Consorci de Biblioteques Universitàries de Catalunya (CBUC) por el área de bibliotecas respectivamente.

El CSUC<sup>75</sup> nace en el año 2013 de la fusión del Consorci de Biblioteques Universitàries de Catalunya (CBUC) y Centre de Supercomputació de Catalunya (CESCA) con la finalidad de ofrecer y mancomunar los servicios académicos, científicos y de gestión de las entidades en consorcio como lo es la Generalitat de Catalunya y diez universidades catalanas integradas por la Universidad de Barcelona, Universidad Autónoma de Barcelona, Universidad Politécnica de Cataluña, Universidad Pompeu Fabra, Universidad de Lérida, Universidad de Girona, Universidad Rovira i Virgili, Universidad Oberta de Catalunya, Universidad Ramon Llull y la Universidad de Vic-Universidad Central de Cataluña, para potenciar las eficacia y eficiencia de colaboración de sinergias en escala. Dentro de los diferentes servicios que se ofrecen en CSUC, el repositorio de Tesis Doctorals en Xarxa (TDX) representa el grupo de contenidos replicados en la red MetaArchive Cooperative.

Después de evaluar y llevar a cabo diferentes mecanismos de respaldo con objetivos preservación digital o a nivel local (CPD) y regional (backup en otras ciudades de Cataluña), se decidió orientar un mecanismo que ofrecería eficiencia y solidez bajo un sistema cooperativo que replicará en diferentes nodos de almacenamiento de manera geográfica (del otro lado del atlántico) copias seguras de las tesis almacenadas en TDX, por lo que se optó por los servicios de la red MetaArchive Cooperative (mencionada ampliamente en el capítulo tres) que utiliza el sistema LOCKSS como base de sus servicios de preservación digital distribuida e integra una capa de coordinación con aplicación web *Conspectus*<sup>76</sup> con la que se administra la replicación segura, implementación de plugins y colecciones de AU (unidad de archivo) en los demás participantes de la red.

---

<sup>75</sup> CSUC <http://www.csuc.cat>

<sup>76</sup> Metaarchive Conspectus: <http://conspectus.metaarchive.org/archives>

### ***Elementos conceptuales y funcionales***

A nivel conceptual y funcional el modelo de trabajo de preservación digital distribuida del que participa el CSUC con el objetivo de proteger, preservar y mantener el acceso permanente a sus contenidos preservados (tesis doctorales de TDX) en la red de MetaArchive Cooperative, puede definirse a nivel funcional como un *nodo de almacenamiento y coordinación*, ya que además de ser un nodo socio en la red, a su vez coordina a una *subred* colaborativa de universidades que aportan contenidos al repositorio digital de contenidos (TDX). El modelo de trabajo e integración de CSUC a MetaArchive Cooperative destaca como elementos fundamentales en la red a las capas de coordinación, ya que estas tienen dos funciones específicas de actuación ante los nodos de almacenamiento preservado e instituciones u organizaciones que la integran. A nivel de elementos conceptuales de organización, MetaArchive Cooperative como servicio de preservación distribuida cooperativa, integra conceptos propios dentro de su sistema de organización, los cuales aplican a los socios participantes de la red bajo diferentes posibilidades de integración (sustaining member, preservación member, collaborative member). En el caso del CSUC en la red, este participa como miembro colaborativo (collaborative member) ya que coordina a una serie de instituciones bajo un único nodo de almacenamiento y cooperación.

En lo que refiere a las capas de coordinación como elementos funcionales, en primer lugar debemos hablar de la *capa central* de coordinación que gestiona MetaArchive Cooperative que tiene la función principal de administrar a los diferentes socios representados por diferentes cajas o nodos distribuidos con base en la integración del sistema LOCKSS y la aplicación web Conspectus combinando los elementos conceptuales y funcionales de organización y técnicos. MetaArchive Cooperative administra la dirección (caja o nodo) a donde cada una de las réplicas debe hacerlo, así como de las gestiones de solicitud a cada una de las mismas para conectarse bajo la implementación de los plugins necesarios de cada socio para crear grupos de replicación de contenidos. En un segundo momento encontramos como segunda capa de coordinación a la de CSUC, que como lo hemos mencionado mantiene una función doble de nodo de almacenamiento y coordinador de la subred de miembros que aportan contenidos a TDX. Para tener una visión panorámica, a continuación definiremos el modelo de trabajo de CSUC en MetaArchive Cooperative para identificar los elementos de definición conceptual y funcionales. En la representación e integración del modelo de

participación de CSUC en la red de preservación distribuida MetaArchive Cooperative podemos observar que se integran definiciones, conceptos y protocolos técnicos del sistema LOCKSS y Conspectus, los de administración y coordinación cooperativa de MetaArchive, así como los protocolos de coordinación y participación del CSUC como coordinador de contenidos de la subred de universidades o instituciones de educación superior que forman el nodo o caja de preservación digital distribuida en la red. Este ejemplo concreto deja clara que la operatividad en modelos de organización y participación de servicios de preservación como MetaArchive Cooperative, facilitan la proyección y definición de protocolos de participación de grupos de IES en el marco de sus servicios de preservación por los protocolos existentes a niveles de organización y solución técnica. Esta labor no es del todo fácil, ya que requiere de diferentes tipos de infraestructuras de organización humana, instrumentación técnica y de recursos de sostenibles.

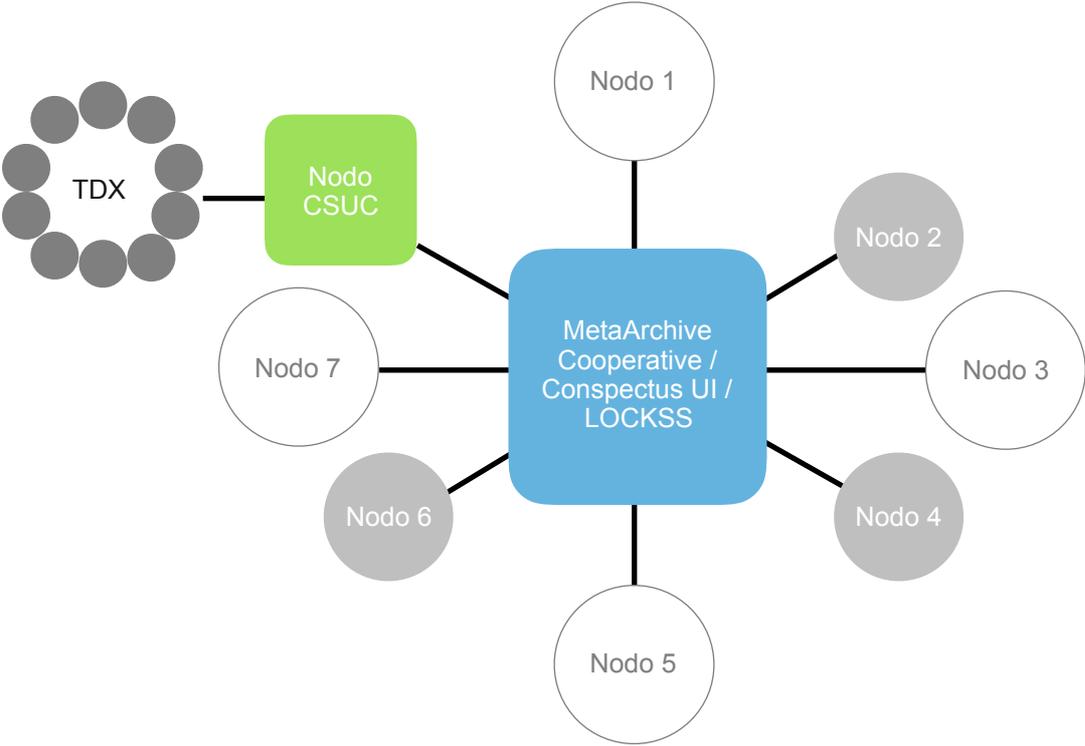


Figura 15. Modelo de participación colaborativa de CSUC en red de PDD MetaArchive Cooperative.

En este caso el CSUC como *intermediario* o *broker* coordinador de servicios universitarios, este aporta una ventaja competitiva que parte de la especialización, evolución y experiencia en la integración colaborativa de ámbitos técnicos (CESCA) y bibliotecas (CBUC) que van más allá de ser una simple capa de coordinación, ya que el CSUC cuenta con los elementos esenciales en todo procedimiento de preservación digital tanto a nivel de roles técnicos, bibliotecarios y de organización humana. Esto define un punto de partida sólido y el cual ayuda considerablemente a crear una normalización integral y centralizada de infraestructuras compartidas como el caso de los repositorios digitales. También, aporta una línea natural para dar el siguiente paso de procuración digital, la preservación digital distribuida o compartida. Hay que recordar que el hecho de contar con un servicio externo de preservación digital distribuida no garantiza tener la solución completa, ya que aun así es necesario incluir todo un sistema de organización de infraestructuras técnicas y humanas (institucionales) paralelas para cumplir plenamente con los objetivos requeridos, por lo que en este ejemplo deja claro que la preservación digital no se trata solo de un *plug and play* técnico, sino que se trata invariablemente de acuerdos, consorcios, colaboración e integración.

#### **4.4.3 Contraste FODA de elementos funcionales en casos de modelos de PDD**

En función de los casos de modelos revisados y su integración a un modelo de entendimiento conceptual en la actuación de figuras de roles coordinadores y ámbitos funcionales tanto de protocolos de sistemas de preservación digital distribuida, analizaremos las posibilidades de integración de elementos de buenas prácticas acorde a las necesidades de IES-MX con base en los resultados de los estudios NDSA-LDP y OCMC con conceptos orientados a la posibilidad de crear el modelo PDDIM. Para atender a la visibilidad y concentración de la revisión de modelos en cuanto a sus elementos de definición conceptual y funcional, realizaremos un balance y consideración general de las fortalezas, oportunidades, desventajas y amenazas de los mismos para definir un punto de partida general en modelos de trabajo con fines de preservación digital distribuida para obtener un patrón de elementos que puedan ser aplicados a el modelo PDDIM.

Tabla 49. Comparativo FODA en casos de modelos PDD.

Caso	Fortalezas	Oportunidades	Desventajas	Amenazas
NDBR	<ol style="list-style-type: none"> <li>1. <i>Desarrollo propio</i> de sistema y protocolo de preservación digital distribuida. <i>BitRepository Project</i></li> <li>2. <i>Red privada basada en confidencialidad, seguridad y privacidad</i> de documentos y datos de orden nacional.</li> <li>3. <i>Capa de coordinación y protocolo de organización</i> interinstitucional de tres instituciones.</li> </ol>	<ol style="list-style-type: none"> <li>1. <i>Flexibilidad de integración de nodos</i> o instituciones adicionales.</li> <li>2. <i>Prevención de dependencias tecnológicas</i> al usar diversos sistemas y soportes de almacenamiento como estrategia.</li> <li>3. <i>Personalización de sistema PDD</i> a requerimientos extras y adecuación a cada institución participante.</li> </ol>	<ol style="list-style-type: none"> <li>1. <i>Alto coste de inversión</i> de desarrollo y mantenimiento de sistema e infraestructuras técnicas y humanas.</li> <li>2. <i>Poca cantidad de nodos replicados</i> geográficamente en Dinamarca.</li> <li>3. <i>Nula replicación internacional</i> en nodos distribuidos en otros sitios que aseguren copias fuera del país.</li> </ol>	<ol style="list-style-type: none"> <li>1. <i>Mantenimiento sostenible</i> de red a largo plazo por y costes de operación.</li> <li>2. <i>Posibilidad de pérdida de datos</i> a causa de la migración y refresco por la diversidad de soportes de almacenamiento.</li> <li>3. <i>Riesgo de pérdida de copias</i> por falta de réplicas en más de tres nodos.</li> </ol>
CARI NIANA	<ol style="list-style-type: none"> <li>1. <i>Alianza e integración de sistema pionero</i> en preservación digital distribuida LOCKSS (Lot of Copies Keep Stuff Safe)</li> <li>2. <i>Red privada con método de trabajo, solución tecnológica y políticas propias</i> para gestión interinstitucional de IES participantes de la red PDD</li> <li>3. Cuenta con una <i>Institución coordinadora estratégica</i> (IBIC) que funciona como nodo y capa de coordinación interinstitucional de la red de PDD en Brasil</li> </ol>	<ol style="list-style-type: none"> <li>1. <i>Flexibilidad de integrar nodos</i> o instituciones adicionales a la red privada en América Latina</li> <li>2. <i>Replicación distribuida en seis cajas</i> repartidas en IES de Brasil, que cuentan con una extensión territorial adecuada para su distribución geográfica nacional</li> <li>3. <i>Pioneros en América Latina en implementar el sistema LOCKSS</i> como red de preservación digital distribuida de IES</li> </ol>	<ol style="list-style-type: none"> <li>1. <i>Bajo coste anual</i> y dependencia técnica de soporte externo de sistema LOCKSS</li> <li>2. <i>Enlace a servidores externos</i> de Universidad de Stanford (LOCKSS Prop Server) como propietario del sistema que recoge metadatos de cajas en Brasil como servicios central y reparte los servicios.</li> <li>3. <i>Nula replicación internacional</i> distribuida en nodos fuera del país</li> </ol>	<ol style="list-style-type: none"> <li>1. <i>Complejidad de coordinación</i> a causa de crecimiento de nodos con diferentes necesidades</li> <li>2. <i>Mantenimiento sostenible</i> de red a largo plazo por y costes de operación</li> <li>3. <i>Riesgo de confidencialidad, seguridad y privacidad</i> de red al depender de servidores centrales externos (U. Stanford)</li> </ol>
CSUC MAC	<ol style="list-style-type: none"> <li>1. <i>Uso e integración de de servicio de preservación digital de MetaArchive Cooperative</i> mediante el uso de la aplicación web <i>Conspectus</i> que gestiona el sistema base (LOCKSS).</li> <li>2. <i>Alta replicación distribuida geográficamente</i> fuera de zona de origen de CSUC(Europa) en siete nodos distribuidos entre América norte y del sur.</li> <li>3. Cuenta con <i>dos capas de coordinación</i>: la capa de servicio MetaArchive y la otra capa es la de CSUC que coordina los contenidos preservados de diversas IES en TDX.</li> </ol>	<ol style="list-style-type: none"> <li>1. <i>Especialización y experiencia en cooperación</i> de CSUC como consorcio cooperativo de IES en repositorios digitales y servicios universitarios.</li> <li>2. <i>Distribución geográfica rotativa y periódica</i> de nodos de almacenamiento de réplicas para asegurar nodos activos y sostenibles.</li> <li>3. <i>Bajo coste</i> de inversión en infraestructura técnica y de operación del nodo participante en red MetaArchive Cooperative.</li> </ol>	<ol style="list-style-type: none"> <li>1. <i>Mediano coste de servicio</i> de participación como nodo colaborativo (grupo de universidades).</li> <li>2. <i>Dependencia técnica de integración a red</i> y necesidad de especialización en el desarrollo de plugins personalizados de cada nodo de la red MetaArchive Cooperative.</li> <li>3. <i>Voto nulo en decisiones administrativas o técnicas</i> de la cooperativa (como usuario colaborativo) así como de en que institución o zona geográfica replicar los contenidos preservados en la red.</li> </ol>	<ol style="list-style-type: none"> <li>1. <i>Riesgo de baja de socios</i> colaboradores de la cooperativa del CSUC, así como de MetaArchive Cooperative.</li> <li>2. <i>Mantenimiento sostenible</i> de red a largo plazo por y costes de operación.</li> <li>3. <i>Riesgo de confidencialidad, seguridad y privacidad</i> de red al depender de servidores centrales externos (MetaArchive Cooperative).</li> </ol>

## **4.5 Visiones de necesidades, compatibilidad y lineamientos para modelo PDDIM**

Para definir una visión objetiva de los elementos necesarios y de buena prácticas de modelos PDD internacionales aplicables en una red de preservación en México, debemos contestar a las siguientes preguntas: ¿Qué necesidades tienen las IES-MX con respecto a la preservación digital?, y ¿Que beneficios y compatibilidad tendría una estrategia a nivel distribuido de colaboración institucional con fines de preservación digital?. Para contestar dicha situación de orden complejo (por las implicaciones de niveles de actuación institucionales y diversidad de necesidades particulares de IES) desarrollaremos puntualmente tres visiones necesarias para ayudar a contestar estas interrogantes de forma objetiva a nuestra investigación.

La primer visión corresponde a las necesidades de IES-MX según la concentración de resultados en el estudio NDSA-LDP. La segunda visión integra la visión de elementos de prácticas internacionales compatible con las necesidades de IES-MX. Por último la tercer visión integra las posibilidades de actuación a nivel interinstitucional de las organizaciones e instituciones involucradas en las actividades del sistema universitario de IES-MX ante un programa y modelo de preservación digital, así como la visión de instituciones de gobierno con objetivos de soporte a programas de actividades de ámbito documental, tecnológico, científico e universitario (CONACYT, AGM, etc.). De esta manera pondremos en contexto el mapa de ruta con los elementos necesarios a nivel funcional, de organización y coordinación de roles institucionales en el modelo.

### **4.5.1 Necesidades de preservación digital en IES-MX ante modelo PDDIM**

Para responder a la interrogante de cuáles son las necesidades de IES-MX ante actividades de preservación digital, abstraemos de forma concentrada y sintetizada los resultados de prácticas detectadas en el estudio NDSA-LDP en IES-MX a nivel general y en función a las posiciones referidas en los indicadores de identificación, categorías, actividades y niveles de preservación digital como prácticas con necesidades de implementación o refuerzo acorde a la metodología de recomendaciones NDSA-LDP y buenas prácticas de una estrategia de preservación digital.

### ***Alcance de niveles en preservación digital***

De acuerdo a la posición general de niveles de preservación digital de IES-MX, estas no alcanzan a cubrir más del 50% de la totalidad de las actividades y prácticas de preservación digital requeridas en las recomendaciones de los 4 tipos de niveles de preservación digital, ya que las IES-MX se encuentran en un máximo de cobertura del 36% perteneciente al nivel 1 de protección de datos y un mínimo del 28% en el nivel 2 de conocimiento de datos. Estos resultados nos dejan claro que existen necesidades ampliación de cobertura a nivel práctico (ejecución) y teórico (conocimiento) de las prácticas recomendadas y la normalización de las mismas, lo que representa un requisito esencial a nivel de práctica individual de IES para de integración de un modelo colaborativo de preservación digital en modo distribuido que ubique el nivel de servicios, participación y prestaciones para cada uno de los integrantes de acuerdo a su cobertura y alcance.

### ***Prioridad en tipología documental de contenidos y formatos***

Otra visión de necesidades detectada corresponde a la tipología de archivos o contenidos digitales que gestionan IES-MX federales y estatales dentro de sus servicios digitales de acceso y difusión como repositorios, BiDis o archivo digital, etc. La demanda de IES-MX, se decanta mayormente por la tipología de tesis (70%), artículos de investigación (60%) y libros digitales (71%). Esto deja ver que existe una necesidad prioritaria de gestión para estos tipos de contenidos a nivel de manejo, producción e inclusión en herramientas de difusión o depósito digital, ya que al encontrar una exposición y difusión digital pública de orden institucional, requieren además de garantizar su acceso constante y fluido a corto plazo, el asegurar su acceso permanente y seguro a largo plazo. Para cumplir dichas necesidades prioritarias para la preservación digital de estos tipos de contenidos digitales, es indispensable considerar por igual integrar un procedimiento de definición de normalización de estructuras y formatos digitales que contribuya a mantener un estándar sólido que posibilite a su vez las futuras integraciones de colaboración de contenidos con otras instituciones.

### ***Almacenamiento replicado y distribuido***

En un procedimiento de preservación digital el almacenamiento de los contenidos es uno de los elementos indispensables donde la replicación y la distribución marcan una diferencia estratégica al asegurar copias extras de los contenidos almacenados. En este renglón y acorde a los resultados de las IES-MX, la gran mayoría de ellas solo disponen de una copia de su archivo digital en una localización geográfica distinta o en su caso solo mueven el contenido de un soporte de almacenamiento a otro. En algunos casos manifiestan tener una copia en una localización geográfica distinta, pero en ningún caso el de contar con tres copias completas de sus contenidos y solo una IES marco que cuenta con tres copias en locaciones geográficas distintas fuera de desastres. Estos resultados se contradicen con el 10% de las IES que afirman contar con un sistema o procedimiento de monitoreo de procesos de obsolescencia de sus sistemas de almacenamiento. Estas situaciones que no rebasan los niveles 1 y 2 de las recomendaciones de preservación digital, nos dejan ver que existe una amplia necesidad de definir y planificar formalmente las estrategias de almacenamiento que además de contemplar los puntos cuantitativo sobre la cantidad de copias y sitios donde se van a distribuir, deben asistir cualitativamente a un plan de monitoreo de obsolescencia de sus múltiples sistemas de almacenamiento y calidad de copias. Estas acciones son una de las panaceas necesarias para plantear un modelo trabajo de almacenamiento distribuido en modo independiente o de participación cooperativa con otras instituciones.

### ***Documentación de actividades y gestión de contenidos***

Una de las prácticas esenciales en las recomendaciones de preservación digital de la metodología NDSA-LDP refiere a documentar las actividades de la gestión de un archivo digital como parte del nivel dos de conocimiento de los datos y contenidos digitales. En los resultados del estudio a IES-MX aunque estas no cubren el nivel de conocimiento de datos a más del 31%, por lo que el mediano cumplimiento y realización de actividades puntuales como las de documentar sistemas de almacenamiento, ubicaciones de almacenamiento, accesos restringidos e inventario de formatos, debe fortalecerse. La documentación de actividades y gestión de contenidos, representan prácticas fundamentales prioritarias en una estrategia de preservación digital a largo plazo, ya que los contextos a nivel de creación o producción, estructuras técnicas y modificaciones de documentos digitales deben estar cubiertos plenamente como parte esencial de la estrategia digital en modo distribuido, ya que

sin la documentación e inventarios requeridos sobre los contextos digitales es imposible atravesar el umbral de los almacenamientos distribuidos, las revisiones de integridad de las réplicas y de las posibilidades de reparar o validar el porvenir de los mismos a distancia.

### ***Consensos de colaboración técnica***

La situación asimétrica y de alcance medio de resultados de niveles de buenas prácticas en actividades de preservación digital de IES federales y estatales que las sitúan mayoritariamente en la protección de datos (nivel 2), y donde las IES federales ponderan el nivel 2 de conocimiento de datos como las actividades menos practicadas, deja entrever que a nivel grupal e individual cada IES recoge una serie de prácticas a conveniencia y ventaja de sus necesidades de implementación institucional, y por lo tanto no se corresponde con los objetivos de consensos de colaboración y prácticas recomendadas en la metodología NDSA-LDP. Esta situación define la necesidad de integrar un sistema de lineamientos de colaboración y consensos a nivel de actividades con fines de preservación digital que cubran claramente las recomendaciones de almacenamiento, integridad, seguridad, metadatos y formatos de archivos digitales, para normalizar e intercambiar prácticas en beneficio de la conservación de acervos digitales de instituciones de educación superior para asegurar su acceso a largo plazo reforzando los niveles de seguridad y protección, conocimiento, control y reparación de los datos que los integran.

Es importante por igual definir según un protocolo de consensos de colaboración técnica los alcances, recursos y requerimientos individuales necesarios para cada tipo de institución de acuerdo a su infraestructura institucional y de contenidos, que defina un balance justo y funcional entre la diversidad de posibilidades para normalizar a los mismos niveles de actuación a de prácticas a todas las IES-MX según la diversidad y posibilidades.

### ***Relaciones y coordinación interinstitucional***

Hoy en día gracias a la conectividad de internet, las visiones internacionales de acceso abierto y las colaboraciones en investigaciones a distancia; la existencia de instituciones o organizaciones aisladas es casi nula. Una de las finalidades de la normalización de documentos digitales es facilitar su acceso e intercambio a nivel técnico como de organización. Al contener en el estudio NDSA-LDP a IES-MX una gran cantidad actividades

y recomendaciones que tienen como fin el de alcanzar las buenas prácticas de preservación digital, estas no se logran o se cumplen plenamente en la mayoría de ellas a causa de las diferencias institucionales existentes en su tamaño, conocimiento, presupuestos o subsistema de pertenencia. Es importante considerar que para una estrategia de preservación digital en modo centralizado o en arquitectura distribuida, el intercambio, la colaboración y las alianzas técnicas resultan una ser las estrategias que relacionan y que se emplean para alcanzar las metas en común de un grupo colaborativo para la mejora de los objetivos de preservación a largo plazo, bajo procedimientos y protocolos coordinados de gestión de documentos digitales de las instituciones involucradas en un proyecto de preservación digital colaborativa o distribuida en común.

De acuerdo a los antecedentes de modelos de participación en donde algunas IES-MX tienen como experiencia previa la colaboración interinstitucional como el caso de la red mexicana de repositorios institucionales (REMEDI), algunas de las políticas e integraciones de lineamientos para la creación de repositorios institucionales y repositorio nacional (CONACYT, 2015), podemos entender que la siguiente necesidad después de la integración colaborativa y coordinada de redes de repositorios, repositorios institucionales individuales y repositorios digitales centralizados interinstitucionales, dará paso a colocar a una institución, grupo o representación interinstitucional como eje de coordinación que podría ejecutar las acciones de nodo central o punto de coordinación con lo que el siguiente nivel a cumplir de asegurar el acceso permanente y largo plazo gozaría de un antecedente que más que ser un requisito por cumplir es una necesidad de antesala que garantiza la operatividad coordinada a nivel de preservación digital de acuerdo con las prácticas de modelos internacionales y recomendaciones de las metodologías específicas.

### ***Divulgación formativa sobre importancia de la preservación digital a largo plazo***

Así como las necesidades de establecer políticas de coordinación y alianzas interinstitucionales refieren a beneficiar una estrategia de conjunto, otra de las necesidades notables en el estudio de IES-MX recae en la falta de contar con una estructura sólida de divulgación y formación sobre el impacto colateral que ofrece la preservación digital a las instituciones y a la sociedad al proteger patrimonios digitales en común. En primer lugar los beneficios de impacto a nivel formativo de los miembros de una organización designada o un

grupo de instituciones dedicadas a las actividades de gestión de documentos y contenidos digitales de importancia clave, recaen en el fortalecimiento de una estrategia común a nivel técnico, de organización y funcional que facilite el mantener sus archivos a largo plazo. El realizar una actividad con objetivos de preservar conocimiento o patrimonio digital común de universidades públicas (investigación y conocimiento académico generado con recursos públicos) plantea una serie de alcances que deben ser comunicados a nivel social en sus diferentes capas de interés, para fomentar el conocimiento de su disposición, de apoyo y la posibilidad de interés en posibles mecenas a nivel del sector privado que puedan colaborar indirectamente a mantener y hacer sostenibles los recursos de preservación digital en función a participaciones de imagen de responsabilidad social del proyecto.

A nivel social es importante difundir la importancia de la labor de preservación digital a largo plazo como goce para generaciones futuras sobre el eje productivo de las instituciones académicas de educación superior y expresión de sus resultados materializados en investigaciones y publicaciones de creación docente y de alumnado titular. Esto recrea un proceso de reto ante el conglomerado de la sociedad que más allá de tener impactos mediáticos de resultados o de transparencia, cuente con un registro íntegro de la actividad digital universitaria del presente en modo de perdurable y accesible, y que pueda dar constancia y justificación de las necesidades de sostenibilidad para la creación de más conocimiento nacional, y de la posibilidad de hacer llegar más allá de los públicos especializados el valor de la preservación del conocimiento producido y almacenado digitalmente por la academia e instituciones públicas de educación superior del país.

#### **4.5.2 Compatibilidad de elementos y prácticas de modelos internacionales en modelo PDDIM**

Las iniciativas que desarrollan, implementan o contratan servicios con objetivos de preservación digital distribuida de antemano merecen un reconocimiento especial como referentes de hechos y acciones de colocar en la misma mesa a la colaboración organizada coordinada de instituciones y contenidos con el mismo objetivo de mantener a largo plazo el acceso permanece al conocimiento común producido y representado en soportes documentales de origen o conversión digital. El contraste y acercamiento a las prácticas de organización en los modelos NDBR, CARINIANA y CSUC MAC son un gran ejemplo de

casos de buenas prácticas en modelos de preservación digital distribuida, los cuales nos otorgan de manera general una línea de organización para modelar el trabajo a desarrollar por otras instituciones con intenciones, posibilidades y compromiso de mantener acervos digitales accesibles a largo plazo. A partir del contraste de los tres casos prácticos de modelos de organización en redes de PDD, definimos y sopesamos los elementos que cuentan con la flexibilidad de integración y posibilidades de desarrollo en compatibilidad de las necesidades del modelo de preservación digital distribuida IES México (PDDIM) los elementos indispensables en la definición de una estrategia de este tipo.

### ***Capa de coordinación***

Uno de los elementos del contraste que muestran compatibilidad de integración en el modelo PDDIM es la capa de coordinación, la cual funciona como eje central de unión y organización de un grupo de participantes de una red con fines de preservación digital pues puede tomar diferentes personalidades de acuerdo a las necesidades y requerimientos de la red. En primer lugar encontramos la personalidad de institución coordinadora que mantiene tres niveles de actuación que se definen como la administración de servicios de preservación, la administración de contenidos y la administración de participantes. Otra personalidad es la de un sistema central coordinador de participantes de acuerdo con protocolos informáticos donde ambas personalidades de coordinación actúan como ejes centrales de monitoreo, distribución, organización y unión de los miembros de una red para operar de manera armónica con base en coordinación, protocolos, sistemas de organización y los acuerdos de niveles de servicios definidos para cada uno de los miembros. En el caso de la propuesta para el modelo PDDIM la diversidad de participantes del subsistema federal y estatal de IES, esta figura es necesaria para poder balancear la participación y requerimientos de las instituciones en una estrategia que ayude a proporcionar su tamaño, intenciones y resultados.

### ***Sistemas y tecnología de preservación digital distribuida***

Los elementos a nivel técnico tienen una amplia aplicación en los casos de los modelos contrastados y que se traducen en soluciones de sistemas y tecnología bajo diversas opciones como la de subcontratación de servicios de preservación distribuida de acuerdo a un servicio en red internacional (CSUC - MAC), la realización de una alianza estratégica y técnica de

colaboración internacional con un sistema PDD (CARINIANA - LOCKSS), o por otro lado optar por la creación y desarrollo propio de un sistema de protocolos informáticos de preservación digital (NDBR). Estos ejemplos aportan distintos niveles de beneficios, prestaciones y necesidades. El caso del desarrollo propio nos aporta directamente beneficios de flexibilidad, personalización y adaptabilidad de la red a todos los objetivos planteados, como por ejemplo en los objetivos más sensibles de una estrategia de preservación digital como el control de privacidad, confidencialidad y seguridad de documentos de orden nacional como es el caso de los contenidos del Archivo Nacional de Dinamarca el modelo de desarrollo propio permite estos objetivos plenamente e incluso cumpliendo normativas de gestión y tratamiento de datos dentro de territorio de la UE. En este renglón hay que considerar antes de elegir una solución tecnológica de PDD, la tipología de documentos y nivel de contenidos sensibles a preservar y compartir, ya que estos pueden estar expuestos a diversas concurrencias y factores de riesgo técnico e incluso de jurisprudencia al ser parte de una red altamente distribuida a nivel internacional, como es el caso de las alianzas con tecnologías no propias y que intervengan con servidores externos fuera del país de la red privada de preservación, lo que podría ser en su caso un factor determinante para sopesar el uso de un sistema propio tomando en cuenta que para ello se requiere de una alta inversión para el desarrollo, implementación y mantenimiento, en comparación a una solución tecnológica con objetivos de preservación distribuida en modo de subcontratación o alianza tecnológica que mantiene costos más alcanzables de implementación y uso a cambio de la internacionalización asociada de contenidos preservados.

Para el planteamiento de un modelo PDDIM encontramos una diversidad de opciones que podrían ser compatibles de manera general, sin embargo y como regla general en las experiencias de redes de preservación digital distribuida, no hay trajes a la medida que soporten integrar a un grupo heterogéneo de instituciones de manera forzada, recordando que las necesidades de las instituciones son diversas y la asignación de sensibilidad de documentos jugará un factor determinante.

### ***Replicación y distribución geográfica***

En los modelos contrastados se pueden encontrar ventajas y desventajas por el tipo de solución por la que se opte. Una situación a tomar en cuenta es la distribución de copias

replicadas en diferentes nodos pues esta se puede integrar en diversos niveles de interés y beneficios. En el caso de la solución de integrar un sistema propio con participantes y servidores de un mismo país (NDBR) o mediante adopción de un sistema en alianza internacional (CARINIANA), el desarrollo propio acarrea como beneficio inmediato la alta confidencialidad y privacidad de contenidos sensibles preservados, sin embargo en estos modelos se podrían considerar una mínima desventaja al contar una distribución de sus réplicas dispersas en una misma geografía del país con respecto a las prácticas que llevan a cabo servicios y cooperativas de PDD (MetaArchive Cooperative) que integran una alta replicación distribuida de contenidos preservados en siete sitios geográficamente e internacionalmente dispersos como medidas de seguridad extra.

Estas consideraciones de alta dispersión geográfica de réplicas en modo internacional conllevan un calado directo a nivel de jurisdicción internacional de protección de datos en comparación a la ventaja del tratamiento de datos en un mismo país. Una alta replicación implica apostar por la colaboración internacional bajo ciertas posibilidades de riesgo de compartir contenidos sensibles en esos niveles de dispersión. Para el modelo PDDIM se debe considerar la ubicación de los servidores de datos preservados, del uso de soluciones de tecnología y el tratamiento de los datos (ya sean de mínimos metadatos o paquetes completos), así mismo plantear la correcta utilización de extensión geográfica de México (1,964,375 km<sup>2</sup>) para la distribución geográfica de réplicas que cumpla con los requisitos de sitios fuera de desastres naturales. Reiteramos que las posibilidades siempre estarán por encima de las necesidades, es importante plantear un esquema de alcance realizable a partir de casos de implementación en modelos de buenas prácticas.

### ***Sostenibilidad de proyecto***

La amenaza reiterativa y siempre latente de la preservación digital distribuida más allá del riesgo o fallos en las infraestructuras, se mantiene fija en el mantenimiento sostenible de una red, de sus participantes (operadores y organizaciones), de los recursos económicos y el compromiso de ambos por mantener a largo plazo un modelo de trabajo de preservación digital en modo distribuido. A partir de aquí, hay que destacar que en los casos contrastados podemos observar que modelos de PDD de tipo compacto como el de NDBR bajo la integración de tres instituciones base, mantiene un nivel de gestión y coordinación concreto,

que no significa que por ser solo tres instituciones implicadas significa que es un procedimiento económico y fácil de aplicar, ya que hay que recordar que detrás de esta iniciativa BitRepository project, existen altos antecedentes de trabajo e intercambio en conjunto de respaldos y archivos entre las instituciones, además de la gestión de recursos para llevar a cabo un desarrollo propio y de una iniciativa nacional de compromiso de preservación digital que da soporte a el proyecto. Ello es un ejemplo a seguir en términos de coordinación interinstitucional y un desarrollo por iniciativa propia de país para proteger el patrimonio nacional y cultural del mismo.

En los casos también hay que destacar la participación de CBUC en los antecedentes y altos índices de colaboración previa interinstitucional para gestionar otras capas de coordinación (de un grupo de universidades) con fines de asegurar sus patrimonios documentales. Sin ello no podría ocurrir la colaboración, coordinación y selección de la mejor solución para dar un extra de seguridad a los respaldos de las universidades colaboradoras bajo un programa de apoyo de gobierno que estimula a la integración de diversas actividades encaminadas a mejorar operativamente los servicios de instituciones de educación superior y por ende de preservar sus productos académicos.

Las alianzas tecnológicas internacionales con fines de preservación (LOCKSS) siempre representaran a nivel de identidad e imagen institucional un trabajo de internacionalización del sector que le atañe, pero más allá de ello también consolida la certeza de apostar por tecnología líder del mercado y por ende para gestionar la antecámara interna de las instituciones para la aprobación de apoyos de orden de alianzas, lo que se traduce en la posibilidad de mantener un programa sólido bajo estos elementos en juego. En todo caso y para la compatibilidad del modelo PDDIM, la sostenibilidad de un proyecto de preservación digital en modo distribuido o en alta participación colaborativa podría someterse a el entendimiento de economía de escala y compromiso para su soporte repartido entre los participantes como primera opción. Otros factores fundamentales como los de palpar, divulgar y difundir los beneficios de un programa de preservación digital, siempre serán los mejores motores que den posibilidad a la perpetuidad de los mismos tomando en cuenta que siempre podrían integrarse estrategias que socialicen la información preservada y su valor no radique en activos congelados o en archivos negros, si no en la rentabilidad bajo la personalidad de activos digitales que son preservados a largo plazo.

### **4.5.3 Políticas, lineamientos e instituciones en México ante modelo PDDIM**

A partir del año 2012 en México se han desarrollado y planteado diferentes políticas, reformas lineamientos y requerimientos en diversos niveles de gobierno, organismos e instituciones. Estas tienen la finalidad de normalizar y reformar el sector de la gestión de datos, información, archivos y preservación de los mismos. Actualmente el reto de México con respecto a los rubros y políticas de gestión de información, se entreteteje en articular siete políticas que conviven en los mismos procedimientos de integración e implementación como lo son: Ley de Acceso a la Información, Ley de Transparencia y Acceso a la Información, Ley de Gobierno Digital, Ley de Gobierno Abierto, Ley de Archivos en Gestión Documental, el Sistema Nacional de Información y la reforma a la Ley de Ciencia y Tecnología con el acceso abierto a los recursos de formación científica, tecnológica y de innovación. Lo cual refiere a el estado actual de generación de diferentes acciones y planes de trabajo para atender de manera integrada a los mismos (AGN, 2016).

En lo que se refiere al ámbito de la gestión de archivos, repositorios y preservación digital, diferentes instituciones como el Archivo General de la Nación (AGN), el Consejo Nacional de Ciencia y Tecnología (CONACYT) y la Secretaría de Educación Pública (SEP) integran un efecto directo a el ámbito de la gestión de archivos y patrimonios digitales de las instituciones de educación superior (IES) a nivel de ciencia y tecnología, tratamiento documental y transparencia de datos. Es por ello que es importante analizar las visiones de las instituciones que tienen altas posibilidades de relación y colaboración en el ámbito de una estrategia de preservación digital en modo distribuido con IES públicas de México.

#### ***Archivo General de la Nación***

El Archivo General de la Nación de México es el más importante de América no solo por el volumen de documentos que posee, sino también por la antigüedad de los mismos en obras originales en diversidad de ciencias, su rico acervo cultural y los variados servicios que proporciona. Como institución líder del Sistema Nacional de Archivos en México, tiene la obligación de dar los pasos necesarios para actualizar sus procesos y servicios, de acuerdo a los avances tecnológicos aplicables al manejo de la información documental (Palacios, 2000).

La visión de experiencia en el tratamiento documental, normalización y creación de pautas de archivística en México, lo coloca como un órgano vital en la consolidación de programas de preservación digital en la diversa red de archivos de toda la República Mexicana. El AGN ha pugnado por la integración de la iniciativa de leyes generales y nacionales de archivos que ayudan a dictaminar y formalizar diferentes pautas para proteger, custodiar, ordenar, describir y conservar los documentos que conforman su acervo y los de la red de archivos en México.

Para la integración del modelo de PDDIM el AGN es el órgano de referencia autorizado y obligado para posibilitar tanto a entidades de gobierno (principalmente) como a otras instituciones que lo requieran (como IES), de la orientación sobre pautas archivistas y ayuda en la organización, integración y preservación de sus archivos. La importancia institucional que mantiene AGN en la participación y divulgación de un programa de preservación digital nacional en modo distribuido forma parte de sus objetivos estratégicos de preservación del patrimonio documental, donde su autoridad será de gran ayuda a nivel de orientación de lineamientos archivístico técnicos con fines de preservación, por lo que la visión de importancia del Archivo General de la Nación ante un modelo de preservación digital distribuida de documentos con fines de investigación en México de IES es una visión compartida y sólida que lo apunta como un referente obligado de participación.

### ***Consejo Nacional de Ciencia y Tecnología***

En función de lo estipulado en las últimas reformas de la Ley de Ciencia y Tecnología en México, CONACYT ha desarrollado y publicado los documentos que estipulan los lineamientos generales (2014) y lineamientos técnicos (2015) para la creación del repositorio nacional y los repositorios institucionales. Esta propuesta coloca al Consejo Nacional de Ciencia y Tecnología, como la institución puntera y rectora en lo que se refiere a la estandarización de la gestión digital institucional de procedimientos de almacenamiento, normalización y descripción de documentos digitales con impacto directo en el sistema de instituciones de educación superior en México. Debemos entender los alcances de aplicación que sugieren los lineamientos que estipulan lo siguiente: *“Los presentes lineamientos serán aplicables al Sistema Nacional de Ciencia, Tecnología e Innovación (SNCTI) y a todo aquel ente que lleve a cabo producción académica, científica, tecnológica y de innovación,*

*financiados total o parcialmente con recursos públicos o hayan utilizado infraestructura pública.”* (CONACYT, 2014), lo que deja clara la aplicación, posición y paulatina obligación a la que se incluyen la IES-MX públicas y sus productos de investigación. Así mismo, el documento en su artículo 4o. en la definición de repositorio nacional, da cabida a CONACYT como institución coordinadora del repositorio nacional en modo centralizado, el cual estará alimentado por la red de repositorios institucionales de IES, centros de investigación, etc, de la siguiente forma: *“El Repositorio Nacional de Acceso Abierto a Recursos de Información Académica, Científica, Tecnológica y de Innovación, de calidad e interés social y cultural el cual consiste en una plataforma digital centralizada cuya coordinación y modelos de operación serán emitidos por el CONACYT, que siguiendo estándares internacionales, almacena, mantiene, preserva y disemina la información académica, científica, tecnológica y de innovación, la cual se deriva de las investigaciones, productos educativos y académicos;”*(CONACYT, 2014). El Consejo de Ciencia y Tecnología además de abanderar la iniciativa de creación estandarizada de repositorios, mantiene la definición de eje coordinador y emisor de la estrategia de repositorios bajo una figura interna responsable de las decisiones sobre el contenido, estructuras e integración de los repositorios, la cual se denomina *Comité de Acceso Abierto a la Información Científica, Tecnología y de innovación*, quien fungirá como capa de coordinación en sobre la labor específica.

La visión de integrar un modelo de preservación digital en modo distribuido que complemente las operaciones regulares de la red de repositorios, ocupa un nivel prioritario por la necesidad obligada de preservar más allá del esquema del depósito digital, la información incluida en los mismos y la red que integran. Así tal mantiene una línea directa sobre la definición, gestión o designación de la institución o grupo de instituciones que puedan actuar como coordinadores o capas de coordinación en la estrategia distribuida de preservación digital. Teniendo en cuenta el antecedente de la estrategia de preservación digital de los casos de modelos internacionales estudiados (NDBR, CARINIANA, CBUC-MAC) que incluyen ejes coordinadores, capas y protocolos de coordinación en sus estrategias de PDD, no podríamos dudar que CONACYT bajo su autoridad de gestión y coordinación de recursos destinados a la ciencia y tecnología en México, determinara la figura más idónea tomando en cuenta los parámetros de buenas prácticas internacionales y necesidades de actuación requeridas.

## *Secretaría de educación pública e IES*

En el universo de la Secretaría de Educación Pública (SEP) en México al que pertenecen las instituciones de educación superior es gestionado por el área de la Subsecretaría de Educación Superior (SES), la cual tiene como objetivo a través de diferentes políticas públicas, planes y programas el brindar una educación equitativa, pertinente, flexible, innovadora, diversificada y de amplia cobertura (SEP, 2016). Aunque la SEP y la SES mantienen objetivos a nivel de coordinación de políticas y programas para propiciar mejoras a la educación superior, también tienen injerencia en proporcionar apoyo para colocar las condiciones necesarias en las que la educación superior tenga la mayor cobertura de calidad, entendiendo por ello una educación equitativa, pertinente, flexible e innovadora. Justamente en la innovación es donde podemos colocar a los modelos de sistemas con integración tecnológica y de organización para proveer permanencia, preservación y trazabilidad de los productos finales de las instituciones de educación superior, más allá de elementos cuantitativos de egresados y extendiendo a los registros cualitativos de investigaciones, tesis y productos académicos con posibilidades de impacto institucional y de sistema de educación pública.

Por tanto, la injerencia de atención y coordinación con otras instituciones como CONACYT y el AGN, juega un papel clave para concatenar lineamientos de programas, políticas y experiencia de orden digital documental que enriquecen más las posibilidades de integrar un modelo de participación paulatino de todo el universos de IES-MX. El estudio NDSA-LDP realizado a IES-MX para conocer sus niveles de preservación digital que se integró por instituciones de nivel licenciatura y posgrado a nivel federal y estatal, arrojó datos de consideración y piezas clave para atender actividades concretas de los dos segmentos y niveles estudiados, los cuales otorgan la posibilidad de integrar una visión panorámica sobre la importancia de la preservación digital para otras instituciones Mexicanas y entender los procedimientos necesarios previos, durante y después de la gestión de objetos digitales en repositorios institucionales y su preservación a largo plazo.

Un modelo de preservación digital distribuida de instituciones de educación superior México, no puede ser la realidad de una institución o tecnología aislada, la naturaleza debe radicar en integrar un ADN colaborativo interinstitucional que logre perpetuar la identidad nacional digital de alto valor que proporcionan los ejes de la formación superior pública de México, y

que van más allá de soluciones técnicas o de organización, y se que es necesario plantear como el compromiso de preservar el patrimonio académico digital de México.

#### 4.6 Integración e implementación de modelo de preservación digital distribuida IES México (PDIIM)

El planteamiento general de la utilidad de un modelo de preservación digital distribuida para IES en México tiene una respuesta coherente que enfatiza en los beneficios de la prevención digital sostenible en escalas distribuidas, y es que no sólo anticipa la corrección de experiencias puntuales sobre el uso, gestión y pérdida de documentos e información divergente de patrimonio nacional desde el pasado prehispánico, colonial, independiente, revolucionario o contemporáneo; si no que afianza y complementa las acciones de la preservación análoga y digital con una capa de gestión de objetos digitales en múltiples ubicaciones de acuerdo y en vista de una futura integración de intereses comunes en la agenda digital tanto del modelo educativo, tecnológico y de archivo del que son partícipes diversas instituciones en México. Colocar las piezas y posibilidades de crear una puesta en común con objetivos compartidos de preservar de manera colaborativa los acervos digitales en intereses comunes de IES a largo plazo.

##### 4.6.1 La escalera al modelo de preservación digital distribuida de IES México (PDDIM)



Figura 16. Escalera de acciones para modelo PDDIM

La definición y propuesta de un modelo de preservación digital distribuida para IES México no puede considerarse como un accesorio genérico de integración técnica, ni mucho menos una acción de simple decisión. Tal como hemos estudiado, revisado y analizado se requiere de un amplio y detallado sistema de organización que debe ser validado bajo el cumplimiento de requisitos previos que puedan ser integrados en un modelo de trabajo con fines de preservación digital, por lo que requiere de manera obligada observar y atender a la panorámica que nos ofrece la siguiente escalera de acciones con orden ascendente, la cual correlaciona diversas etapas previas a la integración de la actividad plena de preservación digital distribuida.

### ***Escalón 1.1: Definición de objetivos de gestión de documentos de investigación institucional***

Toda institución debe conocer cuál es el devenir y porvenir de los documentos de investigación que produce. Este conocimiento debe estar fundado en el establecimiento de objetivos y sistemas de organización de documentos e información documental. La organización de la información es necesaria entre otras razones, para evidenciar los hechos y conservar los documentos (Font et al., 2012). La gestión documental en instituciones de educación superior es una actividad madura que cuenta con aplicaciones a nivel administrativo, financiero, académico, bibliotecario, entre otros. Para el ámbito de documentos de investigación, es necesario establecer objetivos y políticas específicas que definan la importancia de su gestión, organización y recuperación bajo la mirada de lograr cumplir objetivos de su creación, los cuales se concentran en el descubrimiento, la difusión y divulgación de conocimiento como resultado de la investigación, así mismo procurar la gestión organizada y cuidado de uno de los más valiosos productos de identidad que una institución de educación superior tiene como producto de su trabajo a nivel social y global, su investigación institucional.

La definición de objetivos que dictan la importancia de la gestión de documentos de investigación institucional, es el primer requisito recomendado y que toda IES-MX debe

contener ante toda intención de integración, depósito y participación colaborativa de las que haga parte a sus documentos de investigación.

### ***Escalón 1.2: Normalización y estandarización institucional de producción digital de investigación***

Bajo el orden de establecer objetivos de uso y gestión de documentos con fines de investigación dentro del escalón inicial de acciones, es importante que estos se relacionan con procedimientos secundarios que tienen por objetivo el normalizar el fondo semántico y la forma sintáctica a nivel digital. Esta definición de parámetros deberá ser estipulada bajo la definición institucional y una serie de estándares de representación y esquemas de contenidos a nivel de documentos digitales. El fondo semántico de un documento de investigación representa todo el contenido que en él se incluye, y el cual debe contar con una línea de organización de contenidos bajo la definición de los elementos significativos necesarios según sea el tipo de documento de investigación, que en el caso de los artículos o publicaciones científicas, la organización de la presentación de contenidos se definiría con elementos como lo son el título, tabla de contenidos, introducción, metodología, resultados y discusión, así como las conclusiones y referencias bibliográficas. En esta línea el fondo sintáctico definirá las estructuras o plantillas institucionales necesarias que sirvan y orienten a la producción de documentos a la integración de fondo y forma en documentos de investigación institucional como lo son las especificaciones técnicas de estructura de un documento digital.

Todo el material digital documental con fines de investigación producido dentro de una IES-MX debe revisar los estándares nacionales y mundiales en los que se acoge la definición más adecuada y pretensiones de participación, para de esta manera definir por igual los elementos de orden técnico de representación digital como lo son el formato digital y los correspondientes metadatos de contexto que garanticen su calidad, usabilidad y accesibilidad de representación digital. Todas estas acciones tienen por objetivo el corregir la heterogeneidad digital de la tipología documental nacida digitalmente, para mantener la producción digital de investigación en un camino homogéneo e integrado a nivel institucional y con vista a el tratamiento futuro de los documentos digitales, con el fin de mantener una línea de producción bajo un estándar que facilite la normalización documental institucional digital.

### ***Escalón 2.1: Políticas de gestión documental, difusión y visibilidad digital***

El segundo escalón protagonista de la escalera inicia con la formalización de los objetivos de gestión documental y normalización documental digital en un documento técnico institucional denominado políticas de gestión documental. En ellos se debe contemplar además de las directrices, estándares, roles y esquemas de metadatos documentales, los alcances y acciones necesarias para la correcta gestión de la difusión y visibilidad digital de los documentos de investigación como objetivo de impacto institucional. Toda política de gestión documental con orden digital, integra diferentes tipos de objetivos dependiendo de los sectores en donde se desarrolle o aplique, en este sentido si ponemos en perspectiva el ejemplo de las políticas digitales de identidad e imagen comercial de una empresa, tanto los fabricantes, las marcas, los productos y los consumidores participan como elementos concatenados e indispensables para las consideraciones del establecimiento de las políticas comerciales digitales. donde el objetivo formalizado en el plan de políticas y estrategias de alcance digital repercutirá directamente en resultados de reputación, ventas, posicionamiento, hábitos, etc. En este programa, no solo un elemento de la cadena se verá afectado, sino este resultado afectará en todos los elementos que la integran.

En el ámbito de las instituciones de educación superior en México y otras latitudes mundiales las políticas de orden digital y de gestión documental con fines de difusión, visibilidad y preservación digital, aún sufren de una baja y paulatina integración como documento estratégico y necesario (Sheldon, 2014). Si transportamos el papel de una estrategia comercial al sector institucional, podríamos colocarla dentro de las coordenadas de las actividades relacionadas con el *neoinstitucionalismo*, donde la productividad y alcances económicos de visibilidad e impacto afianzan la valoración de las mismas.

Las figuras de la institución, investigadores e investigaciones, necesitan formalizar su triplete institucional a nivel de políticas, lineamientos y criterios con objetivos de gestión digital, donde se estipule claramente los objetivos de trabajo colaborativo, compromisos, beneficios de su correcta integración y futura interoperabilidad con sistemas electrónicos o informáticos especializados. Las políticas de gestión documental, difusión y visibilidad son un requisito obligado antes de realizar cualquier implementación tecnológica de sistemas de gestión

electrónica o repositorios digitales, ya que con una guía formal que estipula que se quiere lograr, donde, cómo y bajo qué criterios, el siguiente paso del escalón que define la elección e integración del vehículo o sistema informático más adecuado para conseguir que las acciones de gestión, difusión y visibilidad digital de los documentos de investigación repercutirá en una integración sólida de base formal e institucional de políticas digitales.

### ***Escalón 2.2: Definición e integración de sistemas de gestión electrónica y repositorios digitales***

La actividad de definir la integración de un sistema de gestión electrónica o un repositorio digital se correlaciona directamente con la integración de las acciones previas, consolidadas con base en la organización, toma de decisiones y formalización de las mismas a modo de políticas. Es importante poner en claro que existen diferentes denominaciones y definiciones de un sistema de gestión electrónica, ya que algunos autores lo definen también como repositorios digitales o sistemas de gestión documental y depósito digital. Los sistemas de gestión electrónica de documentos, sirven para automatizar diversos procesos de la cadena documental, desde la entrada de la información, hasta la reproducción del documento original, pasando por su recuperación y consulta (Codina, 1992).

Por otro lado los SGE se definen como herramientas integradas que ofrecen la capacidad de almacenamiento, organización, descripción y difusión de colecciones (Márdero, 2015). En el ámbito bibliotecario, el término repositorio que proviene de la terminología anglosajona “*repository*”, la cual cada vez gana más aceptación en el terreno de la gestión documental electrónica o digital, en donde la diferencia a modo práctico radica en que el término repositorio resume la actividad de gestión electrónica.

En el caso de aplicación de las IES-MX podemos observar que en los lineamientos propuestos por el Consejo Nacional de Ciencia y Tecnología se hace uso del término repositorio institucional y nacional, por lo que su identificación deberá atenderse como tal para evitar confusiones. En el mercado actualmente existe una gran cantidad de soluciones en sistemas y aplicaciones de repositorios para la gestión documental digital, como ejemplo podemos nombrar a DSpace, Eprints, Archivematica, ICA-AtoM, Koha, Invenio, Kora, Content dm, etc. Por lo que la definición e integración de un sistema de repositorio digital, responde tanto a

efectos técnicos como funcionales que se destacarán por las posibilidades de gestión, organización, descripción, acceso, almacenamiento e interoperabilidad de los documentos digitales, lo cual viene como consecuencia de factores de tratamiento técnico en los documentos y colecciones a nivel de formatos, descripciones de metadatos y su peso digital.

Es muy importante revisar las prestaciones con las que un sistema de repositorio digital cuenta para que estos puedan ser herramientas de difusión pública ya sea en acceso abierto o privado, así mismo que puedan integrar la trazabilidad de los documentos digitales según sus descargas y consultas en el mismo repositorio como en fuentes externas, ya que ante la diversidad de soluciones de repositorios recolectores de metadatos y documentos como por ejemplo Google Academics, estos se sirven de la existencia de repositorios en abierto y de sus documentos normalizados para realizar la indexación de los mismos en sus buscadores, acelerando y al mismo tiempo desviando a modo replicado el origen institucional del documento. A nivel de visibilidad institucional la adaptación a propósito o indirecta a los modelos actuales de meta buscadores especializados como los de Google, pueden incrementar la potencialidad de visibilidad de los documentos de investigación siempre y cuando las características institucionales de los documentos se encuentren bien integradas, por lo que la fórmula de contar con documentos bien identificados bajo un sólido esquema de organización, estructuras de fondo y forma reflejadas en políticas institucionales y la adecuada selección una solución de sistemas de repositorios, repercutirá integralmente en la identidad, accesibilidad y visibilidad de la investigación institucional.

### ***Escalón 3.1: Políticas de gestión y preservación digital a largo plazo***

El tercer escalón para definir un modelo de PDDIM debe tener como antecedente obligado una sólida infraestructura de organización y gestión digital de documentos institucionales de investigación, a partir de ello se tiene que responder a la pregunta: ¿Que documentos digitales son los más importantes para preservar a largo plazo y cómo debemos realizarlo?. Esta pregunta se contestara de manera formal con la integración de las políticas de gestión y preservación digital a largo plazo, las cuales de manera deseable se representarán en un documento con especificaciones técnicas y decisiones concretas sobre la importancia de valor y tipo de documentos que serán considerados como patrimonio institucional para su preservación a largo plazo.

En la políticas quedarán documentadas en planes de preservación que definan las áreas de preservación específicas de colecciones, sus formatos de objeto y esquema de metadatos. El plan de preservación definirá por área los criterios técnicos de las políticas como lo son los nombres de los documentos, las estructuras de archivo o carpetas, formatos de los objetos, tipos de control de ingesta, cantidad de copias, transformación de formatos y auditoría que cumplan y sopesen los intereses y necesidades de cada IES. El nivel más básico de decisiones funcionales de importancia de valor y prioridad de documentos a nivel de memorias digitales, es ofrecido por The National Digital Information Infrastructure and Preservation Program como iniciativa colaborativa de The Library of Congress de los Estados Unidos, el cual se basa en lo que podemos nombrar como el esquema IDEOS (identify, decide, export, organize and storage) o IDOM (identify, decide, organize and make copies) y que sugiere un ruta doméstica para la preservación de diferentes tipos de archivos digitales como lo son el audio, video, fotos y documentos. A nivel de IES el esquema IDEOS e IDOM, y que sugiere un ruta doméstica para la preservación de diferentes tipos de archivos digitales como lo son el audio, video, fotos y documentos. A nivel de IES el esquema IDEOS puede contribuir a realizar una prematura aproximación a decidir y distinguir funcionalmente sobre las prioridades de tipología de documentos y colecciones de valor documental con vistas al largo plazo, para posteriormente ser definido formalmente a nivel de políticas de preservación digital a largo plazo con criterios contrastados, revisados y acordados en equilibrio con los objetivos de la institución.

Las acciones necesarias a nivel de organización, roles, tecnología, beneficios y sostenibilidad de una estrategia de este tipo, también pueden integrarse en el documento de políticas o en un apartado de plan de preservación de áreas específicas. Los criterios o recomendaciones de acciones ofrecidas en Levels of Digital Preservation de la National Digital Stewardship Alliance y los resultados del estudio aplicado a IES-MX de cara a la integración de un modelo colaborativo de preservación digital, deberán considerarse de manera individual por cada IES-MX, a fin de que los criterios pueden revisarse en constancia y conocer el alcance de las prácticas de cada institución y así enriquecer la la integración homogénea futura de un modelo de preservación digital distribuido en colaboración con otras IES-MX.

### ***Escalón 3.2: Definición e integración de sistemas de preservación digital***

La integración, selección y definición de uso de un sistema de preservación digital dentro de una estrategia planificada es un procedimiento complejo, ya que requiere sopesar diferentes factores a nivel técnico como de organización en un esquema centralizado o de pertenencia de una institución. Tal como lo hemos anotado anteriormente, no existen trajes a la medida en lo que se refiere a los sistemas de preservación digital, ya que estos aunque pueden ofrecer un gran paquete de aplicaciones o soluciones “*todo en uno*”, siempre será necesario realizar adaptaciones que correspondan a las necesidades de cada organización, ya que e incluso habrán capacidades o prestaciones de los sistemas que puedan quedar sobrando o todo lo contrario.

La definición de integrar un sistema de preservación digital es una tarea muy importante pues necesita lograr la conjunción perfecta de las ventajas de prestaciones a nivel técnico de software, las necesidades de orden operativo de hardware y las posibilidades de organización sostenida de una institución. A nivel de prestaciones de software, puntualmente se debe cumplir con ofrecer procesos completos de ingesta digital que incluya la limpieza o filtro de objetos ingestados, la transformación enriquecida y derivaciones, tratamiento de metadatos, validaciones de integridad de objetos, almacenaje seguro y auditoría programada. El proceso de recuperación de los objetos archivos o colecciones digitales deberá permitir una realización individual o múltiple sopesando con ello la velocidad de la recuperación de acuerdo al hardware y tipo de almacenamiento. Los procesos de auditoría deberán de ir de la mano de prestaciones de monitores de obsolescencia y de migración digital, lo cual cierra el círculo de funciones generales e indispensables a nivel de software sistema.

A nivel del hardware (ordenadores, servidores, sistemas de almacenamiento) necesario para que un sistema de preservación digital funcione con garantía y seguridad, encontraremos diferentes opciones que en algunos casos dependen directamente de los sistemas de preservación que se deberán revisar ampliamente los requisitos técnicos de los sistemas, ya que los diferentes modelos de operación que validan el uso de sistemas de almacenamiento o servidores (NAS, SAN, DAS, Cloud, etc.) genéricos compatibles, en algunos casos requieren de una extensión de *drivers* a software, e incluso de hardware propietario para el funcionamiento pleno de los sistemas de preservación digital. En lo que refiere a las

posibilidades de la organización y la sostenibilidad del uso de un sistema de preservación digital, deberán ser revisadas las circunstancias económicas para sostener un sistema de preservación en una IES, ya que en el caso de los sistemas comerciales (que al día de hoy ofrecen las soluciones más completas del mercado), existen diferentes modelos comerciales que se componen por la compra del sistema completo, o por sistemas de suscripción o alquiler bajo programas de mantenimiento, asesoría y capacitación. En algunos casos, el acuerdo de uso dependerá de los límites estipulados en el sistema sobre la capacidad límite de ingesta de objetos o archivos, la cual supone un gasto extra para ampliar la licencia de capacidad. En el caso de los sistema de código libre debe considerarse por igual los costos beneficios que implica su implementación, adecuación, capacitación y mantenimiento de los mismos. Todos los aspectos en conjunto deberán tomarse en cuenta ante el compromiso de necesidades de preservación a largo plazo e inversión del sistema más adecuado.

#### ***Escalón 4.1: Políticas de colaboración interinstitucional para preservación digital a largo plazo***

El inicio del cuarto escalón hacia un modelo de preservación distribuida IES México comienza con el balance de las actividades y alcances que ofrece una estrategia de preservación digital a largo plazo en modo centralizado, el cual de manera anticipada refleja que en la operación y procesos de implementación de una estrategia centralizada se asume consigo la responsabilidad institucional total a nivel operativo, infraestructuras y recursos de sostenibilidad a largo plazo de su estrategia de preservación digital. Los resultados del balance de cara a la sostenibilidad a largo plazo y opciones referentes a mejorar la viabilidad de asegurar la longevidad de un programa de preservación digital, abren las posibilidades a el planteamiento de una estrategia de orden interinstitucional o de participación múltiple, que proporcione soluciones estratificadas de acuerdo a las necesidades y requerimientos de cada IES mediante una serie de criterios de participación, inversión y responsabilidad compartida de un grupo de instituciones con objetivos de preservación digital a largo plazo en modo colaborativo y distribuido. En este índice de actuaciones, se debe formalizar un documento que estipule los diferentes criterios, acciones y actividades necesarias para que un grupo de IES pueda colocar bajo un mismo fin su cooperación interinstitucional.

Es necesario estipular acciones a nivel técnico, de organización y coordinación que definen puntualmente los beneficios y alcances de cada una de las instituciones dentro de una estrategia de preservación digital distribuida. Una guía de integración que funciona como base es el documento de preservación digital centralizado, el cual al complementarse con los indicadores y elementos de organización resultantes del estudio OCMC, los cuales ocupan y organizan a temas como: la organización y gobernanza, colaboración y preservación distribuida (criterios técnicos), marco legal y políticas de trabajo, coordinación (técnica e institucional) y comunicación. Desarrollando a profundidad estos indicadores con base en acuerdos y alianzas de colaboración cooperativa necesarios, se podrá estratificar un documento de orden formal con el consentimiento de todas las instituciones participantes materializándose en criterios y lineamientos de operación interinstitucional. Este documento será indispensable para la definición de uso de un sistema de preservación digital distribuida y de otras acciones necesarias a nivel técnico, así mismo de las acciones necesarias para compaginar la organización interinstitucional para consolidar un modelo PDDIM.

#### ***Escalón 4.2: Preservación digital distribuida IES México***

Hasta este punto podemos observar que el modelo de PDDIM es el destino final de la conjunción de acciones previas y componentes de tipo técnico e infraestructura (sistemas informáticos, protocolos, hardware y recursos económicos), actividades de organización (coordinación, cooperación y alianzas) y necesidades específicas de preservación digital de un grupo de instituciones (nivel de alcance NDSA-LDP). Todas estas determinaciones deben concatenarse de forma ascendente y ser reflejadas de manera formal en un plan de preservación digital distribuida que integre a su vez la estrategia de preservación digital que puntualice en la participación, responsabilidad y compromiso interinstitucional. El plan y la estrategia de preservación digital distribuida integran el modelo PDDIM, ya que enlazan y sopesan a todos los escalones PDDIM para conocer los niveles (escalones) donde se sitúan las instituciones, y así poder realizar los planteamientos correspondientes de acuerdo a sus alcances y necesidades para situarlas a el tipo de infraestructura, relaciones y participación interinstitucional óptimo para cada una de ellas.

La escalera representa por un lado el camino y las acciones ascendentes necesarias, y por otro lado el momento y visión de situación en la que cada institución se encuentra ante una

estrategia planificada de preservación digital y su participación en modo distribuido. La inexistencia de trajes a la medida en sistemas, estrategias o planes de preservación digital distribuida para todo un grupo de instituciones diversas, obliga a centrarse en los diferentes ámbitos de aplicación del modelo PDDIM, ya de acuerdo con los resultados del estudio NDSA LDP a IES-MX (capítulo 3), no existe una estrategia formal de preservación digital en consenso y acuerdo interinstitucional, además de que existen diferencias abismales entre instituciones por factores como el subsistema de pertenencia, tamaño de infraestructura institucional, prestaciones tecnológicas, normalización de prácticas de preservación y previas, etc. Estas situaciones nos obligan a considerar los ámbitos de aplicación desde un punto de vista regional, del estado y a nivel nacional, para trazar una hoja de ruta sobre cómo integrar grupos uniformes y sólidos.

Conocer los alcances de aplicación y definir estructuras de integración dentro del modelo PDDIM será vital para lograr que la participación de todas las instituciones encuentre un lugar adecuado, balanceado y sostenible tanto por las cantidades y tipos de contenidos digitales, políticas y marco legal, infraestructuras actuales, organización, normalización, etc. Estos son factores determinantes para su participación en conjunto con la selección de la herramienta informática que dé soporte a la organización y conexión de los distintos participantes bajo el protocolo distribuido.

Reiteramos que no se debe confundir la selección e implementación de un sistema de preservación digital distribuida como una actividad aislada y única, ya que requiere de completar todos los escalones anteriores de manera previa, los alcances de aplicación y las estructuras de integración, para así poder elegir la herramienta que más compagine con la estrategia planificada y coordinación del modelo de preservación distribuido.

#### **4.6.2 Ámbitos, estructuras y coordinación de modelo PDDIM**

Para representar una propuesta integral del modelo PDDIM sobre la base de antecedentes y estudios específicos sobre las principales funciones y técnicas aplicadas en los sistemas de PDD, buenas prácticas de modelos referentes en cuanto a su organización y coordinación como redes institucionales distribuidas y el estado de nivel de necesidades de las IES en México ante actividades de preservación digital, debemos integrar, sumar y considerar a su

vez, la diversidad y cantidad de IES-MX que representan la totalidad del universo de educación superior pública en México. La alta heterogeneidad institucional a nivel de subsistema (federal, estatal, tecnológico, etc.), tipología (universidad, tecnológico, etc), infraestructura (pequeña, mediana) e instituciones de influencia en IES-MX (nacionales, regionales, etc.), contribuyen a ser piezas fundamentales para agudizar de forma objetiva la representación integral y equilibrada del mismo. Se deben complementar de manera recomendada con tres perspectivas que influyen directamente a la visión del modelo PDDIM como lo son: el ámbito de aplicación de la estrategia de preservación (municipal, estatal, nacional), las estructuras de integración y de participación (uniforme e integral) y las paridades necesarias de las IES-MX para estimar una integración funcional y sostenida de un modelo interinstitucional PDDIM.

La definición de la visión de ámbito de alcance y aplicación de la estrategia de preservación digital en modo distribuido, refiere y responde a diferentes motivos. El primer motivo es la diversidad geográfica y tipológica que completa el mapa de las IES-MX. Por otro lado a la importancia de alcance equilibrado de una estrategia de preservación, ya que esta puede colocarse solo con cobertura regional o en su caso nacional, y en ambos casos mantendrán diferencias de protocolos, organización y coordinación que deben ser expuestos de forma clara para crear un ejercicio de cordura sostenible en equilibrio con un modelo interinstitucional de alta diversidad.

En el terreno de las estructuras de integración la primer estructura define una propuesta con integrantes similares en características generales de prestaciones, infraestructura y contenidos que denominamos estructura uniforme, ya que uno de los principales problemas y desventajas ocurrentes en los planteamientos de modelos de redes interinstitucionales de preservación digital, es el contar con participantes con un alto grado de disparidad en cuanto a sus responsabilidades, infraestructuras (tecnológicas y de organización) y cantidad de contenidos digitales en archivo a compartir con fines de preservación. A modo metafórico, si una institución cuenta con un archivo digital del tamaño de un castillo y otra del tamaño de una habitación, los esfuerzos de operación y sostenibilidad a largo plazo en algún momento encontrarán un punto de quiebra e inestabilidad. Los nodos que integran una red institucional que busca su sostenibilidad a largo plazo debe no sólo considerar su actitud positiva de compromiso de participación para colaborar con otras instituciones bajo organización,

contenidos y tecnológica compartida, sino que además debe ser capaz de balancear sus necesidades y alcances reales como institución en un punto más cercano a el similar que al lejano de otras en la misma red, para así crear eslabones de tipo sólido que vayan a la par del crecimiento gradual de los contenidos digitales de las mismas.

La segunda estructura contiene una propuesta de modelo integral que incluye a todos los integrantes variables de modelos uniformes dentro de una mapa equitativo tanto en prestaciones, infraestructura y contenidos. Las estructuras uniformes que definen la composición de un modelo integral, a su vez se relacionan con los diferentes ámbitos de alcance y composición, ya que si se considera una visión holística e integradora de todos los universos de IES-MX, la composición deberá considerar su comienzo desde un ámbito regional o municipal para una vez integrado poder colocar un eslabón sólido a nivel estatal que coordine con el modelo de ámbito nacional o el nivel de modelo integral PDDIM.

La visión integradora de grupos uniformes dentro una estructura con modelo integral de participación, requiere necesariamente de una figura de coordinación que no solo centralice las operaciones y gestiones de cada uno de los miembros o grupos de la red de preservación digital distribuida de IES México, sino que además facilitara la designación de roles necesarios para que los grupos puedan considerar subcapas de coordinación que faciliten los procesos de integración y normalización de instituciones con diferencias de adopción de protocolos tecnológicos de preservación digital. Para la integración de los ámbitos, estructuras y coordinación del modelo PDDIM integral, se hace uso de los diferentes elementos, indicadores, funciones y resultados de buenas prácticas previamente definidos según abstracciones conceptuales de análisis de modelos internacionales y correlaciones de las prácticas desarrolladas con este fin.

a) *Ámbito de aplicación y alcance de IES para modelo PDDIM*

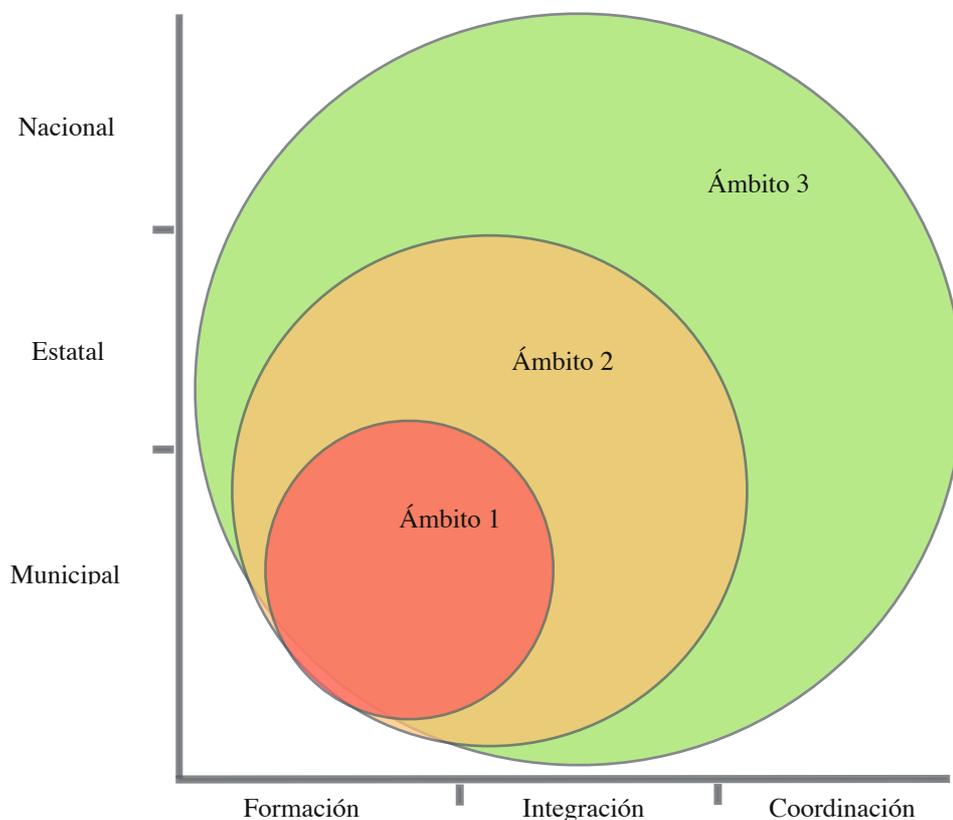


Figura 17. Ámbitos de aplicación y alcance de estrategia de preservación en modelo PDIIM

Para integrar a grupos de instituciones con características compatibles dentro de un modelo de organización con fines de preservación digital distribuida es importante conocer los ámbitos de aplicación y alcance de participación en dependencia del sitio geográfico o país en relación a la composición institucional de IES, objetivos y formación de preservación digital. La estrategia de preservación distribuida necesita un esquema de conocimiento de alcance para considerar y entender la integración de instituciones a estructuras de grupos uniformes para ser agregadas dentro de un modelo integral interinstitucional de preservación digital distribuido. La definición de alcance de la estrategia es primordial y está ligada a la definición de estructuras de grupos de instituciones (uniformes e integrales) y a la coordinación de los objetivos de colaboración en modo distribuido. A continuación describimos la propuesta de

los tres ámbitos esenciales para IES en México y recomendaciones a tomar en cuenta para definir un ámbito de alcance operativo y de posibilidades de integración.

### ***Ambito Municipal - Zona***

Este ámbito refiere y define a las zonas específicas de alcance municipal de un estado que cuentan con IES-MX de subsistema universitario, tecnológico o centros de investigación. Para la integración del ámbito municipal a la estrategia de preservación digital distribuida del modelo PDDIM, se atiende a los resultados por subsistema e individuales del estudio de IES-MX sobre los niveles de actividades de preservación digital y alcance de las mismas, ya que estos resultados definen que IES que lideran en estas actividades por municipio o zona de municipios colindantes, así mismo nos permite conocer en un nivel específico cuales son las IES que realizan o requieren implementar actividades de regulación, formación y apoyo inicial. Dichas observaciones ayudan a definir un esquema inicial formativo de trabajo de ámbito municipal con objetivos de preservación digital en modo distribuido bajo las siguientes recomendaciones.

1. Filtrar, especificar y asignar a las IES-MX que lideran la adopción de acciones de preservación digital en su municipio el rol de mentores municipales de otras instituciones para establecer protocolos de diagnóstico, formación y divulgación seminal sobre las actividades de preservación digital y crear las bases de adopción de actividades de la escalera PDDIM con el objetivo de preparar a las IES a un rango de posibilidades de integración de grupos normalizados sobre actividades de preservación digital.
2. En el ámbito municipal se deberán enfocar los esfuerzos de formación a cubrir de manera íntegra las dos actividades del primer escalón de la escalera PDDIM como lo es la definición de los objetivos de gestión de documentos de investigación institucional, así como la normalización y estandarización institucional de la producción digital de investigación.
3. Podríamos denominar a este ámbito como de divulgación, formación y maduración ya que su objetivo es el de dar conocer, emparejar e implantar acciones y actividades base de

la escalera PDDIM de forma gradual y controlada, así mismo de iniciar el ADN de cooperación de la mano de instituciones punteras en la adopción del tema.

Como ejemplo aplicado podemos nombrar a los municipios colindantes de la zona conurbada de Tampico, Cd. Madero y Altamira, las cuales dentro de su región sur agrupan a IES públicas como la Universidad Autónoma de Tamaulipas Campus Sur, Centro de Rehabilitación y Educación Especial (CREE), Instituto Tecnológico de Ciudad Madero, la Universidad Pedagógica Nacional (UPN), Instituto Tecnológico de Altamira y la Universidad Politécnica de Altamira. Solo una de ellas dentro de los resultados del estudio de IES-MX realiza actividades de niveles básicos de preservación digital, por lo que esta IES podría actuar como institución mentora de la región.



Figura 18. Ámbito municipal de IES México para modelo PDDIM

## ***Ámbito Estatal - Regional***

Una vez establecido el ámbito municipal con actividades seminales y de mentorías, las posibilidades de formar grupos coherentes sobre la importancia, compromiso y gestión de actividades básicas de preservación digital bajo una filosofía cooperativa de colaboración nos dirige al siguiente eslabón de alcance que es el de ámbito estatal, el cual se compone por la representación de la constitución geográfica de México en estados federales por 31 estados federales y la Ciudad de México como capital como número 32. Como base en la visión constitucional geográfica definimos una serie de recomendaciones, actividades y responsabilidades a tomar en cuenta para lograr afianzar la integración de grupos uniformes normalizados o en vías de hacerlo dentro de un ámbito de influencia estatal.

1. El ámbito estatal - regional tiene por objetivo la realización de actividades centradas en unificar, agrupar y coordinar a los grupos regionales formados por mentores e instituciones en formación que serán parte de la estructura o nodo estatal de preservación digital.
2. Para las actividades de unificación y coordinación del ámbito estatal - regional se debe contar con una institución representante estatal o regional con el rol de coordinador, que valide las acciones de las instituciones mentor e instituciones en formación regional sobre sus alcances y actividades centradas en los escalones centrales (2 y 3) de la escalera de preservación, los cuales se enfocan en la definición de políticas de gestión documental, difusión y visibilidad digital. La definición e integración de los sistemas de gestión electrónica y repositorios digitales, que en este caso tienen relación directa con los lineamientos de repositorios institucionales de CONACYT, deberá ser compatible con la creación grupos estatales uniformes y normalizados en los requisitos de prestaciones necesarias, esto para cumplir plenamente con el fin de su rol en la estrategia de preservación distribuida.
3. En este ámbito estatal - regional posteriormente se debe considerar y designar qué institución será la encargada de centralizar las actividades y sistemas de gestión electrónica de los grupos para integrarlos como figura de institución coordinadora posteriormente y cubriendo todos los requisitos de su ámbito de aplicación y alcance al

nodo estatal de preservación digital distribuida, el cual tiene por objetivo conectar a los grupos de IES del estado con el ámbito nacional.

Hay que señalar que el ámbito estatal - regional acoge una gran responsabilidad de selección y definición de participantes según sea la pertinencia, cantidad y calidad de las IES representadas dentro de un estado para la estrategia y modelo PDDIM, ya que cuando esta representación no sea suficiente por la cantidad de participantes o la superficie del estado, se debe considerar la creación de un ámbito estatal compartido que pueda definir dos estados o más dentro de una zona regional en ejecución. Estas recomendaciones se notan en cuenta para hacer más eficiente la estrategia colindante o en su caso crear grupos de mentores por zonas cardinales de los mismos (sur, norte, este, oeste, peninsular, etc.) con estructuras similares.

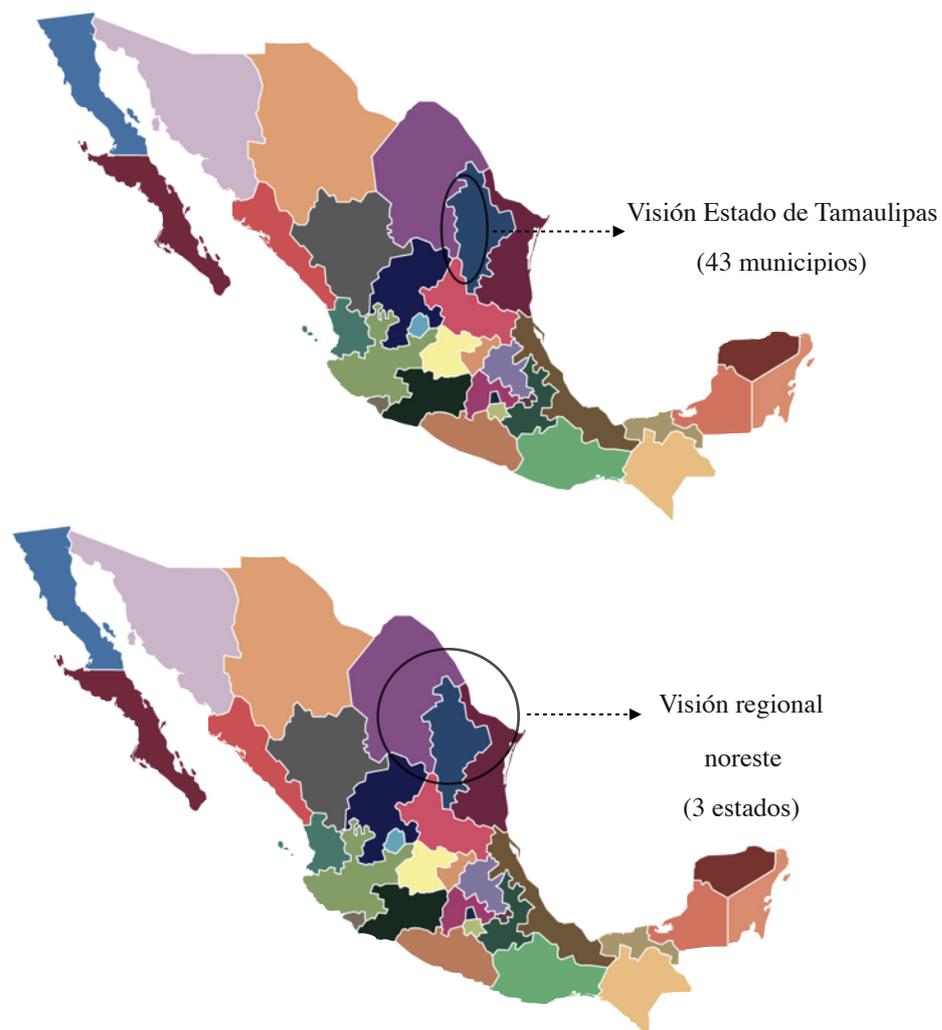


Figura 19. Alcance de ámbito estatal y regional de IES México para modelo PDDIM



Figura 20. Alcance de ámbito regional zona de IES México para modelo PDDIM

### ***Ámbito Nacional***

El ámbito que representa el objetivo final de los diferentes alcances de ámbitos en la estrategia de preservación digital en modo distribuido. Este será el paraguas que acoge a todos los nodos estatales a modo de nodo central y coordinador. En él se realizarán diferentes actividades, gestiones y toma de decisiones de gobernabilidad de la totalidad de la red del modelo PDDIM.

1. El ámbito nacional tiene por objetivo principal la operación general nacional, así como la gobernabilidad bajo políticas de colaboración y relaciones interinstitucionales de los nodos e IES integrantes del modelo PDDIM. Este ámbito es el responsable de supervisar y conectar que todas las IES participantes o prospectos a formar parte de la red PDDIM cumplan con todos los escalones establecidos en los ámbitos municipal y estatal regional, así tal el ámbito nacional está encargado de dar plenitud a las actividades del escalón número cuatro que se enfocan a la definición de las políticas de colaboración interinstitucional para preservación digital a largo plazo, así como de concretar el plan del modelo de preservación digital distribuida IES México a nivel funcional y organizacional.

2. A nivel funcional deberá definir el sistema y la plataforma de infraestructura tecnológica con la que se debe contar tanto en el ámbito municipal, estatal y nacional en las IES participantes, IES mentora, IES nodos estatales y nodos centrales nacionales.
3. El ámbito nacional es responsable de la organización que define la visión de estrategia nacional de preservación digital en modo distribuido, por lo que será necesario estipular en este ámbito la figura de institución coordinadora de la red de PDDIM.
4. En la responsabilidad de organizar y coordinar la colaboración de IES-MX en un modelo de PDDIM, se definirá a nivel nacional con ayuda y retroalimentación de los representantes estatales y municipales a estructurar las políticas de preservación digital por niveles, de relación interinstitucional por ámbitos y de la vinculación con la red nacional de PDDIM.

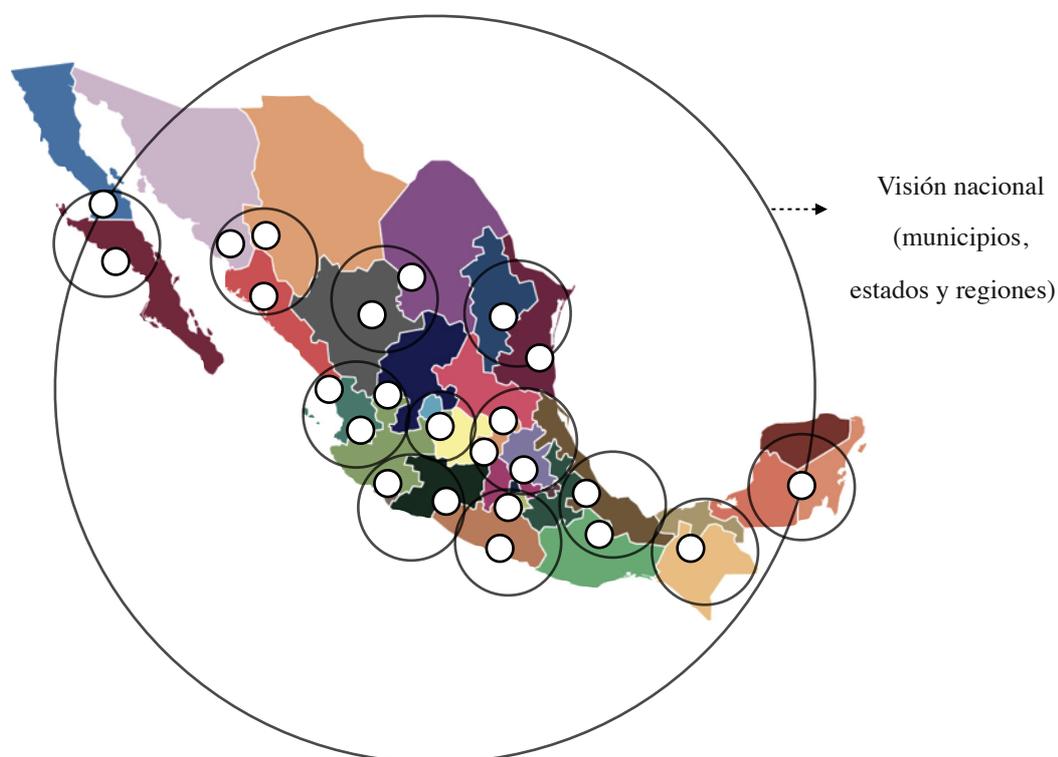
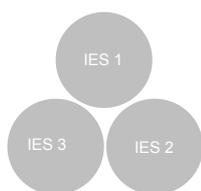


Figura 21. Alcance de ámbito nacional de IES México para modelo PDDIM

Como lo hemos dicho antes, la definición de ámbitos ayuda a visualizar el eje de alcance organizacional de integrantes, posibilidades y necesidades de IES-MX que pueden formar parte de la estrategia de preservación dentro del modelo PDDIM. La definición las estructuras de participantes a nivel uniforme e integral, ayuda a asociar los ámbitos municipales, estatales regionales y nacional de forma estructurada y coherente para el fin de la estrategia de preservación distribuida.

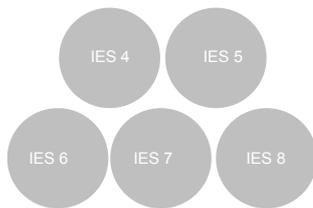
### ***b) Estructura uniforme e integral para modelo PDDIM***

Una de las situaciones que anticipa la creación de una red de preservación distribuida, es que sus integrantes gocen de un grado uniformidad en características y funciones necesarias que garanticen una participación institucional equilibrada a nivel de infraestructura institucional, cantidad y tipología de contenidos, así como de servicios tecnológicos y prestaciones. La estructura uniforme para el modelo PDDIM acoge a instituciones del mismo tamaño e ímpetu operativo ante objetivos distribuidos, por lo que el resultado de crear y relacionar estructuras uniformes tiene como resultado la integración de equilibrada de participantes dentro de un modelo de trabajo. Las características con las que deben gozar los integrantes del modelo uniforme se enfocan a contar con el mismo tamaño o similar en infraestructura (personal, alumnos, profesores, etc), misma o similar cantidad y tipo de contenidos digitales (artículos, preprints, tesis, etc.), mismas o similares prestaciones y servicios tecnológicos (repositorio, almacenamiento digital, BiDi, etc.) e IES con el mismo o similar alcance de actividades de la escalera PDDIM (normalización). A continuación definimos tres ejemplos de estructuras uniformes en escala.



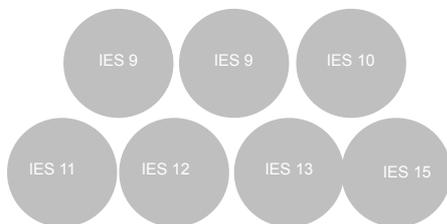
#### **Estructura Uniforme Plena**

- IES con alta infraestructura institucional
- IES con alta cantidad y mismos tipo de contenidos a preservar
- IES con prestaciones y servicios tecnológicos normalizados a plenitud a la escalera PDDIM.



#### **Estructura Uniforme Media**

- IES con media infraestructura institucional
- IES con media cantidad y mismo tipo de contenidos a preservar
- IES con prestaciones y servicios tecnológicos con mediana normalización a escalera PDDIM



#### **Estructura Uniforme Baja**

- IES con baja infraestructura institucional
- IES con baja cantidad y mismo tipo de contenidos a preservar
- IES con prestaciones y servicios tecnológicos con baja normalización a escalera PDDIM

Figura 22. Ejemplos de estructuras uniformes de grupos de participantes para modelo PDDIM.

El modelo de estructura integral se compone y relaciona a diferentes estructuras uniformes de instituciones, donde el enlace de los mismos se define en función del alcance de normalización, prestaciones e infraestructura tecnológica y posibilidades de participación colaborativa. Cada tipo estructura uniforme de acuerdo a su alcance y grado de normalización, podrá integrarse a alguno de los eslabones correspondientes del modelo PDDIM según las figuras de coordinación aplicadas en distintos ámbitos de ocupación.

#### **c) Coordinación de modelo PDDIM**

La integración e importancia funcional de la figura de coordinación en las buenas prácticas de los modelos internacionales de preservación digital distribuida es fundamental. Dentro del modelo PDDIM es necesario designar a una institución que controle, supervise y decida los roles de actuación y responsabilidades interinstitucionales de las IES-MX dentro de la red de

PDDIM. Así mismo, que concrete el protocolo tecnológico más adecuado para las acciones de preservación y la conexión de todos los participantes, por lo que se deben tomar en cuenta las siguientes recomendaciones para definición del modelo.

1. La *figura de responsabilidad institucional para coordinar* a las IES participantes en la red de PDDIM puede corresponder por orden de implicación de manera primaria al Consejo Nacional de Ciencia y Tecnología (CONACYT), ya que en los avances de implementación de lineamientos nacionales de repositorios institucionales y nacional en las IES de México, así como su obligación nacional como promotor y proveedor de infraestructura a programas de investigación en el sistema de instituciones de educación superior, lo colocan como la principal institución nacional con capacidad ejecutiva y de autoridad en el ámbito para dirigir la coordinación del modelo PDDIM a nivel nacional. En un segundo orden de propuesta, la Universidad Nacional Autónoma de México (UNAM), que es la IES de mayor trayectoria, influencia e importancia en el ámbito universitario en México que además cuenta con la experiencia de gestión y coordinación de la Red de Acervos Digitales como un repositorio interinstitucional RAD UNAM que agrupa a las diferentes escuelas y facultades en el mismo, podría ser la IES que basándose en su experiencia y relevancia en la agenda nacional de investigación y educación superior de ser la figura de coordinación institucional de la red PDDIM. Ambas recomendaciones corresponden a fijar el ámbito de coordinación nacional en la flexibilidad de ejecución y especialización de la mismas en el área. Dada la necesidad de plantear además de la coordinación general o nivel nacional, subniveles o subcapas de coordinación de ámbito municipal y estatal, se recomienda establecer un sistema de coordinación municipal y estatal que defina los procesos de selección de IES representantes para realizar las funciones específicas de formación y divulgación a nivel municipal, así como de supervisión e integración al nodo estatal y nacional.
2. Para la *coordinación funcional, sistemática y tecnológica* de la red PDDIM y dado que la propuesta del modelo incluye dentro de su tipología de contenidos a documentos de investigación e IES, consideramos como primer opción el implantar el sistema LOCKSS a nivel de alianza internacional para el funcionamiento de ámbito nacional de la red de PDDIM, ya que además de ser el sistema pionero y líder en el ámbito de la preservación digital distribuida de colecciones digitales de bibliotecas y universidades alrededor del

mundo, mantiene el mismo protocolo con los sistemas de PDD en cooperativas internacionales como es el caso de MetaArchive Cooperative, así tal como el caso y ejemplo de buenas prácticas de una red de preservación digital distribuida de IES en Brasil CARINIANA. A largo plazo ofrece una garantía de interoperabilidad compatible para que la red PDDIM pueda implantarse en coordinación con los lineamientos de repositorios institucionales y nacional (CONACYT, 2015). Así mismo en un futuro pueda ampliar su red de distribución y alianzas tanto a nivel Latinoamérica como internacionalmente con otros sistemas LOCKSS. La opción de implantar un desarrollo a la medida de momento no es compatible con la visión del modelo PDDIM ya que el objetivo principal es el establecer una amplia red que sea compatible por igual con los sistemas de repositorios institucionales y nacional que se implantan actualmente en los programas de IES-MX.

3. La *governabilidad y relaciones* interinstitucionales de la red de PDDIM de acuerdo a la cantidad de IES en probabilidad de integración a él, debe considerar el establecer un modelo de gobernabilidad de la red mayoritariamente de tipo administrativo de red (the network administrative model). A partir de este modelo la institución coordinadora de forma central fungirá como un órgano de administración de la red, participantes y recursos. Tratando los diversos ámbitos de alcance y actividades de composición, podría considerarse por igual el uso del modelo de organización principal (the lead organization model), para una vez que se formen los grupos a nivel municipal, estatal y regional, la organización principal coordine y dé orden a estos grupos descentralizados a la coordinación general. Para ambos modelos de gobernabilidad de la red, será importante designar comités de trabajo (directivo, técnico, contenidos, etc.) con representantes de cada estado e IES, que definan actividades específicas en roles y responsabilidades tanto de la institución coordinadora administradora y los participantes de la red PDDIM.
4. Los *roles, responsabilidades y funciones* de los miembros de la red deberán distribuirse de forma coordinada, equilibrada y con visión flexible de acuerdo con la propuesta de estructuras uniforme e integral y con el modelo de distribución de preservación digital de Eld Zierau (2011). Se recomienda asignar responsabilidades y funciones específicas de preservación digital a los grupos uniformes de IES según sea su alcance e infraestructura, ya que algunas IES pueden fungir sólo como nodos de unidades de almacenamiento

compartido, otras pueden ofrecer solo recursos como servicios y procesos de preservación, así como otras podrán participar con responsabilidades y funciones plenas tanto de almacenamiento, servicios y procesos integrales en la red. Las responsabilidades deben ser bien planteadas de acuerdo a los diagnósticos elaborados en los ámbitos estatales y regionales para definir las composiciones de los grupos participantes en la red PDDIM.

De acuerdo a las recomendaciones de ámbitos, estructuras y coordinación para el modelo PDDIM; a continuación podemos representar las diferentes relaciones de grupos de prospectos y participantes en función a prestaciones de preservación digital distribuida y el rol que desempeñan en el modelo integral PDDIM.

#### **4.6.3 Representación integral de prestaciones, roles y relaciones de modelo PDDIM**

Después de establecer las posibilidades de alcance, ámbitos, estructuras y coordinación funcional del modelo; y con la ayuda de los resultados de alcances y necesidades de IES-MX (capítulo 3); concretamos grupos de relación de IES federales y estatales con estructuras uniformes, así como de instituciones coordinadoras que articulen la diversidad de necesidades, alcances y funcionamiento del modelo PDDIM. Asignaremos prestaciones funcionales de PDD específicas de acuerdo al tipo de necesidades de cada una de las estructuras y sus participantes, para de esta forma y a modo de propuesta flexible, ubicar patrones de grupos uniformes diversos de IES tanto de incorporación progresiva (estructura uniforme baja y media) como directa (estructura uniforme plena) al modelo integral PDDIM.

La representación integral incluye a su vez un mapa de ruta sobre el de ámbito de alcance y tipo relación de participación de nodos institucionales (municipal, estatal y nacional) con ejemplos de integración y posibilidades de distribución. En la representación y composición del modelo PDDIM, encontraremos seis diferentes tipos de relaciones de ámbitos de alcance, participación y prestaciones funcionales de cada una de los grupos e IES participantes, ya que cada uno de ellos, puede representar diferentes roles de actuación tanto a nivel de nodos, como de figuras de coordinación de ámbitos de alcance, e IES con prestaciones funcionales específicas. A continuación reiteramos y describiremos los usos de prestaciones, roles,

instituciones y tipos de relaciones de participación en la representación gráfica del modelo PDDIM.

### ***Prestaciones PDDIM***

Las principales prestaciones funcionales que distinguen a la estrategia de preservación digital distribuida las hemos agrupado por prestaciones funcionales, complementarias y técnicas específicas. Las prestaciones distinguen las cualidades de relación de IES y su representación integral del modelo PDDIM. El primer grupo es el que a nivel conceptual define a la preservación digital distribuida, como lo son: la *independencia*, la *coordinación* y la *replicación*. Los siguientes elementos que complementan a los conceptos anteriores como funciones obligadas son: el *monitoreo*, la *reparación* y el *almacenamiento*. Otro elemento que otorga y proporciona a las prestaciones un sistema de actuación a nivel técnico y funcional, es el sistema de preservación digital distribuida *LOCKSS*. En la representación integral cada una de estas prestaciones funcionales y conceptos tiene asignado un color que distingue su uso en los diferentes roles y relaciones.



Figura 23. Prestaciones funcionales de modelo PDDIM.

### ***Roles PDDIM***

A nivel de actividades y representaciones institucionales en el modelo PDDIM definimos tres grupos de rol y participación que se distinguen por su actuación, los cuales son: *nodos*, *coordinadores* e *instituciones*. En el *rol de nodos* hemos hecho uso de los ámbitos de alcance para designar *nodos de tipo municipal, estatal y nacional*. Los nodos realizarán actividades

puntuales como las de formación, integración y coordinación institucional. A nivel de *rol de coordinación* también se ha hecho uso del ámbito de alcance para definir franjas de actuación, por lo que se definieron roles de *coordinación municipal*, *coordinación estatal / regional* y *coordinación nacional* para dar actuación y seguimiento a nivel operativo como de infraestructuras en la red, lo que ayudará a una integración balanceada y sujeta de una administración central. En la figura de *rol de instituciones* es importante reiterar que en el estudio base de aplicación de niveles de preservación digital en IES-MX se tuvo una actuación directa y concreta del subsistema federal y estatal. La definición de tipo de instituciones de educación superior participantes para el modelo PDDIM refiere principalmente y de manera parcial a las figuras de *IES estatal* e *IES federal* en primer término. Aunado a esto y con la finalidad de la representación integral del modelo PDDIM, también se considera la figura de instituciones nacionales, la cual mantiene actuaciones directas de ámbito nacional.

Tabla 50. Roles generales en modelo PDDIM.

<b>Nodo PDDIM</b>	<b>Coordinación PDDIM</b>	<b>Institución PDDIM</b>
Nodo Municipal	Coordinación Municipal	IES Estatal / IES Federal
Nodo Estatal / Regional / IES	Coordinación Estatal / Regional	IES Estatal / IES Federal
Nodo Nacional	Coordinación Nacional	IES Federal / Inst. Nacional

### ***Relaciones PDDIM***

Las relaciones del modelo integral PDDIM se componen de estructuras uniformes, prestaciones funcionales y roles de actuación que muestran un abanico de integración progresiva y balanceada de acuerdo a diferentes necesidades, infraestructuras y alcances dentro de la estrategia de preservación digital, por lo que para lograr una distinción de relaciones hemos seleccionado seis diferentes tipos de relaciones dentro del modelo.

## Relación tipo “A”

En esta relación se integra a un grupo de *IES estatales* pertenecientes a un *ámbito municipal* específico que cuentan con una *estructura uniforme baja*, y por tanto están en proceso de formación de actividades iniciales de la escalera PDDIM. Este grupo de IES son coordinadas y supervisadas por una *IES estatal o federal* (mentora) de ámbito municipal que mantiene a su vez un *rol de nodo municipal* que centraliza, selecciona y coordina a las IES (municipales) que cumplen con las condiciones de normalización del primer escalón PDDIM (objetivos de gestión documental y estandarización de producción digital) para gestionar sus contenidos normalizados y estandarizados en un sistema de gestión electrónica o repositorio municipal que compartirá su almacenamiento coordinado de contenidos normalizados al *nodo estatal* (IES representante de estado), el cual ejercerá la plena conexión (infraestructura y sistema PDD) con el *nodo nacional* de preservación digital distribuida. Podemos observar que a nivel de coordinación existen dos capas con actuaciones a nivel municipal y otra a nivel estatal para facilitar y hacer más efectivo el proceso de formación y de integración de diversos grupos de IES al modelo integral nacional.

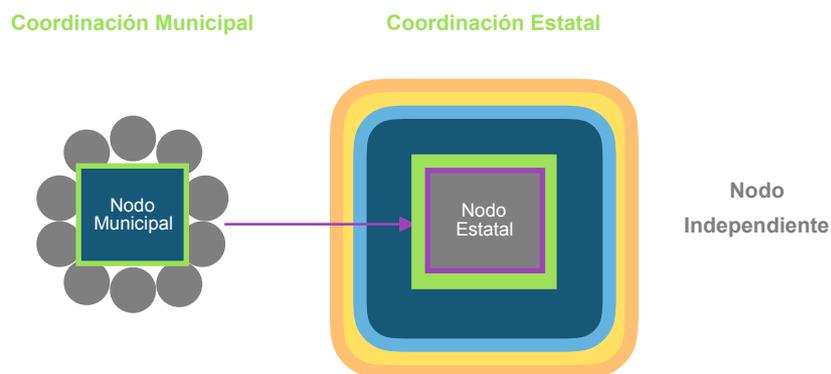


Figura 24. Relación “A” en modelo PDDIM.

## Relación tipo “B”

En esta relación se compone de la actuación independiente de una *IES federal o estatal* que cuenta con una *estructura uniforme plena*, ya que cumple con todos los requisitos de las escalera PDDIM e infraestructura necesaria para llevar a cabo de manera íntegra el *rol de nodo IES* para hacer uso de todas las prestaciones de preservación digital distribuida de manera directa. Esta relación independiente se distingue también por la excepción de coordinación de otras IES dependientes, ya que su actividad es completamente individual y es responsable del funcionamiento total de su nodo de preservación digital distribuida en la red. Este nodo de IES comparte servicios plenos de almacenamiento, replicación, monitoreo y reparación bajo el uso del sistema LOCKSS con el nodo nacional que realiza todas las acciones de administración coordinada de red y servicios plenos de preservación digital distribuida con los demás nodos estatales.

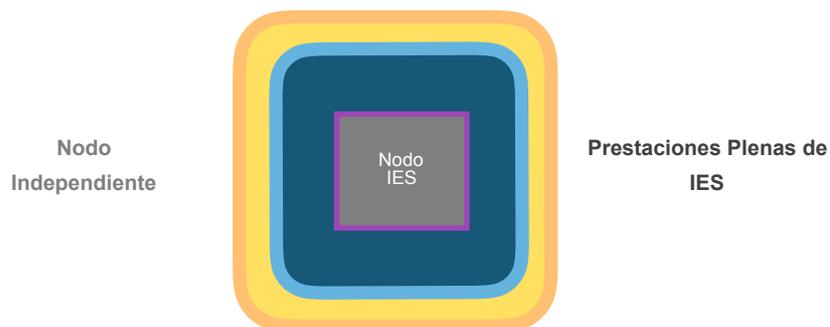


Figura 25. Relación “B” en modelo PDDIM.

## Relación tipo “C”

Este conjunto se integra por dos *IES federales* con una *estructura uniforme media*, y una *IES federal con estructura uniforme plena* que mantiene el *rol de nodo estatal* y *rol de coordinador estatal*. La relación de estas IES se destaca porque las dos *IES federales* (rol institución) solo requieren del uso de la prestación de monitoreo de contenidos almacenados en el *nodo estatal*, por lo que su infraestructura y nivel de servicio está especializado únicamente en tal contratación y solo a nivel informativo, es decir que estas IES con rol de instituciones gestionan sus repositorios digitales institucionales y comparten sus contenidos a el *nodo estatal* para que este realice todas las prestaciones plenas (replicación, reparación, monitoreo, LOCKSS) y les coordine a ellas con el nodo nacional de la red.

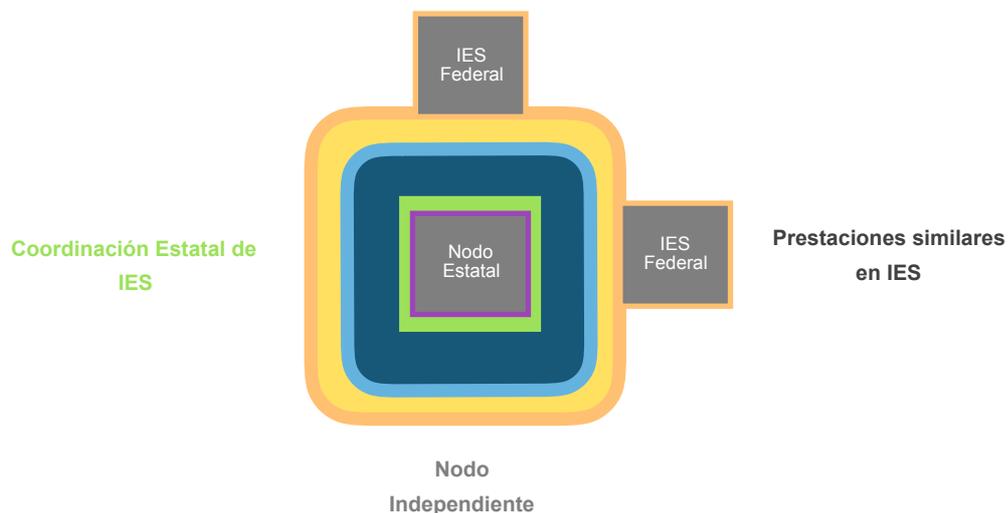


Figura 26. Relación “C” en modelo PDDIM.

## Relación tipo “D”

Esta relación define la participación de una *IES federal o estatal con estructura uniforme plena* que además de contar con las prestaciones plenas de preservación digital distribuida ofrece y se distingue por ofrecer su infraestructura plena como *nodo IES*, su infraestructura a modo de servicios compartidos de coordinación (a diferencia de la relación de tipo B). Está IES requiere mantener una mayor implicación en la red nacional a nivel de toma de decisiones y gestión de infraestructuras similares a la suya, para de esta forma reforzar los servicios e infraestructura de coordinación del nodo nacional. Esta distinción de compromiso extraordinario le otorga un mayor nivel de implicación en la gobernabilidad y administración de la red como socio de infraestructura técnica compartida (coordinación) y *nodo IES*.

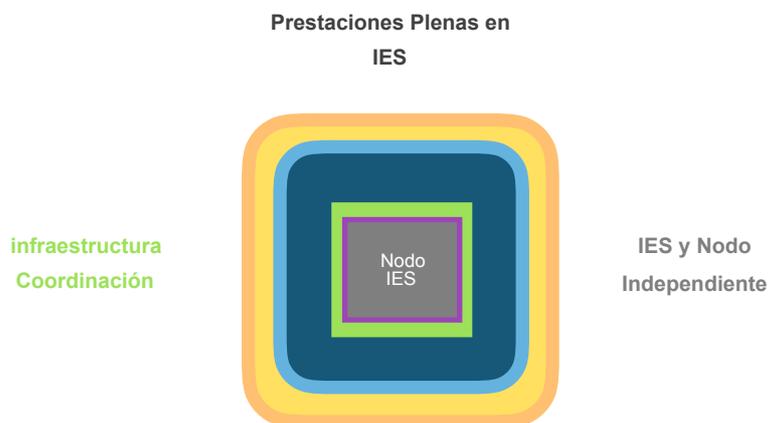


Figura 27. Relación “D” en modelo PDDI.

## Relación tipo “E”

En este conjunto relacional compuesto por tres *IES* con *rol de instituciones* (dos estatales y una federal) y una *IES federal o estatal con rol de nodo regional*. Podemos observar la integración de instituciones con subsistema diferente (estatal y federal) en el *rol de instituciones* pertenecientes a diferentes municipios y estados regionales (norte, sur, centro, etc.) que mantienen a su vez diferentes necesidades de prestaciones y servicios en su *estructura uniforme media*. Por ejemplo, la *IES federal* requiere adaptar de acuerdo a su actualización de políticas de preservación (escalón tres PDDIM) de la prestación de *replicación de contenidos* para complementar las acciones de almacenamiento que realiza de manera interna de sus contenidos en su repositorio institucional. Otra *IES* por su lado, requiere de la prestación específica de reparación a modo preventivo, y otra *IES* sólo requiere tener información de monitoreo sobre el estado de sus contenidos que comparte con el *nodo estatal*. La visión genérica de selección y posibilidades de prestaciones determinadas por cada *IES*, va de la mano de necesidades específicas que se presentan durante las diversas etapas de normalización de las actividades de la escalera PDDIM, ya que esto determinará la dependencia de infraestructura técnica y operativa del nodo estatal, y que mas adelante les posibilite a su independencia como nodo estatal, ya que estas *IES* con rol de instituciones mantienen un avance sustancial en la integración de políticas de preservación digital a nivel institucional y de implementación de infraestructura técnica necesaria para cumplirlo.



Figura 28. Relación “E” en modelo PDDIM.

## Relación tipo “F”

Este conjunto define la emancipación sincronizada de *IES federales o estatales* de la *coordinación municipal* del *nodo municipal* a causa de la integración individual de sus *sistemas de gestión electrónica o repositorios*. Por tal motivo dichas IES con *rol de instituciones* de forma unánime requieren como primera necesidad complementaria a la gestión de almacenamiento de sus contenidos a nivel interno, de compartir su sistema de almacenamiento con el nodo estatal o regional (zona norte, sur, centro, penínsulas, etc) para que este les integre paulatinamente a la estrategia nacional de preservación digital.

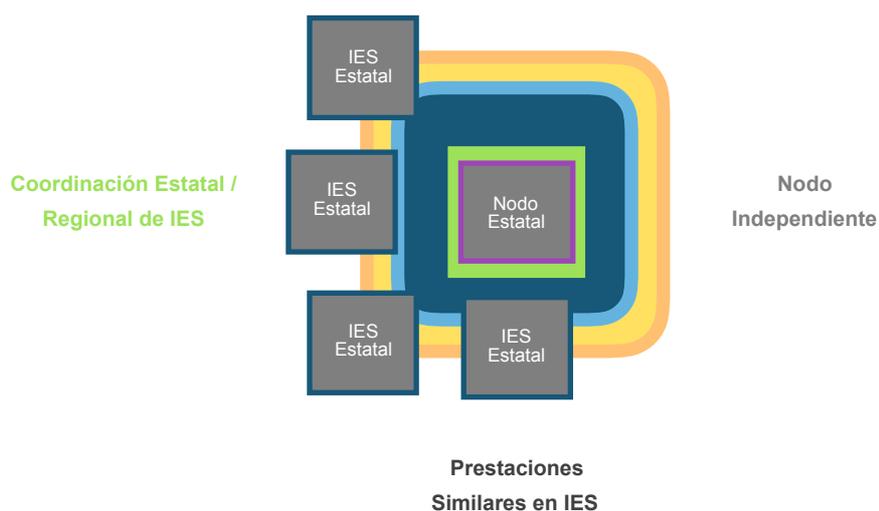


Figura 29. Relación “F” en modelo PDDIM

Este marco de relaciones y opciones de integración de IES a un marco de trabajo de colaboración, va de la mano de las consideraciones de situación en las que se encuentre cada una de las instituciones ante la escalera PDDIM tal como hemos dicho antes. Hay que reiterar que el objetivo de la propuesta es poner en perspectiva la flexibilidad de integración a diferentes tipologías relacionadas entre sí ante un modelo abierto colaborativo y coordinado.

## **Relación Integral Nacional**

Todos los conjuntos de relaciones se mantendrán conectados de forma nuclear al *nodo y coordinación nacional* que administra de manera específica a los *nodos estatales* y referirá las necesidades de los mismos para mantener, equilibrar y sostener su integración permanente a la red. La institución, grupos de instituciones o IES que representen la *coordinación nacional* deberán contar con la *infraestructura plena* necesaria que defina el supuesto de la cantidad de *nodos estatales* a los que dará *servicios compartidos* de coordinación, replicación, monitoreo, reparación y almacenamiento de contenidos. La definición de políticas estratégicas de colaboración interinstitucional con fines de preservación digital en modo distribuido, deben determinarse en esta figura central.

Todas las IES que forman parte de los distintos conjuntos relacionados del modelo PDDIM al momento de integrarse a la célula estatal, se dará por entendido que solicitan servicios y prestaciones de tipo compartido, los cuales el *nodo estatal coordinará* de acuerdo a las necesidades de cada grupo o IES individual, y este a su vez (nodo estatal) dependerá del *nodo nacional* para la operación plena. La forma en la que cada IES con *rol institucional* compartirá contenidos con el nodo estatal es a través del sistema LOCKSS y de la integración de un plugin del mismo en su repositorios de contenidos, el cual será cosechado por el *nodo estatal* y este a su vez creará las unidades de archivo (AU) de tipo estatal, que serán replicadas *nodo nacional* bajo su coordinación y administración.

Este modelo deja claro que tanto las necesidades de integración, como de infraestructuras y necesidades individuales son altamente variables, y por tanto se debe sopesar desde la visión de ámbitos de alcance, estructuras y relaciones para entender qué camino se debe tomar para la integración productiva de un plan de preservación digital nacional en modo distribuido.

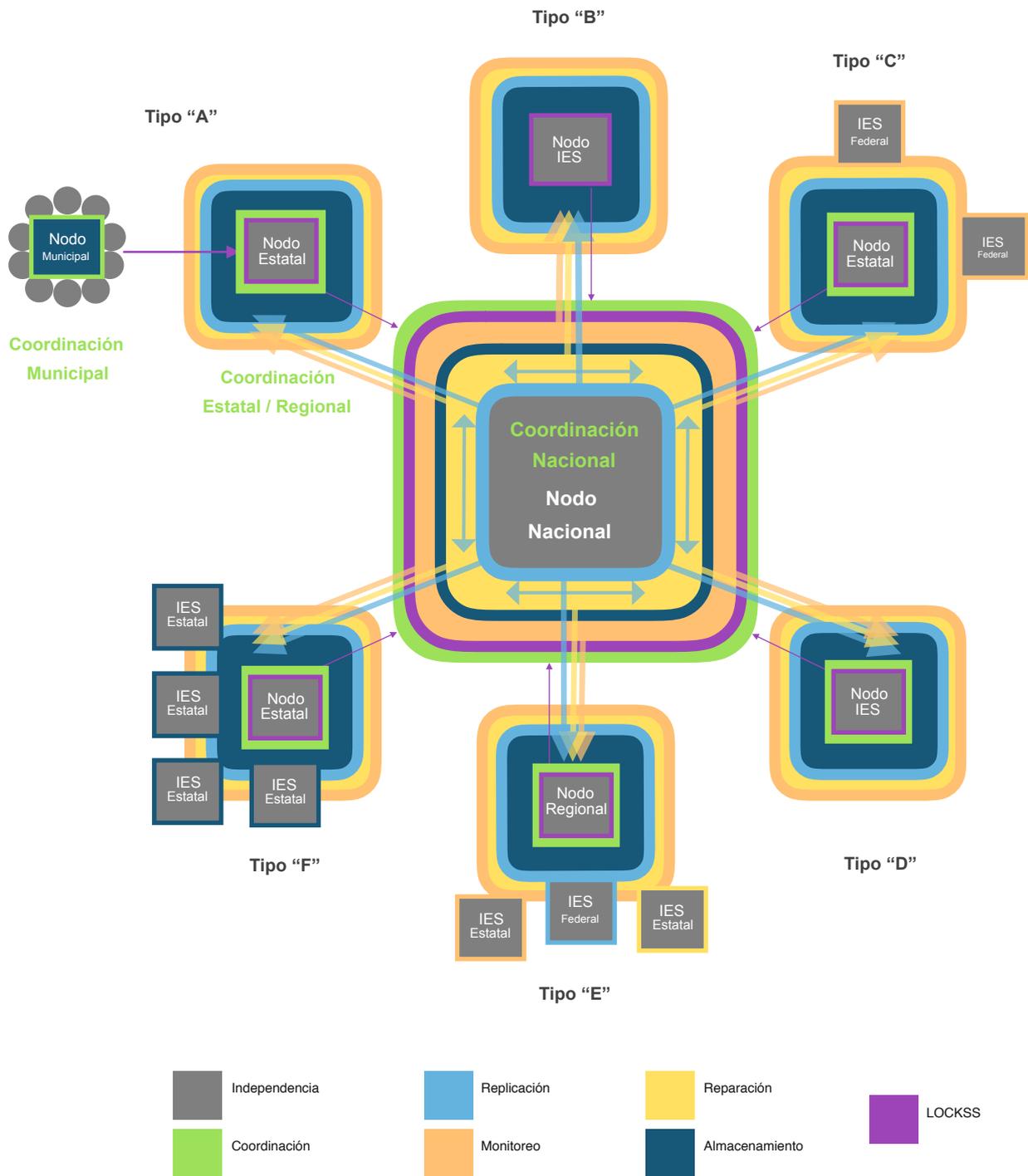


Figura 30. Representación gráfica y relaciones de modelo integral nacional PDDIM

#### **4.6.4 Implementación parcial de modelo PDDIM a IES-MX**

La definición de un modelo de preservación digital distribuido con énfasis en la integración equilibrada y sostenible de relaciones interinstitucionales de IES-MX en el subsistema federal y estatal con ayuda de diferentes estudios de tipo funcional de herramientas, organización de redes PDD y de acuerdo al nivel de alcance en actividades de preservación digital de 43 IES estudiadas como universo representativo del universo total de instituciones de educaciones superior; realizamos la implementación real parcial del modelo PDDIM asignando y tomando en cuenta únicamente a las 36 IES-MX (8 federales y 28 estatales) que atendieron el estudio. La finalidad de esta es entender y contrastar el estado de la cuestión actual de IES con los ámbitos de alcance y aplicación del modelo PDDIM. Esta implementación sirve de antesala sobre las posibilidades futuras de redactar un plan de integración progresiva a una estrategia de preservación digital distribuida a nivel nacional que contemple a todas las IES de México. La propuesta se asigna en función de los estados actuales de infraestructura, resultados específicos del estudio NDSA-LDP a IES-MX, contrastes y advertencias de actuación de modelos internacionales PDDD, así como de los panoramas de instituciones de influencia en el ámbito de actuación de IES-MX ante una estrategia de preservación digital distribuida.

Es importante aclarar que esta propuesta solo pretende mostrar una aproximación parcial a nivel de políticas de organización, requerimientos de tipo técnico informático, así como de las posibles relaciones funcionales de participación preliminar de IES-MX federales y estatales en el modelo PDDIM. La propuesta de implementación no representa una clasificación o calificación de IES de acuerdo a sus resultados individuales ante actividades de preservación digital, ni una propuesta técnica avanzada detallada sobre instalación y programación, ni una alineación y preferencia de instituciones por intereses obligados. La implementación del modelo PDDIM representa puntualmente un ejercicio práctico y panorámico que tiene por objetivo mostrar un borrador para la posible designación de IES dentro de los diferentes roles de actuación como nodos, instituciones colaboradoras y coordinadores de la red. Esta, trata de equilibrar los diferentes ámbitos y alcances de aplicación del modelo PDDIM con relaciones integradas en práctica.

## ***Implementación de políticas de organización, coordinación y gobernabilidad***

Dado que para coordinar una red con participación interinstitucional de acuerdo a las buenas prácticas de organización y gobernabilidad es necesario crear y ceñirse a diferentes acuerdos de organización que atiendan las diversas necesidades en niveles de coordinación y gobernabilidad de IES e instituciones integrantes de la red PDDIM, definimos como necesaria la elaboración detallada de un documento que detalle la política general de la red PDDIM a nivel de gobernabilidad, coordinación, aspectos técnicos y participantes en la cual se detallen los siguientes apartados generales como claves para su ejecución y entendimiento.

### *1. Aspectos de coordinación y gobernabilidad nacional de la red PDDIM*

Este apartado debe integrar el proceso de constitución y composición de las instituciones o IES como coordinadoras nacionales, así como de aquellas que fungirán como coordinadores a nivel estatal regional y municipal en la red, detallando el proceso de selección y designación. De igual forma se deben detallar los procedimientos de administración y gestión de la red para la creación del comité directivo, comité técnico, de contenidos, formación, etc. De forma clara y con énfasis especial, se debe atender y acordar claramente los compromisos de administración general para el seguimiento y supervisión de la red para su sostenibilidad a largo plazo.

### *2. Aspectos de composición técnica de nodos de la red PDDIM*

El correcto funcionamiento técnico de la red PDDIM dependerá en cierta medida de la detallada revisión de requerimientos y prestaciones que quieran cubrirse dentro de los servicios generales de la red PDDIM. Es muy importante integrar en el documento toda la información referente a los requerimientos técnicos como el tipo de tecnología informática, de almacenamiento, servidores, software o instalaciones necesarias, así como los requerimientos a cubrir en el modelo y la escalera PDDIM para el funcionamiento y definición cualitativa de los nodos nacionales, estatales - regionales y municipal para integrar así los diferentes roles de actuación, ámbitos de alcance y requerimientos técnicos.

### *3. Aspectos de organización funcional en red PDDIM*

Las políticas nacionales del modelo PDDIM deben incluir a su vez una detallada relación del todo el mapa de IES de México, así como de sus necesidades puntuales ante actividades de ámbito digital, almacenamiento, depósito y preservación digital. Debe estipular proponer los sistemas de organización más adecuados para que en la *praxis* de composición de la red se puedan crear grupos o subgrupos de diferentes ámbitos regionales para establecer sistemas de creación de células que podrán definir y seleccionar a una IES como núcleo o nodo representante de la misma. Con ello se atenderá a la correcta organización de apoyos en infraestructura de los nodos representantes y crear un sistema de soporte interinstitucional organizado que garantice la funcionalidad equilibrada de la red.

Estos tres aspectos referidos de forma parcial para la implementación de políticas de organización, sugieren como obligada su ampliación detallada a fin de crear un documento sólido, conveniente y de intereses heterogéneos que contribuyan a la definición clara de relaciones, ámbitos de alcance, nodos y aspectos técnicos de la red.

#### ***Implementación de ámbito de alcance, nodos y tipo de relaciones***

Para desarrollar la propuesta de implementación del tipo de relaciones de nodos y alcance del modelo PDDIM a IES-MX del estudio NDSA-LDP hemos asignado e interpretado los roles de participación de la representación integral del modelo PDDIM ante las prácticas resultantes en IES individuales y grupos de ellas de forma general. Para ello en primer lugar hemos integrado cinco grupos de IES de acuerdo al ámbito de zona geográfica regional (norte, centro, sur y penínsulas) que les corresponde e indistintamente del subsistema de pertenencia, para contar con un orden de grupos por zona, y de esta manera visualizar la posibilidad de relación territorial institucional.

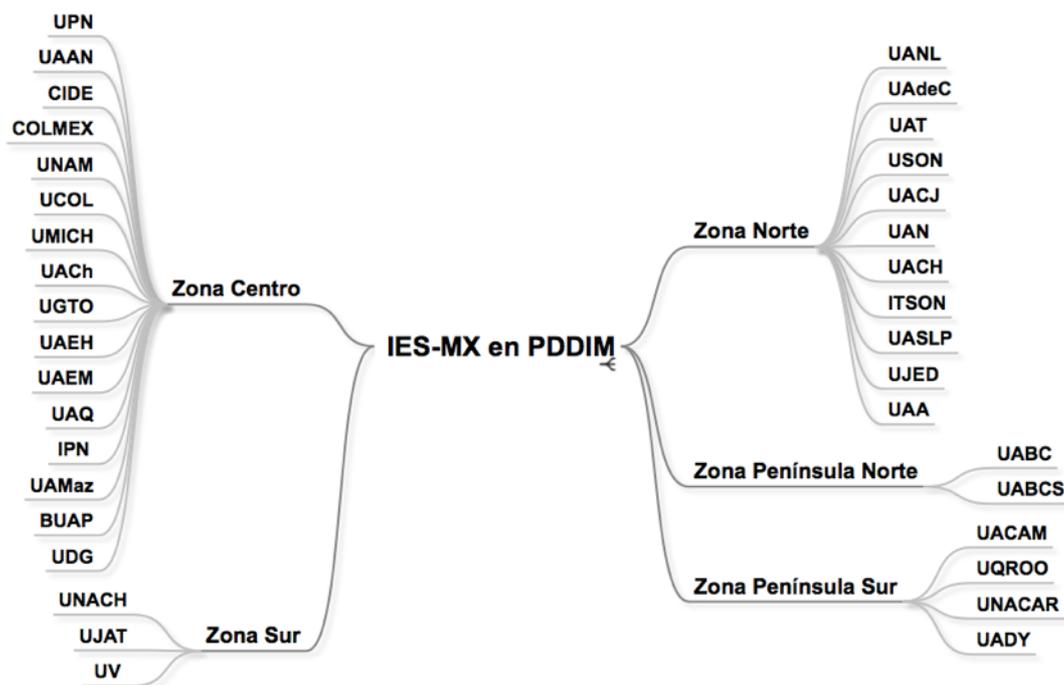


Figura 31. Zonas regionales y pertenencia de IES-MX de estudio.

Una vez distribuidas las IES-MX en grupos por zonas regionales hemos detectado que según las prácticas e infraestructura de algunos de ellos, existe la posibilidad de crear células funcionales directas. En otros casos, será necesario dividir los grupos en diferentes sub grupos territoriales o en otro momento y caso extraer a IES fuera de ellos para su participación individual. Dada la heterogeneidad *per se* de las IES-MX como de sus niveles de alcance individual por subsistema y global en actividades de preservación digital con situaciones mayormente en el nivel 1 (conocimiento de datos) y el nivel 2 (protección de datos) dentro de todas las categorías y actividades de preservación preservación digital como almacenamiento, integridad, seguridad, metadatos y formatos. Decidimos definir grupos y subgrupos de IES con sus respectivos nodos estatales - regionales que les represente o encabece. Esto con la finalidad de realizar una simulación aproximada y funcional de su integración e interacción por ámbito de alcance en su zona geográfica, así como por su similitud o aproximación equiparada de sus niveles de preservación digital con las estructuras uniformes de IES-MX de acuerdo con la escalera PDDIM. Debemos aclarar que al tratarse de una propuesta parcial, hemos recurrido a equilibrar los resultados de prácticas reales de niveles de preservación con

las posibilidades en alcances de participación a nivel funcional en la propuesta, con lo que consideramos simular la probabilidad de participación a un ámbito de recomendable y parcial.

### ***Tipos de relaciones en nodos y grupos de IES-MX para red PDDIM***

#### **Nodo Nacional Coordinador**

Este nodo coordinador se integra por la representación del Consejo Nacional de Ciencia y Tecnología (CONACYT) que será el eje rector de la *relación de tipo nacional*, administrador y gestor de la red PDDIM.

#### **Nodo IES Coordinador**

Como nodo e IES coordinador la Universidad Nacional Autónoma de México (UNAM) cuenta con las infraestructura, influencia y experiencia en el ámbito de gestión de acervos digitales, lo cual la coloca como la principal institución con perfil de integrar el nodo coordinador bajo una *relación de tipo "D"*, y que servirá de IES supervisora de los nodos regionales ya que compartirá con el nodo nacional parte de la infraestructura de coordinación tanto en datos como en decisiones de sostenibilidad de la red.

#### **Nodos Regionales y grupos de zonas**

Dada la cantidad de IES-MX por región hemos decidido situar y hacer uso de las distribuciones iniciales (zonas) para atender el ámbito de distribución geográfica requerida en los protocolos de PDD, y que esta incide de manera productiva y ordenada con el marco de trabajo de los participantes. Hemos definido la integración de grupos, subgrupos y nodos regionales principalmente para las zonas del centro y norte, ya que estas cuentan con una mayor cantidad de IES operativas. En este sentido los grupos y subgrupos mantendrán una dependencia ascendente a su nodo regional, ya que los integrantes de los grupos I actuarán bajo relaciones de tipo "C, D, E y F". Los subgrupos II mantendrán relaciones de tipo "A" con el nodo regional del subgrupo, por lo que por zona contamos con dos nodos regionales coordinadores con experiencia y que pueden atender a alcances de tipo medio en la escalera PDDIM. Hay que aclarar que solo el nodo regional principal de los grupos I será quien esté

conectado a la red nacional. En el caso de los grupos de zona sur y peninsulares, al tratarse de una cantidad compacta de IES operativas consideramos que de manera inicial y estratégica se mantengan como grupos independientes bajo relaciones de tipo “A” y “F” con la posibilidad de que entre estas se designe a una IES o se cree una figura virtual que funja como nodo regional de su zona. Para una propuesta integral se podrían incluso agregar a otras IES de zonas regionales colindantes para crear grupos más equilibrados en cuanto a integrantes y funciones, pero de momento esta propuesta se centra en la designación de grupos coherentes alrededor del modelo PDDIM.

### ***Nodo y Grupo Centro I***

Este nodo se integra por un grupo de ocho IES donde siete de ellas representadas por UPN, CIDE, COLMEX, UAMaz, UAEM, BUAP y la UDG, desempeñan el rol de institución colaboradora bajo una relación de tipo “E”, ya que suponen una similitud de estructuras de tipo uniforme media y de un ámbito regional directo. Dichas IES serán coordinadas por la UAEM, la cual cuenta con un alcance medio a la escalera PDDIM y el respaldo de su experiencia como coordinador y colaborador en la Red Mexicana de Repositorios Institucionales (REMEDI), por lo que fungirá como nodo regional coordinador bajo la designación y apoyo unánime del grupo.

### ***Subgrupo Centro II***

Se integra por siete IES que de acuerdo al modelo seis cumplirán el rol de institución colaboradora como lo son la UAEM, UCOL, UMICH, UACh, UAQ y la UGTO. Estas serán coordinadas por una IES con función de nodo regional y coordinador del subgrupo que se representa por el IPN que cuenta con experiencia técnica en la gestión de repositorios institucionales en abierto, así como de alcances medios a la escalera PDDIM. Este subgrupo cuenta a su vez con algunas IES que requieren normalizar sus actividades a la escalera PDDIM, por lo que este sub grupo mantiene una relación híbrida entre las de tipo “A” y las de tipo “F”, y por lo que su convivencia de subgrupo puede ayudar a la transferencia de conocimiento y mejora.

### ***Nodo y Grupo Norte I***

Siguiendo la estructura propuesta para los grupos de zona centro, el grupo I de la zona norte está integrada por cinco IES con rol de instituciones colaboradoras que son la UAT, UACH, USON, UCJ y el ITSON, y de una IES con la función de nodo regional coordinador representada por la UASLP que se distingue por su experiencia compartida como institución fundadora de la Red Mexicana de Repositorios Institucionales, así como de sus alcances medios sobre actividades de la escalera PDDIM. Esta conjunción nodo y grupo se realizará bajo una relación de tipo “E” ya que el nodo regional se conecta con el nodo nacional y el nodo coordinador IES, centralizando dentro de su nodo regional de almacenamiento compartido los contenidos y material de las IES con rol de institución colaboradora, para así desarrollar de manera nacional las funciones de preservación digital en modo distribuido.

### ***Subgrupo Norte II***

Este subgrupo norte se integra por una IES con rol de nodo regional y coordinador encabezado por la UANL, ya que esta es la principal representante de las universidades del norte de México ante la gestión de repositorios institucionales y cuenta con alcances medios plenos en la escalera PDDIM. La UANL coordina a cuatro IES con rol de instituciones colaboradoras representadas por la UAA, UAdeC, UAN y la UJED bajo una relación de tipo híbrida “A” y “F” ante necesidades de fraguar formación y actividades de especialización en IES colaboradoras.

### ***Nodo y Grupo Sur***

La participación de IES de la zona sur para el estudio converge en solo tres instituciones, por lo que este grupo se integra por una IES como nodo regional coordinador representado por la UV pues esta además de contar con una estructura de alcance medio ante la escalera PDDIM se mantiene como una de las universidades que cuenta con experiencia en la coordinación facultades y campus para la integración de su repositorios institucional en abierto. Es por ello que esta coordinará a dos IES con rol de instituciones colaboradoras definidas en la UACH y la UJAT bajo una relación de tipo “F” o “E para transferencias de conocimiento y mejora.

### ***Nodo y Grupo Península Sur***

Las dos zonas que representan una estrategia territorial importante para la definición de la replicación y distribución geográfica de los nodos de la red PDDIM están definidas en las penínsulas de Baja California y Yucatán. Este grupo inicial de la península sur se integra por tres IES con rol de institución colaboradora que son la UADY, UNACAR y la UQROO, la cuales serán coordinadas por la IES con rol de nodo regional representada por la UACAM con alcances medios a la escalera PDDIM para coordinarse bajo una relación de tipo “A” y “F” a las IES colaboradoras.

### ***Nodo y Grupo Península Norte***

Este grupo se integra por dos IES en las cuales se definirá a una IES como nodo regional (UABC) y una IES con rol de institución colaboradora (UABCS) con relación de tipo “C”. El principal objetivo de esta selección se define puesto que estratégicamente su posición es relevante para la integración de la red, no obstante podríamos encontrar problemas de sostenibilidad o alcance de especialización en las actividades de preservación digital. Se sugiere en caso de que no se llegue a un acuerdo o sea posible la designación de una IES (de estas dos) como nodo regional, el crear un nodo regional virtual que sea en función de la inversión de infraestructura compartida entre las dos IES de la zona y posiblemente de otras IES de la región bajo relaciones de tipo “A”. Esta misma situación puede aplicar de igual forma a la integración de otros grupos y subgrupos de la red.

### ***Implementación de tipo técnico informática***

A nivel técnico informático y a manera de orientación general la implementación refiere principalmente a las necesidades informáticas de hardware y software especializado para la función de la red PDDIM, y que atañe en primer instancia de este ejercicio al *nodo nacional* y *los nodos estatales - regionales (o nodos IES)*, ya que estos son los nodos que realizarán las funciones principales de preservación digital distribuida a nivel técnico informático.

Tal como lo hemos mencionado antes, esta propuesta sugiere a nivel básico general las recomendaciones de soluciones técnicas y funcionales para afrontar las necesidades de integración técnica. En lo que a *software* se refiere nuestra propuesta considero en primer instancia la implementación del software LOCKSS en la red, el cual requiere de *hardware* específico que puede ser consultado en el manual de instalación de LOCKSS (<http://www.lockss.org/support/build-a-lockss-box/>) para su instalación, y correcto funcionamiento. Recordemos que la función principal del *software* LOCKSS es la de convertir al *hardware* (ordenadores o servidores) en una caja o *LOCKSS box*, la cual estará lista para conectar con el resto de los nodos a modo de cajas conectadas que funcionan bajo un mismo protocolo de ejecución técnica y de funciones de preservación digital en modo distribuido. Hay que apuntar que los requerimientos técnicos de implementación mantendrán una diferencia notable entre los tipos de cajas LOCKSS (nodos PDDIM), y esta dependerá principalmente del tamaño de la red, y donde el *nodo coordinador nacional* deberá considerar una capacidad superior de recursos (*hardware*) para mantener y administrar a los demás nodos. Las cajas de nodos estatales - regionales o IES, debe por igual responder a la asignación de cantidad de cajas LOCKSS replicara en su nodo, por lo que esta aproximación se deberá sopesar de manera clara según los tipos de relaciones de la propuesta de implementación.

### ***Representación de implementación parcial de modelo PDDIM en IES-MX***

Dentro de la propuesta parcial podemos observar una alta flexibilidad de integración que es abanderada principalmente por IES con experiencia en el ámbito de gestión de datos y colaboración institucional, factores imprescindibles en toda actividad de tipo distribuido en preservación digital. Hay que destacar que para la balanceada distribución de la cantidad de contenidos replicados entre nodos, se debe sopesar la calidad de los nodos a nivel de infraestructura, participantes y posibilidades de gestión. Para crear una equitativa distribución de contenidos replicados de modo estable y siguiendo las buenas prácticas internacionales de cooperativas de preservación digital altamente distribuidas como MetaArchive Cooperative, la designación de replicación de contenidos a un nodo específico dependerá de protocolos de tiempo, espacio e infraestructura de servicios de preservación que deben ser reflejados en los acuerdos y políticas de distribución a modo equitativo para cumplir con un equilibrio de participa

En el caso de nuestra propuesta de implementación parcial a IES-MX encontramos que el nodo de la zona centro integra la mayor cantidad de IES agrupadas e incluso a subgrupos. Esta situación se traduce en una alta gestión de contenidos que podría ser equiparable a la cantidad de contenidos que se gestionan en el nodo de la zona norte, lo cual a modo cuantitativo y cualitativo representa un espejo claro de replicación. Lo mismo aplicaría con el caso del nodo de IES coordinador que puede sopesar con una replicación de tal envergadura e incluso el nodo península sur. El caso de los nodos de península norte y nodo sur, presenta una compacta cantidad de IES y posiblemente de contenidos que no sean compatibles con el resto de los nodos, aunque en un momento dado la referencia de separar zonas, grupos y subgrupos de tal índole heterogénea (cantidad de IES y contenidos de nodo regional), principalmente se enfoca a crear un orden y visión panorámica de los alcances de una red altamente distribuida por zonas de México, que además obliga a la atención de tipo formativo, infraestructura y de apoyos a iniciativas colaterales como las de los lineamientos de repositorios institucionales y el repositorio nacional, poniendo en clave otra finalidad funcional para involucrarse formalmente en un proceso de colaboración nacional por parte de las IES públicas y que estas adopten de acuerdo a sus posibilidades la cultura de preservación de manera escalonada (escalera PDDIM).

Reiteramos que esta propuesta debe ser atendida como una provisión parcial de participación argumentando que para la integración formal de esta práctica, será necesario establecer formalmente los criterios de acuerdos nacionales o plan de gobernanza de la red interinstitucional, definir de los protocolos técnicos convenidos según las necesidades de los participantes y revisar detenidamente el cumplimiento individual de cada institución sobre los parámetros y requisitos propuestos en la escalera PDDIM. La concatenación formal de todos estos elementos habilita la realización de una propuesta más concreta de participación colaborativa y convirtiendo esta previsión a una propuesta formal.

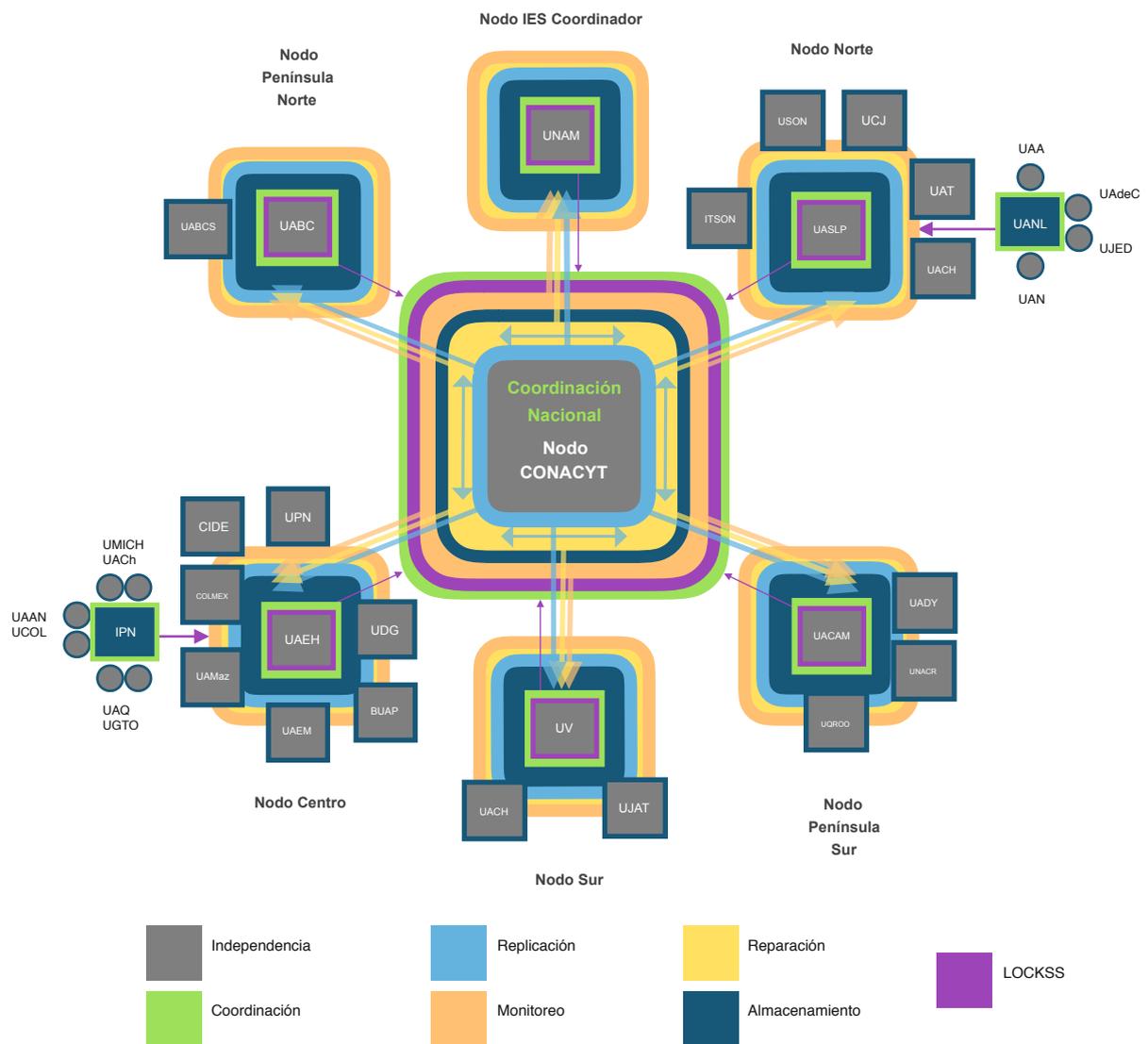


Figura 32. Implementación parcial de modelo PDDIM en IES-MX de estudio.



## **Capítulo 5 - Conclusiones**



## 5.1 Conclusiones Generales

De acuerdo a nuestra investigación y estudio de contextos generales en los distintos ámbitos de la preservación digital, hemos llegado a la conclusión de que la gestión a largo plazo de archivos digitales aun da una alta prioridad a las soluciones de tipo técnico sobre modelos centralizados de preservación, lo que deja en un segundo plano a las estrategias distribuidas por la considerarlas como servicios agregados de la estrategia principal. Esta situación refiere a que las estrategias distribuidas requieren integrar actividades extraordinarias de organización para consolidar la puesta en común de diferentes integrantes con múltiples necesidades dentro de un programa coordinado de esta índole. Podemos afirmar que en la práctica internacional la preservación digital (como tal) aun es percibida como una actividad de lujo o de gama alta por las necesidades conjuntas de organización, tecnología e inversión implicadas para que una o un conjunto de organizaciones integren un programa de preservación digital. Esta percepción debe considerarse como una ventana de oportunidad para divulgar formalmente los beneficios, compromisos y responsabilidades grupales que conllevan la implementación a escala de una estrategia distribuida de preservación digital.

Se ha observado que solo organizaciones e instituciones con posibilidades económicas y localizadas principalmente en países desarrollados (o en vías de ello), sostienen a la fecha este tipo de iniciativas y estrategias de preservación digital tanto en modo centralizado, distribuido o cooperativo. En este sentido cuando la práctica actual de la preservación digital refiere a tal orden y posibilidades económicas debemos preguntarnos: ¿La preservación digital un lujo exclusivo de países desarrollados?, y si es así ¿Que se debería hacer para que la preservación digital pueda expandir y compartir su cultura de adopción a países no desarrollados con necesidades de conocimientos y experiencias que les ayuden a integrar un modelo horizontal, colaborativo y democratizado que apoye a su implementación y sostenibilidad?. Responder a ello es una tarea complicada por la diversidad de intereses y cobertura de necesidades internacionales, lo que nos orienta hacia un orden sin equilibrio de intereses en común. Debemos concretar que de acuerdo al contraste en opiniones de referentes y pioneros de programas de preservación digital distribuida tanto en Estados Unidos, Dinamarca, Cataluña o Brasil; en su mayoría sugieren que para lograr una correcta implementación, seguimiento y sostenibilidad de un programa de preservación digital sea cual sea su estrategia (centralizada

o distribuida), se requiere de forma obligada del compromiso leal de varias organizaciones (públicas o privadas) sobre acuerdos vistos al largo plazo que integren de manera sólida un programa organizado de colaboración técnica y de colaboración. Aunque parece y suena como una tarea fácil, será difícil conseguirlo.

Creemos firmemente que una estrategia colaborativa de tipo distribuido a diferencia de una centralizada de preservación digital es una opción más viable para un conjunto altamente heterogéneo de instituciones, pues permite a las instituciones que no cuentan con suficientes recursos o formación para emprender un programa de preservación de forma individual o centralizada el ser parte de un programa integral donde desde distintos esfuerzos y responsabilidades cooperativas interinstitucionales se pueden aprovechar de los beneficios que trae consigo la economía a escala en modo distribuido tanto en temas financieros como el ahorro en costes y reinversión de los mismos, transferencia de conocimiento entre instituciones y en la replicación diversificada de copias seguras en diferentes sitios geográficos como temas esenciales en la gestión de un programa de preservación digital.

Consideramos que dentro de las prioridades y advertencias de los casos de modelos de preservación digital distribuida estudiados, debemos resaltar que una de las características que debe tomarse muy en cuenta al iniciar un programa colaborativo integrado por diversas instituciones es el de observar minuciosamente las cualidades únicas de cada institución (experiencia, antecedentes de colaboración, visión, infraestructuras, etc.), pues mientras más similares o diferentes sean a las de otras instituciones se facilitan o complican más las soluciones técnicas en conjunto con la creación de acuerdos de cooperación entre ellas, y por otro lado la integración de grupos con un ADN altamente colaborativo que coincidan como responsabilidades y actuaciones concretas a modo de red.

Hemos observado claramente que no existen situaciones ideales para iniciar un programa de preservación digital distribuida. El caso de México y sus instituciones de educación superior públicas ante un programa de este tipo entra en una situación diferente a las de los modelos de preservación digital distribuida de Dinamarca, Brasil o Cataluña tanto por la tipología de los contenidos a preservar, como por la composición institucional de sus integrantes y el nivel de integración a actividades de cooperación interinstitucional por mencionar algunas. La red de

preservación digital distribuida IES-MX (PDDIM) deberá de tomar decisiones propias para establecer un programa formal y un modelo equilibrado de trabajo, ya que la situación de las IES-MX y sus niveles de participación de actividades de preservación digital nos muestran un alcance centrado en niveles básicos (nivel 1 y 2) no normalizados y desiguales parcialmente. Esto se traduce en un problema o barrera para que estas en conjunto puedan iniciar un programa de preservación digital.

Una de las principales aportaciones científicas de nuestra investigación se centra en solucionar la situación o problemática de poner en marcha un programa de preservación digital distribuida de forma equilibrada para un conjunto heterogéneo de IES-MX con alcances básicos o dispares en la materia. La solución que proponemos como producto de investigación es una metodología u hoja de ruta denominada *escalera PDDIM* que tiene como objetivo dar visibilidad, orden y secuencia a la ejecución de los requisitos primarios e intermedios para alcanzar de manera sólida un programa de preservación digital de tipo distribuido. Los requisitos que integran a la *escalera PDDIM* de manera general están representados en la divulgación de la importancia de la cultura de preservación digital, la formación y transferencia de su conocimiento. De forma intermedia continúan con la adopción paulatina de los procedimientos técnicos de normalización y gestión organizacional de documentos digitales para consolidarse como preparativos preliminares dirigidos hacia un programa de preservación digital distribuida a largo plazo. Otra de las aportaciones de nuestra investigación se centra en la propuesta de un modelo conceptual funcional de relaciones, organización y participación de IES-MX (modelo PDDIM) dentro de un programa de preservación digital distribuida. Este modelo tiene como fin el representar las posibilidades y flexibilidad existentes en la integración de relaciones funcionales de IES-MX como parte de una red distribuida y de acuerdo a sus necesidades, alcances y niveles de adopción de actividades de preservación digital. De forma práctica propone un esquema gráfico de relaciones y roles de participación de IES-MX dentro de una red de preservación digital distribuida que pone en panorámica las necesidades individuales, ámbito de alcance y actuación según su zona geográfica de pertenencia.

El modelo PDDIM funciona como propuesta preliminar para visualizar posibles relaciones integrales de participación ante la perspectiva de formalización futura de un programa de preservación digital distribuida equilibrado y sostenible al largo plazo. Tanto la escalera como

el modelo PDDIM representan por igual un esquema de evaluación de IES-MX ante su participación en un programa puntual de preservación digital distribuida.

Debemos reiterar que hasta que no se entienda que la organización y la correcta distribución de responsabilidades están por encima de las prestaciones e infraestructuras tecnológicas, ningún proceso de preservación digital distribuida será viable sea cual sea la institución, país o idiosincrasia pragmática. Una red de preservación digital distribuida debe ser centralizada para los temas de coordinación y administración, distribuida para la organización de responsabilidades técnicas, y colaborativa para temas de sostenibilidad económica.

Ante el escepticismo institucional que refiere a la posibilidad de integrar y sostener un programa de preservación digital de forma individual, debemos decir que todas las soluciones y tendencias apuntan a consolidar un paso gradual a de estándares centralizados a distribuidos como base para democratizar la colaboración y puesta en común de preservar el patrimonio digital institucional. Estas tendencias apuntan a que mientras se cubran los requisitos preliminares de formación y cultura de preservación digital, se organice una correcta distribución de responsabilidades y se establezcan los acuerdos de cooperación con un alto compromiso de responsabilidad; habrá un escenario de suerte para iniciar un programa de preservación digital distribuida pues coincidirá entonces la preparación con la oportunidad.

## **5.2 Limitaciones**

Durante nuestra investigación hemos encontrado limitaciones que definen el espectro de alcance de la preservación digital distribuida como un tópico de estudio abierto. Por un lado todos los desarrollos e implementación de modelos centralizados o de casos de redes distribuidas como estrategia de organización y solución técnica al día de hoy siguen cumpliendo un estado *beta* o *a prueba*, ya que como primeros desarrollos siguen siendo evaluados y mantienen pasos intermedios de conversión a una estrategia consolidada. Esta situación trajo consigo para nuestro objeto de investigación la existencia de escasa literatura de ámbito normalizado o estandarizado a nivel de términos plenamente aceptados, lo que nos apremio a recoger primeras aproximaciones y apreciaciones teóricas que sugieren potenciar su entendimiento al ponerse en práctica plena en conjunto de pruebas pilotos de

implementación e instalación de sistemas de preservación digital distribuida de casos específicos de estudio. La conjunción de las apreciaciones de los modelos de organización teóricos y los juicios de valoraciones de expertos en el ámbito de técnico (programación e ingeniería de redes) consolidarán el estudio y propuesta de un modelo de preservación digital distribuida. Otra consideración oportuna refiere a extender con líneas futuras de trabajo las valoraciones de niveles de preservación digital de las instituciones estudiadas integrando un esquema de análisis en conjunto de auditorías formales de preservación digital. Por último la propuesta de un modelo económico de financiación y sostenibilidad acentuará el complemento de implementación del modelo conceptual funcional.

### **5.3 Líneas futuras de trabajo**

Más allá de las conclusiones, éstas no agotan en sí mismas todo el análisis necesario para que a modo indicativo marquemos a continuación que en el caso de que estas propuestas se quieran llevar a la práctica será importante considerar los siguientes puntos a tomar en cuenta para ello.

#### **5.3.1 Estudio A/B de niveles de preservación digital en IES-MX**

Consideramos que para establecer un contexto de mayor profundidad, personalización y conocimiento pleno de necesidades de IES-MX es necesario llevar a cabo un protocolo de diagnóstico de gestión de archivos ante recomendaciones y niveles de preservación digital con variantes de tipo interno y externo (A/B) que ayuden a contrastar y ceñir los niveles de implicación en actividades de preservación digital de instituciones de educación superior. Dado el acercamiento parcial relacionado a la obtención de participación y resultados las IES-MX participantes en el estudio, esto puede enriquecerse a un nivel memorable que continúe con los protocolos establecidos en el modelo y escalera PDDIM.

### **5.3.2 Modelo de organización de gobernanza y financiación**

Hemos observado que para que una práctica completa de preservación digital distribuida mantenga una cohesión y equilibrio en todos los ámbitos que le atañen debe contar con un modelo sólido de organización de gobernanza. Es por ello que recomendamos realizar estudios especializados y sectoriales en México que definan formalmente un modelo de gobernanza que acondicione en función de números reales de disposición y financiación, un protocolo fiable de integración para diversas instituciones en un programa de preservación digital como requisito vital para la sostenibilidad real que deje ver beneficios reales del mismo en sus participantes.

### **5.3.3 Implementación plena de modelo PDDIM**

Para la implementación plena del modelo PDDIM en los diversos ámbitos de subsistema de IES-MX se recomienda establecer un estudio enfocado a conocer las posibilidades de acuerdos a niveles de organización, cooperación, coordinación y gobernanza. Así mismo de establecer a priori protocolos de relaciones institucionales con fines comunes de preservación digital. La implementación de tipo técnico y de recursos se deberá sopesar bajo el paraguas de los ejes institucionales de interés y responsables para fraguar un esquema equilibrado, alcanzable y sostenible a largo plazo. Hay que dejar claro que el planteamiento de una estrategia de preservación digital en modo distribuido puede encontrar beneficios económicos por las participaciones a escala, por lo que sugerimos por igual atender a la investigación de posibles desarrollos originales que estipulan e integran un marco propio tecnológico que sirva de modelo para otras futuras referencias.

## **6 - Fuentes**



## 6.1 Bibliografía y Referencias

- Abadal, E. (2013). *Acceso abierto a la ciencia*. Editorial UOC. Barcelona, España.
- Adams, I., Miller, E., Storer, M. (2011). *Analysis of workload in scientific and historical long term data repositories*. Technical report - Storage systems research center. University of California, Santa Cruz: Paper.
- ANUIES (2014). *Descripción de proyecto de presupuesto de egresos de la federación 2015 para la educación superior*. México. Publicación Digital.
- Arreola, H. (2012). *Repositorios de tesis digitales de posgrado: Análisis en trece instituciones mexicanas*. Tesis - Biblioteconomía y Estudios de la Información UNAM. México.
- Barnard, A. (2009). *Los Lineamientos para la Preservación de Archivos Digitales: Propuesta del Proyecto InterPARES," Conferencia durante I Seminario Internacional de Archivos y Documentos Electrónicos*. Artículo de conferencia. Bogotá, Colombia.
- Barnard, A., Cardoso, E., Voutssás, J. (2014). *Preservación Digital en México*. SIPAD Symposium de Preservación Audio Visual y Digital. México.
- Barton, M. R.; Waters, M. (2004). *Creating an institutional repository: LEADIRS workbook*. Boston: MIT Libraries.
- Bel, B. (2012). *A Trusted Digital Repository based on the OAIS model with integrated management of access rights*, Cultural Heritage on line – Trusted Digital Repositories & Trusted Professionals. Florence.
- Björnshauge, L. (2003). *The Network :The Reorganisation of Library and Information Services at UGent*. University of Lund, Sweden. Summarising Report.
- Brown, A. (2013). *Practical Digital Preservation: A how-to guide for organisations of any size*. UK. Facet Publishing.

- Borghoff, U.M., Rödiger, P., Scheffczyk, J., Schmitz, L. (2006). *Long-Term Preservation of Digital Documents*. Springer. Germany.
- Bueno, G., Hernández, T. (2011). *Estrategias para el éxito de los repositorios institucionales de contenido educativo en las bibliotecas digitales universitarias*. BiD: textos universitaris de biblioteconomia i documentació, n. 26. España.
- Castillo, J.M., Jorba, F. (2008). *Emmagatzematge distribuït i preservació digital : una panoràmica d'alternatives*. BiD: textos universitaris de biblioteconomia i documentació. n. 2. España.
- CCSDS - OAIS (2002-12). *Reference model for an Open Archival Information System*. Washington, DC, USA. CCSDS Secretariat NASA.
- Chacon, S. (2009). *Pro Git Book: Everything you need to know about git distributed source control tool*. The experts voice in software development. Apress. USA.
- Charlesworth, A. (2012). *Intellectual Property Rights for Digital Preservation*. DPC Technology Watch Report. Digital Preservation Coalition. UK.
- Chen, B., Curtmola, R. (2013). *Towards self-repairing replication-based storage systems using untrusted clouds*. In *Proceedings of the third ACM conference on Data and application security and privacy*. CODASPY ACM. USA.
- COAR (2012). *The Current State of Open Access Repository Interoperability*. Confederation of Open Access Repositories. Germany.
- Conway, P. (2000). *Overview: Rationale for digitization and preservation*. In Sitts, M. K. (Ed.), *Handbook for digital projects : A Management tool for preservation and access*. Andover, MA.. Northeast Document Conservation Center.
- Conway, P. (2010). *Preservation in the Age of Google: Digitization, Digital Preservation, and Dilemmas*. *The Library Quarterly: Information, Community, Policy*. v. 80, n. 1, p. 61-79.
- Corti, L., Van den Eynden, V., Bishop, L., Wollard, M. (2014). *Managing and Sharing Research Data: A Guide to Good Practice*. SAGE Publications Ltd. UK.

- Cunniff, D. (2013). *Presentation Advanced Topics in Digital Curation: Distributed Digital Preservation*. University of North Texas College of Information. Licensed under a CC-BY-NC 3.0 license. USA.
- D'Amato, G. (2012). *Towards distributed digital preservation*. Notes from 2nd LIBER Workshop on Digital Preservation. Italy. DPC.
- Daintith, J. (2004) *Data independence*. Disponible en: <http://www.encyclopedia.com/doc/1O11-dataindependence.html>
- Deegan, M., Tanner, S. (2006). *Digital Preservation*. Facet Ed. UK.
- Del Pozo, N., Stawowczyk, A., Pearson, D. (2010). *Land of the lost: a discussion of what can be preserved through digital preservation*. Library Hi Tech. v. 28, p. 290 - 300.
- Denzin, N., Lincoln, Y. (1998). *Entering the field of qualitative research*. Londres: Sage. UK.
- Dewitt, D. (2012). *Going Digital: Strategies for Access, Preservation and Conversion of collection to digital formats*. Routledge, RLG. UK.
- Dobson, C. (2003). *From bright idea to beta test, the history of LOCKSS*. SEARCHER: The magazine for database professionals.
- Dollar, C., Ashley, L. (2014) *Digital Preservation Capability Maturity Model (DPCMM) background and performance metrics*. Council of State Archivists (CoSA). Albany NY.
- Duarte, M., Márdero, M. (2015). *Preservação do patrimônio científico das humanidades: a emergência da Rede Cariniana (Preservation of the scientific heritage of the humanities: the emergence of Cariniana Network)*. Cadernos de História. Brasil.
- Espinoza, A. (2014) *Repositorios institucionales en el mundo*. CACSLA. México.
- Ferreras, F. (2010). *Preservación digital en RIs: Gredos*. Tesis Maestría. Universidad de Salamanca. España.
- Font, O., Ruiz, A., Mena, M. (2012). *Diagnóstico sobre la gestión documental y de archivos en la Universidad Central Marta Abreu de las Villas. Cuba: Caso de estudio*. Revista Española de Documentación Científica, 35, 4. España.

- García, I. (2000). *El Patrimonio Documental en México*. Centro Universitario de Investigaciones Bibliotecológicas. UNAM. Métodos de Información. v.7, n.40. México.
- Gladney, H. (2007). *Preserving digital information*. Springer. USA.
- Granger, S. (2000). *Emulation as a Digital Preservation Strategy*. D-Lib Magazine. UK.
- Guiaretta, D. (2011). *Advanced Digital Preservation*. Springer -Verlag Berlin Heidelberg. UK/Germany.
- González, A., Barrios, M. (2012). *Métodos y técnicas para la investigación del comportamiento informacional: fundamentos y nuevos desarrollos*. Gijón: Editorial Trea. España.
- Harvey, R. (2011). *Preserving digital materials*. 2nd Edition. Walter de Gruyter GmbH & Co. KG, Göttingen. Germany.
- Hawkins, D., et al. (2013). *Personal Archiving: preserving our digital heritage*. Information today, Inc. USA.
- Hitchcock, S., Brody, T., Hey, J. and Carr, L. (2007). *Digital Preservation Service Provider Models for Institutional Repositories: Towards Distributive Services*. DLib Magazine. UK.
- Huppé, G., Creech, H., Knoblauch, D. (2012). *The Frontiers of Networked Governance*. Paper IISD. Canada.
- Gore, A. (2006). *An Inconvenient Truth*. Film Documentary directed by Davis Guggenheim. USA.
- Kalbach, J. (2007). *Designing Web Navigation*. O'Reilly Media. USA.
- Kaluszka, A. (2011). *Computer emulation history*. Disponible en: <http://kaluszka.com/vt/emulation/history.html>
- Keefer, A., Gallart, N. (2003). *La preservación digital y las universidades : el estado de la cuestión*. Conference paper. 8as Jornadas Españolas de Documentación. España.
- Keefer, A., Gallart, N. (2007). *La Preservación de recursos digitales : el reto para las bibliotecas del siglo XXI*. Editorial UOC. Barcelona. España.

- Kirschenbaum, M., Ovenden, R., Redwine, G. (2010). *Digital Forensics and Born-Digital Content in Cultural Heritage Collections*. Council on Library and Information Resources. Washington, D.C. USA.
- Labastida, I. (2008). *Noves alternatives per gestionar els drets d'autoria en la difusió de continguts. Les llicències de Creative Commons*. Escola d'Administració Pública de Catalunya II. Títol III. Col·lecció: «Manuals i Formularis». Barcelona, España.
- Lindlar, M., et al. (2013). *Benefits of geographical, organizational and collection factors in digital preservation cooperations: The experience of the Goportis consortium*. Conference Paper. International Conference on Preservation of Digital Objects. Lisbon, Portugal.
- Lee, C. (2014). *Up Close and Personal: Individual Digital Traces as Cultural Heritage and Discovery through Forensics Tools*. In *Proceedings of PATCH*. Association for Computing Machinery. Haifa, Israel. New York, NY.
- Lluca, C., Reoyo, S. (2012). *Repositorios sostenibles, reflexiones a partir de la experiencia española*. ETD: Simposio Internacional de tesis y disertaciones electrónicas. Lima, Perú.
- Lluca, C. (2011). *Repositorios digitales: preservación y difusión de las colecciones digitales en bibliotecas, archivos y museos de España*. Presentación no publicada. Universitat de Barcelona (UB) y Universidad Nacional Autónoma de México (UNAM), México DF (México).
- Márdero, M. (2013). *Preservación digital distribuida: un modelo para América Latina*. Conference paper. BIREDIAL 2013 Conferencia Internacional Acceso Abierto, Preservación Digital y Datos Científicos, San José, Costa Rica.
- Maron, N., Ithaka S. (2014). *Guide to the best revenue models and funding sources for your digital resources*. The Strategic Content Alliance. Jisc. UK.
- Molina, J. (2010). *Preservación digital y derechos de autor: ¿un conflicto sin solución?*. V Congreso Nacional de Bibliotecas Públicas. Gijón, España.
- Muñoz, H. (2011). *La universidad mexicana en el escenario global*. Perfiles Educativos | vol. XXXIII | IISUE-UNAM 33. México.

- McGovern, N., Skinner, K., eds. (2012). *Aligning National Approaches to Digital Preservation*. Educopia Institute Publications. Atlanta. GA., USA.
- OECD. (2015). *Estudios económicos de la OCDE México 2015: Vision General*. Organización para la Cooperación y el Desarrollo Económico. México.
- Palacios, J. (2000). *Archivo General de la Nación y bibliotecas digitales hacia el siglo XXI*. Bibl. Univ., Nueva Época. v. 3, n. 1. México.
- Perez, T., Garcia, M. (2013). *Datos abiertos y repositorios de datos: Nuevo reto para los bibliotecarios*. El profesional de la información. v. 22, n.3. Barcelona, España.
- Pérez-Montoro, M. (2003). *El documento como dato, conocimiento e información*. revista tradumàtica Traducció i Technologies de la Informació i la Comunicació. n. 2. La Documentació. Barcelona, España.
- Pérez-Montoro, M. (2010). *Arquitectura de la Información en entornos web*. Gijón: Trea, cop. Barcelona, España.
- Peterson, E. (2013). *Digital Data Distribution Platforms in Action: A New Approach for Driving Real-Time Marketing Optimization*. Web Analytics Demystified. Portland, USA.
- Phillips, M., Bailey, J., Goethals, A., Owens, T. (2013). *The NDSA Levels of Digital Preservation: An Explanation and Uses*. Paper. Library of Congress - NDSA (National Digital Stewardship Alliance). USA.
- Provan, K., Kenis, P. (2007). *Modes of Network Governance: Structure, Management, and Effectiveness*. Journal of Public Administration Research and Theory. n.18, p. 229-252. USA.
- Ribera, M., Splendiani, B., Térmens, M., Salse, M., Llerena, I., Franganillo, J., Centelles, M. (2012). *Accessible teaching documents*. 3er Congrés Europeu de Tecnologies de la Informació en l'Educació i en la Societat: una visió crítica. Barcelona, España.
- Rodriguez, T. (2012). *Diagnóstico de repositorios en las IES de México ¿Sabe Usted donde esta la producción de su Universidad?*. Presentación Estudio. REMERI. Universidad de Guadalajara. México.

- Rosenthal, D.S.H. (2014). *Architectural choices in LOCKSS networks*. Library Hi Tech. v. 32, n. 1, p. 2 - 10. USA.
- Rosenthal, D.S.H.; Vargas, D. (2013). *Distributed Digital Preservation in the Cloud*. International Journal of Digital Curation. USA.
- Rosenthal, D.S.H., Rosenthal, D.C., Miller, E.L., Adams, I.F., Storer, M.W. & Zadok, E. (2012). *The economics of long-term digital storage*. Paper presented at The Memory of the World in the Digital Age Conference. Vancouver, BC.
- Rosenthal, D.S.H. (2010). *LOCKSS: Lots Of Copies Keep Stuff Safe*. Paper presented at NIST Digital Preservation Interoperability Framework Workshop. USA.
- Rothenberg, J. (1999). *Ensuring the Longevity of Digital Information*. (expanded version of the article “Ensuring the Longevity of Digital Documents” that appeared in the January 1995 edition of Scientific American). v. 272, n. 1, p. 42-7. USA.
- Rothenberg, J. (1998). *Avoiding Technological Quicksand: Finding a Viable Technical Foundation for Digital Preservation*. CLIR - Council on Library and Information Resources. Washington, DC. USA.
- Sánchez, S., Melero, R. (2006). *La denominación y el contenido de los repositorios institucionales en acceso abierto: base teórica para la “ruta verde*. CSIC. España.
- Schmidt, E. (2010). *Every 2 Days We Create As Much Information As We Did Up To 2003*. TechCrunch. *Recuperado de:* <http://techcrunch.com/2010/08/04/schmidt-data/>
- Schumacher, J., Thomas, L., VandeCreek, D., et al. (2014). *From Theory to Action: “Good Enough” Digital Preservation Solutions for Under-Resourced Cultural Heritage Institutions*. A Digital POWRR White Paper for the Institute of Museum and Library Services. USA.
- Schultz, M., Gore, E., (2010). *The Importance of Trust in Distributed Digital Preservation: A Case Study from the MetaArchive Cooperative*. Conference Paper. iPRES 2010. Vienna, Austria.
- Schultz, M., Zierau, E. (2013). *Creating a Framework for Applying OAIS to Distributed Digital Preservation*. Conference Paper. iPRES 2010. Lisboa, Portugal.

- Schultz, M., Skinner, K. (2014). *Comparative Analysis of Distributed Digital Preservation (DDP) Systems: Chronicles in Preservation Project*. Educopia Institute. USA.
- Secker, J., Coonan, E. (2013). *Rethinking Information Literacy: A practical framework for learning*. Facet Publishing, UK.
- Secker, J. (2004). *The institutional repository in the digital library. Electronic Resources in the Virtual Learning Environment: a guide for Librarians*. Chandos. Oxford, UK.
- Serra, J. (2014). *Development of a digital preservation policy as a business opportunity for consulting and archivists*. Conference paper. Girona 2014: Arxius i Indústries Cultural. España.
- Sheldon, M. (2013). *Analysis of current digital preservation policies: Archives, Libraries and Museums*. Study Paper. Junior Fellow. USA.
- Sierman, B. (2012). *Preservation Policies: necessary and beneficial*. Document DPC in 2nd Workshop on Digital Preservation. Florence, Italy.
- Skinner, K., Schultz, M., eds. (2010). *A Guide to Distributed Digital Preservation*. MetaArchive Cooperative. Atlanta, GA: Enducopia Institute .USA.
- Smallwood, R. (2014). *Information Governance: Concepts, Strategies, and Best Practices*. Wiley (CIO). USA.
- Sørensen, J. (2014). *From stand-alone preservation to cross-institutional collaboration*. Paper conference, Girona 2014: Arxius i Industries Culturals. España.
- Sosinsky, B. (2011). *¿Que es la Nube? El futuro de los sistemas de información*. Wiley-Anaya Multimedia. USA.
- Stonebraker, M. (1974). *A functional view of data independence*. In *Proceedings of the 1974 ACM SIGFIDET (now SIGMOD) workshop on Data description, access and control (SIGFIDET '74)*. ACM, New York, NY, USA
- Sulé, A., Estivill, A., Gascón, J. (2011). *Interfaces de consulta en las colecciones digitales patrimoniales españolas*. Anales de Documentación. España.

- Térmens, M. (2011). *La sostenibilitat econòmica i tècnica dels repositoris de preservació digital*. Lligall. 31:p.44-62. Barcelona, España.
- Térmens, M. (2012). *La preservació de les revistes digitals: realitats i reptes pendents*. 5 Jornades Catalanes de Revistes Científiques. Barcelona, España.
- Térmens M., Salse, Locher, A. (2012). *Retos técnicos para la preservación de los repositorios institucionales*. Presentación. 5as Jornadas OS-Repositorios. Bilbao, España.
- Térmens, M., Ribera, M. (2009). *El control de los formatos en la preservación digital*. XI Jornadas españolas de documentación, FESABID. Zaragoza, España.
- Térmens, M. (2009). *Investigación y desarrollo en preservación digital: un balance internacional*. El profesional de la información. Barcelona, España.
- Térmens, M. (2010). *Perspectivas para la sostenibilidad y la preservación de los repositorios institucionales*. Presentación. 4as Jornadas OS-Repositorios. Barcelona, España.
- Térmens, M. (2013). *Preservación Digital*. Colección del Profesional de la Información. Editorial UOC. Barcelona, España.
- Thurston, A. (2012). *Preservación digital: preservar el patrimonio y proteger los derechos civiles*. International Records Management Trust (Fundación Internacional de Gestión de Documentos). USA.
- Tramullas, J., Garrido, P. (2006). *Software libre para repositorios institucionales: propuestas para un modelo de evaluación de prestaciones*. El profesional de la información. Barcelona, España.
- Trehub, A., Halbert, M. (2012). *Safety in Numbers: Distributed Digital Preservation Networks*. IFLA World Library and Information Congress. Helsinki, Finland.
- Tzitzikas, Y., Marketakis, Y., Kargakis, Y. (2012). *Conversion and Emulation-aware Dependency Reasoning for Curation Services*. iPress 2012. Toronto, Canada.

- UNESCO (2013). *Concept of digital preservation*. Disponible en: <http://www.unesco.org/new/es/communication-and-information/access-to-knowledge/preservation-of-documentary-heritage/digital-heritage/concept-of-digital-preservation/>.
- UNESCO (2012). *Declaración de Vancouver*. Disponible en: [http://www.unesco.org/new/fileadmin/MULTIMEDIA/HQ/CI/CI/pdf/mow/unesco\\_abc\\_vancouver\\_declaration\\_es.pdf](http://www.unesco.org/new/fileadmin/MULTIMEDIA/HQ/CI/CI/pdf/mow/unesco_abc_vancouver_declaration_es.pdf)
- UNESCO. (2003). *Directrices para la preservación del patrimonio digital*. Librería Nacional de Australia. Australia.
- Van der Hoeven, J. (2014). *The use of emulation tools as part of a strategy for long-term preservation of digital records*. 3rd LIBER workshop. Vienna, Austria.
- Vicente, A., Gozzer, S. (2011). *Nuevos hábitos de lectura, escritura y publicación en la red*. Revista Ábaco. Gijón, España.
- Villa, E. (2014). *Las principales características de la reforma energética en México y su impacto en el sistema nacional de ciencia*. Presentación no publicada. Jornada de Cooperación CONACyT- Cataluña. UPC. Barcelona, España.
- Voutssás, J., Barnard M.Alicia (2014). *Glosario de Preservación Archivística Digital*. Versión 4.0. UNAM. México.
- Voutssás, J. (2012). *Long-term digital information preservation: challenges in Latin America*. Aslib Proceedings. v. 64, n. 1, pp. 83 - 96. UK.
- Voutssás, J. (2009). *Preservación del patrimonio documental digital en México*. Centro Universitario de Investigaciones Bibliotecológicas UNAM. México.
- Waters, D. J., ed. (1997). *Digital archiving: the report of the CPA/RLG Task Force. Preservation and digitization: principles, practices and policies*. OCLC research. USA.
- Walters, T., McDonald, R. (2008). *Creating Trust Relationships for Distributed Digital Preservation Federations*. Paper. Georgia Institute of Technology - Library and Information Center. USA.
- Williams, D., Golden, J. (2014). *Advances in Library Administration and Organization*. Emerald Insight v. 32, p. 107-147. UK.

- Woodward, H., Estelle, L. (2010). *Digital information; Order or anarchy*. Facet publishing, London, UK.
- Zierau, E. (2011). *A Holistic Approach to Bit Preservation*. PhD thesis. Department of Computer Science, Faculty of Science, University of Copenhagen (DIKU). Denmark.
- Zierau, E., Bøgvad, U. (2010). *Cross-institutional cooperation on a shared bit repository*. World Digital Libraries. v. 6, n. 1, p. 25-36. India.



## **7 - Anexos**



## 7.1 Índice de Tablas

Tabla 1. Diferencias de formatos de consulta y formatos de preservación.	11
Tabla 2. Plazos en preservación digital a nivel conceptual DPC.	32
Tabla 3. Dimensiones y necesidades en preservación digital.	38
Tabla 4. Servicios y redes especializadas en preservación digital distribuida.	116
Tabla 5. Estructuras de organización y gobierno en modelos distribuidos.	117
Tabla 6. IES de México por subsistema.	127
Tabla 7. Marcos de análisis PDD.	145
Tabla 8. Identificación de sistema PDD.	148
Tabla 9. Uso de sistemas PDD.	151
Tabla 10. Ingesta de contenido.	154
Tabla 11. Almacenamiento y réplicas.	156
Tabla 12. Monitoreo y mantenimiento.	158
Tabla 13. Recuperación y acceso.	160
Tabla 14. Indicadores generales OCMC.	175
Tabla 15. Indicadores generales y específicos OCMC.	176
Tabla 16. Niveles de Preservación Digital (NDSA-LDP Ver. 1).	191
Tabla 17. Componentes y cantidad de categorías NDSA-LDP.	192
Tabla 18. Categorías y cantidad de actividades NDSA-LDP.	193
Tabla 19. Componentes y cantidad de Nivel - Acción NDSA-LDP.	193
Tabla 20. Nivel - Acción y cantidad de actividades NDSA-LDP.	194
Tabla 21. Actividades y cantidades totales NDSA-LDP.	194
Tabla 22. Nivel - Acción, categorías, actividades y cantidades totales NDSA-LDP.	194
Tabla 23. Línea de tiempo de aplicación de estudio.	199
Tabla 24. Cantidad de IES participantes del estudio.	200
Tabla 25. Disposición de herramientas digitales en IES participantes del estudio.	201
Tabla 26. Tipo de herramientas digitales en IES participantes de estudio.	202
Tabla 27. Tipo de herramientas digitales en IES por subsistema.	203
Tabla 28. Resultados brutos de almacenamiento y localización geográfica en IES.	204
Tabla 29. Porcentajes brutos de almacenamiento y localización geográfica en IES.	205
Tabla 30. Resultados brutos de no alteración de archivos e integridad de datos en IES.	207
Tabla 31. Porcentajes brutos de no alteración de archivos e integridad de datos en IES.	208

Tabla 32. Resultados brutos de no alteración de archivos e integridad de datos en IES.	211
Tabla 33. Porcentajes brutos de no alteración de archivos e integridad de datos en IES.	212
Tabla 34. Resultados brutos de no alteración de archivos e integridad de datos en IES.	214
Tabla 35. Porcentajes brutos de no alteración de archivos e integridad de datos en IES.	214
Tabla 36. Resultados brutos de no alteración de archivos e integridad de datos en IES.	216
Tabla 37. Porcentajes brutos de no alteración de archivos e integridad de datos en IES.	217
Tabla 38. Resultados globales por subsistema en almacenamiento y localización geográfica.	222
Tabla 39. Participación proporcionada por subsistema en almacenamiento y localización geográfica.	222
Tabla 40. Resultados globales por subsistema en no alteración de archivos e integridad de los datos.	223
Tabla 41. Participación proporcionada por subsistema en no alteración de archivos e integridad de los datos.	223
Tabla 42. Resultados globales por subsistema en seguridad de la información.	224
Tabla 43. Participación proporcionada por subsistema en seguridad de la información.	224
Tabla 44. Resultados globales por subsistema en metadatos.	225
Tabla 45. Participación proporcionada por subsistema en metadatos.	225
Tabla 46. Resultados globales por subsistema en formatos.	226
Tabla 47. Participación proporcionada por subsistema en metadatos.	226
Tabla 48. Comparativo de presupuesto en IES federales y estatales de México.	235
Tabla 49. Comparativo FODA en casos de modelos PDD.	261
Tabla 50. Roles generales en modelo PDDIM.	302

## 7.2 Índice de Figuras

Figura 1. Diferencias de preservación y conservación digital a plazos.	9
Figura 2. Entidades funcionales del modelo OAIS	43
Figura 3. Service provider model (Hitchcock et al., 2007)	58
Figura 4. Institutional model (Hitchcock et al; 2007)	59
Figura 5. Software model (Hitchcock et al., 2007)	59
Figura 6. Diagrama de arquitectura de protocolos de preservación digital distribuida	108
Figura 7. Home Page de sistema LOCKSS, The BitRepository Project y IRODS	150
Figura 8. Formularios de estudio de buenas prácticas de organización	178
Figura 9. Diseño de preguntas generales de identificación, herramientas y usos.	197
Figura 10. Modelo de distribución para preservación digital de Zierau (2011)	245
Figura 11. Elementos conceptuales de la preservación digital distribuida	247
Figura 12. Modelo de elementos de definición conceptual y funcionales en preservación digital distribuida	250
Figura 13. Modelo visión conceptual y funcional NBRP	252
Figura 14. Modelo de red privada CARINIANA y misión de integración de subredes.	255
Figura 15. Modelo de participación colaborativa de CSUC en red de PDD MetaArchive Cooperative.	259

Figura 16. Escalera de acciones para modelo PDDIM	276
Figura 17. Ambitos de aplicación y alcance de estrategia de preservación en modelo PDIIM	289
Figura 18. Ámbito municipal de IES México para modelo PDDIM	291
Figura 19. Alcance de ámbito estatal y regional de IES México para modelo PDDIM	293
Figura 20. Alcance de ámbito regional zona de IES México para modelo PDDIM	294
Figura 21. Alcance de ámbito nacional de IES México para modelo PDDIM	295
Figura 22. Ejemplos de estructuras uniformes de grupos de participantes para modelo PDDIM.	297
Figura 23. Prestaciones funcionales de modelo PDDIM.	301
Figura 24. Relación “A” en modelo PDDIM.	303
Figura 25. Relación “B” en modelo PDDIM.	304
Figura 26. Relación “C” en modelo PDDIM.	305
Figura 27. Relación “D” en modelo PDDI.	306
Figura 28. Relación “E” en modelo PDDIM.	307
Figura 29. Relación “F” en modelo PDDIM	308
Figura 30. Representación gráfica y relaciones de modelo integral nacional PDDIM	310
Figura 31. Zonas regionales y pertenencia de IES-MX de estudio.	314
Figura 32. Implementación parcial de modelo PDDIM en IES-MX de estudio.	321

### 7.3 Índice de Gráficas

Gráfica 1. Niveles de preservación digital de IES-MX en categoría de almacenamiento y localización geográfica.	205
Gráfica 2. Niveles de preservación digital de IES-MX en categoría de no alteración de archivos e integridad de los datos	209
Gráfica 3. Niveles de preservación digital de IES-MX en categoría de seguridad de la información	212
Gráfica 4. Niveles de preservación digital de IES-MX en categoría de metadatos	215
Gráfica 5. Niveles de preservación digital de IES-MX en categoría de formatos	217
Gráfica 6. Porcentaje global IES-MX sobre situación y niveles de preservación digital	219
Gráfica 7. Porcentaje absoluto proporcional por niveles NDSA-DPL	222
Gráfica 8. Porcentaje general proporcional de actividades de almacenamiento y localización geográfica	222
Gráfica 9. Porcentaje absoluto proporcional por niveles NDSA-DPL.	223
Gráfica 10. Porcentaje general proporcional de actividades de no alteración de archivos e integridad datos.	223
Gráfica 11. Porcentaje absoluto proporcional por niveles NDSA-DPL.	224
Gráfica 12. Porcentaje general proporcional de actividades de seguridad de información.	224
Gráfica 13. Porcentaje absoluto proporcional por niveles NDSA-DPL.	225
Gráfica 14. Porcentaje general proporcional de actividades de metadatos.	226
Gráfica 15. Porcentaje absoluto proporcional por niveles NDSA-DPL.	226
Gráfica 16. Porcentaje general proporcional de actividades de formatos.	226
Gráfica 17. Porcentaje global por subsistema de IES-MX sobre situación ante niveles de preservación digital	227

