

## Preservación digital distribuida: lecciones de experiencias internacionales

[Versió catalana]

- ① **DAVID LEIJA**  
Profesor de la Facultad de Arquitectura, Diseño y Urbanismo  
Universidad Autónoma de Tamaulipas  
[dleija@uat.edu.mx](mailto:dleija@uat.edu.mx)
- ① **MIQUEL TÉRMENS**  
Profesor del Departamento de Biblioteconomía, Documentación y Comunicación Audiovisual  
Universitat de Barcelona  
[termens@ub.edu](mailto:termens@ub.edu)

Opciones



PDF

<meta />

Metadatos

### Resumen

**Objetivos:** el objetivo de este estudio es ofrecer una panorámica general de los sistemas de organización de las redes de preservación digital distribuida.

**Metodología:** específicamente, el artículo se inicia abordando de manera exploratoria los modelos de referencia teórica y técnica de redes de archivos digitales con fines de preservación digital distribuida a largo plazo. A continuación, se analizan tres casos internacionales en los que se explora la organización interna de sus miembros y cómo ello repercute en el servicio ofrecido. Se termina con un estudio comparativo (FODA) de los tres sistemas.

**Resultados:** los sistemas de preservación digital distribuida, más allá de sus componentes técnicos, son organizaciones basadas en la cooperación en las que sus miembros han de cumplir una serie de requisitos previos.

### Resum

**Objectius:** l'objectiu d'aquest estudi és oferir una panoràmica general dels sistemes d'organització de les xarxes de preservació digital distribuïda.

**Metodologia:** específicament, l'article s'inicia abordant de manera exploratòria els models de referència teòrica i tècnica de xarxes d'arxius digitals amb fins de preservació digital distribuïda a llarg termini. A continuació, s'analitzen tres casos internacionals en què se n'explora l'organització interna dels membres i com això repercuteix en el servei ofert. S'acaba amb un estudi comparatiu (FODA) dels tres sistemes.

**Resultats:** els sistemes de preservació digital distribuïda, més enllà dels components tècnics, són organitzacions basades en la cooperació, els membres de les quals han de complir una sèrie de requisits previs.

### Abstract

**Objectives:** This study offers an overview of the systems used to organize distributed digital preservation networks.

**Methodology:** The article starts by examining the theoretical and technical reference models offered by digital archive networks dedicated to long-term distributed digital preservation. It then analyzes the internal organization of the member institutions of three international networks and considers how their system of organization affects the service offered. It concludes with a feature-oriented domain analysis of the three systems.

**Results:** While the technology they use may vary, the distributed digital preservation networks studied share a common commitment to collaborative practices and their member institutions need to meet certain prior requirements.

**Palabras clave:**

**Keywords:**

Recibido: 07/07/2017. Aceptado: 12/09/2017.

## 1 Modelos de archivo para preservación digital

Han pasado casi veinte años desde que Rothenberg publicó su artículo "Ensuring the longevity of digital documents" (Rothenberg, 1999), en el que visualizaba, analizaba y hacía una previsión técnica sobre la problemática de la accesibilidad futura de soportes y formatos digitales ante el problema de su obsolescencia tecnológica. Desde entonces hasta la actualidad se han concebido distintos modelos de actuación que conforman distintas opciones orientadas a la mejora continua y la sostenibilidad de un programa de preservación digital. Uno de estos modelos es el de la preservación digital distribuida, que intenta ir más allá de simplemente la aplicación de una solución técnica para abarcar también actividades de planificación y de colaboración coordinada entre instituciones.

En preservación digital, los modelos de archivo hacen referencia a los sistemas de almacenaje digital en conjunto, así como a todas las acciones y estrategias colaterales que se requieren para mantenerlos activos. En todo modelo de archivo y gestión de objetos digitales intervienen dos funciones indispensables, que son el acceso y la preservación. Estas funciones, aunque en algún momento puedan ser contradictorias, se han de desarrollar de forma complementaria con distintos procedimientos de gestión de objetos digitales, como son los de control de la procedencia, la autenticidad y la integridad digital. La forma en la que encajan estos distintos procedimientos de gestión puede variar de acuerdo con cómo una determinada organización los implemente o, en su caso, defina como modelo de archivo (D'Amato, 2012).

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

De acuerdo con la definición simplificada, OAIS – Reference Model for an Open Archival Information System es un archivo de una organización (o de una parte de ella) en el que personas y sistemas han aceptado la responsabilidad de preservar información y ponerla a disposición de una comunidad designada. Este compromiso se concreta en el reconocimiento de una serie de responsabilidades o funciones definidas como estándar normativo, las cuales permiten a un archivo OAIS distinguirse de otros usos del término *archivo* (Giaretta, 2011).

El modelo OAIS fue creado en el año 2002 por el Consultative Committee for Space Data Systems, actualizado en el año 2012 (OAIS, 2012) y reconocido como norma ISO 14721:2003. Actualmente es el marco de referencia más ampliamente aceptado para la concepción técnica de archivos de preservación digital.

A partir del modelo de referencia OAIS, con el paso del tiempo han ido apareciendo diferentes alternativas técnicas, de organización y de operación específica para atender las necesidades de la preservación digital de archivos.

Desde el punto de vista de su arquitectura general y de cuál es su enfoque prioritario, podemos categorizar que existen cuatro modelos de aplicación de la preservación digital en un entorno determinado. El primero es el modelo centralizado, que mantiene su interés primario en las herramientas o aplicaciones informáticas para una gestión de la información y los datos centralizada. El segundo es el modelo de la preservación descentralizada o distribuida, que prioriza su enfoque en los datos y su distribución geográfica como prioridad. Ambos modelos coexisten en dos modalidades de acceso al archivo diferentes. El archivo en modo abierto se centra en el acceso y la preservación en conjunto como actividades prioritarias, y el modo de archivo cerrado tiene como prioridad la preservación como actividad realizada por separado sobre los objetos que custodia y deja la función de acceso en un segundo plano y de forma restringida (Térmens, 2013).

Actualmente ya disponemos de un número considerable de sistemas que aplican la modalidad de preservación descentralizada o distribuida, entre los que destacan **APTrust** (<http://aptrust.org>) , **Archivematica**, **Cariniana** (<http://cariniana.ibict.br/>) , **Chronopolis**, **Danish Bit Repository**, **Data-PASS** (<http://www.data-pass.org/>) , **DPN** (<http://dpn.org>) , **DuraCloud** (<http://www.duracloud.org/>) , **HathiTrust** (<http://www.hathitrust.org>) , **Internet Archive** (<http://archive.org/>) , **iRODS** (<http://www.irods.org/>) , **LOCKSS** (<http://www.lockss.org/>) y **MetaArchive Cooperative** (<http://www.metaarchive.org/>) . Las políticas de colaboración en el ámbito de la preservación digital se están extendiendo, pues permiten crear sistemas más sólidos y que generan una mayor confianza en el cumplimiento de su misión (Day, 2008). Pero las diferencias entre estos sistemas son considerables, por lo que merecen ser estudiadas. A continuación, mostraremos dos opciones generales para la colaboración externa en preservación; luego veremos alternativas propias en el caso de repositorios institucionales.

### 1.2 Modelos de colaboración externa para preservación digital

Siguiendo a Brown (2013), observamos que existen diferentes opciones para la colaboración entre instituciones con fines de preservación digital.

#### 1.2.1 Opción subcontratada

Hoy en día la subcontratación representa un mercado en crecimiento formado por servicios externos de repositorios seguros y por sistemas gestores de preservación digital externos a las organizaciones. Este mercado ya está ofreciendo servicios completos que incluyen software propio y las técnicas necesarias para mantener los objetos digitales copiados, al día y accesibles a largo plazo, siempre y cuando se mantenga la relación con los proveedores del servicio. Podemos decir que esta opción es la primera aproximación a la descentralización de un procedimiento de preservación digital. Es una opción fácil de implementar porque permite evitar la inversión en equipo y en sistemas para un sistema propio de la organización. La subcontratación se puede hacer a otras instituciones o bien a empresas privadas de servicios digitales, y existe una gran variabilidad en los niveles de implicación y responsabilidades entre las dos partes, el cliente y el contratista, pero normalmente hay una gran delegación en la segunda.

### 1.2.2 Opción de asociación

La opción de colaborar de manera asociada está planteada principalmente cuando varias organizaciones comparten los mismos objetivos y toman la responsabilidad de participar activamente en el desarrollo de servicios compartidos.

La relación entre las distintas organizaciones cooperantes requiere el establecimiento de acuerdos, los cuales pueden ser informales, formales o, en su caso, con entidades legales separadas. En el modelo de asociación, las organizaciones miembros adoptan distintos roles y funciones, pero el fin conjunto es el de hospedar y copiar todas o parte de las colecciones de los participantes utilizando distintas tecnologías de colaboración que ayuden a mantener la información digital compartida segura, auténtica y siempre disponible con distintas réplicas distribuidas. La asociación mantiene similares ventajas a las de un club, donde los integrantes comparten buenas prácticas y mejora continua.

## 1.3 Modelos de colaboración en repositorios institucionales

La aplicación del modelo OAIS a los repositorios institucionales de las universidades y otras instituciones científicas y culturales es un caso complicado, pues en estos repositorios prevalece la función de acceso sobre la de preservación. Hitchcock y colaboradores (2007) han propuesto distintos escenarios de trabajo en los que se modifica el modelo OAIS para adaptarlo a la realidad de los repositorios institucionales. A continuación, vamos a comentar dos de estos escenarios o modelos, que en este caso están pensados para ser implementados de manera distribuida entre varias organizaciones.

### 1.3.1 Modelo de proveedor de servicio

Este modelo emplea una adaptación concreta del sistema OAIS, en el que las entidades funcionales se modifican de la siguiente forma. La entidad ingestora (*ingest*) se modifica a una interfaz de ingestora personalizada al repositorio institucional (*IR interface ingest*), la gestión de datos (*data management*) toma la personalidad del repositorio institucional (*institutional repository*) y el almacenaje de archivo (*archival storage*) se modifica para convertirlo en un proveedor de servicio de preservación (*preservation service provider*).

### 1.3.2 Modelos de red y federados

Los modelos de red y federados incluyen principalmente la colaboración de varios clientes u operadores, que comparten contenidos y réplicas similares en cuanto a compromiso de socio. Es decir, se comprometen a mantener copias idénticas bajo políticas y normativas establecidas por los participantes. En este caso la modificación del modelo OAIS se centra en el acceso, la ingestora y el almacenamiento.

Como podemos observar, los modelos de archivo con fines de preservación digital a largo plazo corresponden a una alta gama de flexibilidad de acuerdo con opciones, tipología de apertura de archivo, federaciones, distribución y demás componentes. Nuestro caso de interés se centra en estudiar la aplicación de la preservación digital de forma distribuida, asociada, en red y federada como modelo de integración de una estrategia de preservación.

## 2 Casos internacionales de preservación distribuida

A modo práctico, podemos decir que el modelo de preservación digital distribuida (PDD) es una combinación del modelo de opción de asociación de Brown (2013) y de la propuesta de modelo de red y federados de Hitchcock y colaboradores (2007), ya que implementan estrategias de preservación digital descentralizadas en modo de colaboración asociada y sobre una red de distribución organizacional. Skinner y Schultz (2010) estandarizaron la definición de la PDD técnicamente, pues, partiendo del modelo de referencia OAIS, se definieron los procesos y tecnologías que deberían estar presentes en cada uno de los procesos de envío, disseminación y preservación de objetos digitales que definen un modelo distribuido. En este toman preeminencia la existencia de una plataforma abierta de software, de una red de afiliados y el uso de un protocolo de organización que formaliza la estrategia de preservación digital en lo relativo a la colaboración y la distribución dentro de una red privada.

El modelo de preservación digital se basa en el principio de que cualquier sistema responsable de actividades de preservación digital debe mantener y distribuir copias de los ficheros digitales que resguarda en diferentes localizaciones geográficamente dispersas. En estas localizaciones los ficheros no se guardan como un simple respaldo, sino que se gestionan autónomamente siguiendo unos criterios acordados con los que evitar pérdidas de datos o como sostén de las actividades de preservación. En la práctica, la

distribución de los ficheros se realiza en distintas localizaciones que funcionan como servidores o nodos autónomos, siguiendo un software y una metodología que periódicamente realiza una actividad de monitoreo, validación y copia de los ficheros para garantizar su integridad digital.

Si en los sistemas de preservación digital regidos por un único organismo ya se pueden detectar multitud de variantes, cuando hablamos de preservación distribuida las alternativas técnicas y de organización aumentan (Zierau; Schultz, 2013). Por eso el estudio de casos internacionales de implementación del modelo de preservación digital distribuida puede ilustrar las razones que sustentan estas diferentes alternativas y las ventajas e inconvenientes que llevan aparejadas.

El estudio de casos internacionales se llevó a cabo desde un punto de vista exploratorio. La selección de tres redes de preservación digital distribuida se realizó de forma no probabilística y se definió de acuerdo con factores como la tipología de red de preservación digital (solo a redes de preservación digital distribuida), la disposición de comunicación (entrevista personal o vía correo electrónico) y la disponibilidad de información pública institucional (sitios web o repositorios) sobre los modelos de red. A partir de estos criterios se obtuvo un panorama descriptivo de casos, desarrollando individualmente sus antecedentes y objetivos, realizando un diagrama conceptual de su modelo de red y los elementos de composición funcional de aquellos. Una vez realizada la descripción de cada caso, se realizó un análisis e interpretación individual para integrar una matriz FODA de fortalezas, oportunidades, debilidades y amenazas de cada una de las redes de preservación digital distribuida.

Las redes seleccionadas como casos de estudio fueron tres: The National Danish Bit Repository, la Red Cariniana de Brasil y el modelo empleado por el Consorci de Serveis Universitaris de Catalunya (CSUC) en su participación dentro del MetaArchive Cooperative.

## 2.1 Caso National Danish Bit Repository

El objetivo funcional de la red National Danish Bit Repository (NDBR) es el de proveer un protocolo y un sistema de preservación digital con arquitectura distribuida que coordina las tres instituciones danesas participantes, que son The National Archives, The Royal Library y la State and University Library (Sørensen, 2014). Mediante una capa de coordinación se gestionan las diferentes acciones y pilares (unidades de almacenamiento) de cada una de las instituciones participantes, siguiendo los acuerdos establecidos de nivel de servicio (SLA). Los pilares de almacenamiento son evaluados y gestionados por la capa de coordinación de software propio, que mantiene la información sobre el estado del contenido digital almacenado en las instituciones participantes.

El modelo de NDBR conceptualmente utiliza elementos como cliente, protocolo, capa de coordinación y pilares. En lo relativo a las funciones técnicas, estas son definidas en el elemento de protocolo con diferentes acciones y prestaciones para los clientes (nodos de almacenamiento) que representan a los repositorios o el software de las organizaciones, así como la forma en la que pueden acceder a ellas por el elemento de capa de coordinación que realiza las gestiones oportunas de distribución de réplicas digitales en los diferentes pilares (nodo o unidad de almacenamiento independiente) distribuidos en las tres instituciones participantes. El modelo de NDBR utiliza diferentes tipos de medios de almacenamiento como propuesta alternativa a la dependencia de tecnologías únicas, integra unidades de disco magnético, unidades de cinta, unidades de estado sólido y permite la posibilidad de conectar unidades en la nube.

El cliente de este modelo forma parte de los servicios del nodo de almacenamiento a nivel del software de la institución, el cual solicita servicios y procesos al protocolo dentro de la capa de coordinación que desarrolla las actividades de coordinar los diferentes procesos y servicios (incluida la replicación). Las diversas acciones, procesos y servicios refieren por igual a los pilares o nodos de almacenamiento compartidos.

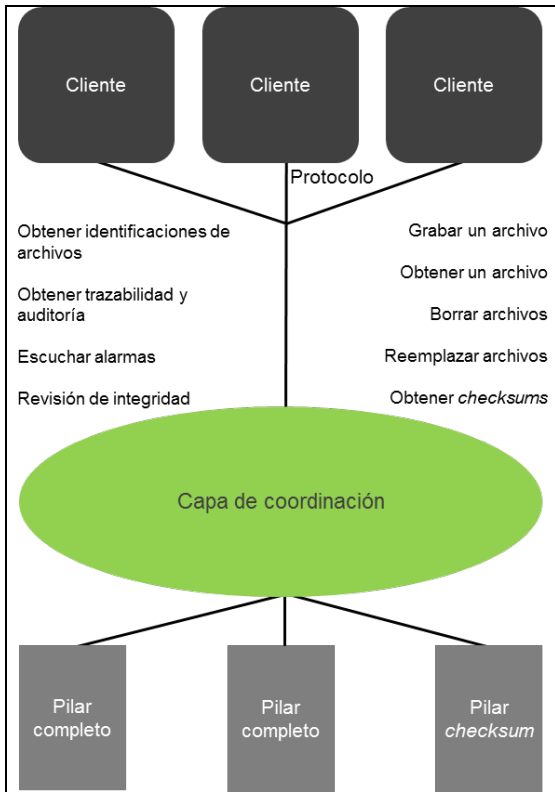
Figura 1. Modelo de visión conceptual y funcional de NDBR

En el modelo de NDBR destaca particularmente que cada una de las instituciones puede gestionar diferentes pilares de almacenamiento independientemente o como unidades compartidas, y se utiliza un pilar exclusivamente para recoger los *checksums* de control de integridad de los ficheros, lo que da mayor autonomía a este tipo de registros y facilita su replicación por seguridad.

Hay que mencionar que en el modelo conceptual y funcional de NDBR los factores de seguridad y confidencialidad y la existencia de diferentes necesidades en las tres instituciones participantes los llevaron a crear un sistema de desarrollo propio. En su caso, dieron una gran importancia al hecho de que estaban custodiando y preservando material de una alta sensibilidad nacional (archivos nacionales) o de gran importancia cultural (The Royal Library y la State and University Library). En este caso se entiende su estrategia de mantener la independencia operativa y funcional en sus actividades de preservación sin hacer uso de servicios de preservación digital externos o cooperativas internacionales, pues esta opción forma parte del nivel de exigencias y necesidades propias de estas instituciones. Mantener la cooperación entre tres instituciones de primer nivel nacional no resulta fácil, y aún menos hacerlo con un desarrollo tecnológico propio y autónomo. Por ello creemos que el modelo de NDBR es un referente que debe seguirse.

## 2.2 Caso Cariniana de Brasil

Cariniana es la Red Brasileña de Servicios de Preservación Digital. Está coordinada por el Instituto Brasileiro de Informação em Ciência e Tecnologia (IBICT), del Ministério da Ciência, Tecnologia e Inovação (MCTI), y está formada por distintas universidades y centros de investigación de Brasil.



Metodológica y tecnológicamente, Cariniana es una red de servicios de colaboración de preservación digital construida sobre el software LOCKSS, desarrollado por la Stanford University (Márdero, 2013). Actualmente la red está integrada por veinte cajas de instituciones socias, cuatro cajas coordinadoras y dos capas de colaboración que gozan de diferentes prestaciones y necesidades en la red. Cada caja es un nodo de hardware con capacidad de almacenamiento y dotado de software de interconexión con el resto de los nodos. Cabe señalar que este modelo de red privada de preservación fue creado inicialmente para la preservación de las revistas científicas de las universidades y los centros de investigación de Brasil, pero después su objetivo se amplió para favorecer la preservación de otros tipos de materiales mediante la creación de diferentes subredes temáticas.

El modelo de organización de la red Cariniana distingue tres grupos de formas de colaboración. En primer lugar, están los colaboradores integrales, formados por instituciones públicas o privadas que aportan y comparten infraestructura de almacenamiento y que poseen

documentos digitales en sistemas electrónicos de gestión. Otro grupo es el de las instituciones colaboradoras, que participan en algún proyecto de preservación digital y poseen documentos digitales compatibles con los sistemas de preservación digital, pero que acaban siendo preservados por los colaboradores integrales. Por último, están los colaboradores individuales, que son personas que colaboran en proyectos e investigación en preservación digital.

Dentro del modelo de red Cariniana podemos observar a tres actores principales. Por un lado, encontramos la integración de una red y subredes privadas de instituciones de educación superior y centros de investigación; por otro, una institución coordinadora de la red (IBICT) y, por último, la adopción de la tecnología del sistema LOCKSS de preservación digital en modo distribuido.

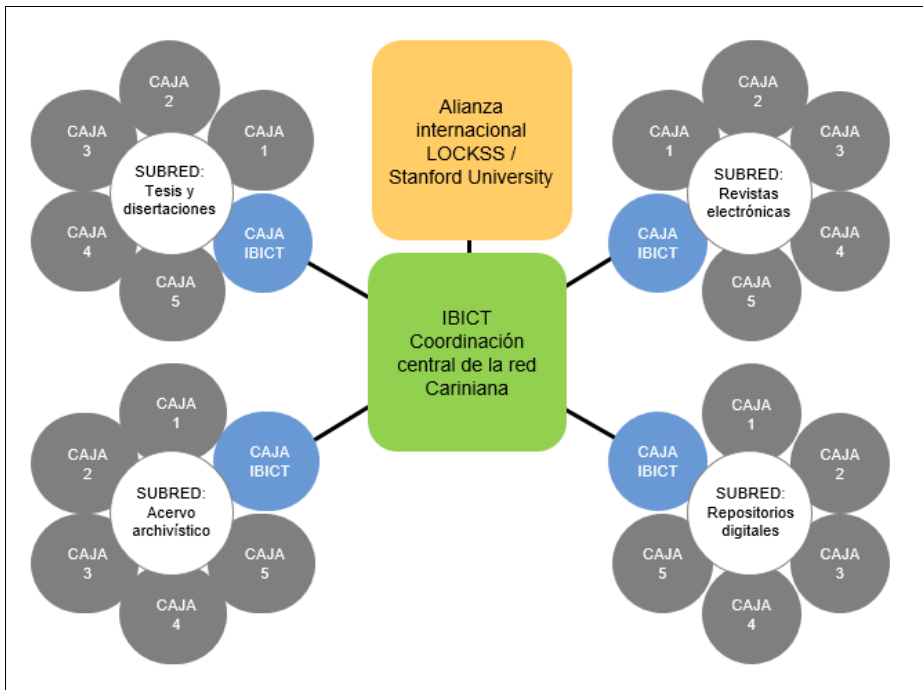


Figura 2. Modelo de red privada Cariniana y de integración de subredes

Específicamente, el modelo de preservación digital distribuida de Cariniana está formado por cuatro subredes con distinto contenido temático, las cuales, a su vez, están integradas 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 IBICT, que funciona como el nodo que conecta con la capa de coordinación IBICT de subredes y de la red general. Esta capa de coordinación IBICT está conectada con la capa de soporte internacional del sistema o Alianza LOCKSS.

Basándonos en la definición del modelo conceptual y funcional de PDD, en la red Cariniana las actividades de coordinación y soporte se realizan de forma centralizada en el núcleo operativo de la capa de coordinación (IBICT). Aquí las cajas de las subredes realizan las actividades de ingreso de contenidos y, gracias a los servicios ofrecidos por el sistema LOCKSS, también pueden copiar, monitorear, reparar y validar sus contenidos almacenados así como las réplicas de las demás cajas de su red, todas ellas conectadas con el protocolo Library Cache Auditing Protocol (LCAP).

En este modelo de servicios de preservación digital distribuida, la alianza internacional con el sistema LOCKSS permite tener acceso a la asistencia técnica de la Stanford University, sin que esta interfiera en la administración de la red Cariniana (Duarte; Márdero, 2015). De esta forma, el sistema de preservación brasileño, por un lado, cuenta con las ventajas de la alianza con el sistema tecnológico pionero en el ámbito de la preservación digital distribuida desarrollado por la Stanford University, lo cual aporta la seguridad, el soporte y la garantía de contar con un sistema sólido y robusto de PDD. Por otro lado, al contar con la figura de una institución coordinadora nacional como el IBICT, con infraestructura propia y experiencia en el trabajo colaborativo a nivel nacional, facilita la integración de las instituciones socias, que disponen de diferentes opciones de colaboración.

Hay que destacar que, en cuanto a la propiedad de los servidores, y a diferencia de la red NDBR que dispone de protocolos y servidores de desarrollo propios, 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 de cada institución, representa un factor importante que debe considerarse según sea el tipo de contenidos que tengan que preservarse y replicarse. Este podría ser un tema sensible que frenase la colaboración y la incorporación de nuevos socios, debido a temas de privacidad de la red o, en su caso, de la seguridad de la información de los registros que se intercambia entre cada una de las cajas y que, en lo relativo a los metadatos, recoge y verifica el servidor propietario.

### 2.3 Caso CSUC en MetaArchive Cooperative

El modelo de red de preservación digital distribuida en el que participa el Consorci de Serveis Universitaris de Catalunya (CSUC) es un modelo que integra diferentes capas de coordinación de tipo colaborativo y de cooperación. El CSUC emplea para la preservación digital de tesis doctorales el servicio externo del sistema MetaArchive Cooperative.

El CSUC nació en el año 2013 de la fusión del Consorci de Biblioteques Universitàries de Catalunya (CBUC) y el 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, formado por la Generalitat de Catalunya y diez universidades catalanas. Dentro de los diferentes servicios que se ofrecen en el CSUC se encuentra el repositorio *Tesis Doctorals en Xarxa (TDX)*, integrado a efectos de preservación en la red MetaArchive Cooperative.

El CBUC ya contaba con experiencia en la utilización de diferentes mecanismos de respaldo con objetivos de preservación digital, bien a nivel local (centro de datos del CESCA) o regional (copia de seguridad en otras ciudades de Cataluña), pero para el repositorio *TDX* se buscó un sistema ya más avanzado que asegurase de forma más robusta la seguridad y la preservación de los contenidos. Así se optó por participar de los servicios de la red MetaArchive Cooperative, que utiliza el sistema LOCKSS como base de sus servicios de preservación digital distribuida e integra una capa de coordinación con aplicación web Conspectus, con la que se administra la replicación segura, y la implementación de conectores y colecciones entre los distintos participantes de la red. MetaArchive es un sistema cooperativo que permite la replicación de las tesis almacenadas en el *TDX* entre distintos nodos de almacenamiento de la red, la mayoría de ellos en el continente americano (Schultz; Gore, 2010).

La participación del CSUC en la red de MetaArchive Cooperative puede definirse funcionalmente como un nodo de almacenamiento y coordinación, ya que además de ser un nodo socio en la red coordina a la subred colaborativa de universidades que aportan contenidos al repositorio digital de contenidos (*TDX*). En cuanto a la organización, MetaArchive Cooperative permite que los socios participen según diferentes posibilidades de integración (*sustaining member*, *preservation member*, *collaborative member*). En el caso del CSUC, 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 se refiere a las capas de coordinación como elementos funcionales, en primer lugar, debemos hablar de la capa central de coordinación que gestiona MetaArchive Cooperative, la cual 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) en donde cada una de las réplicas debe residir, así como de las gestiones de solicitud a cada una de ellas para conectarse bajo la implementación de los conectores 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 hemos mencionado, mantiene una función doble de nodo de almacenamiento y coordinador de la subred de miembros que aportan contenidos al *TDX*.

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 que forman el nodo o caja de preservación digital distribuida en la red.

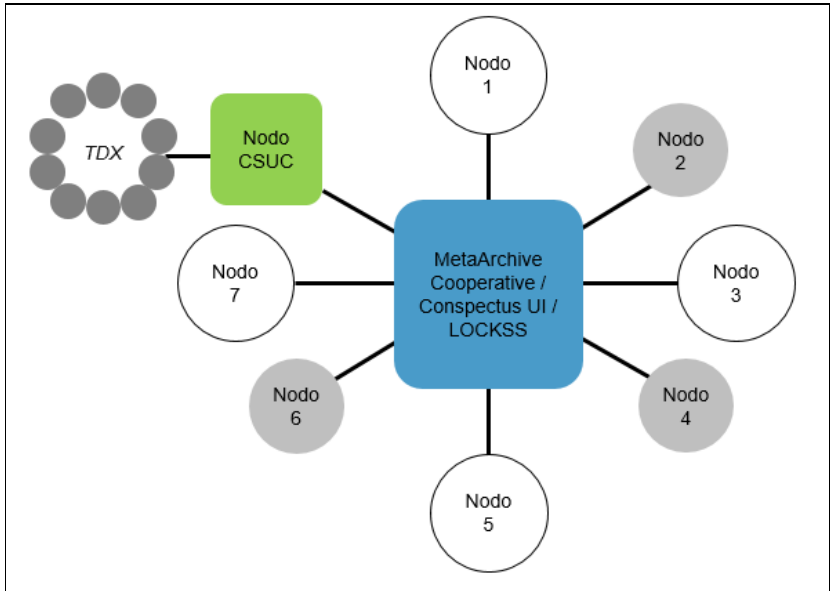


Figura 3. Modelo de la participación colaborativa del CSUC en la red MetaArchive Cooperative

En este caso, el CSUC, como intermediario o bróker coordinador de servicios universitarios, aporta una ventaja competitiva que parte de la especialización, evolución y experiencia en la integración colaborativa de ámbitos tecnológicos (CESCA) y bibliotecarios (CBUC) que van más allá de ser una simple capa de coordinación. Gracias a esto, el CSUC cuenta con elementos esenciales en todo procedimiento de preservación digital tanto en cuanto a roles técnicos como a organización institucional. Hay que recordar que el hecho de contar con un servicio externo de preservación digital distribuida no garantiza tener una solución completa, ya que, aun así, para cumplir plenamente con los objetivos requeridos es necesario disponer de todo un sistema de organización de infraestructuras técnicas y humanas institucionales paralelas.

#### 2.4 Contraste FODA de elementos funcionales en casos de modelos de PDD

En los tres casos presentados se estudió su organización y cómo esta presenta características que pueden favorecer o dificultar la función final de preservación. El resultado es un análisis FODA (tabla 1) con el balance y la consideración general de las fortalezas, oportunidades, desventajas y amenazas que presenta cada modelo de PDD.

Modelo	Fortalezas	Oportunidades	Desventajas	Amenazas
NDBR	<ul style="list-style-type: none"> <li>Desarrollo propio del sistema y protocolo de preservación digital distribuida BitRepository Project.</li> <li>Red privada basada en confidencialidad, seguridad y privacidad de documentos y datos de orden nacional.</li> <li>Capa de coordinación y protocolo de organización interinstitucional de tres instituciones.</li> </ul>	<ul style="list-style-type: none"> <li>Flexibilidad para la integración de nodos o instituciones adicionales.</li> <li>Evita dependencias tecnológicas al usar diversos sistemas y soportes de almacenamiento como estrategia.</li> <li>Personalización de sistema PDD a requerimientos extras y adecuación a cada institución participante.</li> </ul>	<ul style="list-style-type: none"> <li>Alto coste de la inversión en el desarrollo y mantenimiento de sistemas e infraestructuras técnicas y humanas.</li> <li>Poca cantidad de nodos replicados geográficamente en Dinamarca.</li> <li>Nula replicación internacional que aseguren copias fuera del país.</li> </ul>	<ul style="list-style-type: none"> <li>Dificultades para el mantenimiento sostenible de la red a largo plazo debido a los costes de operación.</li> <li>Posibilidad de pérdida de datos en tareas de migración y refresco debido a la diversidad de soportes de almacenamiento.</li> <li>Riesgo de pérdida de datos al no contar con más de tres nodos.</li> </ul>

Cariniana	<ul style="list-style-type: none"> <li>• Alianza e integración con el sistema pionero en preservación digital distribuida LOCKSS.</li> <li>• Red privada con método de trabajo, solución tecnológica y políticas propias para la gestión interinstitucional.</li> <li>• Cuenta con una institución coordinadora estratégica (IBICT) que funciona como nodo y capa de coordinación interinstitucional de la red de PDD en Brasil.</li> </ul>	<ul style="list-style-type: none"> <li>• Flexibilidad para integrar nodos o instituciones adicionales a la red privada en América Latina.</li> <li>• Replicación distribuida en seis cajas repartidas en instituciones de Brasil, que cuentan con una extensión territorial adecuada para su distribución geográfica nacional.</li> <li>• Pioneros en América Latina en implementar el sistema LOCKSS como red de preservación digital distribuida.</li> </ul>	<ul style="list-style-type: none"> <li>• Bajo coste anual, pero con dependencia técnica del soporte externo del sistema LOCKSS.</li> <li>• Enlace a servidores externos de la Stanford University (LOCKSS Prop Server) como propietario del sistema que recoge metadatos de las cajas en Brasil.</li> <li>• Nula replicación en nodos fuera del país.</li> </ul>	<ul style="list-style-type: none"> <li>• Complejidad de la coordinación a causa del crecimiento del número de nodos con diferentes necesidades.</li> <li>• Dificultades para el mantenimiento sostenible de la red a largo plazo debido a los costes de operación.</li> <li>• Riesgo de pérdidas en confidencialidad, seguridad y privacidad de la red al depender de servidores centrales externos (Stanford University).</li> </ul>
CSUC MetaArchive Cooperative	<ul style="list-style-type: none"> <li>• Uso e integración del servicio de preservación digital de MetaArchive Cooperative mediante el uso de la aplicación web Conspectus que gestiona el sistema base LOCKSS.</li> <li>• Alta replicación distribuida geográficamente fuera de la zona de origen del CSUC (Europa) en siete nodos distribuidos entre el norte y el sur de América.</li> <li>• Cuenta con dos capas de coordinación: la capa de servicio MetaArchive y la capa en la que el CSUC coordina los contenidos preservados de diversas instituciones en TDX.</li> </ul>	<ul style="list-style-type: none"> <li>• Experiencia del CSUC como consorcio cooperativo de educación superior en repositorios digitales y servicios universitarios.</li> <li>• Distribución geográfica rotativa y periódica de nodos de almacenamiento para asegurar nodos activos y sostenibles.</li> <li>• Bajo coste de inversión en infraestructura técnica y de operación del nodo participante en la red MetaArchive Cooperative.</li> </ul>	<ul style="list-style-type: none"> <li>• Mediano coste de servicio de participación como nodo colaborativo (grupo de universidades).</li> <li>• Dependencia técnica para la integración en la red y necesidad de que cada nodo desarrolle conectores especializados para conectarse a la red MetaArchive Cooperative.</li> <li>• Voto nulo en decisiones administrativas o técnicas de la cooperativa (debido a la condición de usuario colaborativo). También sin derecho a decidir en qué institución o zona geográfica se replicarán los distintos contenidos que deben preservarse.</li> </ul>	<ul style="list-style-type: none"> <li>• Riesgo de baja de socios colaboradores del CSUC, así como de MetaArchive Cooperative.</li> <li>• Dificultades para el mantenimiento sostenible de la red a largo plazo debido a los costes de operación.</li> <li>• Riesgo de pérdidas en confidencialidad, seguridad y privacidad de la red al depender de servidores centrales externos (MetaArchive Cooperative).</li> </ul>

Tabla 1. Comparativo FODA en tres modelos de PDD

### 3 Conclusiones

Examinando los ejemplos de casos en modelos de redes de preservación digital distribuida y los protocolos de su organización, estos nos muestran lo que muchas organizaciones no quieren ver y que se traduce en que la preservación digital distribuida no debe considerarse solo como el tratamiento de una implementación tecnológica, pues prioritariamente requiere de un trabajo paralelo de coordinación, administración y organización de tecnología, personas e instituciones.

En este estudio, hemos encontrado patrones similares en cuanto a la organización de los modelos de redes de preservación digital distribuida, ya que estas utilizan capas de coordinación representadas por una institución coordinadora, la cual centraliza y distribuye procedimientos, responsabilidades y roles a cada una de las instituciones que forman parte de una red de preservación digital distribuida. También es importante mencionar que hemos observado que las instituciones que forman parte o acceden a las prestaciones de una red de PDD previamente han cubierto e integrado una serie de condicionamientos técnicos necesarios, como la normalización y estandarización de la producción digital documental, la definición de políticas de gestión y difusión, la integración de sistemas de gestión documental y de repositorios, y la integración de sistemas de preservación digital centralizada. Además, se ha comprobado que la mayoría de las



instituciones participantes de la red cuentan con una cultura previa de cooperación interinstitucional o de historial de trabajo cooperativo que les permite gestionar mejor temas como la distribución de réplicas, los acuerdos de servicios y la propia gobernanza de la red.

También observamos que, para implementar un plan o programa de preservación digital distribuida, es obligatoria la existencia de una hoja de ruta personalizada de requisitos, necesidades y prestaciones de cada una de las instituciones participantes, ya que, si una de ellas no cumple con requisitos previos, tanto técnicos como de organización o económicos, será peligroso mantener y sostener un programa colaborativo de preservación digital.

Por último, se observa que una red de preservación digital distribuida tiende a tener un funcionamiento centralizado para los temas de coordinación y administración, distribuido para la organización de las responsabilidades técnicas, y colaborativo para temas de sostenibilidad económica.

## Agradecimientos

Agradecemos la colaboración del personal de Danish Bit Repository, Cariniana y CSUC al prestarse a contestar a nuestras preguntas.

## Bibliografía

Brown, Adrian (2013). *Practical digital preservation: A how-to guide for organizations of any size*. London: Facet Publishing.

D'Amato, Gianluca (2012). "Towards distributed digital preservation". *Notes from 2nd LIBER Workshop on Digital Preservation. Digital Preservation Coalition: Florence, Italy, 7th-8th, May 2012*.

<<http://www.dpconline.org/docman/miscellaneous/advice/755-dpc-liber-may2012/file> (<http://www.dpconline.org/docman/miscellaneous/advice/755-dpc-liber-may2012/file>) >. [Consulta: 06/11/2017].

Day, Michael (2008). "Toward Distributed Infrastructures for Digital Preservation: The Roles of Collaboration and Trust". *International journal of digital curation*, vol. 3, no. 1. <<http://ijdc.net/index.php/ijdc/article/view/60> (<http://ijdc.net/index.php/ijdc/article/view/60>) >. [Consulta: 06/11/2017].

Duarte Tavares, Maria de Fátima; Márdero Arellano, Miguel Ángel (2015). "Preservação do patrimônio científico das humanidades: a emergência da Rede Cariniana". *Cadernos de história*, vol. 16, no. 25, p. 30–45. <<http://periodicos.pucminas.br/index.php/cadernoshistoria/article/viewFile/P.2237-8871.2015v16n25p30/8890> (<http://periodicos.pucminas.br/index.php/cadernoshistoria/article/viewFile/P.2237-8871.2015v16n25p30/8890>) >. [Consulta: 06/11/2017].

Giaretta, David (2011). *Advanced Digital Preservation*. Berlin: Springer-Verlag.

Hitchcock, Steve; Brody, Tim; Hey, Jessie M. N.; Carr, Leslie (2007). "Digital Preservation Service Provider Models for Institutional Repositories: Towards Distributed Services". *D-Lib magazine*, vol. 13, no. 5/6.

<<http://www.dlib.org/dlib/may07/hitchcock/05hitchcock.html> (<http://www.dlib.org/dlib/may07/hitchcock/05hitchcock.html>) >. [Consulta: 06/11/2017].

Márdero Arellano, Miguel Ángel (2013). "Preservación digital distribuida: un modelo para América Latina". *BIREDIAL 2013 Conferencia internacional acceso abierto, preservación digital y datos científicos: San José, Costa Rica, del 15 al 17 de octubre de 2013*. <<http://eprints.rclis.org/20737/> (<http://eprints.rclis.org/20737/>) >. [Consulta: 06/11/2017].

OAIS (2012). *Reference Model for an Open Archival Information System (OAIS). CCSDS 650.0-M-2. Recommended Practice. Magenta Book*. Washington DC: Consultative Committee for Space Data Systems (CCSDS). <<http://public.ccsds.org/publications/archive/650x0m2.pdf> (<http://public.ccsds.org/publications/archive/650x0m2.pdf>) >. [Consulta: 06/11/2017].

Rothenberg, Jeff (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)*. <<https://www.clir.org/pubs/archives/ensuring.pdf>>. [Consulta: 06/11/2017].

Schultz, Matt; Gore, Emily B. (2010). "The Importance of Trust in Distributed Digital Preservation: A Case Study from the MetaArchive Cooperative". *iPRES 2010. Proceedings of the 7th International conference on preservation of digital objects: Vienna, Austria, September 19-24, 2010*. <<https://phaidra.univie.ac.at/view/o:185509>>. [Consulta: 06/11/2017].

Skinner, Katherine; Schultz, Matt (ed.) (2010). *A Guide to Distributed Digital Preservation. MetaArchive Cooperative*. Atlanta: Enducopia Institute. <[https://educopia.org/sites/educopia.org/files/publications/A\\_Guide\\_to\\_Distributed\\_Digital\\_Preservation\\_0.pdf](https://educopia.org/sites/educopia.org/files/publications/A_Guide_to_Distributed_Digital_Preservation_0.pdf)>. [Consulta: 06/11/2017].

Sørensen, Jan Dalsten (2014). "From stand-alone preservation to cross-institutional collaboration". *Archives and cultural industries. 2nd annual conference ICA (International Council of Archives): Girona, 11th-15th October 2014*. <<http://www.girona.cat/web/ica2014/ponents/textos/id119.pdf> (<http://www.girona.cat/web/ica2014/ponents/textos/id119.pdf>) >. [Consulta: 06/11/2017].

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

Voutssás, Juan (2012). "Long-term digital information preservation: challenges in Latin America". *Aslib proceedings*, vol. 64, no. 1, p. 83–96.

Zierau, Eld; Schultz, Matt (2013). "Creating a Framework for Applying OAIS to Distributed Digital Preservation". *iPRES 2013. Proceedings of the 10th International conference on preservation of digital objects: September 3-5, 2013, Lisbon, Portugal*. <<http://phaidra.univie.ac.at/o:377399> (<<http://phaidra.univie.ac.at/o:377399>) >. [Consulta: 06/11/2017].

---

#### **Cita recomendada**

Leija Román, David Alonso; Térmens Graells, Miquel (2017). "Preservación digital distribuida : lecciones de experiencias internacionales". *BiD: textos universitaris de biblioteconomia i documentació*, núm. 39 (diciembre) . <<http://bid.ub.edu/es/39/leija.htm>>. [Consulta: 27-12-2017].

