

IV ENCUESTRO INTERAMERICANO DE BIBLIOTECARIOS DE INFORMACIÓN Y ALFABETIZACIÓN (EIBIA 2026)

Evolución Tecnologías de Virtualización y Repositorios Digitales

Una Historia del Aislamiento de Entornos: La evolución desde el hardware hasta la flexibilidad del software

Expositor: Marcelo Lorca González

Biblioteca del Congreso Nacional de Chile | Programador & Bibliotecario Documentalista

30 de septiembre y 1 y 2 de octubre de 2026

Una Historia del Aislamiento de Entornos

Orígenes (1960-1990)

Aislamiento lógico

Primeras investigaciones teóricas sobre aislamiento de procesos, multitarea y emulación temprana de hardware en mainframes.

Consolidación (2000s)

Hipervisores y VMs

Consolidación de servidores físicos utilizando hipervisores de Tipo 1 y Tipo 2 (ej. VirtualBox) para ejecutar sistemas completos.

Revolución (2013)

Nacimiento de Docker

Lanzamiento de Docker como estándar open-source. Introducción de contenedores de solo lectura en capas ligeras.

Modernidad (Actual)

Ecosistema DevOps

Orquestación en la nube (Kubernetes/ECS) para escalabilidad elástica y preservación reproducible del software.

Conceptos Fundamentales de Virtualización y Preservación



Virtualización

Aislamiento por Hardware

- Creación de servidores virtuales sobre una máquina física usando un hipervisor.
- Útil para emular sistemas antiguos obsoletos (Dspace 5.x/6.x) bajo entornos originales discontinuados.

Cita: (AWS, 2026; Formanek, 2023)



Preservación Digital

Garantía de Acceso Futuro

- Procesos técnicos para que la información digital siga siendo legible a largo plazo.
- Evita la pérdida por obsolescencia física y lógica del hardware y de formatos de archivo.

Cita: (UNESCO, 2003; Hallo Carrasco, 2010)



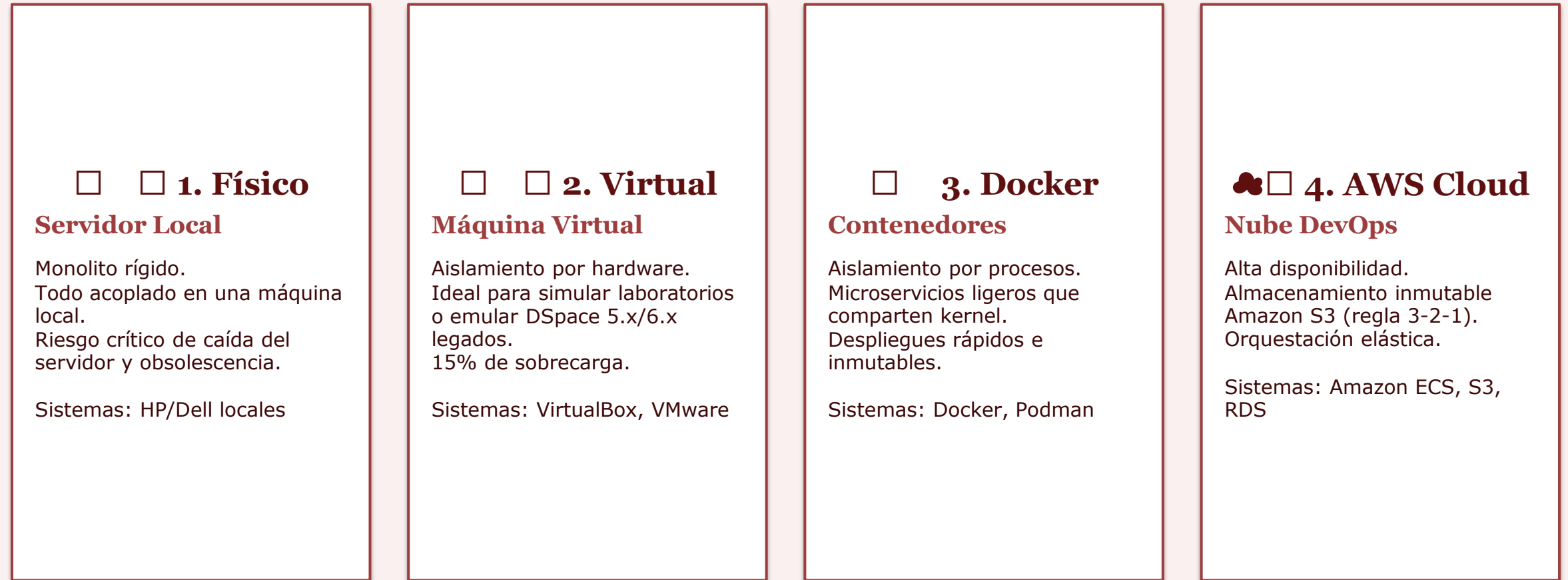
Tecnologías Forenses

Curaduría y Verificación

- Aplicación de análisis técnico (ClamAV, DROID, JHOVE) sobre archivos donados.
- Asegura una ingesta limpia al repositorio sin alterar metadatos originales.

Cita: (Garcia, 2024; IASA, 2023)

Evolución Física, Virtual, Contenedor y Nube



Referencias bibliográficas: Formanek, 2023; OpenWebinars, 2025; Xu et al., 2026

Los contenedores reducen hasta un 80% del almacenamiento frente a máquinas virtuales

Criterio	Máquinas Virtuales (VM)	Contenedores Docker
Nivel de Abstracción	Hardware completo (CPU, RAM, red) simulado a través de un hipervisor de software.	Nivel de sistema operativo (Procesos en espacio de usuario que comparten el mismo kernel).
Consumo de Recursos	Alto. Requiere asignación fija de hardware (vCPU/RAM) para cada SO invitado (sobrecarga de 10-15%).	Mínimo. Comparte recursos del host dinámicamente; sobrecarga ínfima de entre un 2% y un 5%.
Peso y Almacenamiento	Obeso. Imágenes de gigabytes (GB) porque cada VM incluye todo su sistema operativo invitado.	Ligero. Megabytes (MB). Estructura en capas de solo lectura AuFS compartidas, ahorrando entre un 60% y 80%.
Velocidad de Arranque	Lento. Minutos para completar el ciclo de arranque del sistema operativo completo.	Instantáneo. Segundos o milisegundos, ya que solo se inicia un proceso aislado en el kernel.
Aislamiento de Seguridad	Robusto. Aislamiento total de hardware. Una brecha en la VM no accede al host ni a otras VMs.	Ligero. Aislamiento lógico. Comparte kernel, requiere configuraciones extras (no root) para evitar escaladas.

Liderazgo de Mercado: Comunidad Activa a Largo Plazo

□ **DSpace domina el sector académico**

- Comunidad Mundial Activa: Con el 39.6% de adopción según OpenDOAR, DSpace ofrece un ecosistema de soporte robusto frente a alternativas.
- Flexibilidad Total: Su licencia libre BSD permite modificar e integrar el software localmente sin costos de suscripción recurrentes.
- Repositorio Robusto: Ideal para bibliotecas institucionales por su diseño modular específicamente orientado al modelo de preservación digital OAIS.

DSpace 39.6%, EPrints 11%, WEKO 9%, Digital Commons 5%, Otros 33%]

Microservicios y Desacoplamiento en DSpace 8

DSpace 8 consolida la arquitectura en microservicios, separando las interfaces en componentes totalmente independientes que se comunican de forma privada en red.

☐ **Frontend Angular**

Interfaz web rápida, responsiva e indexable por Google Scholar para visibilidad científica.

☐ ☐ **Backend REST API**

Lógica del servidor en Java. Permite interoperabilidad total con otros servicios (OAI-PMH, SWORD).

☐ **Apache Solr**

Motor dedicado para búsqueda facetada rápida, estadísticas de visitas y reportes en tiempo real.

☐ **PostgreSQL**

Base de datos relacional robusta. Resguarda de manera estructurada todos los metadatos de la biblioteca.

Orquestación y Mensajería: Kubernetes y Kafka

❑ **Kubernetes: El Conserje Automatizado**

Autoreparación: Si un contenedor del backend de DSpace 8 falla, Kubernetes lo destruye y levanta uno nuevo en segundos.

Monitoreo Continuo: Vigila el estado de salud de todos los microservicios sin requerir mantenimiento manual.

Control de Tráfico: Ante picos de bots o scrapers de IA, duplica automáticamente los contenedores para distribuir la carga.

❑ **Apache Kafka: El Cartero Veloz**

Mensajería Asíncrona: Permite que los microservicios se comuniquen de forma rápida en segundo plano.

Flujos Libres de Espera: Al registrar un libro, la interfaz no se congela; Kafka genera una cola de eventos y libera la pantalla.

Integración Forense: Al ingresar un archivo, notifica al instante a los contenedores forenses para validar formatos de forma autónoma.

Requisitos de Hardware y Aislamiento de Red

☐ ☐ Hardware Sano

Infraestructura Recomendada

- Servidor Base: 4 núcleos vCPU y 4 GB de RAM son suficientes para DSpace 7/8 departamentales.
- Alto Rendimiento: Procesadores AMD EPYC y discos NVMe SSD eliminan cuellos de botella al indexar metadatos.

Cita: (Hostinger, 2026; Formanek, 2023)

☐ Aislamiento en Red

Seguridad por Arquitectura

- Docker aísla la red interna de la internet pública.
- Backend y base de datos se interconectan mediante una red virtual privada (bridge).
- La base de datos relacional es invisible al exterior; solo se expone el frontend y Nginx.

Cita: (OpenWebinars, 2025; Garcia, 2024)

☐ ☐ Estaciones FRED

Cadena de Custodia Física

- Procesamiento seguro de donaciones mediante estaciones forenses FRED.
- Bloqueadores físicos de escritura (hardware write-blocking).
- Evita la alteración accidental de metadatos del sistema original (ej. fecha de creación).

Cita: (Garcia, 2024; IASA, 2023)

Ciberseguridad: Fin de Soporte y Mitigación

❑ ❑ El Peligro del Fin de Soporte

El fin de soporte de DSpace 5.x (Enero 2023) y DSpace 6.x (Julio 2023) significa que ya no se publican parches. Mantener estas versiones obsoletas deja las bibliotecas expuestas a ataques en internet.

- Log4Shell (CVE-2021-44228) - Severidad 10.0: Vulnerabilidad crítica que permitía control remoto. DSpace 7.1.1 y Solr 8.11.1 implementaron parches inmediatos.
- Spring4Shell (CVE-2022-22965) - Severidad 9.8: Afectaba entornos Java. Se corrigió de urgencia en DSpace backend 7.2.1.

❑ ❑ Mitigación Activa con Docker

- Ejecución sin privilegios (Non-Root): Se configuran los contenedores para correr bajo un usuario común. Un atacante jamás podrá adueñarse del servidor host.
- Aislamiento Forense en Ingesta: Las herramientas antivirus y de validación corren dentro de contenedores independientes. Si un archivo viene infectado con malware, este queda totalmente atrapado y neutralizado en el espacio de usuario aislado del contenedor.

Flujo de Datos bajo el Estándar OAIS

☐ 1. Entrada (SIP)

Submission Information Package

El paquete enviado por el productor o investigador.

- Acuerdo de depósito firmado.
- Metadatos descriptivos iniciales.
- Control de cuarentena antivirus.

☐ 2. Preservación (AIP)

Archival Information Package

El paquete definitivo resguardado para la eternidad.

- Estructura BagIt inmutable.
- Metadatos técnicos y de preservación.
- Enlace con PREMIS y METS.

☐ 3. Acceso (DIP)

Dissemination Information Package

El paquete generado dinámicamente para el público.

- Formato legible web (PDF/MP4).
- Dublin Core para indexación rápida.
- Listo para Google Scholar.

Esquema Flujo OAIS



SIP — Paquete de sumisión de información
AIP — Paquete de archivo de información
DIP — Paquete de diseminación de información

Modelo Premis

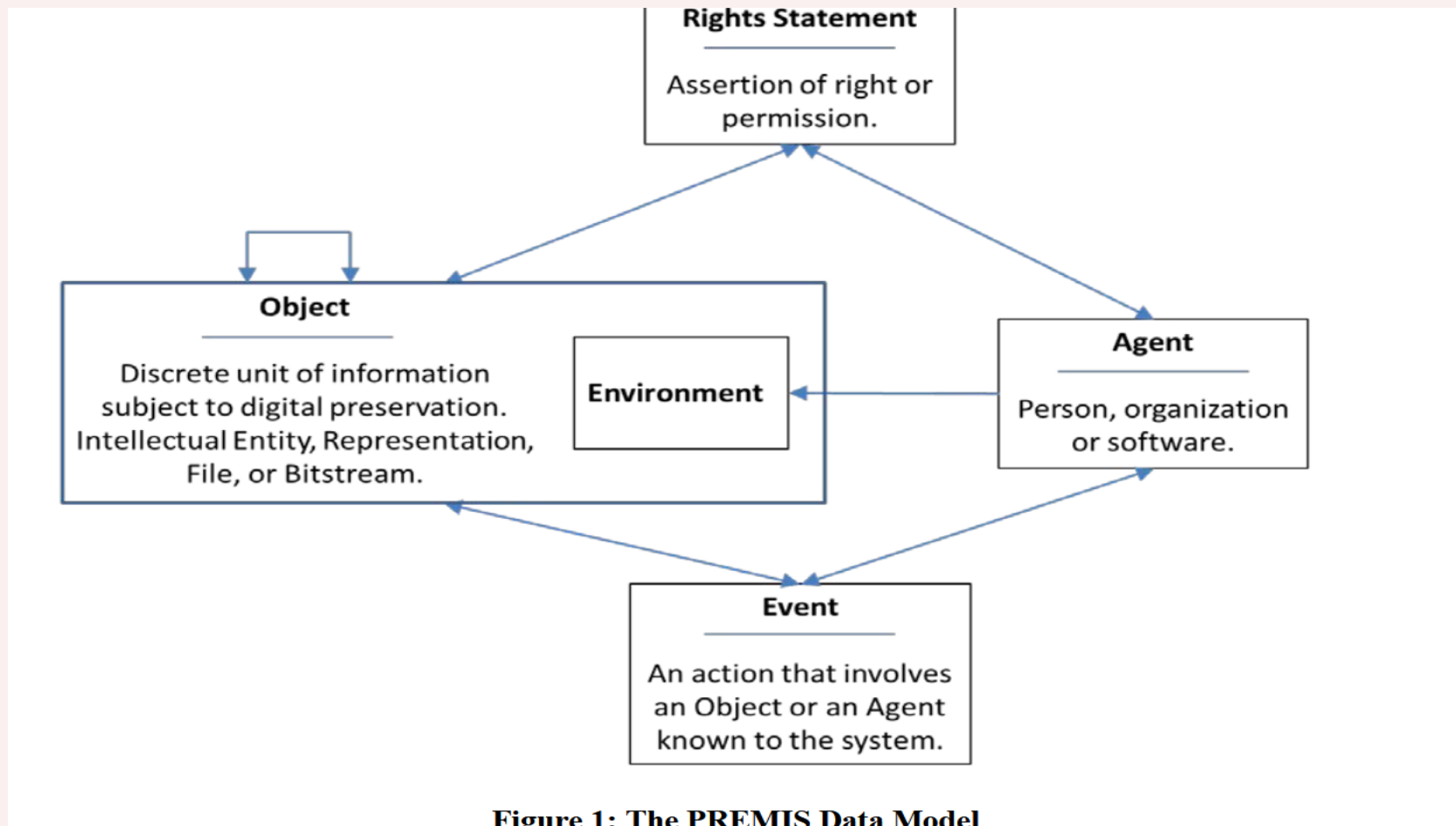


Figure 1: The PREMIS Data Model

Metadatos PREMIS: La Historia Clínica Digital

PREMIS (Preservation Metadata) es el estándar que documenta el historial clínico de un archivo digital para asegurar que siga abriéndose, entendiéndose y considerándose auténtico a largo plazo.

☐ **Modelo de Datos: 4 Entidades Conectadas**

- ☐ Objetos: El archivo físico (ej. PDF o bitstream)
- ☐ ☐ Eventos: Acciones de preservación (ej. ClamAV, JHOVE)
- ☐ Agentes: Personas o software involucrados
- ☐ Derechos: Permisos legales para migrar el formato

☐ **Las 3 Funciones de Preservación de PREMIS**

- ☐ ☐ 1. Integridad Física y Lógica:
Almacena checksums criptográficos para auditar que el archivo no haya sido alterado.
- ☐ 2. Control de Obsolescencia:
Registra el formato y software exactos requeridos para migraciones futuras.
- ☐ 3. Autenticidad y Procedencia:
Documenta el historial de custodia (quién lo creó y modificó) garantizando originalidad.

Flujo de trabajo Forense de Ingesta Segura con Docker

Las Tecnologías Forenses son programas que analizan los archivos antes de guardarlos permanentemente. Aseguran que entren limpios y cumplan con estándares técnicos, sin alterar las propiedades ni metadatos originales del donante.

❑ Etapa 1

ClamAV Virus Scan

Analiza el archivo en contenedor aislado antes de la cuarentena.

Cita: Cisco, 2024

❑ Etapa 2

Quarentena Retención

Espera activa configurable para descartar virus nuevos.

Cita: Garcia, 2024

❑ Etapa 3

DROID Format ID

Identifica el formato exacto en el registro PRONOM.

Cita: Archives, 2024

❑ Etapa 4

JHOVE Validación

Valida que el PDF cumpla la norma de codificación.

Cita: OPF, 2024

❑ Etapa 5

BagIt Empaque

Envasa el archivo descriptivo de forma inmutable.

Cita: Kunze, 2018

Referencias bibliográficas: Garcia, 2024; Kunze et al., 2018; Webb, 2003

Conceptos: PRONOM

☐ **PRONOM: El Registro Técnico de Formatos**

Formatos de archivos digitales del mundo. Es una base de datos o registro técnico creado y mantenido de forma pública por The National Archives (el Archivo Nacional del Reino Unido).

☐ ☐ ¿Cómo funciona de forma fácil?

Cada tipo de archivo (un PDF, una imagen JPEG, un documento de Word, etc.) tiene una estructura o una especie de "huella dactilar" interna única escrita en código binario. PRONOM se encarga de almacenar y listar todas estas huellas oficiales.

☐ ☐ Cuando un repositorio digital recibe un documento, utiliza herramientas de identificación (como DROID) que leen la estructura interna del archivo y la comparan con la base de datos de PRONOM. De este modo, aunque alguien le cambie el nombre o la extensión al archivo (por ejemplo, renombrando un archivo dañino .exe como si fuera un inocente .pdf), el sistema puede identificar con total precisión científica qué tipo de formato es realmente y qué versión exacta tiene.

Enlace oficial: <https://www.nationalarchives.gov.uk/PRONOM/>

Contacto Académico y Referencias Bibliográficas

Contacto Académico

Expositor: Marcelo Lorca González

Bibliotecario Documentalista
Programador
Biblioteca del Congreso Nacional

Datos de contacto:

- mlorca@bcn.cl
- marcelo.lorca@gmail.com

**IV Encuentro Interamericano de
Bibliotecarios de Información y Alfabetización
(EIBIA 2026)
30 de septiembre y 1 y 2 de octubre de 2026**

Referencias Bibliográficas

- Chandra, S., & Gokhale, P. A. (2012). Implementing OAIS Model for Digital Preservation. DESIDOC J., 32(4).
- Cisco Systems. (2024). ClamAV: Open source antivirus engine.
- Donohue, T. (2020). DSpace Resources and Releases Support Wiki. Lyrasis Foundation.
- Formanek, M. (2023). DSpace 7 Benefits: Is It Worth Upgrading? Information Technology and Libraries, 42(3).
- Garcia, M. (2024). Modernizing Digital Preservation with Docker Containerization. KAUST Library.
- Guenther, R. (2008). Flexibility vs. Interoperability When Implementing PREMIS in METS. D-Lib Mag, 14(7).
- Hallo Carrasco, M., & de la Fuente, P. (2010). Preservación digital en repositorios universitarios. RCCI, 4(3-4).
- Hostinger. (2026). DSpace VPS Docker | Configuración de repositorio en un clic.
- Kunze, J., et al. (2018). The BagIt File Packaging Format (V1.0). IETF RFC 8493.
- Masár, I. (2022). DSpace Release 7.0 Status and Roadmap. Lyrasis Wiki Pages.
- Webb, C. (2003). Directrices para la preservación del patrimonio digital. UNESCO.
- Xu, K., et al. (2026). A faster, more resilient digital repository: Migrating DSpace to AWS. AWS Blog.