

Preservación digital con **libsafe**

libnova – junio de 2014

Paseo de la Castellana, 153
28046 – Madrid

Tel: 91 449 08 94
Fax: 91 141 21 21
info@libnova.es

The logo for libnova, featuring the word "libnova" in a lowercase, sans-serif font. Above the letter "o" is a stylized icon of an open book, represented by two curved lines forming the pages and a horizontal line for the spine.

Preservación digital con libsafe

El presente documento es

CONFIDENCIAL/USO SÓLO POR PERSONAL AUTORIZADO

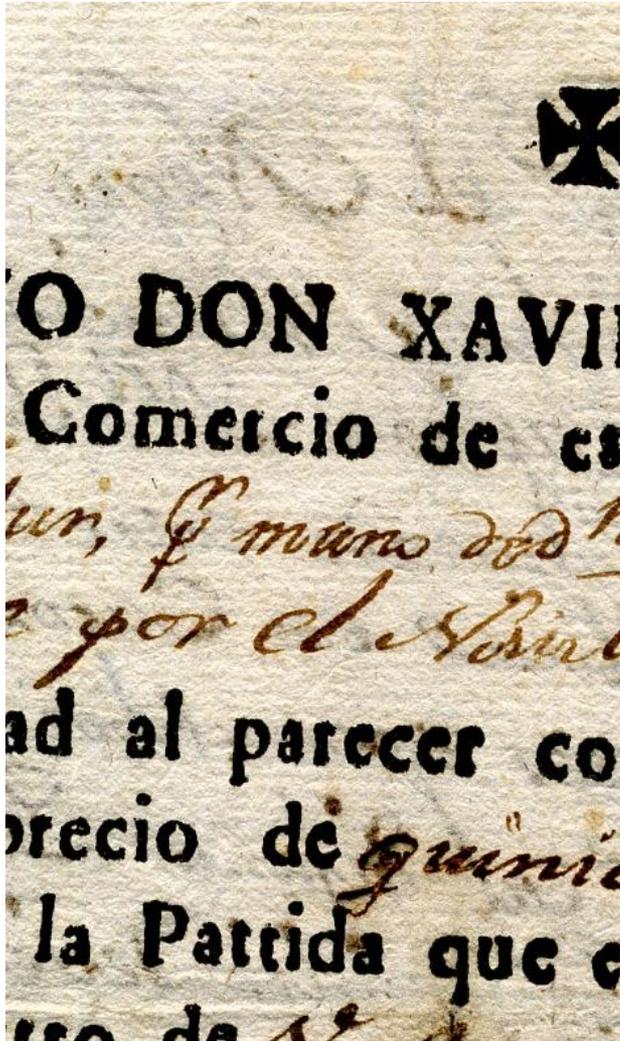
y no debe ser reproducido ni divulgado sin consentimiento previo y por escrito de **libnova**, y en ningún caso excluyendo las consideraciones relativas al objeto y alcance del propio documento.

Este documento y sus anexos contienen información confidencial o legalmente protegida y está exclusivamente dirigido a personal autorizado bajo acuerdo de confidencialidad. Si usted recibe una copia en otro caso no está autorizado a leerlo ni retenerlo.

En ningún caso, además, podrá modificarlo, distribuirlo, copiarlo ni revelar su contenido excepto en las condiciones previstas anteriormente.

Las imágenes contenidas en esta presentación son propiedad de libnova o han sido cedidas al dominio público para su reutilización.

Preservar objetos digitales: un caso real



2007

El fondo histórico se digitaliza masivamente.
Los masters se almacenan en CD/DVD y discos duros.

2014

El 20% de los soportes se han degradado.
Parte de los formatos utilizados han desaparecido.
El catálogo central, los soportes de los masters y los objetos guardados en cada uno no están relacionados.

Han pasado 7 años

¿Cuántos de aquellos máster continúan siendo usables?

Ahora tenemos métodos y tecnología
para que esto no vuelva a ocurrir.

Tabla de contenidos

1.

¿Por qué los máster son diferentes?

Los máster provenientes de procesos de digitalización tienen características especiales que afectan a cómo se manejan y conservan.

2.

Los métodos tradicionales no sirven

Los sistemas de gestión y almacenamiento de la información digital son poco eficaces y muy caros cuando se trata de gestionar y preservar masters a largo plazo

3.

La solución es la Preservación Digital

Para solucionar esta complejidad se han desarrollado metodologías y normativas específicas. Son muy potentes pero, por desgracia, complejas de aplicar.

4.

libsafe es preservación hecha fácil

libsafe implementa un modelo de preservación digital basado en OAIS e ISO14.721 de forma **completa, sencilla y sostenible.**



Características de los masters

- **Generan un gran volumen**

Sólo un libro puede ocupar decenas de gigas, mucho más que los derivados para difusión y que otros datos informáticos.

- **Variedad de formatos**

Tanto en los soportes de almacenamiento como en los contenidos de los objetos.

- **La frecuencia de consulta es baja**

Problemas potenciales de acceso o formato que puedan afectarle se detectarán demasiado tarde.

- **Suelen estar en soportes off-line**

Por ello, son difíciles de documentar, catalogar e incluir metadatos junto con el fichero master.

- **Tienen gran valor**

Tanto para la conservación del artefacto físico como para evitar incurrir de nuevo en costes de digitalización

¿Cómo se pueden
gestionar
los masters
para garantizar su
conservación y
usabilidad futura?



¿Cómo se han gestionado tradicionalmente?



Almacenando en cintas de backup, DVD o discos externos desconectados

- Decenas de soportes físicos, sistemas de ficheros y métodos de compresión en el mercado
- Pensados para duración de 5 a 7 años
- Sujetos a obsolescencia y degradación
- En caso de corrupción, esta se arrastra a todas las copias
- Difícil de catalogar, documentar, localizar, extraer y auditar el contenido



¿Cómo se han gestionado tradicionalmente?



Almacenando en cabinas de discos y servidores

- Alto coste: pensados para almacenamiento en producción, rápido y de acceso frecuente
- Al estar accesibles, los ficheros pueden ser fácilmente borrados o modificados
- Backup muy complicados, en algunos casos no son viables, por el volumen de información
- En caso de corrupción o cambios accidentales, estos se transfieren a la copia de backup
- No se mitiga el riesgo de obsolescencia de los formatos ni de falta de metadatos anexos



¿Cómo se han gestionado tradicionalmente?

Realizando un proyecto a medida



- Proyectos muy complejos
- Exigen alta dedicación, tanto en la puesta en marcha como en el mantenimiento posterior
- Por ello, resultan caros y poco sostenibles
- No incorporan el conocimiento y las prácticas de otras experiencias en el mercado
- Suelen aplicar procesos de preservación y soluciones parciales, sin contemplar el ciclo de vida completo del master



La solución: Preservación Digital

Para garantizar el uso futuro*
de los activos digitales se requiere
metodología, tecnología y actividades específicas.

Esto es la **preservación digital**

(*) **NOTA:** Para el uso futuro de un activo digital es necesario que esté disponible, íntegro, seguro, auténtico, localizable, accesible, y que seamos capaces de representar y visualizar sus contenidos.

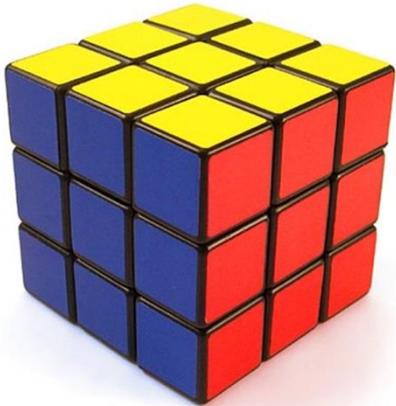


La preservación gestiona riesgos sobre su colección

	Fiabilidad ante...	backup	CD/DVD/HD	Storage	Preservación
Riesgo sobre los soportes	Degradación del soporte	XX	XX	—	✓
	Problemas de producto	X	X	X	✓
	Pérdida del soporte	—	—	—	✓
	Obsolescencia del soporte	XX	X	—	✓
Riesgos sobre los ficheros	Corrupción de datos	XX	X	—	✓
	Modificación/borrado accidental	✓	✓	X	✓
	Obsolescencia de formatos	X	X	X	✓
	Necesidad de migración de formatos	X	X	X	✓
Riesgos sobre los objetos	Objetos originales defectuosos	X	X	X	✓
	Defectos o carencia de metadatos	X	X	X	✓
	Histórico de cambios en los datos	X	X	X	✓
	Catálogo y localización de objetos	X	X	—	✓
	Seguridad y auditoría de acceso	X	X	—	✓



La preservación digital tiene muchas facetas



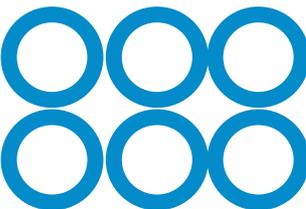
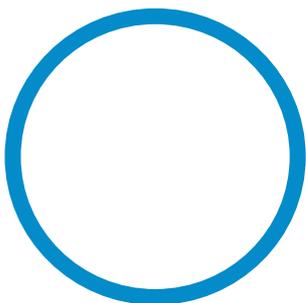
- **Inmenso volumen de información no estructurada**
Cientos de miles de imágenes, videos, audio y documentación escaneada que ocupa centenares de terabytes.
- **Gestión de múltiples copias**
Mantenerlas independientes, sincronizadas, auténticas y auditadas
- **Mantener el control de la colección**
Información periódica y exhaustiva del estado de la colección y capacidad para adaptarla a nuevos métodos y procesos.
- **Comprobar la validez del material a la entrada**
Virus, permisos, formatos, nombres y cualquier otra cosa que comprometa su uso futuro.
- **Evolucionar objetos y formatos**
Aproximadamente cada 7 años, la industria cambia los formatos y soportes mientras su contenido debe permanecer accesible.

La preservación es una tarea compleja que va mucho más allá del almacenamiento informático



libnova: preservation made easy

libnova ha creado una plataforma para **preservación digital**



Sencilla

Porque ha nacido en un proyecto real. Adaptamos los estándares y al día a día real de la gestión de una colección de masters digitales.

Completa

Porque contempla todos los pasos de la preservación digital, cumpliendo OAIS y la normal ISO 14.721,, desde el control de calidad de los objetos cargados, hasta la auditoría y transformación de los ficheros.

Sostenible

Porque la experiencia y la vigilancia tecnológica de **libnova** se reutilizan en cada proyecto. Además, los procesos complejos y repetitivos se automatizan para hacerlos eficientes y seguros.



Sencillez de uso, seguridad de su colección y tranquilidad mental, en un solo software.

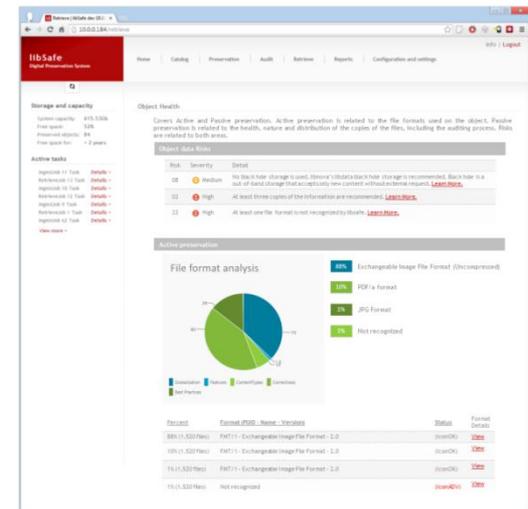
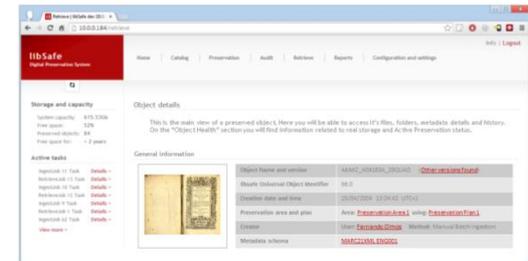
INGESTA Y
DISEMINACIÓN

AUDITORÍA Y
AUTOCURACIÓN

TRANSFORMACIÓN
Y EVOLUCIÓN

CONSULTA Y
VISUALIZACIÓN

- Verifica que la información es correcta según su plan de preservación.
- Ejecuta su política de disseminación en varias copias y sistemas de almacenamiento.
- Verifica que los contenidos siguen siendo válidos e idénticos a los preservados.
- Le notifica de riesgos si se detectan y los soluciona cuando es posible.
- Vigilancia tecnológica de metadatos, formatos y procesos óptimos para cada contenido.
- Asiste en la migración de los formatos cuando llegan a su punto de obsolescencia.
- Dispone de un catálogo integrado para buscar, localizar, previsualizar y extraer los objetos preservados garantizando la seguridad del contenido.





El sistema de almacenamiento complemento perfecto de **libsafe** para su sistema de preservación digital

- **Diseñado para la preservación digital**
Características óptimas de acceso, redundancia y seguridad.
- **Completamente integrado**
Tanto con **libsafe** y como con la gestión de su centro de datos.
- **Sin prestaciones superfluas**
Elimina las prestaciones no necesarias o contraproducentes para preservación (RAID5, compresión, de-duplicación, etc.).
- **¿Más seguridad? Sistemas híbridos**
libnova recomienda arquitecturas con diseminación híbrida en **libdata** y en su proveedor de almacenamiento habitual .
- **Y por supuesto, al mejor precio**
Gracias a su simplicidad, el mejor precio con la mayor fiabilidad.





Un proyecto de preservación con éxito

- **¿Por dónde empezar?**

De lo simple a lo complejo

En preservación, la experiencia es importante, y hacer algo siempre es mejor que no hacer nada. Comience por una colección o fondo que tenga objetos y formatos homogéneos, de forma que pueda concentrarse en poner en marcha la práctica de preservación en su organización. Una vez que esté integrada, podrá focalizarse en objetos heterogéneos de forma mucho más eficiente.

- **¿Cuáles son las primeras acciones para el éxito del proyecto?**

Elección del material, formatos estables y metadatos

Sin ninguna duda, la **elección de material** (si no puede preservarlo todo), la **elección de formatos** – buscando aquellos más estándares en la industria que le garanticen mayor longevidad, y la **inclusión de metadatos**. Ello le permite tener la colección de masters bajo control y acometer acciones más complejas en el futuro.

libsafe le asiste en estas tareas.

- **¿Cómo garantizar mis masters *para siempre*?**

Colección digital: control y evolución

En el mundo digital, que evoluciona rápidamente, es difícil que una tecnología sea para siempre. Para que su información sí dure para siempre por encima de la tecnología hay dos acciones que debe tomar. La primera, mantener la colección de masters controlada y documentada. La segunda, tomar decisiones que garanticen la colección hasta el próximo cambio tecnológico, en el que –gracias al control– podrá tomar las nuevas decisiones con garantías de éxito. **libsafe** es la herramienta perfecta para asistirle en este proceso.



Plataforma de preservación **libsafe**: licenciamiento

- **libsafe**

- Licencia de uso + volumen (Tb) de objetos preservados
- En función del tamaño de los masters de la colección que se preserva, independiente del número de copias
- La licencias base incluye 5Tb de preservación

- **libdata**

- En función del volumen total de almacenamiento requerido
- Sistemas de 36 y 72 discos (min. 12) de 4Tb
- Capacidades brutas desde 48Tb hasta 288Tb por unidad
- Número ilimitado de unidades dentro de un pool de almacenamiento
- Alta densidad: hasta 2.8Pb por rack



Plataforma de preservación **libsafe**: características

Procesos de ingestión	Limpieza del material	<p>La limpieza fija aspectos formales del material que se va a ingestar:.</p> <ul style="list-style-type: none">• Verificación y corrección de permisos de ficheros• Verificación de caracteres no permitidos en nombres de ficheros y carpetas• Verificación del tamaño máximo de las sendas de carpetas• Borrado de ficheros y carpetas de sistema y temporales de aplicación• Identificación e inventario de formatos de ficheros con DROID• Ampliable con controles definidos por el usuario para materiales específicos
	Comprobaciones en fase de ingestión	<p>Las comprobaciones verifican la validez del contenido que se va a ingestar:</p> <ul style="list-style-type: none">• Verificaciones a nivel objeto, fichero o carpeta• Comprobación de objeto existencia y rangos de tamaño válido• Comprobación de nomenclatura y caracteres según el plan de preservación• Comprobación de validez de formato con JHOVE• Ampliable con controles definidos por el usuario para materiales específicos
	Metadatos	<ul style="list-style-type: none">• Precargado con esquemas Dublin Core, Marc21 e ISAD(G)• Posibilidad de incluir esquemas de metadatos definidos por el usuario• Posibilidad de leer desde ficheros XML u otros con formato definido por el usuario• Capacidad para conectar y cargar metadatos desde catálogo o base de datos
	Diseminación y archivo	<ul style="list-style-type: none">• libsafe es capaz de diseminar y auditar los objetos sin limitación en el número de copias• Las copias pueden estar en almacenamientos de diferentes tecnologías y en diferentes ubicaciones geográficas



Plataforma de preservación **libsafe**: características

Catálogo y recuperación	Criterios de búsqueda	<ul style="list-style-type: none">• Tres métodos de búsqueda de objetos: Navegación por la colección, búsqueda simple y búsqueda avanzada• La búsqueda simple permite buscar por textos en el nombre del objeto o en cualquier campo de metadatos.• La búsqueda avanzada permite especificar criterios de búsqueda en descriptores de metadatos concretos y combinar varios criterios• Los resultados de las búsquedas pueden filtrarse y ordenarse por cualquiera de los campos del resultado
	Ficha de objeto y previsualización	<ul style="list-style-type: none">• Una vez localizado un objeto, se accede a una ficha detallada del estado de preservación del mismo, incluyendo: nombre, metadatos, copias almacenadas y su estado, estructura de carpetas y ficheros, versiones, riesgos potenciales e historia de acciones.• Desde la ficha de objeto se pueden ejecutar acciones de visualización, auditoría y recuperación.
	Recuperación de objetos	<ul style="list-style-type: none">• El material preservado está disponible para recuperación a nivel de objeto, área de preservación o colección completa.• Siempre se recupera una copia del objeto, quedando la información preservada fuera del acceso directo y libre de riesgos de modificación accidental.



Plataforma de preservación **libsafe**: características

Gestión de los datos, auditoría y seguridad	Versionado, colisiones y borrado	<ul style="list-style-type: none">• Definición de conjuntos de metadatos únicos (por ejemplo, código de barras). En caso de conflicto, se solicita decisión por el operador.• Definición de conjuntos de metadatos de versión (por ejemplo, título). En caso de conflicto, se preserva el objeto como una nueva versión del mismo.• Los descriptores pueden estar en esquemas de metadatos distintos• Los objetos preservados no se pueden borrar
	Datos de seguridad	<ul style="list-style-type: none">• libsafe almacena información del objeto, incluida su huella digital y la ubicación de cada copia, en una base de datos central y en cada una de las copias.• Gracias a ello, en caso de error la colección puede recuperarse completamente desde cualquiera de las copias• libdata incluye capacidad de redundancia interna que permite recuperar los datos dentro de la propia cabina incluso en caso de pérdida de 2 discos
	Auditorías	<ul style="list-style-type: none">• libsafe audita automáticamente la integridad de toda la colección. El usuario recibe el informe que le garantiza que su colección sigue en perfecto estado de preservación y gestión.• Las auditorías pueden ser de discos, objeto, área o toda la colección.• Adicionalmente, el usuario puede ejecutar auditorías manuales.
	Procesos excepcionales	<ul style="list-style-type: none">• Los datos se almacenan de forma que en caso excepcional toda la colección y sus metadatos pueden recuperarse directamente desde los discos de preservación, incluso aunque use el sistema de redundancia interna de libdata (al contrario de los sistemas RAID tradicionales).

Preservación digital, en resumen...

1.

Los máster son diferentes

Tienen un gran valor. Sus características de gran volumen, formatos heterogéneos y baja frecuencia de consultas, hacen complicada su conservación.

2.

Los métodos tradicionales no sirven

Los backup, CD, DVD y otros sistemas de almacenamiento desconectado son económicos pero muy inseguro. El almacenamiento tradicional y los proyectos a medida son más eficaces pero poco sostenibles.

3.

La solución es la Preservación Digital

La solución es aplicar técnicas de preservación digital OAIS / ISO14.721. Sin embargo, estas metodologías tienen muchas facetas y son difíciles de aplicar en la práctica.

4.

libsafe es preservación hecha fácil

libsafe modeliza, automatiza y simplifica los procesos de preservación digital de forma **completa, sencilla y sostenible**.



Paseo de la Castellana, 153
28046 – Madrid

Tel: 91 449 08 94
Fax: 91 141 21 21
info@libnova.es

