




libnova
the digital preservation company


libsafe

Digital Preservation Made Easy.

LIBSAFE: DISEÑO Y EJECUCIÓN DE SU PLAN DE PRESERVACIÓN DIGITAL.

La plataforma **libsafe** es la solución de **libnova** para preservación digital y está especialmente indicada para preservación de **masters de procesos de digitalización** de patrimonio histórico y cultural que, por sus características de **gran valor**, **baja frecuencia de consultas** y **alto volumen** de almacenamiento, requieren de procesos específicos.

Principales características:

- **Enfoque completo:** Contempla el proceso de la preservación de forma completa y en todas sus facetas.
- **Aplicación eficiente y flexible:** Se rige por el estándar de preservación mundial más aceptado (ISO 14.721 – OAIS), pero mantiene el enfoque práctico y eficaz procedente de la experiencia de su aplicación en proyectos reales.
- **Modelo transparente:** El cliente mantiene total control sobre el estado y ubicación de su colección digital.
- **Plataforma activa:** **libsafe** permite tomar el control de la colección digital, posibilitando evolucionarla en el futuro para mantenerla siempre accesible.

libsafe modeliza y ejecuta los procesos siguientes, cuyas características y posibilidades técnicas se explican a continuación:

- Gestión e incorporación de metadatos.
- Ingestión y control de calidad del material a preservar.
- Diseminación en múltiples copias controladas y auditadas.
- Auditoría y control de estado de su colección.
- Catálogo y recuperación de objetos preservados.
- Evolución y transformación de formatos.
- Informes sobre el estado, riesgos, uso y estadísticas de la colección.

CARACTERÍSTICAS TÉCNICAS DE LOS PROCESOS DE PRESERVACIÓN CON LIBSAFE

Áreas y planes de preservación

Planes y áreas de preservación ilimitados y personalizados de forma independiente.

Las áreas y planes de preservación engloban la definición de los formatos de metadatos, controles de ingesta, procesos de auditoría automática y análisis y transformaciones de formatos que se aplicarán a un conjunto de objetos.

- Un área de preservación agrupa materiales, objetos y fondos conceptualmente coherentes (del mismo tipo y nivel de relevancia, con estructura de objetos y metadatos común).
- Un plan de preservación define la estructura de los objetos preservados, sus metadatos, los controles de ingestión y las auditorías automáticas de control que se le aplican.
- **libsafe** permite crear tantas áreas de preservación como requiera.
- **libsafe** permite asignar planes de preservación diferentes para cada área de preservación. Ello permite aplicar políticas óptimas para cada uno de sus tipos de materiales digitales dentro de la misma plataforma de preservación.

Formatos y objetos preservables

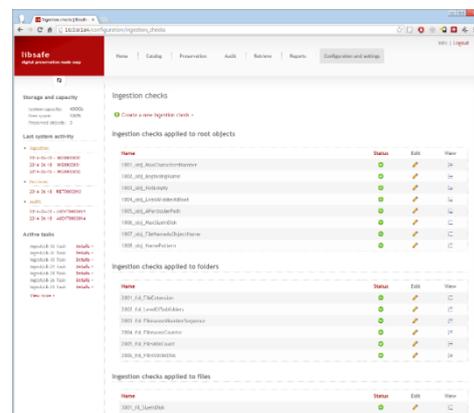
Preservación de cualquier formato de objeto digital.

- **libsafe** es neutral respecto al contenido preservado, siendo por lo tanto capaz de preservar cualquier contenido y formato de ficheros manejados habitualmente. Entre otros: imágenes, audio, video, textos, documentos de ofimática y PDF.
- Identifica más de 1.100 tipos de formatos diferentes. Realiza un perfil del fichero mediante DROID y verifica que los nombres, extensiones y formatos reales coinciden.
- Comprueba la integridad de los contenidos con JHOVE.
- Soporte a UNICODE para los nombres de objetos, metadatos, ficheros y estructura de directorios en los que se almacenan.

Ingesta y control de calidad

En el proceso de ingesta, **libsafe** recibe, analiza y registra los objetos digitales proporcionados por el usuario, garantizando su calidad, integridad y estado óptimo para iniciar un proceso de preservación digital.

Los trabajos de ingesta son creados por el usuario a través de una interfaz web muy intuitiva que le guía para proporcionar fácilmente la información necesaria. A partir de esta información, un **flujo de trabajo automatizado** se encarga de ejecutar las siguientes acciones:



Limpieza del material de ingesta.

- **Limpieza del material:** borrado de ficheros temporales o del sistema operativo, control de permisos, control de nombres que pueden generar inconsistencias en la plataforma actual o futura.
- **Identificación y control de validez de los ficheros:** **libsafe** identifica el formato del fichero con DROID y comprueba su validez con JHOVE.

Identificación de formatos y verificación de la estructura.

- **Comprobación de la estructura de los objetos:** verificación de que coincide con la definida en el plan de preservación en cuanto a:
 - Nombres de carpetas y de ficheros.
 - Presencia de ficheros, carpetas o datos críticos específicos.
 - Verificación de tamaños máximos y mínimos de los archivos y objetos.
- **libsafe** es capaz de ingestar objetos generados por herramientas de formación de paquetes OAIS SIP, bajo especificación bagIt y comprimidos en formato ZIP.

Una vez finalizado el proceso, el usuario recibe un informe detallado de resultados en pantalla, con posibilidad de recibirse también automáticamente en su correo electrónico.

Gestión e incorporación de metadatos

Neutral al esquema y representación de los metadatos.

libsafe analiza, incorpora y gestiona tanto esquemas estándar de metadatos como otros diseñados por el usuario.

- **libsafe** es neutral a la implementación. Almacena en base de datos los descriptores, semántica y valores de los metadatos, y es capaz de encapsularlos en cualquier formato estándar para su transferencia y consulta.
- Precargado con la definición de los esquemas Dublin Core, Marc21 e ISAD(G).
- Pueden configurarse otros esquemas estándar, o incluso modelos de metadatos personalizados por el usuario.
- Capaz de ingerir cualquier conjunto de metadatos encapsulado en XML y convertirlo a cualquier esquema compatible mediante hojas de transformación XML/XSLT.
- Permite la creación de filtros específicos para esquemas de metadatos particulares que pueden incluir incluso consultas Z39.50, OAI y otras.
- En el momento de la recuperación de una copia del objeto preservado, genera su METS asociado, utilizando los metadatos originales o los últimos almacenados y actualizados.

```

1 <?xml version="1.0" encoding="utf-8" ?>
2 <xsl:stylesheet version="1.0" xmlns:xsl="http://www.w3.org/1999/XSL/Transform"
3 <xsl:template match="/">
4
5 <object>
6 <object_metadata>
7 <xsl:for-each select="/doc/author">
8 <xsl:element name="metadata">
9 <xsl:element name="description"><![CDATA[author]]</xsl:element>
10 <xsl:element name="value">
11 <xsl:value-of select="/doc/author"/>
12 </xsl:element>
13 </xsl:element>
14 </xsl:for-each>
15
16 <xsl:for-each select="/doc/title">
17 <xsl:element name="metadata">
18 <xsl:element name="description"><![CDATA[title]]</xsl:element>
19 <xsl:element name="value">
20 <xsl:value-of select="/doc/title"/>
21 </xsl:element>
22 </xsl:element>
23 </xsl:for-each>
24
25 <xsl:for-each select="/doc/libra">
26 <xsl:element name="metadata">
27 <xsl:element name="description"><![CDATA[libra]]</xsl:element>
28 <xsl:element name="value">
29 <xsl:value-of select="/doc/libra"/>
30 </xsl:element>
31 </xsl:element>
32 </xsl:for-each>
33 </object_metadata>
34 </object>
35
36 </xsl:template>
37 </xsl:stylesheet>

```

Transforma e incorpora esquemas estándar o personalizados.

Gestión de datos de la colección

Detecta objetos duplicados o en conflicto incluso entre distintos esquemas de metadatos.

- **Grupos de colisión:** **libsafe** detecta objetos duplicados o conflictos en los contenidos de metadatos con significado equivalente. Por ejemplo, es capaz de determinar títulos coincidentes o referencias ISBN repetidas, incluso con esquemas de metadatos diferentes.
- **Resolución de colisiones:** En función de la configuración, el conflicto puede conducir a rechazar el objeto por duplicación, a generar una nueva versión vinculada a la ya existente o a requerir la intervención de un operador para resolver el conflicto manualmente.

Diseminación y múltiples copias

Múltiples copias controladas, sin ninguna limitación en el tipo y localización del almacenamiento.

- **libsafe** realiza múltiples copias en diferentes almacenamientos definidos en el sistema de forma transparente para el usuario. El número, destino y prioridad de las copias se define en el plan de preservación de forma que el técnico de preservación no tiene por qué conocer los detalles de la implantación de la tecnología, una vez estos han sido decididos.

- **libsafe** es compatible con la práctica totalidad de los sistemas de almacenamiento líderes en el mercado que sean accesibles desde un sistema Microsoft Windows Server (CIFS/SMB).
- Los almacenamientos de destino pueden ser de tecnologías y tamaños diferentes, así como estar ubicados en distintas localizaciones geográficas, siendo de hecho esta la configuración recomendada.
- Cada copia almacena información tanto de integridad y huella digital del contenido de los objetos almacenados (hash MD5) como información de la localización de las demás copias controladas.
- La información almacenada en los discos finales queda organizada de forma que es accesible por completo incluso sin el uso del software **libsafe**, en una situación de emergencia o mantenimiento.

Para colecciones críticas, **libnova** recomienda la realización de 3 copias de cada objeto preservado, manteniendo al menos una de ellas en unidades **libdata**. Para más información, vea el apartado “*requerimientos técnicos – almacenamiento de preservación*”.

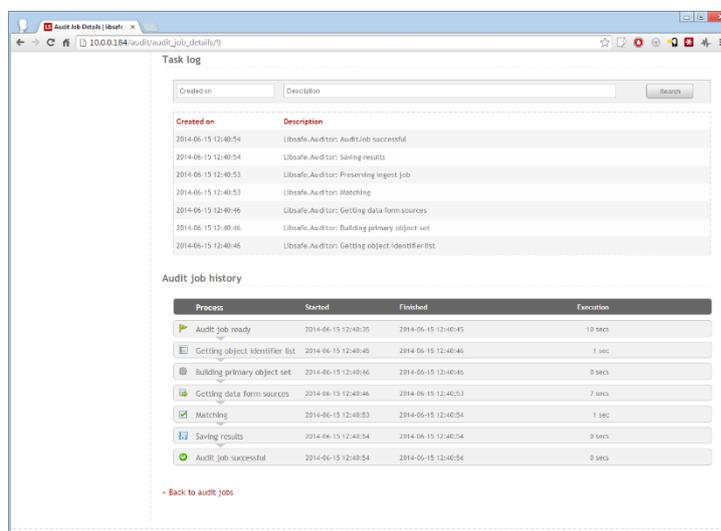
Auditoría y control de estado

Auditorías automáticas y programadas: garantía de integridad de su colección con el mínimo esfuerzo.

En los procesos de auditoría, **libsafe** analiza y verifica periódica y automáticamente que su colección continúa siendo accesible, que está en el estado en el que fue preservada y que está libre de riesgos de integridad y formato.

Cada copia almacena información de integridad y huella digital del contenido de los objetos almacenados. Esta información es almacenada también en la base de datos central de **libsafe**. Los trabajos de auditoría comprueban que todas las copias de esta información, así como los propios objetos, continúan siendo consistentes.

- **Tipos de auditoría.** **libsafe** dispone de auditorías automáticas preconfiguradas, auditorías automáticas programables por el usuario y auditorías manuales.
- **Cobertura de las auditorías:** **libsafe** ejecuta auditorías sobre un objeto preservado, un área de preservación, un disco del sistema o toda la colección completa.
- **Autocuración:** Siempre que sea posible garantizando la integridad de la solución, **libsafe** corrige los errores detectados. En otro caso, emite un informe detallado por correo electrónico y por la interfaz web, indicando los objetos afectados, los riesgos detectados y las posibles soluciones para que el usuario pueda ejecutar las acciones necesarias.



The screenshot shows the 'Task log' section of the libsafe web interface. It contains a table with columns for 'Created on' and 'Description'. Below this is the 'Audit job history' section, which includes a table with columns for 'Process', 'Started', 'Finished', and 'Execution'.

| Created on | Description |
|---------------------|---|
| 2014-06-15 12:40:54 | Libsafe_Auditor: Audit job successful |
| 2014-06-15 12:40:54 | Libsafe_Auditor: Saving results |
| 2014-06-15 12:40:53 | Libsafe_Auditor: Presenting largest job |
| 2014-06-15 12:40:53 | Libsafe_Auditor: Matching |
| 2014-06-15 12:40:46 | Libsafe_Auditor: Getting data form sources |
| 2014-06-15 12:40:46 | Libsafe_Auditor: Building primary object set |
| 2014-06-15 12:40:46 | Libsafe_Auditor: Getting object identifier list |

| Process | Started | Finished | Execution |
|--------------------------------|---------------------|---------------------|-----------|
| Audit job ready | 2014-06-15 12:40:35 | 2014-06-15 12:40:45 | 10 secs |
| Getting object identifier list | 2014-06-15 12:40:45 | 2014-06-15 12:40:46 | 1 sec |
| Building primary object set | 2014-06-15 12:40:46 | 2014-06-15 12:40:46 | 0 secs |
| Getting data form sources | 2014-06-15 12:40:46 | 2014-06-15 12:40:53 | 7 secs |
| Matching | 2014-06-15 12:40:53 | 2014-06-15 12:40:54 | 1 sec |
| Saving results | 2014-06-15 12:40:54 | 2014-06-15 12:40:54 | 0 secs |
| Audit job successful | 2014-06-15 12:40:54 | 2014-06-15 12:40:54 | 0 secs |

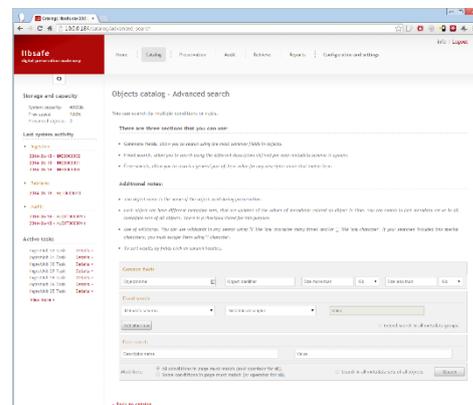
Acceso, catálogo y recuperación

La colección puede ser accedida mediante un catálogo con interfaz de búsquedas para su listado y pre-visualización, o bien mediante la recuperación de objetos preservados concretos.

Catálogo y previsualización

El catálogo es idóneo para la localización de objetos y la pre-visualización de sus contenidos.

- **Navegación** por la colección o por un área de preservación concreta, a través del listado de objetos.
- **Búsqueda simple** por el nombre de objeto o por cualquier metadato del mismo.
- **Búsqueda avanzada** por campos específicos de metadatos, tamaños de objetos, nombres y combinaciones de los anteriores.
- Una vez localizado un objeto, se presenta la ficha detallada del mismo (ver apartado "informes sobre la colección").



Catálogo y visualización de acceso rápido y flexible.

Recuperación segura, el objeto preservado siempre se mantiene aislado y protegido.

Recuperación de objetos

La recuperación está pensada para la extracción de los objetos y su manipulación fuera de **libsafe**.

- Posibilidad de recuperar objetos individuales, áreas de preservación completas o toda la colección.
- Los objetos recuperados se copian en una zona intermedia accesible para el usuario, desde la que se pueden mover, copiar o modificar.
- El objeto preservado siempre queda fuera del alcance de los usuarios, lo que evita errores de manipulación o borrado accidental.
- Si una vez modificado se desean preservar los cambios, deberá volver a ejecutarse un trabajo de ingestión. El objeto se preservará como una nueva versión o como uno nuevo; en ningún caso se sustituye el objeto original.

Informes sobre la colección

Informes de situación y estadísticos

Presentados para toda la colección o para un área de preservación concreta.

- Informes sobre el estado y resultados de los trabajos de ingestión realizados en **libsafe**.
- Informes sobre el estado y resultados de los trabajos de auditoría realizados en **libsafe**.
- Informes sobre el contenido de la colección: áreas de preservación, objetos preservados, perfil y formatos utilizados.

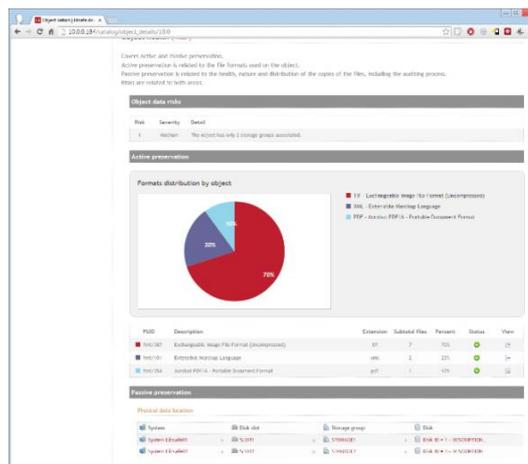
Ficha de detalle por objeto digital preservado

Realizados para un objeto específico de la colección, incluye:

- Datos generales y previsualización del objeto (nombre, trabajo de ingestión, área de preservación y otros).
- Esquema de metadatos utilizados y contenido de metadatos.

Informes flexibles y exhaustivos sobre su colección.

- Información de preservación pasiva: Estructura de carpetas y ficheros, perfil de formatos, localización de diseminación todas sus copias.
- Información de preservación activa: Acciones de auditoría y recuperación y acciones de transformación y evolución.
- Versiones del objeto (objetos preservados que se almacenan como versión de este)
- Historia de preservación del objeto (ingesta, auditorías, modificaciones de metadatos y recuperación del objeto).



Vigilancia tecnológica e informes de riesgos

Ejecutados sobre la colección completa, un área de preservación o un objeto, indica el estado de salud y riesgo del contenido preservado.

- Basándose en la base de datos de conocimiento de **libnova**, **libsafe** genera informes de riesgos sobre la colección: número de copias óptimas, formatos obsoletos o en riesgo de obsolescencia, formatos no reconocidos.
- Mediante procesos de vigilancia tecnológica, la base de datos de conocimiento y formatos de libnova estará siempre actualizada y a su disposición.

Nota: es necesaria conexión a Internet para la funcionalidad de vigilancia tecnológica.

Integración y ampliación de funcionalidades

libsafe está diseñado mediante un sistema de plugins que permite añadir nuevas prestaciones por parte de los usuarios, así como integrarse con otras aplicaciones y sistemas de información de la entidad.

El cliente puede añadir plugins. Estos pueden desarrollarse en cualquier lenguaje de programación y ejecutarse en Microsoft Windows, y deben cumplir las especificaciones de entrada/salida suministrada por **libnova** a los desarrolladores.

Los puntos principales de integración en el flujo de trabajo de **libsafe** son:

- Llamadas a programas diseñados o propiedad del cliente en la fase de ingestión para realizar limpieza o verificaciones de integridad avanzadas y personalizadas.
- Llamadas a programas diseñados por el cliente en la fase de carga de metadatos, bien para la incorporación de formatos y esquemas de metadatos personalizados como para la adquisición de los mismos desde un catálogo u otra base de datos o gestor documental.

Los informes de riesgos aplican la experiencia de libnova a sus objetos preservados.

Ampliable mediante plugins programables por el usuario.

REQUERIMIENTOS TÉCNICOS DEL SISTEMA

Almacenamiento de preservación

libsafe almacena su colección en dispositivos propiedad del cliente, lo que maximiza el control y seguridad de su colección. Es compatible con la práctica totalidad de los sistemas de almacenamiento y hardware de los fabricantes líderes en el mercado que estén certificados por Microsoft para su uso con Windows Server 2008 o superior o bien certificados por **libnova** para su uso junto con **libsafe**.

Entre otros, y de forma no limitada, acepta:

- Almacenamiento directamente conectado al servidor.
- Servidores NAS.
- Unidades accesibles a través de redes de almacenamiento SAN.
- Cualquier otra tecnología de almacenamiento accesible por CIFS/SMB.

libsafe es independiente de la tecnología subyacente en cuanto a tipos de disco, gestión de redundancia y seguridad (RAID) o ubicación física de los sistemas, siempre que la conectividad y ancho de banda sea la adecuada.

libsafe permite combinar almacenamiento de diferentes tipos en el mismo plan de preservación, consiguiendo que sus copias estén diseminadas de forma transparente y sencilla entre equipos de diferente tecnología y ubicación física, maximizando la seguridad de la información preservada.

Requerimientos del servidor

libsafe funciona sobre un sistema basado en Microsoft Windows. De forma orientativa, las necesidades básicas de este sistema son:

- Microsoft Windows Server 2008R2 con Internet Information Server (IIS).
- Procesador de cuatro núcleos de 64 bits.
- 12 Gb de memoria RAM.

- 700 Gb de disco para sistema y base de datos.

Los requerimientos dependen del régimen de uso, volúmenes de los trabajos de ingesta y políticas de auditoría, ya que los procesos necesarios para garantizar la seguridad de la información son intensivos en cálculos, acceso a memoria y a disco, siendo los anteriores los aconsejables para un caso de uso habitual.

Almacenamiento en disco final en función de la información a preservar y el número de copias. Almacenamiento intermedio en función de la cantidad y tamaño de los trabajos de ingestión o recuperación a realizar de forma concurrente.

- Conexión a Internet para mantenimiento de la base de datos de conocimiento y de formatos, actualizaciones, soporte y control de licencia, con agente de soporte y monitorización remota de **libnova** instalado.

libsafe puede instalarse en un entorno virtualizado y entregarse como una máquina virtual preinstalada. En caso de que desee esta configuración, consulte con su proveedor de virtualización la configuración equivalente.

Por favor, póngase en contacto con un representante comercial o distribuidor oficial de **libnova** para más detalles sobre las funcionales, características y requerimientos de su plataforma de preservación digital **libsafe**.