Preservación Digital

Para mercados de bienes culturales









Líder en preservación digital

Trabajamos continuamente para posibilitar que las organizaciones sean capaces de mantener vivos altos volúmenes de datos durante largos periodos de tiempo de una manera eficiente.



Trabajamos continuamente

> El equipo: de 5 a 27 en 3 años.> Clientes: de 0 a 77 en 5 años.> Ingresos: x3 en el último año.

para posibilitar que las organizaciones

> Clientes de primer nivel como la Biblioteca Nacional de España, La real Academia Española de la Lengua, La Casa de las Américas, entre otros muchos, usan nuestras soluciones a día de hoy.

sean capaces de mantener vivos

> Mantener vivo: Reproducción precisa de contenido genuino a lo largo del tiempo.

grandes volúmenes de datos

> Mantener unos pocos gigabytes es sencillo. La dificultad está en cómo gestionar hoy en día colecciones de Archivos y Bibliotecas con tamaños multi-petabyte.

durante largos periodos de tiempo

> Para nosotros 10 años es un corto periodo de tiempo, donde la obsolescencia de los formatos de archivo y la degradación de soportes físicos está muy presente.

de manera eficiente.

> utilizando para ello la menor cantidad posible de recursos . Compartiendo software y conocimiento con nuestros clientes. Factoría de software en España, enfoque internacional.

Core team



Antonio Guillermo Martínez Founder & CEO



Antonio Monís Partner & CFO



Juan Manuel Pérez
Partner & Commercial Director



Fernando Iglesias Partner & Board Member



Juan Carlos Romero USA Country Manager

Funda Libnova en 2010. CFO desde 2014.

Más de 18 años de experiencia en puestos de gerencia de diversas compañías tecnológicas. Más de 10 años de experiencia en preservación digital.

ICEX MBA. Combina capacidad estratégica y de ejecución.

CFO desde 2013

Más de 25 años de experiencia en puestos de gerencia, principalmente en el sector financiero.

Anteriormente CFO de ING Nationale Nederlanden y ocupando previamente cargos directivos en Citibank, The Chase Manhattan Bank y British Life.

CSO desde 2013

Más de 23 años de experiencia como director comercial y en otros puestos directivos en compañías de digitalización de fondos culturales.

Profundo conocimiento proyectos masivos de digitalización de fondos culturales, imagen digital y métodos de captura.

Miembro del Consejo de administración desde su fundación.

Experto en Business transformation y coaching.

IE Business School Executive MBA e ingeniero Senior por ICAI.

Mentoring de pequeña compañía a líder del mercado USA Country Manager desde 2015.

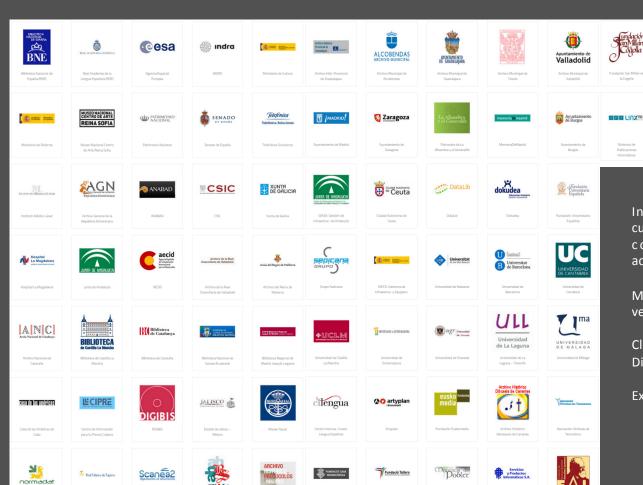
Abogado de Negocio Internacional con más de 20 años de experiencia en desarrollo de negocio de empresas foráneas en los Estados Unidos.

Master en International Business & Trade Law y N Y S E & N A S D Investment Advisor Licenses.

30 personas en el equipo en 2015.Q4



Clientes



Instituciones de gestión de fondos culturales de primer nivel, Universidades y compañías tecnológicas utilizan actualmente nuestras soluciones.

GUREAK

Más de 14Pb procesados, cargados o verificados durante 2014.

Clientes en 4 países Distribuidores en 6 países

Expandiéndonos a

EEUU (2015) África (2015) LATAM (2014)

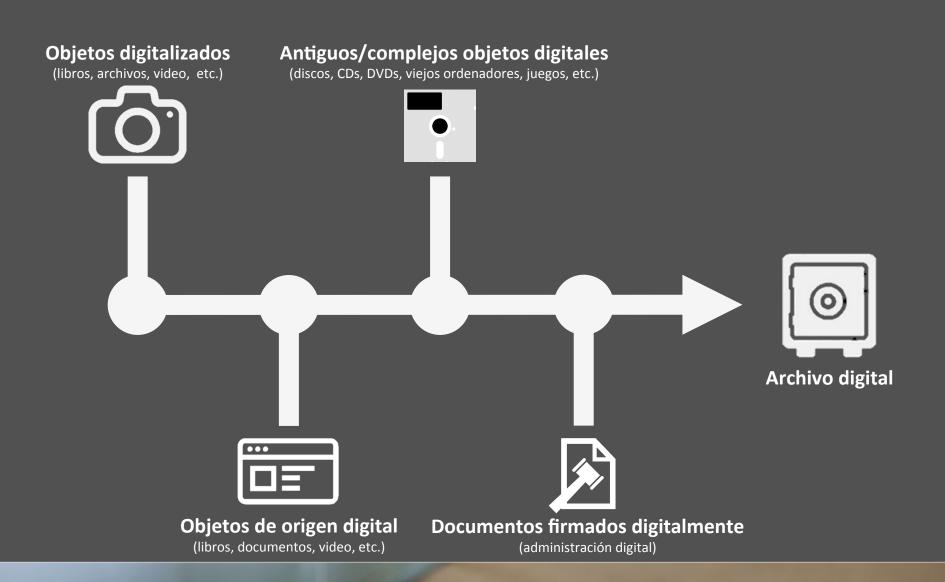






Obligatoriedad de preservar, incluyendo actualmente el formato digital

Mucha mayor complejidad que nunca antes





Alto volumen

Libros digitalizados en alta resolución pueden requerir varios gigabytes (masters).

Alto valor

No solo por su valor intrínseco (y el riesgo asociado de perderlo) del objeto original, sino también por los costes asociados de digitalización.

Baja frecuencia de acceso

Los problemas potenciales serán probablemente descubiertos cuando ya es demasiado tarde para solventarlos.

Diversidad de formatos y estructura de datos a lo largo del tiempo

Dado que estos proyectos cubren largos periodos de tiempo, los objetos pueden presentarse con diversos tipos de archivos, formatos de metadatos y estructuras varias.

Necesita estar "viva" a lo largo de periodos de tiempo realmente largos
 Los objetos preservados pueden necesitar preservación para siempre.



Las soluciones IT habituales, utilizadas de manera aislada, no son suficientes. Una aproximación combinada y especializada es imprescindible para evitar o mitigar los riesgos.

El personal IT sin una formación específica en preservación digital, no está acostumbrado a gestionar grandes volúmenes de información que va a ser escasamente consultada, sin un uso definido ni un periodo de retención prefijado.



¿Está seguro?

¿Backups consistentes?

Si la colección sufre cambios o se corrompe, los backups tradicionales realizados hoy no lo apreciarán hasta e día que el objeto se necesite. El periodo de retención del backup es mucho más corto que la ventana temporal necesaria para la detección del daño. Si la cinta de datos está todavía disponible, ¿es legible?, ¿existe aun el software y el hardware necesario para hacerlo posible?, ¿está la cinta en buenas condiciones?. ¿cómo combinar el backup con los objetos actuales?.

¿Control de carga - ingesta?

Cuando se trabaja con gran cantidad de objetos, ¿cómo estar seguro que estos están estructurados de la manera acordada?. Cuando múltiples objetos, con estructuras diferentes y tipos de datos se digitalizan al mismo tiempo, ¿cómo rastrear y verificar la consistencia del resultado? No sólo la estructura de archivos y carpetas, sino también los formatos de archivos, esquemas de cifrado y metadatos. Algunos formatos pueden también contener virus y malware, ¿cómo verificarlo?

¿Corrupción silenciosa?

Los datos digitales se corrompen por sí mismos si no se hace nada para evitarlo (efecto *bit rot*). Existen cuantiosos estudios que apuntan que el problema es mucho más común de lo que aparenta. El CERN calculó la corrupción sobre 8.7Tb (33.700 archivos) de datos durante 6 meses y encontró 22 archivos corruptos. Para grandes archivos esto significa 1 archivo corrompido de cada 750 cada año (acumulativo). La tasa de corrupción real calculada para el hardware actual es de un byte cada 3x10⁷. Todos estos datos corrompidos pueden mantenerse ocultos hasta el día que sea necesario consultarlos.



¿Está seguro?

• ¿Accidentes?

¿Es posible detectar?:

¿El borrado accidental de un archivo u objeto?

¿Migraciones incompletas (cambios de servidores, sistemas operativos, etc... que llevan a la pérdida de datos)?

¿Errores de administración?

¿Datos almacenados en soportes offline (cintas o discos USB)?

Según KPMG, más del 70% de los escenarios de pérdida de datos están relacionados con USB/ almacenamiento en disco duro, cintas y DVDs. Algunos estudios (Jones & Beagrie, 2001, página 130) destacan que a 25ºC (77ºF) la expectativa de vida de una Cinta magnética de datos es menor de 3 años (Lo mismo para CD/DVDs). Para grandes colecciones, los discos externos y las cintas son propensos a los fallos (mecánicos, caídas, accidentes de uso, etc.), es difícil mantener varias copias, etc. Para datos de alto valor el almacenamiento externo debería evitarse.

¿Evolución de los datos?

La tecnología evoluciona; los formatos datos también. Hoy, a veces se hace complicado abrir un archivo con 7 años de antigüedad. Cuando se gestionan altos volúmenes de información, es importante asegurarse de qué formatos están obsoletos y evolucionarlos periódicamente a nuevos formatos.



¿Está seguro?

¿Búsqueda? ¿Metadatos?

¿Es posible encontrar un objeto cargado hace 10 años usando sus metadatos?

¿Están los metadatos correctamente relacionados con el objeto?

¿Si los metadatos cambian en el catalogo, cómo replicarlos en el sistema de preservación?

• ¿Estándares?

Hay estándares y buenas prácticas seguidas por la mayoría de instituciones, ¿por qué entonces reinventar la rueda?

¿Es su repositorio auditable usando metodologías de auditoría estándar (TRAC – ISO 16.363, etc.)? Si se pierde información y no se seguían estándares ampliamente aceptados, quién debe asumir la responsabilidad?

3



Si existe información digital de alto valor desprotegida, son necesarias acciones de preservación digital.

Existen procedimientos estándar

Estos problemas, procedimientos y tareas han sido ya previamente analizados por instituciones especializadas. Hay marcos de trabajo conocidos y estables para la evaluación y la implementación: OAIS/ISO 14.721 / DRAMBORA / TRAC 16.363 / PAIS-PAIMAS / etc.

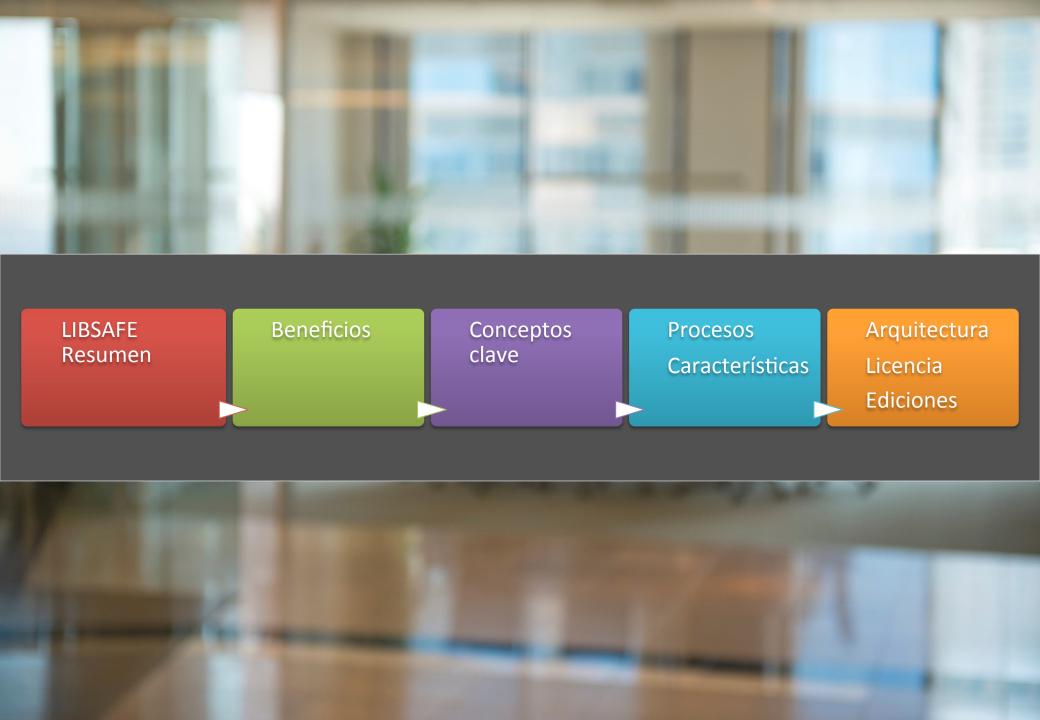
Es una tarea compleja

Hay involucrados una gran cantidad de tareas, riesgos y procedimientos. Sin el conocimiento ni las herramientas correctas, esta no es una tarea sencilla o sostenible. Para hacerlo eficazmente es necesario utilizar software de preservación.

LIBSAFE es un software que ayuda

Basado en experiencias propias y ajenas, hemos desarrollado una herramienta para ayudar alas instituciones a crear Archivos Digitales de largo plazo, automatizando muchas de las actividades necesarias. En proyectos de baja a media complejidad (que no involucren análisis forense), es posible que esta sea la única herramienta necesaria. En proyectos más complejos, puede aumentar la eficiencia.







Solución de Archivo Digital a largo (conforme a OAIS)

- Basado en estándares

Conforme a OAIS Certificable en ISO 16.363 (TRAC)

- Capaz

+2Pb peta bytes en instalaciones operativas. Imagen, audio, video.

- Potente

Validación, conformidad, evolución de formato de datos, migración de hardware y mucho más.

- Fácil de usar

HTML moderno, basado en web, rápido, UX, no usa java.

- Asequible

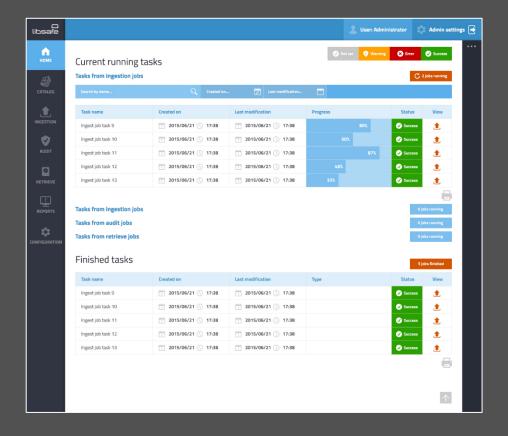
Licencia asequible. Pague sólo por lo que usa.

- Seguro

El sistema se verifica a si mismo. Creado conforme a directivas de software médico e ISO 27.001

- A prueba de futuro

Open Source si no estamos. Code escrow que convertirá LIBSAFE en open source si LIBNOVA desapareciera.





- Rápido de instalar. Solución llave en mano

Requisitos bajos o medios, nada que implantar- Instale y ya está listo para usar. Multitud de validadores y herramientas open source listos para usar. Interface muy intuitiva y fácil de usar (basado en navegador web).

- Reduce los costes y permite la preservación a bajo coste.

- > Desde el punto de vista de la preservación: Automatiza procedimientos, workflows y evolución de datos, asegura una buena calidad de preservación, fácil de certificar (TRAC).
- > Desde el punto de vista de IT: Proporciona al responsable del archivo autonomía para planificar la preservación. Por ejemplo, la migración de datos entre servidores es automática (instale el nuevo servidor, lance la tarea de migración del servidor A al servidor B y desconecte el servidor A cuando haya acabado; LIBSAFE migrará y verificará todos los datos). Permite combinar infraestructura de diferentes fabricantes bajo una única herramienta.

- Una única solución. Todo incluido

Excepto para algunas tareas tempranas de preservación digital (aquellas de carácter forense), la mayor parte de las necesidades están ya implementadas.

Fácil de conectar con otras plataformas (arquitectura *plug-in*, *hot folders* para carga automática, *URL queries* y recuperaciones, full API (80% de las funcionalidades para principios de 2016, 100% para la versión de finales de 2016).

- Tranquilidad

La plataforma es segura gracias a su diseño. Nada es eliminado, los archivos de control no son sobrescritos, todos los datos se guardan en al menos dos localizaciones (por ejemplo, base de datos y ficheros XML), sistema *WORM*.



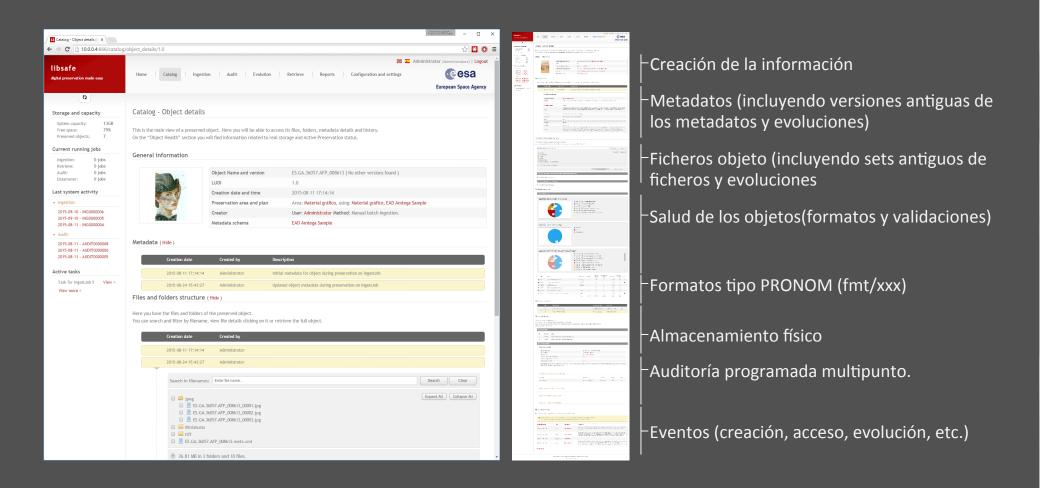
Objeto digital antes de la ingestión

■ igital objet
🛮 📗 metadata
metadata MARC21-METS
metadata MARCXML
metadata RDF
other XML metadata
Imaster images
.TIFF images
derivative files
JPG images
PDF-A derivatives
■ in the contents ■ in the
related audio-video
OCR results
TXT-RTF version

- Un objeto digital es una carpeta en la que los ficheros digitales y sus metadatos asociados son almacenados.
 - Masters, derivados y otros.
 - Metadatos en cualquier esquema y encapsulados en cualquier formato para identificar el objeto
 - El objeto es almacenado en una carpeta del sistema de archivos para su ingestión.

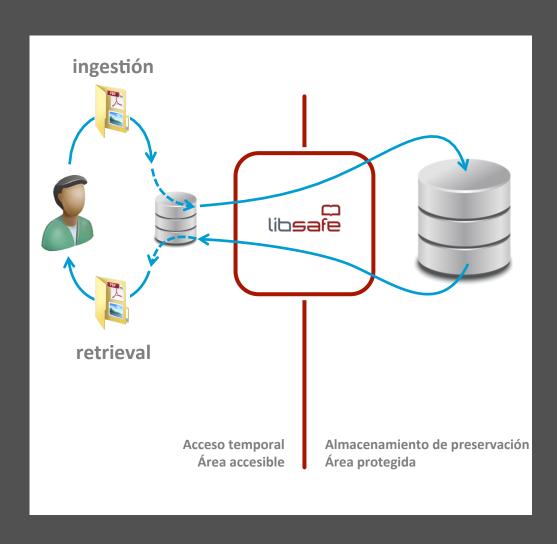


Objeto digital después de la ingestión





Almacenamiento aislado



Aislamiento para la seguridad

- En libsafe el almacenamiento de preservación siempre está aislado.
- En los trabajos de ingestión una copia del objeto es preservada y controlada por libsafe
- En los trabajos de recuperación el usuario sólo accede a copias temporales de los objetos, nunca a la copia preservada real.
- Este modelo incrementa la seguridad y elimina la posibilidad de operaciones peligrosas o error humano.



Política de preservación

Área de Preservación

 Un área de preservación identifica una colección coherente respecto al formato del objeto, esquema de metadatos y plan de preservación.

Plan de preservación

Un plan de preservación
 documenta las políticas de
 preservación: estructura, nombres,
 formato de los objetos, chequeos
 de ingestión, políticas de auditoría y
 transformación de formatos.



Nuestra oferta: Conceptos clave



Transparencia y libertad

LIBNOVA ha decidido crear la plataforma bajo un modelo de licencia, porque creemos que es mejor para nuestros clientes en el largo plazo, pero ha procurado que la plataforma sea tan abierta y libre de riesgos como sea posible:

- Code escrow. Open source si LIBNOVA fracasa

Todo nuestro código fuente e instrucciones de compilación son custodiadas en el despacho de un Notario y actualizadas cada seis meses, junto a él se custodian las instrucciones y procedimientos para cambiar el tipo de licencia a Open Source si fracasamos como negocio.

- Base de datos abierta, función full export y AIPs estable y auto-descrita

>El modelo de base datos (basado en MySQL) es abierto y está debidamente documentado..

>Hay una funcionalidad de "exportación total" en el software, que extrae todos los ficheros objeto y sus metadatos.

>Los objetos preservados son almacenados en disco sin ningún cambio, excepto algunos ficheros XML extra que reproducen el contenido de la base de datos para cada objeto= Únicamente teniendo la unidad de almacenamiento, sin el software o la base de datos, el acceso a todos los objetos está garantizado.

- Integración y expansión sencilla

> Hot folders para ingestión automática, pluq-in adaptables por el usuario para ingestión, transformación y evolución.

Nuestra oferta: Procesos



Descripción de los procesos



Ingestión

Es un workflow usado para cargar objetos digitales en el sistema de preservación. Incluye: saneamiento (limpieza), pre-proceso (transformación y enriquecimiento), extracción de metadadatos (descriptivos y administrativos, también caracterización), validación (estructura, formatos de ficheros, jhove, etc.), almacenamiento (realizando varias copias gestionadas) y una auditoría final (verificando la consistencia en todo el proceso). Lanzado desde el interface web , hot-folders (copia, después renombra) o llamada a RESTful API.



Recuperación

Un fichero, un objeto, múltiples objetos o la colección completa. LIBSAFE seleccionará la mejor fuente de datos en base a las prioridades de lectura del almacenamiento (servidores locales—más rápidos- antes que servidores remotos-más lentos-) y creará una carpeta temporal en la que el objeto será colocado (se evita por tanto el acceso directo). Se realizará un chequeo completo checksum del objeto para asegurar que la copia recuperada del objeto es una copia exacta del original. Lanzado desde el interface web o por URL usando catálogos URL.



Auditoría

LIBSAFE se auto chequea y chequea todos los objetos preservados de forma periódica. Cada N días (el intervalo es configurable), LIBSAFE lanzará una tarea de auditoría sobre un grupo de objetos, comparando los logs de ingestión, registros de la base de datos, ficheros objeto (varias copias) y archivos espejo XML (varias copias). Por ejemplo, si un pixel de una de las imágenes preservadas cambia, se generará una alerta.

Nuestra oferta: Procesos



Descripción de los procesos



Evolución

La tecnología evoluciona, haciendo obsoletos los objetos dependientes de la plataforma original. El motor dinámico de preservación de libsafe detecta formatos potencialmente obsoletos y permite la evolución de formatos (PDF 1.3 a PDF 1.6, por ejemplo). Este motor también permite la evolución del esquema de metadatos a lo largo del tiempo, así como la gestión de firma electrónica para e-administración y documentos electrónicos e-gobierno. Los procesos de evolución no sobrescriben información, sino que generan una nueva versión evolucionada del objeto.



Migración de almacenamiento

Tradicionalmente, cuando un departamento de IT quiere migrar de infraestructura, deben ocuparse de la migración de datos entre dos o más servidores. Esta funcionalidad de LIBSAFE permite migraciones de datos más sencillas y acordes con los estándares, ya que el sistema no sólo copiará la información del servidor antiguo al nuevo (y cambiará los punteros de los objetos a su nueva localización) sino que verificará fichero a fichero (usando hashes MD5), que la migración ha sido completada y es 100% correcta. Todo ello sin abandonar el navegador web.



Proceso Ingestión

Preprocessor

Aplica procesos definidos por el generación de nuevos derivados, renombramiento de ciertos archivos extracción de OCR, etc.

Validator

Practica chequeos sobre los objetos, su carpeta y archivos, para verificar su conformidad. Verifica la existencia de una carpeta con un nombre particular, nombres de archivo sin validación jhove, etc.

Auditor

Este módulo verifica todo el trabajo realizado por los módulos previos. Si alguna información debería estar en la base de datos, la verifica. Si un archivo almacenado debería tener ciertos hash MD5, le aplica el hash módulo se implementa la detección de corrupción end to end.



Desde este punto todos los trabajos sin intervención del usuario

>>>

Pre-SIP:









Realiza una limpieza inicial, borrando archivos no deseados, reparando atributos de archivo y permisos y verificando caracteres extraños en nombres de archivos o rutas largas

Explorer

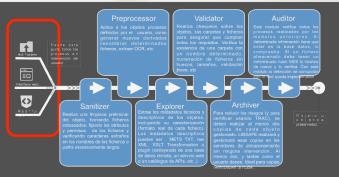
Extrae de los objetos metadatos descriptivos y técnicos, incluyendo su caracterización (formato real de archivo de un archivo). Los metadatos descriptivos XML podrían un METS, TXT, raw, XML, Transformación XSLT o plug-in customizado (extravendo desde base de datos remota, webservice, catalog API, etc.)

Archiver

certificar utilizando TRAC), at least objeto deberían ser realizadas. LIBSAFE hará y gestionará esas en los servidores de almacenamiento sin intervención. Al menos dos y tantas como el cliente quiera. Ideal para copias remotas/

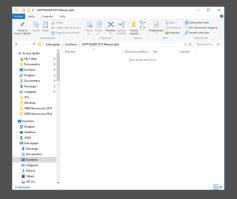
Objeto u objetos preservados. Se genera AIP





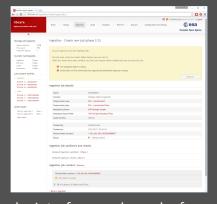
Trabajo de Ingestión: Creación





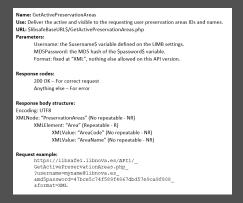
Hay una carpeta de ingestión automática por cada área de preservación. Esta funcionalidad se utiliza para automatizar la ingestión de objetos (de 1 a 100.000) provenientes de otros sistemas. La fuente debe copiar todos los objetos a una carpeta y ponerlos en el hot-folder. Todos los objetos serán ingestados y se generará una alerta si algo va mal..





Usar la interface web es la forma más sencilla para ingestar objetos. Selecciones el Área de Preservación y LIBSAFE creará una carpeta temporal para la ingestión, copie los objetos a ingestar y seleccione "Start". Se puede cerrar la ventana del navegador y el proceso continuará.

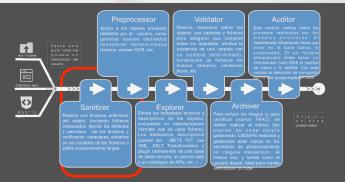




Es posible lanzar un trabajo de ingestión usando la API. Esta opción se utiliza cuando es necesaria una profunda integración entre sistemas.

LIBSAFE se integra de esta forma con el software de procesamiento de imágenes LIMB, de I2S.



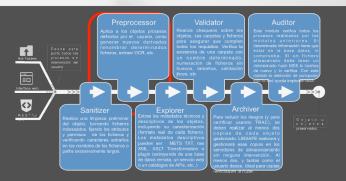


Trabajo de Ingestión: Sanitizer

Realiza una limpieza inicial del objeto, borrando ficheros indeseados, fijando los atributos y permisos y verificando caracteres extraños en el nombre del fichero o paths excesivamente largos. Es posible cargar *plug-in* específicos por parte del usuario.

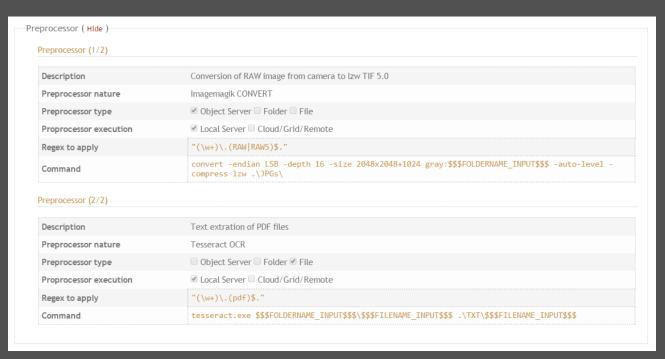
$lue{1}$ A sanitizer verifies some condition (like all folders and files are accesible by syste	m) and it may clean files or folders.	
1 A mandatory sanitizer verifies that all folders and files for all objects in the ingestisabled.	tion job are accesible by the system. It is manda	tory and it cannot be
f n (Optional) The sanitizers for folders are used to delete undesirable folders befor	e preserving the objects.	
🀧 (Optional) The sanitizers for files are used to delete undesirable files before pre	serving the objects.	
Mandatory Sanitizers: check long path and permission problems.		
Name	Parameter	View
Sanitizing FS		E
Sanitizers triggered by the applicaton of regular expressions for folders.	Sanitizers triggered by files	
No ingestion sanitizers were found.	Ingestion sanitizers are not set for this type in this preservation plan.	
Sanitizers triggered by the applicaton of regular expressions for folders.	Add ingestion sanitizer	
No ingestion sanitizers were found.	DeleteFilesExtension	•
	Select an ingestion sanitizer	
	DeleteFile DeleteFilesExtension	
	Mgester James description beter	tension match with -parameter-
	Parameters	
	File extension	





Trabajo de ingestión: Preprocessor

Aplica sobre los objetos procesos definidos por el usuario, como la generación de nuevos derivados, renombrar determinados ficheros, extraer OCR, etc.



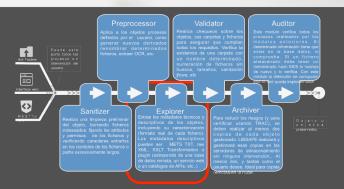
Hay tres tipos de preprocessors:

- Para objetos: Lanzado una vez por cada objeto en el trabajo de ingestión.
- **Para carpetas:** Se lanza una vez por cada carpeta para cada objeto en el trabajo de ingestión.
- Para ficheros: Se lanza una vez por cada fichero de cada carpeta para cada objeto en el trabajo de ingestión

Para cada uno, si el objeto/carpeta/nombre de fichero coincide con la expresión regex proporcionada, el comando será lanzado contra el objeto, carpeta o fichero.

De este forma podemos generar fácilmente y de forma automática JPGS para todos los ficheros TIFF almacenados en la carpeta "MASTERS".





Trabajo de Ingestión: Explorer

Extrae los metadatos técnicos y descriptivos de los objetos, incluyendo su caracterización (formato real del fichero). Los metadatos descriptivos XML pueden ser METS, TXT, raw XML, XSLT Transformation o un plug-in personalizado (extrayendo de una base datos, servicio web, API catalog, etc.)

E

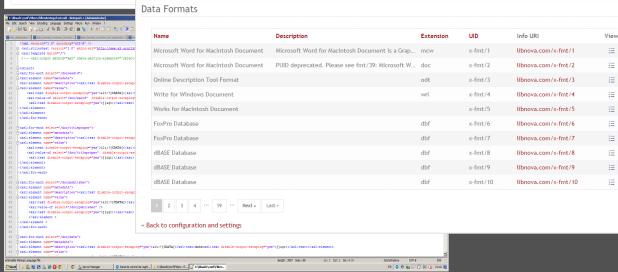
Ξ

1=

Ξ

=

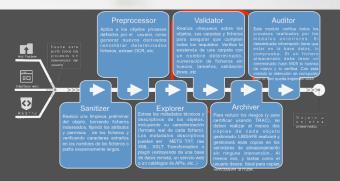


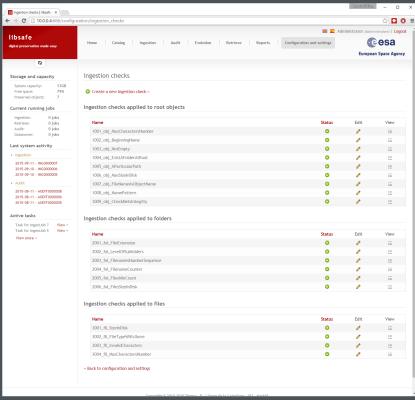


El módulo EXPLORER extrae información técnica v administrativa de los objetos:

- Administrativa: Para una buena preservación, se espera que los objetos vengan con metadatos, tales como el nombre del autor del objeto digital. LIBSAFE mantiene un esquema de metadatos para cada área de preservación e importa metadatos para cada objeto. Es posible importarlos desde cualquier clase de fichero XML, creando un fichero XSLT, pero también es muy fácil (aunque requiere programación por parte del usuario o un consultor) la integración con otras herramienta, lanzando una petición para que una API externa obtenga los metadatos, por ejemplo.
- Técnica: LIBSAFE caracterizará el formato real de los ficheros usando un caracterizador (DROID instalado por defecto, otros pueden ser usados).







Trabajo de ingestión: Validator

Realiza chequeos sobre el objeto, sus carpetas y ficheros para verificar su corrección. Verifica la existencia de una carpeta con un nombre determinado, numeración sin huecos, tamaños, validación jhove, etc...

Tal y como se define en el modelo OAIS el Productor y el Archivo deben tener la misma estructura de ficheros a ingerir.

Usando el Validator, LIBSAFE ayudará a asegurar que los objetos llegan como se espera.. El usuario puede cargar otros *plug-in*, pero a continuación se incluye una lista de los más representativos ya instalados:

Chequeos para objetos:

- Verificar que el nombre del objeto comienza con una "string" determinada, incluyendo que algunos ficheros comiencen con el nombre del objeto.
- Verificar que varias carpetas existen.
- Verificar la integridad METS MD5 (Verifica el hash MD5 incluido en el fichero METS (si ha sido generado) para asegurar la integridad del fichero).

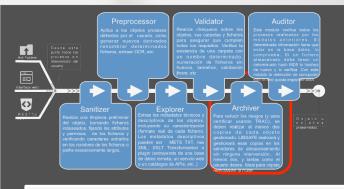
Chequeos para carpetas:

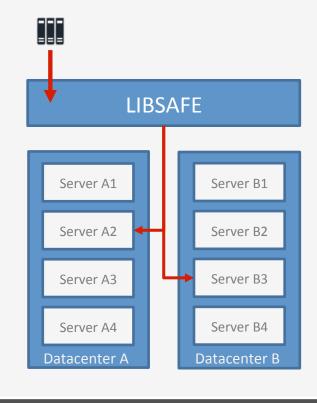
- Verificar que no hay un anidamiento excesivo.
- Verificar que, dentro de una misma carpeta, todos los ficheros están numerados de una forma determinada, y no hay huecos (ej: image001.jpg, image002.jpg, image004.jpg).

Chequeos para ficheros:

- Verificar un tamaño determinado, nombre del fichero, etc.
- Verificar el tipo de fichero con jhove.







Trabajo de ingestión: Archiver

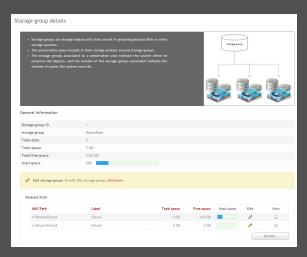
Para reducir el riesgo (y para certificar el uso de TRAC), se deben realizar al menos dos copias de cada objeto. LIBSAFE realizará esas copias y las gestionará en los servidores de almacenamiento sin intervención. AL menos dos y tantas como el usuario desee. Ideal para copias en la nube o remotas.

LIBSAFE gestionará el almacenamiento subyacente para guardar de forma segura múltiples copias de los objetos. Esto es una ventaja importante sobre otros sistemas o la simple replicación, porque las copias son gestionadas (LIBSAFE las gestiona siendo capaz de detectar una corrupción en cualquiera de ellas sin propagarla a las demás como en la sincronización tradicional).

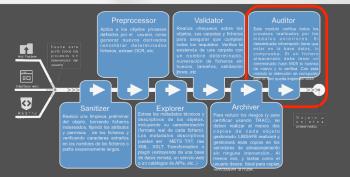
Esta característica también es muy conveniente ya que permite gestionar múltiples sistemas de almacenamiento de distintos proveedores de forma unificada.

Los Bloques de Almacenamiento se agrupan en pools de almacenamiento. Cada Área de preservación está relacionada con dos o más pools. Cuando un objeto es ingestado en una determinada Área de Preservación, LIBSAFE encuentra, para cada Grupo de Almacenamiento cual es el Bloque de Almacenamiento con prioridad más alta y que tiene suficiente espacio libre para la copia del objeto sin fracturarlo.

LIBSAFE es capaz de utilizar el 95% de los sistemas de almacenamiento disponibles en el mercado (montables usando CIFS), por lo que es independiente del fabricante.







Trabajo de Ingestión: Auditor

Este módulo verifica todo el trabajo realizado por los módulos anteriores. Si una determinada información debería estar en la base de datos el módulo lo verifica, si el fichero almacenado debería tener un determinado MD5 le aplica el hash de nuevo y lo verifica . Con este módulo queda implementada la detección de corrupción end-to-end.

Un vez que el objeto es finalmente preservado en LIBSAFE, la siguiente información está disponible:

- Fuentes del objeto que aún están en la carpeta temporal.
- Base de datos, describiendo sus ficheros, hashes de los ficheros, paths de almacenamiento de los mismos y metadatos administrativos.
- Detalles del Plan de Preservación (cuantas copias, donde, esquema de metadatos a usar (incluyendo campos obligatorios, etc.)
- Copia de los ficheros objeto en el sistema de almacenamiento.
- XML describiendo ficheros objeto, hashes de los fichero, paths de almacenamiento de los ficheros del sistema y metadatos administrativos almacenados en la misma carpeta.
- XML describiendo el contenido del bloque de almacenamiento en el que está almacenado el fichero (XML) que contiene información de todos los objetos almacenados en el mismo bloque de almacenamiento).
- Copia del XML describiendo los ficheros objeto, hashes de los ficheros, paths de almacenamiento de los ficheros del sistema y metadadtos administrativos, almacenados en la misma carpeta.
- XML describiendo el contenido del segundo bloque de almacenamiento en el que el objeto es almacenado (XML conteniendo información de todos los objetos almacenados en el bloque de almacenamiento).
- (y más si se preservan más copias del objeto)

El módulo de auditoria es un módulo independiente, escrito por expertos desarrolladores diferentes e independientes de los que han desarrollado los módulos precedentes, que carga de nuevo toda esta información y verifica su coherencia, detectando las menores discrepancias. Si el proceso de auditoría falla, los objetos no se consideran preservados y el objeto es devuelto al usuario.

Por ejemplo, si hay un fichero de más en una de las copias del sistema de archivos, o falta alguno, o tiene un hash MD5 diferente, el proceso fallará. Si hay un metadato obligatorio para determinados objetos y el metadato está vacío (debido a un fallo de otro módulo o a corrupción de la base de datos) se generará una alerta.

Este módulo asegura que todos los objetos ingestados sean finalmente preservados.



Trabajo de Ingestión

Preprocessor

Aplica a los objetos procesos definidos por el usuario, como generar nuevos derivados renombrar determinados ficheros, extraer OCR, etc.

Validator

Realiza chequeos sobre los objetos, sus carpetas y ficheros para asegurar que cumplen todos los requisitos. Verifica la existencia de una carpeta con un nombre determinado, numeración de ficheros sin huecos, tamaños, validación ihove, etc

Auditor

Este modulo verifica todos los procesos realizados por los módulos anteriores. Si determinada información tiene que estar en la base datos, lo comprueba. Si un fichero almacenado debe tener un determinado hash MD5 lo hashea de nuevo y lo verifica. Con este módulo la detección de corrupción end-to-end queda implementada



todos los procesos sin intervención del usuario

Desde este punto



>>>



Interface web



Realiza una limpieza preliminar del objeto, borrando ficheros indeseados, fijando los atributos y permisos de los ficheros y verificando caracteres extraños en los nombres de los ficheros o paths excesivamente largos.

Explorer

Extrae los metadatos técnicos y descriptivos de los objetos, incluyendo su caracterización (formato real de cada fichero). Los metadatos descriptivos pueden ser METS TXT, raw XML, XSLT Transformation o plug-in (extrayendo de una base de datos remota, un servcio web o un catálogos de APIs, etc..)

Archiver

Para reducir los riesgos (y para certificar usando TRAC), se deben realizar al menos dos copias de cada objeto gestionado. LIBSAFE realizará y gestionará esas copias e n los servidores de almacenamiento sin ninguna intervención.. Al menos dos, y tantas como el usuario desee. Ideal para copias remotas/en la nube..

Objeto u objetos preservados.



Trabajo de recuperación (Retrival)

Recuperación usando el catálogo

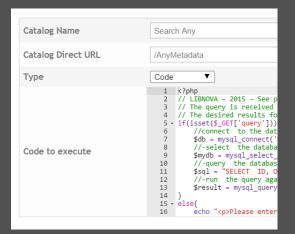
- Browse Browsing into catalog with a page that

 Simple search Making a simple search to find s

 Advanced search (Beta implementation) Using an a
 - Additional notes:
 - The object name is the name of the folder given d
 - The LUOI acronym means "libsafe universal object
 - The value of any metadata descriptor in an object
 - Each object may have different metadata sets, w should use advanced search if you want to search i

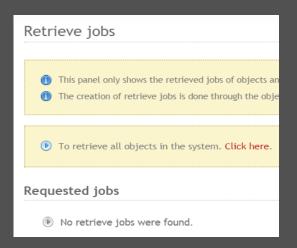
Navegue o busque en la colección completa y recupere ficheros individuales u objetos completos.

Recuperación por URL



Usando la funcionalidad de catálogo URL, es posible recuperar un objeto mediante una llamada a través de una URL El usuario puede añadir código para una necesidad particular.

Recuperación total



Es posible lanzar un trabajo de recuperación total, para extraer todos los objetos de la plataforma de forma estándar.



Recuperación usando el catálogo

Common fields										
Object name			€3	Creation date	"from"				€3	
LUOI			€3	Creation d	ate "to"				€3	
Any object metadata										
Any object metadata			€⊠							
Fixed search										
Listed only metadata schem	nas containing at least one descriptor	·.								
Metadata schema	▼	Metadata descriptor		▼	Value			С	lear	
Add condition										
Add colidition										
Free search										
1 If Descriptor name is em	npty, the condition will not be used.	You can use the "Any object meta	adata" field to search	h the value in al	l metadata of all objects.					
Descriptor name			Value					С	lear	
Add condition							6 1 1 0			
Add condition							Catalog - O	bject details		
Final conditions and searc	<u>ch</u>									its files, folders, metadata details and history.
Exact match (use case sSearch in all metadata set	_						On the "Object r	leatth" section you will fi	ind information related to real sto	rage and Active Preservation status.
Search in all metadata ser	ts of all objects.						General info	rmation		
Es nosible loca	lizar un objeto med	diante cualquiera	de sus me	tadatos	Únicamente h	av			Object Name and version	ES.GA.36057.AFP_008613 (No other vi
	obligatorio (Nombi								LUOI	1.0
	onigatorio (Normal	re dei objeto –	nombre u	ie ia cai	peta dei obje	ιο		(2)	Creation date and time	2015-08-11 17:14:14
ingestado).									Preservation area and plan	Area: Material gráfico, using: Material
							i i		Creator	User: Administrator Method: Manual b
la magunamasiá	ا ماممه مامانه م	a nácina da datall	ام مام ییم ما	المخملة					Metadata schema	EAD Amtega Sample
	n es posible desde l						Metadata (H	de)		
carpeta tempo	ral para el jtrabajo	de recuperación	y pondrá i	una copi	a del objeto (en				
ella.								Creation date	Created by	Description
								2015-08-11 17:14:14	Administrator	Initial metadata for object during preservation o
Para cada Plai	n de preservación,	es posible confi	gurar un g	enerado	or DIP XML. aı	ie		2015-08-24 15:43:27	Administrator	Updated object metadata during preservation or
	chero METS actualiza							Metadata set details		



Recuperación por URL

/catalogs/author/bell



El acceso a ficheros objeto mediante una URL es posible y fácil de configurar. Hay dos opciones disponibles:

- Defina un nombre de URL y la correspondencia con sus metadatos (/catalogs/mycatalog/title/alice); cuando se reciba esta petición los objetos serán mostrados.
- Introduzca código de gestión avanzada en PHP, permitiendo consultas avanzadas, descarga directa de ficheros, Elaboración y descarga XML o json, etc. Las peticiones recibidas serán enviadas al código de ese catálogo. Esta opción permite una integración sencilla y rápida con cualquier otro sistema o plataforma.



Recuperación total

To retrieve all objects in the system. Click here.

Si es necesario, la recuperación total de toda o parte de la colección es posible para su migración a otras herramientas.

Nuestra oferta: Procesos



Trabajo de auditoría (Audit)

Audit jobs with integrity to Some content audited is not correct	(1 jobs	with integrity faults)			
Audit job		Finished date from	Finished date to		Search
Audit job	Туре	Finished date	Handler	Status	View
AUDIT0000010		2015-09-12 14:46:03	MANUAL	(2)	E

Este módulo verifica periódicamente la coherencia de todos las fuentes de datos enviados a preservación, generando alertas cuando es necesario.

En el momento de la ingesta, LIBSAFE genera el hash MD5 para todos los ficheros ingeridos y los almacena en dos repositorios: las base de datos del sistema y ficheros XML de seguridad almacenados en el sistema de archivos vinculado al objeto que representan.

Cada N días (1 semana, 1 mes, etc.) el sistema enumera los objetos que requieren ser auditados y lanza un trabajo de auditoría sin intervención del usuario. Este trabajo regenera cada MD5 de cada copia de cada fichero y la chequea contra el correspondiente MD5 almacenado en la base de datos y el archivo de seguridad XML. El cambio, borrado o adición de cualquier dato en cualquier lugar sin seguir los métodos de ingestión, conllevará un error de integridad y generará la correspondiente alerta.



Trabajo de Evolución (Evolution)

Preserved content evolution

Technology evolves, rendering original content obsole signatures up to date, keeping digital material alive s

File format evolution

- Recaracterize file formats Files are characterize file. From here it's possible to launch re-charac
- File formats at risk report Detect potential obso
- File format evolution job File format evolution

Metadata schemas evolution

Metadata schema evolution Evaluate the impact

Digital signature evolution

- Digital signatures report View the digitally signe
- 🔯 Digital signature evolution Launch massive signa

La tecnología cambia haciendo que el contenido original quede obsoleto. El motor dinámico de preservación de LIBSAFE permite mantener actualizados los formatos de ficheros, esquemas de metadatos y firmas digitales, manteniendo vivo el material digital de forma que siga siendo utilizable a lo largo del tiempo.

Evolución del formato de ficheros

Durante la ingestión se extrae el formato real del fichero. Para algunos formatos la migración es posible. Esta funcionalidad permite evolucionar de TIF 4.0 a TIF 6.0 o PDF 1.3 a PDF/A, por ejemplo.

Evolución de esquemas de metadatos

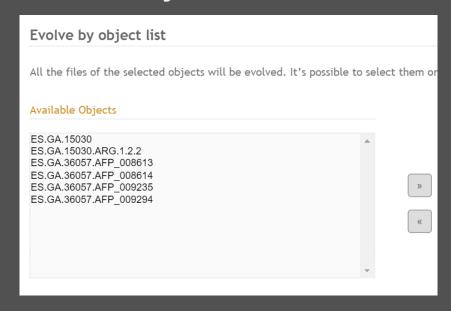
No sólo los ficheros requieren migración. Los esquemas de metadatos cambian a lo largo del tiempo. Esta funcionalidad permite la evolución esquema a esquema.

Evolución de firma digital

La firma digital de documentos legales puede caducar. Esta funcionalidad permite la refirma de documentos sin intervención del usuario.



Trabajo de evolución: Evolución de formatos de ficheros



La evolución de formato de ficheros se realiza de una forma restringida.

Sobre FUIDs (Format Unique Identifiers)

Seleccionando el formato a evolucionar, LIBSAFE enumerará los objetos afectados

Sobre Objetos

Seleccionando objetos concretos.

Sobre Áreas de Preservación

Seleccionando todos los objetos ingestados en una o más áreas de preservación.

Sin restricción, sobre la colección completa.

Nuestra oferta: Procesos



Trabajo de evolution: Evolución de formato de ficheros

JID	Name	Extension	Action	Validate	Bit Image Compar
-fmt/43	JPG File Interchange Format	jpg, jpeg	Do not evolve ▼		
-fmt/353	Tagged Image File Format	tif, tiff	Do not evolve BMP (Provided by Imager	,	
-fmt/394	DS_store	.ds_store	BMP2 (Provided by Image BMP3 (Provided by Image	emagick CONVERT)	
-fmt/18	Acrobat PDF 1.4 - Portable Document Format	pdf	JP2 (Provided by Imagem PCX (Provided by Imagen	nagick CONVERT)	
-fmt/101	Extensible Markup Language	xml	TIFF (Provided by Imagen PDF (Provided by Imagen	nagick CONVERT)	
-fmt/111	OLE2 Compound Document Format	ole2, cdf	PNG (Provided by Imager Soft delete	nagick CONVERT)	

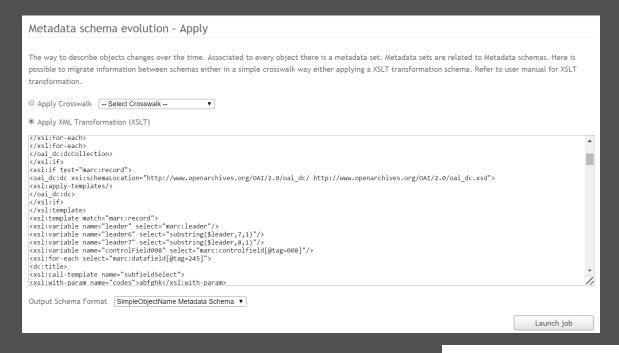
Los FUIDs seleccionados disponen de uno o más *plug-in* de conversión de salida. El cliente puede cargar nuevos *plug-in*.

La evolución llevará a un nuevo objeto del sistema, manteniendo la versión nueva y la antigua vinculadas y disponibles.

La validación y la comparación de imagen binaria está disponible también para algunos formatos de imagen.

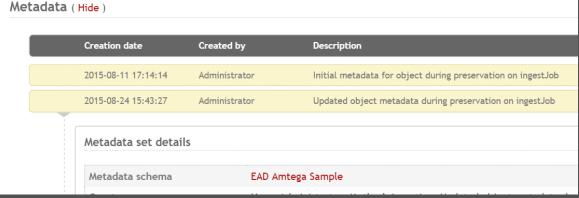


Trabajo de evolución: Evolución de metadatos



Es posible evolucionar esquemas de metadatos.

Defina el nuevo esquema, a crosswalk o XSLT entre ellos y todos los objetos relativos al antiguo esquema serán actualizados con el nuevo set de metadatos.





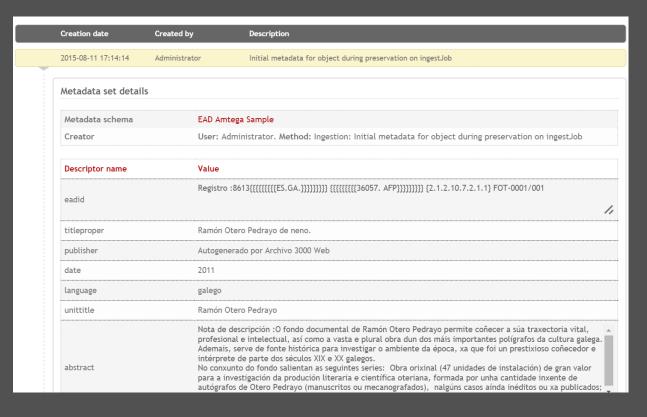
Migración de almacenamiento

Sour	ce disk					
U	INC Path	c:\Discos\DiscoA				
St	torage group name	DiscosAlpha				
L	abel	DiscoA				
D	escription	Disco A				
Si	ize	2 GB				
F	ree space	0.63 GB				
U	Ised space					
Р	riority	1				
А	llow more objects					
L	ocation	Remote				
Dest	ination disk					
	Danisa di Galda					
	Required fields					
	UNC Path					
		D: A				
	Label	DiscoA				
	Description	Disco A				

Tradicionalmente, cuando un departamento de IT quiere migrar infraestructura, deben ocuparse de la migración de datos entre dos o más servidores de almacenamiento. Esta característica de LIBSAFE permite una migración de datos más sencilla y conforme a los estándares, ya que el sistema no sólo copia la información del antiguo servidor al nuevo (y cambia los punteros a la nueva localización), sino que también verifica, fichero a fichero, utilizando hashes MD5, que la migración es completa y 100% correcta. Todo ello sin abandonar el navegador web.



Gestión de metadatos



LIBSAFE incluye un potente gestor de metadatos.

Permite gestionar distintas versiones (sets de metadatos) para cada objeto

Los metadatos son gestionados de acuerdo con el esquema de metadatos asociados con Plan de preservación del objeto.

Hay cuatro tipos distintos de descriptores:

Único: Sólo se permite un objeto con una combinación metadato/valor determinada.

. **Versionado:** Usado para detectar nuevas versiones del objeto seleccionado. **Obligatorio:** Los objetos deben tener este descriptor para ser válidos. **Normal:** Descriptores normales.



Ingestión de metadatos

Metadata (Hide)

Metadata schema EAD Amtega Sample

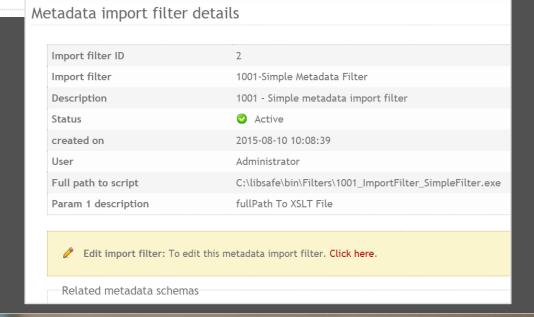
Metadata import filter 1001-Simple Metadata Filter

Filter parameters of 1001-Simple Metadata Filter: 1001 - Simple metadata import filter

Metadata file pattern *.xml

fullPath To XSLT File c:\libsafe\conf\Filters\filtroAmtegaTest.xslt

La carga de de información de metadatos de un objeto se realiza durante la ingestión utilizando los filtros. Varios filtros de importación están incluidos, pero el usuario puede incluir más, incluyendo la posibilidad de realizar consultas a otros sistemas, confeccionar el metadato y asociarlo con el objeto.





Esquemas de metadatos

Metadata schema details

Metadata schema ID	4
Metadata schema	Marc21 Metadata Schema EN
Created by	Administrator
created on	2015-08-10 10:08:41
Status	Active

Edit metadata schema: To edit this metadata schema. Click here.

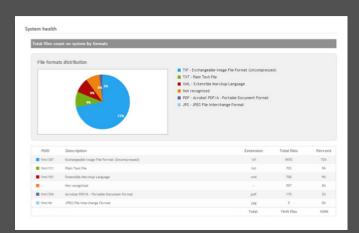
To create a new metadata schema based on the current. Click here.

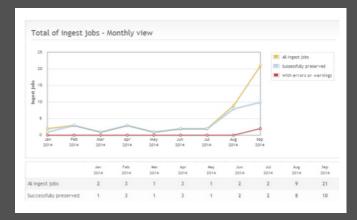
Related metadata descriptors

(2399metadata descriptors)

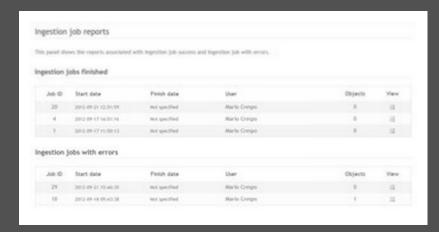
Name	Order	IE Code	Mandatory	Unique	versioning	View
001: CONTROL NUMBER (NR)	2	0001:001:#:#				∷ ≡
003: CONTROL NUMBER IDENTIFIER (NR)	2	0001:003:#:#				∷ ≡
005: DATE AND TIME OF LATEST TRANSACTION (NR)	2	0001:005:#:#				ŧΞ







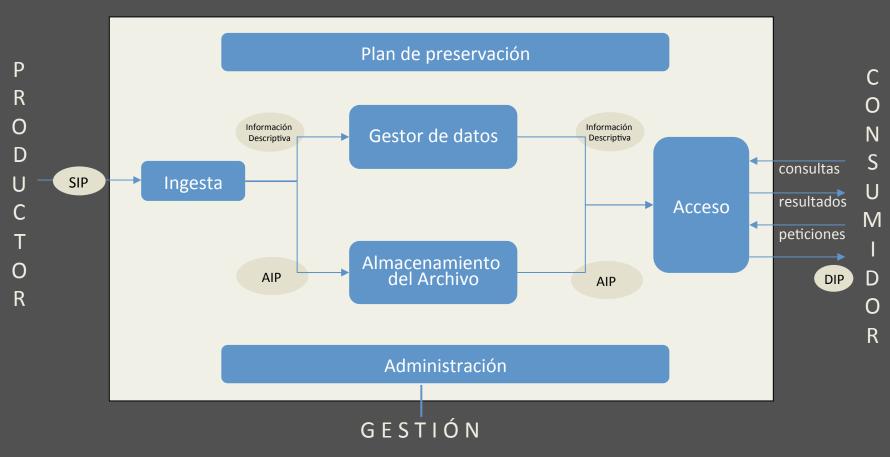
Informes



Varios informes ya están disponibles. Más disponibles en breve



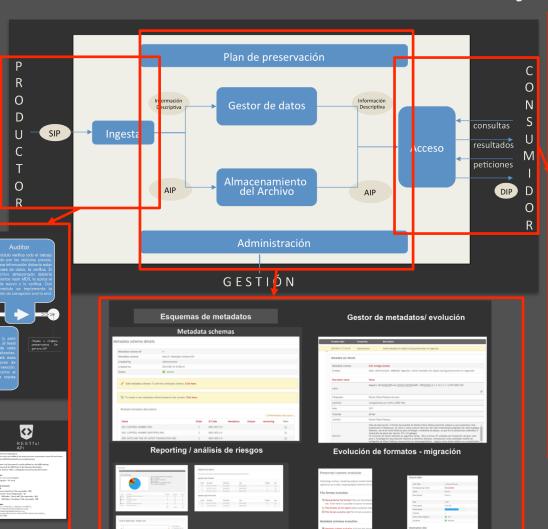
Conceptos OAIS

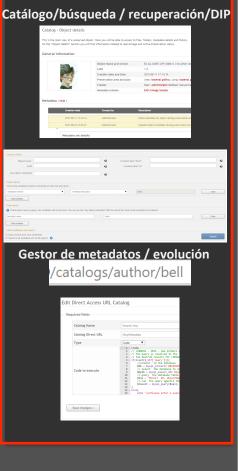




Solución FULL OAIS + Aditivos de eficiencia y seguridad

LIBSAFE es una solución integrada de Archivo Full OAIS, con extensiones que hacen la gestión del archivo más eficiente, sencilla y segura.







Solución FULL OAIS + Aditivos de eficiencia y seguridad

Nucleo Full OAIS +

Almacenaje /
Gestor de
infraestructura

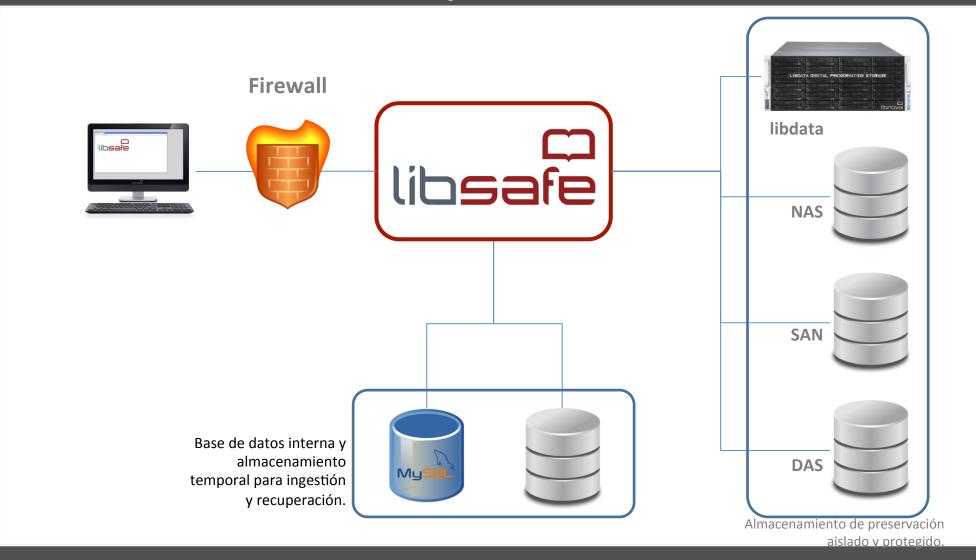
Validación y transformación de objetos

Motor de firma digital

Ingestión y transformaciónd e metadatos multiesquema



Arquitectura



Nuestra oferta: Licencias



Licencia

Coste de licencia inicial (5Tb incluidos) + Upgrades de licencia para ingestión adicional

LIBSAFE se licencia en base a la información total ingestada. La licencia inicial permite ingerir hasta 5Tb. Los packs para ingestión adicional son asequibles y son necesarios para continuar ingiriendo pero no para continuar usando el software.



- Producto basado en estándares, pero también en proyectos reales.
- Implantación rápida. Solución llave en mano.
- Reduce los costes de preservación o permite la preservación a bajo coste.
- Solución all-in-one. Completa.
- Tranquilidad.

LIBNOVA EMEA & LATAM

Paseo de la Castellana, 153 Madrid – España Tel: +34 91 449 08 94 Fax: +34 91 141 21 21 contacto@libnova.com libnova USA 14 NE First Ave (2nd Floor) Miami, Florida 33132, USA Tel: +1 844-894-6532 jcromero@libnova.com contact@libnova.com

