

libsafe – Preservación digital

XIX Jornadas de Archivos Universitarios, marzo de 2013

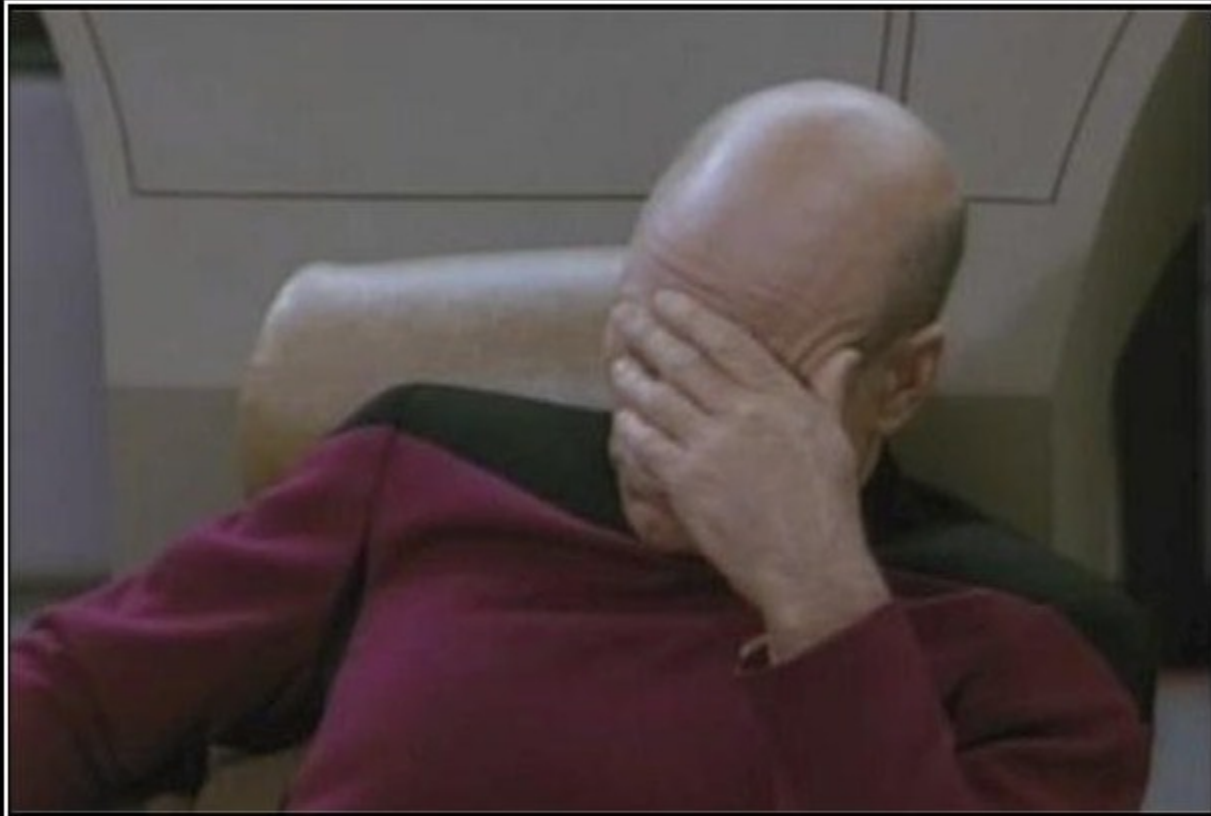
Paseo de la Castellana, 153
28046 – Madrid

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



Antes de nada...

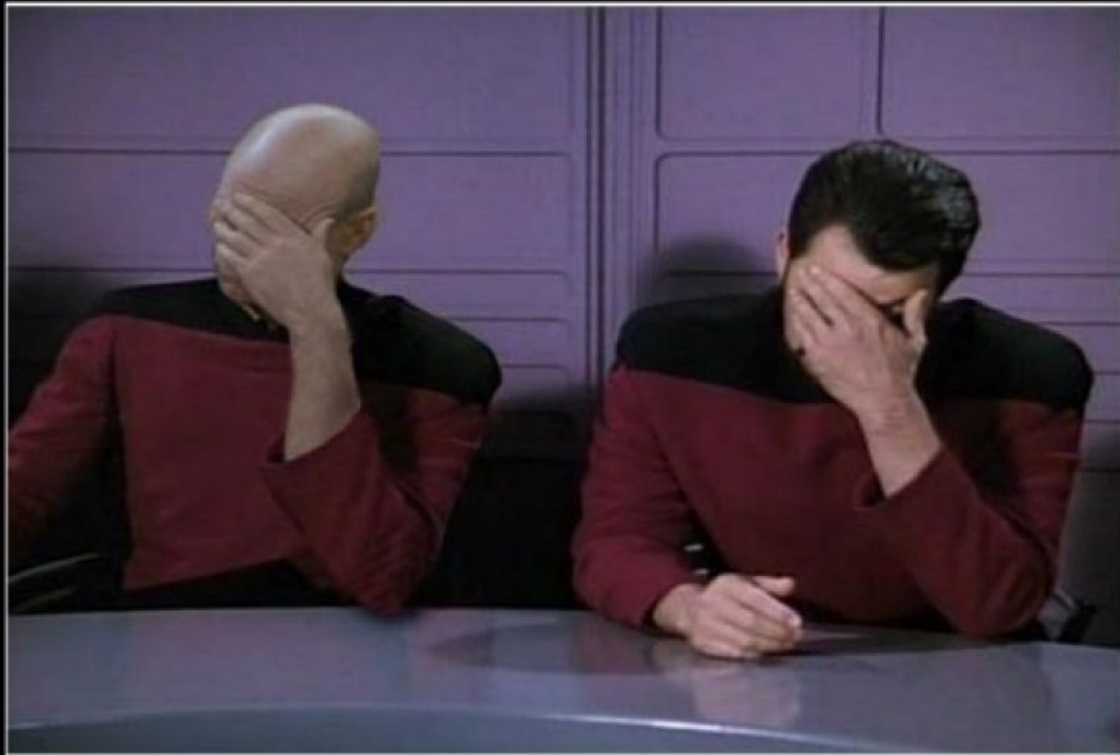
... corre por internet



F A C E P A L M

Because expressing how dumb that was in words just doesn't work.

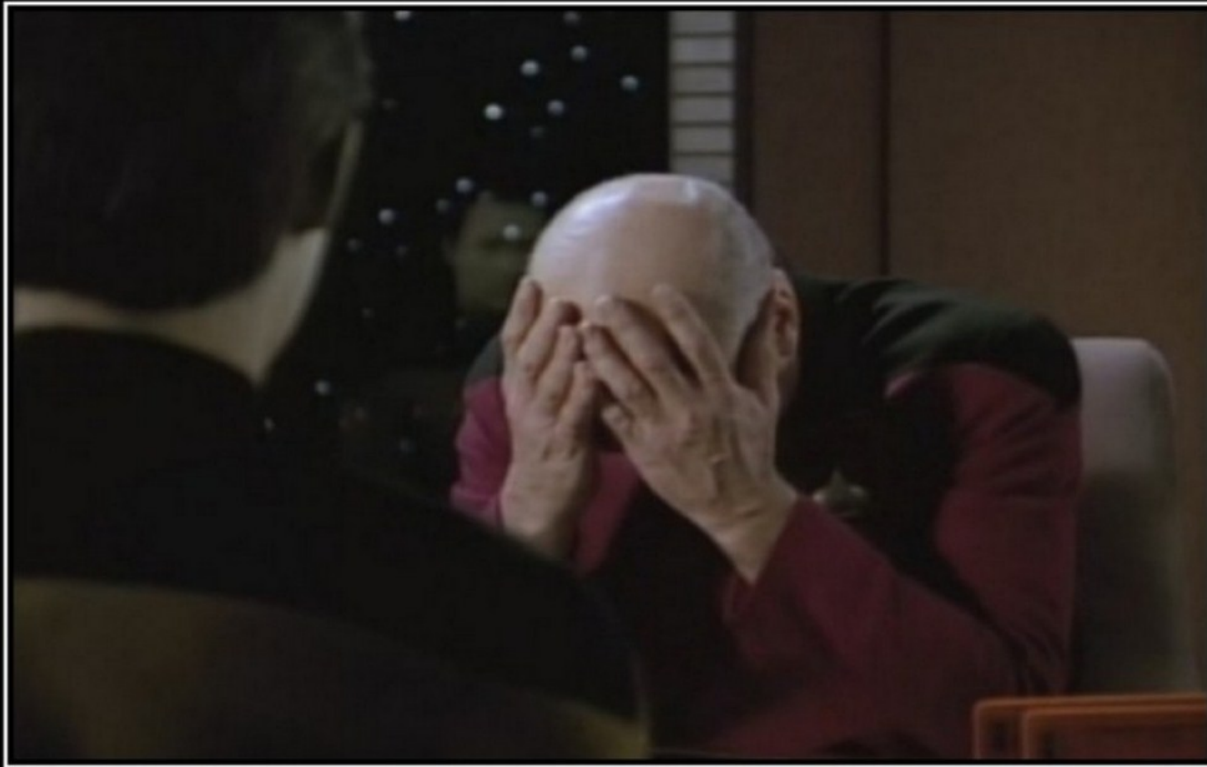
gesto universal



DOUBLE FACEPALM

When the Fail is so strong, one Facepalm is not enough.

dobles



FACEPALM

When words fail to describe the dismay, there is always Facepalm.

de mayor gravedad



en el reino animal



desde hace muchos años



e incluso presidenciales.



- > Antonio Guillermo Martínez – Director de desarrollo en libnova
a.guillermo@libnova.es
- > libnova nace en el año 2009, adquiriendo tecnología de una empresa anterior.
- > Desarrollamos tecnología, pero conocemos bien la realidad de las bibliotecas y archivos.
 - > Especialistas en desarrollo y nuevas tecnologías
 - > Especialistas en bibliotecas y archivos
- > Ejecutamos junto con Vinfra, Telefónica y otras empresas el proyecto de digitalización masiva de la Biblioteca Nacional.

- > Orientados a día de hoy a la **preservación digital**, con varias patentes en este ámbito.
- > Desarrollamos tecnología orientada al mercado de la digitalización y organización de bibliotecas y archivos.



Certificados por SGS España y Reino Unido en ISO9.001, ISO14.001 e ISO 27.001

ISO 9.001 – Gestión y mejora de la calidad

ISO 14.001 – Medio ambiente y tecnologías verdes

ISO 27.001 – Seguridad de la información



ISO 9001



ISO 14001



ISO 27001

Y mantenemos partnertships y acuerdos con algunos de los principales fabricantes de IT



La preservación digital

¿Qué es?

El conjunto de actividades a realizar sobre la información digital que asegure la posibilidad de su utilización futura en, al menos, las mismas condiciones que en las que se creó.

Parece razonable. Pero entonces

¿Por qué no tiene una adopción masiva?

Desconocido

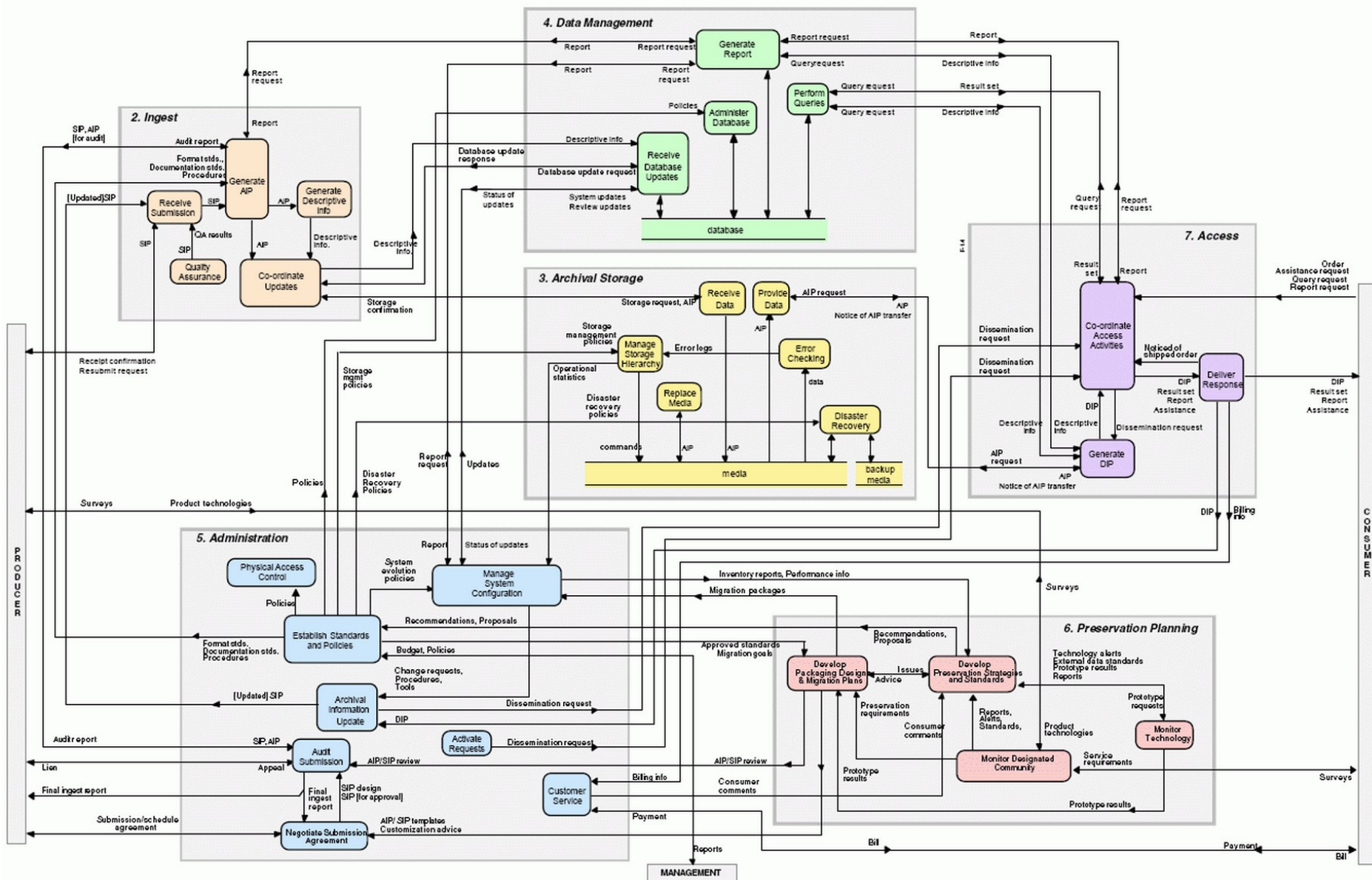
hasta ahora digitalizar o almacenar = preservar
no existía el marco tecnológico ni operativo

+

Complejidad

+

Coste



Desconocido

hasta ahora digitalizar o almacenar = preservar
no existía el marco tecnológico ni operativo

+

Complejidad

El modelo OAIS “parece” sencillo

+

Coste

Software a medida, incierto, gasto adicional
Al alcance sólo de grandes instituciones

~~Desconocido~~

hasta ahora digitalizar o almacenar = preservar
no existía el marco tecnológico ni operativo

+

~~Complejidad~~

El modelo OAIS “parece” sencillo

+

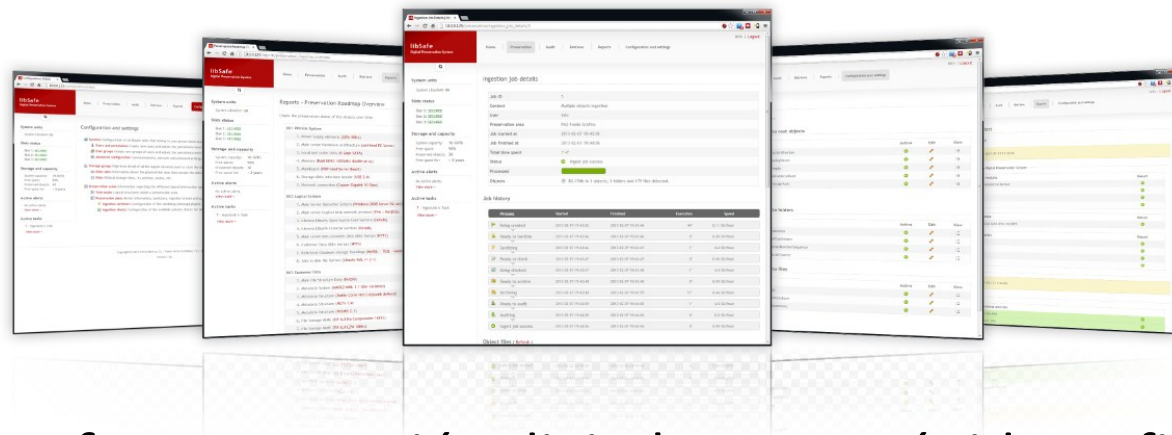
~~Coste~~

Software a medida, incierto, gasto adicional
Al alcance sólo de grandes instituciones

*Fácil de entender. Ayuda
para ponerlo en marcha.*

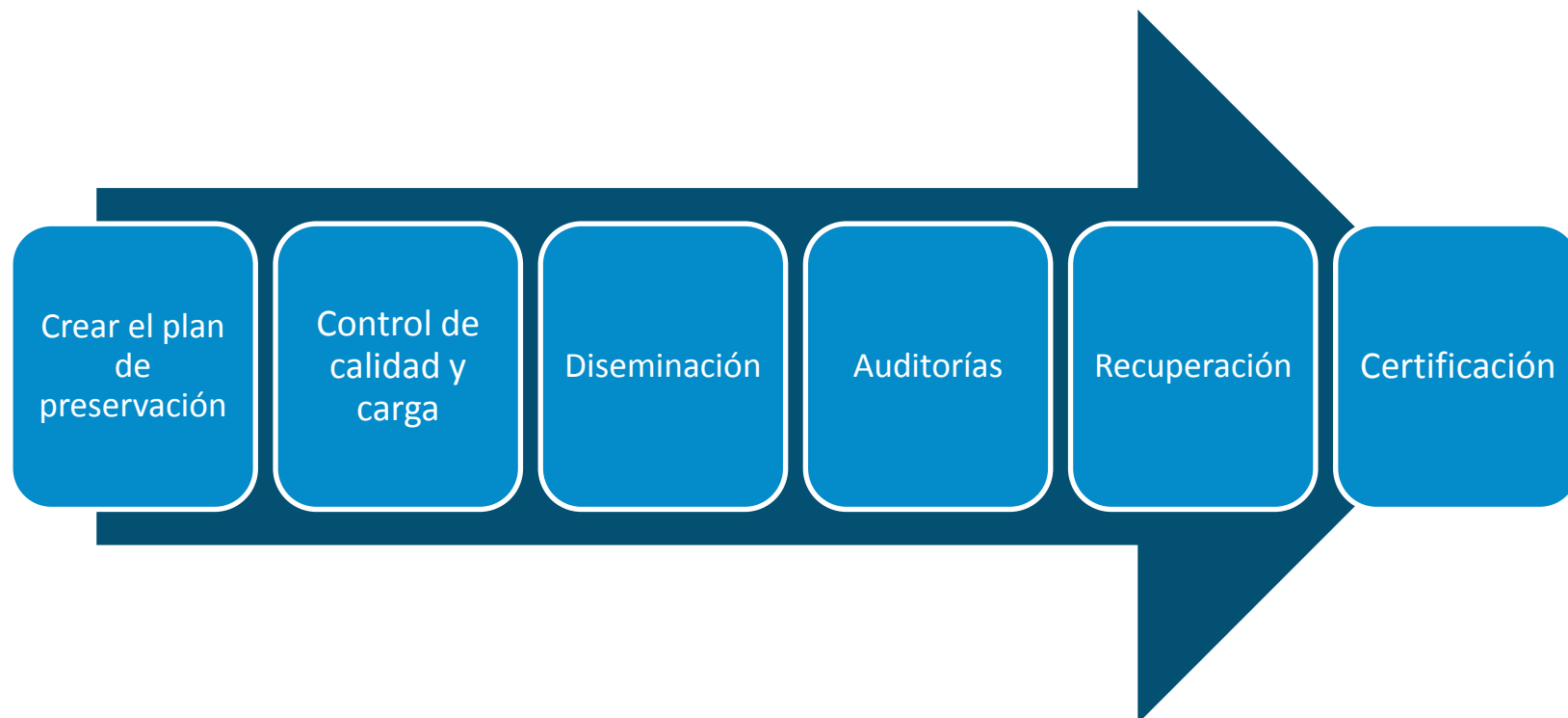
*Fácil de usar. Probado bajo
“fuego real”. 100% basado
en web.*

*Licencia de software, coste
conocido y razonable.*



libsafe: Preservación digital segura, rápida y eficaz

Puntos clave del proceso de preservación







Creación del plan de preservación

- > Sistema centralizado de gestión documental de fácil uso. Permite crear y mantener todo lo relativo al plan de preservación en un único lugar.
 - > Control y autorización de cambios
 - > Control de usuarios
 - > Muy fácil de usar, edición sobre el propio contenido
- > Incluye una guía de todos los puntos a los que se debe prestar atención en el plan, así como consejos para su elaboración.
- > Nuestra experiencia en preservación también está a su disposición.



Control de calidad y carga

Motor de auditoría de calidad

(Definido en el plan de preservación)

> Comprobaciones básicas

- > Nombre de ficheros
- > Estructura de carpetas
- > Metadatos
- > Validación con jhove
- > Comprobación de formatos
(TIF 6.0 Sin Compresión, PDF/A, etc)
- > Comprobación antivirus
- > Huella y firma digital

The screenshot displays the libSafe Digital Preservation System interface. The left sidebar contains navigation links: Home, Preservation, Audit, Retrieve, Reports, and Configuration and settings. The main content area is titled 'Ingestion job details' and shows information for a specific job (Job ID: 5). Below this, a 'Job history' table lists various processes and their execution details.

Process	Started	Finished	Execution	Speed
Being created	2013-02-07 19:43:02	2013-02-07 19:43:06	44"	8.11 Gb/hour
Ready to Sanitize	2013-02-07 19:43:46	2013-02-07 19:43:46	0"	0.00 Gb/hour
Sanitizing	2013-02-07 19:43:46	2013-02-07 19:43:47	1"	4.8 Gb/hour
Ready to check	2013-02-07 19:43:47	2013-02-07 19:43:47	0"	0.00 Gb/hour
Being checked	2013-02-07 19:43:47	2013-02-07 19:43:48	1"	4.8 Gb/hour
Ready to archive	2013-02-07 19:43:48	2013-02-07 19:43:48	0"	0.00 Gb/hour
Archiving	2013-02-07 19:43:48	2013-02-07 19:43:59	11"	0.44 Gb/hour
Ready to audit	2013-02-07 19:43:59	2013-02-07 19:44:00	1"	4.8 Gb/hour
Auditing	2013-02-07 19:44:00	2013-02-07 19:44:06	6"	0.8 Gb/hour
Ingest job success	2013-02-07 19:44:06	2013-02-07 19:44:06	0"	0.00 Gb/hour

> Extensible por el usuario, integrable con herramientas Open Source, scripts, etc.

Control de calidad y carga

Motor de auditoría de calidad

(Definido en el plan de preservación)

> Comprobaciones avanzadas

- > Páginas en blanco
- > Archivos vacíos o demasiado pequeños
- > Páginas borrosas
- > Archivos dañados



Carga en el sistema

Proceso de ingestión

(Si pasa el control de calidad)

libSafe
Digital Preservation System

Home | Preservation | Audit | Review | Reports | Configuration and settings

Multiple Objects preservation process (phase 2/3)

Create new ingestion job | Copy object files to the libsafe unit | Ingestion job review and start

You have created the new ingestion process & now you should copy the files associated to the objects to this system using one of the following methods:

- Network Share:** Open a file browser window and write `libSafe@192.168.1.100` in the address bar to open the shared folder location of the Ingest Job where you can put the files. The shared folder is available to the user who created the job.

When you have copied all the objects files to the network location, please click next.

Object Files (Refresh)

30,000 in 1 objects, 5 folders and 170 files detected so far.

30,000 in 1 objects, 5 folders and 170 files detected so far.

30,000 in 1 objects, 5 folders and 170 files detected so far.

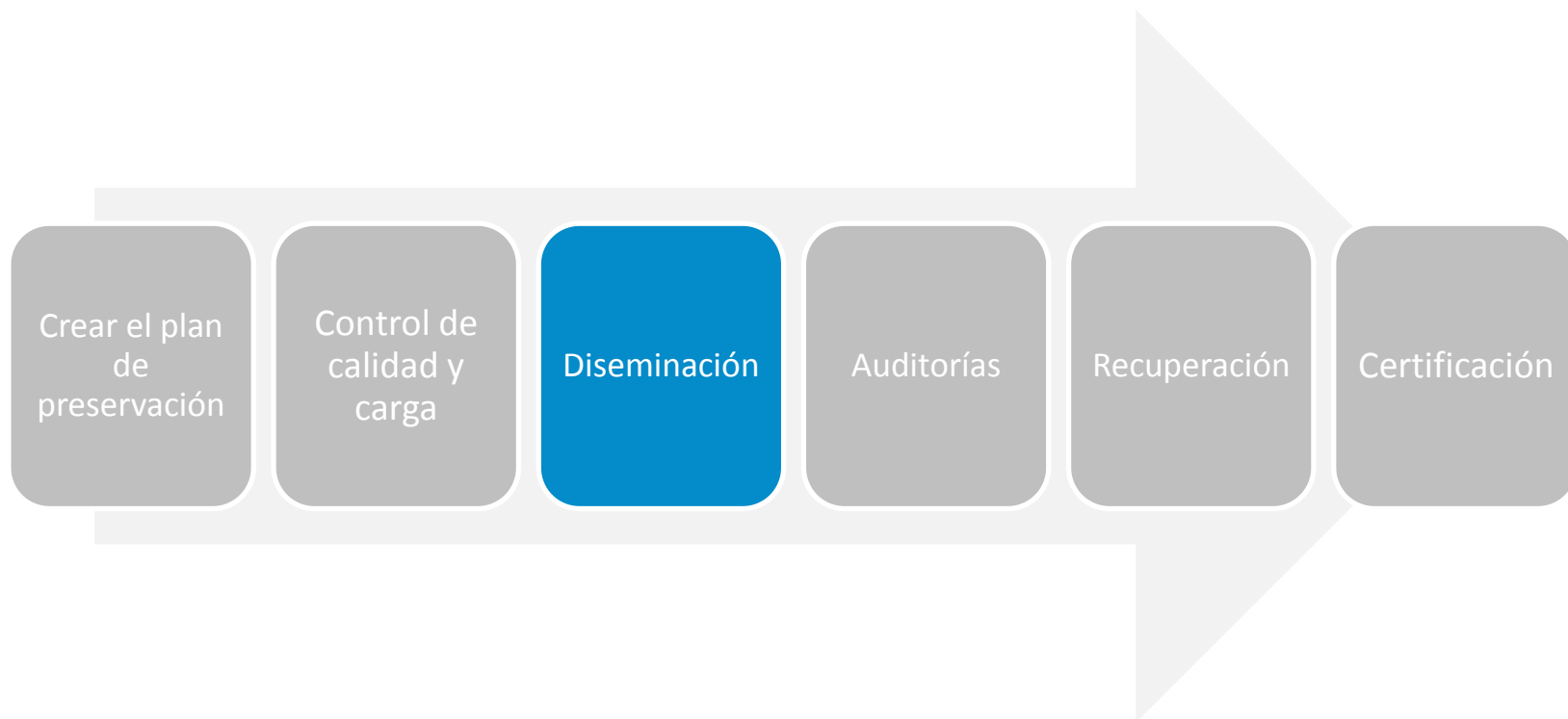
30,000 in 1 objects, 5 folders and 170 files detected so far.

Multiple Objects preservation process (phase 2/3)

Create new ingestion job

Copy object files to the libsafe unit

Ingestion job review and start





Diseminación

libsafe realiza copias a los distintos dispositivos de almacenamiento

- > Cuantas más copias (controladas) de la información, menor es el riesgo de pérdida.
- > libsafe copia de forma automática a servidores, NAS o equipos libdata
- > Posteriormente realiza una verificación de cada copia.

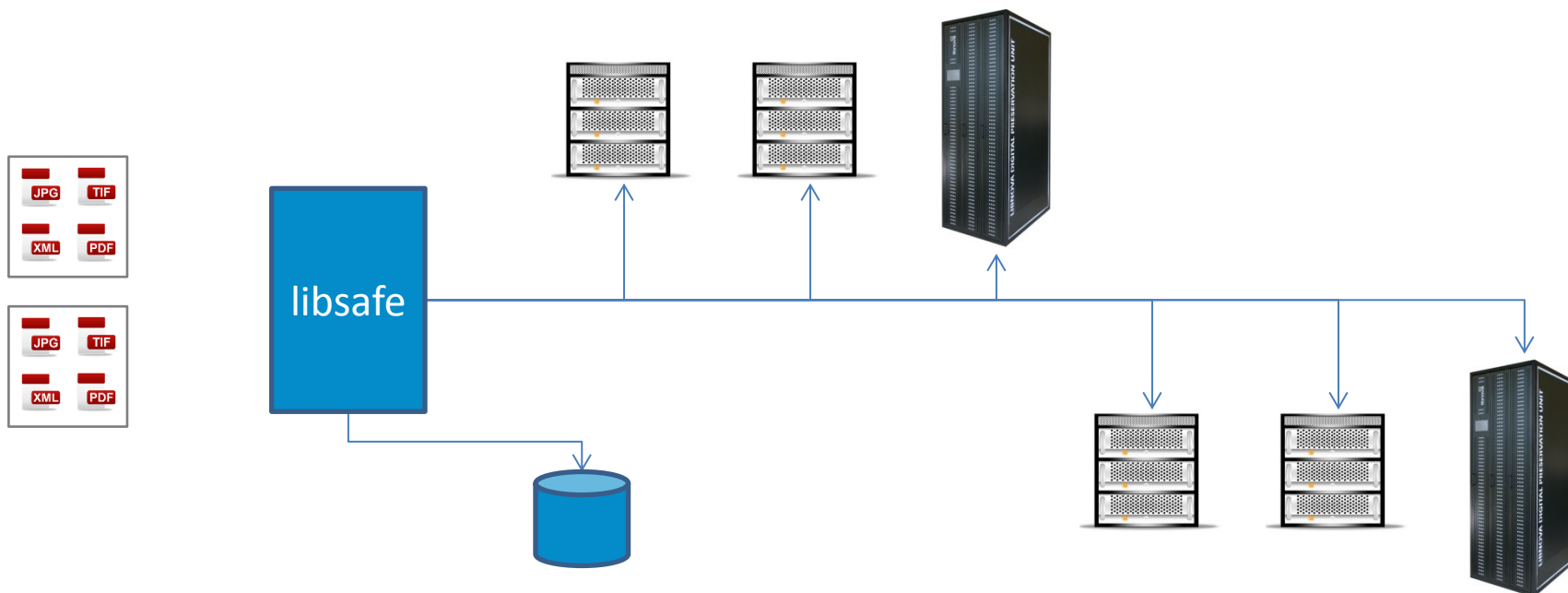


Diseminación

libsafe realiza copias a los distintos dispositivos de almacenamiento

Diseminación

- Después de la comprobación inicial, libsafe carga localmente los objetos digitales (conjuntos de archivos, estructuras y metadatos)
- Carga e indexa los metadatos y genera un HASH MD5 de cada fichero y lo almacena en la BBDD
- En función de lo indicado en el plan de preservación, realiza copias idénticas de los archivos en los distintos dispositivos de almacenamiento.
- Después de la copia, audita y verifica que ésta se ha hecho bien.



Diseminación

Unidades de la familia libdata – Almacenamiento masivo seguro



1,6 Pb
1.600Tb
En cada armario

(el equivalente
a 1.600 discos
duros externos
de 1Tb cada
uno)

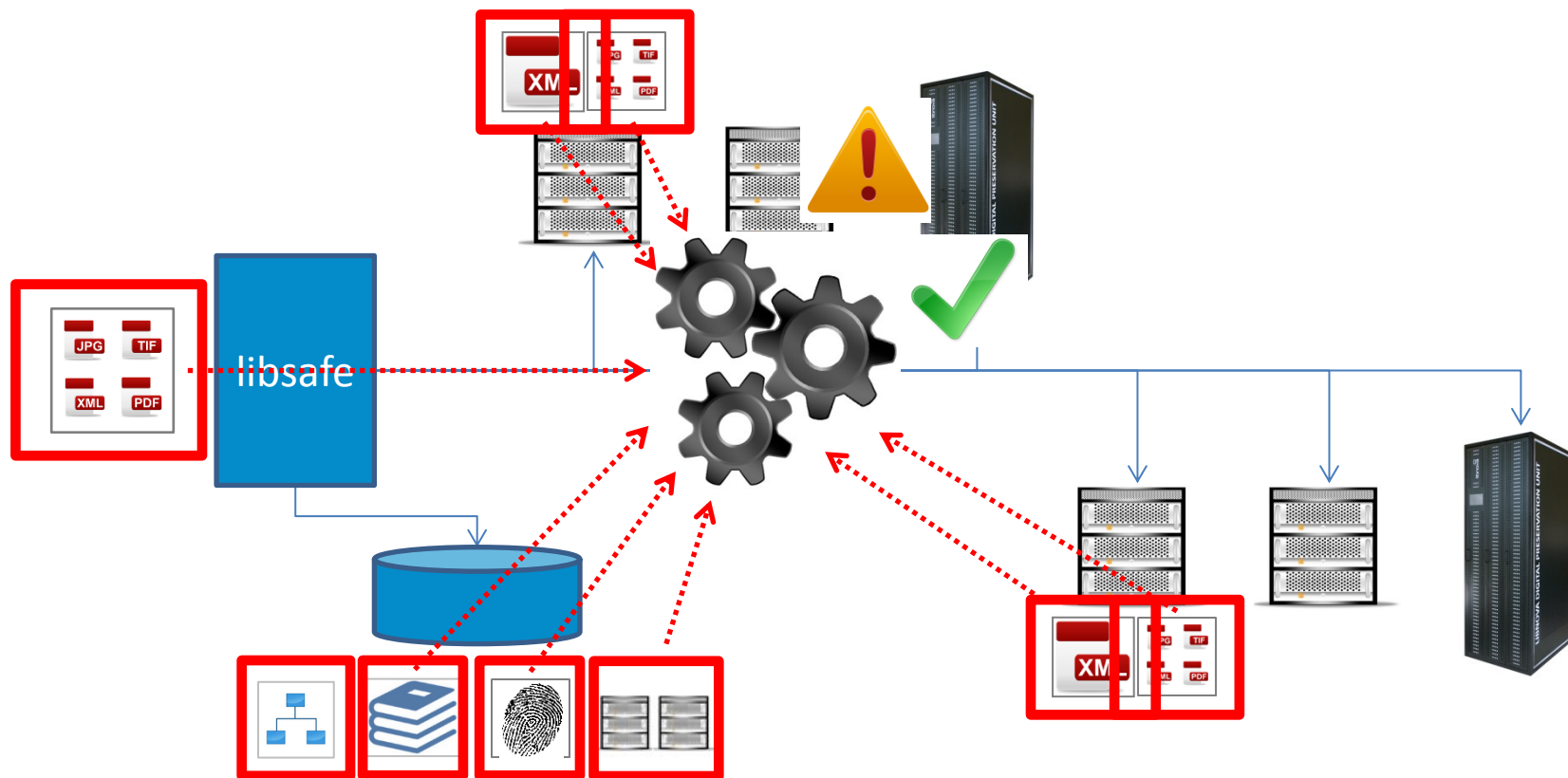




Auditoría

En el proceso de ingestión guardamos información sobre:

- > Cómo era el objeto cuando se cargó en libsafe
- > Qué estructura de ficheros y carpetas tiene
- > Sus metadatos
- > Las huellas dactilares de cada fichero (hash MD5)
- > Los dispositivos de almacenamiento donde debe estar guardado
- > Los ficheros del objeto, en sus distintos dispositivos
- > Y un XML adicional de seguridad que genera libsafe junto a cada objeto





Recuperación

Acceso a los objetos preservados – Por catálogo

libSafe
Digital Preservation System

Home | Preservation | Audit | **Retrieve** | Reports | Configuration and settings

info | Logout

System units
System Libsafe01 OK

Slots status

Storage and capacity
System capacity: 10.92Tb
Free space: 98%
Preserved objects: 20
Free space for: > 2 years

Active alerts
No active alerts.
View more »

Active tasks
IngestJob 7 Task
IngestJob 6 Task
View more »

Retrieve jobs
Create new retrieve job »

Requested jobs
No retrieve jobs were found.

Retrieved jobs
Created on: [input] Object or file name: [input] Select status: [dropdown] Search

Created on	Content	Object / File name	Clear	Status	View
2013-02-13 11:12:42	musica	U_000578_V01_1000360477	[icon]	[icon]	[icon]
2013-02-11 19:38:17		U_000578_V01_1000360477	[icon]	[icon]	[icon]
2013-02-07 19:52:35		U_000578_V01_1000360477	[icon]	[icon]	[icon]

Status reference

[icon] Data Requested	[icon] Waiting	[icon] Retrieving
[icon] Data Retrieved	[icon] Retrieve Error	[icon] Clear Requested
[icon] Cleared		

Copyright © 2010-2013 Libnova, SL / Paseo de la Castellana, 153 - Madrid
Version: 1.02



Recuperación

Acceso a los objetos preservados – Por enlaces

1. <http://libsafeserver/catalog01/IDITEM/1233>

2. <http://libsafeserver/catalog01/IDITEM/1233/files/FILEID/5>

3. <http://libsafeserver/catalog01/IDITEM/1233/files/FILENAME/fichero1.tif>

4. <http://libsafeserver/catalog01/IDITEM/1233/files/FILENAME/fichero1.tif&VERB=DIRECTDOWNLOAD>

5. <http://libsafeserver/catalog01/IDITEM/1233&VERB=SHOWXML>



Certificación

Cuatro ojos ven más que dos







Antonio Guillermo Martínez Largo
a.guillermo@libnova.es

www.libnova.es
Teléfono +34 91 449 08 94

info@libnova.es

Paseo de la Castellana, 153
28046 – Madrid

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

**libnova**