



Grupo de
Preservación
Digital



APLICACIÓN DE ESTRATEGIAS DE PRESERVACIÓN DIGITAL EN LA BIBLIOTECA Y HEMEROTECA NACIONALES DIGITALES DE MÉXICO

Agosto 19 del 2021

Mtra. Ana Yuri Ramírez Molina

Coordinación de Innovación y Estrategia Digital



Temas

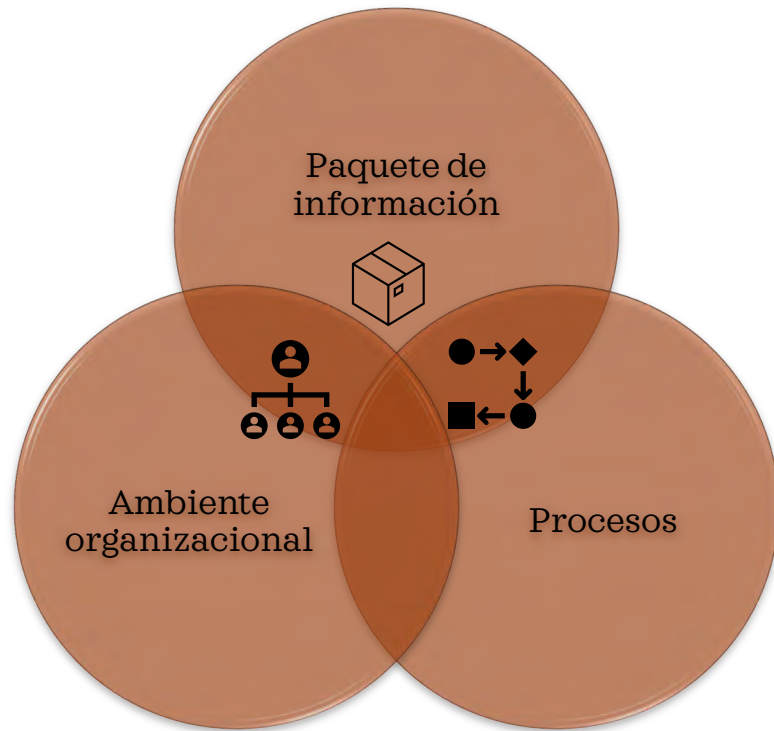
- Bases teóricas
- Ambiente organizacional
- Procesos
- Paquete de información
- Herramientas



BASES TEÓRICAS

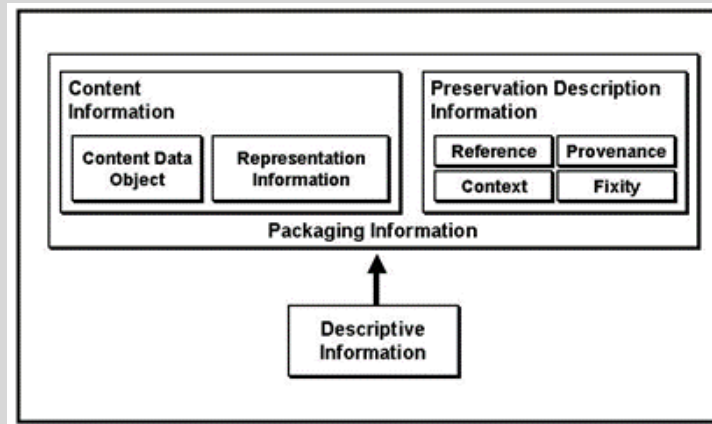
Modelo OAIS

Agosto 19 del 2021
Mtra. Ana Yuri Ramírez Molina
Coordinación de Innovación y Estrategía Digital

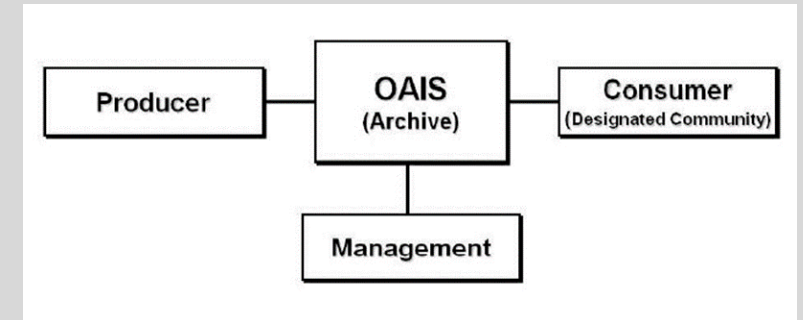


Bases teóricas

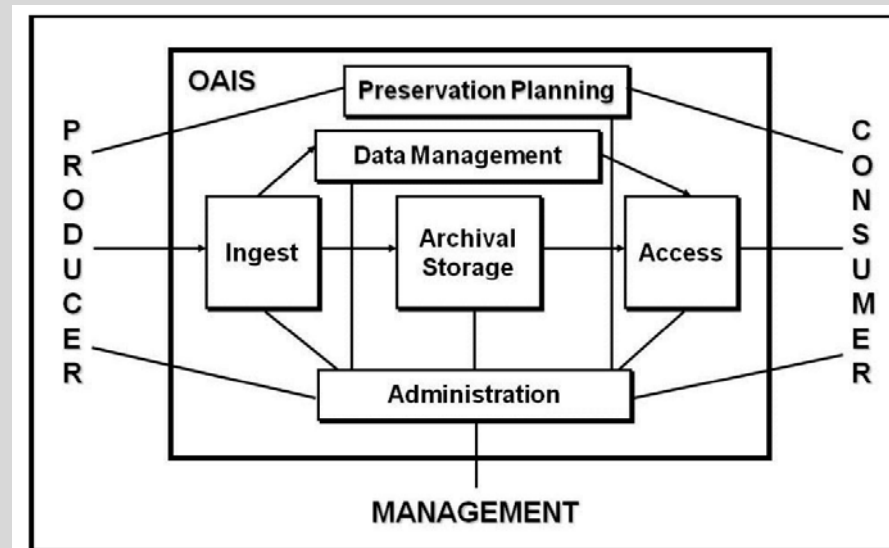
Modelo OAIS



Paquete de Información



Ambiente



Modelo funcional



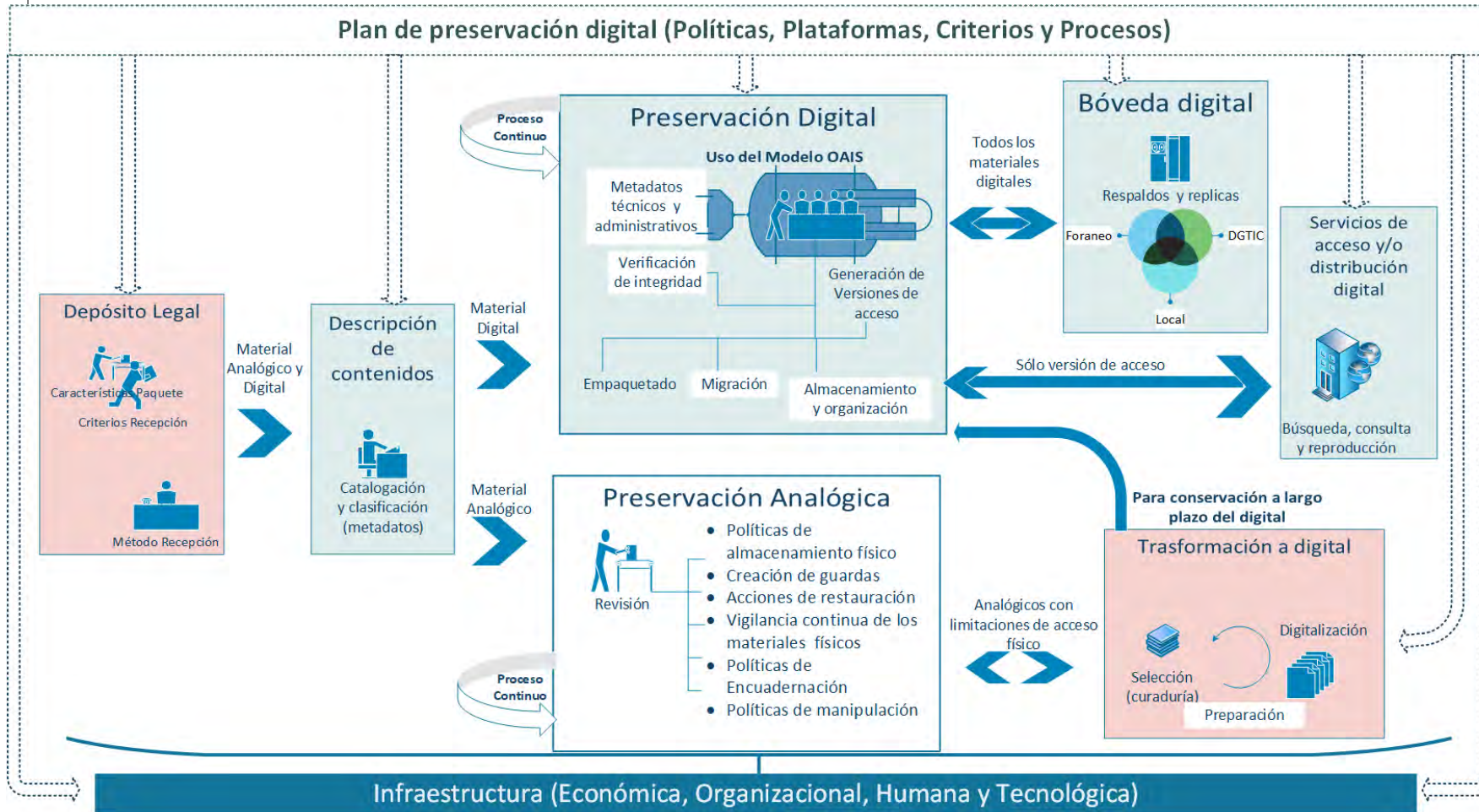
AMBIENTE ORGANIZACIONAL

Necesario para preservación digital

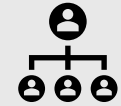
Agosto 19 del 2021
Mtra. Ana Yuri Ramírez Molina
Coordinación de Innovación y Estrategia Digital



Identificación elementos



Ambiente organizacional



¿Cómo se inserta en la organización?

¿Quiénes son los productores?

¿Quiénes conforman la comunidad designada?

¿Qué se requiere para la gestión de la preservación digital?

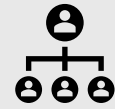
¿Qué se requiere para la gestión de la preservación digital?

IDENTIFICAR



TRANSFORMAR

Ambiente organizacional





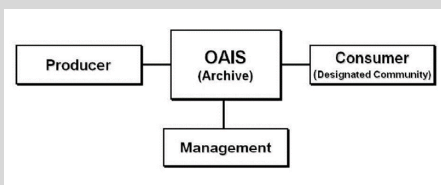
PROCESOS

Identificación

Agosto 19 del 2021
Mtra. Ana Yuri Ramírez Molina
Coordinación de Innovación y Estrategia Digital



Elementos identificados de acuerdo con el ambiente



- Depósito Legal
- Digitalización

Productores



- Paquetes de información
- Herramientas (infraestructura)
- Políticas, normas y estrategias

Control y gestión



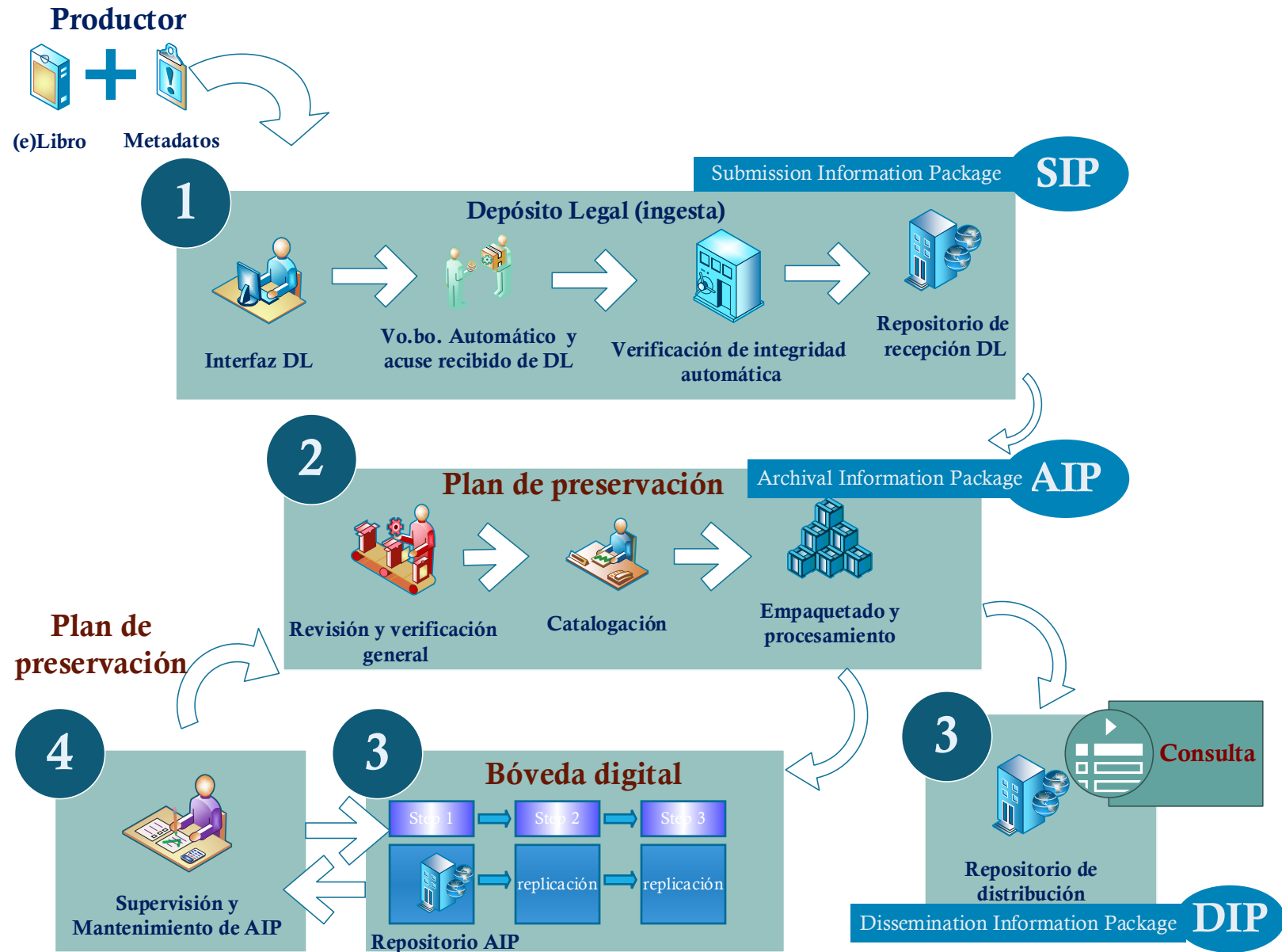
- Gestores
- Usuarios

Consumidores

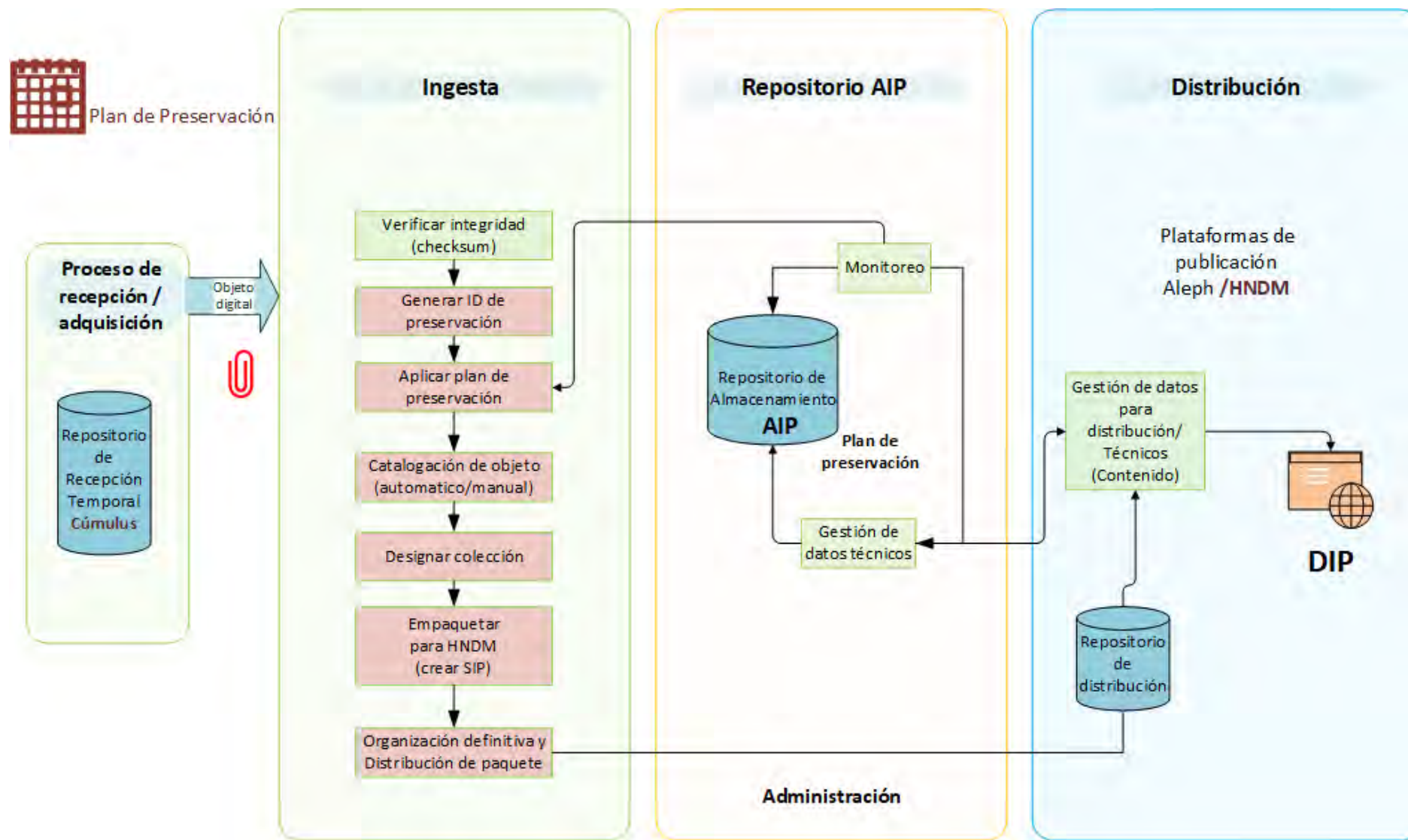


Depósito Legal

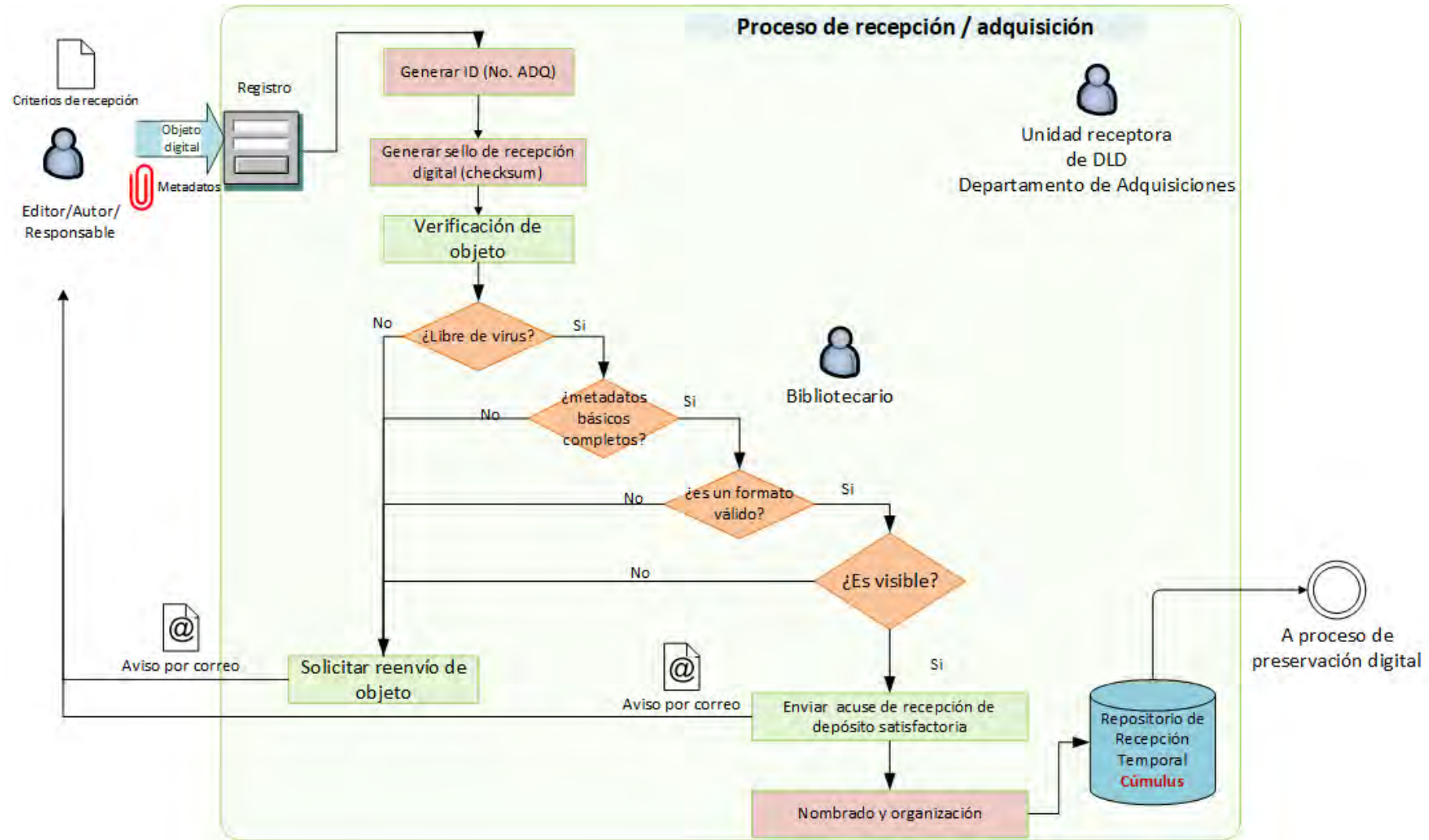
- Definición de componentes
- Identificación de etapas
- Identificación de entidades



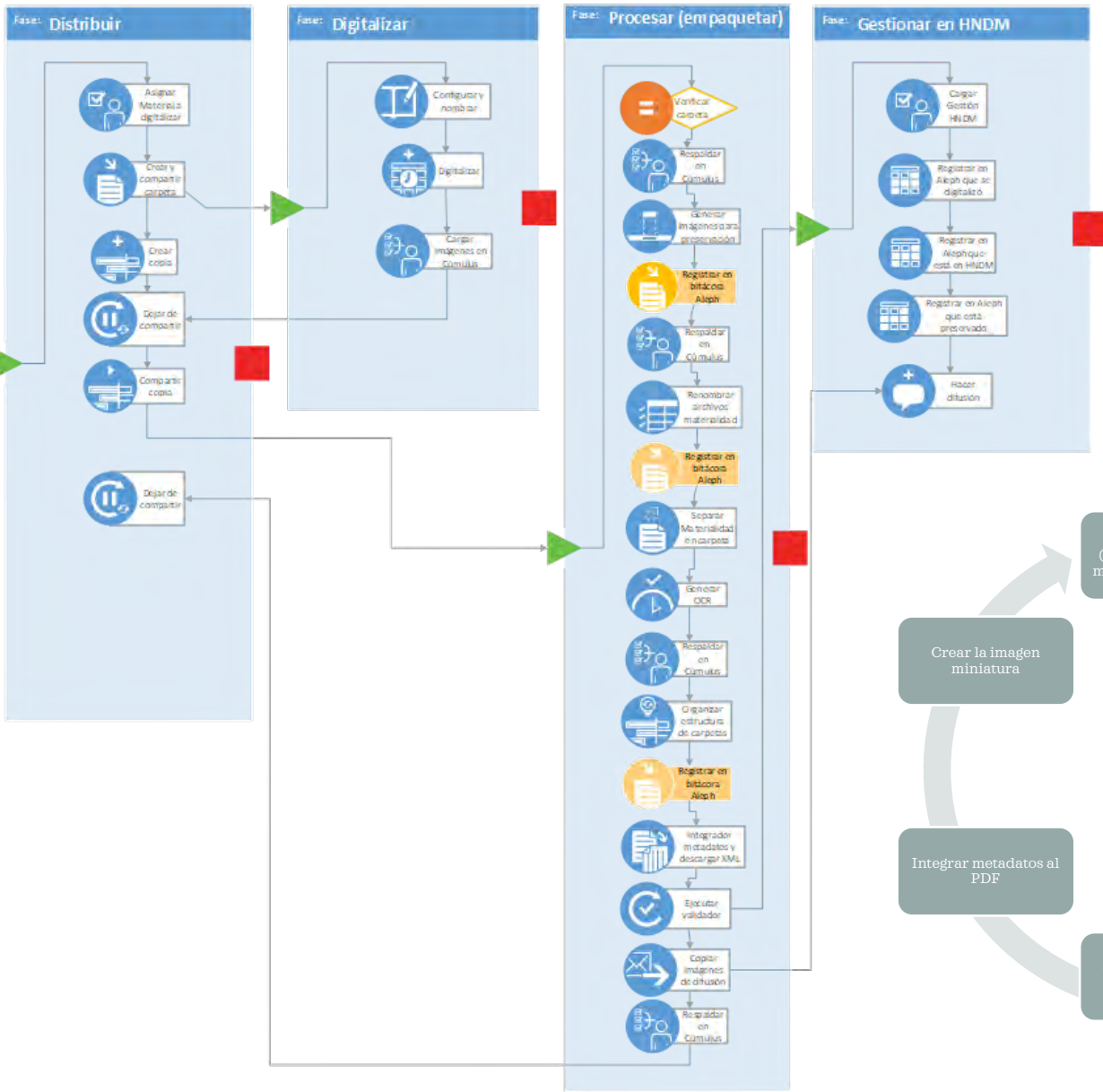
Arquitectura para Depósito Legal



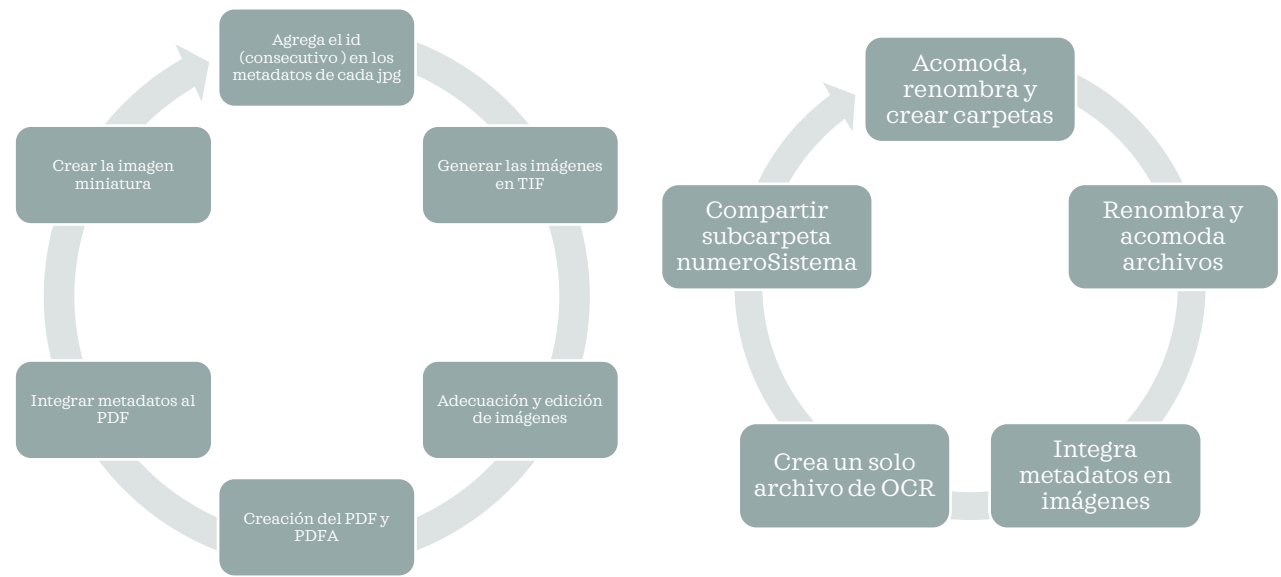
Recepción y verificación del Depósito Legal



Digitalización



- ❑ Metadatos embebidos
 - ❑ Descriptivos
 - ❑ Administrativos
 - ❑ Preservación
- ❑ Versiones de preservación y distribución.
- ❑ Conexión con Aleph y HNDDM

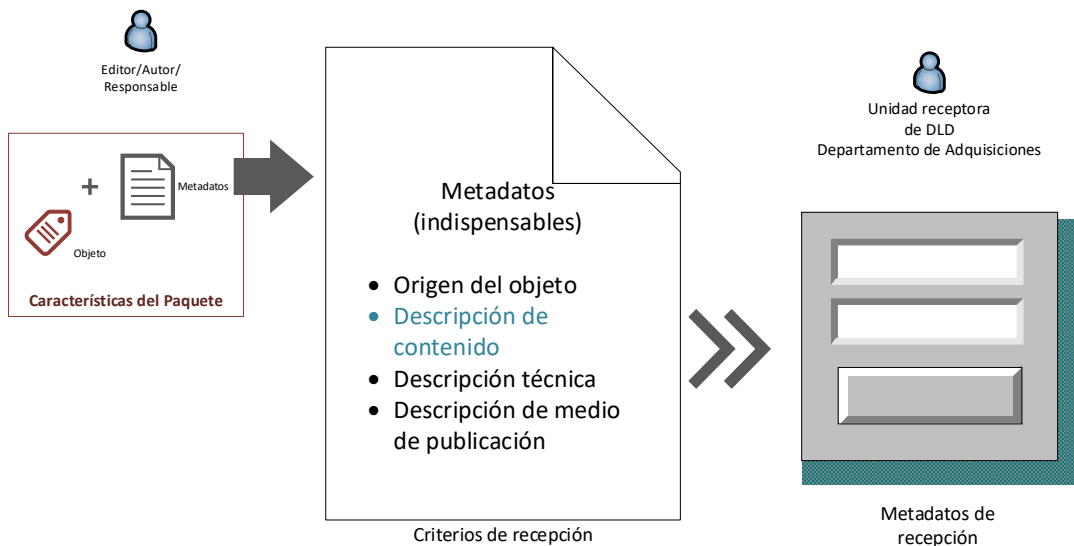




PAQUETE DE INFORMACIÓN

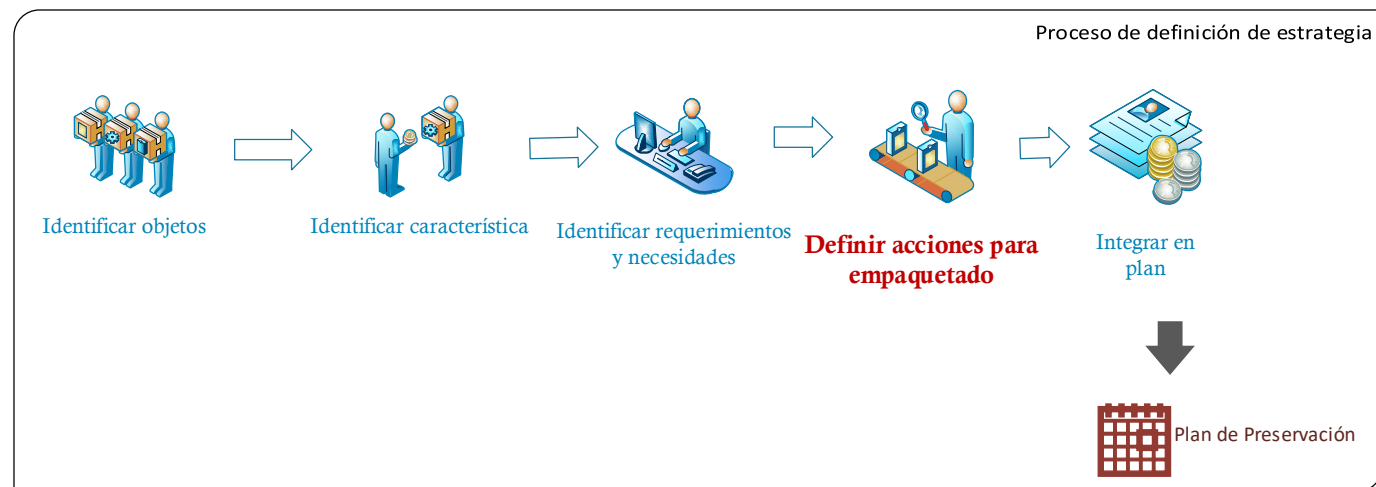
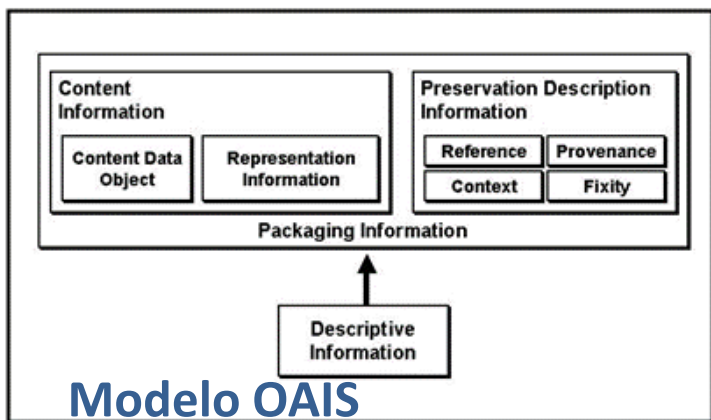
Para preservación

Proceso de recepción / adquisición

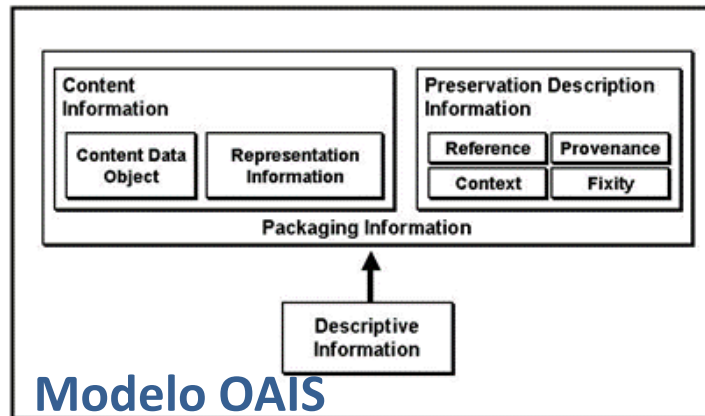


Paquetes de Información Depósito Legal

Estrategias técnicas de preservación por tipo de objeto (contenidas en el plan de preservación)



Paquete de Información (digitalización)



Representación

- Origen
- Proceso de captura

Contenido

- Estructura
- Metadatos

Comunidad designada

- Necesidades
- Alcance y expectativas

Representación - origen



Catalogación completa en Aleph



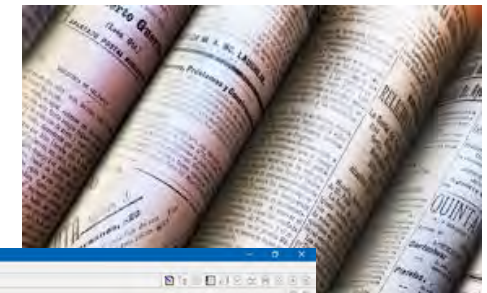
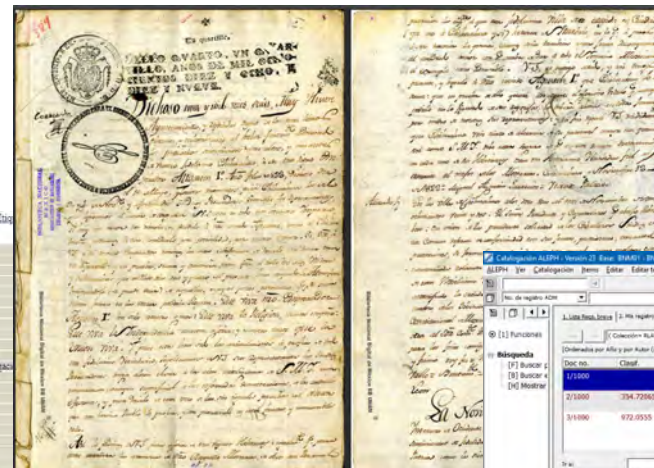
Registro de existencias (códigos de barra)



Materialidad incluida

Seleccione formato: Formato entidad: Tarjeta catalográfica Cita Numeral de epígrafe: Epi
Registro 26 de 5518

FM1	SE
LDR	00770an 2200199 a 4300
001	000081300
005	2015008142016.0
006	m d
007	cr ca
008	950912c20140999aditf pas 0 b rupa d
040	1
0410	1
08204	1
1300	1
24510	1
24603	1
340	1
310	1
521	1
3670	1
500	1
509	1
500	1
500	1
500	1
504	1
5100	1
534	1
553	1
563	1
563	1
563	1
65014	1
65024	1
7001	1
7001	1
7001	1
7102	1
77608	1
77608	1
85440	1
904	1



Catálogo: ALPH Versión 21 Eje: BR01 (B) y 48 (BR01) Servidor: 132.240.77.54001 Usuario: ALPH

111 Particulares

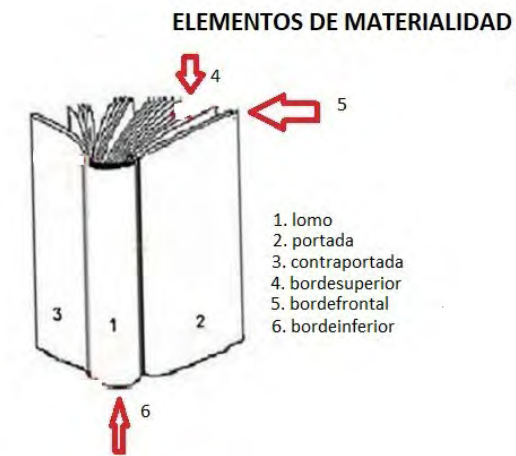
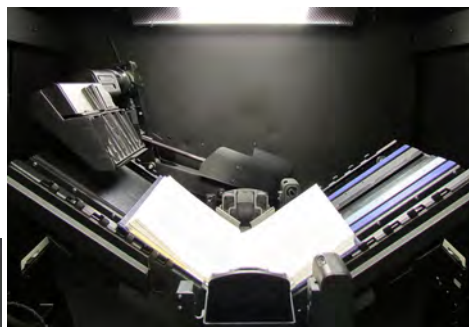
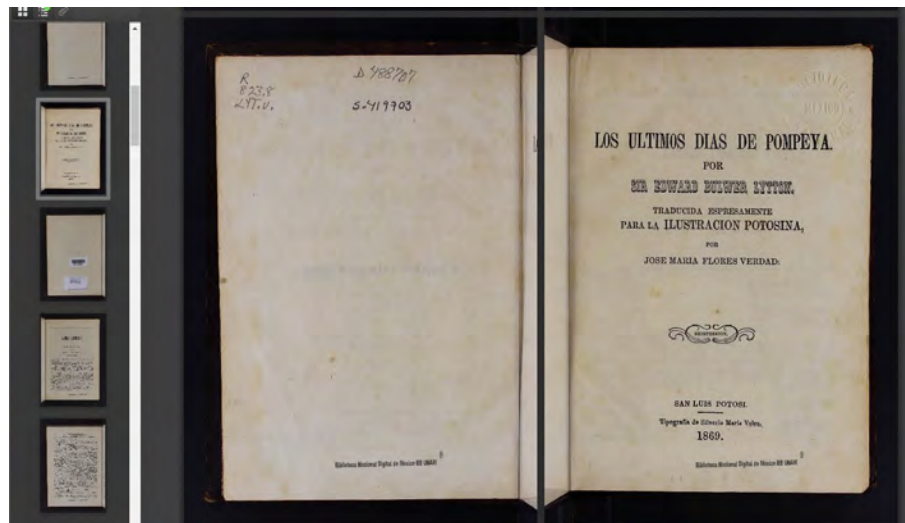
Doc. no.	Clasif.	Autor	Título	Año	Pa. Impronta	Num. Sig.
17500			Suscripción:	1700	[Méjico : s. n.,	000174325
21500	204.72055	México,	Memoria:	1800	[Méjico,	000116070
31900	672.0355	Soubis, Maurice	La aventura de Roucaut Bouber en Méjico /	1700	[Méjico] :	000282261

11111

FM1	SE
001	000081300
005	2015008142016.0
006	m d
007	cr ca
008	950912c20140999aditf pas 0 b rupa d
040	1
0410	1
08204	1
1300	1
24510	1
24603	1
340	1
310	1
521	1
3670	1
500	1
509	1
500	1
500	1
504	1
5100	1
534	1
553	1
563	1
563	1
563	1
65014	1
65024	1
7001	1
7001	1
7001	1
7102	1
77608	1
77608	1
85440	1
904	1

Representación – proceso de captura

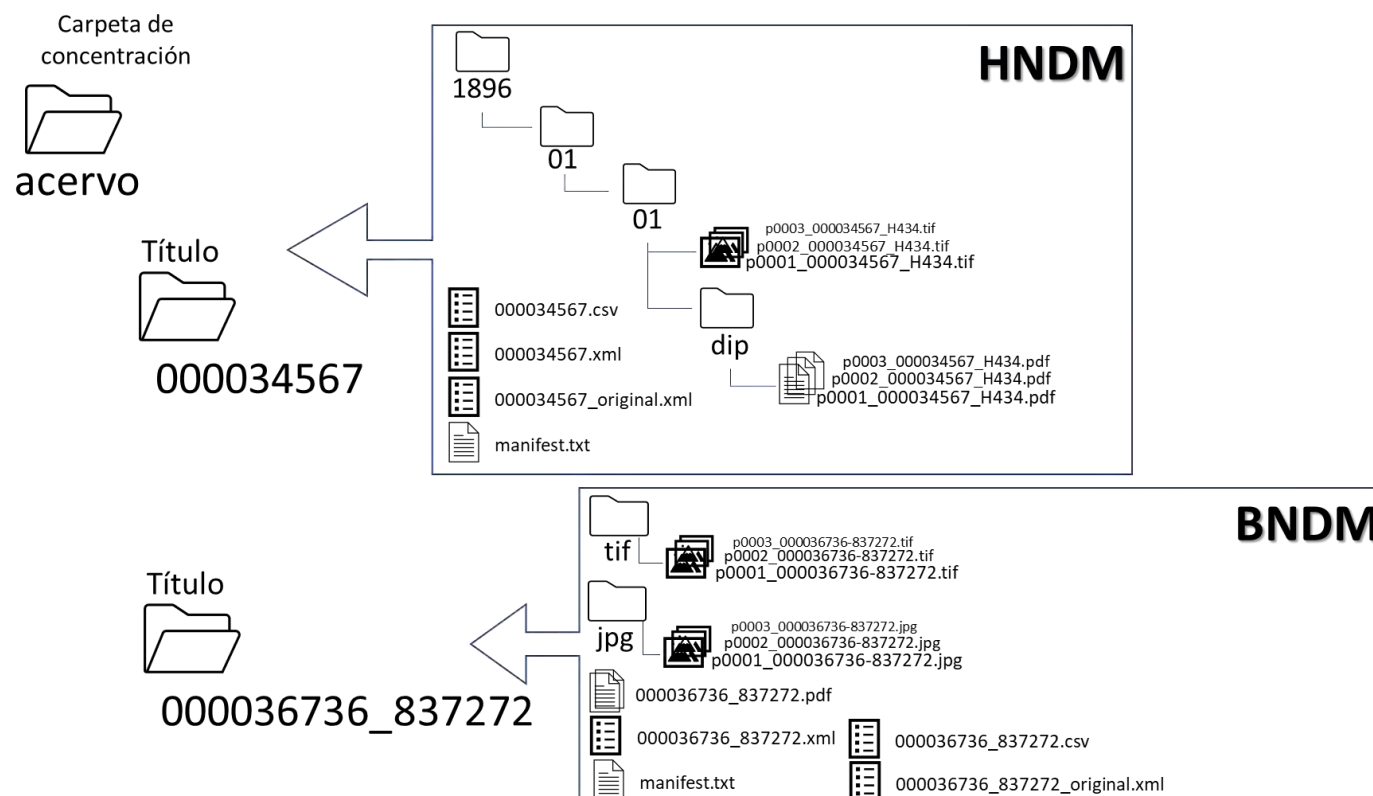
- Registro del proceso de captura



desplegable01 ... desplegable15



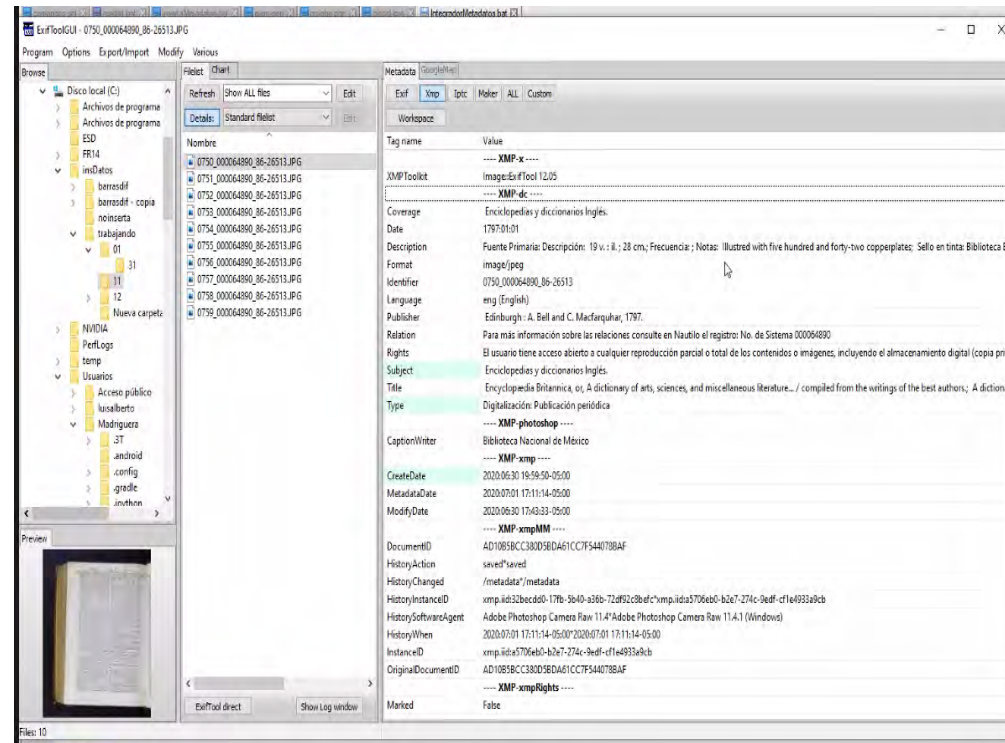
Contenido - estructura



- OCR embebido en PDF
- Carpetas por No. de sistema
- Uso de códigos de barra y números de sistema Aleph

Contenido - metadatos

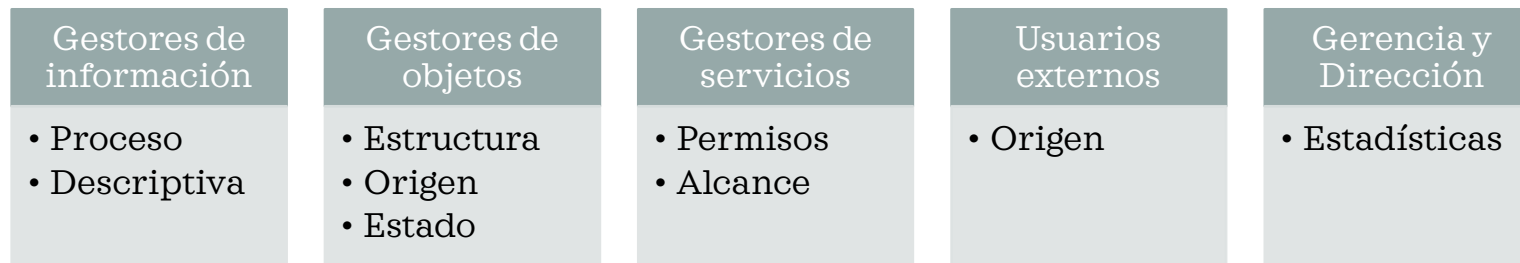
- ❑ MARC 21 - Dublin Core (embebido)
- ❑ MarcXML
- ❑ Fomentar el registro de bitácoras de proceso.



Campo Marc 21	Dublin Core	Notas
700, 710, 711	<u>Contributor</u>	Asiento secundario del contenido.
651, 662, 751, 752	<u>Coverage</u>	Cobertura espacial y/o temporal del contenido.
100, 110, 111	<u>Creator</u>	Asiento principal del contenido.
008/07-10, 260\$ScSg, 264\$Sc	<u>Date</u>	Fecha de publicación del contenido.
300, 310, 500-599, <u>except</u> 506, 530, 540, 546	<u>Description</u>	Descripción física del contenido.
(Este dato se asigna en automático)	<u>Format</u>	Formato del objeto digital.
<u>pNumeroProgresivo_NumSistema_CodigoBarras</u>	<u>Identifier</u>	Permite ubicar al ítem físico del que proviene la digitalización, así como el número de sistema registrado en Aleph.
008/35-37, 041\$a\$b\$d\$e\$f\$g\$h\$j, 546	<u>Language</u>	Código de lenguaje del contenido (ver si se puede poner el equivalente completo)
260\$a\$b, 264\$a\$b	<u>Publisher</u>	Pie de imprenta del contenido (lugar y editor)
530, 760-787\$o\$t	<u>Relation</u>	Consultar con BN y HN si es pertinente este campo.
506, 540	<u>Rights</u>	<i>Derechos del contenido:</i> 506, 540 <i>Objeto Digital:</i> Es la leyenda de los derechos de acceso que tiene el objeto digital.
534\$t, 786\$o\$t, (Todos los datos se tomarán del ítem)	<u>Source</u>	Clasificación, volumen y ejemplar se tomarán del ítem. <i>Fuente del contenido:</i> 534\$t, 786\$o\$t.
600, 650, 651	<u>Subject</u>	Palabras clave o materia del contenido (temas)
245, 246	<u>Title</u>	Título del contenido.
Leader06, Leader07, 655	<u>Type</u>	<i>Tipo de objeto digital:</i> Digitalización (puede cambiar de acuerdo al origen del objeto digital) <i>Tipo de contenido:</i> Leader06, Leader07, 655.

Comunidad designada

Necesidades de Información

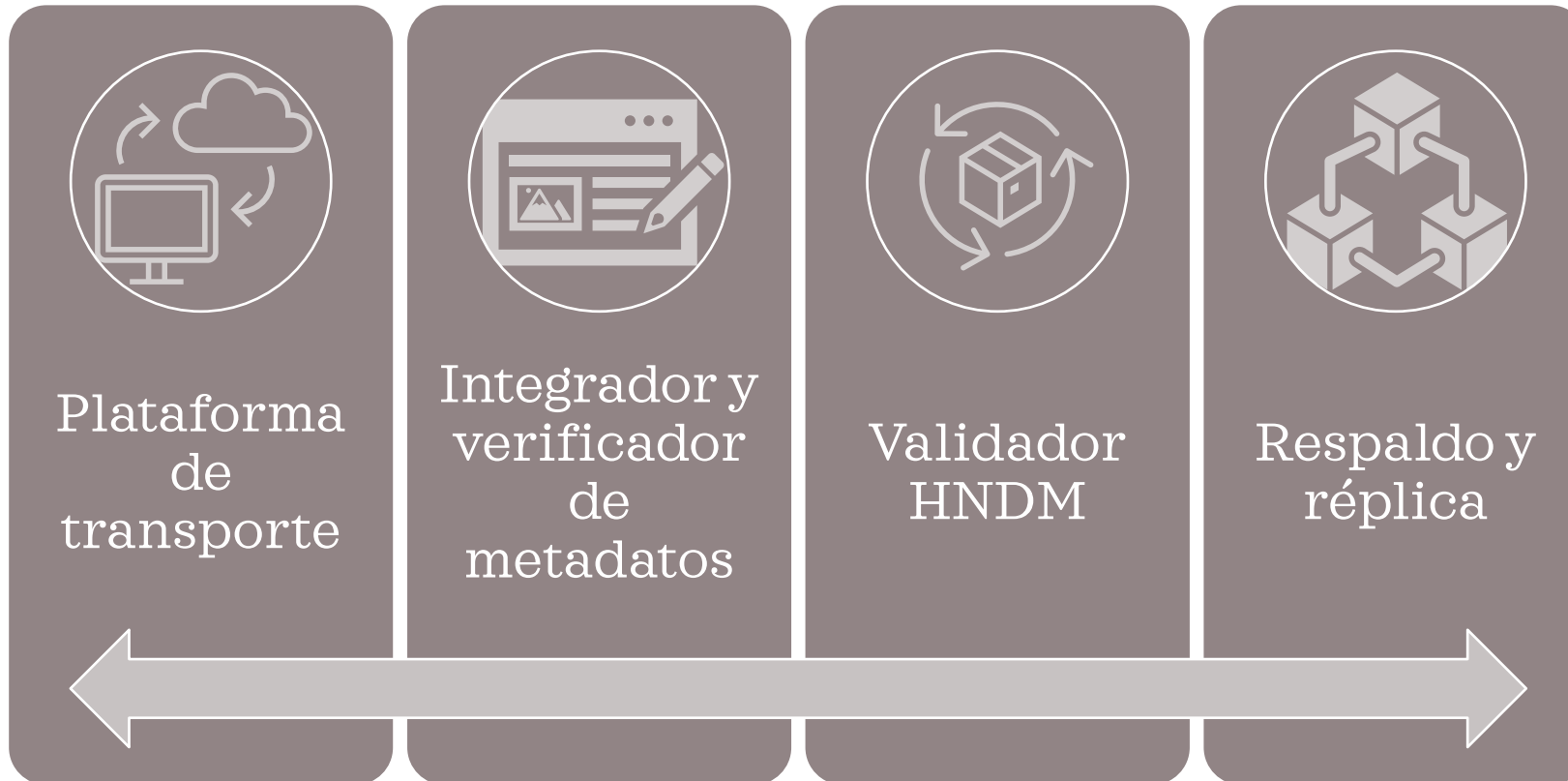




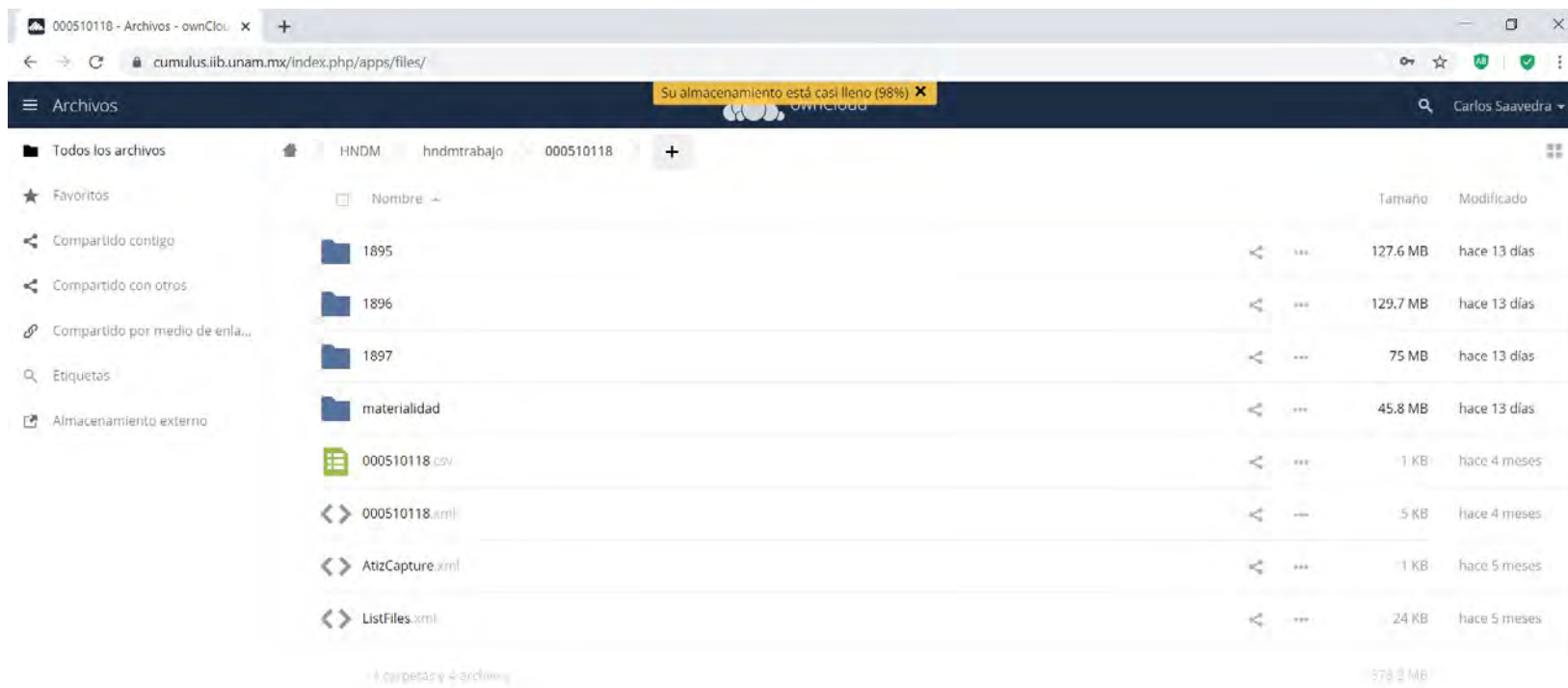
HERRAMIENTAS

Uso y desarrollo

Adaptación y desarrollo de herramientas



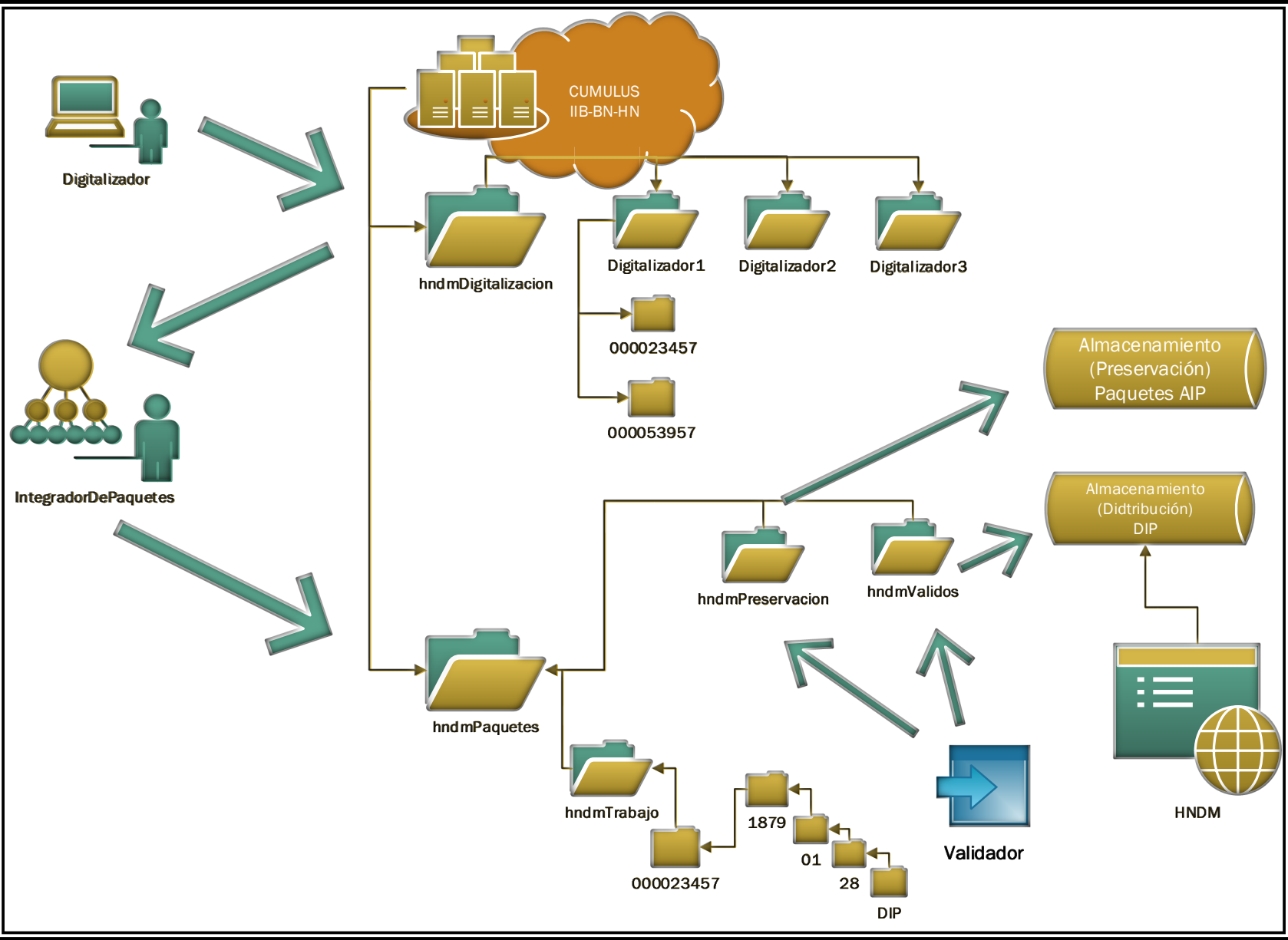
Cúmulus (nube IIB-BN-HN)



Plataforma de transporte

- Software de código abierto
- Estable
- Con posibilidades de soporte y mantenimiento
- Control interno

Trabajo realizado por : Mtra. Erika Arreguin



Plataforma de transporte

- Para el manejo de objetos digitales
- Evitar uso de discos duros externos.
- Asegurar el respaldo de información
- Optimizar el espacio de almacenamiento.

Integrador y verificador de

- Desarrollo propio
- Uso de exiftool
- Programación por lotes.
- Programación PHP
- Verifica que el material esta registrado en Aleph

```

C:\WINDOWS\system32\cmd.exe
*****
*          Script para la integracion de metadatos          *
*          version 2.7 - CIED - DIT                        *
*  Aviso: Los textos a continuacion pueden no contener acentos o *
*  caracteres especiales.                                  *
*  *
*  Precondiciones:                                       *
*  *
*  - Los archivos a trabajar ya deben estar nombradas con *
*  p<progresivo>_<numeroDeSistemaDeAleph>_<codigoDeBarras *
*  *
*  - La publicacion debe estar dada de alta en el catalogo *
*  contar con un numero de sistema y al menos un codigo d *
*  *
*****

El directorio de trabajo es: C:\insDatos\

Ingrese el nombre del directorio con archivos a procesar

Buscando la informacion del recurso en: http://132.247.131.26/sc_json.php?

Se recuperaron correctamente los datos del servidor

-----
Se localizo el titulo:
"Encyclopædia Britannica, or, A dictionary of arts,
compiled from the writings of the best authors. A dictionary of arts, sciences, and miscellaneo
us literature..."

-----
Presione ENTER para iniciar con el ingreso de metadatos, si desea cancelar el proceso, cierre la
ventana desde la opcion "cerrar" en la superior derecha:

Presione una tecla para continuar . . .

```

```

REM Autor: Luis Alberto Hernández Macías
REM Desc: Script de procesamiento por lotes para la integracion de metadatos en archivos variados (Spq,tiff,pdf)
REM Version: 2.7
REM Fecha: 17/06/2021

REM Se declaran las rutas a utilizar, seccion de configuracion
REM Asegurarse que existen los directorios definidos
REM Directorio raiz
set dirTrabajo=C:\insDatos\
REM URL para la descarga del archivo JSON con metadatos
set servAleph=http://132.247.131.26/extraccion/metadatos_marc_dublin_misc_json.php?
REM URL para la descarga del archivo marc-xml
set servMarc=http://catalogo.iib.unam.mx:1851/rest-dif/record/RNM01
set servMarcCom=view=full

REM Codigo de error
REM Error al consultar el archivo JSON generado
set codigoError="Error-JSON"
REM Error al consultar el archivo MARC-XML generado
set codigoErrorMarc="Record does not exist"

setlocal enabledelayedexpansion

REM Se declaran los contadores que se imprimiran al final del proceso
set /a suceFile=0
set /a errFile=0
set /a suceJPG=0
set /a errJPG=0
set /a suceJIF=0
set /a errJIF=0
set /a sucePDF=0
set /a errPDF=0
set /a errArchivo=0
set /a datosCorrectos=0

REM Se imprimen las indicaciones en pantalla
echo *****
echo          Script para la integracion de metadatos          *
echo          version 2.7 - CIED - DIT                        *
echo  *  Aviso: Los textos a continuacion pueden no contener acentos o *
echo  *  caracteres especiales.                                  *
echo  *
echo  *  Precondiciones:                                       *
echo  *  - Los archivos a trabajar ya deben estar nombradas con el formato: *
echo  *  p"<progresivo>_<numeroDeSistemaDeAleph>_<codigoDeBarras">.<extencion"> *
echo  *
echo  *  - La publicacion debe estar dada de alta en el catalogo y debe *
echo  *  contar con un numero de sistema y al menos un codigo de barras valido *
echo  *
echo  *
echo *****
echo.

```

Trabajo realizado por:

- Lic. Luis Alberto Hernández
- Ing. José Luis Estrada

Integrador y verificador de metadatos

```

Presione una tecla para continuar . . .

Se utilizara la extension .jpg

Integracion de metadatos exitoso en: C:\insDatos\trabajando\01\31\0750_000064890_86-26513.JPG
Integracion de metadatos exitoso en: C:\insDatos\trabajando\01\31\0751_000064890_86-26513.JPG
Integracion de metadatos exitoso en: C:\insDatos\trabajando\01\31\0752_000064890_86-26513.JPG
Integracion de metadatos exitoso en: C:\insDatos\trabajando\01\31\0753_000064890_86-26513.JPG
Integracion de metadatos exitoso en: C:\insDatos\trabajando\01\31\0754_000064890_86-26513.JPG
Integracion de metadatos exitoso en: C:\insDatos\trabajando\01\31\0755_000064890_86-26513.JPG
Integracion de metadatos exitoso en: C:\insDatos\trabajando\01\31\0756_000064890_86-26513.JPG
Integracion de metadatos exitoso en: C:\insDatos\trabajando\01\31\0757_000064890_86-26513.JPG
Integracion de metadatos exitoso en: C:\insDatos\trabajando\01\31\0758_000064890_86-26513.JPG
Integracion de metadatos exitoso en: C:\insDatos\trabajando\01\31\0759_000064890_86-26513.JPG
Integracion de metadatos exitoso en: C:\insDatos\trabajando\12\31\0750_000064890_86-26513.JPG
Integracion de metadatos exitoso en: C:\insDatos\trabajando\12\31\0751_000064890_86-26513.JPG
Integracion de metadatos exitoso en: C:\insDatos\trabajando\12\31\0752_000064890_86-26513.JPG
Integracion de metadatos exitoso en: C:\insDatos\trabajando\12\31\0753_000064890_86-26513.JPG
Integracion de metadatos exitoso en: C:\insDatos\trabajando\12\31\0754_000064890_86-26513.JPG
Integracion de metadatos exitoso en: C:\insDatos\trabajando\12\31\0755_000064890_86-26513.JPG
Integracion de metadatos exitoso en: C:\insDatos\trabajando\12\31\0756_000064890_86-26513.JPG
Integracion de metadatos exitoso en: C:\insDatos\trabajando
Integracion de metadatos exitoso en: C:\insDatos\trabajando
Integracion de metadatos exitoso en: C:\insDatos\trabajando
Integracion de metadatos exitoso en: C:\insDatos\trabajando
  
```

Se utilizara la extension .tif

Se utilizara la extension .pdf

```

Total de archivos .jpg integrados exitosamente: 4
Total de archivos .jpg con errores: 0

Total de archivos .tif integrados exitosamente: 0
Total de archivos .tif con errores: 0

Total de archivos .pdf integrados exitosamente: 0
Total de archivos .pdf con errores: 0
  
```

```

Se solicita el archivo MARC-XML en URL:
http://catalogo.iib.unam.mx:1891/rest-dlf/record/BNM01000315215?view=full
  
```

```

-----
 % Total   % Received % Xferd  Average Speed   Time    Time     Time  Current
          Dload  Upload   Total     Spent    Left     Speed
100 2750     0 2750    0     0    2750    0  ---:--:--  ---:--:--  ---:--:-- 19503
-----
  
```

```

Archivo MARC-XML descargado en: C:\insDatos\barrasdif\marcxml.xml
  
```

```

Termina la ejecucion del programa.
  
```

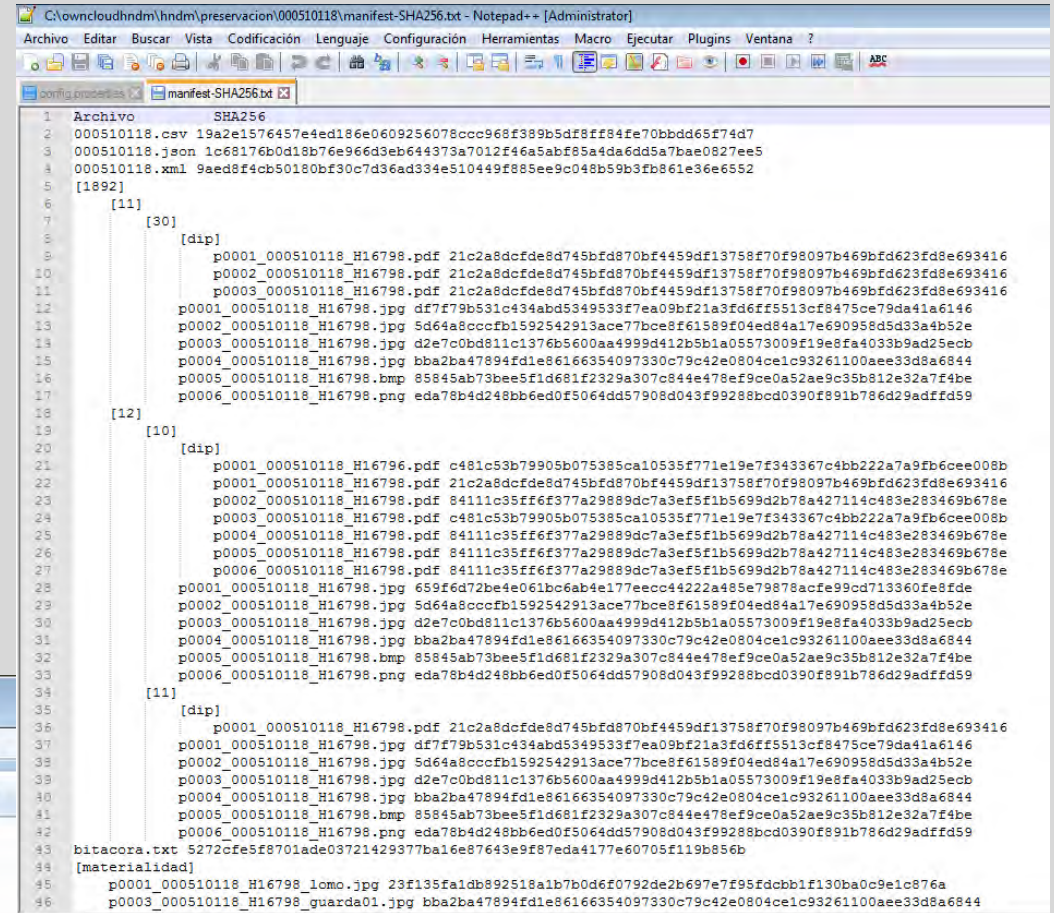
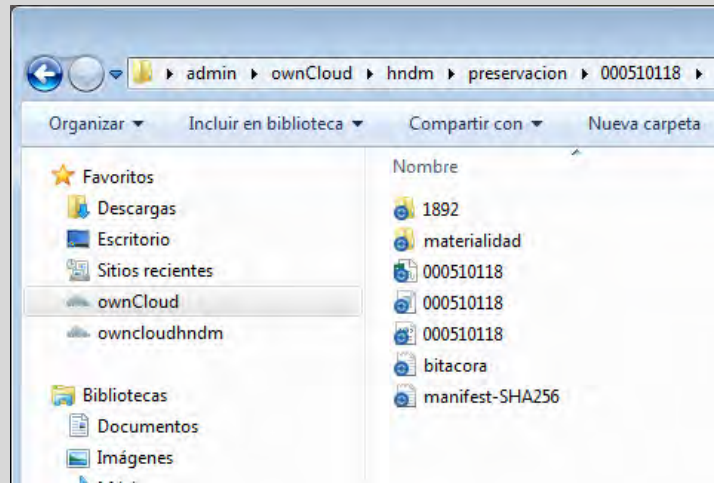
```

Presione una tecla para continuar . . .
  
```

- Inserta metadatos en archivos:
- PDF, JPG y TIFF
- Recorre todas las ramas del directorio contenedor
- Descarga archivo XMLmarc

Validador HNNDM

- Desarrollo propio
- Lenguaje Java
- Manejador de base de datos MongoDB
- Configurable



- 07/10/2020 11:34 a... Archivo de valores...
- 07/10/2020 11:34 a... Archivo JSON
- 07/10/2020 11:34 a... Documento XML
- 07/10/2020 11:34 a... Documento de tex...
- 07/10/2020 11:34 a... Documento de tex...

Trabajo realizado por :

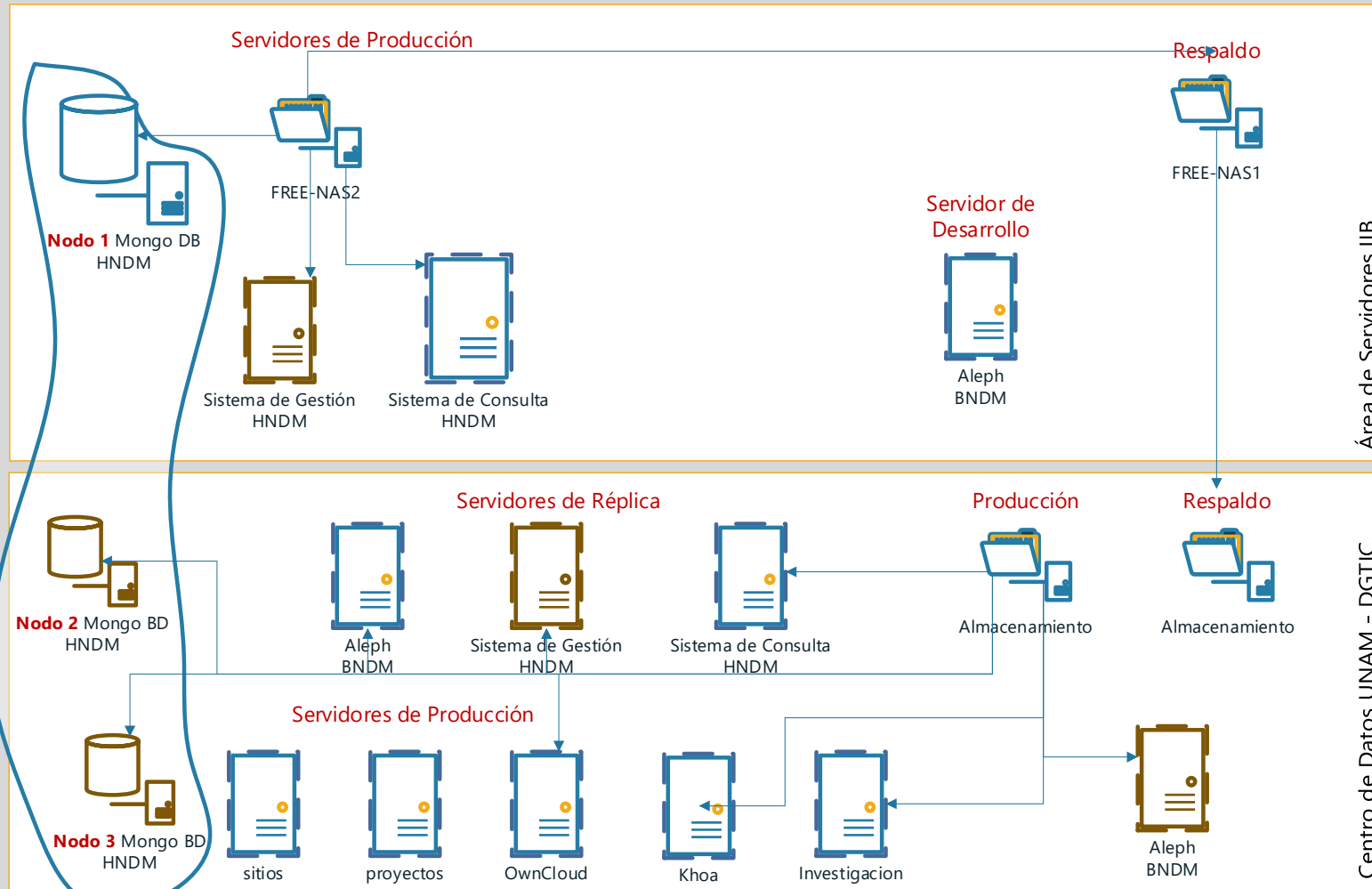
- Mtro. Lisandro Pablo Olivares

Validador HNDM

- Verifica existencia de XML descargado de Aleph
 - Número sistema y título
- Comprueba en Aleph
 - No. De Sistema
 - Código de barras
- Revisa estructura de carpetas
 - Nombrado de archivos
 - OCR
- Genera paquete DIP
- Genera paquete SIP con huella digital

The image displays three overlapping screenshots of the 'Validador HNDM' software interface. The main window shows a file list with columns for file names and counts (e.g., [12], [10], [dip]). A 'confirmacion' dialog box is visible, asking for confirmation to proceed with 6 of 10 PDF files. A 'Mensaje' dialog box shows an error: 'Error en la estructura de carpeta o nombrado de archivos, revise el archivo bitácora.txt'. The bottom screenshot shows a successful validation process with green checkmarks and status messages: 'La carpeta está bien formada.', 'Validación terminada exitosamente', 'Creando carpeta DIP...', 'Carpeta DIP creada.', 'Creando carpeta SIP...', 'Carpeta SIP creada.', 'Insertando registro en la base de datos...', and 'Registro en la base de datos realizado.' Buttons for 'Abrir Bitácora' and 'Cerrar' are present at the bottom of the main window.

Respaldo y réplica



- ✓ Datos
- ✓ Servicios

Trabajo realizado por :

- Mtro. Luis Miguel Estrada
- Mtro. José Antonio Salazar



GRACIAS

TODO EL TRABAJO AQUÍ PRESENTADO ES INICIATIVA DE LA COORDINACIÓN DE INNOVACIÓN Y ESTRATEGIA DIGITAL (CIED) Y ES RESULTADO DEL ESFUERZO DE SUS INTEGRANTES, SIN EMBARGO, SE HA PUESTO EN MARCHA SÓLO GRACIAS A LA PARTICIPACIÓN DE TODAS LAS ÁREA DE PROCESOS TÉCNICOS Y ADMINISTRACIÓN DE LA BNM Y HNM

Agosto 19 del 2021

Mtra. Ana Yuri Ramírez Molina

Coordinación de Innovación y Estrategia Digital