

RAIMON NUALART

# PLAN DE PRESERVACIÓN DIGITAL

# HOJA DE RUTA Y

# SOLUCIONES



2022

Logroño  
9-11 de noviembre

XXVII Jornadas de  
**Archivos Universitarios**

Preservación digital:  
retos, experiencias, soluciones

**30** UNIVERSIDAD  
DE LA RIOJA  
1992-2022

**crue** Universidades Españolas Secretarías Generales Archivos

obsolescencia  
política de preservación  
disponibilidad  
**Transferencia**  
planes de preservación  
soportes  
ciclo de vida Integridad patrimonio  
archivo electrónico recursos  
formatos emulación PREMIS  
autenticidad expediente ENI sistema de preservación METS  
conversion- modelo OAIS  
e-ark archivo electrónico único estrategias de preservación  
cloud  
análisis de riesgos  
accesibilidad migración ENS  
capacidad  
virtualización  
Planes de acciones



# PRINCIPIOS DE LA PRESERVACIÓN DIGITAL

La **preservación digital** es el conjunto de operaciones orientadas a compilar toda la información posible de cada activo digital para que pueda disponer de todos los elementos y componentes necesarios a corto, medio y largo plazo al fin de favorecer su accesibilidad y su reutilización en cualquier momento.

La **conservación digital** es la aplicación de medidas de protección y recuperación con el objetivo de garantizar la perdurabilidad, accesibilidad y usabilidad de los activos digitales a lo largo del tiempo.

(Joan Soler, dixit)

**1 INTEGRIDAD**

**2 ACCESIBILIDAD**

**3 EQUIVALENCIA**

**4 ECONOMÍA**

**5 ACTUALIDAD**

**6 COOPERACIÓN Y NORMALIZACIÓN**

# QUÉ ES UN PLAN DE PRESERVACIÓN DIGITAL

Es un documento estratégico que establece el conjunto de acciones a corto, medio y largo plazo que tienen como fin implementar las estrategias, procesos y procedimientos, encaminadas a asegurar la preservación a largo plazo de los activos digitales de una institución



# CARACTERÍSTICAS PLAN DE PRESERVACIÓN DIGITAL

**1** COMPLETO

**2** REALISTA

**3** INTEGRADO

**4** DINÁMICO

**5** ANÁLISIS  
NECESIDADES

# CÓMO IMPLEMENTAR EL PLAN: HOJA DE RUTA

**FASE 1 - ESTABLECER LAS BASES**

**FASE 2 - DIAGNÓSTICO**

**FASE 3 - EVALUACIÓN ESTRATEGIAS**

**FASE 4 - PLAN DE ACCIÓN**



# FASE 1 - ESTABLECER LAS BASES

La preocupación principal de la preservación digital es **mantener la capacidad de acceder al contenido** de los activos digitales a lo largo del tiempo.

Adicionalmente, hay que garantizar su integridad, autenticidad, fiabilidad y seguridad

Aspectos a tener en cuenta:

- Tipologías activos digitales
- Obsolescencia formatos
- Integridad y autenticidad de los activos
- Metadatos de preservación
- Gestión del riesgo
- Monitorización y trazabilidad
- Aplicación estrategias de preservación
- Archivo digital y sistema de almacenamiento
- Estandarización y buenas prácticas
- Capacitación y formación continua

## FASE 1.1- OBJETIVOS

Definición de los objetivos de alto nivel y adaptados en un entorno determinado. Han de ser:  
Concretos, medibles, asumibles, relevantes y tienen un tiempo

## FASE 1.2 ALCANCE

Se debe enfocar en definir y controlar qué se incluye y qué se excluye dentro del mismo.  
Identificación fondos y variedad de activos digitales a conservar

## FASE 1.3- ALINEAR

El plan de preservación digital ha de estar alineado con las diferentes políticas, planes y programas de la entidad

## FASE 1.4 ROLES Y RESPONSABILIDADES

La entidad debe definir y asignar los roles y responsabilidades para efectuar el desarrollo, ejecución, actualización y monitorización o del plan

# **CÓMO IMPLEMENTAR EL PLAN**

**FASE 1 - ESTABLECER LAS BASES**

**FASE 2 - DIAGNÓSTICO**

**FASE 3 - EVALUACIÓN ESTRATEGIAS**

**FASE 4 - PLAN DE ACCIÓN**



# FASE 2 - DIAGNÓSTICO



## FASE 2.1 IDENTIFICACIÓN ACTIVOS DIGITALES A PRESERVAR

Selección responsable y fundamentada  
Identificación tipologías, volumetría, software relacionado, procedencia y formas de captura

## FASE 2.2 DIAGNÓSTICO Y MUESTRA REPRESENTATIVA

Selección de una muestra representativa de los activos digitales del archivo y determinar las características esenciales de cada uno de los elementos de la muestra y que serán los retos de preservación a mantener a lo largo del tiempo

## FASE 2.3- ANÁLISIS DE RIESGO

Identificar las amenazas y las debilidades de los activos (analizar las vulnerabilidades) para determinar el impacto y evaluar las medidas que se pueden implementar con el fin de mitigarlos al máximo

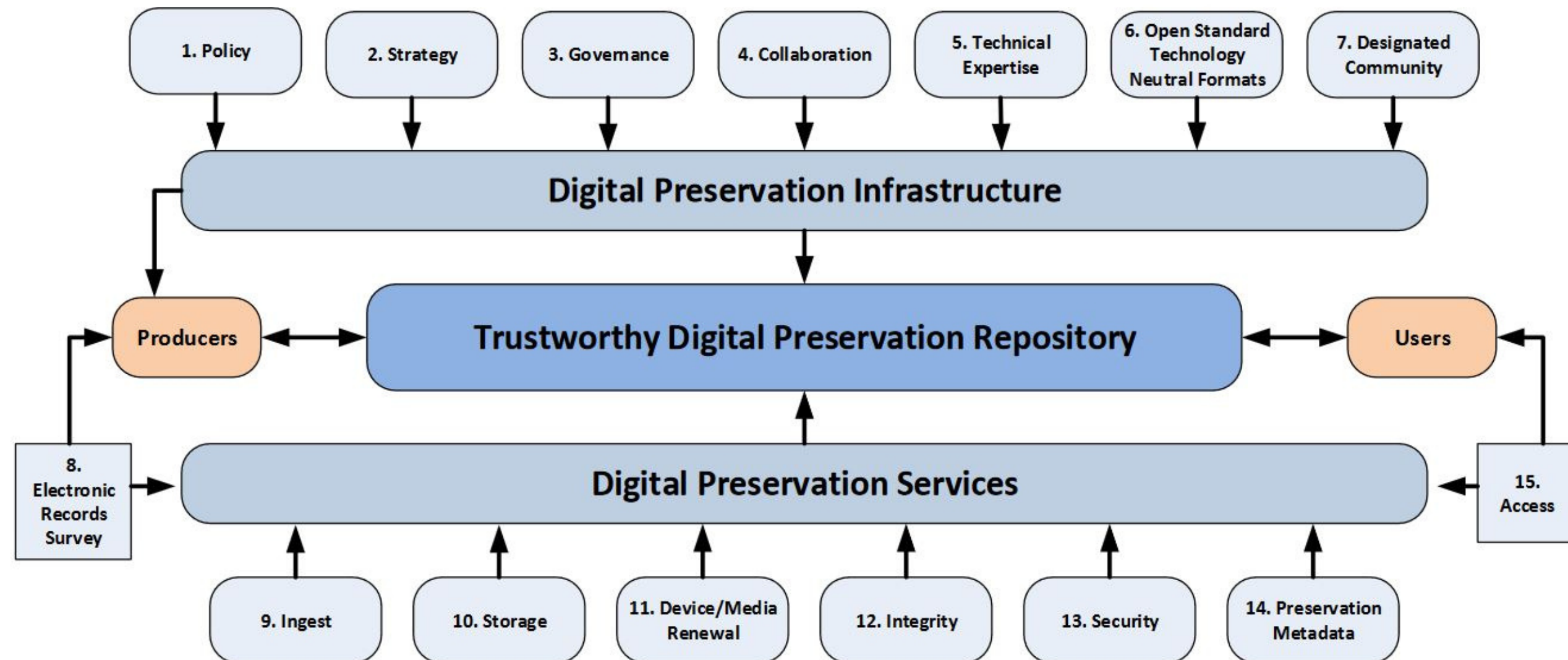
## FASE 2.4 EVALUAR MADUREZ DE LA ENTIDAD

La entidad debe realizar un diagnóstico y medir el nivel y capacidad de preservación digital que tiene frente a los temas de preservación digital, aplicando modelos de madurez de referencias



# DPCMM -DIGITAL PRESERVATION CAPABILITY MATURITY MODEL

## 1 - COMPONENTES A EVALUAR



# DPCMM -DIGITAL PRESERVATION CAPABILITY MATURITY MODEL

## 2 - MÉTRICAS DE CAPACIDAD

Level	Ingest Capability Metrics
0	The organization does not have a digital preservation repository capable of receiving or ingesting long-term and permanent electronic records.
1	The preservation repository receives electronic records from records producers based on ad hoc agreements without regard to format, integrity, virus checks, and metadata quality. None of this rises to the level of an ISO 14721 conforming SIP.
2	The repository receives surrogate SIPs that are held in a staging area while virus checks and format validations are manually executed. Surrogate AIPs are manually created and transferred to archival storage.
ISO 14721 Conformance	
3	The preservation repository ingests SIPs through <b>semi-automated means</b> that validate the completeness of Administration, Technical, Provenance, Content Description, and Preservation Description significant properties. The significant properties are extracted from SIPs and written to Preservation Description Information (PDI). Archival Information Packages (AIPs) are created and transferred to the repository's storage function.
4	The preservation repository ingests SIPs through <b>automated means</b> that validate the completeness of Administration, Technical, Provenance, Content Description, and Preservation Description significant properties. The significant properties are extracted from SIPs and written to Preservation Description Information (PDI). Archival Information Packages (AIPs) are created and transferred to the repository's storage function.

# DPCMM -DIGITAL PRESERVATION CAPABILITY MATURITY MODEL

## 3 - RESULTADO



# DPCMM -DIGITAL PRESERVATION CAPABILITY MATURITY MODEL

## 4 - PLANIFICAR

DPCMM Components	Current Capability	Year 1	Year 2	Year 3	Year 4	Year 5	Difficulty
<b>INFRASTRUCTURE</b>							
Policy	1	2	⇒	⇒	3	⇒	LOW
Strategy	0	2	⇒	⇒	3	⇒	MEDIUM
Governance	0	1	2	3	⇒	⇒	MEDIUM
Collaboration	1	2	⇒	3	⇒	⇒	MEDIUM
Technical Expertise	1	⇒	⇒	2	⇒	3	MEDIUM
Open Standard Technology Neutral File Formats	1	⇒	2	⇒	3	⇒	MEDIUM
Designated Community	1	⇒	2	⇒	3	⇒	MEDIUM

# **CÓMO IMPLEMENTAR EL PLAN**

**FASE 1 - ESTABLECER LAS BASES**

**FASE 2 - DIAGNÓSTICO**

**FASE 3 - EVALUACIÓN ESTRATEGIAS**

**FASE 4 - PLAN DE ACCIÓN**

# FASE 3 - EVALUACIÓN ESTRATEGIAS



## FASE 3.1- PRIORIZACIÓN

Crterios para priorizar el proyecto

Identificar los flujos de información y formas de captura

## FASE 3.2 IDENTIFICACIÓN CARACTERÍSTICAS A PROTEGER

Del contenido

- Captura y representación
- Integridad
- Autenticidad
- Accesibilidad
- Fiabilidad

Del sistema

- Plataforma preservación digital
- Seguridad
- Almacenamiento

## FASE 3.3- IDENTIFICACIÓN Y EVALUACIÓN ESTRATEGIAS

La entidad debe definir, evaluar y documentar las estrategias de preservación digital dependiendo de las necesidades y los requisitos identificados para planificar las acciones de preservación que se van a ejecutar

# **CÓMO IMPLEMENTAR EL PLAN**

**FASE 1 - ESTABLECER LAS BASES**

**FASE 2 - DIAGNÓSTICO**

**FASE 3 - EVALUACIÓN ESTRATEGIAS**

**FASE 4 - PLAN DE ACCIÓN**



# FASE 4 - PLAN DE ACCIÓN



## FASE 4.1- DEFINICIÓN DE ACCIONES Y ESTRATEGIAS

Definir y establecer qué acciones se van a implementar y su cronograma, en función de las bases establecidas, diagnóstico y evaluación de estrategias.

## FASE 4.2 DEFINICIÓN RECURSOS Y CRONOGRAMA

Realizar un cronograma a corto, medio y largo plazo, donde se incluyan los hitos y actividades específicas a realizar para la aplicación de estas estrategias, así como, la asignación de los recursos necesarios.

## FASE 4.3- IMPLEMENTACIÓN DEL PLAN



Más información

## FASE 4.4 EVALUACIÓN Y MEJORA CONTINUA

# ELEMENTOS BÁSICOS DEL PLAN DE ACCIÓN

1

## SISTEMA DE PRESERVACIÓN

- Creación sistema de preservación y archivo siguiendo modelo OAIS
- Implementación medidas de seguridad - ENS

2

## PROCEDIMIENTOS Y GUÍAS

- Identificación y análisis de activos digitales a preservar
- Transferencia al sistema de preservación
- Intercambio de AIP entre sistemas
- Puesta disposición y remisión (consulta)
- Administración del sistema de preservación
- Guía de conservación
- Instrucciones técnicas de seguridad



[Guía conservación](#)

[Guía de transferencia](#)

3

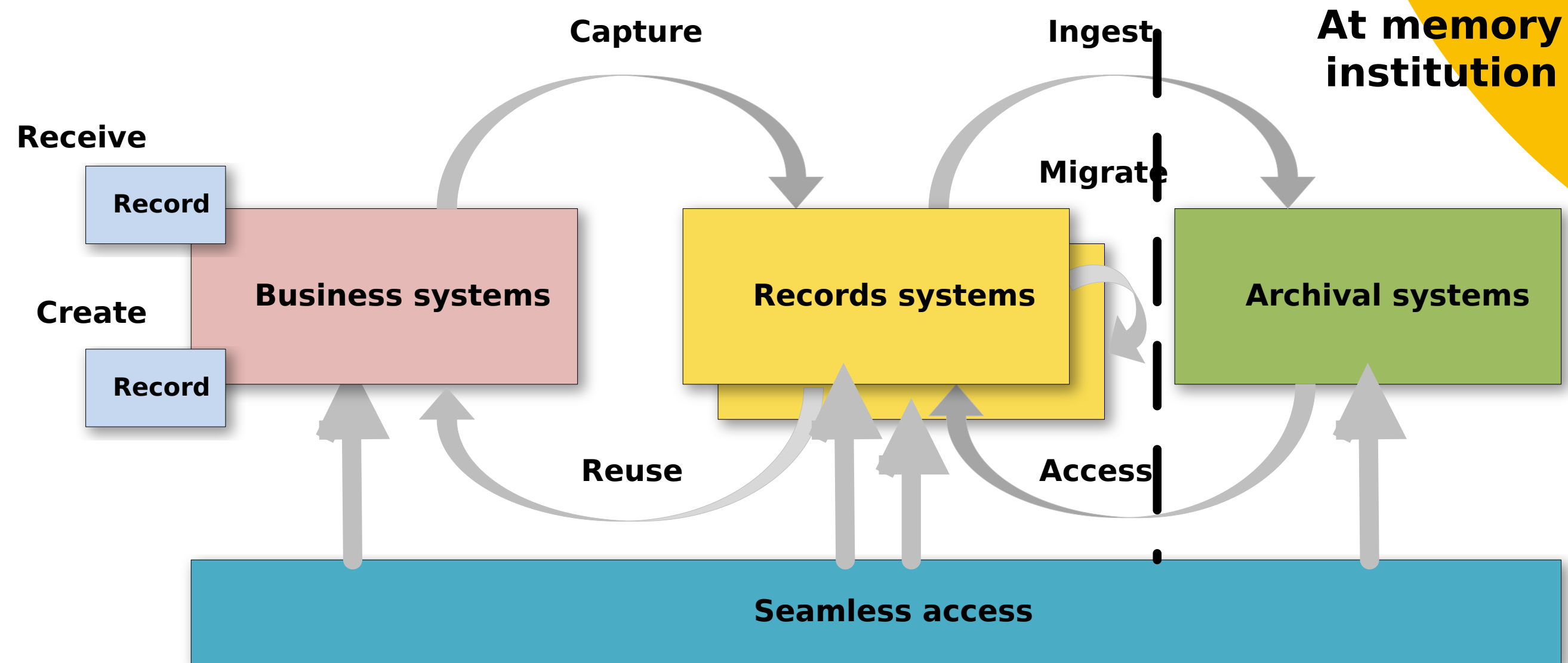
## ESTRATEGIAS Y ACCIONES DE PRESERVACIÓN

- Captura activos digitales
- Mantener integridad y autenticidad
- Detección formatos
- Normalización formatos
- Servicios de emulación y virtualización
- Metadatos de preservación
- Sistema almacenamiento

# EL ARCHIVO DIGITAL COMO EJE DEL SISTEMA

## Reflexiones

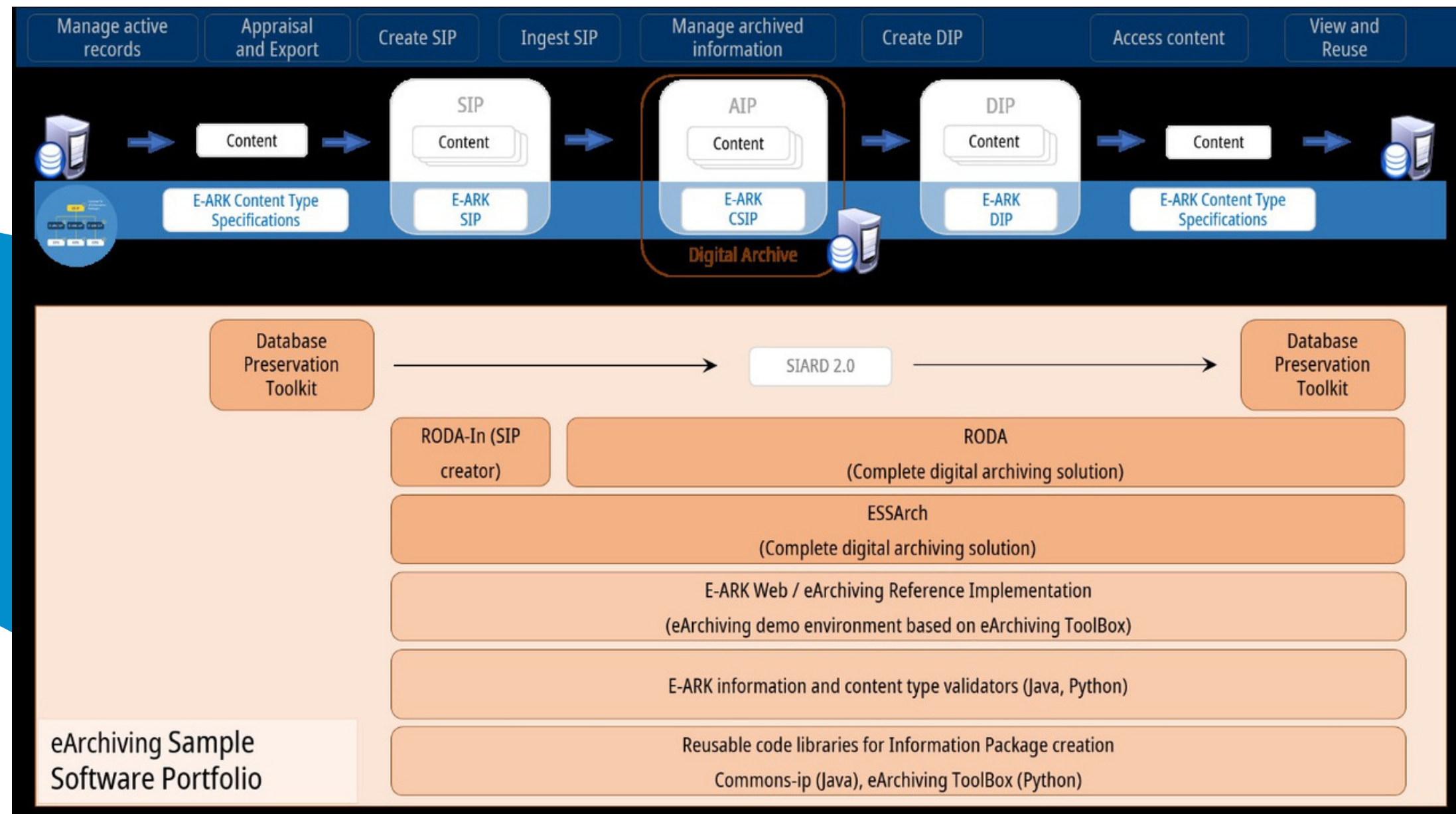
- El modelo OAIS es la referencia para implementar sistemas de archivo orientados a la preservación a largo plazo. Existen alternativas?
- Es necesario establecer estos tres sistemas? Se puede simplificar ya que los procesos de transferencia entre sistemas de información son costosos
- Podríamos crear microservicios de preservación y aplicarlos en función del ciclo de vida y las necesidades detectadas sin necesidad de transferir?



# SOLUCIONES ARCHIVO DIGITAL

## Modelo europeo:

- Proporcionar software independiente de código abierto que se pueda instalar y utilizar para tareas específicas del archivo digital.
- Proporcionar componentes de software de código abierto para integrar las especificaciones de archivo electrónico en sistemas y flujos de trabajo



[Github Software E-ARK](#)

## Soluciones globales

[Roda](#)

[ESSArch](#)

[Database Preservation Toolkit](#)

## Herramientas creación y validación SIP/AIP/DIP

[eArchiving ToolBox \(eATB\)](#)

[E-ARK Information Package Validation](#)

[E-ARK IP Viewer](#)

# SEGURIDAD, EL "NUEVO ENS"

**Objetivo principal:** crear las condiciones necesarias de seguridad en el uso de los medios electrónicos, a través de medidas para garantizar la seguridad de los sistemas, los datos, las comunicaciones, y los servicios electrónicos que el sistema de preservación ofrece a su comunidad de usuarios.

## MEDIDAS SEGURIDAD (73)

- **Marco organizativo:** conjunto de medidas relacionadas con la organización global de la seguridad
- **Marco operacional:** medidas a tomar para proteger la operación del sistema como conjunto integral de componentes para un fin
- **Medidas de protección:** Se centran en resguardar activos específicos, conforme su naturaleza y la calidad demandada por el nivel de seguridad de las dimensiones perjudicadas

[Portal ENS](#)

[Real Decreto 311/2022, "el nuevo ENS"](#)

## PRINCIPIOS BÁSICOS

- Seguridad como proceso integral
- Gestión de la seguridad basada en los riesgos
- Prevención, detección, respuesta y conservación
- Existencia de líneas de defensa
- Vigilancia continua.
- Reevaluación periódica
- Diferenciación de responsabilidades.

## CATEGORÍAS Y DIMENSIONES DE SEGURIDAD

- Confidencialidad (C), Integridad (I), Trazabilidad (T), Autenticidad (A) y Disponibilidad (D)
- BAJO/MEDIO/ALTO (Nivel seguridad requerido de las dimensiones de seguridad)
- BÁSICA/MEDIA/ALTA (Categoría de seguridad del sistema)

# ELEMENTOS BÁSICOS DEL PLAN DE ACCIÓN

1

## SISTEMA DE PRESERVACIÓN

- Creación sistema de preservación y archivo siguiendo modelo OAIS
- Implementación medidas de seguridad - ENS

2

## PROCEDIMIENTOS Y GUÍAS

- Identificación y análisis de activos digitales a preservar
- Transferencia al sistema de preservación
- Intercambio de AIP entre sistemas
- Puesta disposición y remisión (consulta)
- Administración del sistema de preservación
- Guía de conservación
- Instrucciones técnicas de seguridad



[Guía conservación](#)

[Guía de transferencia](#)

3

## ESTATEGIAS Y ACCIONES DE PRESERVACIÓN

- Captura activos digitales
- Mantener integridad y autenticidad
- Detección formatos
- Normalización formatos
- Servicios de emulación y virtualización
- Metadatos de preservación
- Sistema almacenamiento

# ESTRATEGIAS Y TÉCNICAS DE PRESERVACIÓN DIGITAL



# MODELOS REPRESENTACIÓN "OBJETOS" A PRESERVAR

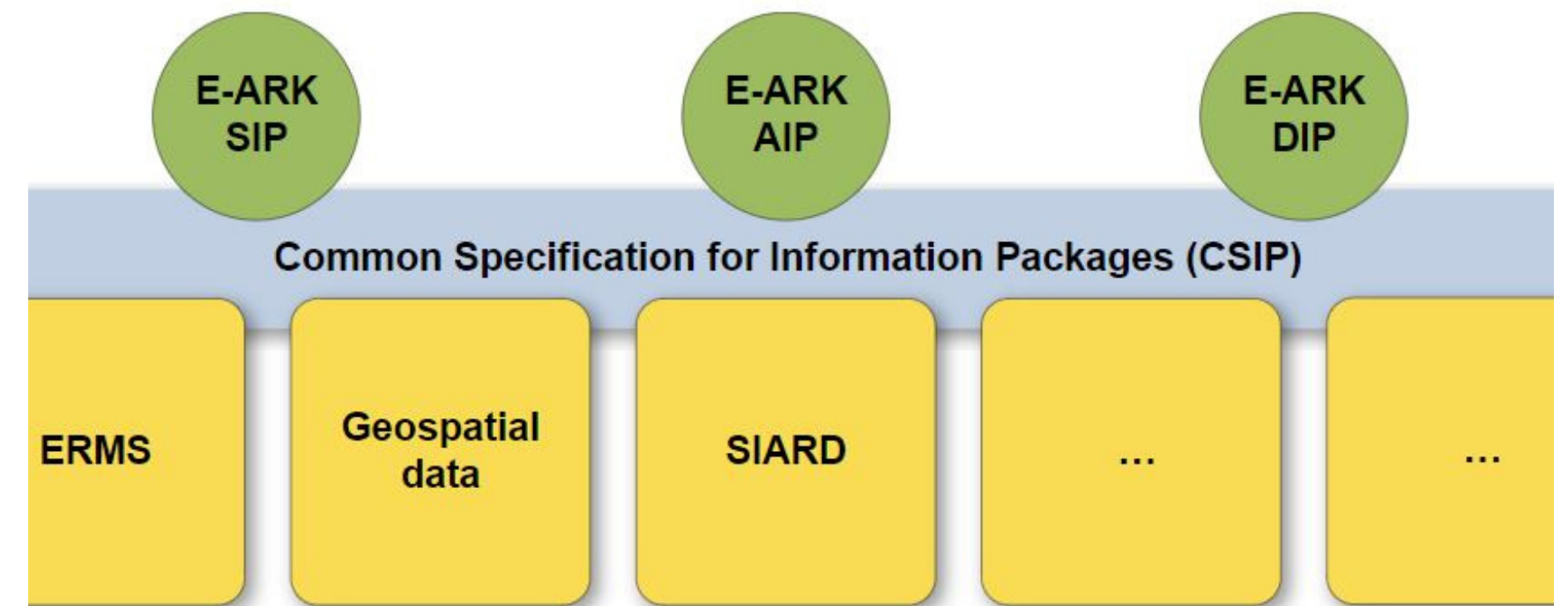
## ESPAÑOL



[NTI](#)

[Manual intercambio XML](#)

## EUROPEO



[Especificaciones](#)

[Contenedor formato METS](#)



# OTROS ACTIVOS DIGITALES A PRESERVAR

## REDES SOCIALES



### SOLUCIONES

- Descargar contenido desde las APP (titular). Capturas temporales
- Descarga via API (titular y de otras cuentas/hastags)

[TNA: mirror web](#)

[Twitter: TWARC](#)

## WEB



### SOLUCIONES

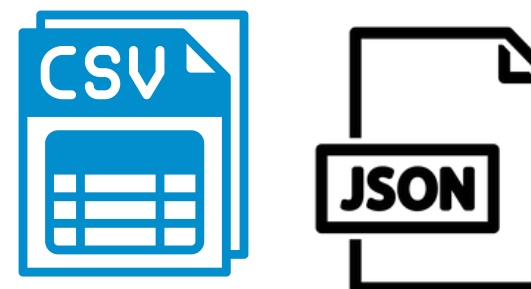
- WARC Web ARChive format
- Capturas "snapshot" temporales

[warcreate](#)

[Padicat](#)

[TNA: web archives](#)

## DATASETS



### SOLUCIONES

- Comprobaciones calidad datos
- Metadatos de los datasets
- Acumulativo vs historificación
- Formatos abiertos

[DCAT-AP](#)

[Guia metadatos \(cat\)](#)

## BBDD



### SOLUCIONES

- Documentar y estructurar BBDD
- Historificación
- SIARD (no válido para BBDD NoSQL)

[SIARD](#)

[Database Preservation Toolkit](#)

# ESTRATEGIAS PARA GARANTIZAR LA INTEGRIDAD Y AUTENTICIDAD

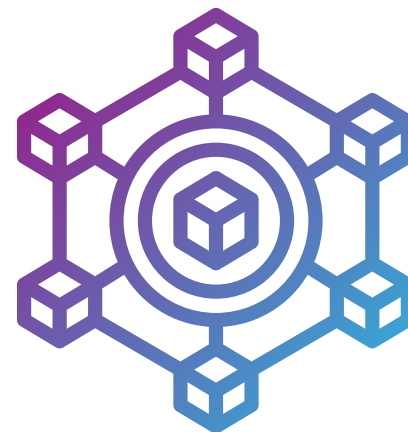
## INTEGRIDAD (1)



### SOLUCIONES

- Calcular hash de los objetos en el momento de ingreso
- Comprobaciones periódicas
- Confianza en la institución

## INTEGRIDAD (2)



### SOLUCIONES

- Almacenamiento hash en la DTL
- Depósito descentralizado de las evidencias del contenido
- No archiva contenidos

[Archangel - github](#)

[Archangel](#)

## AUTENTICIDAD

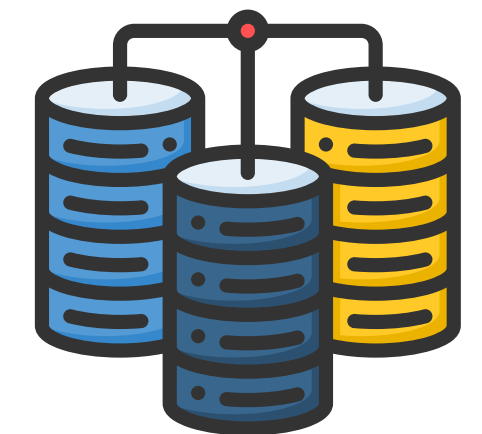


### SOLUCIONES

- Uso formatos firma longeva (XAdES-A, PAdES-LTV)
- Implica resellado continuo
- Sostenible?

[Digital Signature Service \(DSS\)](#)

[Trusted List](#)



### SOLUCIONES

- Confianza en el archivo
- Extracción metadatos firmas en el proceso de ingreso
- Firma de los AIP?
- Necesario validar firmas?

[NTI firma digital](#)

# SISTEMAS DE ALMACENAMIENTO

1

**RENOVACIÓN MEDIOS**



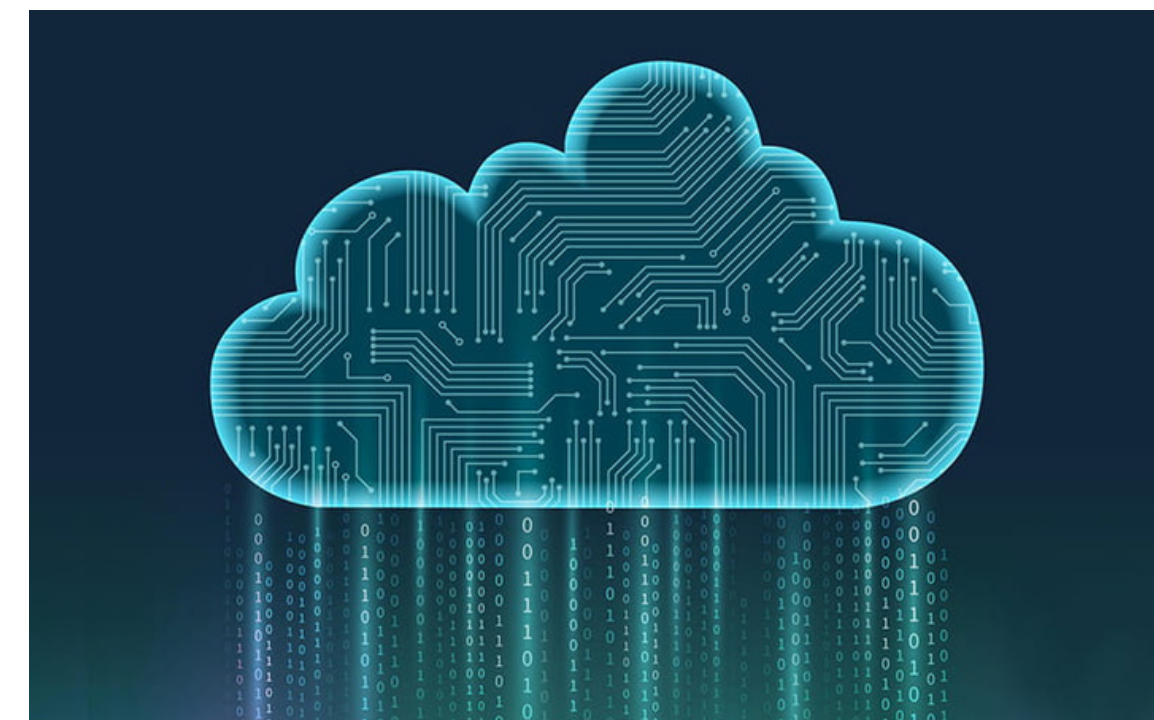
2

**CPD PROPIO**



3

**CLOUD STORAGE**



[UNE-ISO/TR 22428-1:2021 IN](#)

[Check-list Interpares Trust](#)

# ESTRATEGIAS PARA GARANTIZAR LA ACCESIBILIDAD

## FORMATOS



### SOLUCIONES

- Establecer política de formatos
- Registro formatos
- Herramientas detección de formatos

[NTI catálogo estandares](#)

[PRONOM](#)

[DROID](#)

[JHOVE](#)

## CONVERSIÓN



### SOLUCIONES

- Conversión a formatos estandarizados
- Herramientas validación de formatos
- Regulación jurídica

[veraPDF](#)

[DPF Manager](#)

[MediaConch](#)

[JPYLYZER](#)

[NTI copia](#)

## EMULACIÓN



### SOLUCIONES

- La idea es recrear el entorno tecnológico original con tecnología actual sin necesidad de modificar o alterar el contenido original

[Emulation Framework](#)

[DosBox](#)

## VIRTUALIZACIÓN



### SOLUCIONES

- Creación a través de un software de una versión virtual de cualquier recurso informático hardware o software
- Se mantienen aplicaciones obsoletas

[Olive Executable Archive](#)

[Virtual Box](#)

# ESTRATEGIAS PARA GARANTIZAR LA ACCESIBILIDAD

## DIFUSIÓN



### SOLUCIONES

- Protocolo OAI-PMH
- Metadatos descriptivos  
Dublin Core
- Solo contenidos públicos
- Portales de difusión



[OAI-PMH](#)

# ESTRATEGIAS PARA GARANTIZAR LA FIABILIDAD: EL PAPEL DE LOS METADATOS

El **PREMIS Data Dictionary for Preservation Metadata** es el estándar internacional de metadatos para respaldar la preservación de objetos digitales y garantizar su fiabilidad a largo plazo.



## QUÉ INCLUYE?

- Diccionario de datos e información sobre los metadatos de preservación
- Define 4 entidades y sus metadatos (objetos, eventos, agentes, derechos)



## CÓMO USUARLO?

- Proporciona un esquema XML para describir los objetos, eventos, agentes y derechos
- Importar y exportar paquetes entre archivos
- Generar el paquete de información de archivo



[Esquema metadatos](#)

[Guia implementación](#)

[Unidades semánticas](#)

**P R E M I S**

# CONCLUSIÓN

IMPOSIBLE NADA ES.  
DIFÍCIL, MUCHAS COSAS SON

Yoda dixit





# MUCHAS GRACIAS

RAIMON NUALART MERCADÉ  
RNUALART@AOC.CAT