

Repository of Authentic Digital Objects (RODA)

Armida Escalante.
Antonio Contreras.

Carlos Román.
Antonio Salazar.

¿Qué es RODA?

RODA es una solución de repositorio digital que permite la conservación y autenticidad a largo plazo de objetos digitales, para lo cual se basa en estándares y especificaciones internacionales, lo que lo convierte en un repositorio confiable y que ofrece funcionalidad para las unidades principales del modelo de referencia OAIS.

Tiene funcionalidades para llevar a cabo el control de la ingesta, la administración y acceso a diferentes tipos de contenido digital.



Historia

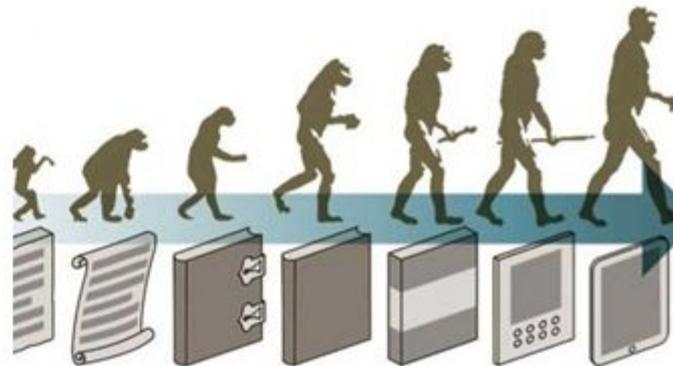
La iniciativa surge en el año 2006 por parte de la Dirección General de Archivos Portugueses (DGARQ) y la participación de la Universidad de Minho.

En 2008 se liberó la primera versión y el repositorio fue hospedado por la propia DGARQ.

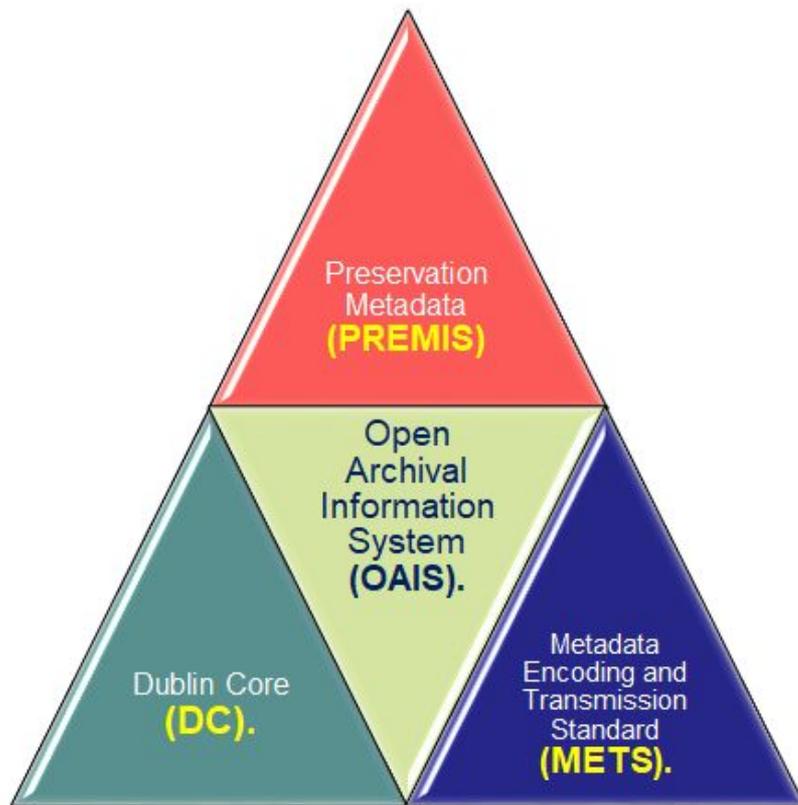
Entre 2011 y 2014 se continuó con el desarrollo cofinanciado por Keep Solutions y la Comisión Europea.

Entre 2014 y 2016, recibió apoyo financiero y técnico del proyecto E-ARK (cofinanciado por KEEP SOLUTIONS y el Programa Marco de Competitividad e Innovación de la Comisión Europea

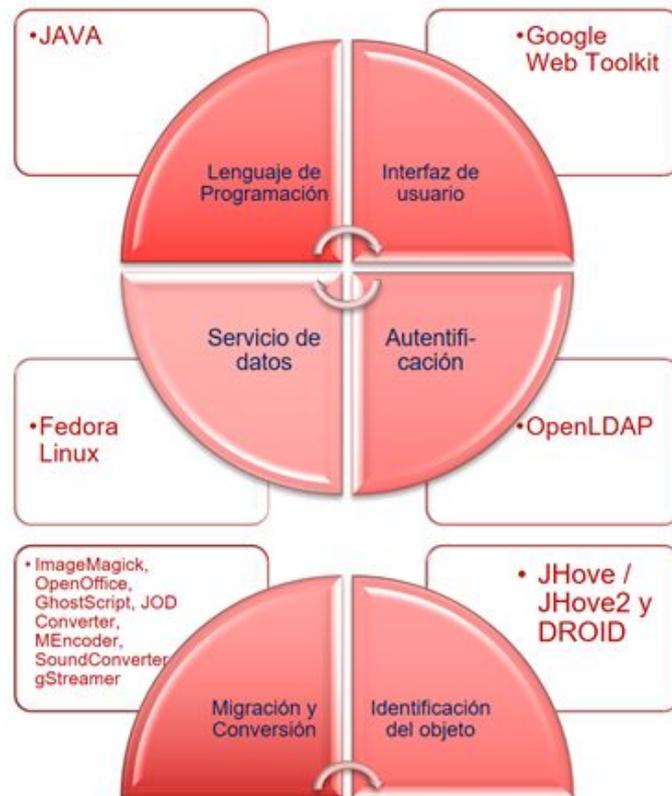
En 2015, el software recibió fondos adicionales de la Dirección General del Libro, Archivos y Bibliotecas (DGLAB, oficialmente conocido como DGARQ), también respaldado por el programa de financiamiento nacional Compete 2020.



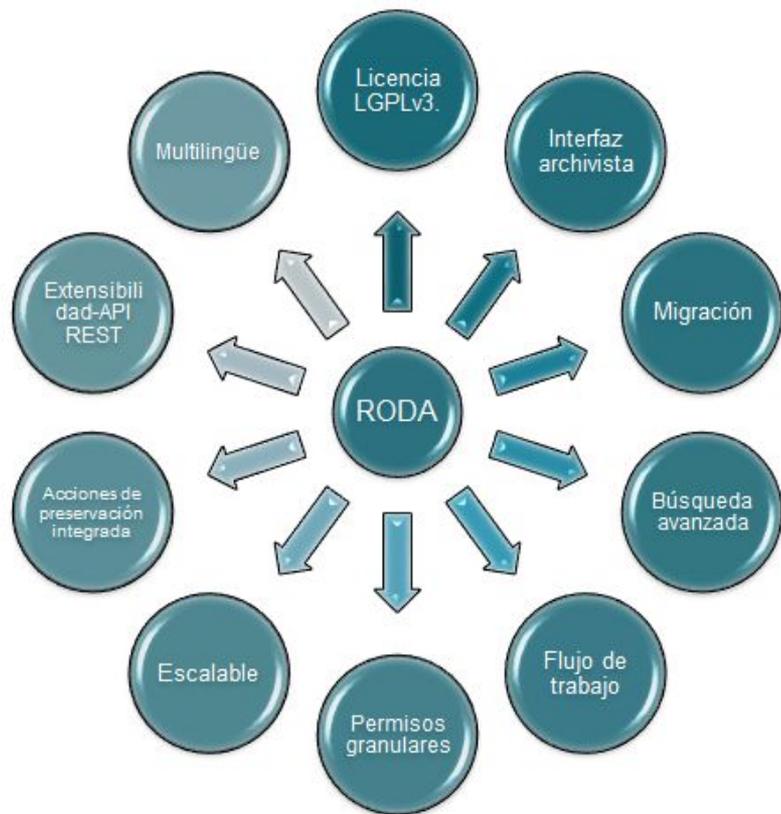
Estándares utilizados



Principales tecnologías utilizadas



Principales Características

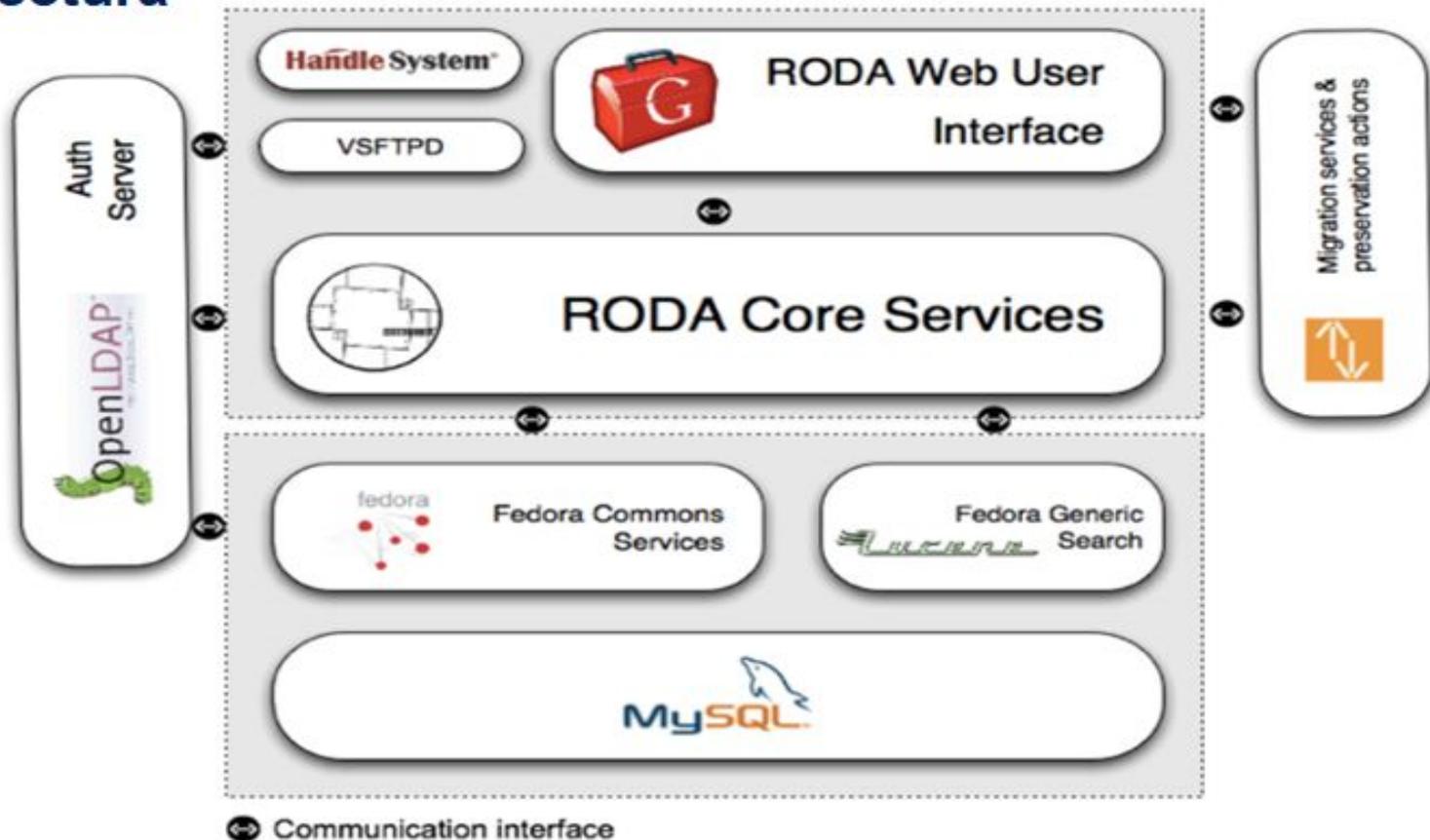


Formatos soportados

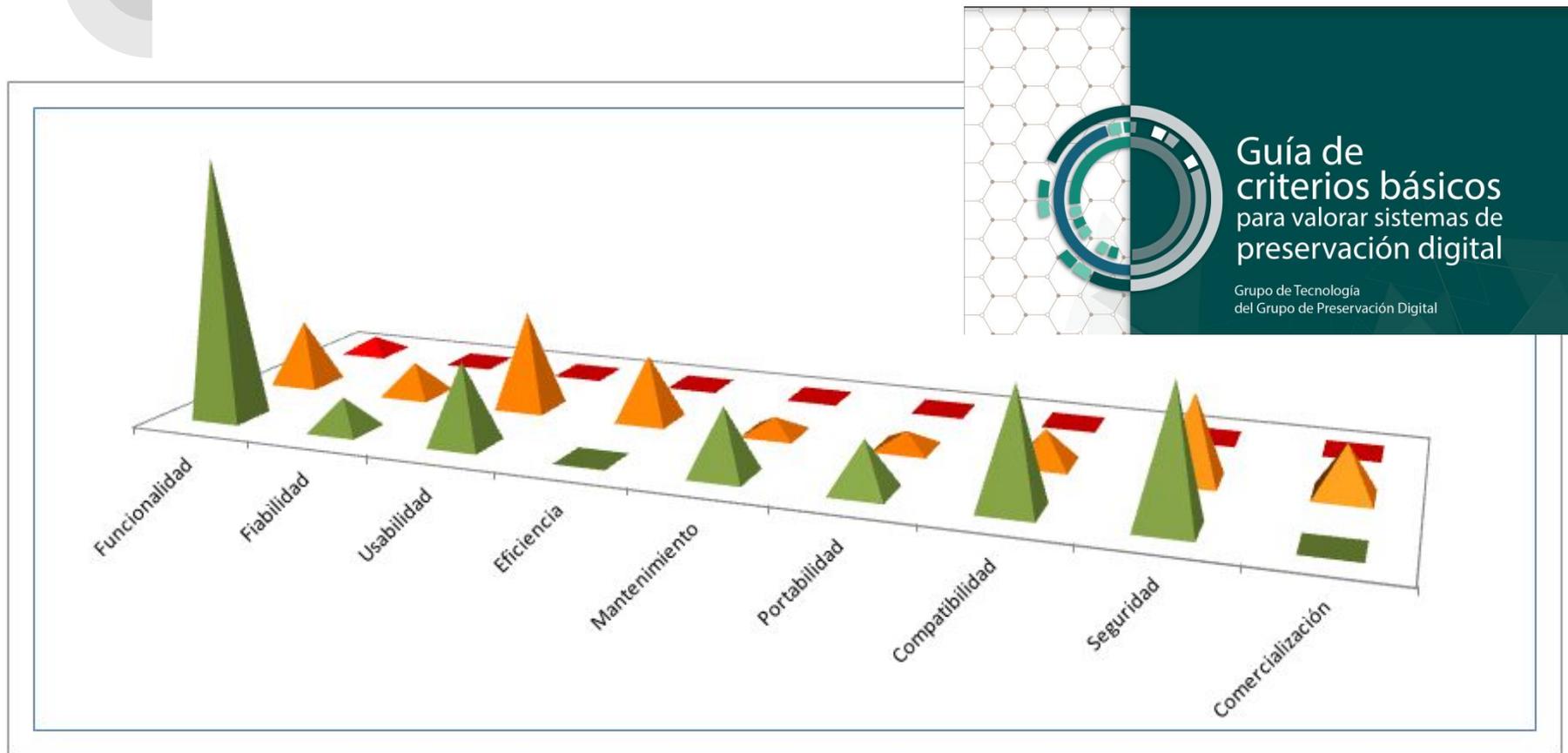


Fuente: Imagen obtenida de Open Repositories 2019, Atlanta, Georgia, USA.
Recuperado de https://smarttech.gatech.edu/bisstream/handle/1853/2846242_608-1-PB.pdf

Arquitectura



Aplicación de la guía de criterios de preservación digital a RODA.



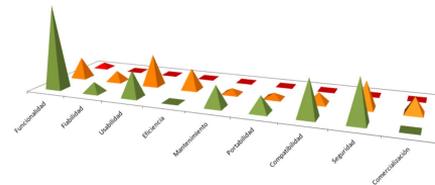
Aplicación de la guía de criterios de preservación digital a RODA.

Funcionalidad	Fiabilidad	Eficiencia
Traslado del acervo	Mecanismos de respaldo	Elementos comprimidos
Respaldos	Replicación del acervo	
Especificación de tipologías	Estadísticas específicas	
Validación en ingesta		
Supervisión de objetos		
Acciones en acervo		
Cosecha de objetos y metadatos		
Depósito de objetos		
Generación de metadatos		
Captura de metadatos		
Personalización de metadatos		
Uso de metadatos estándares		
Almacenamiento de objetos		
Personalización de interfaz de consulta		
Interoperabilidad		

Portabilidad	Compatibilidad	Seguridad
Descarga de metadatos estándares	Estándar OAIS	Políticas de preservación
	Paquetes SIP	Integridad de objetos
	Paquetes AIP	Avisos y acciones automáticos
	Paquetes DIP	Configuración de acciones
	Herramientas para ingesta	Identificadores únicos de objetos
		Metadatos descriptivos, legales, técnicos y de preservación
	Formato de objetos	
	Importar y exportar metadatos	Revisiones de integridad
		Verificación de duplicidad y autenticidad
	Acceso a objetos	



Software

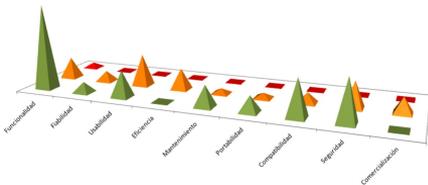


Aplicación de la guía de criterios de preservación digital a RODA.

Funcionalidad	Fiabilidad	Usabilidad	Eficiencia	Mantenimiento
Personalización de interfaces	Registro de bitácora	Capacitación especializada	Herramientas adicionales	Manual de usuario
Reportes y estadísticas		Diseño estandarizado	Tiempos de respuesta	Manual técnico
Control de objetos y procesos		Atajos a funcionalidades	Límites del sistema	Mantenimiento
		Personalización a detalle		Soporte técnico
		Configuración de idioma		
		Mensajes para usuario		
		Mensajes de alerta		
		Mensajes de error		
		Ayudas		
		Ejemplos de información		

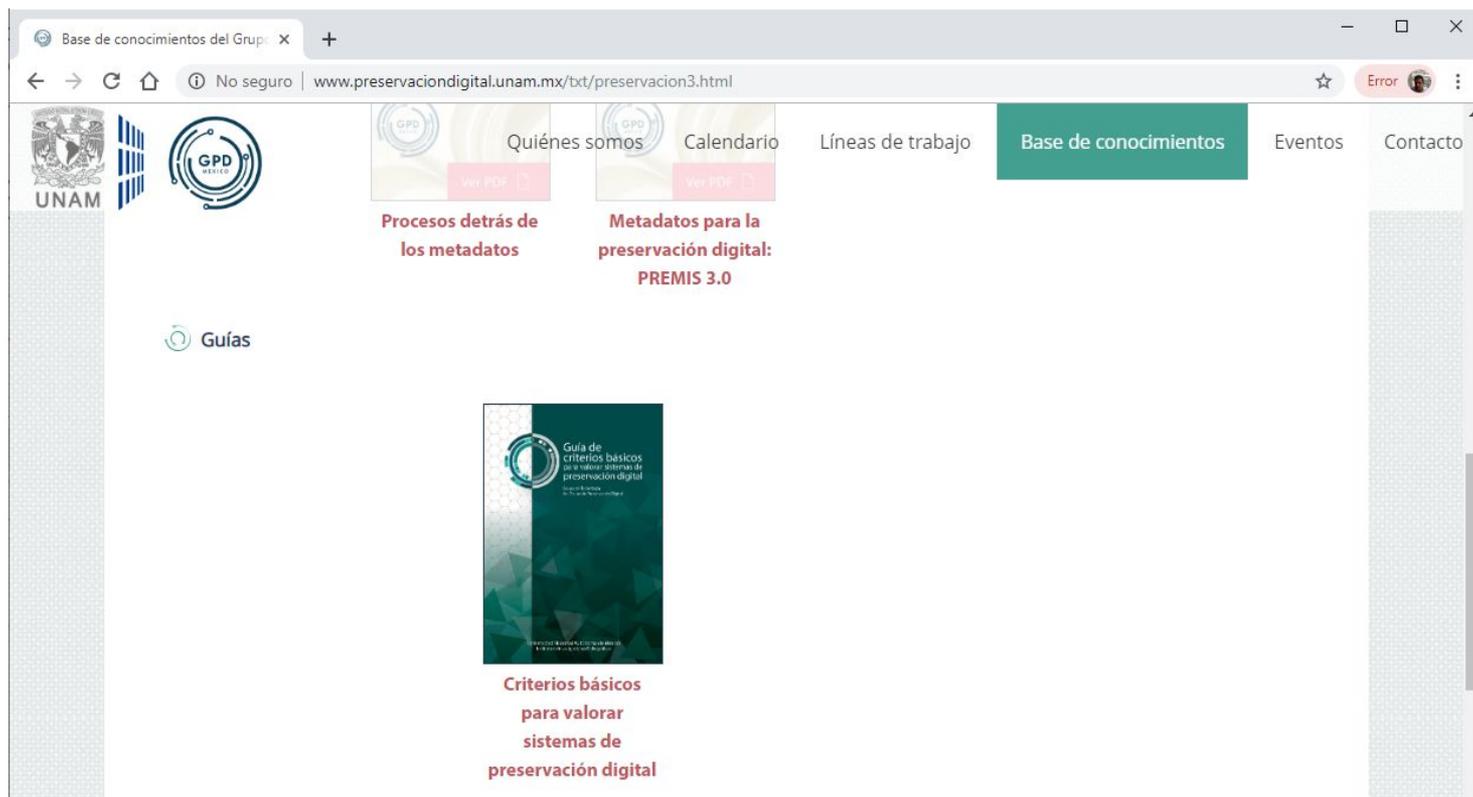


Portabilidad	Seguridad	Distribución y comercialización
Traslado entre plataformas	Permisos de acceso	Años en el mercado de proveedor y producto
Servicios nube	Responsabilidad por bitácora	Zonas geográficas de comercialización
Gestor de base de datos		Cartas de recomendación de usuarios del sistema
		Nuevas versiones
		Madurez del proveedor
		Instituciones que usan el sistema
		Costos, licenciamiento y almacenamiento
		Tiempos de instalación y actualización del sistema
		Tiempos para mantenimiento y soporte técnico
		Requerimientos para instalación



Aplicación de la guía de criterios de preservación digital a RODA.

<http://www.preservaciondigital.unam.mx> >> Base de conocimientos >> Nuestros productos



The screenshot shows a web browser window with the address bar displaying "www.preservaciondigital.unam.mx/bxt/preservacion3.html". The page features a navigation menu with the following items: "Quiénes somos", "Calendario", "Líneas de trabajo", "Base de conocimientos" (highlighted in green), "Eventos", and "Contacto". On the left side, there are logos for "UNAM" and "GPD". Below the navigation menu, there are two product cards: "Procesos detrás de los metadatos" and "Metadatos para la preservación digital: PREMIS 3.0". At the bottom center, there is a large card for "Guías" with a cover image titled "Guía de criterios básicos para valorar sistemas de preservación digital".

Base de conocimientos del Grupo

No seguro | www.preservaciondigital.unam.mx/bxt/preservacion3.html

Quiénes somos | Calendario | Líneas de trabajo | **Base de conocimientos** | Eventos | Contacto

UNAM | GPD

Procesos detrás de los metadatos | Metadatos para la preservación digital: PREMIS 3.0

Guías

Guía de criterios básicos para valorar sistemas de preservación digital

Criterios básicos para valorar sistemas de preservación digital

RODA-In.

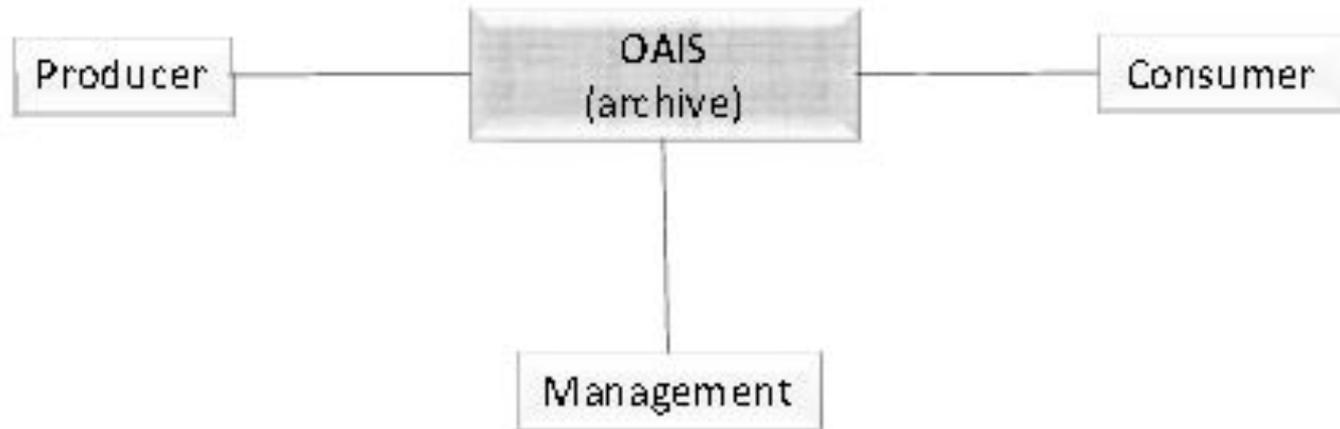
Es una herramienta que facilita la creación de SIP apropiados para ser enviados a un Archivo del tipo OAIS, obteniendo los objetos de datos de carpetas y archivos en un computador local y está dirigido a los productores de objetos de información y los Archivistas.



¿Quién produce un SIP?

En el modelo OAIS existen diferentes vistas, una de ellas es la relación entre los actores involucrados para la preservación de objetos de información en soporte digital . Y que se reconoce como el ambiente de trabajo para un Archivo del tipo OAIS o un OAIS.

Producir o proveedor, es el rol que realizan las personas o instituciones que proveen la información a ser preservada. Y están fuera del dominio del Archivo tipo OAIS.



¿Que es un SIP?

Un **paquete de información** es un **contenedor conceptual** que alberga dos tipos de información:

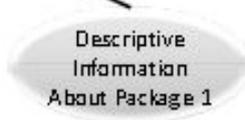
- Content Information
- PDI Preservation Description Information

El **paquete de información** es visible por la información que lo describe. El “target” de la preservación es el Content Information, que consiste de:

- un Objeto de Datos con el Contenido y
- los requisitos para realizar su representación o reproducción.



Package 1



La información de descripción para la preservación incluye:

- Procedencia
- Contexto.
- Referencia.
- Un valor o información que protege de la alteración.
 - Especificación de los derechos de uso, distribución y preservación.

RODA-In. Herramienta para creación de SIP (Submission Information Package)



Con el fin de obtener SIP que cumplan los estándares de calidad para la aceptación en el Archivo del tipo OAIS que usa RODA

RODA-In acepta dos tipos de estándares, para conformar la información de descripción para la preservación.

BagIt (corriente americana)

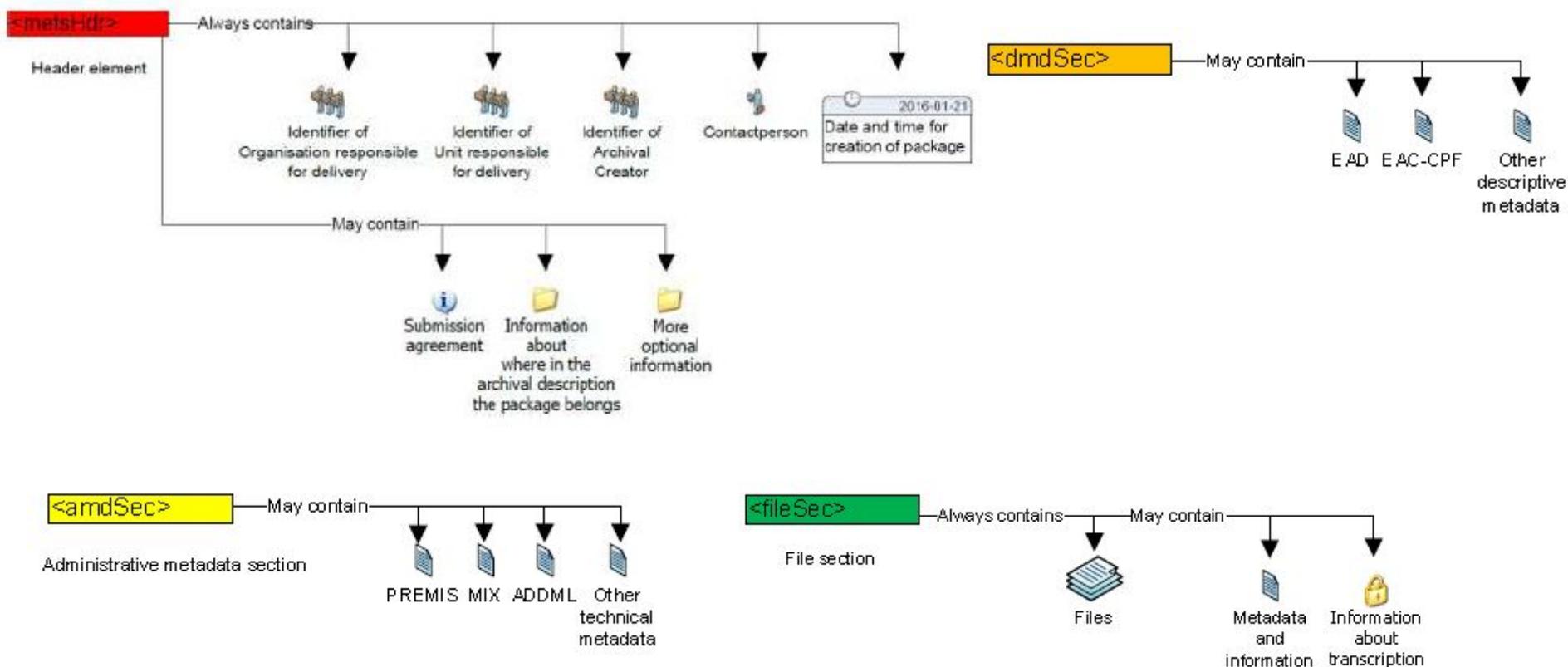
Serie de convenciones para almacenar y transferir una carpeta con archivos nombrados organizados en subdirectorios. La “bag” consiste en un directorio de archivos “payload” acompañados por archivos de metadatos denominados “tag”. Procesar una “bag” no requiere entender el contenido de los archivos y puede ser accesada sin procesar los metadatos de la BagIt.

Estándar E-ARK SIP (corriente europea)

Estructura el paquete de información con base en la presentación desarrollada por el proyecto E-ARK financiado por la UE. Y está basado en METS (Metadata Encoding and Transmission Standard), usa EAD para descripciones ISAD-G, EAC-CPF para describir a los individuos que preservan, PREMIS para soportar los metadatos de preservación y establece las relaciones con otros objetos de información con METS. La identidad del SIP se establece con un identificador UUID o GUID.

el SIP tipo E-ARKs Se basa en METS (Metadata Encoding and Transmission Standard)

Root. OBJID="(UUID o GUID)" | LABEL="Paquete SIP de algo" | TYPE="SIP:Texto"



el SIP tipo BagIt

<base directory>/

```
|
|      BagIt-Version: 1.0
+-- bagit.txt      Tag_File_character_encodin : UTF-8
|
+--manifest-<algorithm>.txt  lista cada uno de los archivos y su checksum para
|                             permitir la validación de la integridad
+-- [additional tag files]
|
+--data/           pueden estar organizados en una estructura de subdirectorios arbitraria
|
|      +--[payload files]
|
+--[tag directories]/
|
|      +-- [tag files]
```

```
myfirstbag/
  manifest-md5.txt
  (49afbd86a1ca9f34b677a3f09655eae9 data/27613-h/images/q172.png)
  (408ad21d50cef31da4df6d9ed81b01a7 data/27613-h/images/q172.txt)
  bagit.txt
  (BagIt-version: 1.0 )
  (Tag-File-Character-Encoding: UTF-8 )
  \--- data/
    27613-h/images/q172.png
    (... image bytes ...)
    27613-h/images/q172.txt
    (... OCR text ...)
    ....
```

RODA-In. Conjuntos de archivos para integrar en un SIP



Un TAG (archivo de metadatos) y un CONTENT (archivo de contenido), forman un SIP

Un TAG y muchos CONTENT en una carpeta

Un TAG y muchos CONTENT en una carpeta con varias subcarpetas

Un conjunto de TAG y CONTENT (estructura de una serie archivística)

RODA-In. Tipos (templates) de conjuntos de metadatos TAG que se pueden asociar a los contenidos CONTENT a integrar en un SIP

EAD 2002

Dublin Core

Clave-Valor

Ejemplos

Un TAG y muchos CONTENT en una carpeta



Un TAG y un CONTENT en una carpeta

ExpRH_786945JuanPerezSanchez

id_7869JPSps.pdf

RH - uuid-1c21a03e-bba8-434d-90be-8040963a25bf.zip

- metadata
- representations
- schemas
- METS.xml

SerieFoto - SerieFotograficaRomuloDias2019 - uuid-ad4b3da8-c729-47aa-8eca-4ce3e74fbd1d.zip

- metadata
- representations
- schemas
- inventory_report - 2019.08.21 15.50.30.802.csv
- METS.xml

SerieFotografica



figura1.JPG



figura2.JPG



figura3.JPG



figura4.JPG



figura5.JPG

Metadatos incluidos en el template EAD 2002



Identity

Description level
Title
Description date
Country code
Material specification
Origination
Physical description
Extent and medium
Dimensions
Appearance
Quote

Context

Administrative and biographical history
Name of creator(s)
Producer
Repository code
Repository
Custodial history
Immediate source of acquisition or transfer

Content and structure

Scope and content
Appraisal, destruction and scheduling
Accruals
System of arrangement

Conditions of access and use

Administrative and biographical history
Conditions governing access
Conditions governing reproduction
Language of material
Language and script notes

Allied materials

Existence and location of originals
Existence and location of copies
Related units of description
Bibliography
Find aid

Notes area

Notes
Description control area
Rules or conventions
Status description
Level of detail
Date of creation or revision
Sources
Archivist notes



Instalación del repositorio.

Instalando RODA en CentOS 7

El ejemplo aquí mostrado se realizó en un host con 1.8 TB de almacenamiento, con 131 GB de memoria y S.O. Centos 7

Prerrequisitos:

- Instalar Java versión 8 (JDK 1.8)

La versión de Java que más convenga, puede ser OpenJDK 8 JDK (para desarrollo) u OpenJDK 8 JRE (para ejecución)

```
sudo yum install java-1.8.0-openjdk-devel (para compilar y ejecutar)
```

o

```
sudo yum install java-1.8.0-openjdk (para ejecutar)
```

- Instalar git

```
sudo yum -y install git
```



Instalación del repositorio. (Continuación)

- Instalar maven, vía wget

```
cd /usr/local/src
```

```
wget
```

```
http://www-us.apache.org/dist/maven/maven-3/3.5.4/binaries/apache-maven-3.5.4-bin.tar.gz
```

- Instalar Apache Tomcat 8 o superior

```
tar -xvf apache-tomcat-8.5.38.tar.gz
```

Teniendo todos los prerequisites anteriores, se procede a la descarga de RODA:

Se sugiere cambiarse al directorio /opt y ahí ejecutar el comando de git:

```
cd /opt
```

```
git clone https://github.com/keeps/roda.git
```



Instalación del repositorio. (Continuación)

Una vez descargado, creó un directorio roda. Hay que ejecutar los siguientes comandos:

```
cd /opt/roda
```

```
mvn clean
```

```
mvn compile
```

```
mvn package -Dmaven.test.skip=true
```

Con lo anterior, se genera un archivo en la ruta `roda-ui/roda-wui/target/roda-wui-VERSION.war` y solo hay que colocar ese archivo en el subdirectorio webapps de Tomcat e iniciar el Tomcat para que se realice el deployment:

```
/opt/apache-tomcat-8.5.38/bin/startup.sh
```



Instalación del repositorio. (Continuación)

Posteriormente en un navegador web abrir la siguiente dirección:

`http://localhost:8080/roda-wui-3.2.0-SNAPSHOT/`

La dirección va a variar, dependiendo de qué versión de RODA se haya descargado.

The screenshot shows a web browser window titled "RODA - WELCOME - Mozilla Firefox". The address bar displays "localhost:8080/roda-wui-3.2.0-SNAPSHOT/#welcome". The page features a blue header with the RODA logo and navigation links for "Welcome", "Login", and "English". The main content area includes a "Welcome to RODA!" heading, a sub-heading "An open-source digital repository designed for preservation", and a paragraph describing RODA's capabilities. Below this, there are two columns of features: "Conforms to open standards" (with a trophy icon) and "Vendor independent" (with a briefcase icon).

RODA - WELCOME - Mozilla Firefox

RODA - WELCOME x +

localhost:8080/roda-wui-3.2.0-SNAPSHOT/#welcome

 **RODA**

Welcome Login English

Welcome to RODA!

An open-source digital repository designed for preservation

RODA is a digital repository solution that delivers functionality for all the main units of the OAIS reference model. RODA is capable of ingesting, managing and providing access to the various types of digital objects produced by large corporations or public bodies. RODA is based on open-source technologies and is supported by existing standards such as the Open Archival Information System (OAIS), Metadata Encoding and Transmission Standard (METS), Encoded Archival Description (EAD), Dublin Core (DC) and PREMIS (Preservation Metadata).

 **Conforms to open standards**

RODA follows open standards using EAD for description metadata, PREMIS for preservation metadata, METS for structural metadata, and several standards for technical metadata (e.g. NISO Z39.87 for digital still images).

 **Vendor independent**

RODA is 100% built on top of open-source technologies. The entire infrastructure required to support RODA is vendor independent. This means that you may use the hardware and Linux distributions that best fit your institutional needs.



¿Preguntas?



iGracias!