

RODA

An open-source digital repository designed for
preservation

Preserving and Archiving Electronically Generated Materials

A joint project of the Leon Levy Foundation and the Archivists Round Table of Metropolitan New York, Inc.

The Morgan Library & Museum, NY

19th November, 2013

What is RODA?

A digital repository designed for **archives** with the following characteristics

Open-source

Long-term **Preservation** and **Authenticity**

Based on **standards**

EAD, PREMIS, OAIS, METS, ...

Secure

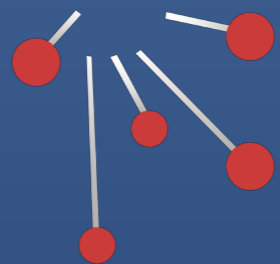
fine-grain permissions, no anonymous usage, LDAP

Scalable architecture (SOA)

Clean web user **interface**

Vendor independent

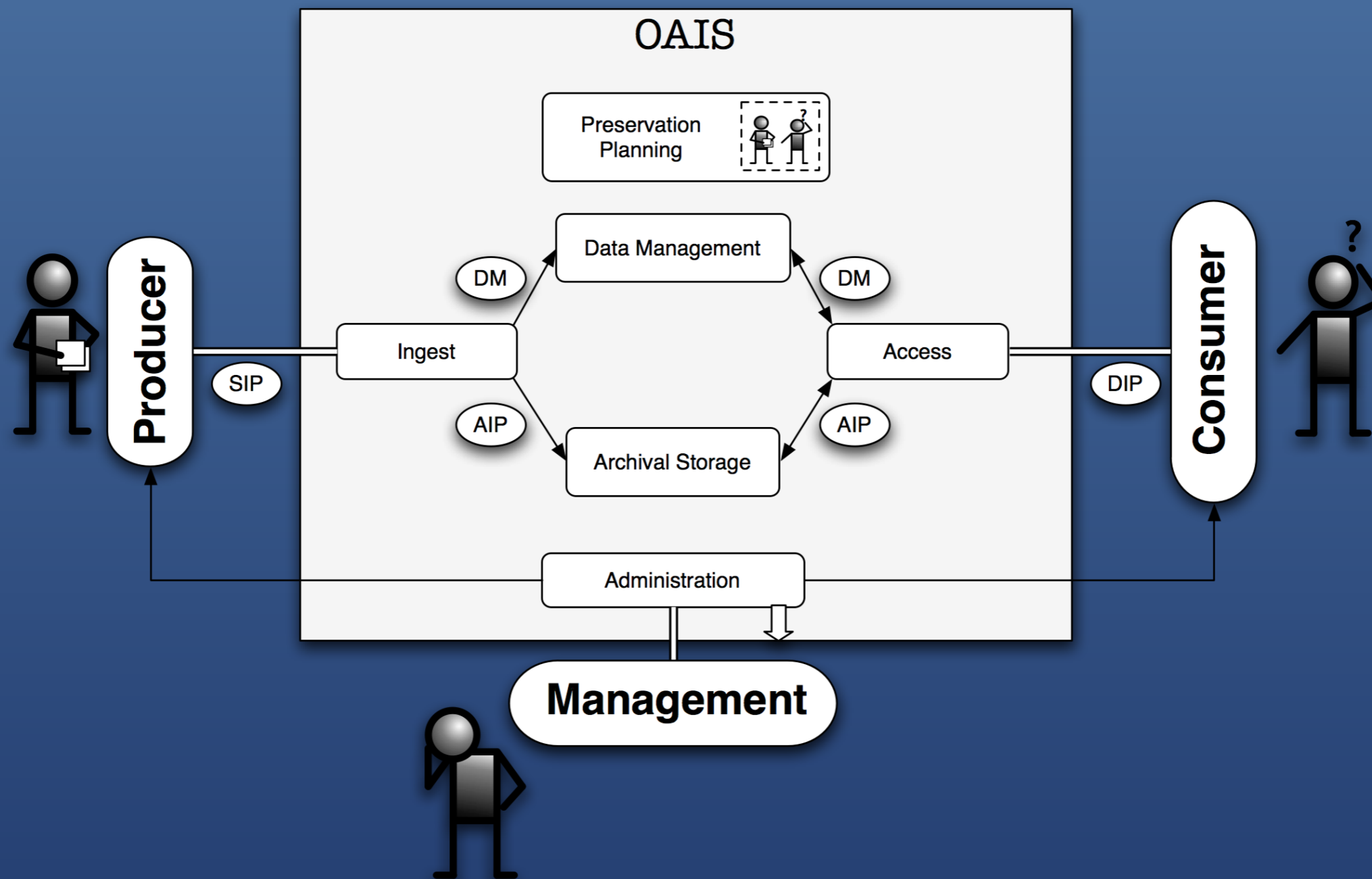
Built on top of



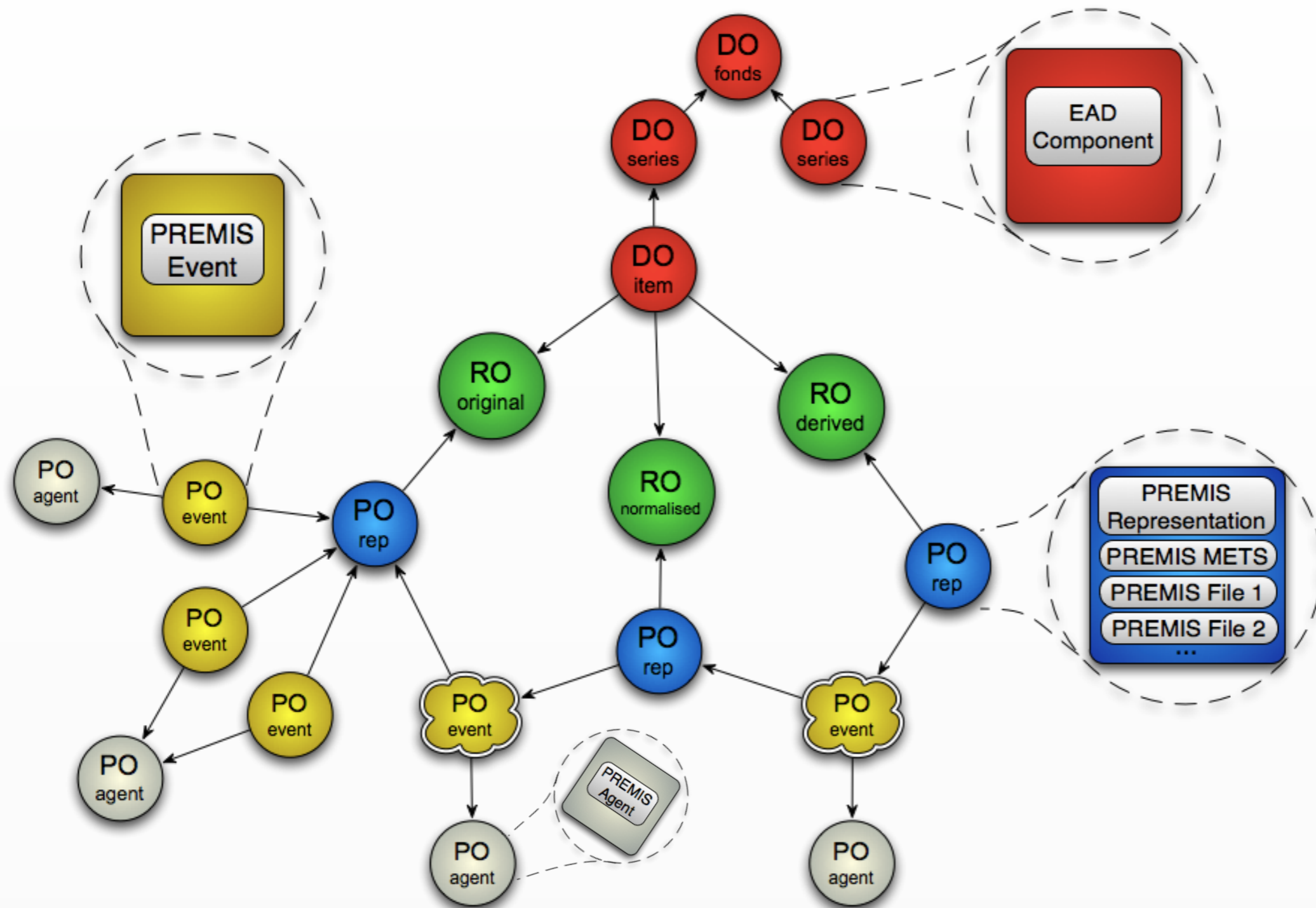
FedoraCommons

OAIS compatible

but... who isn't anyway?



Atomistic data model



- DO Description object (EAD component)
- RO Representation object (files and bitstreams)
- PO Preservation object (PREMIS object entity)
- PO Preservation object (PREMIS event entity)
- PO Preservation object (PREMIS linking event)
- PO Preservation object (PREMIS agent)

Supported media types



Text documents (PDF)



Relational databases (DBML)



Video (MPEG-2)



Audio (Wav)



Digitised works (TIFF unc.)

Supported media types



Text documents (PDF)



Relational databases (DBML)



Video (MPEG-2)



Spreadsheets (CSV, PDF)



Audio (Wav)



Presentations (PDF)



Digitised works (TIFF unc.)



E-mail

All other media types can also be ingested

All other media types can also be ingested
But will only be preserved at the bit level

Default ingest workflow

Step	Name	Description
1	Upload	Upload SIP to the ingest stage area by HTTP, FTP or local copy
2	Unzip	Open ZIP package
3	Virus Check	Check all files for viruses with Clam Anti Virus
4	Syntax and fixity check	Check if metadata is well formed and integrity of all files against checksums in METS envelope
5	Authorization check	Check if user has authorization to ingest into the defined classification plan
6	Ingest	Ingest original into Fedora. This ingestion is not final, can be undone in next steps.
7	Normalization	Check if representation needs normalization, and enforce it as needed
8	Semantic check	Manual semantic check by an archivist, then object is marked active, indexed and published.

RODA ingest tool

The screenshot displays the RODA ingest tool interface, titled "RODA-in 0.0.14". The main header reads "REPOSITÓRIO DE OBJECTOS DIGITAIS AUTÊNTICOS" and "RODA". The interface is divided into several sections:

- Left Panel:** A tree view showing a hierarchy of folders: TEST, LFARIA, sf1, c1, sr1, NG, and NG. Buttons for "Update plan" and "Send" are visible at the top of this panel.
- Center Panel:** A list of image files under the heading "Select and organize images:". The files include:
 - R2009-05-11T13.38.58.46Z
 - ocelot-california-1198172-xl.jpg
 - right-whales-1189052-xl.jpg
 - haw-creek-falls-1190091-xl.jpg
 - mexican-gray-wolf-1198152-xl.jpg
 - gray-fox-dusk-1201638-xl.jpg (highlighted)
 - coconut-island-fish-xl.jpg
 - bee-eater-hungary-118899.jpg
 - moose-mayfield-lake-1191505-xl.jpg
 - atlantic-walrus-MM7468-xl.jpg
 - toad-river-canada-1191509-xl.jpg
 - elephant-seal-pups-1192160-xl.jpg
 - crystal-cave-1191533-xl.jpg
 - chicago-lights-1191505-xl.jpg
- Right Panel:** A preview window showing a close-up of a gray fox yawning. The preview window has tabs for "Preview" and "Metadata".
- Bottom Panel:** A toolbar with buttons for "Create", "Remove", "Save", "Validate", "Create simple document", "Remove simple document", "Create representation", and "Remove representation".

The interface also features a breadcrumb path "TEST/sf1/c1/sr1/NG" and a set of radio buttons for image format selection: "Only TIFFs", "Only JPEGs" (selected), "Only PNGs", "Only BMPs", "Only GIFs", "Only ICOs", "Only XPMs", and "Only TGAs".

STATUS

38 packets PDF CSV

File	Submission date	Current state	%	Producer	
TEST_sf1_c1_sr1_NG.sip	2009-05-11 14:41:47	accepted	100%	Ifaria	
TEST_sf1_c1_sr1_fox.sip	2009-05-07 12:19:36	accepted	100%	Ifaria	
TEST_LFARIA_MP48298.sip	2009-03-20 17:50:39	accepted	100%	Ifaria	
TEST_LFARIA_PNG2.sip	2009-03-20 15:06:14	accepted	100%	Ifaria	
TEST_LFARIA_PNG492.sip	2009-03-20 15:05:28	accepted	100%	Ifaria	
TEST_LFARIA_XPM.sip	2009-03-20 11:40:25	accepted	100%	Ifaria	
TEST_LFARIA_WAV.sip	2009-03-20 11:40:24	accepted	100%	Ifaria	
TEST_LFARIA_TXT.sip	2009-03-20 11:40:16	accepted	100%	Ifaria	
TEST_LFARIA_TIFF.sip	2009-03-20 11:40:15	accepted	100%	Ifaria	
TEST_LFARIA_TGA.sip	2009-03-20 11:40:15	accepted	100%	Ifaria	
TEST_LFARIA_RTF.sip	2009-03-20 11:40:14	accepted	100%	Ifaria	
TEST_LFARIA_PDF.sip	2009-03-20 11:40:13	accepted	100%	Ifaria	
TEST_LFARIA_OGG.sip	2009-03-20 11:40:06	accepted	100%	Ifaria	
TEST_LFARIA_ODT.sip	2009-03-20 11:40:05	accepted	100%	Ifaria	
TEST_LFARIA_MP3.sip	2009-03-20 11:40:01	accepted	100%	Ifaria	
TEST_LFARIA_JPEG.sip	2009-03-20 11:39:28	accepted	100%	Ifaria	
TEST_LFARIA_ICO.sip	2009-03-20 11:39:25	accepted	100%	Ifaria	
TEST_LFARIA_GIF.sip	2009-03-20 11:39:24	accepted	100%	Ifaria	

Filter

- processed
- processing
- accepted**
- rejected
- All

Producer

REPORT

VIEW

ACCEPT

REJECT

File

TEST_sf1_c1_sr1_NG.sip
TEST_sf1_c1_sr1_fox.sip
TEST_LFARIA_MP482.sip
TEST_LFARIA_PNG2.sip
TEST_LFARIA_PNG49.sip
TEST_LFARIA_XPM.sip
TEST_LFARIA_WAV.sip
TEST_LFARIA_TXT.sip
TEST_LFARIA_TIFF.sip
TEST_LFARIA_TGA.sip
TEST_LFARIA_RTF.sip
TEST_LFARIA_PDF.sip
TEST_LFARIA_OGG.sip
TEST_LFARIA_ODT.sip
TEST_LFARIA_MP3.sip
TEST_LFARIA_JPEG.sip
TEST_LFARIA_ICO.sip
TEST_LFARIA_GIF.sip

Ingest report

Original file name: TEST_sf1_c1_sr1_NG.sip

- ▶ received (HTTP)
- ▶ unpacked
- ▶ virus-free
- ▶ well formed
- ▶ authorized
- ▶ ingestd
- ▶ normalized
- ▼ accepted

Date and time: 2009-05-11 15:10:44.000**Task:** Semantic Check**Result:** Success**Details:** SIP aceite pelo utilizador Ifaria: SIP valid for demoSIP objects marked active [ro

CLOSE ✕

Access

ADVANCED SEARCH

Select the level of description you want to search

- all levels
 choose levels

- | | | | |
|---|---|---|--|
| <input type="checkbox"/> F fonds | <input type="checkbox"/> C class | <input type="checkbox"/> SR series | <input type="checkbox"/> DC file |
| <input type="checkbox"/> SF subfonds | <input type="checkbox"/> SC subclass | <input type="checkbox"/> SSR subseries | <input checked="" type="checkbox"/> DS item |

Select the time limits in which to search

- without time restrictions
 specify interval

from to


Fill out the words you want to find

- all the fields
 choose fields

- all words exact phrase at least one of the words without the words

- | | | |
|---|--|--|
| <input type="checkbox"/> Complete reference | <input type="checkbox"/> Quote | <input type="checkbox"/> Physical characteristics and technical requirements |
| <input type="checkbox"/> Title | <input type="checkbox"/> Administrative and biographical history | <input type="checkbox"/> Access restrictions |
| <input type="checkbox"/> Origination | <input type="checkbox"/> Custodial history | <input type="checkbox"/> Reproduction restrictions |
| <input type="checkbox"/> Acquisition number | <input type="checkbox"/> Acquisition information | |
| <input type="checkbox"/> Material specification | | <input type="checkbox"/> Other find aids |
| <input type="checkbox"/> Physical description | <input checked="" type="checkbox"/> Scope and content | <input type="checkbox"/> Related Materials |
| <input type="checkbox"/> Dimensions | <input type="checkbox"/> Organization and ordering | |
| <input type="checkbox"/> Facet or appearance | <input type="checkbox"/> Appraisal | <input type="checkbox"/> Bibliography |
| <input type="checkbox"/> Extent | <input type="checkbox"/> Accruals | |

All the words **"tests"** must appear in one of the fields *Scope and content* of the results.

SEARCH 

ABOUT RODA

SERVICES · POLICY AND TECHNICAL DOCUMENTS · R&D · CONTACTS · ABOUT THE REGISTER · HELP

CATALOG

BROWSE · SEARCH · ADVANCED SEARCH · HELP

INGEST

PRE-INGEST · SUBMIT · STATUS · HELP

ADMINISTRATION

USERS · SCHEDULER · STATISTICS · LOG · HELP

catalog / browse / 50218

BROWSE

4 fonds

- Reference
- AACC
 - DGARQ
 - TEST
 - 692
 - LFARIA
 - sf1
 - c1
 - sr1
 - 101
 - 815
 - 1030
 - 50192
 - 50218
 - fbarbedo

Description Visualization Preservation Permissions

Identification

Reference	50218
Complete reference	PT/DGARQ/TEST/sf1/c1/sr1/50218
Handle	http://hdl.handle.net/10384/50218
Title	National Geographic Photos
Level	item
Initial date	2008-01-01
Final date	2009-01-01
Country code	PT
Repository code	DGARQ
Origination	National Geographic
Acquisition number	NG
Acquisition date	2009-05-11

Content and structure

Scope and content	tests for or09
-------------------	----------------

Notes

Note	Pacote de submissão validado semanticamente por Ifaria
------	--

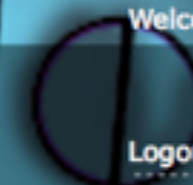
EDIT

SUBLEVEL

MOVE

EAD

CLOSE



ABOUT RODA

SERVICES · POLICY AND TECHNICAL DOCUMENTS · R&D · CONTACTS · ABOUT THE REGISTER · HELP

CATALOG

BROWSE · SEARCH · ADVANCED SEARCH · HELP

INGEST

PRE-INGEST · SUBMIT · STATUS · HELP

ADMINISTRATION

USERS · SCHEDULER · STATISTICS · LOG · HELP

catalog / browse / 50218

BROWSE

4 fonds

- Reference
 - AACC
 - DGARQ
 - TEST
 - 692
 - LFARIA
 - sf1
 - c1
 - sr1
 - 101
 - 815
 - 1030
 - 50192
 - 50218
 - fbarbedo

Description Visualization Preservation Permissions

Disseminassions of 50218 - "National Geographic Photos"



JPEG
(original)



TIFF
(normalized)

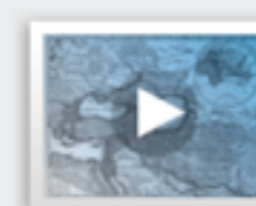


Photo pre-viewer



Book pre-viewer

CLOSE



NATIONAL GEOGRAPHIC

Photograph by Matthias Klum

© COPYRIGHT NATIONAL GEOGRAPHIC SOCIETY. ALL RIGHTS RESERVED.

National Geographic Photos



(5/13) gray-fox-dusk-1201638-xl.jpg.tiff

SIMPLEVIEWER

ABOUT RODA

SERVICES · POLICY AND TECHNICAL DOCUMENTS · R&D · CONTACTS · ABOUT THE REGISTER · HELP

CATALOG

BROWSE · SEARCH · ADVANCED SEARCH · HELP

INGEST

PRE-INGEST · SUBMIT · STATUS · HELP

ADMINISTRATION

USERS · SCHEDULER · STATISTICS · LOG · HELP

catalog / browse / 50218

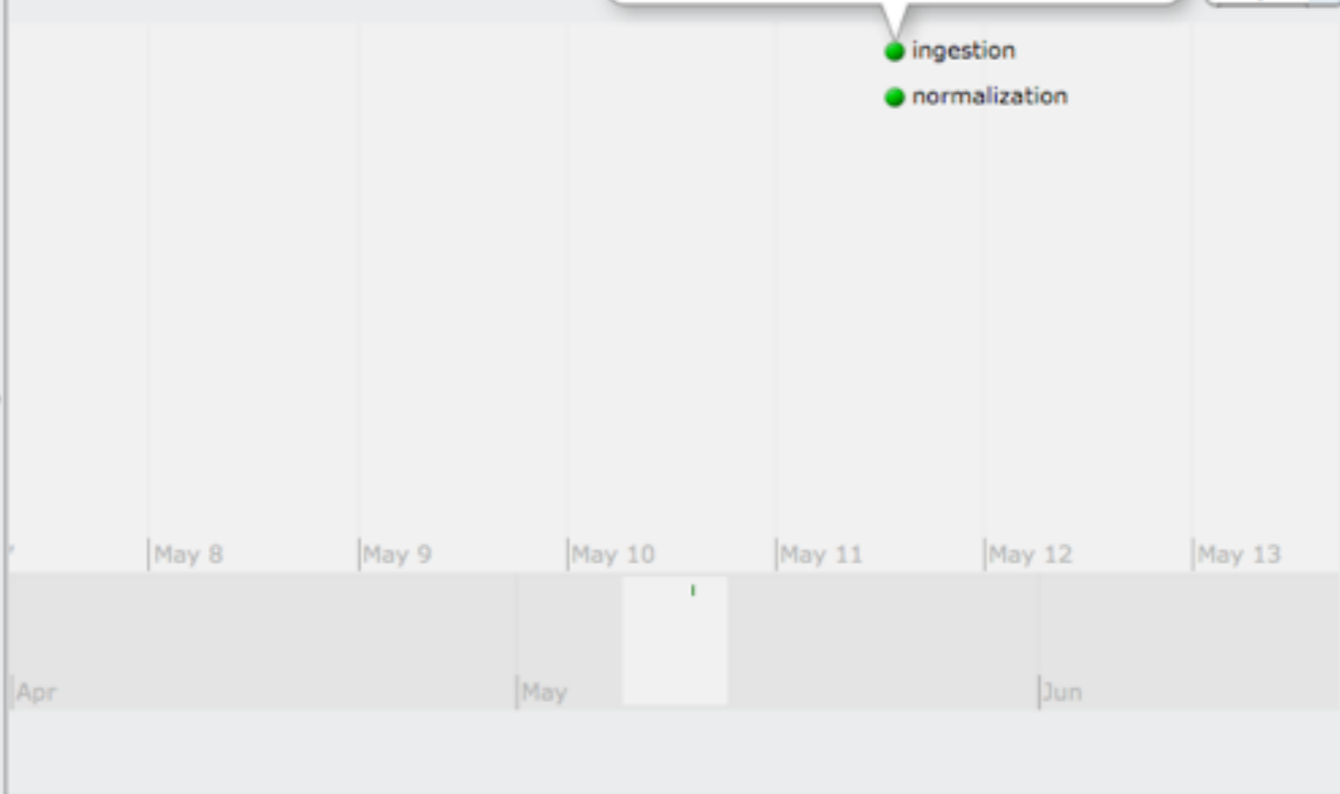
BROWSE

4 fonds

- Reference
 - AACC
 - DGARQ
 - TEST
 - 692
 - LFARIA
 - sf1
 - c1
 - sr1
 - 101
 - 815
 - 1030
 - 50192
 - 50218
 - fbarbedo

Description Visualization Preservation Permis

roda:50219 (original) roda:50222 (normalized)



ingestion

Description: The objects inside the SIP were added to the preservation repository

Result: success

Ingest details:

1 SIP file inside.com success.com

PREMIS

CLOSE

Lets talk about **preservation**, shall we?

Normalisation on ingest

Configurable normalisation rules

Automatic generation of preservation metadata

PREMIS events, representation information, agents

Preservation plans can be executed on the demand

New preservation components are easily installed as add-ons

Original representations are kept unchanged

Enables future approaches, e.g. emulation on-demand

Fixity checks

Notifications are sent when bit rot is detected

Management features

Advanced user management

administration / users / list of users

LIST OF USERS

ALL A B C D E F G H I J K L M N O P Q R S T U V W X Y Z		19 user(s)
	admin Administrador RODA	
	arodrigues Ana Maria do Rosário Silva Rodrigues	
	chenriques Maria Cecília de Jesus Henriques	
	fbarbedo francisco barbedo	
	guest Utilizador anónimo	
	jmaferreira José Ferreira	
	lcorujo Luis Miguel Nunes Corujo	
	lfaria Luís Francisco da Cunha Cardoso de Faria	
	guests	
	administrators	
	mariosantana Mário Henrique Marçal Sant'Ana	
	mferreira José Miguel Araújo Ferreira	
	mfrias Miguel Frias	
	qwert qwertyuiop	
	rcastro Rui José Gonçalves de Castro	
	roda-disseminator RODA	
	roda-handle RODA Handle	
	roda-ingest-task RODA	
	roda-preservation-task RODA	
	roda-wui RODA	
	test001 Test	

List

[Users](#)

[Groups](#)

Search

Users

REPORT

NEW

EDIT

INACTIVATE

REMOVE

Groups

NEW

EDIT

REMOVE

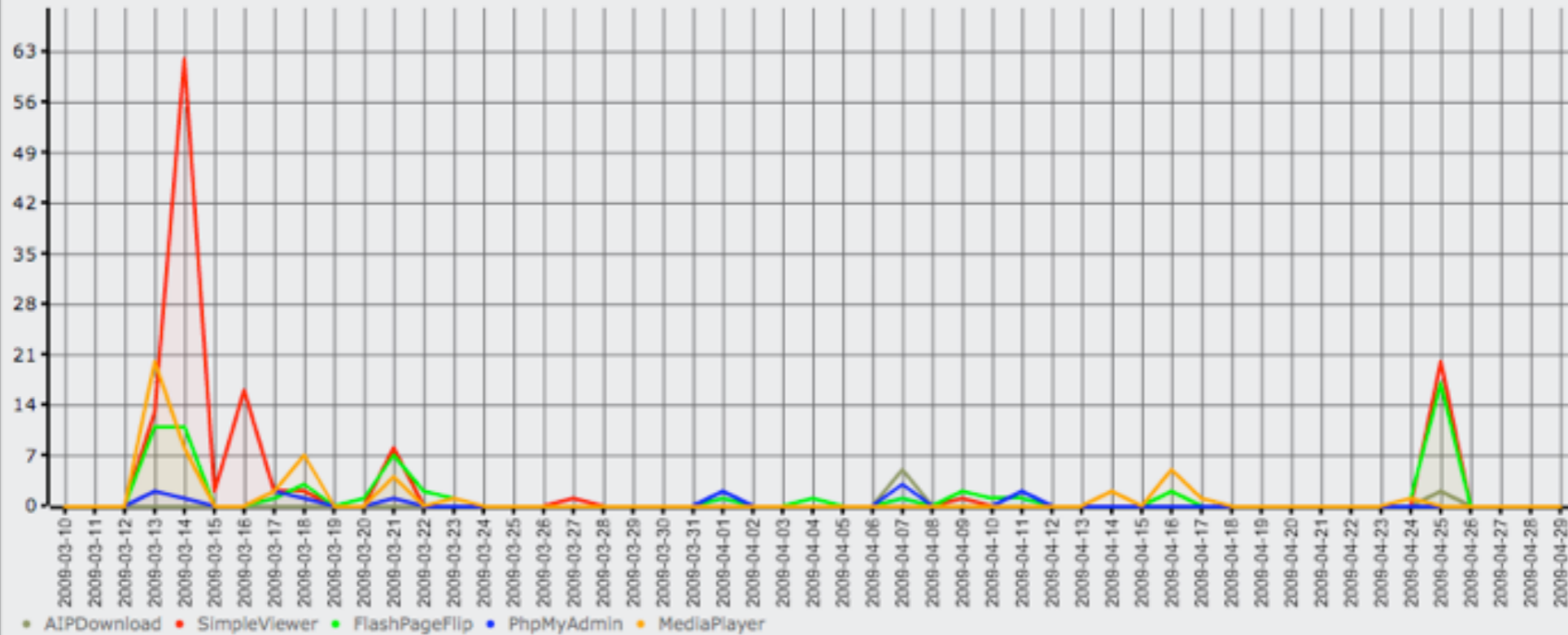
Reporting & statistics

Distribution of disseminators accessed

Graph Table

Unit of time: years months days

2009-03-10 at 2009-04-28



CLOSE

confirmed.

Total No. of confirmed email addresses

Count of confirmed email addresses. After registration, an email is sent to the email address submitted. Within this email a confirmation code will be used



14

10

LOG

Comprehensive logging system auditing purposes

Action: Begin Date: January End Date: January

327210 record entries  

Date and Time	Action	Parameters	User
2009-05-12 12:07:35.0	RODAWUI pageHit	hostname: 193.136.19.138 address: 193.136.19.138 user: beta pagename: administration.log	lfaria
2009-05-12 12:05:10.0	Browser getProducers	doPID: roda:2	lfaria
2009-05-12 12:05:10.0	UserBrowser getUser	name: mferreira	lfaria
2009-05-12 12:05:10.0	UserBrowser getGroup	groupName: administrators	lfaria
2009-05-12 12:05:10.0	UserBrowser getGroup	groupName: producers	lfaria
2009-05-12 12:05:07.0	UserBrowser getUser	name: admin	lfaria
2009-05-12 12:05:07.0	UserBrowser getUser	name: fbarbedo	lfaria
2009-05-12 12:05:07.0	UserBrowser getUser	name: lfaria	lfaria
2009-05-12 12:05:07.0	UserBrowser getUser	name: mferreira	lfaria
2009-05-12 12:05:07.0	UserBrowser getGroup	groupName: administrators	lfaria
2009-05-12 12:05:06.0	Browser getDescriptionObject	pid: roda:2	lfaria
2009-05-12 12:05:06.0	Browser getRODAObjectUserPermissions	pid: roda:2	lfaria
2009-05-12 12:05:06.0	Browser getRODAObjectPermissions	pid: roda:2	lfaria
2009-05-12 12:05:05.0	RODAWUI pageHit	hostname: 193.136.19.138 address: 193.136.19.138 user: beta pagename: dissemination.browse.2	lfaria
2009-05-12 12:05:05.0	Logger addLogEntry	logEntry: LogEntry (datetime=null, address=192.168.111.154, username=lfaria, action=RODAWUI pageHit, parameters=	roda-wui

Take up so far...

Running at

National Archives of Portugal

National Archives of Brazil (will go live soon)

Organisations that have requested the VM

University of Michigan, USA

National Library of Wales, UK

The Bavarian State Archives, Germany

King's College London, UK

University of Dundee, UK

Casa de las Americas, Cuba

Marine Biological Laboratory, USA

ELDA: Evaluations and Language resources Distribution Agency

Unicamp, Brasil

National Library of Australia

University of St Andrews, UK

Community support



RODA

The world's most advanced **open-source** digital repository



An open-source **digital repository** designed for preservation

RODA is a complete digital repository 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 content produced by large corporations or public bodies. RODA is based on open-source technologies and is supported by existing standards such as the OAIS, METS, EAD and PREMIS.



Conforms to open standards

RODA follows open standards using EAD for description metadata, PREMIS for preservation metadata, METS for structural metadata, 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.



Scalable

The service-oriented nature of RODA's architecture allows the system to be highly



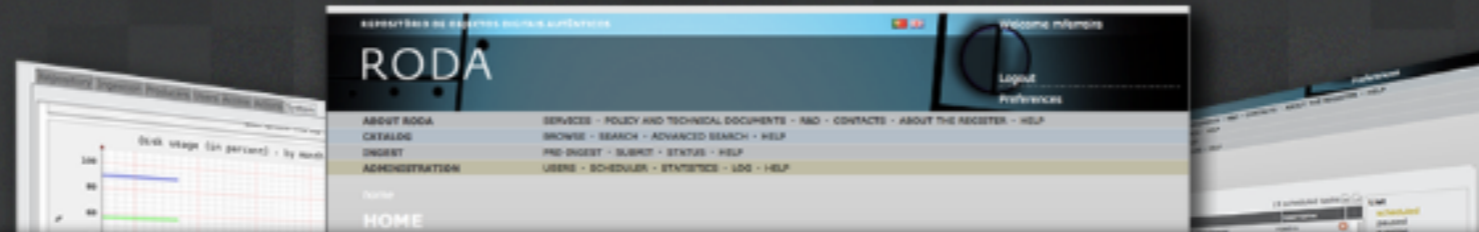
Embedded preservation actions

Preservation actions and management within RODA is handled by a task scheduler. The task



RODA

The world's most advanced **open-source** digital repository



www.roda-community.org

An open-source **digital repository** designed for preservation

RODA is a complete digital repository 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 content produced by large corporations or public bodies. RODA is based on open-source technologies and is supported by existing standards such as the OAIS, METS, EAD and PREMIS.



Conforms to open standards

RODA follows open standards using EAD for description metadata, PREMIS for preservation metadata, METS for structural metadata, 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.



Scalable

The service-oriented nature of RODA's architecture allows the system to be highly



Embedded preservation actions

Preservation actions and management within RODA is handled by a task scheduler. The task

Know more about the project and product

Download RODA

Source-code, binary, VM

Online demo

demo.roda-community.org

Find help

Documentation, developers guide, articles, **mailing lists**

Find vendors for paid support

KEEP SOLUTIONS

University of Minho SPIN-OFF

Developing the future today

Miguel Ferreira
mferreira@keep.pt

KEEP SOLUTIONS, LDA
<http://www.keep.pt>