RODA: A Service-Oriented Digital Repository

Database Archiving

José Carlos Ramalho
jcr@keep.pt

José Carlos Ramalho
jcr@di.uminho.pt

Dep. of Informatics
University of Minho
**Context**

**RODA (2006-2009)**
- Metadata management (EAD based)
- Digital object management (...)
- Digital preservation protocols and policies
- National project with public funds (Portuguese National Archives)

**CRiB: Preservation Services for Digital Repositories (2005-2008)**
- Distributed Migration Service
- Migration Adviser Service
- Phd thesis at U.Minho (Miguel Ferreira)
RODA (2006-2009)
• Metadata management (EAD based)
• Digital object management (...)
• Digital preservation protocols and policies
• National project with public funds (Portuguese National Archives)

CRiB: Preservation Services for Digital Repositories (2005-2008)
• Distributed Migration Service
• Migration Adviser Service
• Phd thesis at U.Minho (Miguel Ferreira)
Premisses

• RODA:
  ✴ OAIS as reference model
  ✴ EAD as descriptive metadata
  ✴ Open Source
  ✴ SOA
  ✴ Preservation Policy: Migration
  ✴ Authenticity

• CRiB:
  ✴ Distributed Migration
  ✴ SOA
  ✴ Quality control
  ✴ Migration Adviser
Authenticity

“O Codex 632” by José Rodrigues dos Santos

Subject: Who really was Cristophoros Colombus?

Was he italian? Spanish? Or a portuguese belonging to a jewish family?
We must trust our sources: in ancient History there are no direct speech or evidence.

EX: the bible
We must trust our sources: in ancient History there are no direct speech or evidence. 
EX: the bible
We must trust our sources: in ancient History there are no direct speech or evidence.

EX: the bible

How do we become trustful?

• Reputation
• Documenting every action taken upon DOs
• Sometimes the law will be an obstacle
DO CLASSES: NORMALIZATION

terça-feira, 7 de Fevereiro de 2012
Open Archival Information System

ISO 14721

terça-feira, 7 de Fevereiro de 2012
OAIS (Information Packages)

- Submission Information Package (SIP)
  - Digital Object
  - Metadata created by producer
    - too open...
- Archival Information Package (AIP)
  - Digital Object to be stored
  - Metadata: enough to ensure DO’s preservation and access
    - model defined by PREMIS
- Dissemination Information Package (DIP)
  - DO transformed into the format that will be delivered to the consumer
  - Metadata
Ingestion

Producer

Negotiation

Archive

Submission Contract
Ingestion

Submission Contract
- SIP specification
- Ingestion workflow specification
SIP Structure (example)

one still image
SIP Structure (example)

one still image

creation properties:
- date
- hardware
- ...
SIP Structure (example)

one still image

creation properties:
- date
- hardware
- ...

Technical Metadata:
- color
- dimensions
- ...

terça-feira, 7 de Fevereiro de 2012
SIP Structure (example)

one still image

tool properties:
  - date
  - hardware
  - ...

Technical Metadata:
  - color
  - dimensions
  - ...

Descriptive Metadata:
  - producer
  - collection
  - ...

terça-feira, 7 de Fevereiro de 2012
SIP Structure (Example)

One still image

Creation properties:
- date
- hardware
- ...

Technical Metadata:
- color
- dimensions
- ...

Descriptive Metadata:
- producer
- collection
- ...

Manifest
SIP Structure (Example)

One still image

Technical Metadata:
- color
- dimensions
- ...

Descriptive Metadata:
- producer
- collection
- ...

Creation properties:
- date
- hardware
- ...

Manifest

Compressed File
SIP Structure (+complex)
SIP Structure (+complex)
SIP Structure (+complex)

DO = Image+
SIP Structure (+complex)

DO = Image+ Properties
SIP Structure (+ Complex)

DO = Image+  Properties  Technical Metadata
SIP Structure (+complex)

DO = Image+

Properties

Technical Metadata

Descriptive Metadata
SIP Structure (+complex)

DO = Image+
Properties
Technical Metadata
Descriptive Metadata

Manifest
Ingestion Workflow

SIP ➔ AIP

• integrity check
• virus check
• generation of preservation metadata (PREMIS)
• conversion to a normalized format
• generation of technical metadata
• generation of preservation metadata (PREMIS)
Roda Data Model
Roda Data Model

Description Objects