La aplicación de herramientas para la generación de paquetes de información OAIS en el contexto europeo. La experiencia de implementación de EARK2 en el Archivo Nacional de Hungría.

Zoltán Szatucsek

Sevilla, 19th February, 2020
A vision

• Resolve the problem of Digital Dark Age
• Renew the regulation on records management
• Fit standards
• State of the art IT infrastructure
• To be financially sustainable,
• To support education
And the result
Tandem

[Logos for Tandem, Scope, and Tessella]
Modernization

200+ PC scanners
storage
data center
experts
partners
knowledge
Local data center
Integration with regional archives since 2012
Special thanks to
Standards

OAIS
PAIMAS
METS
OAI-PMH
Moreq
ISADg – EAD
ISAAR(CPF) – EAC(CPF)
ISDIAH - EAG
Archives Portal Europe
Components

Label

Users Manual

[Diagram showing components labeled as 'Components', 'Label', and 'Users Manual' being directed towards a box]
Structure

- SIP_20131201_KSH_001.txt
- SIP_20131201_KSH_001.zip
  - Header
    - Metadata.xml
  - Content
    - Data
      - 001.pdf
      - 002.pdf
    - Documentation
      - Detailed description.doc
      - Transfer agreement
Specifications

- 34/2016. (XI. 30.) EMMI
- vs. EARK specs
  - Provides core principles for any IP
    - Most crucial: a common structure
    - Influenced by the Swedish CS
  - Aims to support automation of high-level package identification and validation (integrity, technical validity, etc.)
  - Is built on widely used XML standards (METS, PREMIS)
    - Concentrates on structural metadata, includes elements of administrative and technical metadata
  - Is flexible enough to be used for any type of data
  - Allows for further specification and localisation if needed
Steps in regulations

SIP definition 2013

SIP transfer mandatory 2016

ERMS SIP export mandatory 2018
Alternatives for SIP creation (HU)

- **ERMS automatic export**: different by applications
- **Individual codes**: Python scripts
- **SIP Creator applications**: Preservica XIP Creator, ELEV SIP Creator, Roda In
eArchiving
Store and preserve digital information cost-efficiently over the long term
Database Preservation

Relational databases are one of the most important technologies supporting today's information management activities. They are designed to store, organize and explore digital records that not only support but also document day-to-day business operations. Very often, these records are irreplaceable or prohibitively expensive to reacquire by other means rendering the preservation of databases a serious concern.

This page focuses on workflows, tools and standards to allow information managers to extract, archive and preserve records of information currently managed by relational databases.

The most relevant initiatives in this context are the Database Preservation Toolkit, the Database Visualization Toolkit and the SIARD 2.0 preservation format.

The following screencast aims to illustrate how all these tools fit together in a full-cycle archiving and preservation workflow for relational databases.
Pre Ingest process workflow
1. EXPLORADOR DE ARCHIVOS

2. PAQUETES DE INFORMACIÓN

3. INSPECTOR

4. PASSO FINAL

Carga tu esquema de clasificación

o crear uno Nuevo Esquema de Clasificación

Selecciona un elemento del paquetes de información para inspeccionarlo

Cargar
Who is using it?

- Record creators
- Contractors (digitization)
- Archivists (irregular transfers)
- Reprography (digitization)
Conclusion: Digital Preservation is not an IT challenge
Fasten your seat belt!
2nd digital revolution is here

financing
new task for archives
management
Artificial Intelligence