The Digital Repository of the Swiss Federal Archives

International Conference on Digital Library Management
Kolkata, 11-13 January 2011 – Jean-Marc Comment
Program

- Introduction
- Digital Repository
- Standards
- Pilot project
- Open issues
The Swiss Federal Archives

- Is mandate by the federal law on archiving to store the documents of the Swiss administration.
- Has a Digital Repository and the trained staff to manage it.
- Accept documents from records management systems, databases and collections of files.
- The government departments must deliver digital documents in data packages compliant with the structure and formats we have defined.
- Introduction
- Digital Repository
- Standards
- Pilot project
- Open issues
Base - OAIS reference model

SIP = Submission Information Package
AIP = Archival Information Package
DIP = Dissemination Information Package
Application Landscape

 Archive

 Finding Aids

 Data to support search

 Access to archived data

 Metadata on archived data, provenance, order

 AIP archival goods

 Infrastructure Digital Archive

 Digital Information Repository

 Archival Information System

 Queries

 Government Units

 Archival Interfaces

 Submission

 SIP

 Transfer

 SIP

 DIP

 Dissemination

 Customers

 Access

 DIP
IT Landscape

Digital Information Repository (DIR)

Archiv Informations System (AIS)

Ingest
Preservation
Access
Package Workflow Control

Tools Integration
Data Management
Reporting
Administration

Storage Adaption

Package Handler
SIARD
Data Conversion
etc.

Submission Workbench

Dissemination Workbench

Application
Infrastructure (HW&SW)

NAS
Tape library
Tape library
Adic StoreNext
HSM-Software

DIP

IT Landscape

Federal Department of Home Affairs
Swiss Federal Archives
Unit Technology for Information

International Conference on Digital Library Management, Kolkata
© J.-M. Comment, 11-13 January 2011
Archiv Information System: scopeArchiv
DIR - Network and Security

secured network

PRONOM
Format Registry

http

DIR

Storage Adaption

https

AIS

xml

PC clients

HSM

NAS

Tape Library
Library

International Conference on Digital Library Management, Kolkata
© J.-M. Comment, 11-13 January 2011
Digital Information Repository (Tessella)
IT Infrastructure

Network Layout, Servers and Storage Units
DIR & Package Handler

Digital Information Repository (DIR)

Ingest  Preservation  Access

Package Workflow Control

Tools Integration
PackageHandler
Antivirus
DROID
Jhove
...

Data Management  Reporting  Administration

Storage Adaption
Package Handler

- **Package Handler – Release 2008:**
  Tool for archivists at the SFA:
  Quality check SIPs, viewing und editing (Ingest)

- **Package Handler – future development 2011:**
  Tool for agencies:
  Quality check SIPs, preparation of digital submissions
  
  Tool for users (agencies, general public):
  Viewing and working with DIPs (Access)
Ingest / Access: Package Handler
• Introduction
• Digital Repository
• Standards
• Pilot project
• Open issues
Base – SIP

- Digital files from Records Management Systems
- Data from relational databases
- Files collections

Metadata

Primary data
Logical View and Package Metadata

Registry
(Dossiers and Documents)

View of SIP Metadata

```xml
<paket>
  <inhaltsverzeichnis>
    <ordner>
      <datei id=…>
        ...
      </datei>
    </ordner>
  </inhaltsverzeichnis>
  <ablieferung>
    <unstrukturierterAnhang/>
    <ordnungssystem>
      <ordnungssystemposition>
        <dossier>
          <dokument>
            <dateiRef/> <dateiRef/>
          </dokument>
        </dossier>
        ...
      </ordnungssystemposition>
    </ordnungssystem>
  </ablieferung>
</paket>
```
Package Metadata and Physical Structure

View of SIP Metadata

View of SIP structure (Folders and Files)

SIP/
header/
metadata.xml
old/
unstructured/
content/
d000001/
p000001.tif
p000002.tif
d000002/
d000001/
p000003.txt
p000004.pdf
Versioning of Packages

SIP

AIP (= version 1)

AIP' (= version 2)

Existing & new metadata

Existing & modified metadata

1:1 copy

Format conversion

Local AIP ID
From AIP to DIP
E-Access
Data Formats suitable for digital archiving

- SFA defined the following data formats as suitable for electronic archival:
  - Text data (unstructured): TEXT
  - Printable documents: PDF/A
  - Data in tabular form: CSV
  - Relational databases: SIARD
  - Halftone images: TIFF
  - Audio data: WAVE
  - Video data: MPEG/4
SIARD Suite 1.0
Introduction

Digital Repository

Standards

Pilot project

Open issues
Pilot project

- In the past, SFA has run two pilot projects:
  - DAp: Database Archival Process
  - GAp: GEVER Archival Process
- SFA has started a new preliminary study with the Federal Office of Topography swisstopo about geographical data (geodata) coming from Geographical Information System (GIS) to examine the robustness of the SIP specification and to define a file format for geodata.
Pilot Project Geodata

- Was based on the principles of digital archiving used by the Swiss Federal Archives (SFA), and the official geodata of the federal government.
- The study also took account of the international context.
- It wasascertained that no solution for the archiving of geodata currently exists.
- The aim was to define criteria not only for the development of an integral solution for the entire federal administration, but also for securing long-term availability.
- In addition, the intention was to secure nationwide co-ordination of all aspects of geoinformation, including the archiving of geodata.
Pilot Project Geodata

- By studying a variety of models it was found that the existing digital archiving solution in use at the SFA could (and should) serve as the basis for the archiving of geodata.
- The option of adapting the SIARD model is found to be inadequate for the declared purpose.
- A first prototype of a method of delivering geodata to the archive was developed.
- The preliminary study put forward proposals for formats that would be suitable for archiving (for both vector and raster data).
Introduction

Digital Repository

Standards

Pilot project

Open issues
Interfaces with the federal agencies

Agency → Building a SIP → Transfer of the SIP → SFA
Preparation of the SIP in the agency
Life cycle: from technique to business process
Points of intervention

- Filing plan / Definition of the processes
- Metadata / Retention period
- Prospective appraisal
- Prescriptions of organisation
- Proof of completeness and conversion
- Disposal / SIP-Building
- Transfer to SFA

Archival value
Seamless and streamlined flow of records
New standards for file formats

- With the apparition of new formats like Open Office XML (OOXML) and Open Document Format (ODF), we have now production format which can be also archival format.
- So if we choose these formats as archival format, we don’t need to convert the produced documents.
- The SFA has made in 2010 a preliminary study with the format XLSX. The first results of this study are very positive.
Links

- Archival formats:

- Digital Archiving Policy:

- Specification SIP:

- SIARD Suite Software:
Thank you for your attention!