Digital Preservation System

ExLibris Rosetta

OAI6 | Geneva | June 2009

Dr. Axel Kaschte, Strategy Director Europe
Copyright Statement

All of the information and material inclusive of text, images, logos, product names is either the property of, or used with permission by Ex Libris Ltd. The information may not be distributed, modified, displayed, reproduced – in whole or in part – without the prior written permission of Ex Libris Ltd.

TRADEMARKS
Ex Libris, the Ex Libris logo, Aleph, SFX, SFXIT, MetaLib, DigiTool, Verde, Primo, Voyager, MetaSearch, MetaIndex and other Ex Libris products and services referenced herein are trademarks of Ex Libris, and may be registered in certain jurisdictions. All other product names, company names, marks and logos referenced may be trademarks of their respective owners.

DISCLAIMER
The information contained in this document is compiled from various sources and provided on an "AS IS" basis for general information purposes only without any representations, conditions or warranties whether express or implied, including any implied warranties of satisfactory quality, completeness, accuracy or fitness for a particular purpose.

Ex Libris, its subsidiaries and related corporations ("Ex Libris Group") disclaim any and all liability for all use of this information, including losses, damages, claims or expenses any person may incur as a result of the use of this information, even if advised of the possibility of such loss or damage.

© Ex Libris Ltd., 2009
Ex Libris – a library automation company
Rosetta heritage

2003 - National Library Act of New Zealand
Enable the National Library to meet its mandate to collect, make accessible, and preserve in perpetuity, New Zealand’s digital heritage.

2004 – New Zealand NDHA founded
2005 – Requirements according OAIS
2006 – A digital repository is not enough!
International peer review group

New Haven
Ithaca
London
Glasgow
Helsinki
Beijing
The Hague
Jerusalem
Los Angeles
Singapore
Wellington
True preservation is needed

- Deterioration of the storage medium
- Obsolescence of the storage medium
- Obsolescence of the data format
- Obsolescence of the software
- Obsolescence of the hardware required to run the software

“The goal of digital preservation is the accurate rendering of authenticated digital content over time.” (ALA)
ExLibris Rosetta

- Version 1.0 released January 2009
- New digital library software package
  - Based on preservation needs
  - Takes repository solutions one step further
What makes it “preservation“?

Figure 1
The Preservation Pyramid.

based on an article by Pricella Caplan (Premis)
Availability (capture/selection)

Definition:

It is truism that you can not preserve digital objects that you do not control.

What is supported in Rosetta:

• Deposit module
• Software Development Kit
• Submission applications
**Identity (description)**

**Definition:**

The digital object must be described in sufficient detail to allow future access and/or use.

**What is supported in Rosetta:**

- Supporting persistent identifiers (internal and external)
- Storing descriptive information on the AIP:
  - Using the out of the box DC editor
  - Integrating with descriptive tools
Understandability (documentation)

**Definition:**
The repository is responsible for providing and preserving enough information to ensure future users to understand the preserved objects.

**What is supported in Rosetta:**
- Rosetta stores descriptive and preservation metadata
- The data model is based on a standard model - PREMIS (IE, Rep, File) and METS file.
- The elements are standard - Rosetta supports the PREMIS elements as preservation metadata
Fixity (secure storage)

Definition:
Protect digital objects from unauthorized changes, whether deliberate or in-deliberate.

What is supported in Rosetta:
- On going fixity checks
- Fixity stored for both the metadata and the files.
- Support for a write once policy – any change is a new version.
- Fixity validation upon receiving.
Viability (media management)

**Definition:**
Viability is the quality of being readable from media.

**What is supported in Rosetta:**
- Ongoing fixity and virus checks on the files stored on the permanent repository.
- Abstraction of the application and the storage layer.
- All information stored on disk, which is proven technology.
Definition:
Ensuring that a digital file is renderable (displayable, playable...) may be the heart of the digital preservation process.

What is supported in Rosetta:
- Characterization.
- Preservation planning module.
- Support for different preservation strategies.
Definition:

Authenticity means that the integrity of both the source and the content of the object can be verified.

What is supported in Rosetta:

- Producer management.
- Auditing.
- Human assessment is possible.
- Write once repository.
ExLibris Rosetta system characteristics

- Fully OAIS compatible
- Scalable
- Secure
- Flexible and extensible
OAIS and Rosetta

OAIS

- Ingest
- Archival Storage
- Data Management
- Access
- Preservation Planning

Rosetta

- Deposit
- Working Area
- Permanent
- Operational Management
- Publishing
- Delivery
- Search Tools
- Preservation
Scalable

- Manages tens of millions of objects, petabytes of storage
- Supports ingest of and access to tens of thousands of objects per day
- Scalable both vertically and horizontally
- Deployed in a distributed manner
- Has no single point of failure; all software components are redundant
Secure

- Manages audit trails:
  - Every event is audited
  - Events can be logged, reported on, or stored as provenance
  - Audit information is configurable
- Performs routine checks to ensure system integrity e.g., virus and fixity checks
Flexible and extensible (1)

- Embedding of third-party tools:
  - Validations tools, e.g., JHOVE, DROID, NLNZ extractor tool, antivirus, checksum.
  - Enrichment tools
  - Delivery conversion tools
- Embedding of pre-transforms
- SDK to allow external Submission Applications
Flexible and extensible (2)

- Embedding of on the fly conversion tools for DIP creation
- Enabling the customization of processes through configurable rules and workflows
- Forming an open architecture through API and Web services
- Integrating with existing library applications
Working Preservation Management

Manual Automatic process process

Deposit SIP DIP

Transformers SDK

Search tools Primo

Publishing Delivery

Conversers SDK Viewers OAI/SRU

Risk identifiers Pres. Tools

Validation Format enrichment Metadata enrichment

Working area

Operational

Permanent Repository
Ingest: Receive Submission (deposit module)

Use web interface to load info

Use FTP to load info

Use submission applications

Integrate with existing applications
Metadata Enrichment

1. Automatic extraction of Tech MD
2. Edit MD manually
3. Automatic import of External MD
Access Viewers

Descriptive Metadata - IF DC

creator: Yaniv Levi
title: State archive
description: bla bla
type: Photographic Archive

Representation (REP68388) > Table of Contents > Koloriertes Titelblatt

© Ex Libris Ltd., 2009
Thank you

Axel.Kaschte@exlibrisgroup.com