Digital Building Blocks: Rosetta’s Plug-and-Play Ability

Euwe Ermita
Manager, Digital Library Systems and Services
State Library of New South Wales
1 million microforms
500,000 bound serials
65,000 newspaper volumes

2 million books
13 kms manuscripts

13k hours Oral History
10k posters
130k ephemera

26k realia, coins, tokens, medals

1.5 million photographs
0.5 million prints, drawings, paintings and maps
Digital Excellence Program:

$72 million over 10 years

1. Digitisation Program
   - $22 million over first 4 years - 62 million over 10 years
   - Fast tracking digitisation program

2. Digital Infrastructure Program
   - Renew ageing infrastructure
   - Integrate archive and collection management
   - Digital preservation, access and delivery
Digitisation Program

Total Collection estimated at 750,000,000 pages/images/items

Previously digitised around 30,000 p.a

DEP around 12,000,000 over 10 years
Integrated Collections Management System

AdLib
Hierarchical archival collections and findings aids management, circulation, tracking and display for both digital and physical objects

Ex Libris Alma
- Collection Management Print and Electronic resource management
- Resource sharing
- Authority control
- Serials management
- Fulfillment

Ex Libris Rosetta
- Digital Asset Management
- Ingestion
- Deposit
- Digital Preservation
- Digital Legal Deposits

Ex Libris Primo
Discovery, End user search and content delivery

Corporate Systems
- HR
- Email
- EDRMS
- LDAP
- Finance
Rosetta: SLNSW Vision

- Digital Preservation
- DAM
- Digital Deposit
- Access & Delivery
Rosetta Implementation Options v4.x

- Turnkey/OOTB
- Lightly Configured & Customised
- External Feature Development

Plug and Play

Heavily customised
To implement and integrate Rosetta into the Library’s Digital Ecosystem, comprising corporate and Library applications.
Plug: Rosetta architecture v4.x
Plug: Rosetta Storage (3 copies)
To configure, customise and where appropriate, extend Rosetta’s functionality to meet business objectives
Play: Develop SIP Scripts

- Required for mass ingestion
- Python scripts configurable by JSON files and command line switches
- Ingest variety of collections and formats
- Requires technical capability
- Integration with catalogues

Usage Instructions

generate.py [-h]
  --assetpath ASSETPATH
  --outputpath OUTPUTPATH
  --mappings MAPPINGS
  [--metaenrichment [METAENRICHMENT [METAENRICHMENT ...]]]
  [--embed [EMBED [EMBED ...]]]
  [--embed-mappings MAPPINGS_PATH]
  [--logfile-suffix LOGFILE_SUFFIX]

For example:

```
python3 ./generate.py \
  --mappings mappings/mappings-slnsw-scenario-a.json \ 
  --assetpath /cygdrive/u/sips/input/UAT/2A/SIP_D34826_900937/ \ 
  --outputpath /cygdrive/u/sips/output/UAT/2A/SIP_D34826_900937_TEST/ \ 
  --metaenrichment AdlibMetaEnrichment
  --logfile-suffix sip-test

python3 ./generate.py \
  --assetpath /mnt/sip_gen/input/20170130_hazlewood_a/ \ 
  --outputpath /mnt/sip_gen/output/20170130_hazlewood_a/ \ 
  --mappings ./mappings/mappings-slnsw-turned-digital-image.json \ 
  --embed XMPFileEmbed \ 
  --embed-mappings ./embed_mappings/mappings-slnsw-turned-digital-image-xmp.json \ 
  --metaenrichment AdlibMetaEnrichment
  --logfile-suffix hazlewood-gen-a
```
Play: Develop Download Tool
Play: Customise API access

UTS & SLNSW IMAGE AUTO-TAGGING PILOT

JUNJIE ZHANG (UTS PHD), A/PROF. JIAN ZHANG

- 30K IMAGES FROM THE SAM HOOD COLLECTION
- 20K TRAINING >> 10K TESTING
- 119 LABELS

RESULTS:
- 78.1% OVERALL PRECISION
- 49.3% OVERALL RECALL
- 60.4% OVERALL F1 SCORE

www.bdt.uts.edu.au
Play: Develop Data Visualisation

Play: Develop External Viewer

TITLE:
Album of cloud studies, mountain, bush and harbour scenes, ca. 1841-1850 / drawn by Conrad Martens

COPYRIGHT AND PERMISSIONS:
Out of copyright
Creator died before 1955
Please acknowledge:
Dixson Library, State Library of New South Wales

FILE NUMBER:
FL651541

FILE TITLE:
Album of cloud studies, mountain, bush and harbour scenes, ca. 1841-1850 / drawn by Conrad Martens

CALL NUMBER:
DL PX 28

IE NUMBER:
IE651311

DOWNLOAD:

JPEG image 1.27 MB

SHARE:
| **Play: Develop Adlib UI integration** |

| **TITLE** | Series 14.03: Letter received by Banks from William Wright, 6 November 1778 |
| **CALL NUMBER** | SAFE/Banks Papers/Series 14.03 |
| **LEVEL OF DESCRIPTION** | item |
| **DATE** | 6 November 1778 |
| **TYPE OF MATERIAL** | Textual Records |
| **REFERENCE CODE** | 1003458 |
| **ISSUE COPY** | Digitised |
| **GENERAL NOTE** | CY 3680 / 158 |
| | CY 3680 / 159 |
| **NAME** | Wright, William, 1735-1819 |
| | Clarke, Doctor |
| | Fothergill, John, 1712-1780 |
| | Franklin, Mr |
| | Hope, John, 1725-1786 |
| | Solander, Daniel Charles, 1733-1782 |
| | Webster, Charles |
| **SUBJECT** | Tetanus -- Treatment |
| **OPEN ROSETTA VIEWER** | IE3113858 |
Rosetta Architecture

Adlib: Integration points

Deposit

Manual / Automatic

SIP

Preservation

Management

Working Area

Operational Repository

AIP

Permanent Repository

Search Tools (Primo)

Publishing

Delivery

World leading library: Collect Connect Community
Play: Develop use case 1
Deposit and Ingest from Adlib

- Adlib: Archival Record
- Adlib: Attach Digital Object(s)
- Adlib -> Rosetta: Deposit to Rosetta
- Rosetta: Validate & Ingest
- Rosetta: Move to Preservation
- Rosetta -> Adlib: Update Adlib with Identifiers
- Adlib -> Primo: Publish to Primo
- Primo: Harvest Adlib & Update Indexes
- Alma -> Adlib: Acquisition
- Adlib: Update Adlib with Identifiers
- Primo: Publish to Primo

World leading library: Collect Connect Community

STATE LIBRARY
NEW SOUTH WALES
Play: Develop use case 2
Suppress images from Adlib

- Adlib: Archival Record
- Adlib: Suppress Digital Image(s)
- Adlib -> Rosetta: Update Access Rights
- Rosetta -> Adlib: Update Adlib FL’s
- Adlib -> Primo: Publish to Primo
- Primo: Harvest Adlib & Update Indexes
Play: Develop use case 3
Batch ingest with Adlib/Rosetta
Rosetta Plug-and-Play ability

- High-availability: architecture considerations
- Storage: Enterprise vs fit-for-purpose preservation storage
- Mass ingestion: SIP scripts required
- Delivery UX: OOTB viewer(s) replaced with external viewer
- Integrations: Workflow, catalogue(s) synchronisation
- Capabilities: staff skills (to be discussed in lightning session)
Findings

- Turnkey/OOTB
- Lightly Configured & Customised
- External Feature Development

SLNSW
Future plans

- Alma integration: digital inventory
- DAM use case for non-collection items
- Discovery: integration & context with Primo (full-text transcripts & hierarchies)
- Digital storage strategy
- Streamlined end-to-end workflows
- Improve Adlib archive integration
General discussion

Integration with Archival Management Systems

Teresa Soleau & Euwe Ermita
Open Questions:

• How would you like to map Rosetta with archival/collection management systems?

• Who is currently integrating?

• Who’s interested?
Open Questions:

Integration points:

- Access Rights - especially mixed or tiered access
- One to one? i.e. if you have more descriptive info about digital than you do physical
- Ingest - initiate from CMS? Or Rosetta? Or both?
- Discovery - where is discovery system getting information

Bi-directional synchronization essential?