Ingesting for Preservation & Access

Euwe Ermita
Manager, Digital Library Systems and Services
State Library of New South Wales
Digital Infrastructure
2015

ExLibris

AIELL
ARCHIVES LIBRARIES MUSEUMS
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

Ex Libris Primo
Discovery, End user search and content delivery

Corporate Systems
- HR
- Email
- EDRMS
- LDAP
- Finance

MD Synch, updates MD and item notes
APIs, Web Services, MD Synch
Publishing
Deposit and publishing
MD Synch
Publishing
APIs, Web Services, MD Synch
"Whenever I walk in a room, everyone ignores me."

Digital Preservation Elephant 2015
From ingest to access

People

Process

Tools
Process: Governance

Recent key decisions
- Access copy replacement for MP3s – AAC codec in MP4 wrapper
- Item-level records for born-digital material

Collaboration and input
- Specification for high-resolution images
- Digital preservation and collection lifecycle
- Digital preservation resourcing
## Process: Content models

<table>
<thead>
<tr>
<th>Digital Preservation Development Group</th>
<th>Content Model Working Group</th>
</tr>
</thead>
<tbody>
<tr>
<td>Master TIFF</td>
<td>Co-master TIFF</td>
</tr>
<tr>
<td>Access JPG</td>
<td>Rep (n) Sidecar XML</td>
</tr>
</tbody>
</table>

specifications, metadata, access rights, formats
Draft digital collection management workflow

Process: Workflow

- Appraisal
- Acquisition
- Storage
- * Transfer
- * Ingest
- * Pre-ingest
- * Technical Analyst Workbench (TAW)
- Preservation
- Delivery
- Access
- Retrieval
- Description
- * Ingest
- Workbench (TAW)
People: Roles

Inter-divisional collaboration
- Technical Lead
- Developer
- Systems/Applications support officer
- Technical Analyst
- Digital Curation Specialist
- Digital Collections Analyst
- Digital Collections Librarian
- Digital Media Technical Analyst

* Technical Analyst Workbench (TAW)
* Transfer
* Pre-ingest
* Ingest
Tools: Project PanDA

Preservation and Digital Access

AUTOMATED
- Process
- Notify
- Retry

TRANSPARENT
- Progress
- Logs
- Reporting

FLEXIBLE
- Configurable
- Priority
- Scalable
Tools: Project PanDA

Asset Preparation

Prepare rep folders → Generate IE meta.json files → Generate rep access rights files → Generate file_meta.json files

SIP Generation

Enrich json files with metadata from Alma/Adlib → Embed metadata from Alma/Adlib into files → Load mapping file → Generate SIPS

Rosetta Ingestion

Load SIPS → Virus Check → DROID file format identification → Move to permanent storage → JHOVE file validation → Search indexing

Technical metadata extraction → Format risk analysis → Move to permanent storage → Search indexing
Tools: Project PanDA
Tools: Solutions Business Manager

- End to end process
- Workflow automation
- Digitisation request workflow
- Integration with PanDA
- Metrics and reporting
### Tools: Adlib UI integration

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

**VIEW MEDIA FILES**

- ![Image 1](image1.png)
- ![Image 2](image2.png)

**BROWSE COLLECTION HIERARCHY**

- **Aggregated collection (Contains 96 series)**
  - Sir Joseph Banks Papers, 1767-1822. Includes some papers of Sarah Sophia Banks, 1766-1773, 1775-1779; and Dorothea, Lady.
Tools: External Image 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
Dixon 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:
Next 12 months

- Process: Separate access and preservation workflows at ingest
- People: Digital preservation team – roles, responsibilities, capabilities and capacity
- Tools: Digital storage strategy
Future: 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
Future: Insights via data visualisation

Thank you