Managing Digital Assets from Curation to Exhibition

Reintroducing MAGPIE
Metadata Annotation GUI Providing Ingest and Export

William Welling, James Creel, Jeremy Huff, Jason Savell, Douglas Hahn, Michael Bolton
Challenges

• Annotating digital assets
  – Accuracy
  – Efficiency
  – Normalization
  – Synchronization

• Managing digital assets
  – Multiple collections
  – Varying sources
  – Varying destinations

• Exhibition
  – Exhibit multiple collections with annotations
Use Cases

• Legacy dissertations
• Agricultural Bulletin
• Image Collections
• Newspapers and articles
• Existing Collections
• Etc.
Ecosystem

- MAGPIE Service and UI
- Pelican Service
- Authorization Service
- Scanning workstations
- Abby OCR
- DSpace
- Fedora
- Archivematica
- Spotlight
Components

- Projects
- Observers
- Authorities
- Automatic Suggestions
- Exporters
- Repositories
- Exhibits
Projects

• Defined by JSON
• Configurable
  – metadata input
    • label
    • gloss
    • input type
  – authorities
    • CSV
    • Voyager
  – suggestors
  – repositories
    • DSpace
    • Fedora
    • Archivematica
Observers

- listen to directories
  - new projects
  - new documents
  - ingest notification from IR
- local file system
- network drive
Authorities

• CSV
  – populated with document name and metadata
  – maps to document’s metadata

• Voyager
  – requests MARC record
  – maps to document’s metadata
Automatic Suggestions

• Pelican
  – semi-automatic suggestion
  – provides detailed suggestions to annotator
Exporters

- SAF
  - DSpace
- CSV
  - DSpace
  - Archivematica
  - Spotlight
Repositories

• DSpace
  – single item ingest
  – batch ingest
• Fedora
  – single item ingest
  – PCDM ingest
• Archivematica
  – initiates archival transfer
Exhibits

• Spotlight
  – batch upload using CSV
  – items from Fedora *duplicated in Spotlight
Workflow

• Documents
  – scanned with OCR
  – existing in IR
  – images from physical exhibit

• Annotations
  – automatic from authorities
  – semi-automatic from suggestions
  – manual input
Workflow

• Exports
  – CSV
  – SAF

• Ingests (Batch or Single)
  – DSpace
  – Fedora
  – Archivematica
Demo – Legacy Dissertations

Legacy Dissertations Project

File System
- Legacy Dissertations
- Agricultural Bulletin
- Spotlight Exhibit

MAGPIE
- Observer
- Observer
- Observer

Repository
- Archivematica
- Fedora
- DSpace

Exporters
- CSV
- SAF

Suggestor
- Authority

System Components
- Pelican
- CSV
- CSV
- Voyager
- Spotlight
- CSV
- SAF
Demo – Agricultural Bulletin

TAES Agricultural Bulletin Project

File System
- Legacy Dissertations
- Agricultural Bulletin
- Spotlight Exhibit

Pelican
- Suggestor
- Authority

CSV
- Authority

Voyager

MAGPIE
- Observer
- Observer
- Observer

Repository
- Repository
- Repository
- Repository

Archivematica

Fedora

DSpace

Exporter
- Exporter

Spotlight

CSV
- SAF

Exporter
Demo – Spotlight Exhibit

Spotlight Exhibit Project

File System
- Legacy Dissertations
- Agricultural Bulletin
- Spotlight Exhibit

MAGPIE
- Observer
- Observer
- Observer

Pelican
- Suggestor
- Authority
- Authority

CSV

Voyager

Spotlight
- CSV
- SAF

Repository
- Archivematica
- Fedora
- DSpace

Exporter
- Repository
- Repository
- Repository
Future Direction

• UI managed projects
• UI annotation to support PCDM
• Heuristics to guess metadata
Thank You.

Any Questions?