

Between Millennium & CONTENTdm

Stuart Hunt

EIUG conference
London
September 2008

THE UNIVERSITY OF
WARWICK

Overview

- Background
- Digitisation issues
- Metadata issues
- Workflows
- The future

Background

- Marandet Collection
- 4000+ French plays
 - 1720 to 1900
- Acquired 1974 & 1979
 - 1720 to 1830 & 1830 to 1900
- Guide published 1979
- Bibliographic records in Millennium

Project

- 2005/6
- £45K internal funding
- Library
- Dept of French Studies
- Revolutionary Drama
 - 1789 to 1800
 - 301 plays

Objectives

- Full-text searching
- Cross-searching
- Proof-of-concept
- Transition from project to service
- Integration with existing systems
 - Millennium
 - Web

Options

- Existing solutions
 - Millennium
 - In-house web publishing tool
- Separate product
 - Digital collection management software

Media Management

- Link digital content
 - Existing bibs
 - New bibs
- No full-text
- No cross-collection searching (unless in separate scope)
- Not a total solution

CONTENTdm

- Windows server
- Software
- Acquisition Station
- Web Admin
- Web UI
- Full-text & cross-collection searching
- Initial licence <50K images

Digitisation

- Image files
 - Outsourced
 - Tiff & jpeg derivative
- Text files/transcripts
 - OCR quality not acceptable
 - Re-keying
 - Outsourced
- File-naming conventions

Metadata context

- Separate bibs
 - Standard local practice
 - Print
 - Electronic
- 793 for local added entry/uniform title
 - Collection name

Metadata options

- Play-by-play from already existing Millennium bibliographic data
- Play-by-play & harvest from CDM into Millennium via XML Harvester
- Batchload via delimited file from Create Lists, cross-walking MARC21 to DC
 - Directory structure

Metadata workflow

- Create separate bibs for e-versions
- Export print records via Data Exchange
- MarcEdit to remove extraneous tags (907, etc)
- Insert 007, GMD, 533, revise 008 (/23)
- Re-import into Millennium as new bibs
- 856 reference url added to bibs after added to CONTENTdm

Metadata in CONTENTdm

- Compound objects
- Document level
- Page level
 - Page level less rich than document level (local decision)
- Hospitable to multiple schemas
- Administrative metadata
 - Later feature

Document level

- AACR in DC wrapper
- All descriptive metadata from bib (except LDR, 008, GMD)
- Authority control (names, subjects, uniform titles)
- Rights
- Identifier (.b number)
- Mapped to DC for OAI harvesting

Page level

- Basic descriptive metadata (creator, title, publisher, date)
- Rights
- Identifier (.b number)
- Transcript (dc:description)
- No OAI harvesting at page level
 - Local decision

Accessing content

- CONTENTdm
- CQR
- WebPAC
 - E-resources scope
- Bibs contributed
 - RLUK & COPAC
 - WorldCat
- OAI

Additional projects

- Napoleonic/Empire period (1801-1815)
 - Education Innovation Fund
 - Joint bid with Dept of French Studies
 - 2 yrs (2006/7, 2007/8)
 - Student-selected content
 - Undergraduate research driven
- Additional Revolutionary content
 - Mopping-up outstanding funds (2008)

Present

- Established service
 - Revolutionary Drama: 339 plays
 - Empire Period Drama: 123 plays
- Licence increased to unlimited
- Further funding from Dept of French Studies
- Links from CÉSAR
- *Marandet Matters*

Future

- JISC Digitisation Programme: Enriching Digital Resources
 - 2008/9
- Platform for delivery of digital content
 - Use for other projects & services
- Integration with Encore

Links

- <http://contentdm.warwick.ac.uk>
- <http://webcat.warwick.ac.uk>
- <http://go.warwick.ac.uk/fac/arts/french/marandet/>

stuart.hunt@warwick.ac.uk