Using Primo to discover e-research repositories

Stefania Riccardi, Susan Lafferty, Tom Ruthven, Sue Harmer
Outline

1. Background – Library Repository Services and Data Librarianship

2. UNSW Multi-repository model

3. Primo customisations

4. Next Steps
UNSW Library Repository Services (LRS)

**Created**: January 2009.

**Aim**: support online environments for research, learning and teaching

**Position description**:

UNSW Library…is actively working with research projects, Schools and Research Centres to develop infrastructure and services that explore and meet this new role for libraries. The Library is focussed on standards-based, interoperable services and systems
Multirepository Model

October 2010

The suite:

• **One Fedora repository** per faculty
• **One Primo front end**, with multiple Views – i.e. a separate View for each group
• A **deposit tool** (Valet) permitting independent deposit of metadata and digital objects
• An **editing tool** built in-house, which we have not named.
• **Citation builder** – developed in house.

…..Intention is to move from Primo Views to Primo Institutions – greater flexibility in customising for each faculty.
The model: Single Fedora per faculty

Primo views and Fedora PID/Sets in Multirepository environment

1 Fedora per Faculty
1 Fedora PID or 1 Fedora OAI-PMH SET per subject repository (e.g. 'Water' within Engineering Fedora)
1 search scope (PRIMO view) = Many source scopes (Fedora subject repositories)
The Model (single Fedora instances per faculty)

DRAFT – Multirepository model – Primo Views and Scopes (based on multiple Fedora PID)

**Faculty Server**

- Fedora
  - Measles (001)
  - Rubella (002)
  - Chicken Pox (003)
  - Flu (004)

**Primo Scopes**
- 001
- 002
- 003
- 004

**Primo Views**
- Measles/Rubella
  - Tab 1 – Scope 1 only
  - Tab 2 – Scope 2 only
- Medicine
  - Tab 1 – All scopes
  - Tab 2 – Scope 1 only
  - Tab 3 – Scope 2 only
  - Tab 4 – Scope 3 only
  - Tab 5 – Scope 4 only
- Measles
  - 1Tab – Scope 1 only
- Exanthematos diseases
  - 1Tab – Scopes 1, 2 & 3

** SCOPE** – Specific group of records
**VIEW** – Different look and feel for one or more scopes; View has 1 default institution mandatory for end-user authentication.
**INSTITUTION** – Has many PDS and IP ranges; PDS are needed so that users not on Active directory can sign-on. Has one e-shelf

Single Sign-On only works on EL products but can be customised to include other applications as long as they support Shibboleth (not MemRE).
The Model (multiple Fedora instances per faculty)

DRAFT – Multirepository model – Primo Views and Scopes (based on single Fedora PID)

Primo View
Measles/Rubella
- Tab 1 – Scope 1 only
- Tab 2 – Scope 2 only

Primo View
Medicine
- Tab 1 – All scopes
- Tab 2 – Scope 1 only
- Tab 3 – Scope 2 only
- Tab 4 – Scope 3 only
- Tab 5 – Scope 4 only

Primo View
Measles
- 1 Tab – Scope 1 only

Primo View
Exanthematous diseases
- 1 Tab – Scopes 1, 2 & 3

Primo View
Scope 1
Fedora Measles

Primo View
Scope 2
Fedora Rubella

Primo View
Scope 3
Fedora Chicken Pox

Primo View
Scope 4
Fedora Flu

Faculty Server

SCOPE – Specific group of records
VIEW – Different look and feel for one or more scopes; View has 1 default institution mandatory for end-user authentication.
INSTITUTION – Has many PDS and IP ranges; PDS are needed so that users not on Active directory can sign-on. Has one e-shelf

Single Sign-On only works on EL products but can be customised to include other applications as long as they support Shibboleth (not MemRE).

Riccardi 23/8/11
Fedora: a bucket of digital objects – and each object a bucket in its own right

Image: By compn http://www.flickr.com/photos/compn/5431773976/#/
....and each object a bucket in its own right
....and each object a bucket in its own right
....and each object a bucket in its own right

**DC descriptive metadata** – OAI-PMH compliant – so anyone can harvest the record

**JHOVE preservation metadata** – includes info about size of attachment, - elements ‘size’ and ‘size format’
....and each object a bucket in its own right

DC descriptive metadata – OAI-PMH compliant – so anyone can harvest the record

JHOVE preservation metadata – includes info about size of attachment, elements ‘size’ and ‘size format’

Licence – may be PDF, word, text, any format at all

Image: By compn
http://www.flickr.com/photos/compn/5431773976/#/
...and each object a bucket in its own right

**DC descriptive metadata** – OAI-PMH compliant – so anyone can harvest the record

**JHOVE preservation metadata** – includes info about size of attachment, - elements ‘size’ and ‘size format’

**Licence** – may be PDF, word, text, any format at all

**Attachment** – may be PDF, audio, AV, etc

Image: By [compn](http://www.flickr.com/photos/compn/5431773976/#/)
...and each object a bucket in its own right

**MODS metadata record** – Provides better data granularity to support application such as citation builder and editing tool.

**DC descriptive metadata** – OAI-PMH compliant – so anyone can harvest the record

**JHOVE preservation metadata** – includes info about size of attachment, elements ‘size’ and ‘size format’

**Licence** – may be PDF, word, text, any format at all

**Attachment** – may be PDF, audio, AV, etc

Image: By [compn](http://www.flickr.com/photos/compn/5431773976/#/)
Primo Customisations

Tile Customisation – using controlled vocabularies

Citation Builder (in house)
Reads the MODS record, builds the (modified Harvard) citation based on resourcetype, then puts it into the DC record

Use of Primo File Splitter:

• Harvests the Citation and displays it in Primo

• Providing information about size of attachment in Fedora

• Converting PDF attachments to text (indexing will mine only first 10,000 words for searching, the setting can be changed on Primo Back Office) Ex Libris has advised 10,000 word limit.

Harvesting wikis

• Currently only available in User Acceptance Testing
Customisation

Link to standard URL(s) – source: DC metadata record)
Customisation

Link to object (note file size – source: JHOVE metadata record)
Customisation

Library

UNSW WORKS

Search for UNSWorks

Conference Paper

Strategies for large scale blended learning initiatives: Training, teaching and management
McIntyre, Simon, College of Fine Arts, UNSW; Watson, Karen, College of Fine Arts, UNSW; Larsen, Soren, University of Copenhagen
The Fourth International Blended Learning Conference "Engaging Students in the Curriculum", 2009, June

Details

Title: Strategies for large scale blended learning initiatives: Training, teaching and management
Author/Creator: McIntyre, Simon, College of Fine Arts, UNSW; Watson, Karen, College of Fine Arts, UNSW; Larsen, Soren, University of Copenhagen
Resource Type: Conference Paper
Date: 2009
Conference names: The Fourth International Blended Learning Conference "Engaging Students in the Curriculum", 2009, June
Publisher: University of Hertfordshire, Hatfield, UK

Links

- Additional links
- McIntyre, Watson, Larsen. Blended Learning. Paper draft 1.0
- Add/Remove attachment

Link to Editing tool

Link to deposit tool
Editing and Valet deposit tools
Handle, built from combination of Fedora PID and Primo PID
Modifying PRIMO Search tile:
Knowledge Centre

Welcome to the Human Resources for Health Knowledge Centre, a resource that lists publications used in our research on the topic of Human Resources for Health. Where possible, the library will link you to full-text papers.

If you have any questions about the Knowledge Centre please contact our Knowledge Manager Denise Quey.

Links to other resources:
HRH Global Resource Centre
Human Resources for Health
TOPICS in Advanced Search
Text mining first 10,000 words…

We use Primo’s File Splitter to

• grab Fedora objects which are the **Active** attachments (eg fulltext)

• Convert PDFs to text and add the text to a free text data element of the Primo PNX Extension (Primo’s metadata schema) to enable full text searching of the first 10,000 words of each object in Primo

… we don’t text mine licences, embargoed documents, or other elements of a Fedora object which are not ‘active’.
Harvesting wikis …if only
Issues

• Each time we install a new Service Pack – we need to test all customisations.

• If a stored file is too big, or the format difficult (even .mov files) we can’t access them. This is a Fedora issue, not a Primo issue. We’ll get to it.

• Consolidation and sustainability – now becoming an issue

• Controlled vocabularies – time consuming to modify once they are in place

Next steps

• Consolidate

• Evaluate

• Resource
Summary

1. UNSW Repository support environment
2. UNSW Multi-repository model
3. Primo customisations
4. Next Steps
Authors

TECHNICAL CONTACT:
Stefania Riccardi       s.riccardi@unsw.edu.au
Susan Lafferty         susan.lafferty@unsw.edu.au
Tom Ruthven            t.ruthven@unsw.edu.au
Sue Harmer             s.harmer@unsw.edu.au