

FEZ development at the University of Queensland Library

As part of its eScholarship testbed (<http://www.library.uq.edu.au/escholarship>) for the Australian Partnership for Sustainable Repositories project (<http://www.apsr.edu.au/>), staff at the University of Queensland Library are developing software for a flexible, digital repository and workflow management system which will be released under open source license by the end of 2005. The digital repository software is PHP- and MySQL-based and works as a front-end and administration tool using Fedora (<http://www.fedora.info/>). Designed around Fedora 2, it uses RELS-EXT to describe relationships between objects for communities, collections, and records. Records will have associated datastreams which may be documents such as PDFs. Records can belong to both collections and communities. All records will also be associated with a controlled subject thesaurus. The software will also be able to handle multiple metadata schema, i.e. Dublin Core, EAD and so on.

The digital repository can handle different content models which can be customised through the web-based administration interface. A conference paper, for instance, could have multiple datastreams - one for the actual document, e.g. the PDF of the paper, one for the Dublin Core data, one for preservation metadata, one for provenance, one with conference-specific data such as location, date and so on. New content models can be added at any stage. Each object type is governed by XSDs that map those XSDs against HTML inputs, static variables or dynamic variables and search keys. The create, edit and view forms, searches and the consequent XML objects are controlled dynamically by these mappings.

Each object will have layers of associated metadata – structural, rights, preservation and so on.

The system is designed to be open access, with records harvestable under OAI-PMH. Some records can be of restricted access if so desired, for example, embargoed archival material.

Records can be added to the system singly. For bulk ingest, existing scripts on the eScholarship page that convert OAI records to METS/XML for ingest can be used (OAI to METS XML Converter (PHP) (<http://www.library.uq.edu.au/escholarship/tools/oai2mets.zip>)). There is also a script for converting records currently in ADT into an OAI stream that can then be converted to METS/XML (ADT OAI HTML Stripping Service Provider (PHP) (<http://www.library.uq.edu.au/escholarship/tools/adt2oaisp.zip>)).

The system is also designed to handle a number of different roles – creator, editor, viewer, approver, commenter, comment-viewer, annotator and so on. The permissions and authentication assigned to each role will be governed by rules set up in the FEZ-ACML.

Authentication can be by groups such as LDAP/Active Directory, the III Innopac library system, or FEZ internal groups, as well as by any other groups defined in the system.

The workflow system will also be covered by an XSD to dynamically describe workflows for particular document types. The workflow will cover the life cycle of an object in Fedora – from the create or deposit stage, through approval and edit, to the publish stage. It will also cover security and the creation of an archival copy.

Achieved:

The digital repository set up and authentication system are completed. The workflow system is currently being written. Workflow will allow repository managers to configure all the different stages through which a person or object moves in the carrying out of common tasks. For users, this might be to register or to upload materials, or to comment on, or annotate, existing objects within the system. Objects in their turn can be ingested, viewed, updated, archived and so on.

Simple and advanced search are completed, with further 'Google-isation' being planned. Browse views by author, year, controlled vocabularies and recently added items are available. Browse by recently accessed items will be added. JHOVE is being used for automatic preservation metadata extraction. Simple statistics are available for top 50 authors, top 50 papers and by item.

In the works:

- Dynamic customisable Javascript input validation
- Generic FEZ branding for banners, footers etc

Planned for releases soon after initial release (within 1-3 months):

- Annotation and commenting system
- Registration process
- Full text indexing
- Dynamic Canned Search links (eg for Authors, Subjects)
- Integrate the VTLS Fedora open-source modules for Handles, Google web crawler
- Conversion to XML format for Word/PDF files
- More manual & automatic provenance metadata extraction
- Creative Commons licensing tie-in

Planned for later releases (within 3-6 + months):

- Federated searching, and federated authentication with MAMS
- Batch import for files with associated MARC21 metadata
- Ill Inopac catalogue authentication
- Test and Integrate with Fedora 2.1+ features, especially OAI sets and searches
- RSS feeds

Belinda Weaver
Coordinator, APSR at UQ
4 October, 2005