Using Hydra to Model Sustainable Collections

Chris Awre

Sustaining content and collections

Repository Fringe

2nd August 2016
Fedora/Hydra @ University of Hull

Welcome to Hull's digital repository, Hydra

Search the University's digital collections using the search box, or browse them via the links below...

The Hydra repository is a digital archive for the University of Hull. It has been developed to hold, manage, preserve and provide access to the growing body of digital material generated through the research, teaching and administrative activities of the University.

Open Access publication

Submit your journal articles to Hydra

Email the accepted version to repository@hull.ac.uk

See http://libguides.hull.ac.uk/openaccess for more information.

Browse

Collections
Managing digital collections

1. Transparency is in the eye of the beholder: the effects of identity and negative perceptions on ratings of transparency via surveys.

Creator: Park, Heungsik; Blenkinsopp, John
Resource type: Journal article
Subject: Transparency; Public administration; Administration and democracy; e-Government; Trust
Language: English

2. Forms, dialectics and the healthy community: the British idealists' receptions of Plato
Five principles (leading to our implementation)

- Fedora is content agnostic
- Fedora is (open) standards-based
- Fedora is scalable
- Fedora understands how pieces of content relate to each other
- Fedora is manageable with limited resource
  – With help from the community
Five principles (leading to our implementation)

- Fedora/Hydra is content agnostic
- Fedora/Hydra is (open) standards-based
- Fedora/Hydra is scalable
- Fedora/Hydra understands how pieces of content relate to each other
- Fedora/Hydra is manageable with limited resource
  - With help from the community
What is Hydra? Community

• Hydra Connect

• Mailing lists, Slack, Skype/Hangouts

• Meetings – manager and technical focus
...are individuals, institutions, corporations or other groups that have committed to contributing to the Hydra community; they not only use the Hydra technical framework, but also add to it in at least one of many ways: code, analysis, design, support, funding, or other resources.

**Hydra Partners collectively advance the project and the community for the benefit of all participants.**
Metadata modelling in MODS

• Use MODS as our baseline for describing different types of content

• Create profiles for dissemination on the fly
  – DC
  – UKETD_DC
  – RIOXX

• Could apply a different metadata schema if preferred
  – E.g., PBCore for multimedia
Adapt to the content

Journal article

CLIF: moving repositories upstream in the content lifecycle

Authors
Waddington, Simon; Green, Richard A.; Awe, Christopher L.

Subjects
CLIF, JSC: Content lifecycle; institutional repository; Sakai; eBridge; Microsoft SharePoint

Abstract
The UK JSC-funded Content Lifecycle Integration Framework (CLIF) project has explored the management of digital content throughout its lifecycle from creation through to preservation or disposal. Whilst many individual systems offer the capability of carrying out lifecycle stages to varying degrees, CLIF recognised that only by facilitating the movement of content between systems could the full lifecycle take advantage of systems specifically geared towards different stages of the digital lifecycle. The project has also placed the digital repository at the heart of this movement and has explored this through carrying out integrations between Fedora and Sakai, and Fedora and SharePoint. This article will describe these integrations in the context of lifecycle management and highlight the issues discovered in enabling the smooth movement of content as required.

Published
Published in Journal of Digital Information, 2012

Dataset

HMAP Dataset 06: Newfoundland, 1675-1698

Description
Fishermen, settlers and cod catches in 17th-century Newfoundland.

Citations
The map below gives an indication of the extent of the Newfoundland-Labrador shelf; the View as map link in the download panel at the right will show a much more detailed representation. The kmz file download, when used with Google Earth, will render the extent of the Newfoundland-Labrador shelf in detail.

Coverage
Newfoundland

Temporal
1675-1698

Geo-data

Download

Article
(3 MB pdf)

QR code link to this page

Download

Database - ASCII format
( zip)

Database - Access 2000 format
( zip)

Database - csv format
( zip)

Documentation - PDF format
( pdf)

Documentation - text format
( txt)

Documentation - Word (.docx) format
( docx)

Newfoundland-Labrador Shelf - kmz file
(495 KB vnd.google-earth.kml+xml)
View as map

QR code link to this page
Archivematica as part of a Hydra preservation workflow in Hull

Where are we now?
Hull has a well established Hydra repository but we need to be able to preserve research data and other content for the long-term.

Why the need?
We have always intended that the repository should offer the option of long-term preservation but UK universities now have a mandate to preserve research data in particular.

Why Archivematica?
Archivematica is a well respected, open-source tool which seemed to offer much of the functionality that we needed. With the University of York we received a Jisc grant to test it out.

What are we doing now?
We and York now have a Phase 2 grant (ending November 2015) which is enabling us to work with Archivematica to improve its applicability to research data.

What do we hope to do?
We hope we shall be successful in bidding for a Phase 3 grant (January - June 2016) which would enable both Hull and York to build “proof-of-concept” systems.

It looks something like this...

- Workflow deposit tool
- Proto-queue
- Metadata record
- QA-queue
- Repository
- Ingest folders
- Arrangement tool
- DIP processor
- Archivematica
- AIP store
- Fedora PID
- Legacy content for AIPs
- DIP
- Hydra directory
- UUID
- Content files
- Batch ingest
- "Preservation only" content
- External systems (History Centre etc).
- Reprocess for DIP

Your thoughts and questions are welcome!
c.awre@hull.ac.uk
r.green@hull.ac.uk
hydra.hull.ac.uk
Move to Linked Data with Fedora 4

PCDM (Portland Common Data Model), for data and code interoperability
Summary

• Sustainability comes through community
  – And flexibility in software development and delivery
• A repository cannot do everything
  – So link it to tools that can assist with sustainability
• Datasets come in all shapes and sizes
  – Still finding common themes, but treating each on case by case basis for now
• Looking ahead, RDF will foster further content sustainability
Thank you

c.awre@hull.ac.uk