

Meeting institutional needs for digital curation through shared endeavor

The application of Hydra/Fedora at Hull

Chris Awre

Head of Information Management, Library and Learning Innovation

IDCC, 10th February 2015

To cover



- Background/context
- Digital repository development at Hull
- Open source and community
- Applying Fedora at Hull
- Hydra development
- Hydra and Fedora
- Hydra and Hull
- Reflections and looking ahead

Background/context

- 2005
 - The Jisc FAIR (Focus on Access to Institutional Resources) programme was finishing
 - Lots of emphasis on having a stable platform upon which to base management of institutional assets
 - Led to first Digital Repositories programme, 2005-7
 - University of Hull started down its path of implementing a digital repository
 - Desire to address multiple use cases together through a single solution
 - Practical and financial basis and limitations

An institutional repository

One institution = one repository?

- Repositories are infrastructure
 - Maintaining infrastructure requires resource, which we need to minimise to justify costs in the long-term
- Content doesn't sit in silos
 - One repository facilitates cross-fertilisation of use
- Integration with one system
 - Embedding the repository means linking to one place

Five principles

- A repository should be content agnostic
- A repository should be (open) standards-based
- A repository should be scalable
- A repository should understand how pieces of content relate to each other
- A repository should be manageable with limited resource

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



Open source and community

- University of Hull had recognised that engaging in open source (community source) projects would enable more local implementation and innovation
- uPortal
 - Institutional portal solution
 - Came out of JASIG community
- Sakai
 - VLE solution
 - Community built as part of Mellon grant conditions
- Now both part of Apereo Foundation
 - Hull contributed Executive Director!

Implementing Fedora at Hull

- RepoMMan project, 2005-7
 - Examining how a repository can be made part of day-to-day activity
 - Not just somewhere to put stuff at the end of a process
- REMAP project, 2007-9
 - Applying a repository to serving records management and preservation by starting to capture items from their creation
- Institutional developments
 - Creation of local service in parallel (October 2008)
 - Initial use cases were e-theses and exam papers, not OA

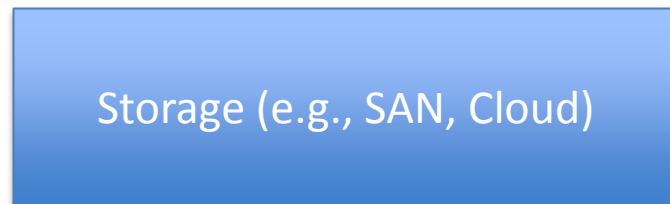
Fedora and end-user interfaces



Need an end-user interface



Fedora is the digital repository system, holding the content in a highly structured way



The content is stored either locally or in the Cloud (currently a slice of the SAN)

Fedora and end-user interfaces



Muradora

Adopted open source Muradora system from Macquarie University in Australia

Fedora

Fedora is the digital repository system, holding the content in a highly structured way

Storage (e.g., SAN, Cloud)

The content is stored either locally or in the Cloud (currently a slice of the SAN)

Fedora and Hydra



- Fedora can be complex in enabling its flexibility
- How can the richness of the Fedora system be enabled through simpler interfaces and interactions?
 - The Hydra project has endeavoured to address this, and has done so successfully
 - Not a turnkey, out of the box, solution, but a toolkit that enables powerful use of Fedora's capabilities through lightweight tools
- Hydra 'heads'
 - Single body of content, many points of access into it

Fedora and Hydra



Hydra

Hydra

Hydra

Hydra

Fedora

Storage (e.g., SAN, Cloud)

Hydra provides user interfaces and workflows over the repository

Concept of multiple Hydra 'heads' over single body of content

Fedora is the digital repository system, holding the content in a highly structured way

The content is stored either locally or in the Cloud (currently a slice of the SAN)

Hydra

- A collaborative project between:
 - University of Hull
 - University of Virginia
 - Stanford University
 - Fedora Commons/DuraSpace
 - MediaShelf LLC
- Unfunded (in itself)
 - Activity based on identification of a common need
- Aim to work towards a reusable framework for multipurpose, multifunction, multi-institutional repository-enabled solutions
- Timeframe - 2008-11 (but now extended indefinitely)





Fundamental Assumption #1

No single system can provide the full range of repository-based solutions for a given institution's needs,

...yet sustainable solutions require a common repository infrastructure.

Fundamental Assumption #2

No single institution can resource the development of a full range of solutions on its own,

...yet each needs the flexibility to tailor solutions to local demands and workflows.

Eight strategic priorities

1. Solution bundles
2. Turnkey applications
3. Vendor ecosystem
4. Training
5. Documentation
6. Code sharing
7. Community ties
8. Grow the User base

Hydra partnership

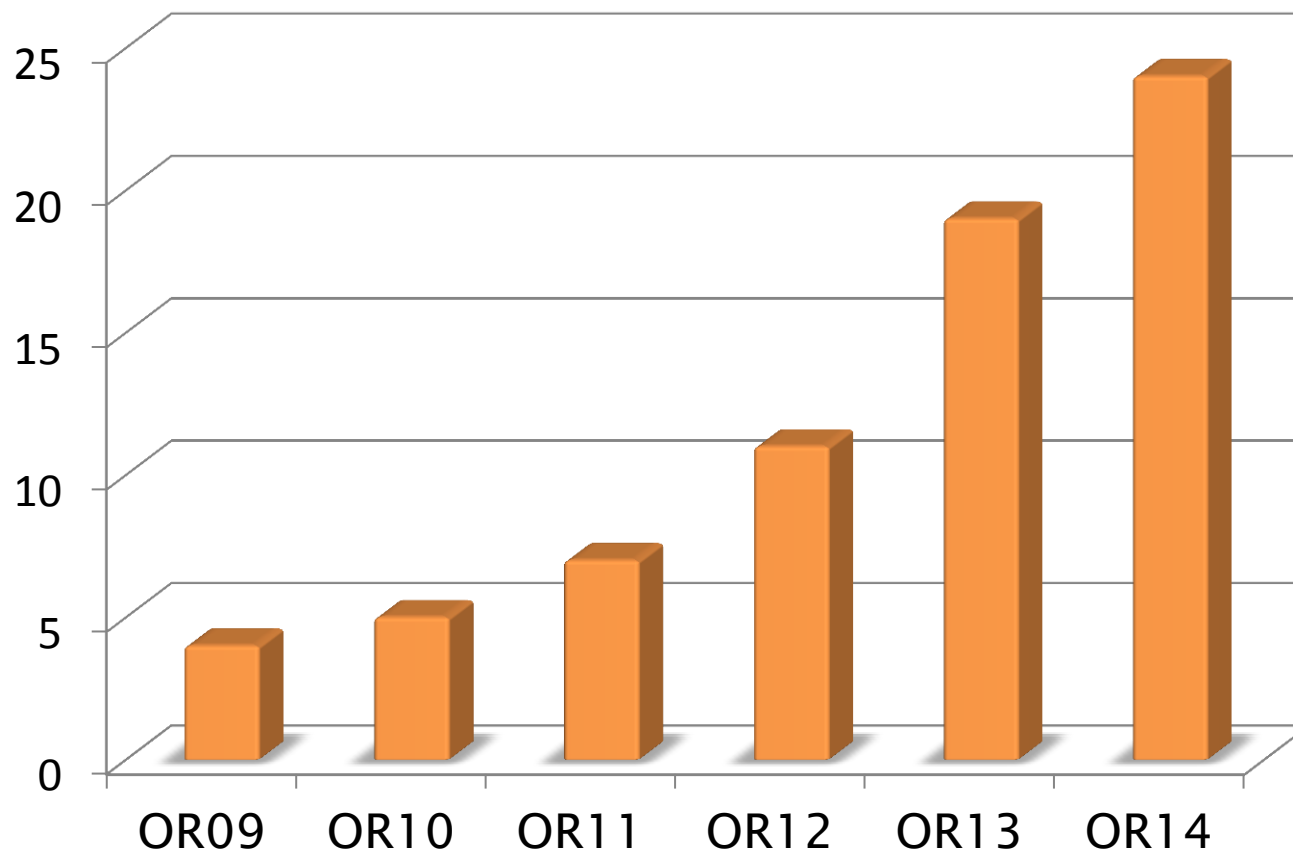


- From the beginning key aims have been and are:
 - to enable others to join the partnership as and when they wished (Now up to 27 partners)
 - to establish a framework for sustaining a Hydra community as much as any technical outputs that emerge
 - Establishing a semi-legal basis for contribution and partnership

“If you want to go fast, go alone. If you want to go far, go together”

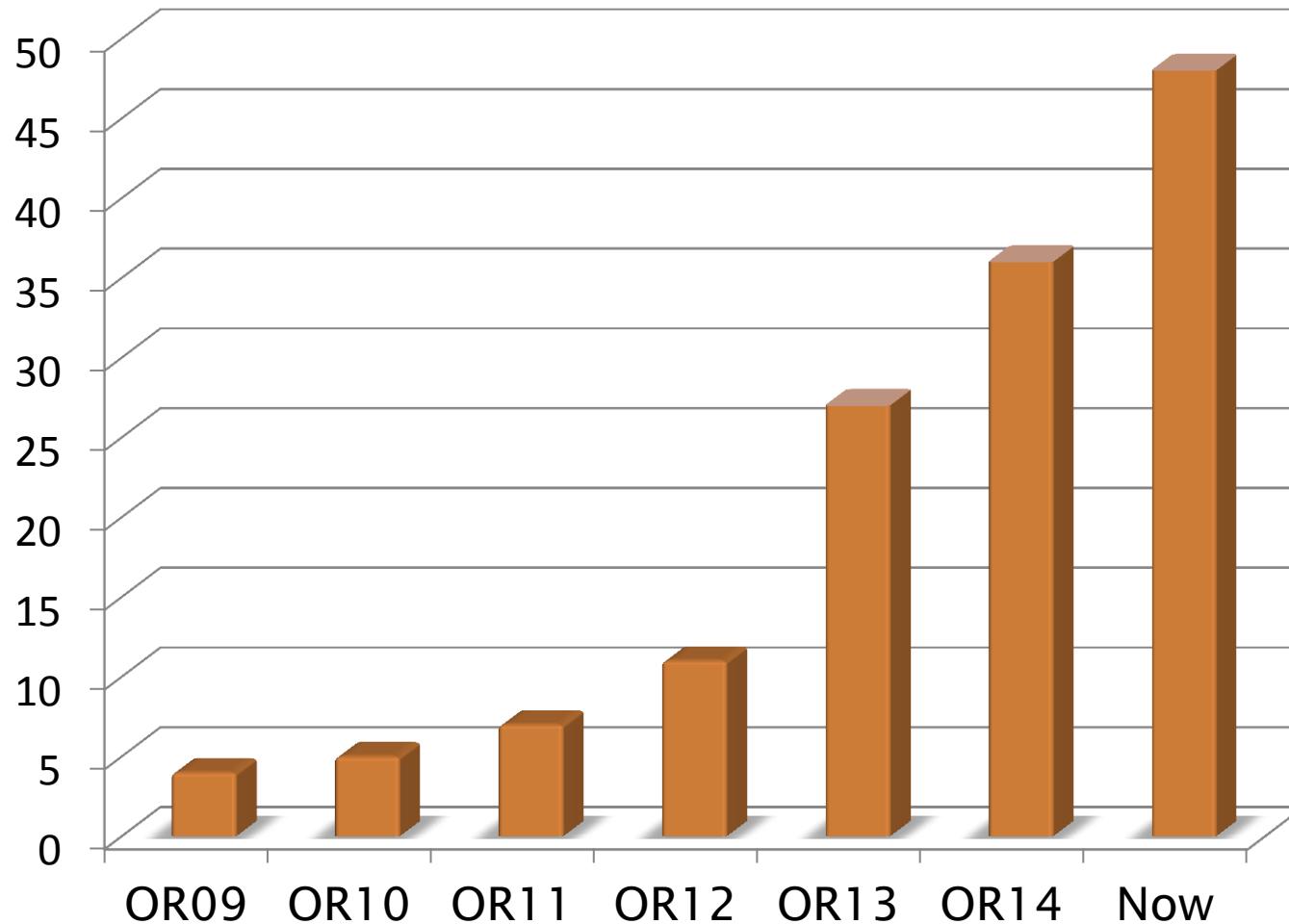
(African proverb)

Hydra Partners



OR = Open Repositories Conference

Hydra Partners *and Known Users*



OR = Open Repositories Conference

A Worldwide Presence



Software development

In a Nutshell, Project Hydra...

... has had 9,623 commits made by 78 contributors representing 67,625 lines of code

... has a codebase with a long source history maintained by a very large development team with decreasing Y-O-Y commits

... took an estimated 17 years of effort (COCOMO model) starting with its first commit in October, 2009 ending with its most recent commit 9 days ago

Hydra technical implementation



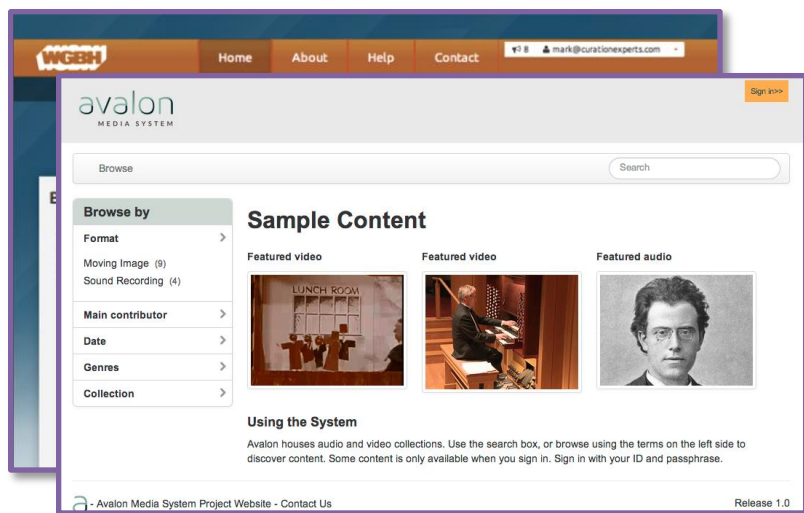
- Fedora
 - All Hydra partners are Fedora users
- Solr
 - Very powerful indexing tool, as used by...
- Blacklight
 - Prior development at Virginia (and now Stanford/JHU) for OPAC
 - Adaptable to repository content
- Ruby
 - Agile development / excellent MVC / good testing tools
- Ruby gems
 - ActiveFedora, Opinionated Metadata, Solrizer (MediaShelf contributions)

Four Key Capabilities

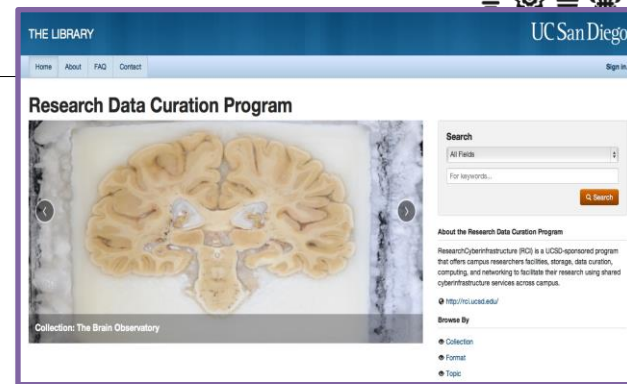
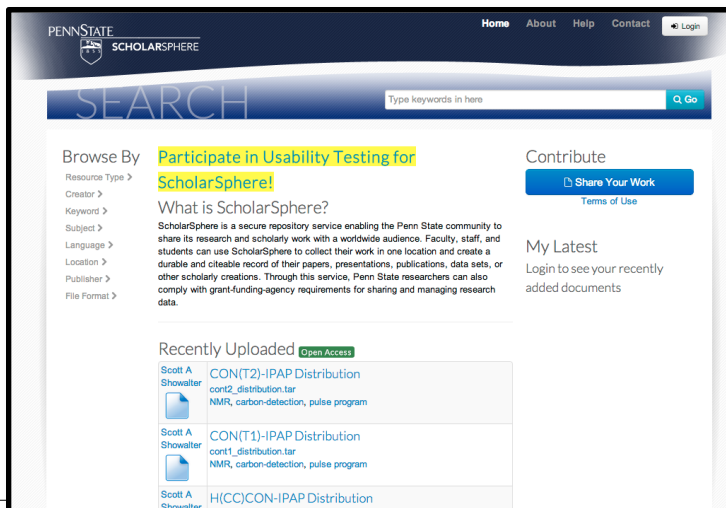


1. Support for any kind of record or metadata
2. Object-specific behaviors
 - Books, Articles, Images, Music, Video, Manuscripts, etc.
3. Tailored views for domain or discipline-specific materials
4. Easy to augment & over-ride with local modifications

Hydra Heads of Note



Avalon & HydraDAM for Media



UCSD DAMS



BPL Digital Commonwealth

See a full list at:
<https://wiki.duraspace.org/display/hydra/Partners+and+Implementations>



Hydra and Fedora/DuraSpace

- Hydra designed to work with and complement Fedora
 - And is somewhat dependent on it
- Hydra project has tracked Fedora 4 developments
 - Every release has been tested against Hydra (and Islandora)
 - New Hydra implementations now recommended to start with Fedora 4.0
 - Hydra becoming RDF-native alongside Fedora 4
- DuraSpace
 - Sit on Hydra Steering Group
 - Hydra is working with DuraSpace on support services

Hydra @ Hull

- Implemented during 2011
 - Read interfaces went live Sep 2011
 - Create, update and delete functions went live Feb 2013
- Developer had to learn Ruby technology base from scratch
 - But came to want to do everything based on it
- Upgraded in 2013
 - Now planning move to RDF, probably in 2016
- Use cases expanding – and including data!
- <http://hydra.hull.ac.uk>

Hydra @ Hull


[Digital Repository](#)
[Home](#)
[About](#)
[Contact](#)
[Sign in](#)


Welcome to Hull's digital repository, Hydra

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

All Fields

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.

Browse

You can browse our repository by...

Based on our latest statistics these are currently the most popular searches in the repository. [Sign in](#) to discover the most popular searches for University only content.

[Resource type](#)

[Subject](#)


Collections

Have a look at our collections...

[Collections](#)

Published in Journal of digital information, 2012
ISSN (Electronic) 1368-7506
Publisher Texas Digital Library
Volume 13
Issue 1

Download

 **Article**
(3 MB pdf)

QR code link to this page



Person Pope, P. (Author)

Subjects Population census; History of marine animal populations; Fishing effort; Cod fishery

Rights Creative Commons Licence: Attribution-Noncommercial-Share Alike 2.0 UK: England and Wales. See: <http://creativecommons.org/licenses/by-nc-sa/2.0/uk/>

Citations (a) The dataset: please cite as follows: P. Pope, ed. 'Newfoundland, 1675- 1698' in M.G Barnard and J.H Nicholls (comp.) HMAP Data Pages (www.hull.ac.uk/hmap); (b) Supporting documentation: please cite as follows: P. Pope, 'HMAP dataset 6: Newfoundland, 1675-1698, Supporting Documentation', in M.G Barnard and J.H Nicholls (comp.) HMAP Data Pages (www.hull.ac.uk/hmap)

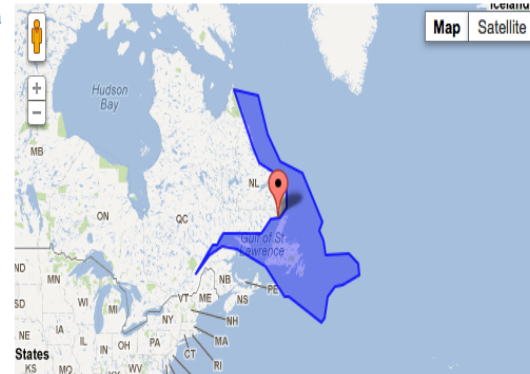
Description Fishermen, settlers and cod catches in 17th-century Newfoundland.

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 kml 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




Downloads


Database - ASCII format
(zip)


Database - Access 2000
format
(zip)

Database - csv format
(zip)

 Documentation - PDF format
(pdf)

Documentation - text format
(.txt)

 Documentation - Word
(.doc) format
(doc)

 Newfoundland-Labrador Shelf - kml file
(499 KB vnd.google-earth.kml+xml)
[View as map](#)

QR code link to this page



Reflections

- Only could have provided local solution through collaboration
- Hydra can't do everything, but provides capability and confidence we can adapt and implement solutions to meet needs, now and in the future
 - Our digital curation journey is ongoing, and we know where we are going
- Relationship between Hydra and Fedora is vital
 - Community support required for each and across them
- Success has come through good software design and patterns as much as from ability to address digital curation use cases

Looking ahead

- RDF!
- Hydra has been successful in stimulating activity across institutions in a dispersed way
 - Now looking to explore better cohesion across all Hydra initiatives
- Locally, do we address additional use cases through the same of separate Hydra heads?
 - Born-digital archives
 - Multimedia content (cf. Avalon)
- Development of Hydra community in Europe
 - Hydra Europe (23-24 Apr) and Hydra Camp (20-23 Apr)
 - LSE, London, UK

Thank you

c.awre@hull.ac.uk

(with thanks to Tom Cramer and Simon Wilson for use of their slides)