

CASE STUDY

A Digital Curation Centre Case Study
November 2013



Storing and sharing data in an institutional repository – Hydra@Hull

Chris Awre, (University of Hull) and Monica Duke, (DCC)

Section of the How to guide that this supports
Data Repositories

Introduction

Hydra@Hull is the institutional repository running at the University of Hull. Hydra is a repository solution based on the Fedora software combined with other open source software components, developed over a number of years by the Hydra Partners. The repository has a flexible content model and hosts Hull's research outputs alongside other content. The system offers workflows to enable deposit from other systems, content access control, collection management and navigation features, and can support data citation. The repository has recently been used to hold research data.

Background context

Hydra was developed as a reusable framework for multi-purpose, multi-functional, multi-institutional repository-enabled solutions for the management of digital content collections, in a project started in 2008 by the University of Hull, Stanford University, University of Virginia, and DuraSpace. The founding Hydra Partners were driven by a common vision when starting the Hydra project: that a repository should be an enabler for managing digital content collections, not a constraint or simply a silo of content. Their aim was to identify a repository solution that could be applied flexibly to meet the requirements of different use cases and content types. This led to the idea of a single repository with multiple points of interaction – Hydra – and the concept of individual 'Hydra head' solutions. The Hydra Project is informed by two main principles: 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; and 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.

The University of Hull has been a partner in the development of Hydra, and has deployed the repository to contain a variety of content. A solution to managing research data is being considered at Hull as at other institutions, and Hydra has recently started hosting some of the University's research data outputs.

The Hydra Philosophy and Community Structure

Hydra is a repository solution: it can be taken as is and used for managing digital content collections, recognising that work is involved in meeting your particular needs.

Hydra is a community: the project is based on the principle that working together enables us to address more use cases than working individually.

Hydra is a technical framework: the components can be applied as required to meet repository needs.

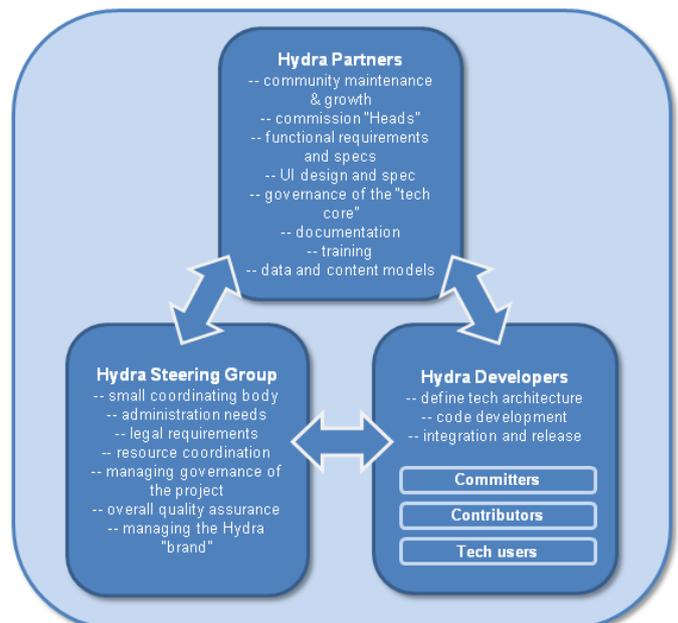


Figure 1 The Hydra Community.

The dataset display demonstrated in the example above is based on the History DMP project. Guidance on how to cite the dataset is also displayed.

Through integration with the institutional CAS system Hydra at Hull allows granular access control levels (e.g. internal, open, restricted to groups), which can be expanded to use additional criteria within the local directory service where available.

A common workflow can be employed although specialised workflow instances have also been developed. Deposit of content from other systems (e.g. research outputs via CRIS and publications from Open Journal Systems) are in development alongside direct deposit workflows into Hydra. At Hull all workflows have been configured so that content deposited is queued for QA purposes prior to publication in the repository.

Lessons learned

The implementation of Hydra at Hull was started by launching the search interface, moving on a few months later to make available *create*, *update* and *delete* functions.

As an early adopter, Hull had at times some painful learning experiences with early upgrades to subsequent versions of Hydra. This experience led to focused development by the Hydra community, and has resulted in a more streamlined development and upgrade process.

Future plans for Hydra at Hull include the addition of image management, integrating the repository with other library search services for staff and students, and ongoing work on integrating Fedora with the CRIS (Hull is using the Converis research information management system).

"Hull's experience of being a Hydra partner has been a fruitful one, and we have been able to contribute in a variety of different ways (code, architecture, Web site maintenance, documentation). It has also at times been a painful birth as we have seen the system come to life. The work has been well worth it"
Richard Green, consultant to Library and Innovation at the University of Hull. [1]

"Working with Hydra has been a steep learning curve at times, not least because the community and technology have been developing fast over the past 18 months, and Ruby on Rails training had been a

necessary, though invaluable, starting point. Once up and running, the flexibility of the framework and agility in making changes has enabled rapid progress to be made."

Caroline Webb, Spoken Word @ Glasgow Caledonian University. [1]

Other implementations of Hydra in the UK

- London School of Economics (LSE) has used Hydra components selectively (ActiveFedora (part of the Hydra-head plugin), Solrizer, Hydra Community input), creating a local user interface solution that suits their current needs. LSE became a Hydra Partner in November 2012.
- The University of Oxford is exploring use of Hydra to manage interaction with data objects over DataBank, rather than Fedora, for example for deposit of materials and image viewing.
- At Glasgow Caledonian University, Hydra is being adapted to support the Spoken Word service, a repository of recordings from the BBC that are made available to support teaching and learning. The focus has been on delivering the audiovisual material using progressive download or pseudo-streaming, providing easy access to them through embedded media players, and holding all the metadata that can be useful for such materials.

Future Developments

Gemification

- Use of Ruby on Rails leads to discrete pieces of software being produced, labelled 'gems'. As Hydra Partners have adopted the software they have created different workflows and functions as part of Hydra. These local adaptations are now being extracted as gems for sharing with others to avoid duplication of effort.

Community development

- The Hydra Partners are keen to work with others. The initial three Partners have developed into 20 Partners as of October 2013. Whilst partnership is a commitment to the sustainability of Hydra, engagement can be at any level.

Further information

Where to get more info

- The Hydra project main website:
<http://projecthydra.org>
- Browse Hydra@Hull
<http://hydra.hull.ac.uk>

Technical details

- Technical information on the wiki <https://wiki.duraspace.org/display/hydra/Technical+Framework+and+its+Parts>
- Github
<https://github.com/projecthydra>

Hydra support and communication is managed in several ways including the website, a wiki, Skype calls, mailing lists, face to face meetings and an IRC channel.

Why not join the user / developer communities?
<https://wiki.duraspace.org/display/hydra/Connect>
Besides the international Hydra community, European Hydra Partners are in the process of developing a European support network.

References

[1] Chris Awre. "Hydra UK: Flexible Repository Solutions to Meet Varied Needs". November 2012, Ariadne Issue 70 <http://www.ariadne.ac.uk/issue70/hydra-2012-11-rpt>

Follow DCC on Twitter:
@digitalcuration and #ukdcc