22 July 2014

Present

<table>
<thead>
<tr>
<th>Institution</th>
<th>Attendees</th>
</tr>
</thead>
<tbody>
<tr>
<td>University of Hull</td>
<td>Chris Awre, Richard Green, Simon Lamb</td>
</tr>
<tr>
<td>University of York</td>
<td>Liz Waller, Julie Allinson, Jenny Mitcham</td>
</tr>
<tr>
<td>Lancaster University</td>
<td>Masud Khokhar, Michael Dunne, Hardy Schwamm</td>
</tr>
<tr>
<td>University of Durham</td>
<td>Matthew Phillips, Paul Dixon</td>
</tr>
<tr>
<td>London School of Economics</td>
<td>Neil Stewart</td>
</tr>
</tbody>
</table>

Background

The aim of the meeting was for those working with Hydra and those proposing to adopt Hydra to come together and discuss how we can work together for mutual benefit. This is in keeping with the operation of the Hydra community in general, where everyone contributes what they can and benefits from the work of others.

A preliminary list of topics to address was circulated prior to the meeting. These were:

- Current state of Hydra - a brief update overview of developments to ensure discussions are well-informed
- Moving to Hydra - implications, routes, how to mutually support each other in doing so
- Working together on Hydra developments - how do we make the most of the resources we have?
- Hydra and the HEFCE Open Access policy

In a review of other areas those attending wished to cover the following were also highlighted:

- Hydra as an institutional repository
- Hydra for research data
- Hydra integration with a CRIS

Drivers for moving to Hydra and attending the meeting were also flagged up:

- The benefit of forming a local community, and specifically within this a local group of developers to bounce idea around within. Hydra is a different IT environment to the regular institutional IT environment, and some adjustment will be necessary. Support for this will be beneficial.
- The current emphasis on managing research data and the need to identify a flexible repository option that can accommodate shifting requirements in this area.
- Understanding better the relationship between a CRIS and a repository so as to better define their respective architectural and functional roles.
- Understanding how Hydra fits within a broader digital archiving architecture.
- A desire to bring multiple systems managing content together within a single system that can accommodate different types of digital content.
- Understanding the role of a central service and the cultural change this may require.
- A desire to better embed the digital library within other library services.
- How Hydra collections can feed into discovery tools such as Primo (in use at Lancaster, York, Durham and LSE).
- A desire to make better use of controlled vocabularies to describe digital content.
- Interest in embedding ORCID identifiers for authors of materials within Hydra.
- Interest in developing a common approach and implementation of a page turner to meet different institutional needs.

It was noted that an initial library commitment to Hydra at Lancaster and Durham had the scope to move beyond this within the institution. Hydra allowed the Library to take the lead on managing digital collections in a way that could be applied to other areas of the University.

Current state of Hydra

CA presented a verbal update on the current state of Hydra and its development, both as a technology and as a community.

Implementing Hydra

Four options for implementing Hydra were described:

- Implement the Hydra gem and build your own Hydra head from this - the Hydra gem is a tested combination of all of the core component Ruby gems that are used within Hydra. It can be used to provide a vanilla implementation of Hydra. A fuller Hydra head can be built from this using the tools within the different gems. This option provides the greatest flexibility, but also requires the greatest amount of effort in defining specific workflows and interfaces to meet local needs.
- Implement an existing generic Hydra head gem, for example Sufia - Penn State University implemented their self-deposit Hydra head, ScholarSphere, based on the component core gems. They then generalised this as Sufia so others could make use of it as the basis for their own Hydra head. For example the Curate Hydra implementation at University of Notre Dame is based on Sufia. Hence, if you have a set of functional requirements that are matched by gems like sufia then using an existing gem and adapting it to meet local needs is a route that can be followed. A lot of the development work has been done, and the main effort is in assessing and carrying the local adaptation required.
• Take an existing Hydra head and adapt to local needs - this option offers a broader range of options than using a gem like Sufia as not many sites have yet been able to generalise their implementations. However, the approach is much the same, in that an existing Hydra head is adapted to local needs. Matching functional requirements is key to establishing which head is the best match. The majority of Hydra heads are available openly on github allowing use of the code, and the main effort is in unpicking any localisation so that your own implementation can be built from the head. For example, Hydra@Hull is an example of an IR head that can support self or mediated deposit to capture varied digital collections. The code is available on github, but some effort would be required to remove Hull's localisations to make it suit your own needs.
• Use a Hydra turnkey solution - The intention of the Hydra community is to produce and share a number of turnkey solutions for specific areas of repository activity. The single solution currently available is Avalon, a Hydra head that incorporates Matterhorn software to provide a complete package for the ingest, processing, delivery and archiving of audiovisual content. There is also an initiative based at Notre Dame called Hydramata that is planning to build a flexible Hydra institutional repository turnkey solution, encompassing much of the learning from existing Hydra head developments.

Key to any implementation of Hydra is the intellectual effort in organising the collections to be managed. This will then inform the workflows and interfaces required, which will then feed into your functional requirements and the implementation approach that will work best. A use case of managing objects sitting within multiple systems currently was highlighted: in such a case it is necessary to think of the workflows involved in bring those objects together and how local or generic these are likely to be.

Working together

It was agreed to initiate steps to help the institutions present work together on Hydra developments. The following steps were agreed:

• Set up a wiki space on the main Hydra wiki (RG)
• Consider the role of Docker and a sandbox for common developments to help bounce ideas around and get feedback (All)
• Develop a contents list of points that can usefully be taken into account when getting going with Hydra (RG to start, and then All)
• Break-up Hydra@Hull into a series of functional area descriptions to facilitate awareness of what is in the Hydra head and what might be useful to have (Hull)

Alongside the areas listed above, it was also felt to be useful to explore a Hydra exit strategy to provide institutional assurance that any repository would last beyond the current technology.

The common interest in the Internet Archive page turner at Hull, Durham and LSE will be explored further and a common implementation developed. It was noted that there are other page turning initiatives within the Hydra community it will be useful to keep in touch with (e.g., Stanford and Indiana).

The use of controlled vocabularies is an area that the Boston Public Library has spent some time exploring. They have created the questioning_authority gem to allow them to embed any structured vocabulary available through a RESTful API.

The use of Hydra for research data has been piecemeal to date, largely as the precise needs are still unclear. It was noted that a number of Hydra heads are able to manage research data in varying degrees (including Hydra@Hull) and that further developments would be forthcoming. Hydra is confident the system can be adapted to meet the variety of use cases that will emerge, and this will be an area to keep talking about to ensure we work together efficiently on it. Ongoing efforts within the N8 consortium to address research data (including York, Durham and LSE) were reported. Virginia Tech will be using Hydra to manage the outputs of a big data experiment from the autumn (as reported at OR2014).

There are a number of ongoing conversations about how Hydra fits best within a digital archiving solution. LSE has an architecture that includes preservation and access repositories (as does Stanford), but the precise place of Hydra within a digital archiving/preservation workflow is both a topic of ongoing discussion and a matter of local preference. Two working groups exist:

• The Hydra Archivists Working Group (HAWG) - focusing on an archival workflow view of Hydra and the link between Hydra and ArchivesSpace
• The Hydra Digital Preservation Working Group (HDPWG) - focusing on Hydra and the variety of digital preservation tools available, and working with Open Planets to identify ways of embedding these

York and Hull are both exploring how Archivematica can be used in this context.

Hydra and open access

The use of Hydra to meet open access requirements was discussed. Recent work at Hull has worked with two gems that could be used by any Hydra head:

• The irus_analytics gem to enable submission of statistics to IRUS-UK. This is work in progress but should be implemented by the end of August
• Enhancing and rolling out the blacklight_oai_provider gem to provide an alternative OAI interface to the PROAI service (to be released in due course once tested)

Hull is leading on the HHuLOA Jisc Open Access Pathfinder project, working with Huddersfield and Lincoln, and will be developing Hydra to help meet the HEFCE Open Access Policy. MK reported that Lancaster is also a partner in one the Pathfinder projects, the e2eoa project, and is also committed to producing a Hydra solution. Hull and Lancaster will work together towards this.

CA described work at the University of Notre Dame to develop an ORCID gem for Hydra. This would be made generally available in the autumn.

Next steps

It was agreed to hold a further meeting in October. Lancaster and LSE offered themselves as hosts for future meetings. All attendees were urged to join the hydra mailing lists (focusing on hydra-community and hydra-tech) to engage with ongoing discussions, and raise questions were required.