Hydra Community Annual Report 2016

table of contents

from the community

Throughout 2016, the Hydra Project saw a year of steady growth, major updates, and thriving Community activity.

On the Community side, the Project reached new heights in October with the largest ever Hydra Connect, the annual, worldwide community meeting, at the Boston Public Library, with more than 260 delegates from scores of institutions across the world. We saw a flowering of regional events, with locally organized gatherings of Hydranauts in the UK and across the US.

Technically, the Project continued to make great strides with five releases of Sufia (including a major update to 7.0), and multiple releases of Avalon (the Hydra solution for audio-video materials). Looking ahead, early 2017 will see the beta release of Hyku, the IMLS-funded effort to produce a "Hydra-in-a-Box" turnkey solution for libraries, archives and museums everywhere. Hydra's technical framework and functionality has never been richer, nor easier to employ.

Hydra continues to be a distributed, open, and inclusive project. Through its interest and working groups, the Project participants continue to push the envelope of best practices in everything from data modeling to service management, from managing geospatial assets to digital preservation.

Outside the Project, Hydra Community members are major contributors to and drivers in broader Community activities such as the Fedora Repository, the Portland Common Data Model (PCDM), and new initiatives such as the Open Science Framework. We continue to live by our slogan, "if you want to go far, go together."

Looking forward, we see 2017 as a watershed year for the Project. With the upcoming release of Hyku and a substantial refactoring of some of our gems, we anticipate a spike in interest and adoption from the broader cultural heritage and science data communities. We will be rebranding the Project, and reflecting the Community ethos and core principles that have carried us this far. We will continue to optimize how we organize the effort among the Hydra Partners – the backbone and owners of the Project – to get the best of both distributed innovation and contributions, as well as focus on a more centrally-organized effort to support efficiency and sustainability of the Project.

Hydra Connect 2016

Boston Public Library
October 3-6, 2016

Photo courtesy of Boston Public Library

community & partner structure

2016 saw five more institutions sign up as formal Partners. We welcomed UC San Diego, the University of York, Lafayette College, Washington University in St. Louis and the Digital Repository of Ireland bringing us to a total of 34.

- Boston Public Library
- Case Western Reserve University
- Columbia University
- Cornell University
- Data Curation Experts
- Digital Public Library of America (DPLA)
- The Digital Repository of Ireland
- Duke University
- Duoc UC
- DuraSpace
- Indiana University
- Lafayette College
- London School of Economics and Political Science
- Northwestern University
- Oregon State University
- The Pennsylvania State University
- Princeton University Library
- Rock and Roll Hall of Fame and Museum
- Royal Library of Denmark
- Stanford University
- Tufts University
- University of Alberta
- University of California, San Diego
- University of Cincinnati
- University of Hull
- University of Michigan
- University of Notre Dame
- University of Oregon
- University of Virginia
- University of York
- Virginia Tech
- Washington University in St. Louis (WUSTL)
- WGBH
- Yale University

The Project receives support, not just from Partners, but also from many others within the Community. At the time of this report, there were 65 institutions and more than 250 individuals holding contributor license agreements.

partner profiles: from the partners

Digital Repository of Ireland

What did DRI need from the software it adopted?

DRI was created to provide a national digital repository for Ireland's social, cultural, and heritage data. We were tasked with the dual roles of providing long-term digital preservation of these collections as well as access to the digital objects to the public, students and scholars. To address both these mandates, we needed a flexible, robust, scaleable software platform on which to implement our specific requirements. After a period of evaluation and experimentation, we decided that Hydra provided us with the best set of tools to deliver our repository.

Why did DRI choose to become a Hydra Partner?

DRI has been developing with Hydra since early 2012, and has been involved in the Community since then. We have presented our work at various Open Repositories and Hydra Connect meetings, as well as hosting Hydra Europe events in Dublin. In a way we have been acting as partners for many years without formally being one. In 2016, DRI was awarded long-term funding from the Irish government, and so we decided that it would be appropriate now to sign up to partnership and formally contribute to the project where we can.

What interesting features have DRI implemented in its repository?

The DRI repository was formally launched in June 2015, having already migrated to Fedora 4, meaning we were one of the first sites worldwide to have Fedora 4 in production. However, we do not house our digital assets in Fedora, rather we ingest the metadata and a reference to the file assets into Fedora, storing the files on an S3 compatible storage layer provided by CEPH. Recently we took advantage of this separation to migrate our digital objects to the Moab standard without needing major changes to the data model within DRI. We also had a mandate to provide a multilingual repository (English and Irish languages) at launch and so built everything with this in mind from day one.

What's next for DRI?

As we are now running a live service, changes are more incremental. However, we continue to track developments in the Hydra Community. We have a legacy data model that supports a wide variety of file formats and metadata standards. Moving to the Portland Common Data Model would be a substantial body of work but might allow us to leverage other developments in the Community. As an active research project, we continually publish reports and papers about our work so keep an eye on our website, find us at a conference or just drop us an email. We're always on the look out for smart collaborators who are doing interesting things and the Hydra community has no shortage of those!

https://repository.dri.ie/

Map of Hydra Partners and Adopters, WorldWide, Fall 2016

Red = Partners
Blue = Adopters
Green = Solution Bundle Users

partner profiles: from the partners

Open Access at Oregon State University

Oregon State University (OSU) Libraries and Press and University of Oregon (UO) Libraries began working with Hydra in 2012 as part of the joint project, Oregon Digital, for cultural heritage content. The goal of this project was to provide a better media viewing experience for users and more flexibility in how our system dealt with metadata.

Oregon Digital is an example of the push to make digital collections available widely and promote open content. In addition, OSU's institutional repository, ScholarsArchive@OSU, has been opening access to scholarly materials from OSU for over a decade as the venue for implementation of our campus-wide Open Access Policy and for research datasets in need of a home for preservation and dissemination. In 2016, ScholarsArchive@OSU began a migration path towards Hydra, allowing OSU Libraries and Press to align our repository applications onto a single platform, with a single set of technologies and skills, and to prioritize the user experience.

Where does Hydra fit into all of this? In some ways, the Hydra applications we are using (Oregon Digital, Hyrax) are not "magic bullets" that solve some unsolved problem of open repositories. Rather, the strength of the Hydra Community has been the draw for OSU - robustness in software development, ideation, and communication. We have felt, as an institution, that our ability to participate in the process of creating and supporting open content (be it collections, articles, or data) has been catalyzed by the work and workings of the Hydra Community. The extent to which the community works together towards common goals of access to our cultural heritage has been the magic bullet for OSU.

http://oregondigital.org/
http://ir.library.oregonstate.edu/
https://github.com/OregonDigital/oregondigital
https://github.com/osulp/Scholars-Archive

governance, finance & expenditures

The Steering Group, which has a specific responsibility for stewardship and central administration of the Project, has continued its pattern of monthly meetings online supplemented by a small number of face-to-face meetings: an annual winter planning meeting and meetings alongside other events and conferences that members of the group have attended.

The Steering Group manages the Project's financial and legal affairs through a Memorandum of Understanding (MoU) with a fiscal sponsor. Following a successful first year with DuraSpace in that role, a new MoU was signed for another year's support. We are now working with DuraSpace to align the period of future MoUs with their financial year in order to streamline our financial processes. A summary financial report is found on page 8 of this report.

In 2016, Hydra embarked on a fundraising campaign to accumulate funds intended to address a number of important tasks that were unlikely to be completed by Community volunteers alone - among these a complete redesign of the Project website, Community management, Marketing and Communications, and Technical Coordination (including documentation). We are grateful to more than 20 of our Partners who together contributed over $60,000.

Using funds raised in 2015 for this purpose, one of the Project's initiatives last year sought to trademark Hydra's name and logo in the US, Canada and the European Union. During this process we became aware of a German company that has a wide ranging trademark on the use of "Hydra" for computer software and their claim to the word considerably predates ours. Our lawyers, backed by a second opinion from counsel at Indiana University, advised that we could not successfully fight this requirement and this resulted in the rebranding program which is currently in progress.

strategic funding priorities

After extensive conversations throughout 2014 and 2015 within the Community, the Hydra Steering Group in 2016 identified four areas for potential spending confirmed at the spring 2016 Power Steering meeting:

- Project administration
- Community communications / marketing
- Technical coordination
- Building a financial reserve

At their March 2017 meeting, Hydra Partners discussed the spending priorities for the coming year.

accounts
1st January 2016 - 31 December 2016

	Income ($)	Expenditure ($)
Fundraising and donations	66,045	
DuraSpace staff time		4,569
Outsourced financial, accounting and tax work		6,935
Contractors (website redesign)		4,175
Legal and IP services (mainly trademark work)		13,200
Office expenses		315
Website hosting		73
Total	**66,045**	**29,267**

Net income from above	$36,778
Less invoices outstanding	($6,045)
Cash balance forward from 2015	$12,499
Cash balance 31 December 2016	**$43,232**

progress on key initiatives

Avalon

Development of Avalon Media System, the Hydra solution bundle focused on audio and video access, continued in 2016, led by Indiana University and Northwestern University, with support from the Andrew W. Mellon Foundation. Avalon version 5 was released in mid-2016, with new features for creation of private and shareable playlists, enhanced access control capabilities, and improved accessibility, including support for captions, along with tools to help institutions migrate from the Variations digital music library system to Avalon. Avalon version 6, released in March 2017, provides the same feature set as 5.x but runs on Fedora 4 rather than Fedora 3. Planning is underway for piloting a cloud-hosted SaaS Avalon offering later in 2017, along with additional opportunities for Community engagement in development.

HyBox

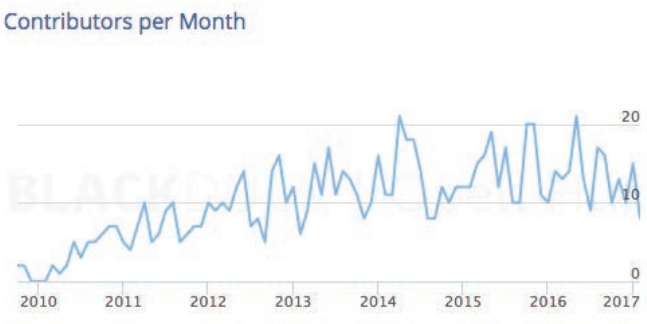
The Hydra-in-a-Box project, the Institute of Museum and Library Services (IMLS)-funded effort to produce a "turnkey" Hydra repository application for institutions large and small, hit its stride and midpoint in 2016. It completed an intensive, user-centered design process and launched engineering in earnest, with frequent releases and enhancements being made throughout the Hydra codebase. It also made significant investments in a cloud-ready platform, and Hydra for the first time can now run in a multi-tenant environment — essential positioning for the future where many institutions will be running repositories as remote services. With contributions from more than ten institutions, the Project has been taken up as a shared initiative in the Hydra Community. A Beta release that pulls the essential features into an easily-deployed package is due in the Spring of 2017 under the name of Hyku.

Sufia

2016 was a productive year for Sufia, with five new versions released. Included among these new releases was Sufia 7, the product of a set of Community sprints with the goal of adding the concept of works to Sufia based on the Portland Common Data Model. Subsequent releases have added the ability to create multiple work types, to create hierarchical works (or "complex objects"), to organize works into administrative collections, and to support configurable workflow (as needed for mediated deposit workflows). The Community invested significantly in the Sufia 7.0.0 release, having worked on the release for a little over a year, comprising 1,181 commits from roughly 60 developers.

Community Code Contributions

Project Hydra has had 14,356 commits made by 134 contributors representing 66,788 lines of code, has a well established, mature codebase, is maintained by a very large development team with increasing Y-O-Y commits and took an estimated seven years of effort (COCOMO model) starting with its first commit in October, 2009.



New Website

A significant initiative during the past year has seen a complete redesign of the Project's website. The new site is fully responsive (designed with a flexible layout to support many devices) and its content has been rewritten to better reflect the Project's growth. The new site will be launched under our new name as soon as the rebranding program is complete; a review of our wiki will follow.

Vendors

Hydra has long recognized the value that a robust vendor community brings to an open source project. Commercial firms help with implementation, consulting, training and maintenance, and can often do the heavy lifting and bootstrapping for institutions that are seeking to start up with the project. In 2016, we saw more than a dozen companies and individual contractors provide Hydra-related services to the Community, with strong clusters in the US and UK, and we see this growth continuing into the future.

Hydra Camp at University of California , Santa Barbara, February 2016

2017 initiatives

Hyku

Hyku will be useful to any institution of any size managing collections of digital assets, particularly libraries, archives, museums, and historical societies with collections of cultural heritage materials. The Hydra-in-a-Box team,with the support of IMLS, will produce a repository application that is easy to install and easy to maintain over time, and to provide software that is compelling to hosted service providers. In these ways, Hyku will be a solution for any institution seeking a next-generation digital repository solution for its digital collections, particularly if technical resources and infrastructure are limited. Services will also be available for Hydra hosted as a cloud service.

Advanced Hydra Camp

In response to training requests from the Community more advanced level Hydra training will be offered, starting with the May 2017 Advanced Hydra Camp in Minneapolis, Minnesota.

Hyrax

In order to streamline development and take advantage of features across Community development, Curation Concerns and Sufia will be integrated into one architectural level and Ruby gem, Hyrax.

events

The Community continues to find new ways in which to collaborate and grow. The past year has seen a wide range of mixed-content meetings ranging in scale from the annual, international Hydra Connect in Boston, which attracted more than 260 delegates, to regional meetings and trainings in the US and the UK. Our developer Community has also held three "Code Congresses" (Ann Arbor, Michigan; State College, Pennsylvania; and San Diego, California) to progress specific areas of the code base.

Face-to-face meetings can be a challenge with scarce resources, particularly time and money. In July 2016 we held the first Hydra Virtual Connect, an opportunity for Project participants to gather online to touch base on the progress of Community efforts at a point roughly halfway between face-to-face Hydra Connect meetings. A second event is being planned for 2017.

West Coast Regional Group

UC Santa Cruz hosted the second West Coast Regional Group meeting on February 10, 2017. All near the West Coast who are interested or invested in Hydra were invited to attend, including Hydra implementers, developers, and interested onlookers. The meeting was organized as an unconference in order to create an informal collaborative atmosphere.

Outcomes:

- 28 individuals from 11 organizations in California and Oregon attended the meeting
- Unconference session topics included audiovisual media, metadata ingest workflows, linked data implementation, Hydra-in-a-Box and Hyrax, and lightning talks
- The group plans to hold a third meeting in 2018

Hydra Connect Mentoring Program

Our first ever mentoring program kicked-off at HydraConnect 2016 with 19 mentor-mentee pairs; 15% of attendees participated. The goals of the program are to:

- Bring people together to build individual capacity and empowerment
- Increase diversity in participation, interaction and presentation at Hydra Connect
- Break down barriers to entry and provide a welcoming orientation to the Hydra Community
- Strengthen the Hydra Community by learning from mentees and diverse voices

Penn State Developer Code Congress

During the week of Sept. 19-23, a group of 16 developers from eight institutions gathered in State College to work on three initiatives with the goal to provide desired functionality back to the Community.

Workflow
The workflow team extracted the database-backed workflow implementation from Notre Dame's Sipity application into CurationConcerns, using Princeton's Plum workflow as an initial target for modeling multiple configurable workflows.

Fedora Import/Export
The import/export team started working on a BagIt implementation design including Bag Profile support for APTrust and MetaArchive and transferring between Fedora implementations.

Admin Dashboard
The administrative dashboard team added a configurable, extensible admin dashboard to CurationConcerns. The dashboard design allows flexible control over what appears in the dashboard menu, and in what order, in addition to what views are rendered and what data sources are used.

United Kingdom

Members of the Hydra Community in the UK have met several times to discuss their existing and potential involvement with Hydra. In June 2016 these nine organizations signed a consortium agreement which sets out their intention to work together in order to further the take-up and use of Hydra and Fedora installations in their region. Later in the year, the group sought to broaden participation by inviting other European users to join them. The intention is that, going forward, there will be three meetings each year.

Developer Code Congress (Concentrate)
Penn State, September, 2016

planned training & events for 2017-2018

Winter 2017	Sandi Metz – Practical Object-Oriented Design, San Diego, January 25-27
Winter 2017	Workshops – Code4Lib, Los Angeles, March 6
Spring 2017	Hydra Camp, Atlanta, April 17-20
Spring 2017	Advanced Hydra Camp, Minneapolis, May 8-10
Spring 2017	Hydra Camp, Europe, Details TBD
Summer 2017	Workshops – Open Repositories, Brisbane, June 26-30
Summer 2017	Virtual Hydra Connect, Dates TBD
Fall 2017	Hydra Camp, Austin, Dates TBD
Fall 2017	Hydra Connect, Evanston, November 5-9
Spring 2018	Hydra Camp, Tentative location: St. Louis, Dates TBD
Spring 2018	Hydra Camp, Europe (European partner input welcome)
Spring 2018	Advanced Hydra Camp - TBD
Fall 2018	Hydra Camp - TBD

West Coast Regional Meeting, University of California, Santa Cruz, February, 2016

progress on digital preservation

Digital Preservation Tools

A number of Hydra Partners are investigating the use of our software for dealing with the long-term preservation of research data, audiovisual collections, and other digital information. In the UK, the Universities of York and Hull have been integrating the open-source preservation system Archivematica into their Hydra workflows. Their Jisc-funded project was a finalist in the international Digital Preservation Awards 2016. In addition to the actual integration work, York and Hull commissioned Archivematica's development staff to add a number of helpful features to their software, and worked with the National Archives in the UK to extend the research data capabilities of their PRONOM file format identification tool.

Fedora Repository

Throughout 2016, the Community continued to invest and co-evolve with the Fedora Repository, the major open source component at the core of Hydra's architecture. There continues to be a deep strategic and structural link to Fedora, with 35 Hydra institutions active in Fedora's membership, and two committers who are substantial contributors to both projects. Fedora's roadmap and enhancements are heavily informed by Hydra's direction (and vice versa), and the Fedora 4.5.0 through 4.7.0 versions were released in lockstep with corresponding updates to Hydra's gems.

contributions to the annual report

Content

Chris Awre
Tom Cramer
Jon Dunn
Karen Estlund
Dermot Frost
Michael Giarlo
Richard Green
Steven VanTuyl
Andrew Woods

Images

Hydra Connect 2016 (cover photo), photo courtesy of Boston Public Library

Ned Henry, West Coast Regional Meeting

Suprita Srinivas, West Coast Regional Meeting Dinner
contributed by Chrissy Rissmeyer

Heather White, Penn State Hydra Developer Code Congress (Concentrate)
contributed by Karen Estlund

Design & Editing

Karen Hackett
Michael Tribone