Hydra: One Body, Many Heads for Digital Repositories

Jon Dunn, Indiana University
Karen Estlund, Penn State University

EDUCAUSE 2016 - Anaheim - October 26, 2016
What is Hydra?
What is Hydra?

A community

And an open source technology framework

for digital repository applications
Digital Repositories

- Content management / digital asset management for the long term
  - Access
  - Digital preservation
- Around since the early to mid 2000s
- Variety of technology/product options
- Open source: Eprints, Fedora, DSpace, ...
- Commercial: Digital Commons, ContentDM, ...
Repository-Powered Approach

ETDs (Theses)  Books, Articles  Images  Audio visual  Research Data  Maps & GIS  Documents

Digital Repository

Scalable, Robust, Shared Management and Preservation Services
One Body, Many Heads...

Scalable, Robust, Shared Management and Preservation Services
Hydra Community
Hydra Community Framework

- Hydra Partners
- Hydra Power Steering
- Hydra Interest Groups
- Hydra Working Groups
- Hydra Steering

Hydra developers
Hydra Partners as of October 2016

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

Getting involved with Hydra can be at different levels:

- Adopter
- Developer
- Partner

Opportunities for non-technical leaders:

- Community Advocacy
- Strategic Planning
- Meetings – Facilitate, Record and Speak Out!
Wide Community Support

- Community sprints and development events
- Tech calls
- Google groups
- Working and interest groups
- Power Steering and Steering
- Partners and Hydra adopters Worldwide
- Code of conduct [https://wiki.duraspace.org/display/hydra/Anti-Harassment+Policy](https://wiki.duraspace.org/display/hydra/Anti-Harassment+Policy)
Financial Framework

- **Individual Institution**
  - Software Free/Open Source
  - Staffing (varies)
  - Training / Community Events / Developer Congresses
  - Storage
  - Vendor support
  - Proprietary Integrations?
  - Fundraising for Hydra Project

- **Hydra Connect Conference**
  - Self-organized/funded
  - Try to keep costs low

- **Hydra Project**
  - Spending Plans Approved By Hydra Partners
  - Financial home currently through MOU with DuraSpace
Legal Framework

- Apache 2.0 license
- All contributions are made under a Contributor Licensing Agreement, granting a perpetual license to the Hydra Project
- All documentation is shared using a Creative Commons Attribution Share-Alike License
- Partners sign an MoU highlighting expectations and requirements of Partnership
Hydra Partners

OR = Open Repositories Conference
Hydra Partners and Known Users

OR = Open Repositories Conference
Hydra Partners and Known Users

OR = Open Repositories Conference
A Worldwide Presence
Hydra Showcase
Hydra in South America
Hydra at a Public Library
Hydra as the Digital Repository of Ireland
Hydra at Hull
Self-Deposit at the Royal Library of Copenhagen

Digital Archive for Student Assignments contains assignments passed at the University of Copenhagen.

The archive offers you inspiration when writing assignments - get ideas for subjects, see examples of assignment structures, or simply read an interesting title.

Search for student assignments from the search bar above or upload your own assignments by logging in.

Enjoy.
Hydra Framework
Major Hydra Components

- hydra-head Rails Plugin (CUD)
- Blacklight (Read Only)
- Fedora
- Solr
- Solrizer
Fedora as an Implementation (4.7+)

- REST Framework
- Fedora Services
- ModeShape

Access & Preservation Services

Repository Services

Caching, Clustering & Storage Services

Storage (Objects and Datastreams)
Fedora 4 Core Features and Standards
1. Create/Read/Update/Delete - Linked Data Platform ✔
2. Versioning - Memento
3. Authorization - WebAC ✔
4. Batch Operations
5. Fixity
Fedora 4 - Web Resources

Everything is a web resource with a URI

Resources have properties expressed as RDF triples

Resources can contain other resources (containers) or files (binaries)
http://www.slideshare.net/jpstroop/a-more-worthwhile-sufia-now-with-pcdm/
Additional Gems:

- Controlled Vocabs
- SHARE
- LTI Integration
- Derivatives

Hydra Application

GEO | Books | ETDs | Media

Curation Concerns

Hydra Works

Active Fedora

Layers integrated via APIs
# Select Gems

<table>
<thead>
<tr>
<th>Project Hydra</th>
<th>Hydra Labs</th>
</tr>
</thead>
<tbody>
<tr>
<td>• Browse-Everything</td>
<td>• GeoMash</td>
</tr>
<tr>
<td>• Hydra-Editor</td>
<td>• IIIF Manifest</td>
</tr>
<tr>
<td>• Hydra-Derivatives</td>
<td>• Batch Edit</td>
</tr>
<tr>
<td>• File Characterization</td>
<td>• ORCID</td>
</tr>
<tr>
<td>• Opinionated Metadata (XML)</td>
<td>• HyBag</td>
</tr>
<tr>
<td>• Role Management</td>
<td>• Share-Notify</td>
</tr>
<tr>
<td></td>
<td>• Questioning Authority</td>
</tr>
</tbody>
</table>

What Level of the Stack?

- Sufia
- Curation Concerns
- Fedora

MY APP
Possible Paths to Your Heads

- **Turn-key**
  - Use Existing Framework with Customizations
  - Contribute to Development of Feature(s) - Existing Framework
  - Contribute to Development of Feature(s) - Gem
  - Custom Head

- **Hydra-in-a-Box Repository Product**
  - Avalon
  - Sufia

- **SHARE Notify**

- **Curation Concerns**

**Requirements Gathering and Resource Assessment**
PCDM (Portland Common Data Model)

PCDM Ontologies
- PCDM Models Ontology
- PCDM Use Ontology
- PCDM Rights Ontology
- PCDM Works Ontology
- PCDM File Formats Ontology

PCDM Wiki


By Tony Webster from Portland, Oregon (Portland Oregon White Stag Sign) [CC BY 2.0 (http://creativecommons.org/licenses/by/2.0)], via Wikimedia Commons
Book Objects in PCDM

Acerca del chocolate

Inscripción
Custom Hydra Heads
Plum Use Case: Reorder Page Images - Editor
Plum Use Case: Structure - Editor
Plum Use Case: Bulk Label - Editor
Plum Use Case: Rights Statements - Editor

Visibility: Who should be able to view or download this content?
- Open Access: Visible to the world.
- Princeton University Library: Visible to all Princeton University Library users.
- Private: Only visible to you and your delegates.

Choose Rights and Holding Location

Rights Statement
- No Known Copyright

- In Copyright
  - In Copyright: Rights-holder(s) Unlocatable or Unidentifiable
  - In Copyright: Educational Use Permitted
  - In Copyright: Non-Commercial Use Permitted
- No Copyright: Contractual Restrictions
- No Copyright: Other Known Legal Restrictions
- No Known Copyright

Create Scanned resource  Cancel
Plum Use Case: Rights

<table>
<thead>
<tr>
<th>State</th>
<th>Complete</th>
</tr>
</thead>
<tbody>
<tr>
<td>Rights</td>
<td>No Known Copyright</td>
</tr>
</tbody>
</table>

Princeton University Library reasonably believes that the item is not restricted by copyright or related rights, but a conclusive determination could not be made. You are free to use this item in any way that is permitted by the copyright and related rights legislation that applies to your use.

Princeton University Library claims no copyright governing this digital resource. It is provided for free, on a non-commercial, open-access basis, for fair-use academic and research purposes only. Anyone who claims copyright over any part of these resources and feels that they should not be presented in this manner is invited to contact Princeton University Library, who will in turn consider such concerns and make every effort to respond appropriately. We request that users reproducing this resource cite it according the guidelines described at http://rbsc.princeton.edu/policies/forms-citation.

Princeton users with copyright questions or concerns should refer to http://www.princeton.edu/copyright/.

<table>
<thead>
<tr>
<th>Identifier</th>
<th>ark:88435/d504rk397</th>
</tr>
</thead>
<tbody>
<tr>
<td>Source metadata identifier</td>
<td>5184576</td>
</tr>
<tr>
<td>Collections</td>
<td>Islamic Manuscripts</td>
</tr>
</tbody>
</table>
Oregon Digital

https://github.com/OregonDigital/oregondigital_2
https://github.com/OregonDigital/ControlledVocabularyManager
Oregon Digital Use Case: Linked Data - Editor
Oregon Digital Use Case: Linked
Oregon Digital Use Case: Linked Data - RDF


Hydra Solution Bundles
Avalon

http://avalonmediasystem.org/
Avalon

- Hydra Solution Bundle for audio/video access
- Goals:
  - Easily installable and configurable
  - Serve a variety of use cases
- Hosted options emerging
- www.avalonmediasystem.org
Avalon Use Case: Archival Film

Repository | audio + video

Big Ten, Big Time, 1971

Main contributor
Northwestern University (Evanston, Ill.)

Date
1971

Summary

Contributors
Wright, Al G.; Kelso, William D.; Flamming, Bill; 1926-2007

Genre
Sports

Subjects
Northwestern University (Evanston, Ill.)--Football--History; Northwestern Wildcats (Football team); Big Ten Conference (U.S.)--Football--History; Agase, Alex, 1922-2007
Avalon Use Case: Online Exhibits
Avalon Use Case: Publishing
Avalon Use Case: Consortial Licensed Video
Avalon Use Case: Video Reserves in LMS
Sufia

Need help with research data management?

Consult the Libraries' guide on RDM services. Questions? Contact the Research Data Management Team.
Sufia Use Case: Self Deposit

My Dashboard

Upload

Create Collection

View Collections

View Files

Your Statistics
- Files you've deposited
- Views
- Downloads
- Collections you've created
- People you follow
- People who are following you

User Activity (last 24 hours)
- User has no recent activity

User Notifications
- User has no notifications

Manage Proxies

Authorize Proxy

Search for a user

Current Proxies

Transfer of Ownership

Sent
You haven't transferred any files

Received

Select files to transfer

Hello,
Sufia Use Case: Feature Researcher

Dr. Laura Guertin, Professor of Earth Science at Penn State Brandywine.

I have a pedagogical research agenda, looking at how new and emerging technologies can enhance teaching and student learning in introductory-level geosciences courses.

Why does Dr. Guertin use ScholarSphere?

At scientific conferences, more and more presentations are being given in a poster format, on 4 foot by 6 foot boards. A poster is an excellent way for me to capture highlights of my work with text and images, but it is difficult for colleagues and other researchers to take away all of that information after viewing the poster. Fortunately, I can include a QR code and/or URL to a ScholarSphere entry with my uploaded poster as a PDF file to be shared post-conference and to serve as a reference for years to come.

See the research Dr. Guertin has deposited into ScholarSphere.

For more about Dr. Guertin's teaching and research, please visit her Penn State Brandywine website and her ePortfolio.
Sufia Use Case: Analytics
Sufia Use Case: Citation Export

Export to Zotero

Exporting to Zotero is supported via embedded metadata. If Zotero does not automatically pick up metadata for files deposited in ScholarSphere, please report the issue via the Contact Form.
# Use Case: SHARE Notify

**Publisher:**

**File Details**

- **Depositor:** Laura Guertin
- **Date Uploaded:** 2014-03-31T13:04:00+00:00
- **Date Modified:** 2014-09-10T10:37:05+00:00
- **Audit Status:** passing
- **Characterization:**
  - File format: pdf (Portable Document Format)
  - Mime type: application/pdf
  - File size: 2783690
  - Last modified: 2014-03-31 09:04:13-04:00

- **Filename:** PenroseChapmanPoster.pdf
- **Original checksum:** 4d2829187ab758594a4fb7e7bc8e5e4
- **Well formed:** true
- **Valid:** true
- **Status message:** File header gives version as 1.4, but catalog dictionary gives version as 1.3
- **Page count:** 1

**User Activity**

<table>
<thead>
<tr>
<th>File was successfully sent to SHARE Notify</th>
<th>Date</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>4 months ago</td>
</tr>
</tbody>
</table>
Hydra-in-a-Box

http://hydrainabox.projecthydra.org/
<table>
<thead>
<tr>
<th>Repository Product</th>
<th>Aggregator Product</th>
<th>Hosted Service</th>
</tr>
</thead>
<tbody>
<tr>
<td>Create a turnkey Hydra application based on Sufia 7</td>
<td>Development / integration of DPLA metadata aggregation &amp; enrichment tools based on Heiðrún</td>
<td>Cloud-based version of the repository product for use across multiple domains</td>
</tr>
</tbody>
</table>

Connections across DPLA hubs, current Hydra partners, and prospective Hydra adopters
Design Artifacts Published, Survey Data Archived

All available on the project wiki: http://bit.ly/hybox-design-docs
Community Sprints

- Complex objects (Princeton, Oregon State, Chemical Heritage Foundation)
- Mediated deposit (Oregon State, Cincinnati, Indiana, Michigan, Cal State, Penn State, Notre Dame, Cincinnati)
- Admin dashboard (Princeton, Penn State, Cincinnati, Cornell)
- Fedora Import/Export (York, Columbia, Princeton, Virginia, Penn State, DuraSpace)
- FileSets WG (Art Institute of Chicago and others)
For More Information

projecthydra.org

github.com/projecthydra

jwd@iu.edu

karen.estlund@psu.edu