Hydramata

Building a Nimble Solution with Hydra to Transcend the Institutional Repository

Rick Johnson, University of Notre Dame
Robin Ruggaber, University of Virginia
Open Repositories, 2014
Discovering Common Challenges

• Aging, limited or absence of Institutional Repositories (IR)
• “IR” - a limiting descriptor of need
• Limited resources & turnover
• Scholarship now demands a flexible solution
  – Thesis and Dissertations are no longer simple text
  – Digital scholarship blows away traditional limits
Hydra Community Goals

• Identifying successful working methods to support multi-institutional projects

• Solution Bundles – software leveraging the Hydra framework that can be easily configured, tailored and integrated with an institution’s architecture
Alignment Toward a Solution

• Hydra Community
  – Known partners
  – Experienced with common technologies
  – Emerging solution bundles developed by more than 1

• Critical needs by 6 institutions with desire to work together

• Unique opportunity and willingness to solve a problem together with a chance to meet community goals

• Solution Flexibility
  – Modular – architectural components would need to do specific functions for a variety of “intellectual works”
  – Configurable – 6 institutions with varying Enterprise systems
Mixed Architecture
Timeline

- OR 2013:
  - Began discussions with interested partners (Northwestern U., U. of Notre Dame, U. of Virginia, Indiana U.)

- August, 2013:
  - Northwestern and Notre Dame partnered together and hired DCE to make 3 contributing institutions

- November, 2013:
  - Alpha release of Curate gem and CurateND at Notre Dame

- December, 2013:
  - Added U. of Virginia, Indiana U., and U. of Cincinnati

- Jan-Apr, 2014:
  - Group of 5 partners Extended Curate

- May, 2014:
  - Notre Dame deployed Beta Update of CurateND
March, 2014
Houston We Have a Problem

➢ Hard to understand
  ○ Unclear data flow
  ○ Hard to assign story points
➢ Hard to maintain
  ○ Engine design hard to test
  ○ Slow tests
➢ Hard to configure
  ○ Configurable only at install time
  ○ Monolithic; lacking modularity
➢ Ingest unreliable and opaque
  ○ Fedora/Solr sync problems
  ○ Orphaned Fedora objects
➢ Not set up for advanced work
  ○ Role-based Access Control
  ○ Batch ingest
  ○ Multi-stage workflow

Decreasing velocity
Increasing errors
Diminishing ROI
Preservation failure
How We Have Worked

- Lead Project Director
  - Goal & Priority
- Scrum Master
  - Operations
- Lead Product Owner
  - Stories
  - Product Owners
  - Subject Matter Experts
- UX Lead
- QA Lead
- Technical Lead
  - Implementation
  - Developers
Refreshed Process

• Focus on smaller teams with fewer decision makers
• Split into parallel smaller teams in the future
• Tighter involvement of Development and QA and Sprint Planning and Story Writing
  – 3 Amigos/Story Triage process
  – Tasking happens before sprint planning

<table>
<thead>
<tr>
<th>Queue</th>
<th>TCC</th>
<th>DEV</th>
<th>QA</th>
<th>Done</th>
</tr>
</thead>
<tbody>
<tr>
<td>Story H</td>
<td>Story D</td>
<td>Story C BLOCKED - Story G</td>
<td>Story A</td>
<td>Story B</td>
</tr>
</tbody>
</table>
Our Successes

• Teamwork - new people working alongside seasoned Hydranauts and RAILS engineers benefited both
• Work surfaced challenges with the core Hydra components that can now be explored further with the community for resolution
• Increased awareness for project complexity and how that impacts how we work together and how much planning is warranted
• CurateND
CurateND

1. Article
   MyLibrary: A Digital Library Framework and Toolkit
   
   Abstract:
   This article describes a digital library framework and toolkit called MyLibrary. At its heart, MyLibrary is designed to create relationships between...

2. Article
   Ethical and economic issues surrounding freely available images found on the Web
   
   Abstract:
   By reverse engineering Google queries and by tracing back the referrer values found in Apache log files, the use of images made available from info...

3. Article
   Creating and managing XML with open source software

June 12, 2014
Wide Content Support

What are you uploading?
Before we can begin we need to know a little about what you’re uploading.

Article
Deposit or reference a preprint or published article.
Add New

Dataset
One or more files related to your research.
Add New

Document
Deposit any text-based document (other than an article).
Add New

Image
Any image file: art, photograph, poster, etc.
Add New

Senior Thesis
PDFs and other Documents for your Senior Thesis
Add New

Video
Upload your video here.
Add New

Coming Soon
Creating a Work

Required Information

- Title
  - Chris DeLuca
  - + Add

Creator
  - Chris DeLuca
  - + Add

Description

Content Editors

- Editor Individual
  - At least one person should be able to edit this work
  - Chris DeLuca
  - + Add

Group
  - Add groups to allow others to collaboratively edit this work.
  - + Add

Additional Information

- Contributor
  - + Add

Date created
  - 2014-05-12
  - + Add

Subject
  - + Add
Assign a DOI

Add Your Content

The subject of your dataset can either be a file or a URL.

Upload one or more files
Sample Dataset.csv

OR

External link

Assign a Digital Object Identifier (DOI)

A DOI is a permanent link to your dataset. It's an easy way for other people to cite your work.

Want more information on DOIs? Here's a brief summary and the DOI FAQ.

☐ Yes, I would like to create a DOI for this Dataset.
☐ Yes, I already have one that I want to use: doi:10.5072/Fi
☐ Not now...but maybe later.

Access Rights  Who should be able to view or download this content?
Share with the Right People

Want more information on DOIs? Here’s a brief summary and the DOI FAQ.

- Yes, I would like to create a DOI for this Dataset.
- Yes, I already have one that I want to use: [do:10.5072/Fi]
- Not now but maybe later.

Access Rights Who should be able to view or download this content?

Please note, making something visible to the world (i.e. marking this as Open Access) may be viewed as publishing which could impact your ability to:

- Patent your work
- Publish your work in a journal

Check out SHERPA/RoMEO for more information about publisher copyright policies.

- Open Access Visible to the world.
- Open Access with Embargo Treated as Private until 2014-05-15 then it is Open Access.
- University of Notre Dame Visible to all University of Notre Dame users.
- Private Only visible to you and your delegates.
User Profiles

Linda Newman

Selected Works

Cincinnati Subway
Learning what a Profile Section does

- Andy’s Test Collection 1
  Lorem ipsum dolor sit amet, consectetur adipiscing elit. Aenean a erat id tellus imperdiet malesuada sit amet vel sem. Ut consectetur, sapien in tempor suscipit, tortor orci fringilla eros, vitae lacreset nisl nibh ac nunc. Aenean a ultrices odio.

- Street Improvement Photographs -- Box 32, Folder 35 (Calvin Cliff) -- negative, 1930

Email: newmanld@ucmail.uc.edu

Update Linda Newman’s Information
Updated Product Vision

• Flexibility
• Modularity
• Extensibility
• Wide Format Support
• Support IR use cases
• Swappable Viewers, Workflows, Other Plugins
Hydramata Deep-Dive Teaser

A set of separate hydra-based automata, “hydramata”, that can be combined to form a single Hydra Head. Each hydramataton can be mixed and matched with other hydramata to provide wide format support tailored to fit your institutional or digital repository needs.
Tighter OAIS and Hydra Alignment
Contact Info

• Rick Johnson, rick.johnson@nd.edu
• Robin Ruggaber, rsl6m@eservices.virginia.edu
• Hydramata Project wiki:
  – https://wiki.duraspace.org/display/hydra/Hydramata+Project
• Curate code base:
  – https://github.com/projecthydra-labs/curate
• CurateND: http://curate.nd.edu
• Hydra Community: http://projecthydra.org/