

# 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

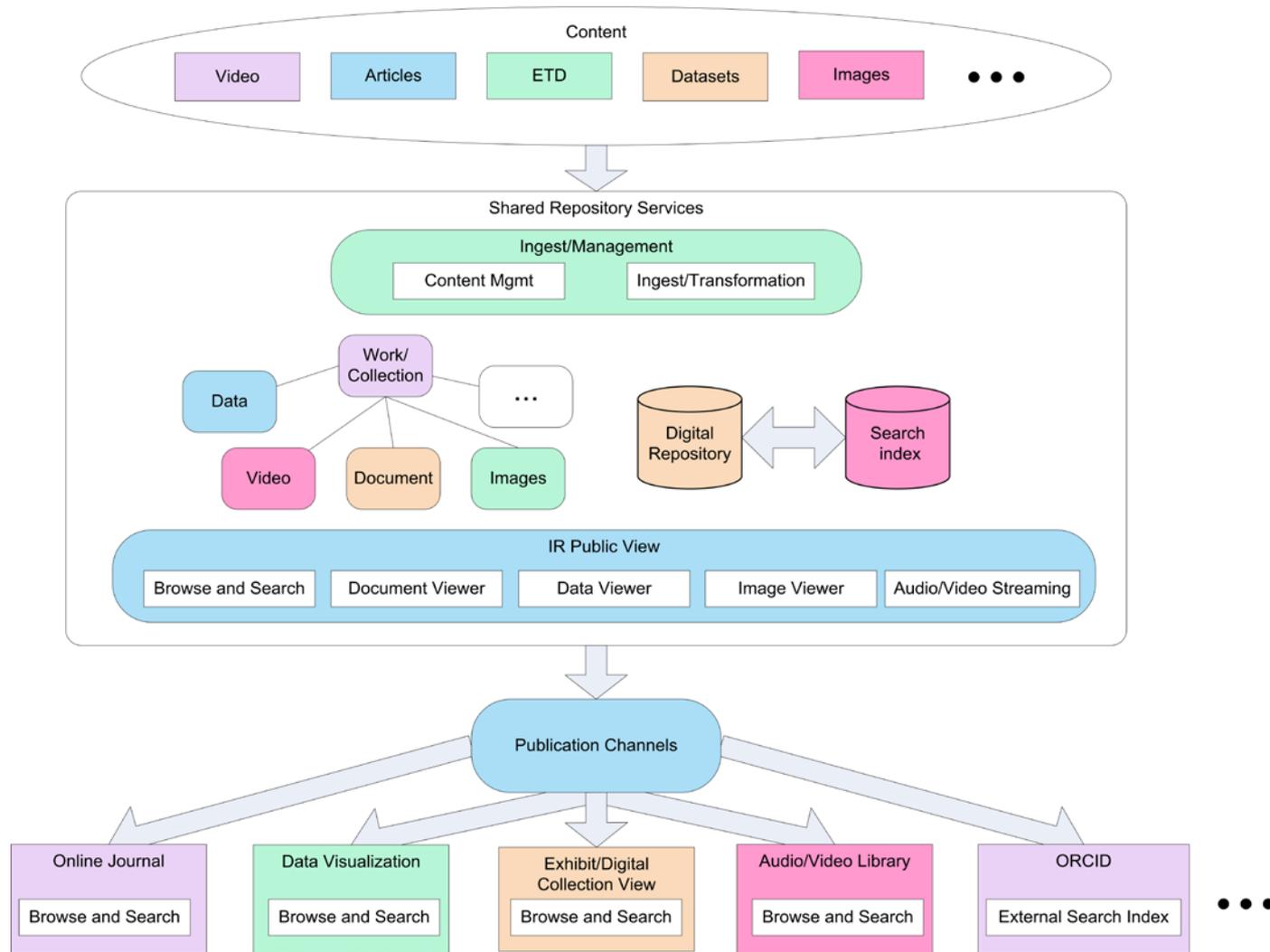
# Hydra Community Goals

- 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



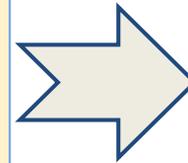
# Timeline

- 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



# 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

Queue	TCC	DEV	QA	Done
<i>Story H</i>	<i>Story D</i>	<i>Story C</i>	Story A	Story B
<i>Story I</i>	<i>Story E</i>	<i>BLOCKED - Story G</i>		
<i>Story J</i>	<i>Story F</i>			

# 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

CurateND Search Results

UNIVERSITY of NOTRE DAME HESBURGH LIBRARIES

Search CurateND CurateND Rick Johnson Help!

Works Collections People All

View:

- All Content
- My Content

Type of Work >

Creator >

Language >

Publisher >

« Previous Next » 1 2 3 4 5 ... 18 19

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](#)   

# Wide Content Support

The screenshot shows the CurateND website interface. At the top, there is a navigation bar with a search box labeled "Search CurateND", the "CurateND" logo, and user information for "Rick Johnson" with a dropdown menu and a "Help!" button. The dropdown menu is open, showing options: "My Works", "My Collections", "My Profile", "My Proxies", and "Log Out".

The main content area is titled "What are you uploading?" and includes the text: "Before we can begin we need to know a little about what you're uploading." Below this, there are six content type categories, each with a description and an "Add New" button:

- Article**: Deposit or reference a preprint or published article.
- Dataset**: One or more files related to your research.
- Document**: Deposit any text-based document (other than an article).
- Image**: Any image file: art, photograph, poster, etc.
- Senior Thesis**: PDFs and other Documents for your Senior Thesis.
- Video**: Upload your video files. A red "Coming Soon" stamp is overlaid on this category.

# Creating a Work

## Required Information

★ Title

Creator  
 – Remove  
 + Add

Description

## Content Editors

Who should be able to make changes to this work?

★ Editor Individual At least one person should be able to edit this work  
 – Remove  
 + Add

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

## Additional Information

★ Contributor  
 + Add

Date created  
 – Remove  
 + 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

Choose File

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:
- 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:
- 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  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

The screenshot shows a web browser window displaying a user profile for Linda Newman. The browser's address bar shows the URL: <https://curatevanilla.library.nd.edu/people/x059c7885>. The page header includes "YOUR INSTITUTION" and "YOUR DIVISION AT INSTITUTION". A search bar with the text "Type keywords in here" and a "Curate" logo are visible. The profile name "Linda Newman" is prominently displayed. Under the heading "Selected Works", there are three entries, each with an "Edit" button and a "Remove" button:

- 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 laoreet nisl nibh ac nunc. Aenean a ultrices odio.
- Street Improvement Photographs -- Box 32, Folder 35 (Calvin Cliff) -- negative, 1930**

To the right of the works is a profile picture of Linda Newman, a woman with red hair and glasses, wearing a teal shirt. Below the photo is her email address: [newmanld@ucmail.uc.edu](mailto:newmanld@ucmail.uc.edu). At the bottom of the profile section is a button labeled "Update Linda Newman's Information".

At the bottom of the page, it states: "Powered by [Curate](#) and [Sufia](#). Made possible by the [Hydra](#) project. Build: 2014-03-26 14:07:47"



# Updated Product Vision

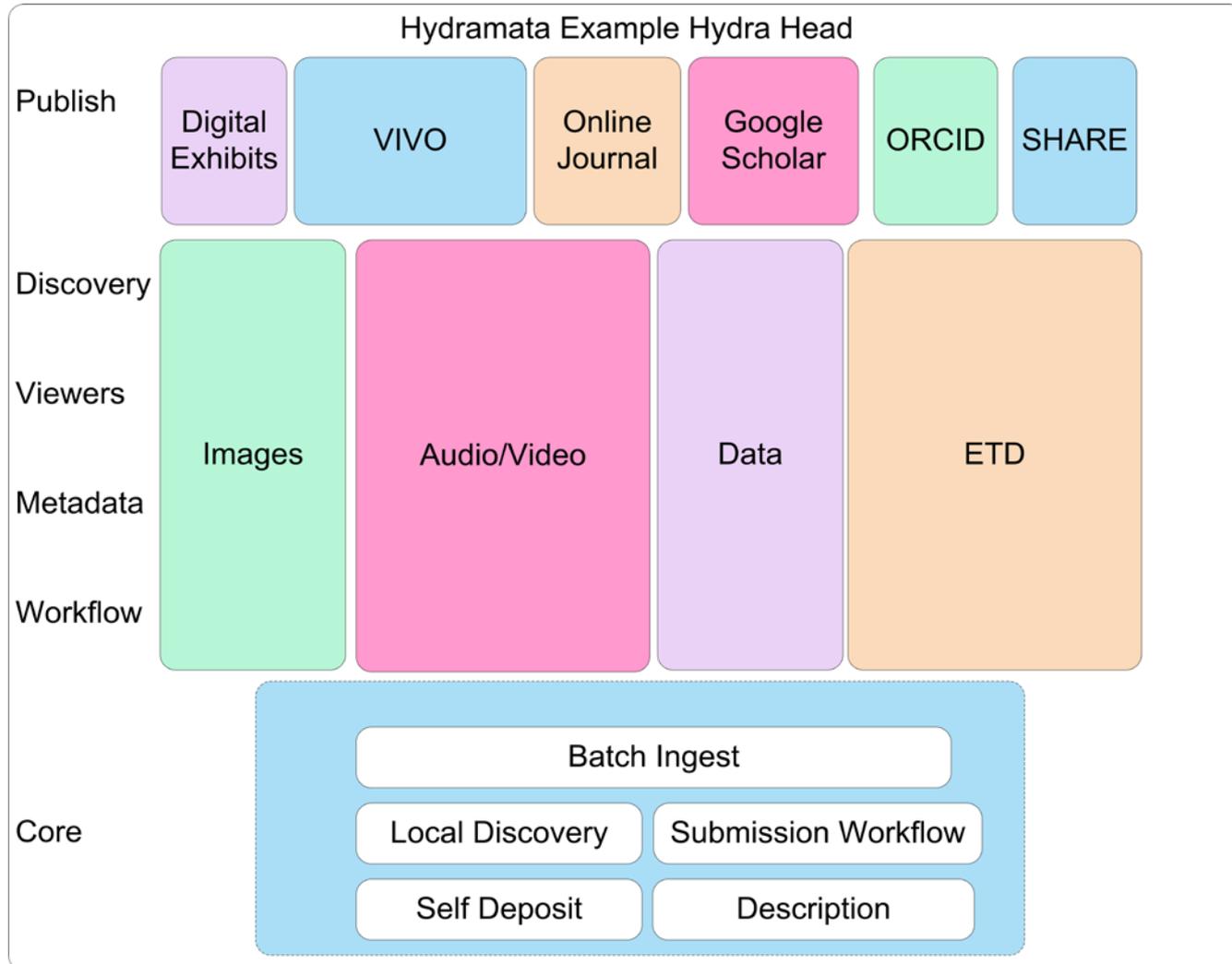
- Flexibility
- Modularity
- Extensibility
- Wide Format Support
- Support IR use cases
- Swappable Viewers, Workflows, Other Plugins



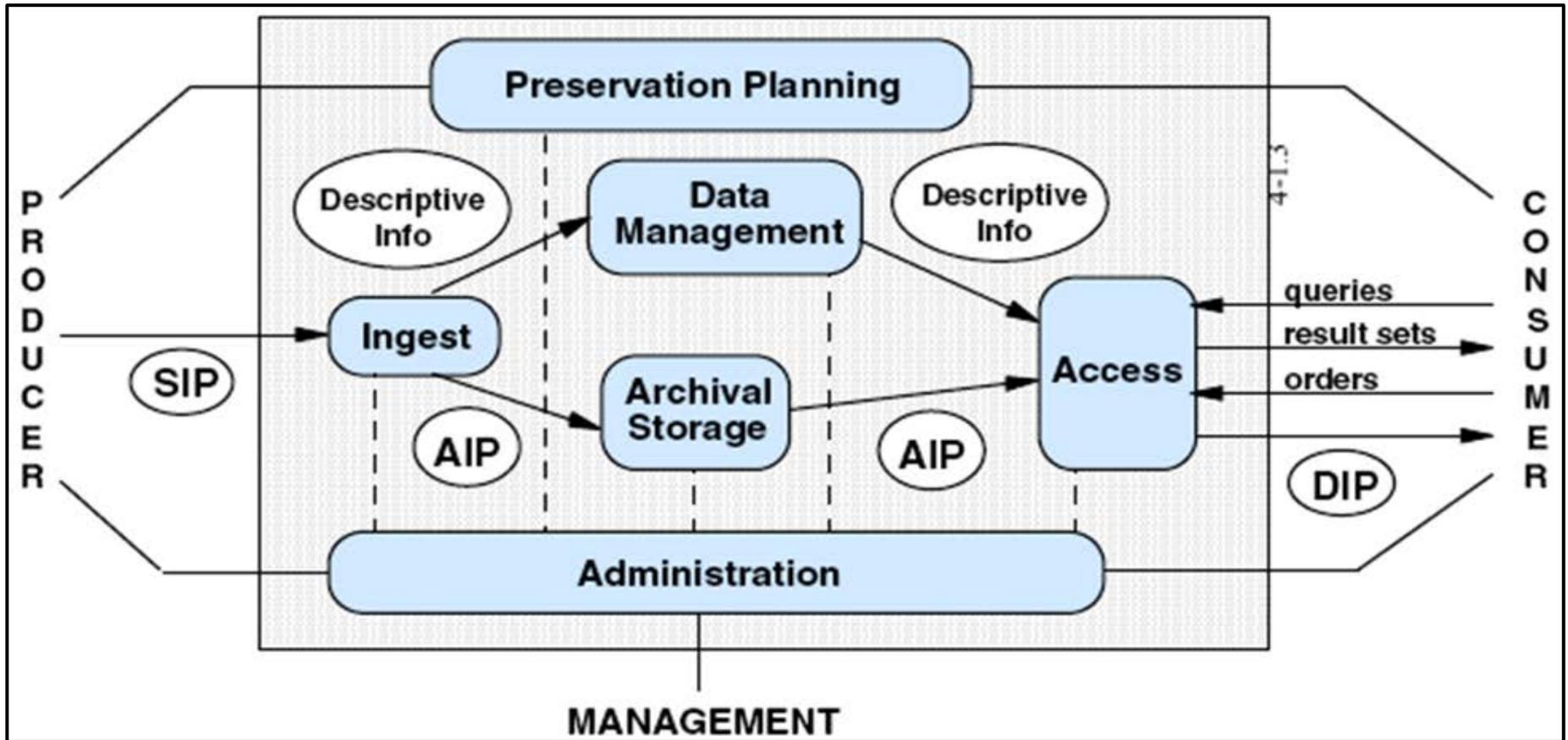
# Hydramata Deep-Dive Teaser

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

# Transformation



# Tighter OAIS and Hydra Alignment



# Contact Info

- Rick Johnson, [rick.johnson@nd.edu](mailto:rick.johnson@nd.edu)
- Robin Ruggaber, [rsl6m@eservices.virginia.edu](mailto: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/>

# Q & A

