



# **A Technical and Community Framework For Customized, Shared Repository Applications**

Tom Cramer  
Lynn McRae  
Willy Mene  
Bess Sadler  
*Stanford University*

Chris Awre  
Richard Green  
*University of Hull*

Tim Sigmon  
*University of  
Virginia*

Thornton Staples  
*DuraSpace*

# **We have a ~~problem~~ ~~need~~ opportunity**

**At Open Repositories 2009, the notion of “durable objects, ephemeral applications” was a major theme.**

**Repositories are well established to provide valuable infrastructure to manage and preserve digital objects.**

**But we need a framework to quickly produce rich, user-facing applications that interface between repositories and users.**

# Fundamental Assumption #1

**No single system can provide the full range of repository-based solutions for a given institution's needs,**

**...yet sustainable solutions require a common repository infrastructure.**

## For instance...

### An ETD solution...

- **Single PDF**
- **With auxiliary data files**
- **Simple, prescribed workflow**
- **Integrated with student administration system**
- **Streamlined UI for depositors, reviewers & readers**

### A digitization workflow system...

- **Potentially hundreds of files type per object**
- **Complex, branching workflow**
- **Sophisticated operator (back office) interfaces**

### A general purpose institutional repository

- **Heterogeneous file types**
- **Simple to complex objects**
- **General purpose user interfaces**

# Distinct Application Needs

More than one dozen distinct repository application needs across three institutions.

- **Electronic theses & dissertations**
- **Open access articles**
- **Data curation application(s)**
- **General purpose institutional repository**
- **Manuscript & archival collection delivery**
- **Library materials accessioning tools**
- **Digitization workflow system**
- **And more...**

# Shared, Primitive Functions

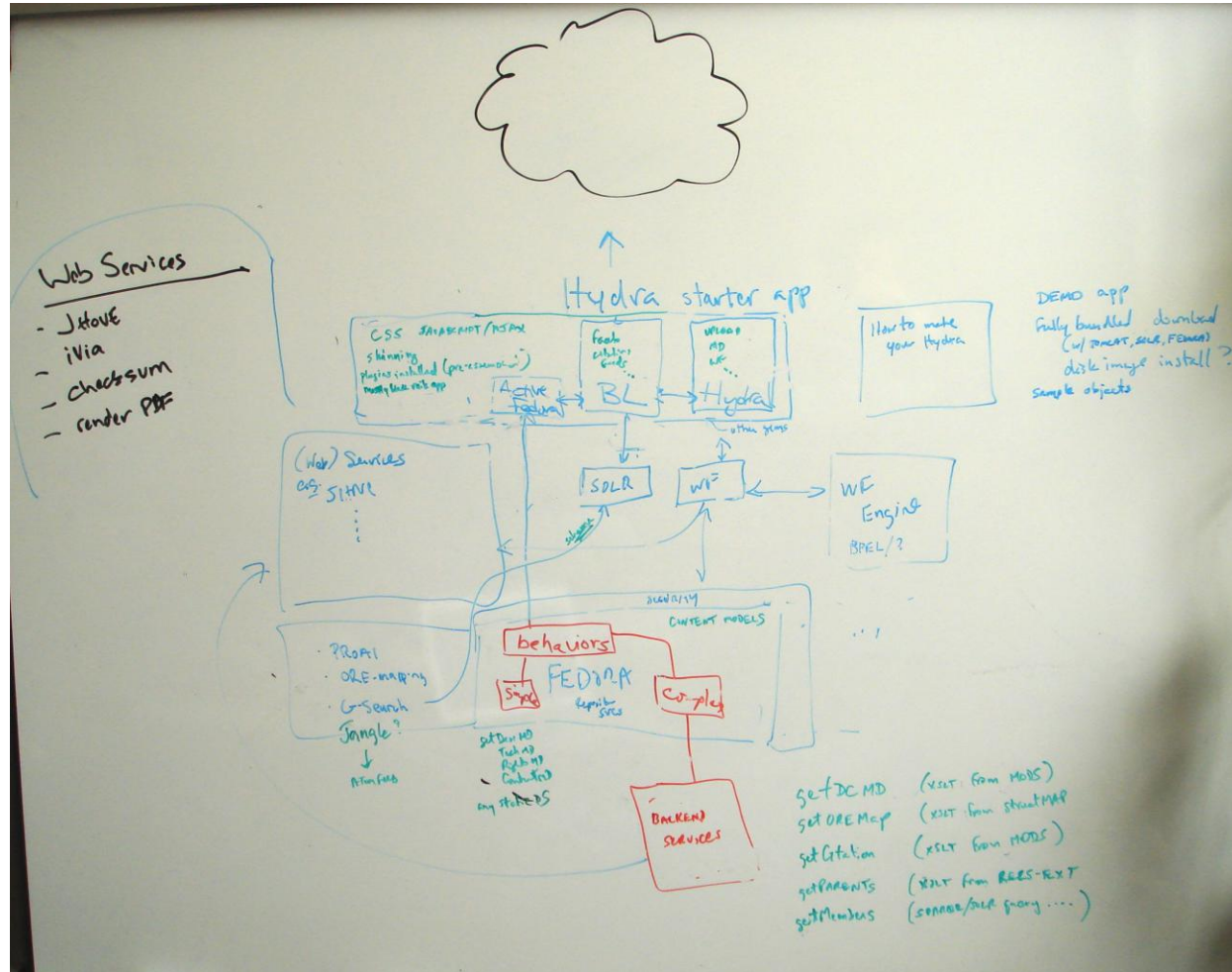
- **Deposit** – uploading simple or multipart objects, singly or in bulk
- **Manage** – editing an object's content, metadata and permissions
- **Search** – full text and fielded search supporting both user discovery and administration
- **Browse** – sequential viewing of objects by collection, attribute or ad hoc filtering
- **Deliver** – viewing, downloading & disseminating objects through user and machine interfaces

# Hydra Philosophy – Technical

- Tailored applications and workflows for different content types, contexts and user interactions
- A common repository infrastructure
- Flexible, atomistic data models
- Modular, “Lego brick” services
- Library of user interaction widgets
- Easily skinned UI

*One body, many heads*

# Technical Framework





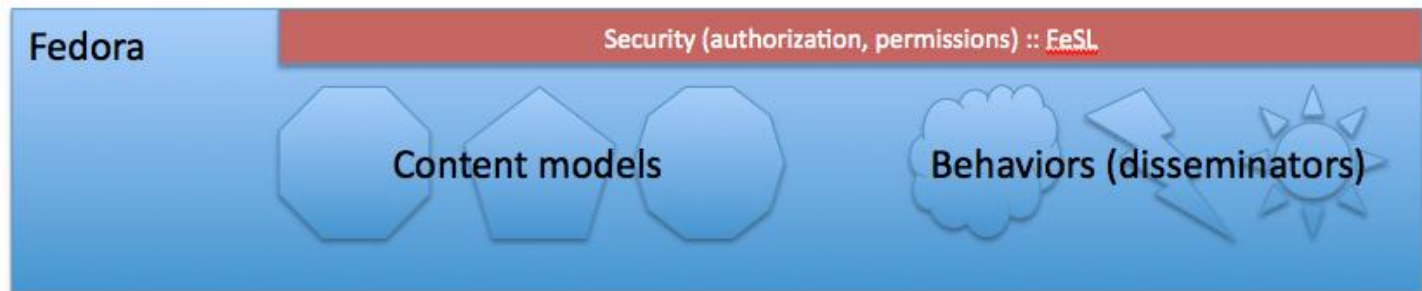
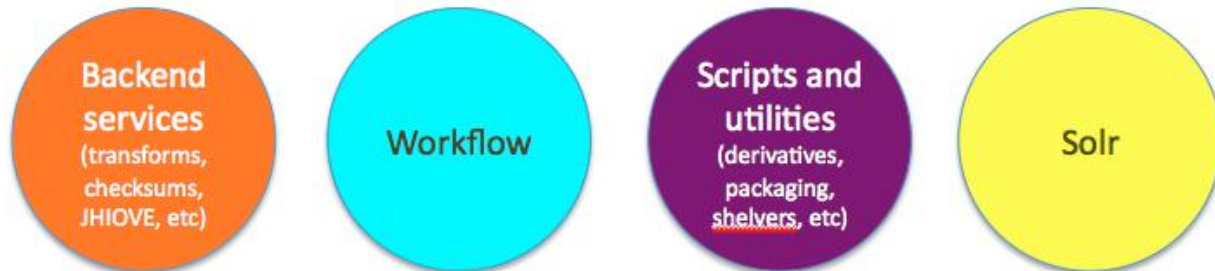
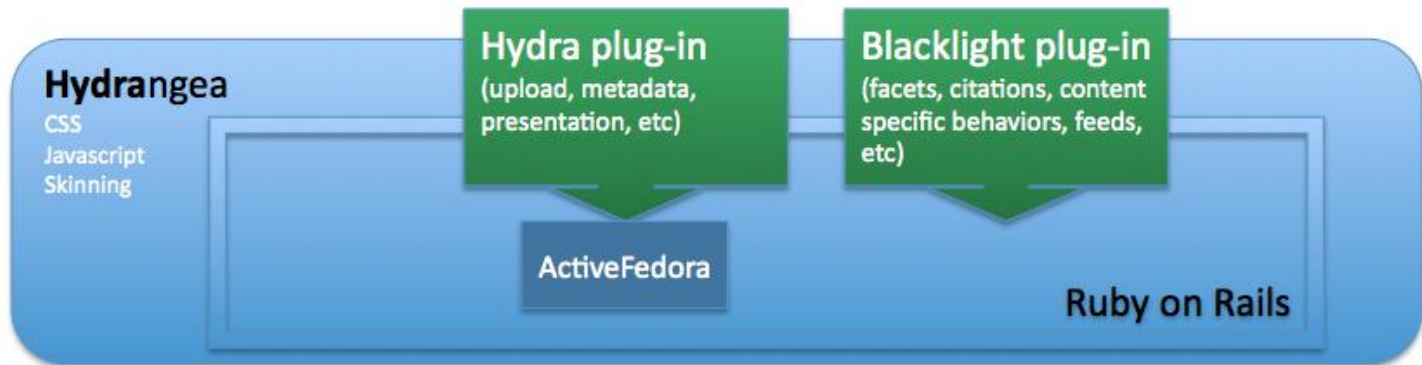
# Technical Framework - Components

- **Fedora** provides a durable repository layer to support object management and persistence
- **ActiveFedora**, a Ruby gem for creating and managing Fedora objects
- **Solr**, provides fast access to indexed information
- **Blacklight**, a Ruby on Rails plugin that sits atop solr and provides faceted search & tailored views on objects

# Technical Framework - Components

- **Hydra Plugin**, a Ruby on Rails library that provides create, update and delete actions against Fedora objects
- **Services**, providing discrete, reusable Lego bricks of functions (e.g., indexing, checksumming, MD transforms, etc.)
- **Hydrangea**, a web application that bundles all the RoR components, user interaction and hooks to services into a single, adaptable package

# Technical Framework



# Fundamental Assumption #2

**No single institution can resource the development of a full range of solutions on its own,**

**...yet each needs the flexibility to tailor solutions to local demands and workflows.**

# Hydra Philosophy – Community

- An open architecture, with many contributors to a common core
- Collaboratively built “solution bundles” that can be adapted and modified to suit local needs
- A community of developers and adopters extending and enhancing the core
- “If you want to go fast, go alone. If you want to go far, go together.”

*One body, many heads*

# Community

- **Conceived & executed as a collaborative, open source effort from the start**
- **Joint development project among Stanford, University of Virginia, University of Hull, DuraSpace, and MediaShelf**
- **20+ active contributors across 5 organizations**
- **Half dozen (plus?) potential adopters (& contributors!) waiting in the wings**

# Ultimate Objective

**Hydra's ultimate objective is to effectively intertwine its technical and community threads of development, producing a community-sourced, sustainable application framework.**

**This framework will ideally provide rich and robust repository-powered solutions as an integrated part of an overall digital content management architecture.**

**Such solutions can meet the distinct needs of digital library, institutional repository, discipline repository, research, preservation and publishing workflows.**

# Hydra Heads

- **Electronic Theses & Dissertations**
- **SALT (digital archival materials)**
- **EEMs (accessioning born digital library materials)**
- **IR @ Hull**
- **Images & Manuscripts at UvA**
- **Hydrangea**
  - **For articles**
  - **For data**

## On Deck:

- **Preservation Repository – administrator’s interface**
- **Digitization Workflow & Management Tool**



# ETDs

Logout

Search for student, advisor, title, department...

[Start over](#) | [Back to search results](#)

### Actions

**Submitted** 2/22/2009 3:20pm

**Rejected** 2/24/2009 12:39pm  
Margin spacing is incorrect. Please reformat so margins are 1" on all sides, then resubmit.  
- John R. Doe

**Resubmitted** 2/26/2009 2:55pm

### Next step

- [Reject](#) (return to student to resubmit)
- [Comment](#) (no change in status)
- [Verify](#) (sign off and send to library for final processing)

Your comment here...

#3 of 10 results for Degree: PhD, Status: Submitted

[« Prev](#) [Next »](#)

Name: [Shester, Geoffrey G.](#)  
Degree: PhD  
Major: School of Earth Sciences: Interdisciplinary Program in Environment and Resources (IPER)  
Advisor: [Loague, Keith](#)

Dissertation Title: Sustainability in small-scale fisheries: An analysis of ecosystem impacts, fishing behavior, and spatial management using participatory research methods

### Thesis or Dissertation (PDF)

<a href="#">Shester-Sustainability-in-small-scale-fisheries.pdf</a>	1,060K	2/26/2009 2:51 pm
---	--------	-------------------

### Supplemental files

<a href="#">cover.jpg</a>	High-res version of cover image	125K	2/26/2009 2:52 pm
<a href="#">video.avi</a>	Video of interviews	2,325K	2/26/2009 2:52 pm
<a href="#">data.xls</a>	Data for survey appendix B	325K	2/26/2009 2:52 pm
<a href="#">illustration.jpg</a>	Full-size illustration from page 75	600K	2/26/2009 2:52 pm

### Abstract

Anthropogenic impacts to ocean ecosystems have been well documented throughout the globe, highlighting commercial fishing as a primary threat to marine life populations and habitats. Small-scale commercial fishing operations provide more than half of the world's seafood, employ over 90% of all fishermen, and take place along a vast proportion of the world's coastline. Their local, decentralized nature poses unique challenges for research, data collection, and management, particularly in developing nations. As a result, a focus on large, industrial fisheries has discounted the potential impacts of small-scale fisheries, leading to an assumption that they are generally more sustainable. My research on two remote fishing cooperatives in Baja California Sur, Mexico asks to what extent these small-scale fisheries are sustainable, and in cases where significant ecological impacts do occur, I offer and compare several distinct conservation approaches. While engaging fishermen to participate directly in the data collection, I begin by conducting a series of studies that examine how small-scale fisheries interact with marine ecosystems. I quantitatively assess potential ecological effects including bycatch of non-target species and damage to seafloor habitats resulting from several common fishing practices and find that the type and magnitude of

## Components

- File upload (PDF & aux. files)
- Metadata entry
- Access & license settings
- Workflow & in page status panel
- Dissemination (to ILS, to preservation core, to Google)

# SALT: Self Archiving Legacy Toolkit

The screenshot shows the SALTworks web interface. At the top, there is a search bar with the text 'artificial intelligence' and a 'SEARCH' button. Below the search bar, it says 'Start Over Your Search: artificial x + intelligence x Limited to: (no limits selected)'. There is also a 'Logout [ttramer]' link. The main content area displays the document 'Artificial Intelligence 1973' with a thumbnail image of a handwritten document. The document text is partially visible, starting with 'THE FOLLOWING ARE SOME THOUGHTS ON MAN'S NEW IMPACT IN APPLICATION-ORIENTED RESEARCH...'. To the right of the document, there is a metadata section with the following information: Title: Artificial Intelligence 1973, Type: Paper Document, Access: private, Series: Accession 1986-052, Subseries: Correspondence, Box: 3, Folder: 33, ID: druid:yz604wq6818. On the left side, there is a sidebar with 'Edward A. Feigenbaum Papers' and a list of 'Artificial Intelligence 1973' documents. Below the sidebar, there is a section 'This document refers to:' with a list of series, boxes, and folders.

## Components

- Gated discovery (authz for search)
- Direct edits to repository objects
- Crowd-sourced description & annotation
- Embedded page viewer

# EEMs: Accessioning Born Digital Materials

SUOLAIR

Tom Cramer | Logout

Search

Filter by  
Selector  
Glen Worthey (19) [remove]

Status  
Purchase  
Paid (10)  
Free (9)

Fund  
LATAM (6)  
DATAFILEZ (2)  
MIDEAST (1)  
RUEEUR (1)

Copyright  
Requires request (11)  
Public access OK (7)  
Stanford access OK (1)

Language  
Creator (organization)  
Creator (person)

« Previous Next »  
Sort by relevance  
Show 10 per page

Title	Created	Submitted	Selector	Copyright	Status
История советских органов государственной безопасности	24-Jun-2010	24-Jun-2010	Glen Worthey	Public access OK	Submitted
FIRST CENSUS RESULTS BY MUNICIPALITIES AND SETTLEMENTS OF THE REPUBLIC OF SERBIA	24-Jun-2010	24-Jun-2010	Glen Worthey	Public access OK	Submitted
ITHAKA'S 2006 STUDIES OF KEY STAKEHOLDERS IN THE DIGITAL TRANSFORMATION IN HIGHER EDUCATION	24-Jun-2010	24-Jun-2010	Glen Worthey	Requires request	Submitted
FACULTY SURVEY 2009	24-Jun-2010	24-Jun-2010	Glen Worthey	Public access OK	Submitted
أدب ونقد، عدد إبريل	29-Jun-2010	29-Jun-2010	Glen Worthey	Requires request	Submitted
Први резултати пописа по општинама и насељима	29-Jun-2010		Glen Worthey	Requires request	Created
Convención de las Naciones Unidas contra el Crimen Organizado Transnacional y sus Protocolos	29-Jun-2010	29-Jun-2010	Glen Worthey	Public access OK	Submitted
Convención de las Naciones Unidas contra la Delincuencia Organizada Transnacional	29-Jun-2010	29-Jun-2010	Glen Worthey	Public access OK	Submitted
El Aleph	29-Jun-2010	29-Jun-2010	Glen Worthey	Requires request	Submitted

## Components

- File capture
- Metadata entry
- Branching workflow
- Admin “dashboard”
- Dissemination (to ILS, to preservation core, to Google)

# Hydra @ Hull

The screenshot shows the Hydra search interface. At the top, there is a navigation bar with links for 'UoH Home', 'Library', 'eDocs', 'About', and 'Contact'. The University of Hull logo is on the right. Below the navigation bar, there is a search bar with the text 'Green' and a dropdown menu set to 'All Fields'. A search button is next to it. On the left side, there is a 'Limit your search' section with links for 'Subject', 'Author/Creator', and 'Genre'. The main search results area shows a message 'Did you mean to type screen?' and 'You searched for: Keyword Green'. Below this, it says 'Displaying all 5 items' and 'Start again'. The results are sorted by 'relevance' and show 10 items per page. The first result is '1. Content models in the Hydra Project' by Green, Richard A. The second result is '2. An investigation of the factors which influence the degree of patient involvement in the physiotherapeutic consultation' by Green, Angela Jane.

## Components

- Blacklight on top of Fedora
- Rich, robust user search & view interface
- Configured facets to support browse
- Specialized object views
- Asset delivery from Fedora

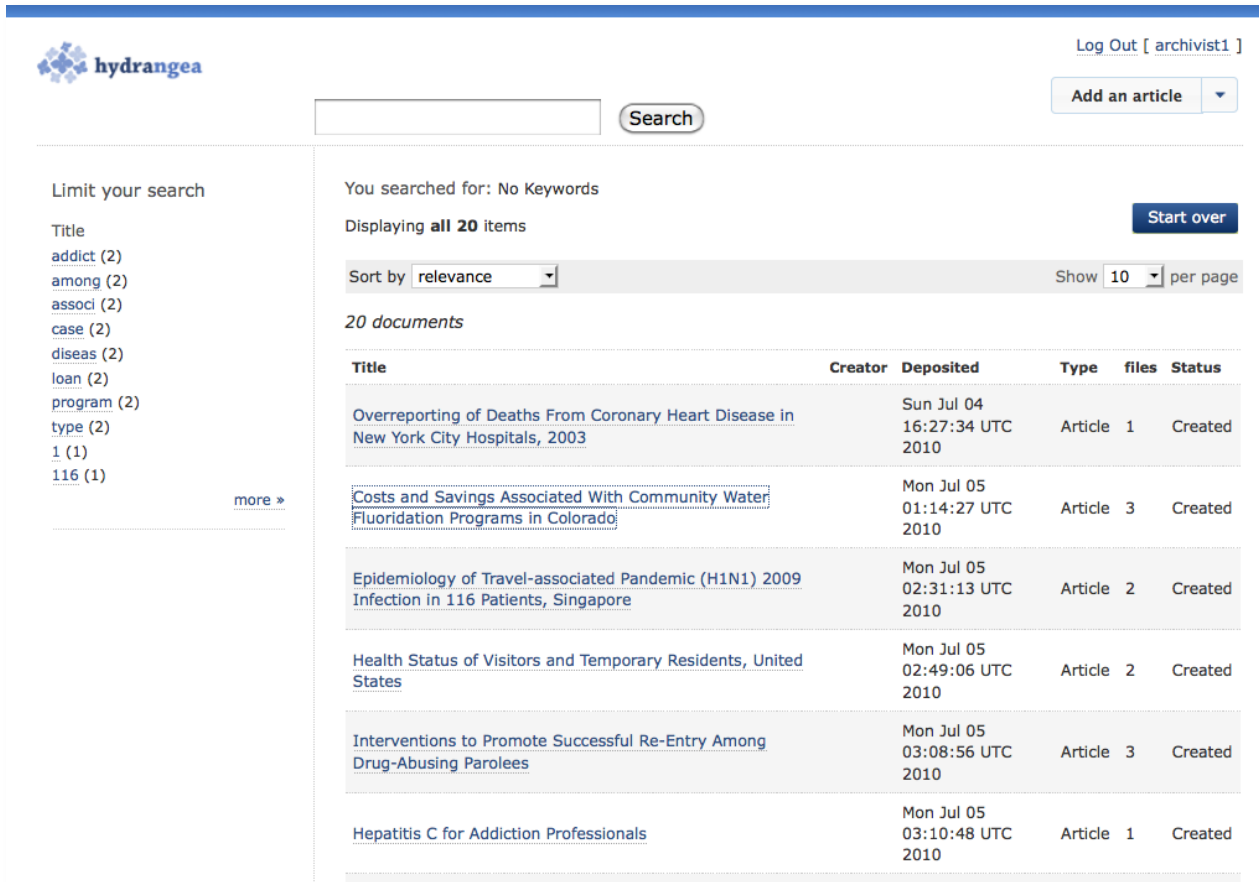
# Manuscripts & Images at UVa

The screenshot displays the VIRGO digital repository interface. At the top, the VIRGO logo is on the left, and 'About the new VIRGO' and 'Login' are on the right. Below the logo, there are tabs for 'General View' and 'Music View', and a link to 'Ask A Librarian'. A search bar with a 'Keywords' dropdown and a 'Search' button is present. A 'Print View' icon is also visible. The main content area shows search results for 'Photography'. It includes a 'You searched for:' section with 'No Keywords' and a list of filters: 'Subject > Photography', 'Digital Collection > Holsinger Studio Collection', and 'Source > UVA Library Digital Repository'. Below this, it says 'Displaying items 1 - 20 of 10,681' with a 'Start over' button. A 'Sort by' dropdown is set to 'Date Received', and 'Show 20 per page' is displayed. A pagination bar shows '1' as the current page, with links for 2, 3, 4, 5, 6, 7, 8, 9, and 'Next'. The first two search results are listed: 1. 'Mrs. C.M. White; Holsinger Studio Collection, 1889 - 1939' and 2. 'Dr. J.S. Grasty; Holsinger Studio Collection, 1889 - 1939'. Each result has a 'Preview' link and an 'Add to Marked List' button. On the left side, there is a 'BROWSE BY:' section with categories: Library (Special Collections (10,681)), Location (Charlottesville (Va.) (10,681)), Format (Photographs (10,681)), Author (Holsinger, Rufus W. (9,780), Holsinger, Ralph W. (901)), and Subject (Holsinger Studio (Charlottesville, Va.) (10,681), Photography (10,681), Portraits (5,626)).

## Components

- Blacklight on top of Fedora
- Complex, aggregate objects
- Asset delivery from Fedora
- Integrated ddatoka image streaming

# Hydrangea – Open Access Articles



hydrangea

Log Out [ [archivist1](#) ]

Add an article

Search

Limit your search

You searched for: No Keywords

Displaying all 20 items

Sort by relevance

Show 10 per page

20 documents

Title	Creator	Deposited	Type	files	Status
<a href="#">Overreporting of Deaths From Coronary Heart Disease in New York City Hospitals, 2003</a>		Sun Jul 04 16:27:34 UTC 2010	Article	1	Created
<a href="#">Costs and Savings Associated With Community Water Fluoridation Programs in Colorado</a>		Mon Jul 05 01:14:27 UTC 2010	Article	3	Created
<a href="#">Epidemiology of Travel-associated Pandemic (H1N1) 2009 Infection in 116 Patients, Singapore</a>		Mon Jul 05 02:31:13 UTC 2010	Article	2	Created
<a href="#">Health Status of Visitors and Temporary Residents, United States</a>		Mon Jul 05 02:49:06 UTC 2010	Article	2	Created
<a href="#">Interventions to Promote Successful Re-Entry Among Drug-Abusing Parolees</a>		Mon Jul 05 03:08:56 UTC 2010	Article	3	Created
<a href="#">Hepatitis C for Addiction Professionals</a>		Mon Jul 05 03:10:48 UTC 2010	Article	1	Created

## Components

- File upload
- Metadata entry
- Aggregate objects
- Access & license settings
- Authz w/ search and browse
- Tailored object views
- Dashboard
- Workflow

# Timeline

- **Currently in Y2 of a three year project**
- **Framework coding began in February '09**
- **First production application released in Nov '09**
- **Hydrangea Beta scheduled for public release this summer**
- **Next:**
  - **Distributed installations & development**
  - **From Beta -> 1.0**
  - **More heads**

# An Open Framework

**Wiki:** <http://wiki.duraspace.org/display/hydra>

**List:** [hydra-tech@googlegroups.com](mailto:hydra-tech@googlegroups.com)

**Code:** <http://github.com/projecthydra/hydrangea>

**Meet:** Hydra & Blacklight Camp, Minneapolis, Oct 4-8  
DLF Fall Forum, Palo Alto (California), Nov 1-3

**Join us!**



**Gracias**



***Whereon this **Hydra** son of war is born:  
Whose dangerous eyes may well be charm'd asleep,  
With grant of our most just and right desires;  
And true obedience of this madness cur'd,  
Stoop tamely to the foot of majesty.***

**-Shakespeare  
Henry IV, Part 2**