One Body, Many Heads for Repository-Powered Library Applications

Chris Awre
Head of Information Management
Library and Learning Innovation
University of Hull

Hydra UK event, LSE, 22nd November 2012
Hydra

• A collaborative project between:
  – University of Hull
  – University of Virginia
  – Stanford University
  – Fedora Commons/DuraSpace
  – MediaShelf LLC

• Aim to work towards a reusable framework for multipurpose, multifunction, multi-institutional repository-enabled solutions

• Timeframe - 2008-11 (but now extended indefinitely)
Multipurpose, multi-institutional approach

- A repository should be an enabler, not a constraint
  - Repositories have been put forward as potential solutions for a variety of use cases
  - Hydra recognises that repositories can be used in the management of digital content at different stages in the lifecycle of that content
  - It is therefore useful to consider how to enable multiple interactions with a repository for different purposes
Hydra (repository) take-up and embedding

- Hydra is about developing flexible interfaces over a repository that allow for the management of different types of content in the same repository
  - Hydra supports embedding by allowing a single repository to serve multiple needs
  - Hydra supports take-up through the flexible development of end user and management interfaces
    - Designed for use according to content type
- Hydra provides a framework to support adaptability
  - Hydra ‘heads’
Sustainability

No animals were harmed in the making of this film.

...but we’ll use them where we can!
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.

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.
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 partnerships

• From the beginning key aims have been and are:
  – to enable others to join the partnership as and when they wished
    • Now up to 10 partners, with two others in process
  – to establish a framework for sustaining a Hydra community as much as any technical outputs that emerge
“If you want to go fast, go alone. If you want to go far, go together”

(African proverb)
Actively Participating Institutions

OR09 | OR10 | OR11 | Now
--- | --- | --- | ---
4 | 6 | 8 | 12

OR = Open Repositories Conference
**Community Model**

**Hydra Steering Group**
- small coordinating body
- collaborative roadmapping (tech & community)
- resource coordination
- governance of the "tech core" and Hydra Framework
- community mtce. & growth

**Hydra Partners**
- shape and direct work
- commission "Heads"
- functional requirements & specs
- UI design & spec
- Documentation
- Training
- Data & content models
- "User groups"

**Founders**
- Duraspace
- Hull
- Stanford
- UVa

**Hydra Developers**
- define tech architecture
- code development
- integration & release

- Committers
- Contributors
- Tech. Users

Currently
- DuraSpace
- Hull
- MediaShelf
- Stanford
- Virginia
Managing the community

• Founding partners have an MoU governing how the community is managed
  • Subsequent partners have signed up to this MoU through a partner agreement addendum

• All code contributions are being managed through Code Licensing Agreements
  • Individual – so each developer is clear about what they are contributing
  • Corporate – so each institution is clear about what they are contributing

• All Hydra code is available under Apache Licence, Version 2.0
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...

**ETD Deposit System**
- Generally a single PDF
- Simple, prescribed workflow
- Streamlined UI for depositors, reviewers & readers

**General Purpose Institutional Repository**
- Heterogeneous file types
- Simple to complex objects
- One- or two-step workflow
- General purpose user interfaces

**Digitization Workflow system**
- Potentially hundreds of files type per object
- Complex, branching workflow
- Sophisticated operator (back office) 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
Content Framework

• Key to enabling re-use of Hydra repository solutions is a common baseline to how objects are structured
  • Objects must have associated rights metadata
  • Objects must include a statement of what content models the objects adhere to
• The Hydra community has developed some basic building block content models (the Lego brick approach)
  • Combine and/or extend these to meet your needs
Technical Framework - Components

• **Fedora** provides a durable repository layer to support object management and persistence

• **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

• **Hydra Plugin**, a Ruby on Rails library that provides create, update and delete actions against Fedora objects
Fedora and Hydra

• Fedora can be complex in enabling its flexibility
• How can the richness of the Fedora system be enabled through simpler interfaces and interactions?
  – Hydra has sought to address this, and has done so successfully
  – Not a complete turnkey, out of the box, solution, but a toolkit that enables powerful use of Fedora’s capabilities through lightweight tools
    • Principles can also be applied to other repository environments
CRUD in Repositories

hydra-head
Rails Plugin

ActiveFedora  OM

Fedora

Solrizer

Blacklight
(Read Only)

Solr
Emerging Solution Bundles

Institutional Repositories
University of Hull
University of Virginia
Penn State University

Images
Northwestern University
(Digital Image Library)
Emerging Solution Bundles

Archives & Special Collections
Stanford University
University of Virginia
Rock & Roll Hall of Fame

Media
Indiana University
Northwestern University
Rock & Roll Hall of Fame
Etc.
Emerging Solution Bundles

Workflow Management (Digitization, Preservation)
Stanford University
University of Illinois – Urbana-Champaign
Northwestern University

Exhibits
Notre Dame
Emerging Solution Bundles

ETDs
Stanford University
University of Virginia
Etc.

(Small) Data
everyone...
Hydra is a Repository Solution

Hydra is a repository solution that is being used by institutions on both sides of the North Atlantic to provide access to their digital content. Hydra provides a versatile and feature-rich environment for end-users and repository administrators alike.

Hydra is a Community

Hydra is a large, multi-institutional collaboration. The project gives like-minded institutions a mechanism to combine their individual repository development efforts into a collective solution with breadth and depth that exceeds the capacity of any single institution to create, maintain or enhance on its own. The motto of the project's partners is "if you want to go fast, go alone. If you want to go far, go together."

Hydra is a Technical Framework

Hydra is an ecosystem of components that lets institutions deploy robust and durable digital repositories (the body) supporting multiple "heads": fully-featured digital asset management applications and tailored workflows. Its principle platforms are the Fedora Commons repository software, Soir, Ruby on Rails and Blacklight. See how you can get started.

Hydra is Open Source Software

Hydra software is free and open source, available under an Apache 2 license.
Philosophies

• Building a framework, not an application (variation is part of the plan)
• Opinionated software
• Invest time & resources into collaborative community (face time!)
• Trainings & workshops
• Openness, transparency (code, designs, discussions)
• Commit to contributing back to core
• Design for re-use
http://projecthydra.org

http://wiki.duraspace.org/display/hydra