What Is Hydra?

• A robust repository fronted by feature-rich, tailored applications and workflows (“heads”)
  ➤ *One body, many heads*

• 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.*
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.
A Tale of Two Workites

Stephen Jay Gould

Museum of Comparative Zoology
Harvard University
Cambridge, MA 02138

Christopher Wren, the leading architect of London’s reconstruction after the great fire of 1666, lies buried beneath the floor of his most famous building, St. Paul’s cathedral. No elaborate sarcophagus adorns the site. Instead, we find only the famous epitaph written by his son and now inscribed into the floor: a monument requites, eumvaxis — if you are searching for his monument, look around. A tall granite perhaps, but I have never read a finer testimony to the central importance — one might even say sanctity — of actual places, rather than replicas, symbols, or other forms of vicarious resemblance.

An odd coincidence of professional life recently turned my thoughts to this most celebrated epitaph when, for the second time, I received an office in a spot laden with history, a place still redolent with ghosts of past events both central to our common culture and especially meaningful for my own life and choices.

In 1973, I spent an academic term as a visiting researcher at Oxford University. I received a cranny of office space on the upper floor of the University Museum. As I set up my books, fossil snails and microscope, I noticed a metal plaque affixed to the wall, informing me that this reconfigured space of shelves and cubicles had been, originally, the site of the most famous public confrontation in the early history of Darwinism. On this very spot, in 1858, just a few months after Darwin published the Origin of Species, T.H. Huxley had drawn his rhetorical sword, and soundly skewered the slick but superficial champions of creationism, Bishop "Nobby Sam" Wilberforce.

As with most legends, the official version ranks as more cardboard before a much more complicated and multifaceted truth. Wilberforce and Huxley did put on a splendid, and largely pompous, show — but no clear victor emerged from the scuffle and Joseph Hooker. Darwin’s other champions, made a much more effective reply to the Bishop, however forgotten by history. See my essay on this debate, entitled "Knight takes Bishop!" and published in an earlier volume of this series: Risky for Brontosaurus.

I can’t claim that the lingering presence of these Victorian giants increased my resolve or improved my work, but I loved the sense of continuity vestigial to me by this happy circumstance. I even traced the etymology — for "circumstance" means "standing around" (...
Point Solution Approach ...Welcome to Siloville

ETDs (Theses) | Books, Articles | Images | Audio-Visual | Research Data | Maps & GIS | Documents
---|---|---|---|---|---|---
ETD | IR | Image DB | DAM | | Geospatial Inf. | Records Mgmt.

Management | Access | Preservation(?)
Effective? | Sustainable?
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...

ETDs (Theses)  Books, Articles  Images  Audio-Visual  Research  Maps & GIS  Documents

Scalable, Robust, Shared Management and Preservation Services
Hydra Technical Framework
CRUD in Repositories

Create/Submit/Edit (CUD)

Search/View (R)

Repository/Persistent Storage
CRUD in Repositories

Create/Submit/Edit (CUD)

Search/View (R)

Repository/ Persistent Storage
Major Hydra Components

- hydra-head
  Rails Plugin (CUD)
- Fedora
- Solrizer
- Blacklight (Read Only)
- Solr
A Note on Fedora

• Flexible, Extensible, *Durable* Object Repository Architecture
  – Flexible: model and store any content types
  – Extensible: easy to augment with apps and services
  – Durable: foundation of preservation repository

• Proven, sustained and successful digital repository
  – 100’s of adopters; 13 years of development, 4 releases
  – Vibrant community & funding under DuraSpace

• Fedora 4.0 released in Dec ‘14; co-evolving with Hydra
A Note on Ruby on Rails

- **Rapid** application development for web applications: “Convention over configuration”
  - 10x productivity
- **Supportable:** MVC (Model-View-Controller) and Rails framework make code well-structured, predictable
- **Testable:** Rspec and Cucumber give powerful, automatable, testing tools
- **Learnable:** Stanford went from 1 to 8 Ruby savvy developers in one year (no new hires)
  - 1 week learning curve to basic proficiency
Hydra’s Core Gems

Rails engine with “everything you need to build a user-facing app”

DSL (domain specific language) for easy mgmt of XML metadata

Applies ActiveModel pattern to working with Fedora objects

Client API for consuming Fedora’s REST API via Ruby
Hydra Heads of Note

Avalon & HydraDAM for Media

UCSD DAMS

BPL Digital Commonwealth

Sufia

Northwestern Digital Image Lib.
Hydra-based Applications at Stanford

ETD’s – Electronic Theses & Dissertations

SALT – Self-Archiving Legacy Toolkit

EEMs – Everyday Electronic Materials

Argo – Repository Reporting and Management

Hypatia – Archives & Special Collections

SDR – Web UI (aka Hydrus)
Portland Common Data Model

pcdm:Collection
  pcdm:hasMember (m:m)
  ore:aggregates (m:m)

pcdm:Object
  pcdm:hasMember (m:m)
  pcdm:hasFile (1:m)
  pcdm:hasRelatedFile (1:m)

pcdm:File

Access
Bitstream
Descriptive
Technical
Book
Why PCDM?

1.) Best practices in data modeling
2.) Code interoperability
3.) Data interoperability
Getting Started

https://wiki.duraspace.org/display/hydra/Developers

On-line Resources and Tutorials

For New and Potential Adopters:

- new adopters and potential adopters may find the pages here useful: http://projecthydra.org/
- See Hydra Heads (Known Implementations) for more information about each of the Hydra Heads.
- A Hydra Head interactive tutorial has been developed - see under the 'For developers' section below.

For Architecture Level:

- See the Duraspace projecthydra wikis for information at the architecture level: http://wiki.duraspace.org/display/hydra/

For Developers:

- Code is hosted on Github: https://github.com/projecthydra
- Github’s wikis for projecthydra git repositories are targeted to developers: https://github.com/projecthydra/heads/wiki
Hydra Community 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 distributed, collaborative, open source effort from the start
- Initially a joint development project between Stanford, Univ of Virginia, and Univ of Hull
- Hydra Partners are the backbone of the project
  - Coalition of the willing
  - No fees or dues
  - Apache-style consensus and governance
- Steering Group provides administration, continuity, and serves as backstop when needed
  - But no central planning, no Project Director, no “one” architect
Hydra Partners

- 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 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 Virginia
- Northwestern University
- Indiana University
- The Royal Library of Denmark
- Boston Public Library
- Virginia Tech
- Cornell University
- Case Western Reserve University
- University of Alberta
Code Licensing

• All Hydra code is available under Apache License, Version 2.0

• All code commitments are being managed through Contributor License Agreements
  • Individual – so each developer is clear about what they are contributing
  • Corporate – so each institution is clear about what it is contributing

• Code contributors maintain ownership of their IP
  • ...and grant a non-exclusive license to the project and its users
Hydra Partners and Known Users

OR = Open Repositories Conference
A Growing Worldwide Presence
Trending

Commits per Month

Contributors per Month
Hydra Heads: Institutional Repositories

University of Hull
Hydra Heads: Institutional Repositories

ScholarSphere / Sufia @ Penn State
Partners & Implementations

https://wiki.duraspace.org/display/hydra/Partners+and+Implementations
Sustainability

No animals were harmed in the making of this film.
If You Want To Go Fast  ...go alone.
  ...use Hydra

1) Notre Dame deployed a video cataloging head in 6 weeks, from scratch

2) Hypatia development at Stanford
   - First 20% of project took 80% of time (modelling, data acquisition and preparation)
   - 80% of deliverables (coding) happened in 8 week spurt, by leveraging Hydra
Hydra Heads: Institutional Repositories

Hydra-in-a-Box

- $2M grant from IMLS, Institute for Museum & Library Services
- Goals:
  - Turnkey solution for institutions large and small
  - Metadata enrichment
  - Hosted service
Sustainability...

- For institutions
  *Partners help shoulder the load*

- For other projects
  *Post-project maintenance & governance*

- For leveraging funders’ investments
  *Adoption by growing user community*