



Developing the Hydra - Blacklight Way

CNI Membership Meeting St. Louis * April 1, 2014

Tom Cramer
Chief Technology Strategist
Stanford University Libraries
@tcramer

Agenda

- What is Hydra?
- What is Blacklight?
- Where are they going?
- What does it all mean?



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.

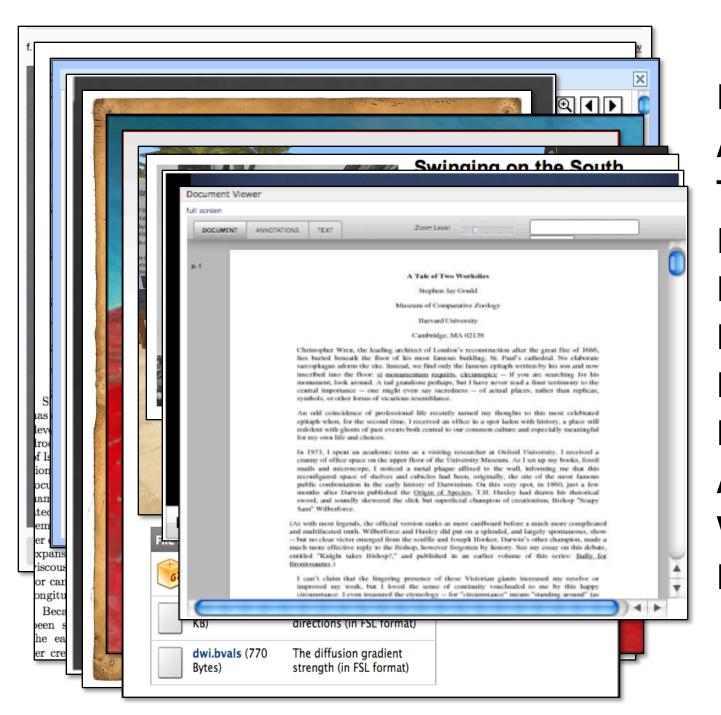


Fundamental Assumption #1

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

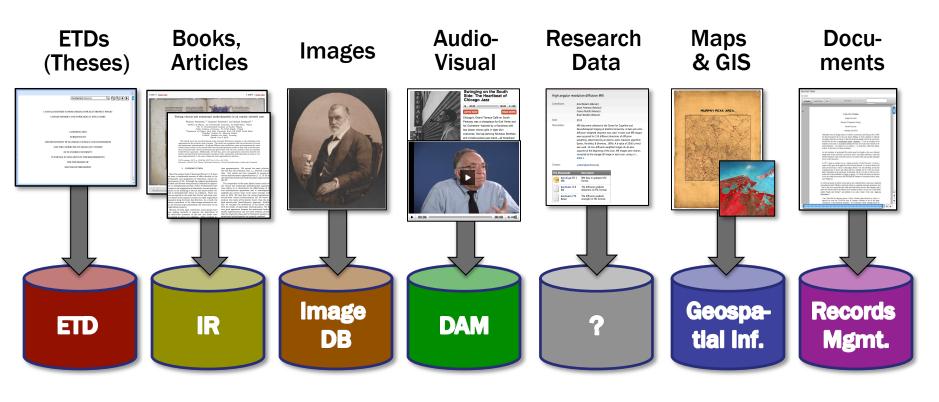
...yet sustainable solutions require a common repository infrastructure.





Books Articles Theses Images Maps Data (Raster) Data (Comp.) Data (Observ.) **Audio** Video **Documents**

Point Solution Approach ... Welcome to Siloville



Management

Access

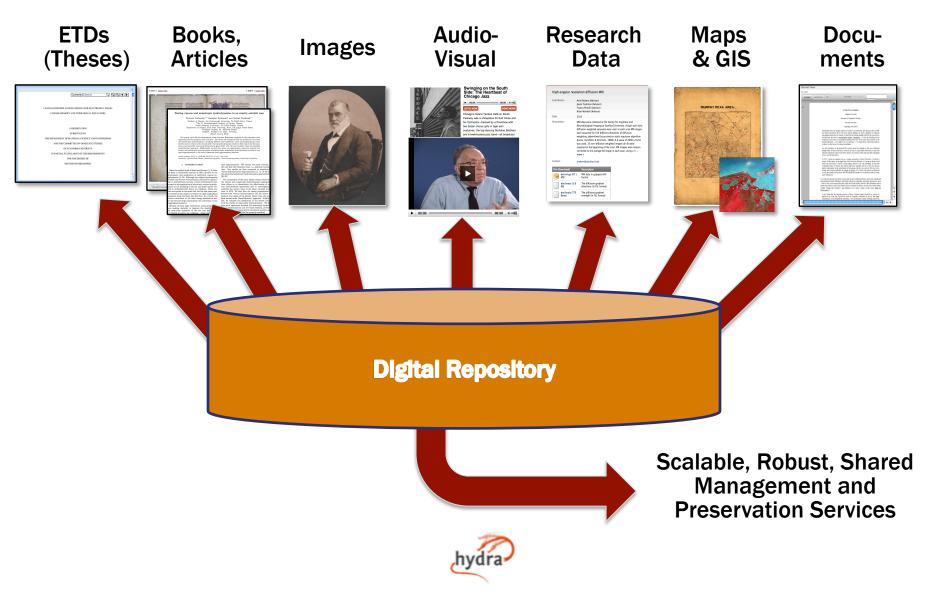
Preservation(?)

Effective?

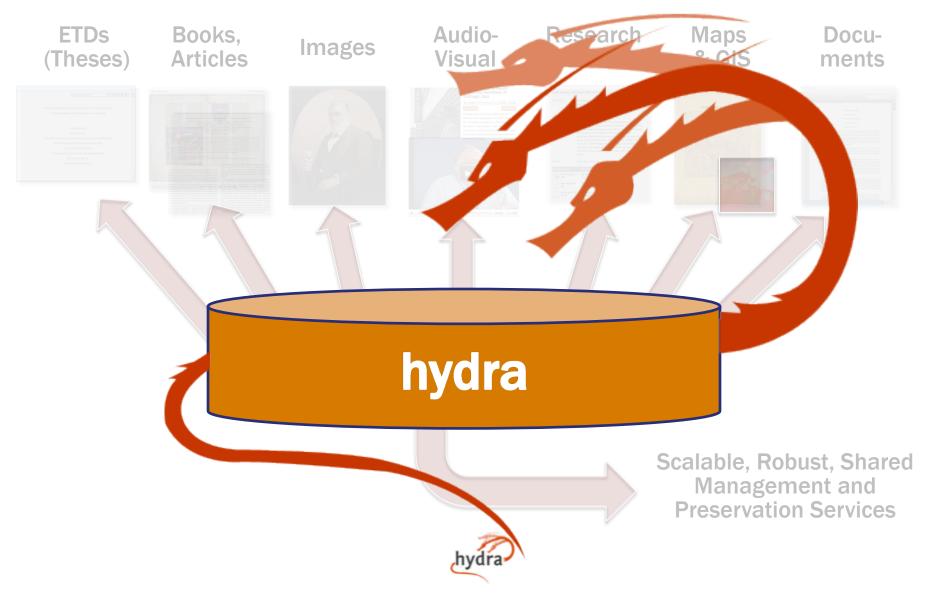
Sustainable?



Repository-Powered Approach



One Body, Many Heads...



Hydra Technical Framework



CRUD in Repositories

Create/Submit/Edit (CUD)

Search/View (R)

Repository/ Persistent Storage

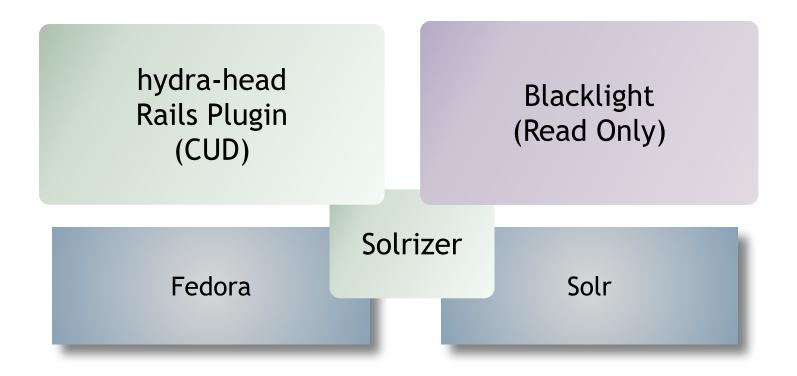


CRUD in Repositories Search/View Create/Submit/Edit (CUD) (R) Repository/



Persistent Storage

Major Hydra Components





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



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; 12 years of development, 3 releases
 - Vibrant community & funding under DuraSpace
- Fedora 4 in development; co-evolving with Hydra



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...

...are individuals, institutions, corporations or other groups that have committed to contributing to the Hydra community; they not only use the Hydra technical framework, but also add to it in at least one of many ways: code, analysis, design, support, funding, or other resources.

Hydra Partners collectively advance the project and the community for the benefit of all participants.





NORTHWESTERN UNIVERSITY





Stanford University























National Library of Denmark and Copenhagen University Librar



Cornell University







Yale University

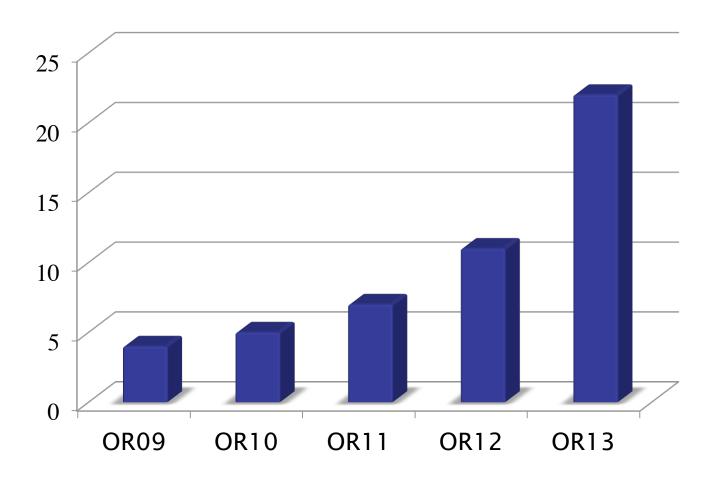
Code Licensing

- 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 Current State

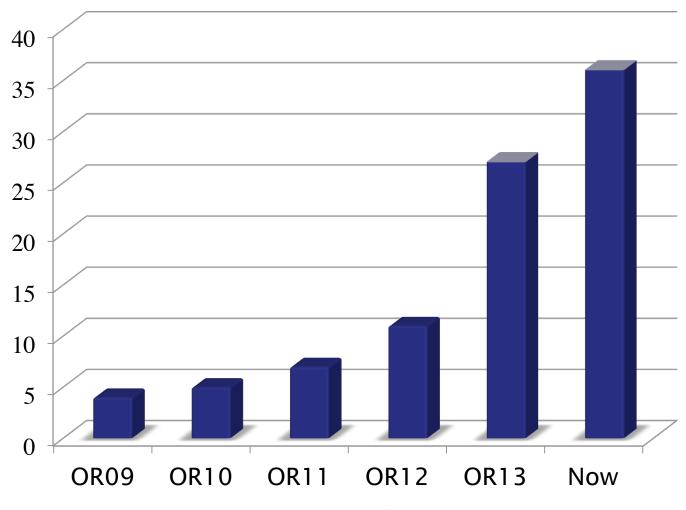


Hydra Partners





Hydra Partners and Known Users







University of Virginia, 2008



LibDevConX^3, Stanford, 2012: (Hydranauts and Fellow Travelers)



Hydra Partners @ UCSD, 2012



First Worldwide Hydra Connect @ UCSD, 2013



Trending

In a Nutshell, Project Hydra...

... has had 7,716 commits made by 62 contributors representing 60,904 lines of code

... has a codebase with a long source history maintained by a very large development team with stable Y-O-Y commits

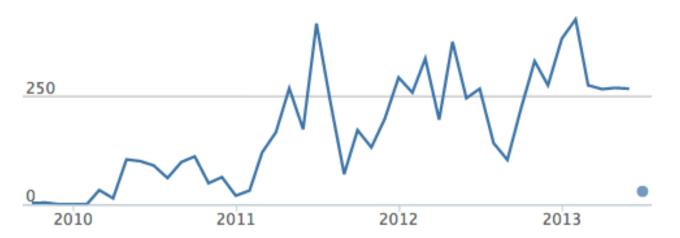
... took an estimated 15 years of effort (COCOMO model) starting with its first commit in October, 2009 ending with its most recent commit 6 days ago



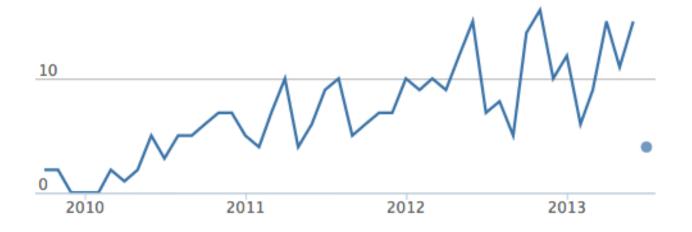
Commits per Month

Trending

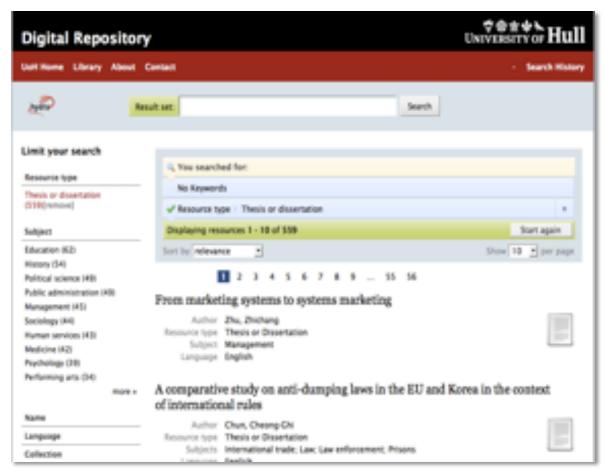
500



Contributors per Month



Hydra Heads: Institutional Repositories



University of Hull



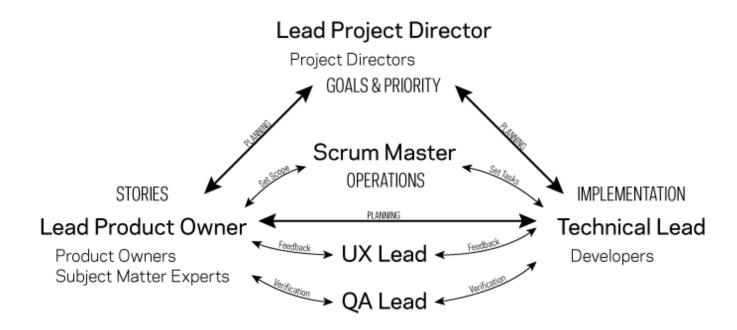
Hydra Heads: Institutional Repositories



ScholarSphere / Sufia @ Penn State



Hydra Heads: Institutional Repositories



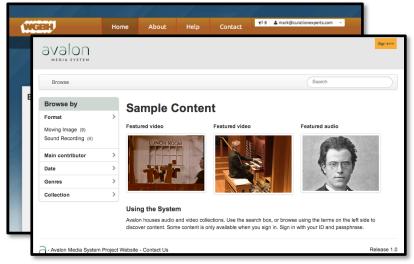
Hydramata

- 6-institution collaboration
- Build a shared Hydra head to serve as an IR
- Distributed, agile team

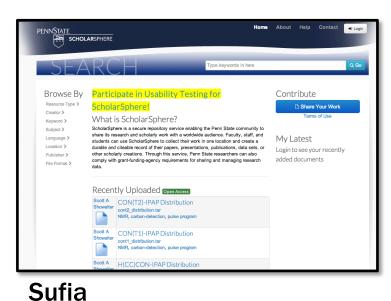


University of Notre Dame Northwestern University Indiana University University of Cincinnati University of Virginia

Hydra Heads of Note

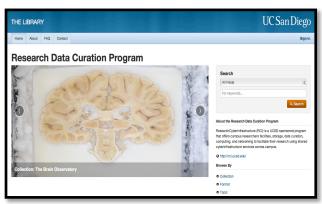


Avalon & HydraDAM for Media





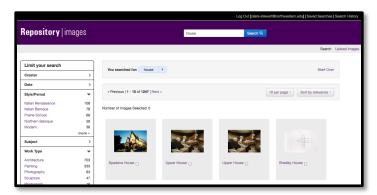




UCSD DAMS

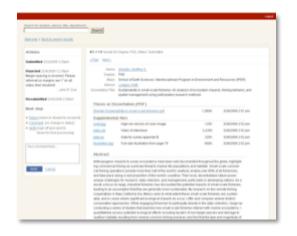


BPL Digital Commonwealth



Northwestern Digital Image Lib.

Hydra-based Applications at Stanford



SACTIVORES

Parker marigues

Bitel City: Two Search and Parkel City I had a feet figures of City I had

State by Selector Comments (in personnel Service Servi

ETD's – Electronic Theses & Dissertations

SALT – Self-Archiving Legacy Toolkit



Argo – Repository Reporting and Management

Hypatia – Archives & Special Collections

EEMs – Everyday Electronic Materials



SDR - Web UI (aka Hydrus)

Sustainability

No animals were harmed in the making of this film.

project / community



Current State of Hydra

- Doubled in size, thrice, since its founding
- Achieved all its original goals
- Entering a new phase



Three critical priorities

- 1. Technical framework continue to advance to enable easy code sharing and reuse
- 2. Community framework evolves to accommodate growth in partners.
- 3. Create a full suite of solution bundles



Eight strategic priorities

- 1. Develop solution bundles
- 2. Develop turnkey applications
- 3. Grow the Hydra vendor ecosystem
- Codify a scalable training framework to fuel community growth
- 5. Develop a documentation framework
- 6. Ensure the technical framework allows code sharing
- 7. Refresh and intensify the community ties
- 8. Grow the community of users and adopters



Blacklight



Blacklight for Repositories

- Repository-agnostic, feature-rich, contentaware, turnkey access interface
- Aggregate content from multiple repositories, with links back to source systems
- Vibrant, multi-institutional, open source community on its own
- Can be used independently, or as the first component of, Hydra



Blacklight

- Who's Using It?
 - Scores (probably hundreds) of institutions
 - Notables: UVa, Stanford, Columbia, Cornell, Princeton, Yale,
 Johns Hopkins, WGBH, BPL, NYPL, Danish Technical University,
 Tufts, Indiana University, Univ. of Wisconsin, et al.
 - Every Hydra institution
- Remarkably strong community
 - Regular workshops and tutorials at Code4Lib, DLF, et al.
 - 328 members on <u>blacklight-development@googlegroups.com</u>



Blacklight Development

In a Nutshell, Blacklight...

- ... has had 2,135 commits made by 49 contributors representing 10,855 lines of code
- ... has a codebase with a long source history maintained by a large development team with increasing Y-O-Y commits
- ... took an estimated 3 years of effort (COCOMO model) starting with its first commit in October, 2009 ending with its most recent commit 1 day ago

30 Day Summary

Mar 2 2014 — Apr 1 2014

79 Commits

6 Contributors including 2 new contributors

12 Month Summary

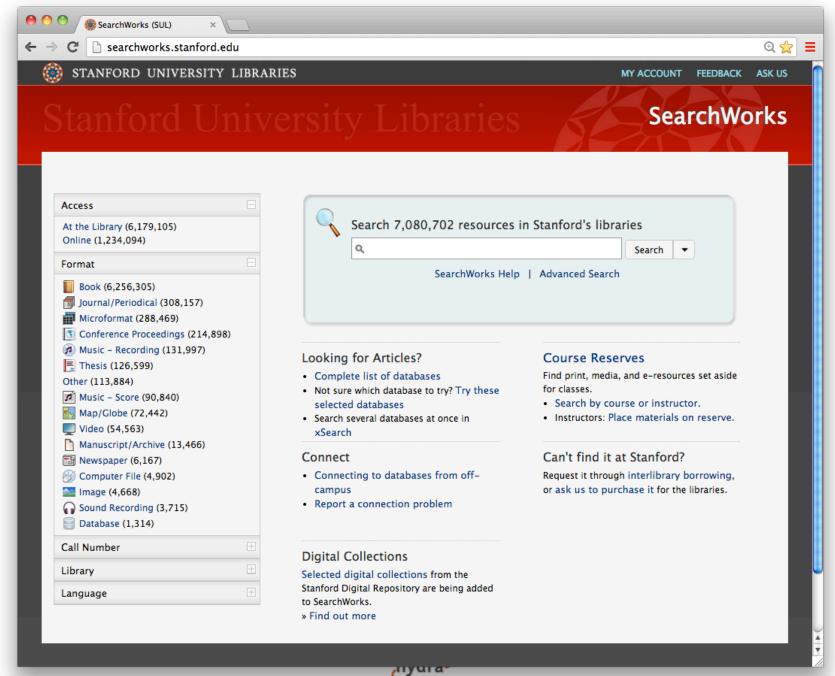
Apr 1 2013 — Apr 1 2014

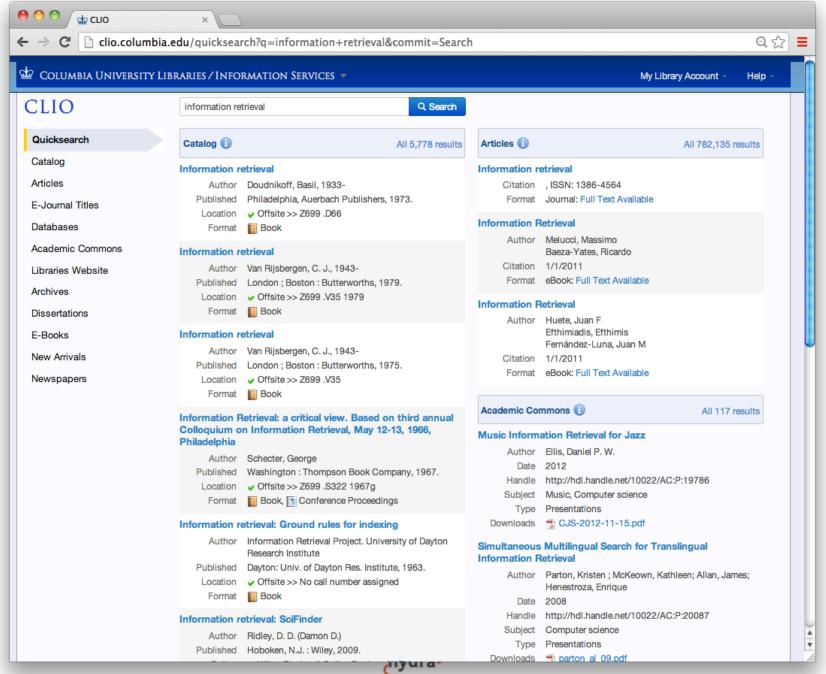
508 Commits

Up +152 (42%) from previous 12 months

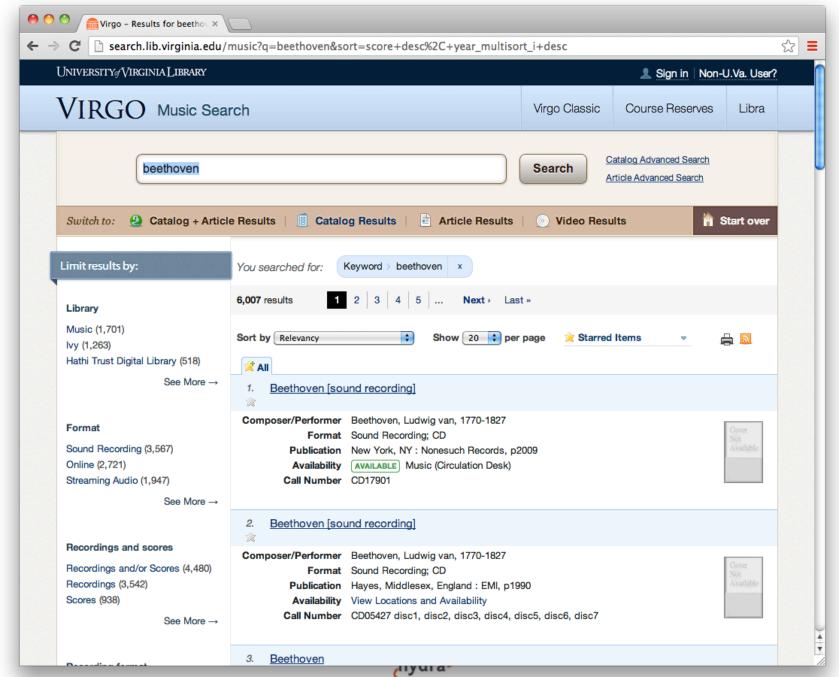
19 Contributors

Down -2 (9%) from previous 12 months

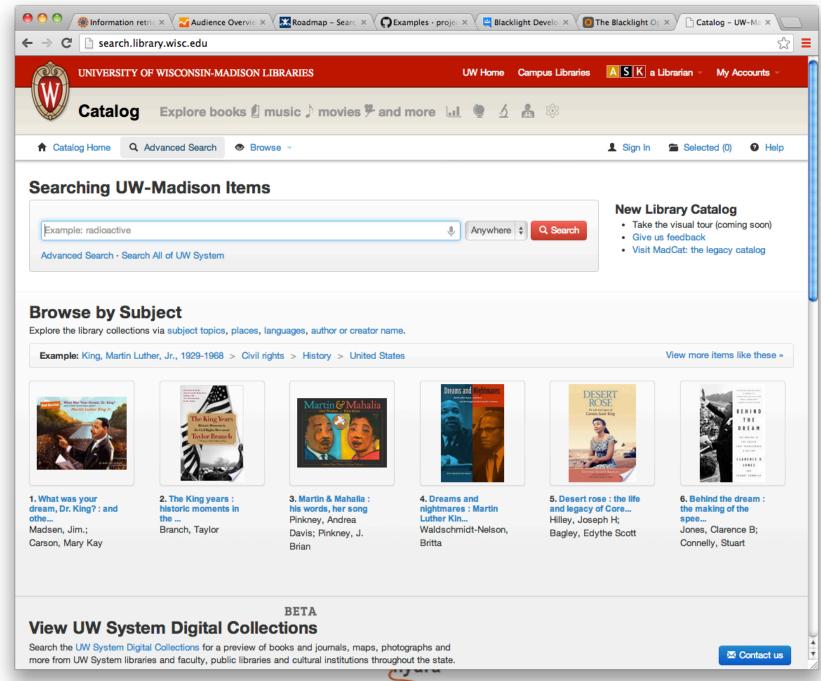




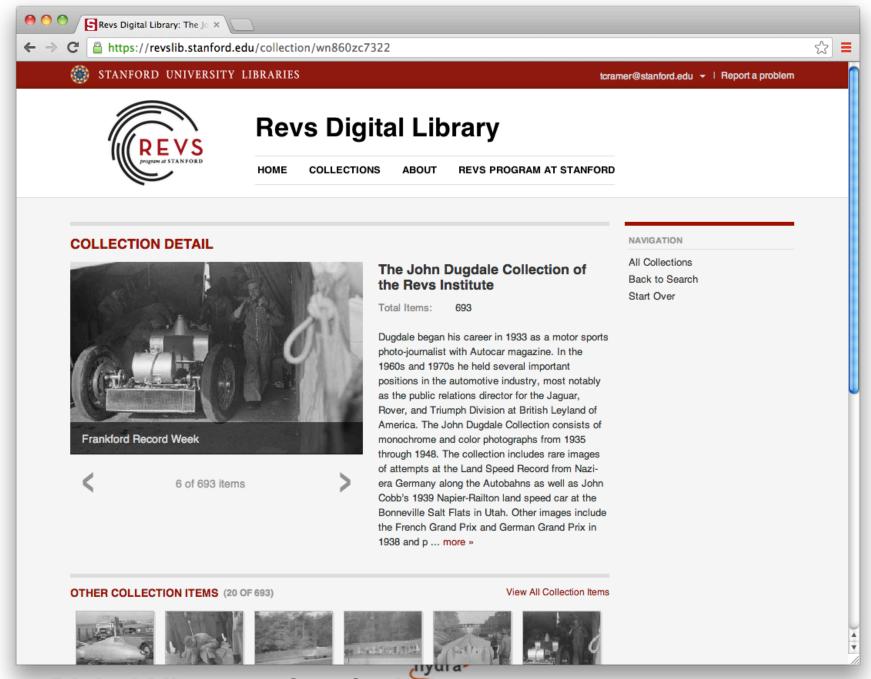
CLIO at Columbia University - Bento Box results w/ article search



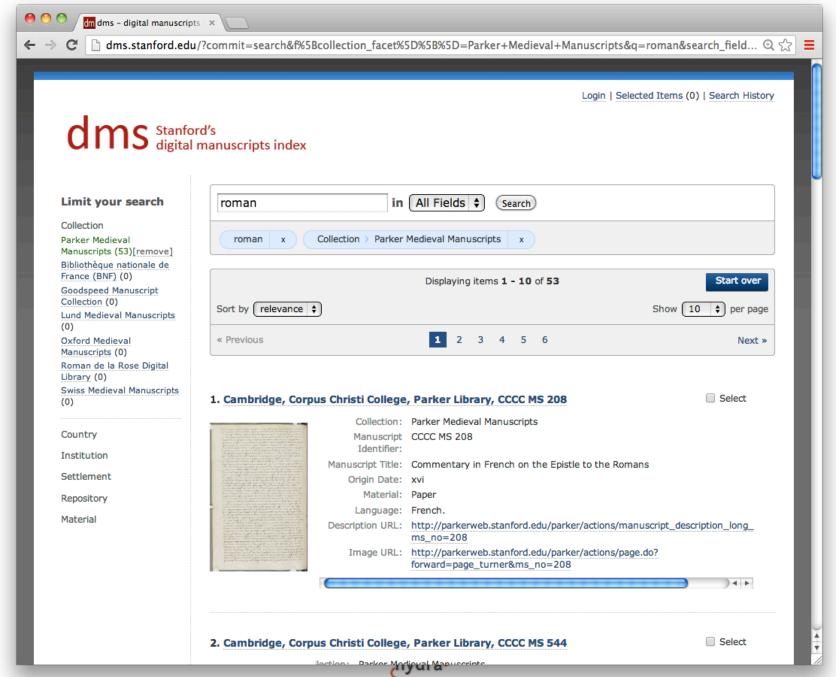
Virgo at University of Virginia – Music Search View



Forward at University of Wisconsin





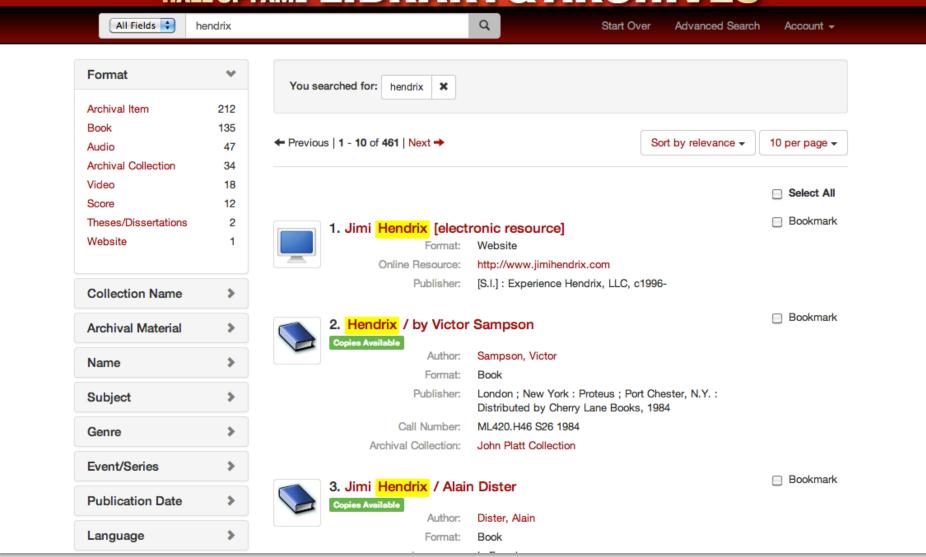


Digital Manuscript Index at Stanford: Blacklight for MS Catalog

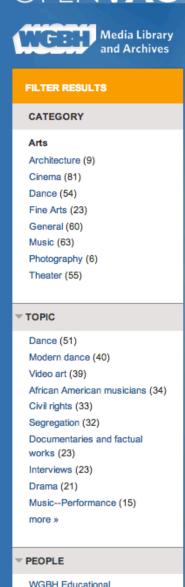


Bassi-Veratti at Stanford: Blacklight for EAD and digital archive

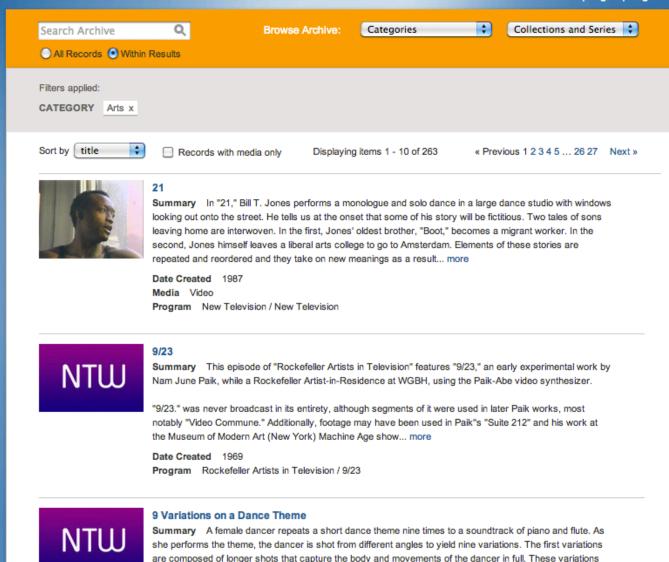
ROCK&ROLL LIBRARY& ARCHIVES



OPENVAULT



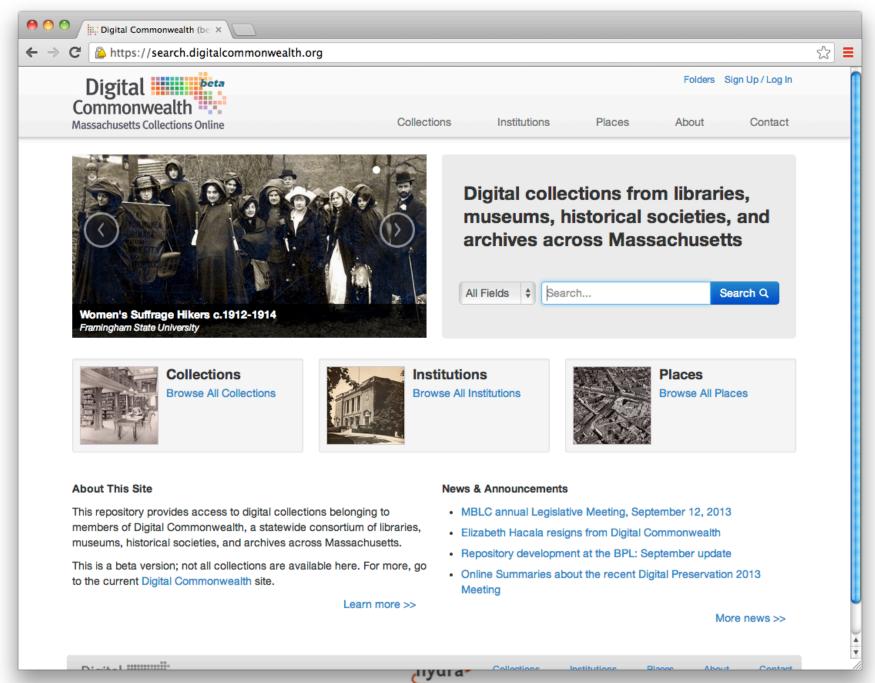
Foundation (263)

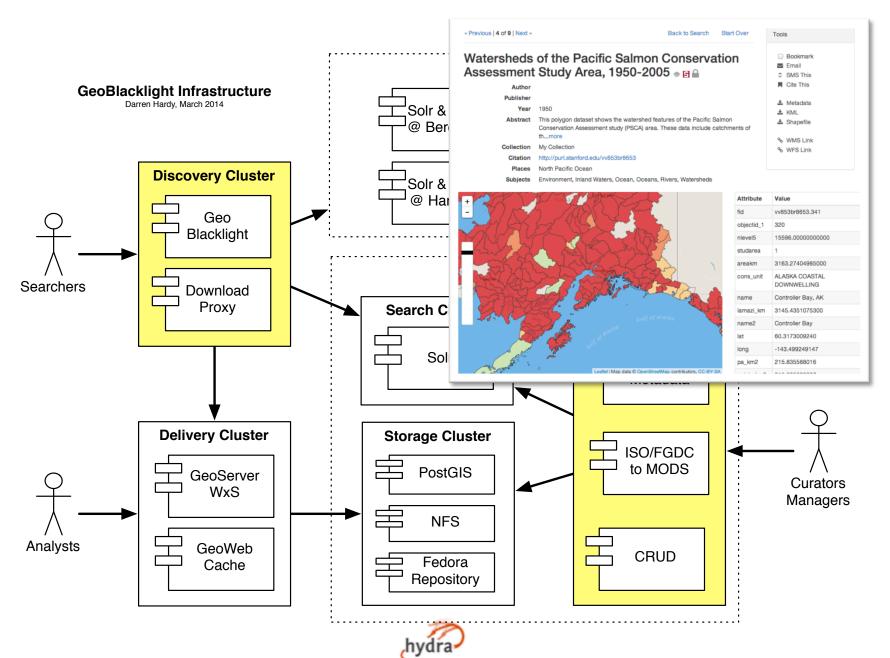


are accompanied by a soundtrack of slow-tempo piano. The later variations are composed of many shorter

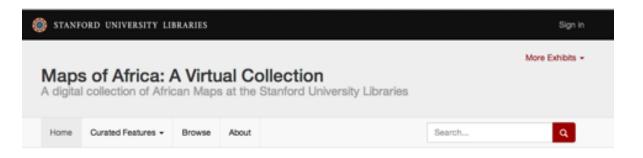
shots that capture the components of the dancer's body and movements... more

OpenVault: Blacklight for Video at WGBH





GeoBlacklight: Blacklight for geospatial resources



Limit your search



Stanford's Maps of Africa Collection became a major research resource in August of 2001 with the acquisition of the Oscar I. Norwich Collection of Maps of Africa and its islands, 1486 – ca. 1900. This acquisition added 316 maps to Stanford's existing holdings of Maps of Africa. The collection was once again augmented in 2010 with the purchase of 294 maps from Caroline Batchelor, and then again in 2011 with a small set of maps from Rodney Shirley. Altogether, the Maps of Africa Collection holds over 700 maps, 632 of which are searchable and viewable through Stanford's Image Gallery.









A Map of Africa

1793

Cape Colony, Natal &c.

Africa from the latest Authorities

1806

Pellentesque habitant morbi tristique senectus et netus et malesuada fames ac turpis egestas. Vestibulum tortor quam, feuglat vitae, ultricies eget, tempor sit amet, ante. Donec eu libero sit amet quam egestas semper. Aenean ultricies mi vitae est. Mauris placerat eleifend leo.



A New and Accurate Map of Africa.

Pellentesque habitant morbi tristique senectus et netus et malesuada fames ac turpis egestas. Vestibulum tortor quam, feugiat vitae, ultrícies eget, tempor sit amet, ante. Donec eu libero sit amet quam egestas semper. Aenean ultrícies mi vitae est. Mauris placerat eleifend leo.



The Hydra-Blacklight Way

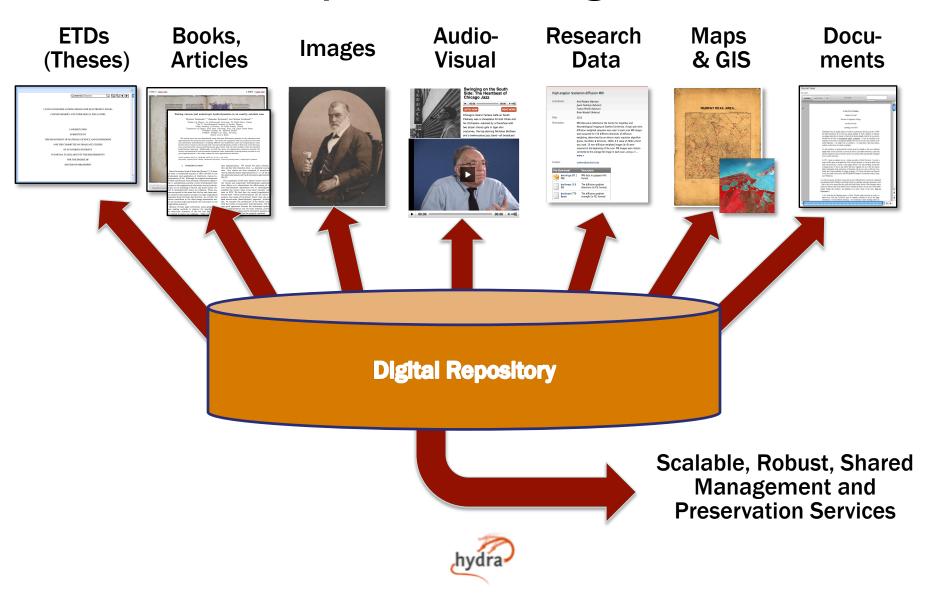


Why Hydra?

- Need a flexible, rich, robust repository solution
- Be part of a community
 - Shared resources, shared risk, shared advances
- "We" is smarter than "me"
 - Especially in this crowd
- Control your own destiny
- Build institutional capacity
- Practice best practices
- The way work is getting done now

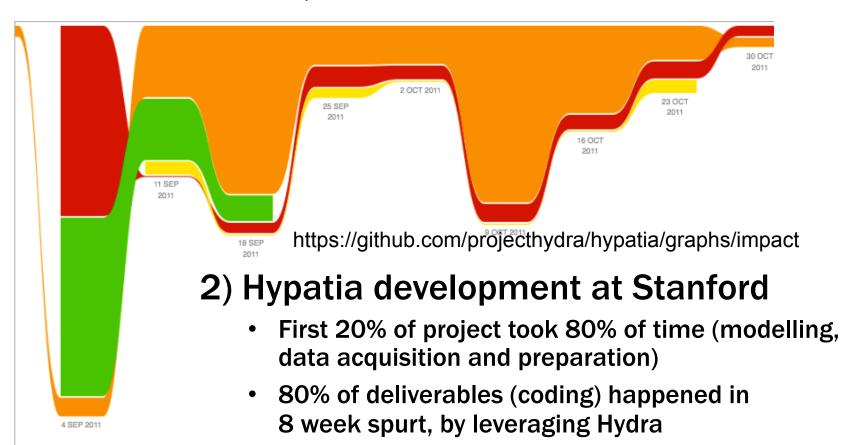


Breadth & Depth: Access, Mgmt & Pres



If You Want To Go Fast ...go alone. ...use Hydra

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



Becoming a Hub

- Working Groups forming
 - Repository Managers
 - Archivsists
 - Hydra RDF
 - Fedora...

Jobs



Sustainability...



- For institutions Partners help shoulder the load
- For other projects
 Post-project maintenance & governance
- For leveraging funders' investments
 Adoption by growing user community

The Hydra Way

- Not a directed project
- Investment in a framework, not an application
 - Contributions back to core code base
- Investment in a community, not a vendor
 - Contributions back to community: Training, documentation, modeling, evangelism, support
- Travel / face-time





http://projecthydra.org