GET A HEAD ON YOUR REPOSITORY

Tom Cramer
Chief Technology Strategist
Stanford University Libraries

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.*

hydra

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.
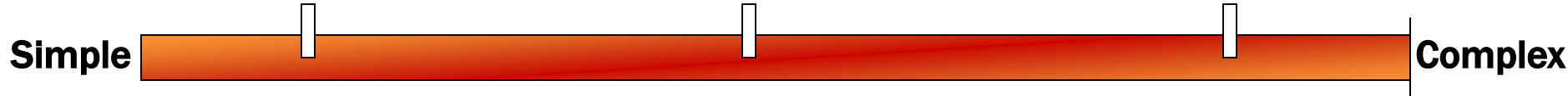
For Instance...

| | ETD Deposit System | General Purpose Institutional Repository | Digitization Workflow System |

Simple ▮▮▮▮▮ Complex

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

Hydra Heads: Emerging Solution Bundles
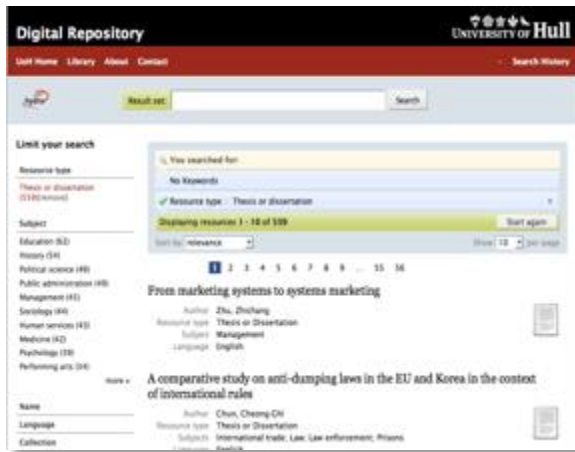
Institutional Repositories

University of Hull

University of Virginia

Penn State University

Images

Northwestern University

(Digital Image Library)

hydra
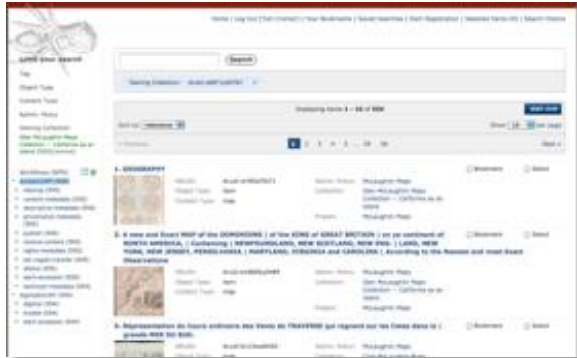
Hydra Heads: Emerging Solution Bundles

Archives & Special Collections

Stanford University

University of Virginia

Rock & Roll Hall of Fame

Media

Indiana University

Northwestern University

Rock & Roll Hall of Fame

WGBH

hydra

Hydra Heads: Emerging Solution Bundles

Workflow Management (Digitization, Preservation)

Stanford University
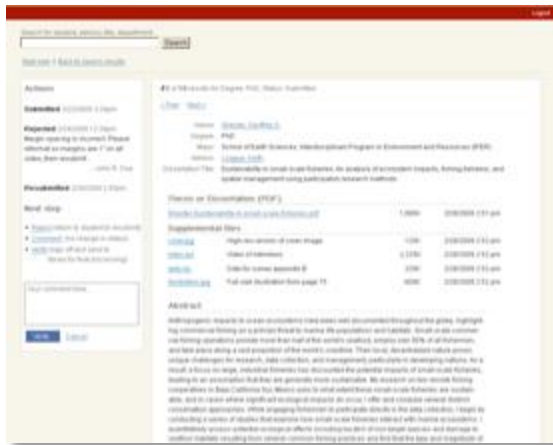**University of Illinois – Urbana-Champagne**
Northwestern University

Exhibits

Notre Dame

Hydra Heads: Emerging Solution Bundles

ETDs
Stanford University
University of Virginia
Etc.

(Small) Data
everyone...

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

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

- Initially a joint development project between Stanford, Univ of Virginia, and Univ of Hull
  - Close collaboration with DuraSpace / Partnership with MediaShelf, LLC

- Complementary strengths and expertise

- Stanford University
- DuraSpace
- Northwestern University
- Indiana University
- The Royal Library of Denmark
- Boston Public Library

- University of Hull
- MediaShelf
- Columbia University
- London School of Economics
- Data Curation Experts

- University of Virginia
- University of Notre Dame
- Penn State University
- Rock and Roll Hall of Fame
- WGBH

Hydra Partners

Bar chart showing Hydra Partners growth: OR09 = 4, OR10 = 5, OR11 = 7, OR12 = 11, OR13… = 16
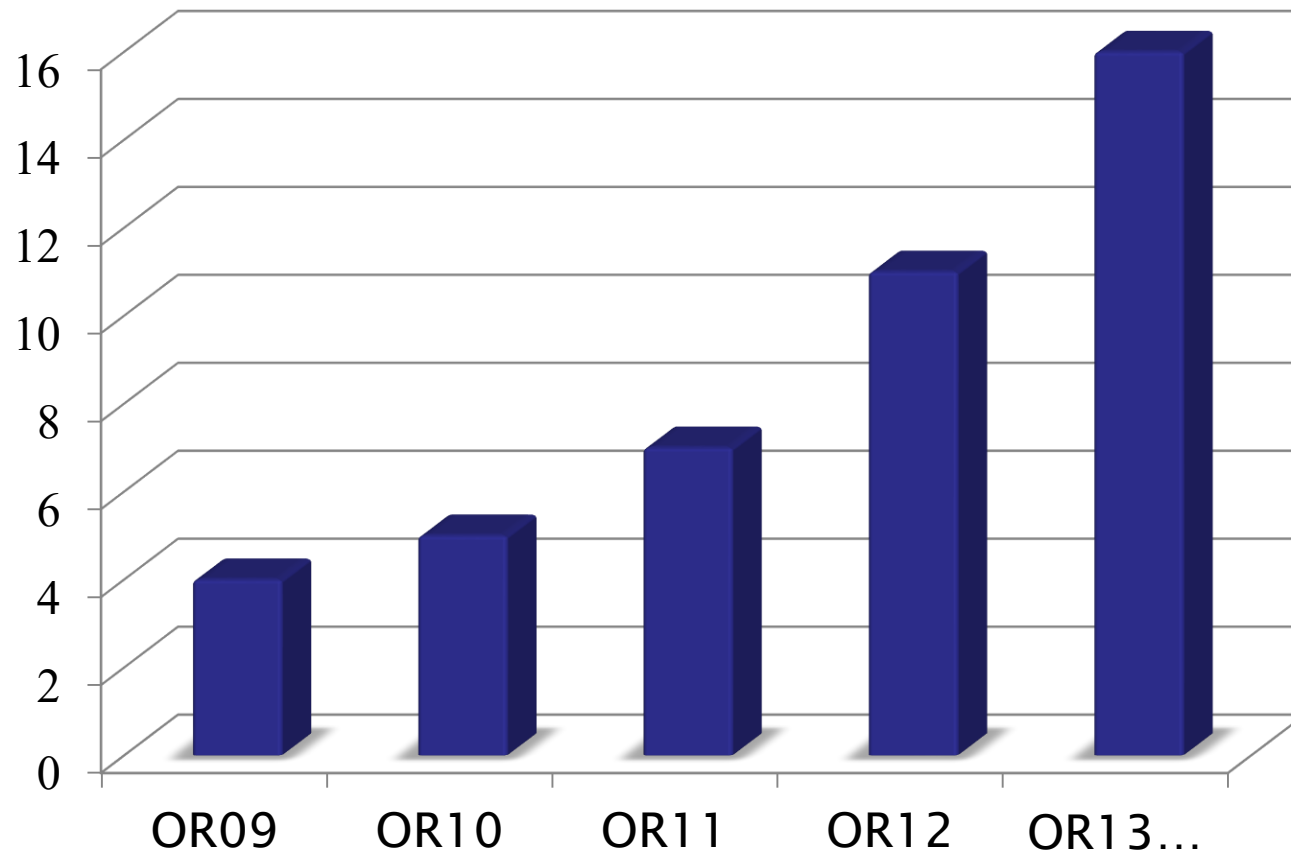
OR = Open Repositories Conference

Bigger things.... (Hydra Partners Meeting, UCSD Dec 2012)

From small beginnings. University of Virginia, 2008

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.

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

Sustainability

No ~~animals~~ *grants* were ~~harmed~~ *abused*
in the making of this ~~film~~ *project / community*.

hydra

Argo – Workflows Reporting & Controls

Workflow: googleScannedBookWF

	Process	Description	Waiting	Hold	Ready	Error	Skipped	Completed
1	register-object	Register new objects in DOR	1	0	1	0	0	187494
2	descriptive-metadata	Make descriptive metadata (MODS & DC) from Symphony/MARC	3	0	2	150	0	187342
3	google-convert	Request GRIN items to be converted for download	185	0	184	4	0	187306
4	google-download	Download content from Google	189	0	36	2032	0	185274
5	process-content	Process content, create tech/rights/provenance/content metadata	2222	0	1	10	0	185263
6	sdr-ingest-transfer	Prepare bagit transfer package, push to SDR-Stage	2529	0	297	0	0	184966
7	sdr-ingest-deposit	Call from SDR on completed ingest	0	0	0	0	0	163424
8	shelve	Shelve contents in the Digital Stacks	2232	0	0	1	0	185262
9	cleanup	Cleanup workspace; make room for more!	2585	0	0	24	0	184886

hydra

Argo – Digital Object Inspection

ScholarSphere

ScholarSphere
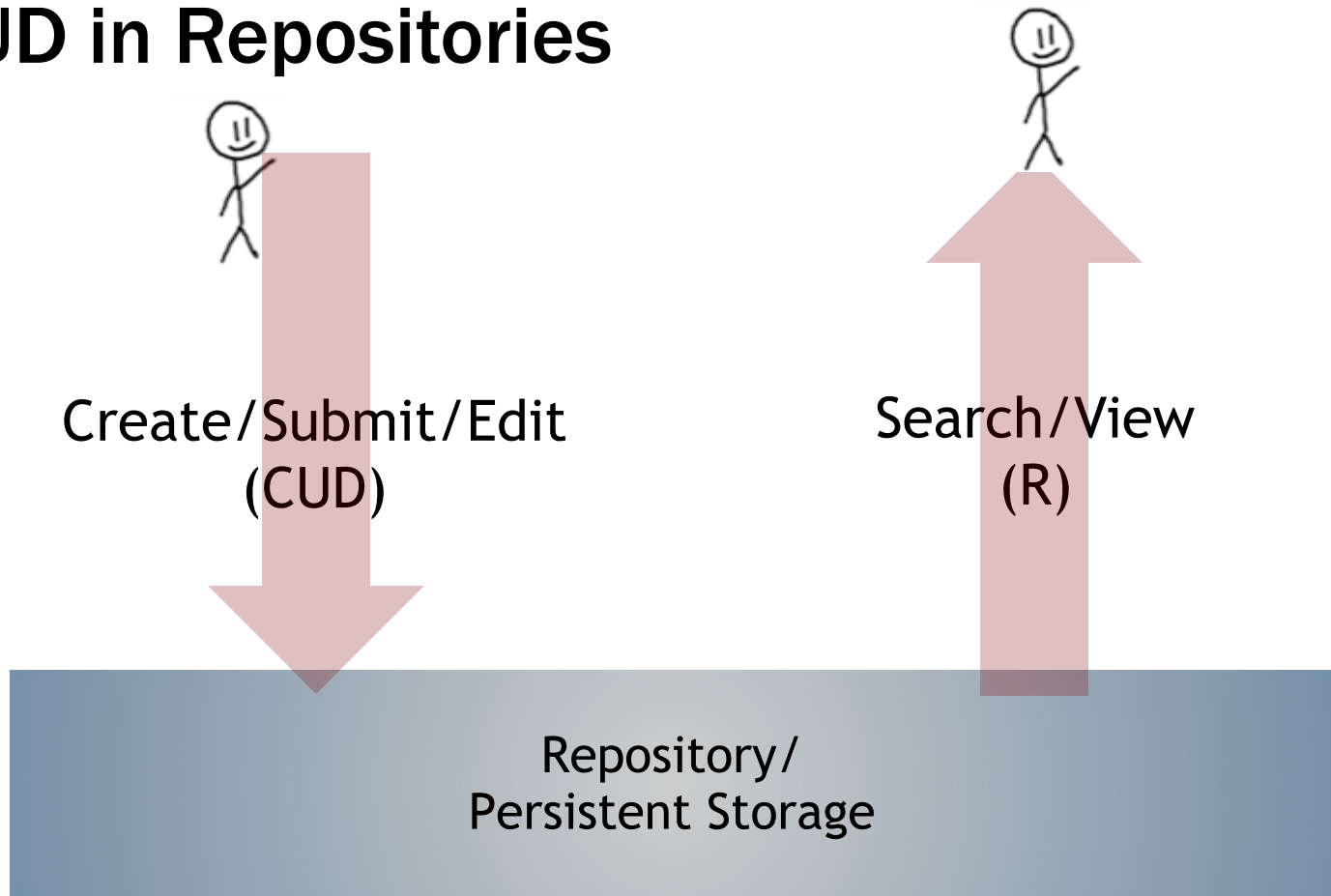
CRUD in Repositories

Create/Submit/Edit
(CUD)

Search/View
(R)

Repository/
Persistent Storage

CRUD in Repositories

Major Hydra Components

hydra-head
Rails Plugin
(CUD)

Blacklight
(Read Only) (RO)

Solrizer

Fedora

Solr

hydra

http://projecthydra.org