
Samvera

**an open source solution for
building digital repositories**

Mark Bussey, Karen Cariani, and Robin Ruggaber



What we'll cover

- What is Samvera?
- Community structure and resources
- Hands on examples exploration
- Technical Overview & Resourcing
- Getting started with Samvera
- How to pitch Samvera: Benefits
- Q & A



Let's Start with You

- Introductions
- What are your goals for the morning?
- Where are you in the process?
- What are your concerns about Samvera?

Samvera Overview

what is it?

what does it do?

why is it different?

What is Samvera?

- 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 extending and enhancing the solutions
 - ⇒ *If you want to go fast, go alone. If you want to go far, go together.*



Why Samvera?

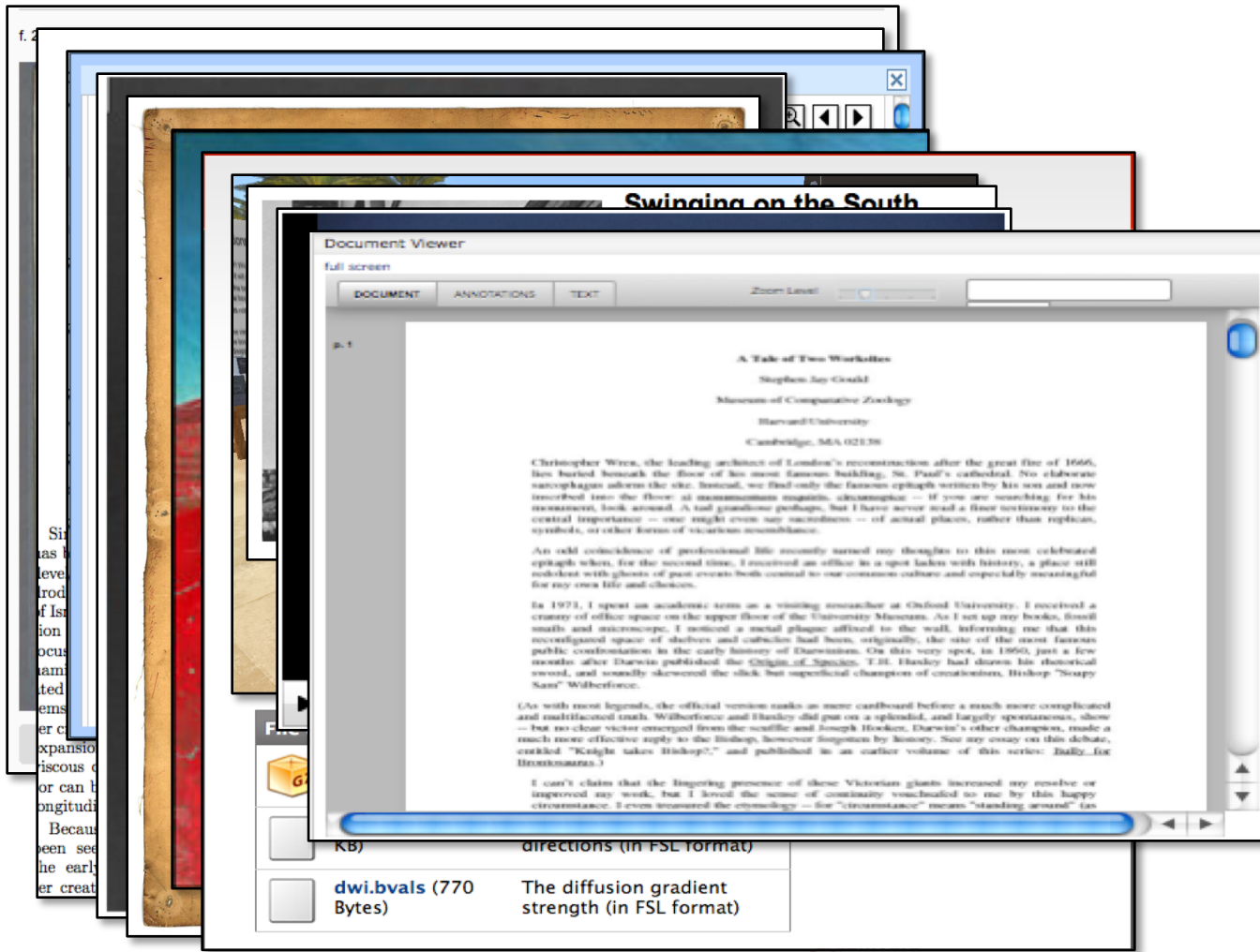
- Need a flexible, rich, robust repository solution
- Be part of a community
 - Shared resources, shared risk, shared advances
- “We” is smarter than “me”
 - Especially in this crowd
- Control your own destiny
- Build institutional capacity
- Practice best practices
- The way work is getting done now



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



ETDs
(Theses)



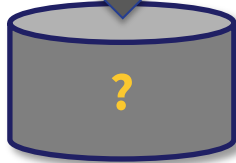
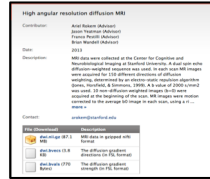
Books,
Articles



Images



Audio-
Visual



Research
Data



Maps
& GIS



Documents

Management

Access

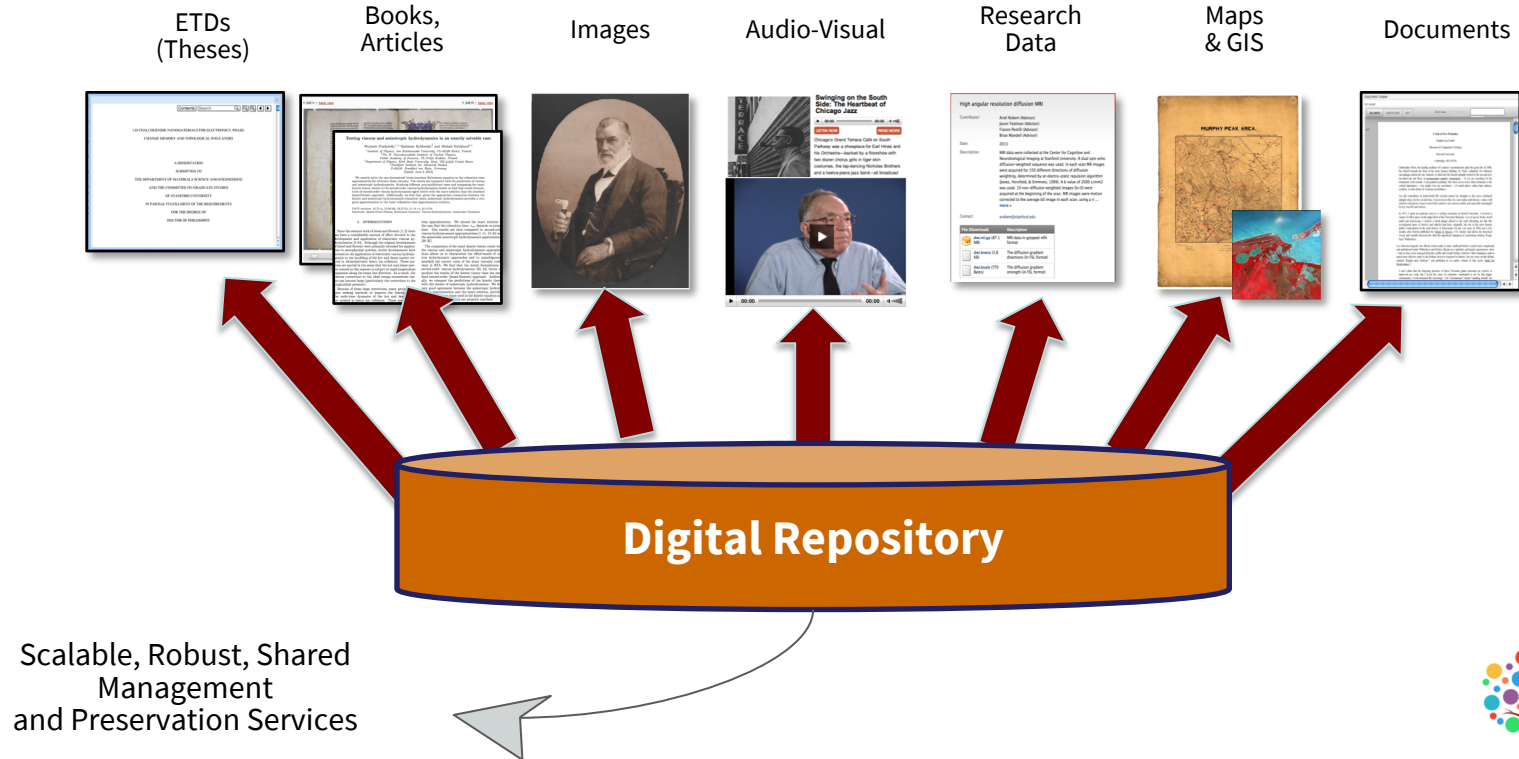
Preservation?

Effective?

Sustainable?



Repository Powered Approach



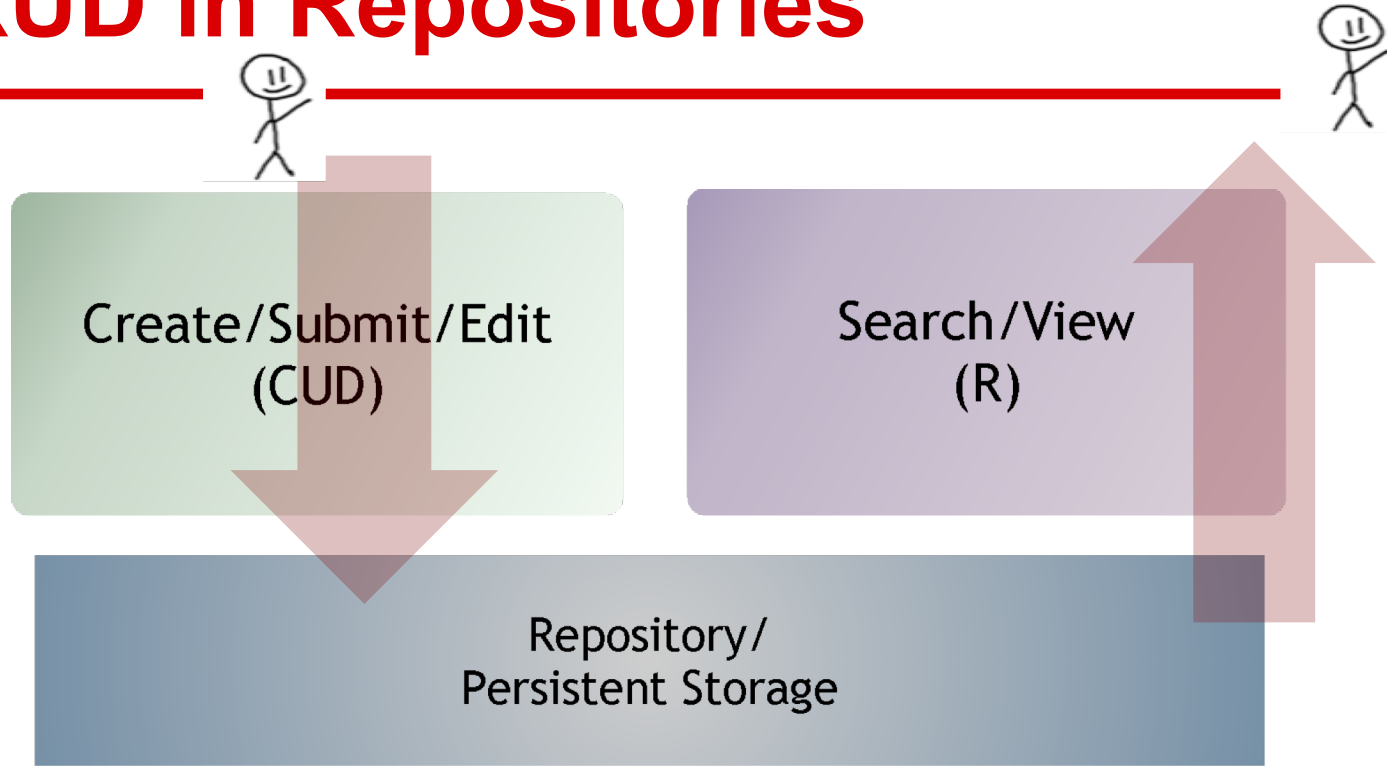
CRUD in Repositories

Create/Submit/Edit
(CUD)

Search/View
(R)

Repository/
Persistent Storage

CRUD in Repositories



Used by:

- Large Universities
- Small Universities
- Colleges
- Public Broadcasting
- Government Ministry
- National Libraries
- National Lab
- Small Research Labs
- National Digital Repository
- Statewide Digital Libraries
- Rock & Roll Hall of Fame
- Chemical Heritage Foundation
- Museum of Performing Arts
- A Shakespeare Festival

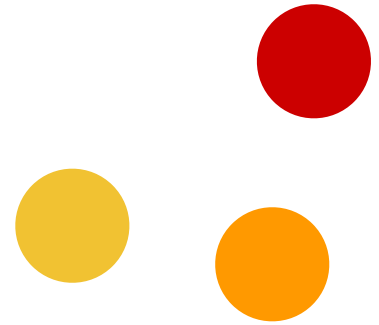
Used for:

- Self-deposit system
- Digital Collections System
 - Sheet Music
 - Architectural resources
- Electronic Theses & Dissertations
- Digital Image System
- Media Management
- Media Preservation System
- Research Data Management
- Digitization Workflow System
- Digital Preservation System
- Digital Archives System
- Library Monograph Acquisitions System
- And more!



Next Up:

How we work



How do I get involved?

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.

Samvera Philosophy: A 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.”



The Samvera 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



Community

- Conceived & executed as a distributed, collaborative, open source effort
- Initially a joint development project between Stanford, Univ of Virginia, and Univ of Hull
- Samvera 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
 - But no central planning, no Project Director, no “one” architect



Community Structure

- **Steering Group** (elected by partners)
Stewards of the project - individual membership
- **Partners** (Institutional)
Formally committed institutional membership
- **Working & Interest Groups**
Typically aligned groups that span organizations & roles
- **Committers**
Developer community
- **Community adopters**



Samvera Partners

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

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



Training & Connecting

- **Good:** read the documentation and learn-by-doing
- **Better:** connect through social channels
 - Email lists, IRC, weekly phone calls
 - Github interactions
- **Best:** meet in person. Face time fuels collaboration!
 - Samvera Connect: annual, worldwide, community meeting
 - Samvera (Partner) meetings: 3-6 ad hoc, topic-based meetings per year (usually)
 - Regional Groups: convening ad hoc

Get Involved: Many Hands Make Light Work

- Getting involved with Samvera can be at different levels:
 - Adopter
 - Developer
 - Partner
- Need to ask:
 - What involvement is required to enable you to meet your needs?
- But also,
 - If we contribute back, what more do we gain?



Stay Connected

- Website (<http://samvera.org>)
- Wiki page (<https://wiki.duraspace.org/display/samvera/>)
 - Partner meeting agendas & notes
 - Events & Details
- Monthly Partner calls
- Partner listserv
- Samvera Partner meetings & Samvera Connect
- Identify a Samvera Partner doing similar work and get in touch
- Identify need for vendor input and make contact to discuss options
- Come to events (like this one!)



Code Licensing

- All Samvera code is available under Apache License, Version 2.0
- All code commitments are being managed through Contributor License Agreements
 - Individual – each developer is clear about what they are contributing
 - Corporate – 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

Interest & Working Groups

- Collection Extensions Working Groups
- Component Maintenance Working Group (Phase 1)
- DSpace/Samvera Interest Group
- Hyrax Analytics Working Group
- Repository Management Interest Group
- Samvera Applied Linked Data Interest Group(formerly Working Group)
- Samvera Archivists Interest Group (formerly Working Group)
- Samvera Code Stability Working Group
- Samvera CONTENTdm Migration Interest Group
- Samvera Contribution Model Working Group
- Samvera Data Mapper Working Group
- Samvera Documentation Working Group(enrolling)
- Samvera EZID to DataCite DOI Working Group
- Samvera Geo Predicates Working Group
- Samvera Geospatial Interest Group
- Samvera Governance Working Group
- Samvera Hyrax Batch Import Working Group
- Samvera Marketing Working Group
- Samvera Metadata Interest Group (formerly Working Group)
- Samvera Newspapers Interest Group
- Samvera Permissioning Analysis and Design Working Group
- Samvera Interest Group for Advising the Hyrax Roadmap (SIGAHR)
- Samvera URI Selection Working Group
- Samvera User Experience Interest Group



Community Resources

website

wiki

slack channels

e-mail lists

meetings- in person and virtual

the people themselves!

Communication Channels

- **Slack** - for real-time interaction
- **E-mail** - for asynch and longer discussions
- **Community Calls** - Samvera Tech & Community calls, Virtual Connect
- **In Person** - Connect, regional meetings, other conferences, & workshops

Community Resources

- **Website** - for the managers

<https://samvera.org>

- **Wiki** - for the doers

[https://**samvera.org/wiki**](https://samvera.org/wiki)

- **Github** - for the makers

<https://github.com/samvera>

<https://github.com/samvera-labs>

<https://github.com/samvera-deprecated>



First Up:

Example Sites



What are folks doing with it?

INDIANA UNIVERSITY

Media Collections Online

- Built using the *Avalon Media System*
- Rich tool for management and discovery of audio and visual resources
- Provides gated discovery and streaming authentication

Live Site: <https://media.dlib.indiana.edu>



Alexandria Digital Research Library

- Fedora 4 + RDF based
- Broadly focused Institutional Repository hosting multiple content types
- Includes a mix of born digital and digitized content

Live Site: <https://www.alexandria.ucsb.edu>



Earthworks

- Heterogeneous GIS Content
- Search and browse augmented by spatial search functions
- Includes content harvested from multiple sources

Live Site: <https://earthworks.stanford.edu>



Tufts Digital Library

- Fedora content pre-dates Hydra & Samvera, currently migrating to Hyrax
- Multi-app ecosystem (management, discovery, image library, others)
- Supports multiple content types

Live Site: <https://dl.tufts.edu>



Emory ETD Repository

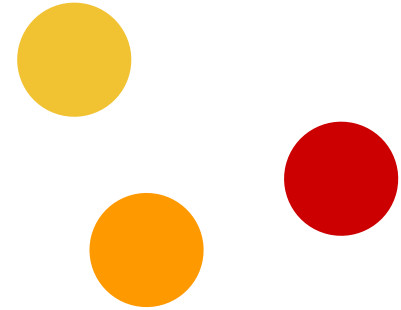
- Greenfield Hyrax Development
- Electronic Theses and Dissertations focus
- Student submission and workflow focused

Live Site: <https://etd.library.emory.edu>



Next Up:

Hands-On



Any questions so far?

Create an account

1. Visit <http://demo.curationexperts.com/>
2. Click the Login link
3. Click on “[Sign Up](#)”
(just under the Login button)
4. Enter your e-mail, choose a password, and click on “Sign Up”

Deposit a new work

1. Click on the logo in the upper left corner
2. Click on the [“Share Your Work”](#) button
3. Select a type of work
4. Add required metadata, attach a file, choose visibility, accept the deposit agreement
5. Save your work

Edit your work

1. Choose “Dashboard” from the menu in the upper right corner
2. Click on [Works](#)
3. Use the “Actions” dropdown to choose edit
4. Enter some data in “Additional Fields”
5. Add a file or two, update versions

Browse content

1. Click on the logo in the upper left
2. Click on the [search icon](#) to browse the repository
3. Use search and facet to narrow results
4. View some items
5. Log-out - how does your experience differ?

Deposit a new work

1. Reminder: Logo, Share, Select
2. Add required metadata, attach a file, choose visibility, accept the deposit agreement
3. Confirm “one step approval” is selected under Relationships
4. Save your work

View your work?

1. Reminder: dashboard, works
2. What do you see?
3. What does your administrator see?
4. Approve some works (Admins)
 - a. All Works
 - b. Review Submissions
 - c. Notifications

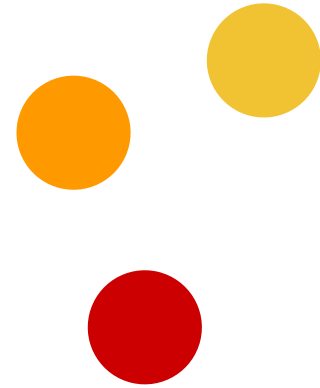
Other features to try

1. Gated discovery: Public, DCE, Private
2. Featured works (Admins)
3. Collections
4. Admin dashboard (Group Walkthrough)
5. Background Processing (Group)

Next Up:

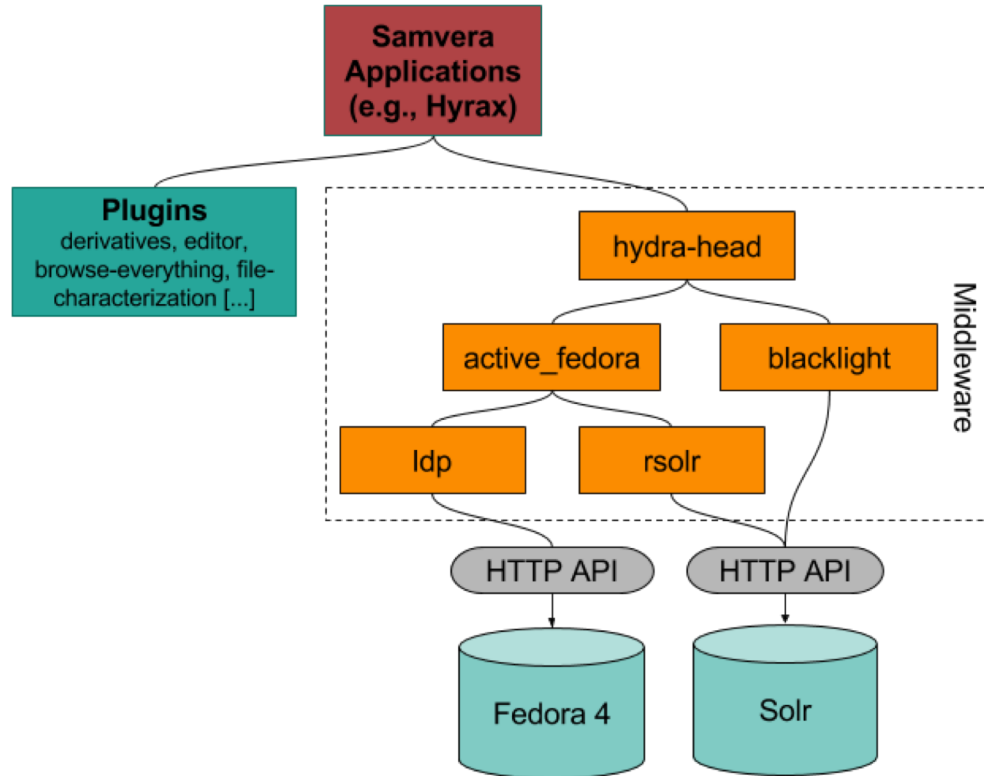
Tech Details

(totally optional)



Any questions so far?

Major Components



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 major releases
 - Vibrant community & funding under DuraSpace
- Fedora 4.0 released in Dec 2014; co-evolving with Samvera
 - Current release: 4.7.5 as-of February 2nd, 2018



A Note on Rails

- *Rapid* application development for web applications: “Convention over configuration”
 - 10x productivity
- Supportable: 14 years of production use with major implementations across the internet - Github, Airbnb, Square, and 1000s of others
- Testable: Rspec & Capybara 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

Hyrax Repository Engine

- Feature Rich
- Complex objects based on PCDM in Fedora 4
- Evolved from Sufia 7, one of most actively maintained and widely used community gems



Hyrax



Hyku

- Availability: DIY=now, Hosted=tbd
- Goal: turnkey ease of installation and management
- Goal: Provide configured version of Hyrax with multitenancy



A Note on Rails

- *Rapid* application development for web applications: “Convention over configuration”
 - 10x productivity
- Supportable: 14 years of production use with major implementations across the internet - Github, Airbnb, Square, and 1000s of others
- Testable: Rspec & Capybara 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



The Stack

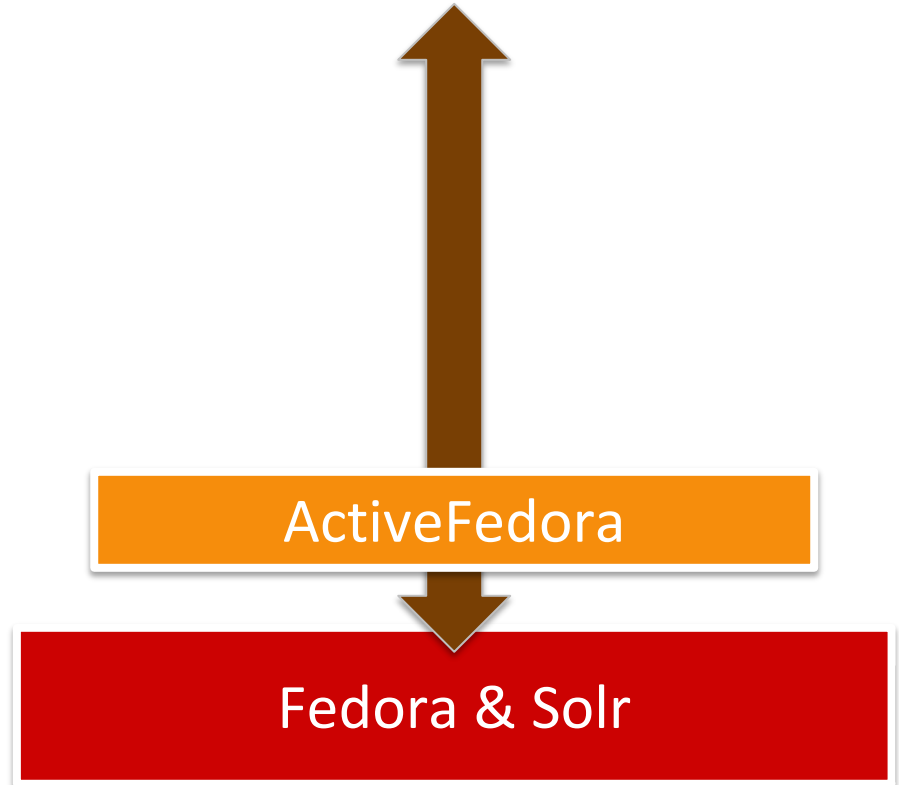
Samvera is designed to provide a rich front-end framework for a Fedora based repository. We use Solr to index, search, and discover content in the application.



Fedora & Solr

The Stack

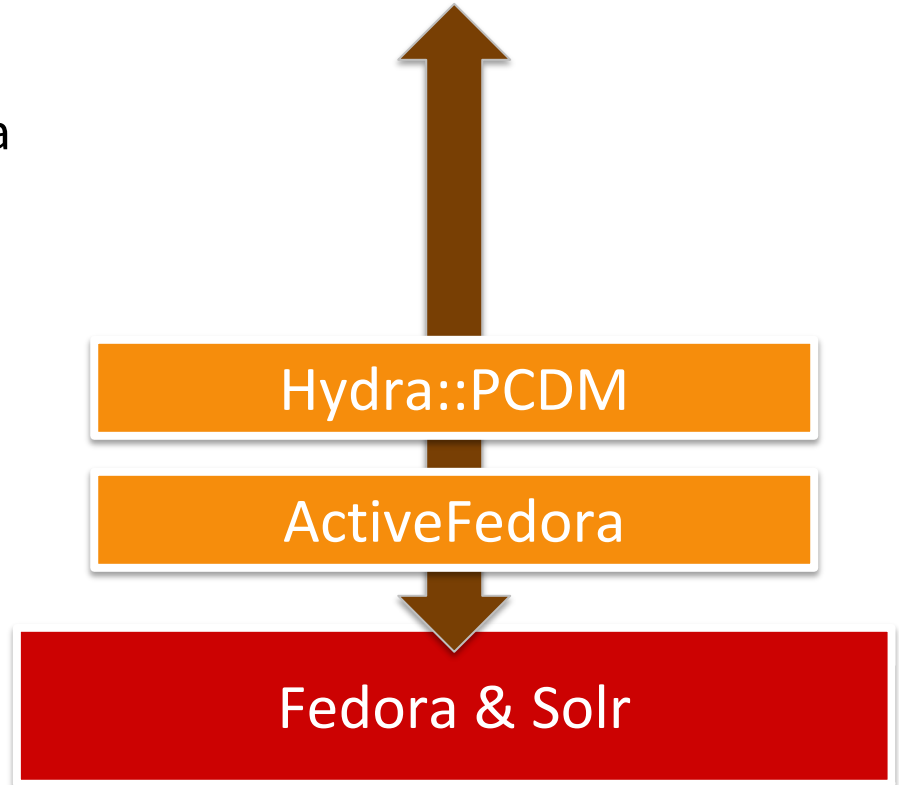
The ActiveFedora gem provides an interface for rails based applications to interact with Fedora and Solr. This allows developers to focus on user facing application features instead of re-inventing low-level repository interfacing.



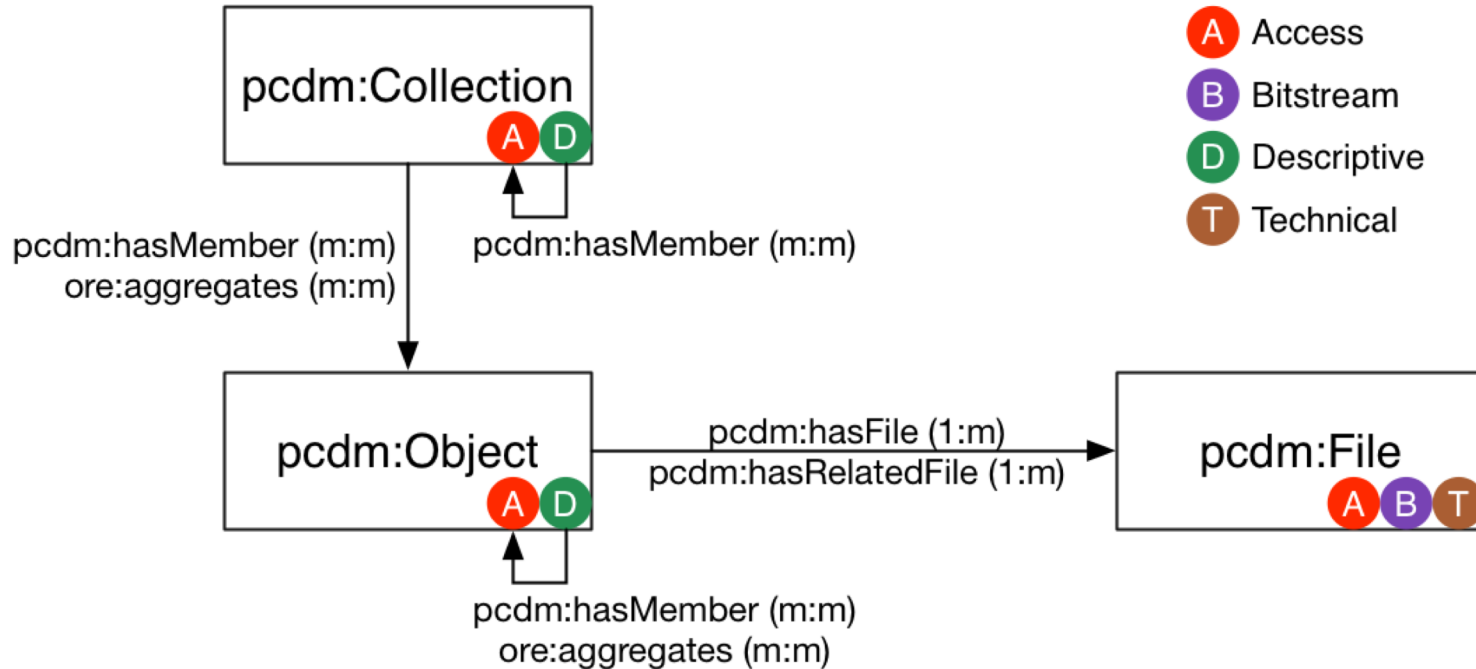
The Stack

Using the Portland Common Data Model (PCDM) gives the community a common data modelling tool-set. This, in turn, supports a much broader range of code reuse at higher levels.

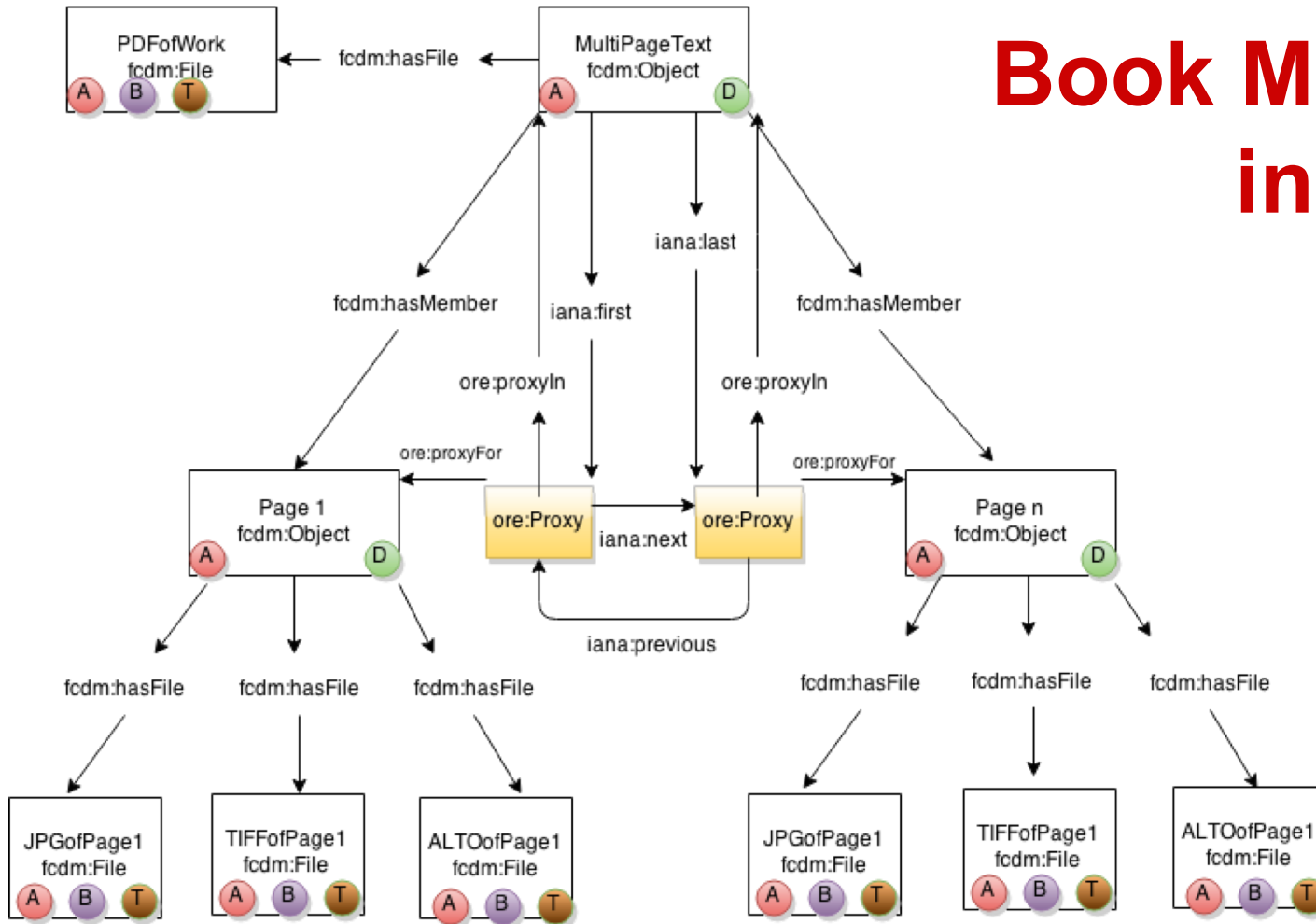
PCDM simplifies the creation of complex works and collections.



PCDM (interlude)



Book Modeled in PCDM

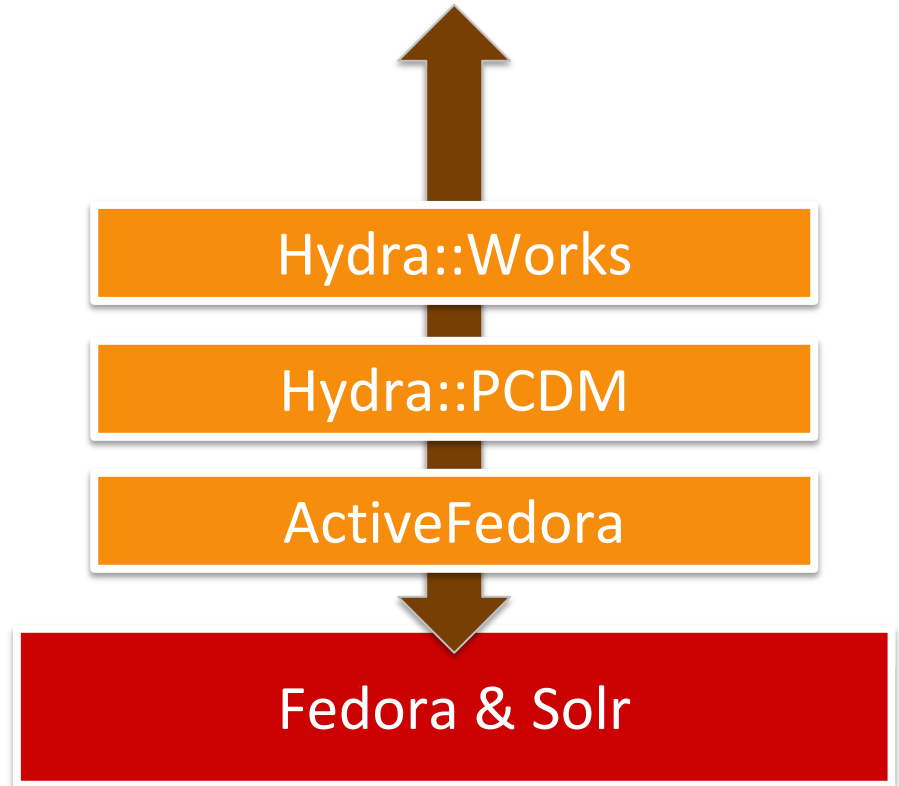


The Stack

The Works gem provides the utility drawer for managing digital content by defining common patterns for:

- File Characterization
- Derivative Generation
- Virus Scanning
- Text Extraction

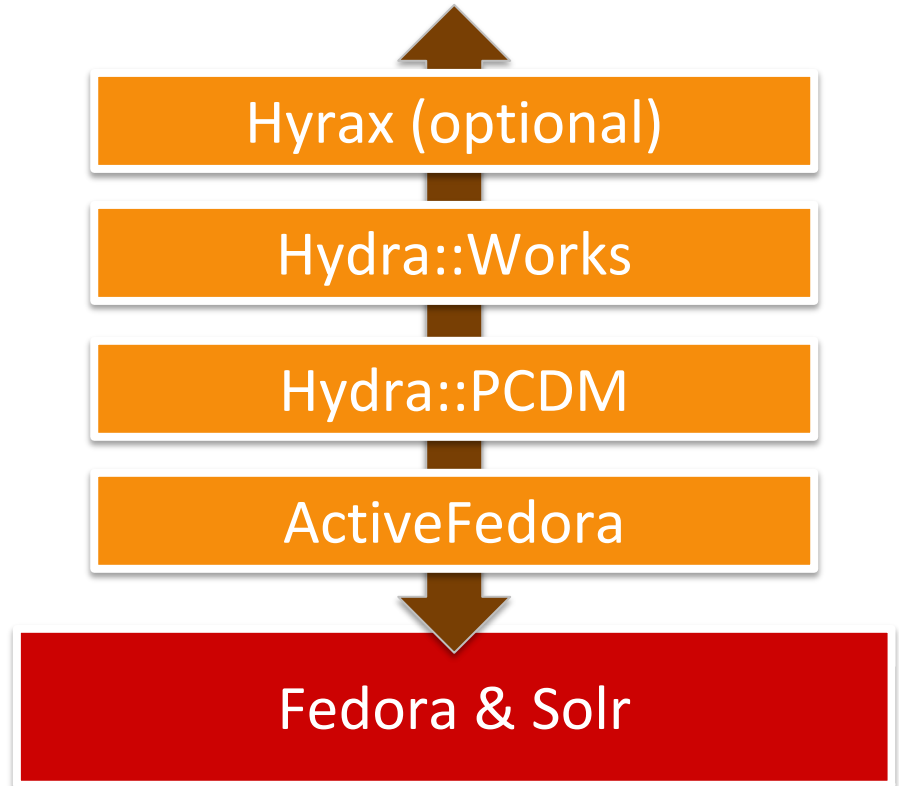
You can tailor any of these to local requirements, but get working solutions out of the box.



The Stack

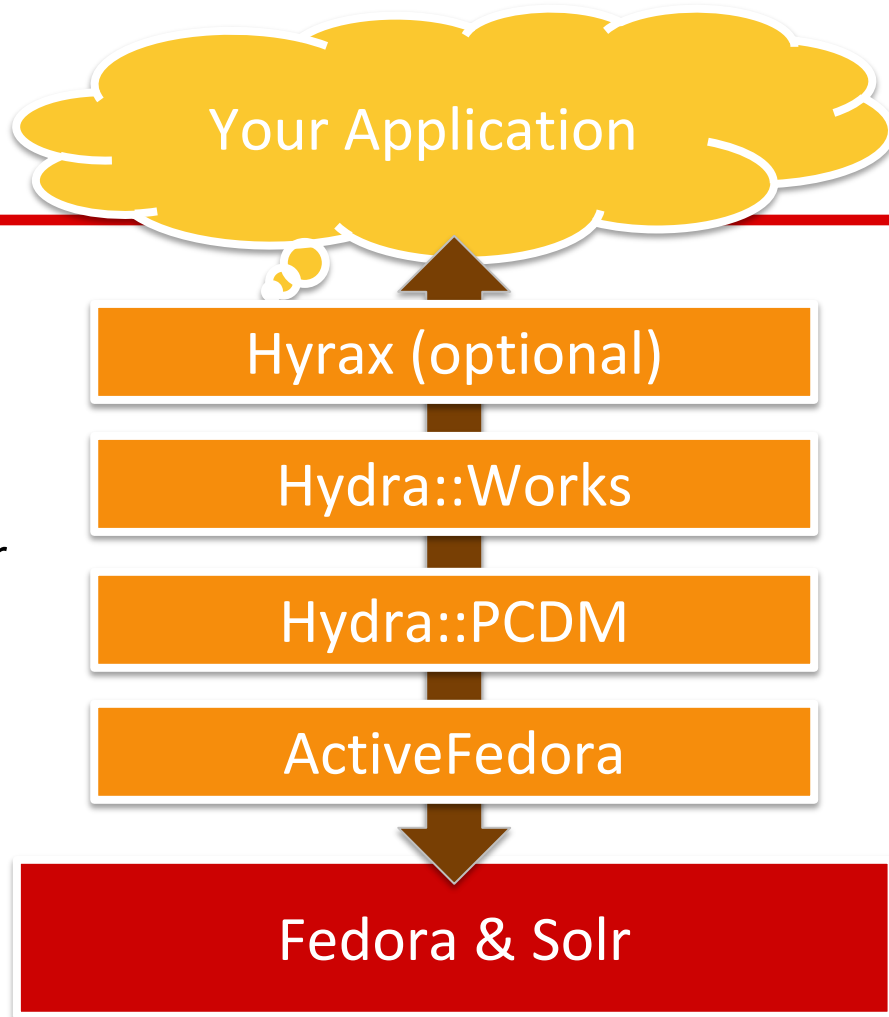
The Hyrax rails engine provides a fully featured repository front end with self-deposit and administrative workflow features along with:

- User & Admin dashboards
- Upload via cloud providers
- Citation integration
- File auditing & versioning
- Access controls
- Leases and embargoes



The Stack

Your own ruby-on-rails application can be a very lightweight wrapper around an existing application like Hyrax or Avalon, or a highly tailored in-house solution depending on your needs and your team's capabilities.



Next Up:

The Samvera Pitch

The Elevator Pitch? Who's Riding?

- UL/Dean or CIO? Resources and Risk - doing more with less while control the runway, maturity
- Content Specialists? Flexibility and Risk- reducing barriers to effective asset management & preservation
- Management - The Churn - Architecture, flexibility, evolution in a vibrant community, staffing impact
- Developers, PMs, DevOPs, -Connection, Time, Training, Visibility - It's a Family Affair and you are not alone

Community Sustainability...

- Core Purpose- We believe it is possible to provide and evolve solutions which reduce cost and risk in stewardship of digital assets for access and preservation.
- Partners - The whole is greater than the sum of its parts - partners join equally to share the load for mutual gain and greater good.
- Flexible Change with Steady Growth



Community Growth - Going Far Together



First Worldwide Hydra Connect @ UCSD, January 2014

2nd Worldwide Hydra Connect @ Cleveland, OH, September 2014

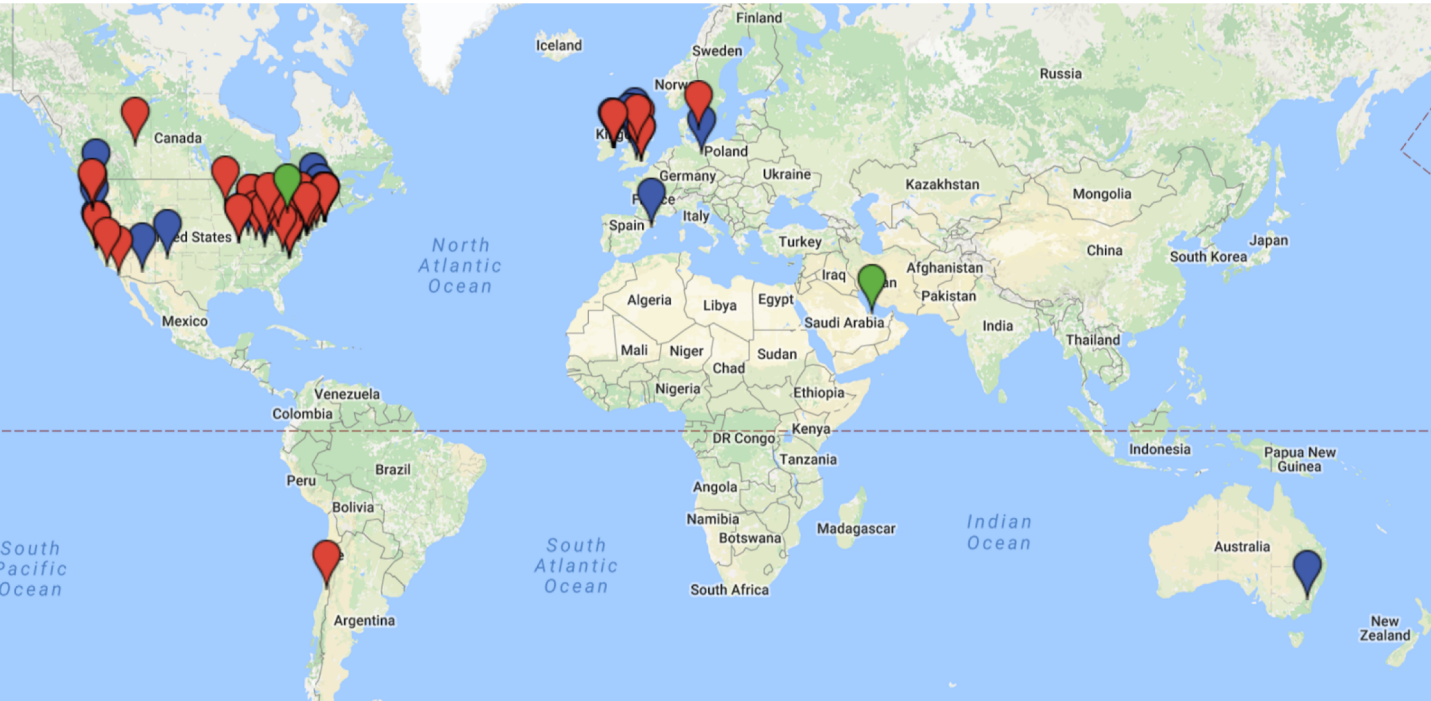
160 participants from 56 institutions

3rd Worldwide Hydra Connect @ Minneapolis, OH, September 2015

200 participants from 60 institutions



A Worldwide Presence



Samvera Partners

- DuraSpace
- Stanford University
- University of Hull
- University of Virginia
- University of Notre Dame
- Northwestern University
- Columbia University
- Penn State University
- Indiana University
- London School of Economics and Political Science
- Royal Library of Denmark
- Data Curation Experts
- WGBH
- Boston Public Library
- Duke University
- Yale University
- Virginia Tech
- University of Cincinnati
- Princeton University Library
- Cornell University
- Oregon State University
- University of Oregon
- Tufts University
- University of Alberta
- Digital Public Library of America (DPLA)
- University of Michigan
- University of California, San Diego
- University of York
- Lafayette College
- Washington University in St. Louis (WUSTL)
- The Digital Repository of Ireland
- University of California, Santa Barbara
- University of Houston Libraries
- Emory University
- CoSector, University of London





Resource Sustainability

- New Approach for Cross Training
- Reducing Turnover Impact
- Time - If you want to go fast ~~...go alone.~~
...use Samvera

Solution Sustainability...

- Protecting the partner & funders' investment - together build, leverage & evolve
- Organization & Governance
- Adopt best practices

Samvera Code Snapshot

1.48M Lines of code

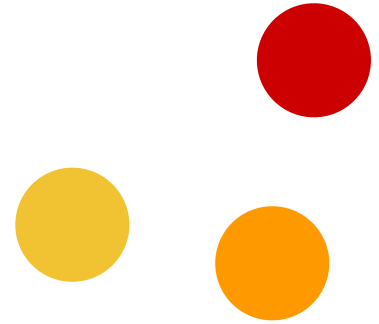
108 Current contributors

2 Days since the last commit

Activity Rating by Open Hub? **Very High**

Next Up:

What next?



What do I do back home?



Community Resources

- **Website** - for the managers

<https://samvera.org>

- **Wiki** - for the doers

[https://**samvera.org/wiki**](https://samvera.org/wiki)

- **Github** - for the makers

<https://github.com/samvera>

<https://github.com/samvera-labs>

<https://github.com/samvera-deprecated>



Try it out

- github.com/samvera
- demo.curationexperts.com
- ask other partners for demos!

Stanford University
Univ. of Notre Dame
Indiana University
Data Curation Experts
Yale University
Cornell University
Tufts University
University of Michigan
University of York

University of Hull
Northwestern University
London School of
Economics
WGBH Boston
Virginia Tech
University of Oregon
Lafayette College

UC San Diego
Digital Repository of
Ireland
University of Virginia
Columbia University
Boston Public Library
University of Cincinnati
Oregon State University
University of Alberta

UC Santa Barbara
DuraSpace
Penn State University
Duke University
Princeton University
Library
Washington U. in St Louis
University of Houston
DPLA





samvera

J. Willard Marriott Library

THE UNIVERSITY OF UTAH



October 9-12

Questions

samvera-community@googlegroups.com
samvera-tech@googlegroups.com

karen_cariani@wgbh.org
robin.ruggaber@virginia.edu
mark@curationexperts.com

Thank you

