

---

# 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



# from proposal submission

---

Hour 1: Introductions, overview of Samvera products and the community – what is it, why is it different, ask audience what they are looking for, show website and wiki, and discuss the how the community at large works.

Hour 2: The second hour will give a general technical overview designed for non-coders. Showcase some applications that exist. The staff skills needed to maintain and contribute a Samvera solution will be discussed, resources that exist to get started, and how to contribute code to the community. *[Give a 10 minute break halfway through]*

Hour 3: *Personas (UL/Dean, Executive Tech Lead, Middle Managers, Software Engineers); Benefits of being part of a community, how is this and advantage; positive and negative risk assessment; how does it differ from other solutions; call to action; finish with an invitation to Virtual and physical connect; how to pitch Samvera and get institutional support. It will discuss the advantages of being part of the community and how that strengthens the sustainability of the tools, the applications, and the community overall.*

Mark says: check out the “What is Samvera” presentation on the right side of this page  
<https://wiki.duraspace.org/display/samvera/Events%2C+presentations+and+articles>



# 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 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.*



# 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

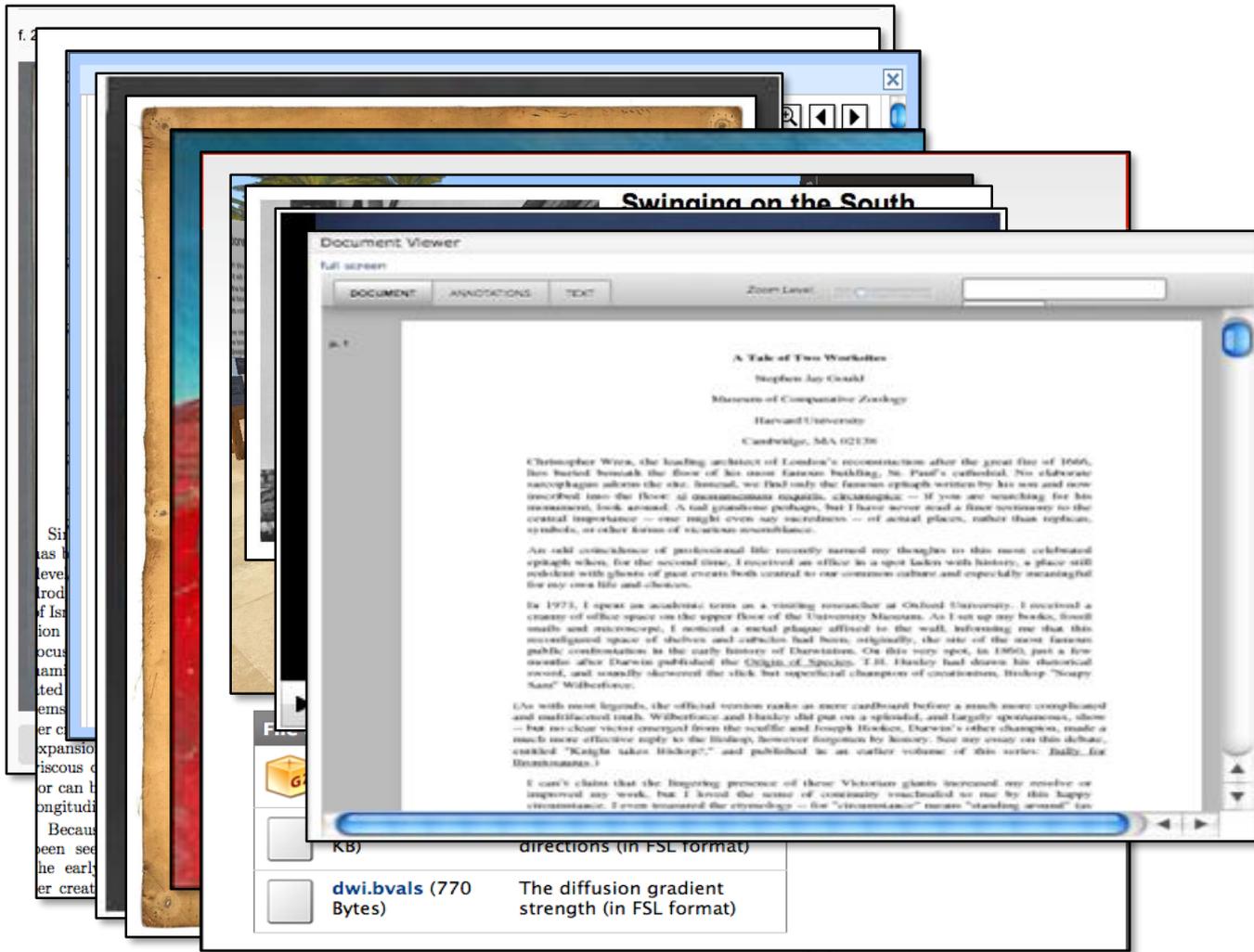


# Fundamental Assumption #1

---

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

## Swinging on the South

Document Viewer

Full screen

DOCUMENT

ANNOTATIONS

TEXT

### A Tale of Two Worksheds

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, has 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 at monumentous expense, circumscribing — if you are searching for his monument, look around. A tall granite one perhaps, but I have never read a finer testimony to the central importance — one might even say sacredness — 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 marked 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 chance of office space on the upper floor of the Teynham Museum. As I set up my books, fossil snails and microscopes, I noticed a metal plaque affixed to the wall, informing me that this neurological 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 1860, just a few months after Darwin published the *Origin of Species*, T.H. Huxley had drawn his rhetorical sword, and soundly skewered the stilted but superficial champion of creationism, Bishop "Sam" Wilberforce:

[As with most legends, the official version reads as more cardboard before a much more complicated and multifaceted truth. Wilberforce and Huxley did put on a splendid, and largely spontaneous, show — but no clear victor emerged from the scuffle and Joseph Hooker, Darwin's other champion, made a much more effective reply to the Bishop, however forgotten by history. See my essay on this debate, entitled "Keight takes Hooker?" and published in an earlier volume of this series. *Daily for Darwinians*.]

I can't claim that the lingering presence of these Victorian ghosts increased my resolve or inspired my work, but I found the sense of continuity meaningful to me by this happy circumstance. I even learned the etymology — for "circumstance" means "standing around" in

KB)

directions (in FSL format)

**dwi.bvals (770 Bytes)**

The diffusion gradient strength (in FSL format)

# Point Solution Approach - Welcome to Siloville



**ETD**

ETDs  
(Theses)



**IR**

Books,  
Articles



**Image  
DB**

Images



**DAM**

Audio-  
Visual



**?**

Research  
Data



**Geo-  
spatial  
Inf.**

Maps  
& GIS



**Records  
Mgmt.**

Documents

Management

Access

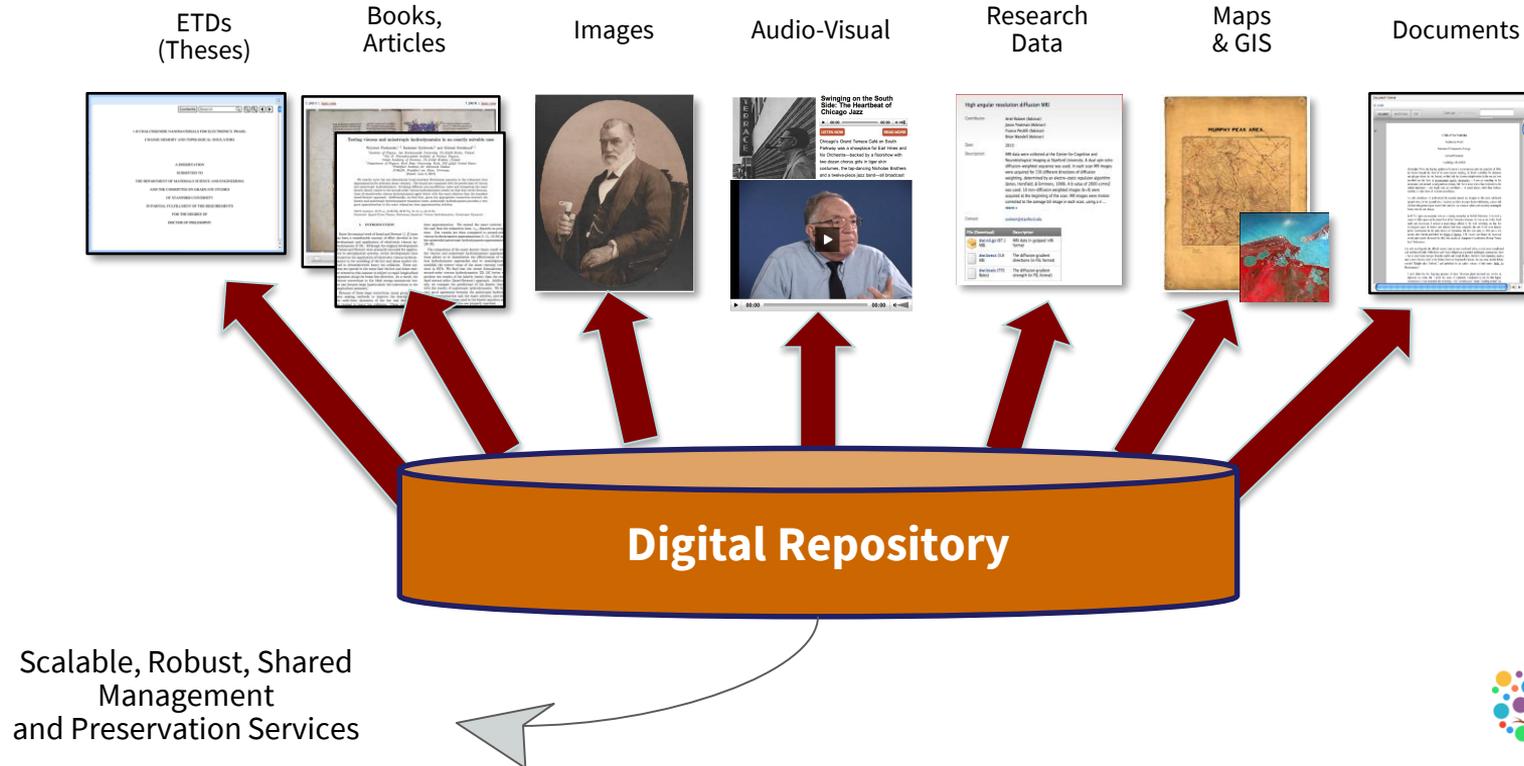
Preservation?

Effective?

Sustainable?



# Repository Powered Approach



# CRUD in Repositories

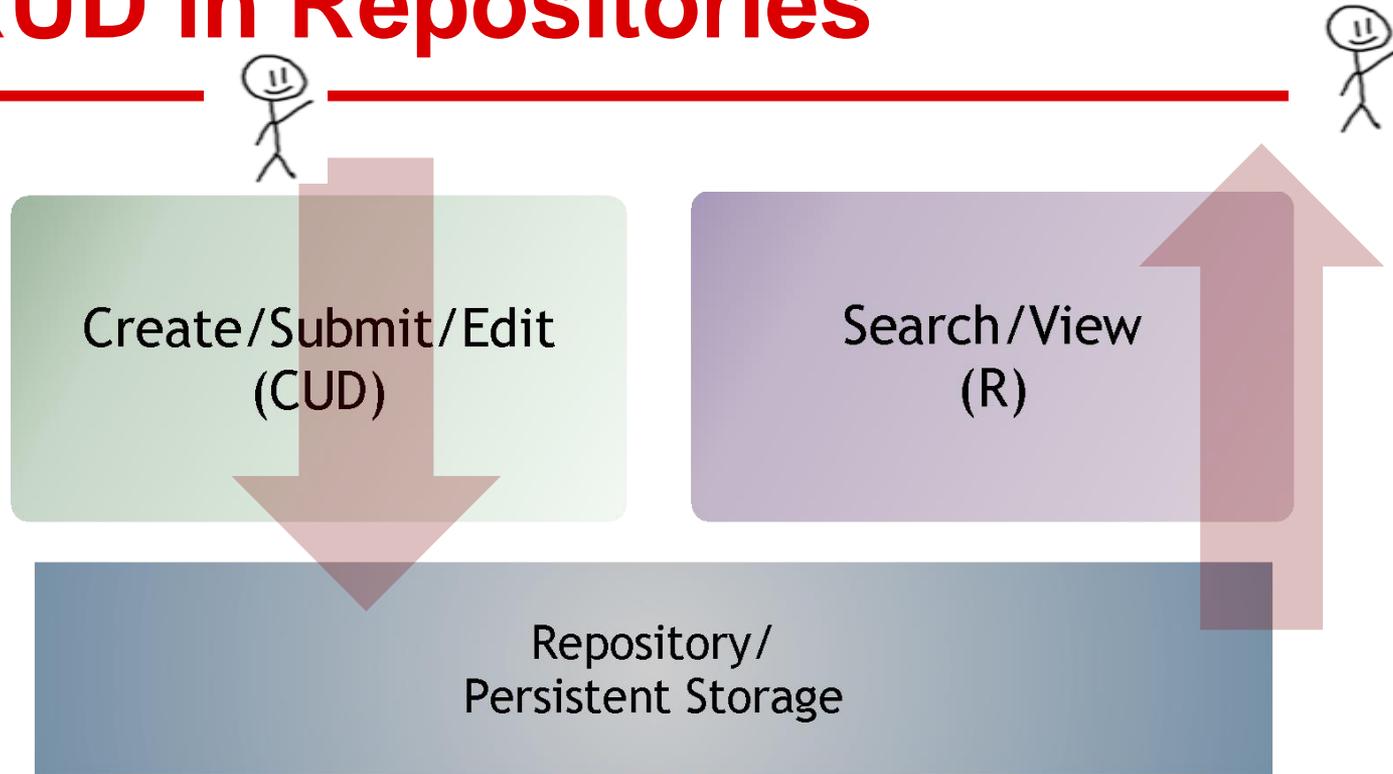
---

Create/Submit/Edit  
(CUD)

Search/View  
(R)

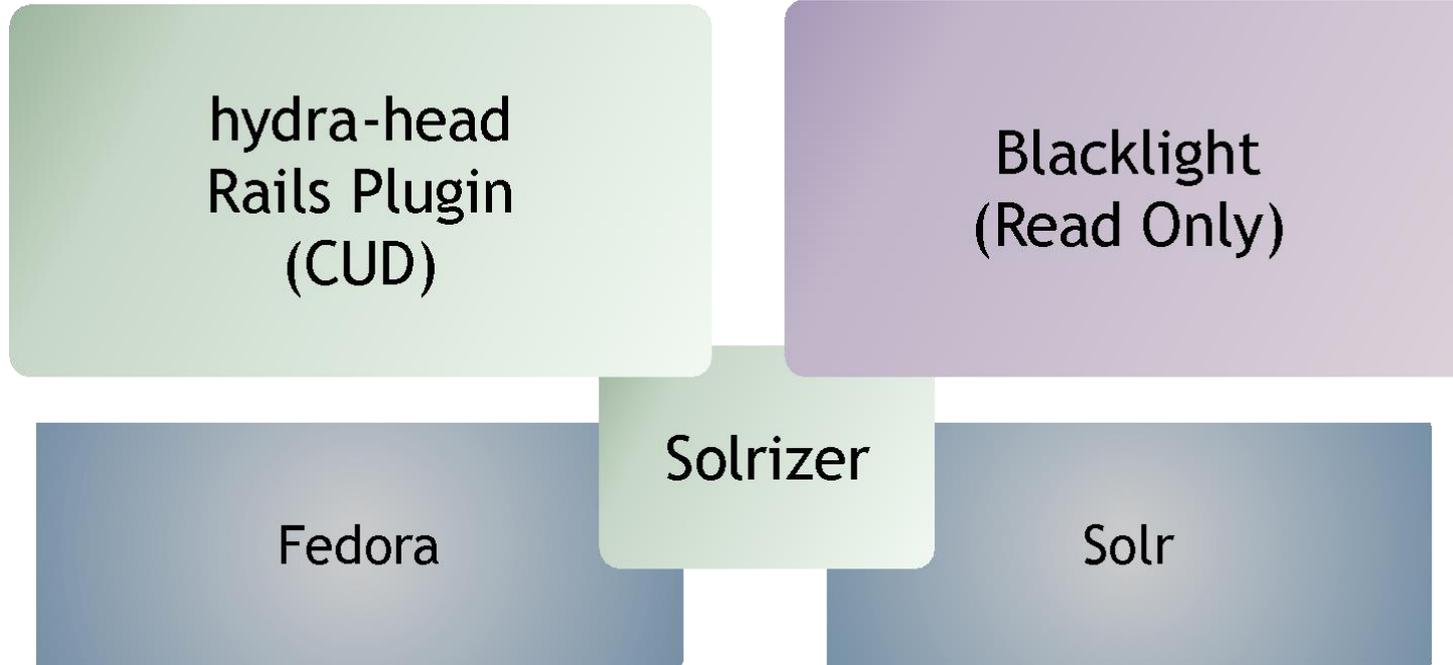
Repository/  
Persistent Storage

# CRUD in Repositories



# Major Samvera Components

---



# 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:

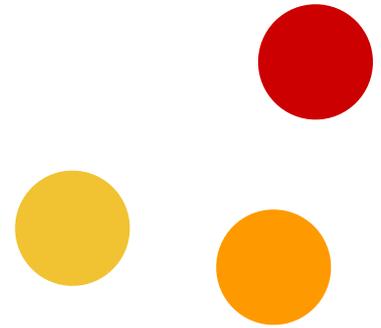
- Self-deposit system
- Digital Collections System
  - Sheet Music
  - Architectural resources
- Electronic Theses & Dissertations
- Digital Image System
- Media Management
- Media Preservation System
- Research Data Management
- Digitization Workflow System
- Digital Preservation System
- Digital Archives 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>



# Exercises

---

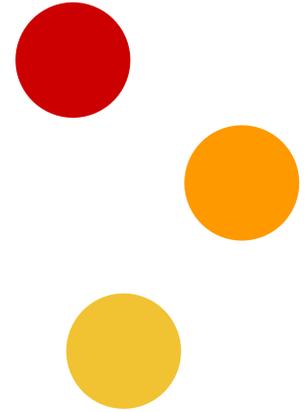
- Search for a topic of interest to you from the samvera-tech or samvera-community Google Groups
- Find a Samvera Interest or Working Group of interest to you; find the most recent meeting notes
- Using the Samvera Wiki...  
<https://wiki.duraspace.org/display/samvera/Samvera>
  - Sign up for the samvera-tech or samvera-community Google Groups
  - Get a Samvera wiki account
  - Join the project-samvera Slack team



---

**First Up:**

# **Example Sites**



---

What are folks doing with it?



# Tufts Digital Library

---

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

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



# 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

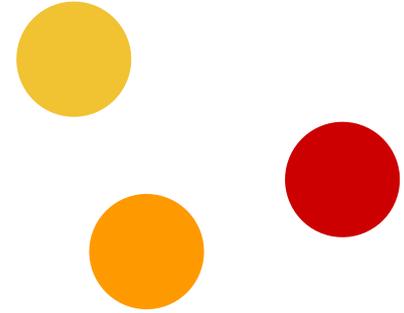
Live Site: <https://earthworks.stanford.edu>



---

**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:**

# **The Samvera Pitch**

---

---

# Repository Elevator Pitch

---

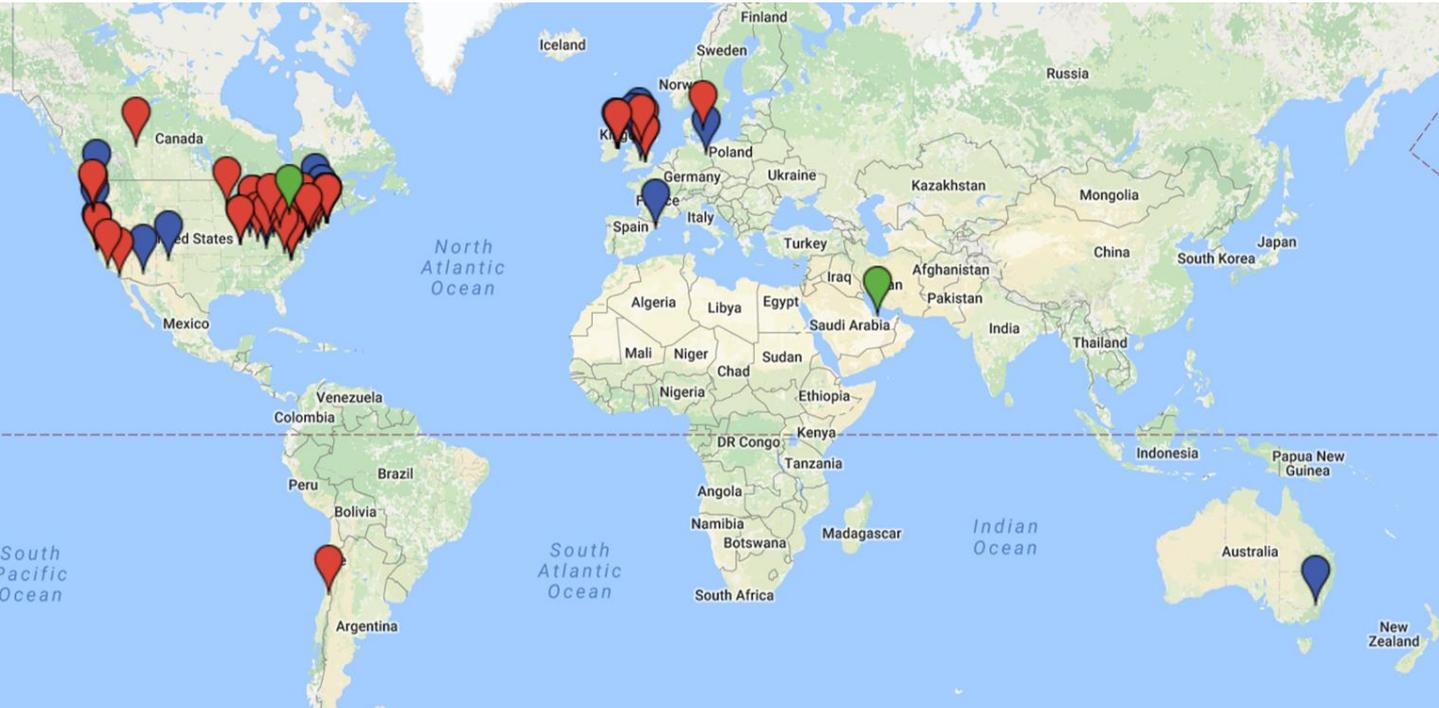
- A (new) way to manage digital content
- The best of Hydra in an easy-to-install, easy-to-configure, feature-rich solution
- Reducing barriers to effective asset management
- & preservation
- For collections and content types of many kinds
- Supports next-generation web technologies
- Is supported by a vibrant and ever-growing open source community.



If You Want To Go Fast ~~...go alone.~~  
*...use*  
*Samvera*



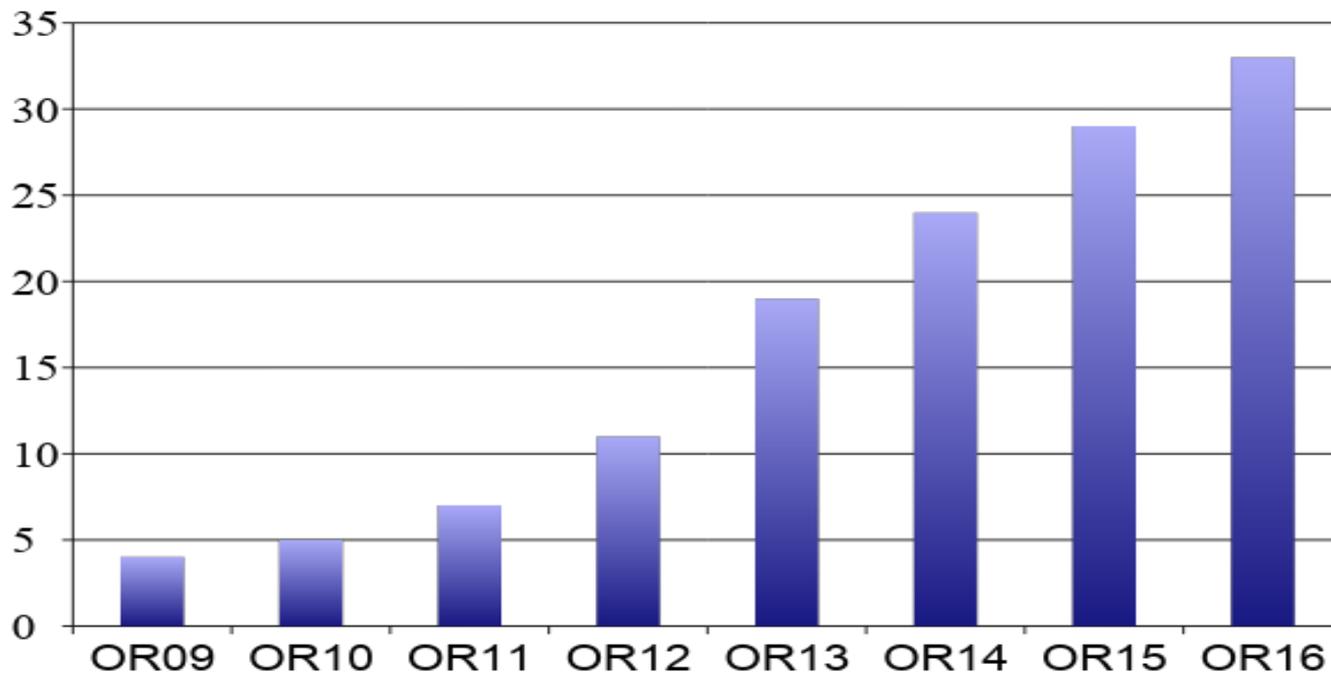
# A Worldwide Presence



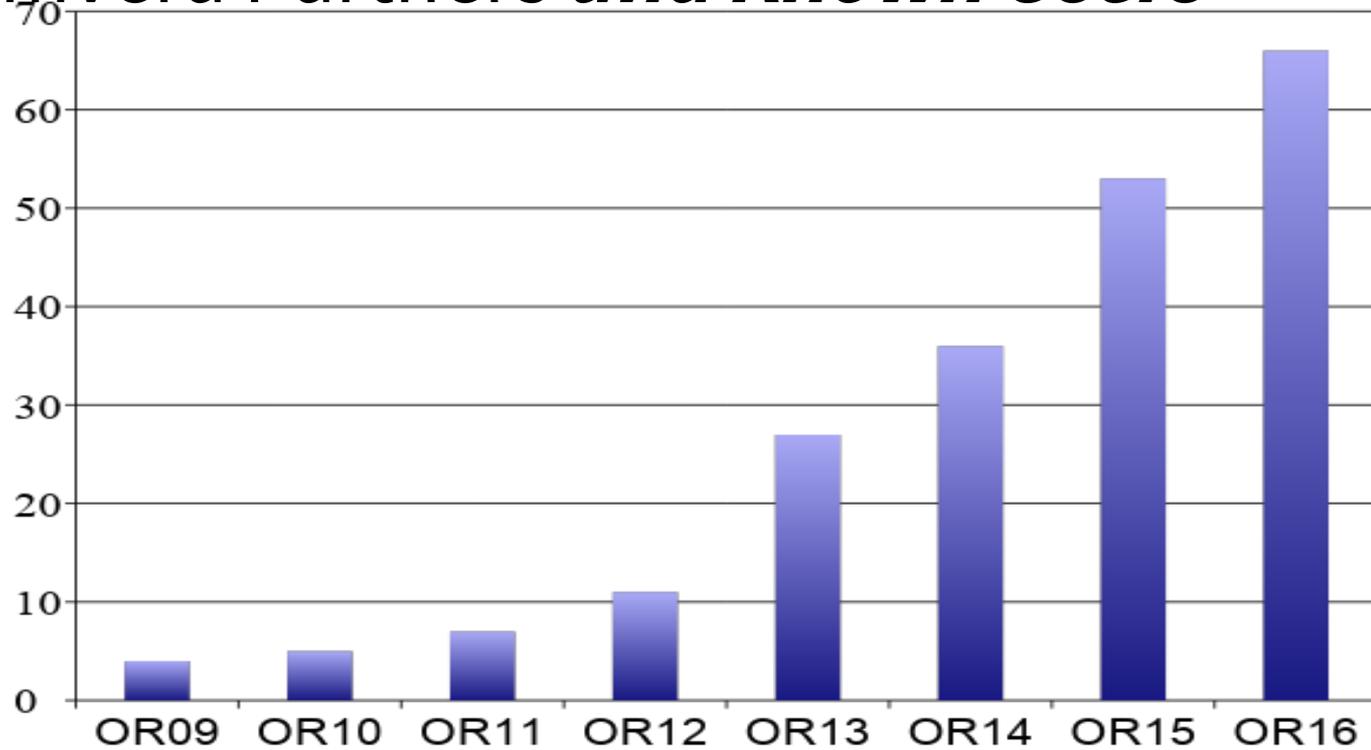
# By the Numbers....



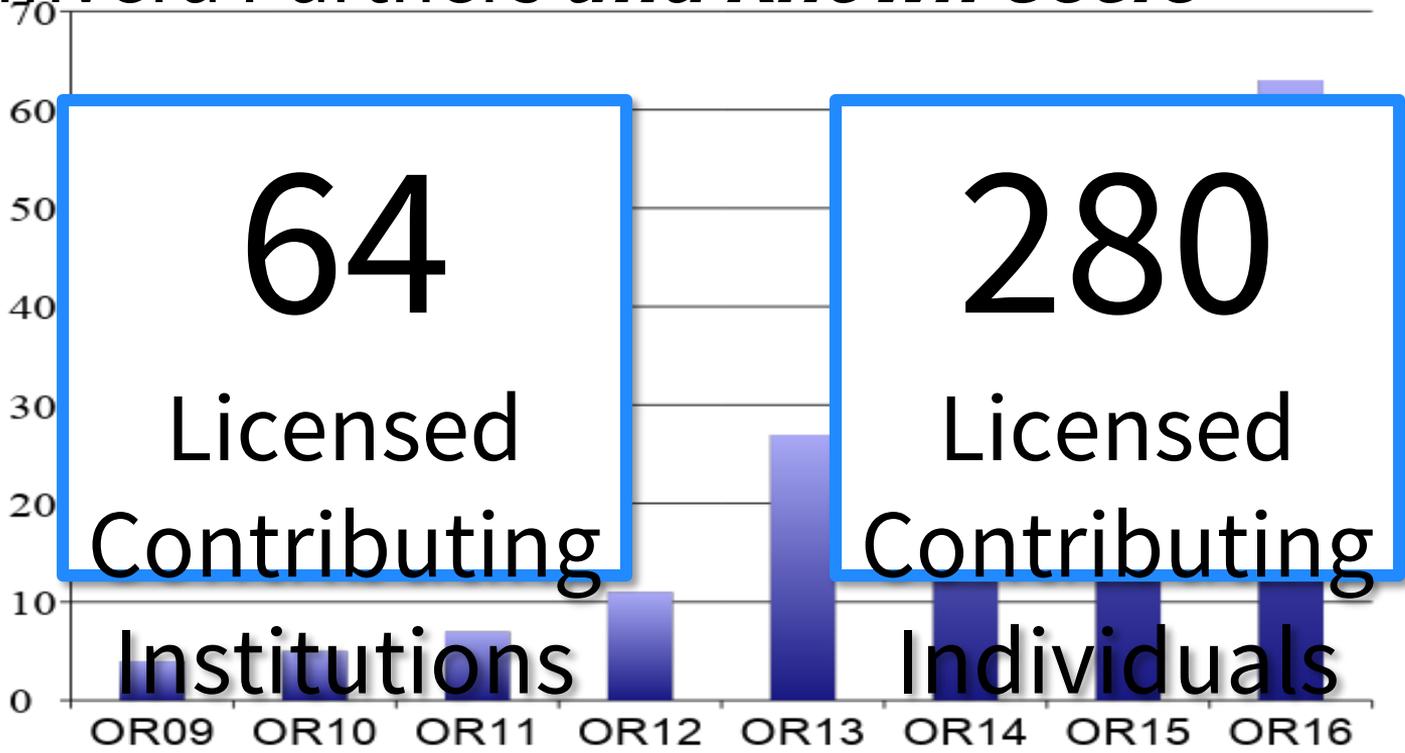
# Samvera Partners



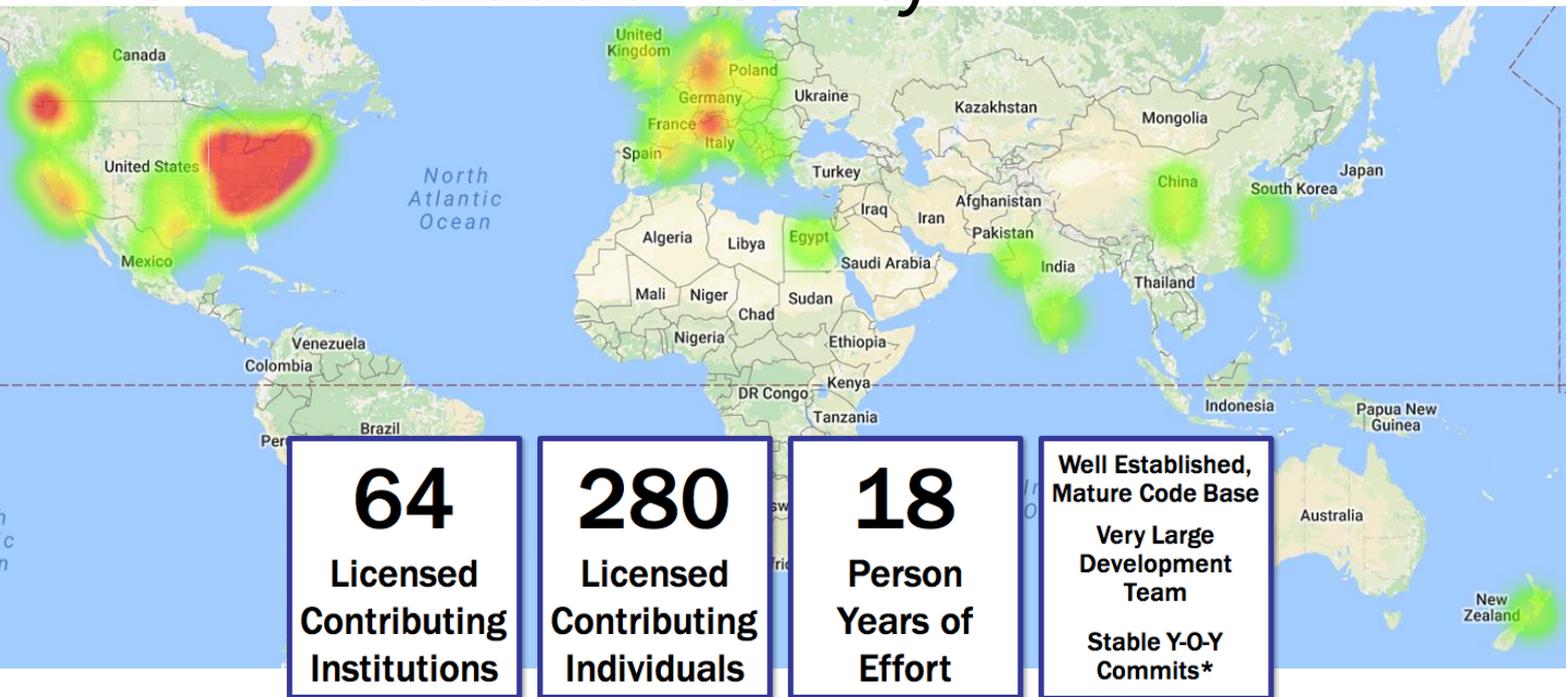
# Samvera Partners *and Known Users*



# Samvera Partners *and Known Users*



# Samvera Code Activity



samvera

132

samvera-

748

samvera-

617

samvera-



478 Slack Team  
Members

[project-hydra.slack.com](https://project-hydra.slack.com)



@HydraSphere  
623 Followers



# Hydra Partners

## 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 Michigan
- 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



# Going Far Together



First Worldwide Hydra Connect @ UCSD, January 2014

2<sup>nd</sup> Worldwide Hydra Connect @ Cleveland, OH, September 2014

*160 participants from 56 institutions*

3<sup>rd</sup> Worldwide Hydra Connect @ Minneapolis, OH, September 2015

*200 participants from 60 institutions*





# Sustainability

*grants*

*abused*

**No animals were harmed  
in the making of this film.**

*project / community*

# Sustainability...

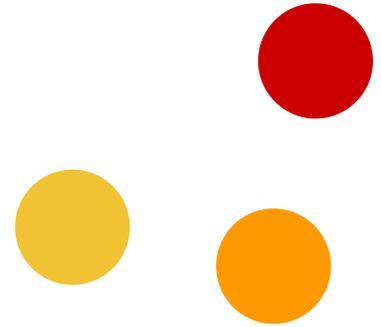
*grants* *abused*  
No animals were harmed  
in the making of this film.  
*project / community*

- For institutions  
*Partners help shoulder the load*
- For other projects  
*Post-project maintenance & governance*
- For leveraging funders' investments

---

**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](https://github.com/samvera)
- [demo.curationexperts.com](https://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  
Duoc UC (Chile)  
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  
University of York

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



# Questions

---

[samvera-community@googlegroups.com](mailto:samvera-community@googlegroups.com)  
[samvera-tech@googlegroups.com](mailto:samvera-tech@googlegroups.com)

[karen\\_cariani@wgbh.org](mailto:karen_cariani@wgbh.org)

[robin.ruggaber@virginia.edu](mailto:robin.ruggaber@virginia.edu)

*Last resort...* [mark@curationexperts.com](mailto:mark@curationexperts.com)

Thank you

---



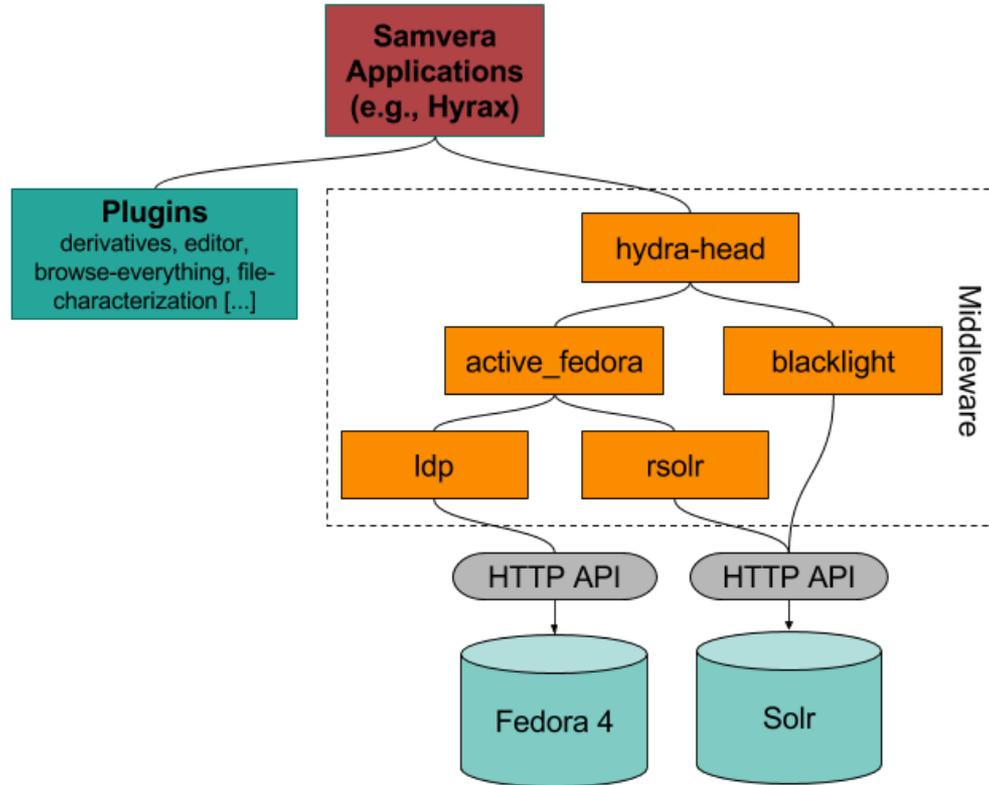
Slides below from last years OR related to tech and may be of use as others were.

# Agenda

- What is Samvera?
- Community Framework
- *Technical Framework*
- Data Modeling
- Getting Started with Samvera
- Discussion



# 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 releases
  - Vibrant community & funding under DuraSpace
- Fedora 4.0 released in Dec '14; co-evolving with Hydra



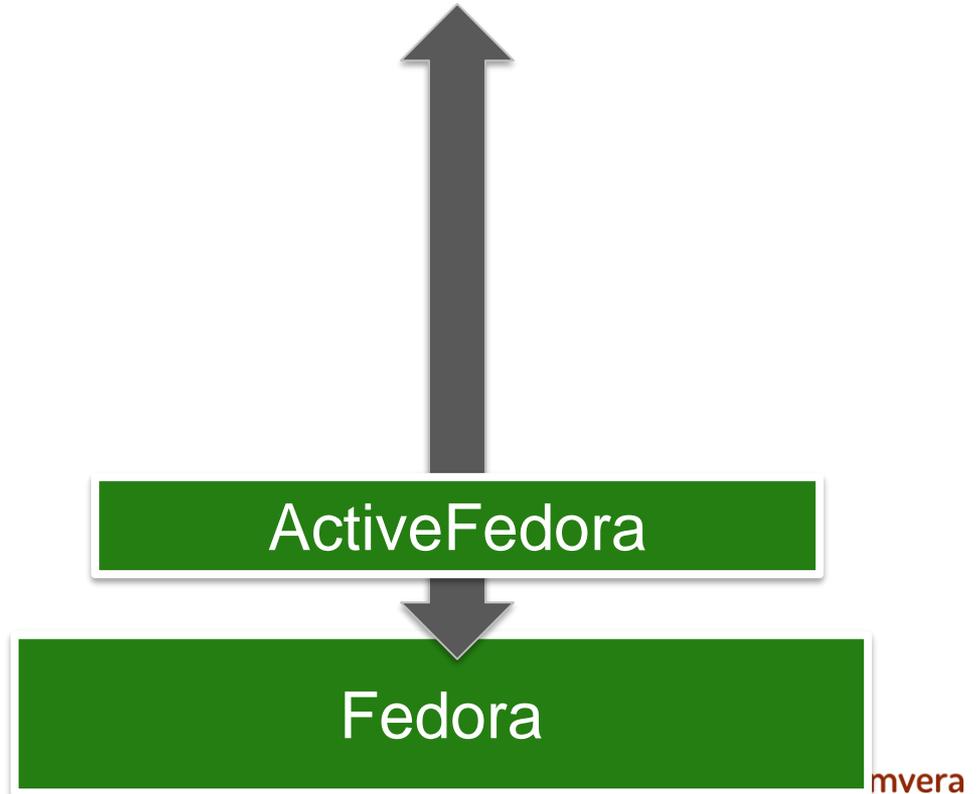
# A Note on Ruby on Rails

- *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)



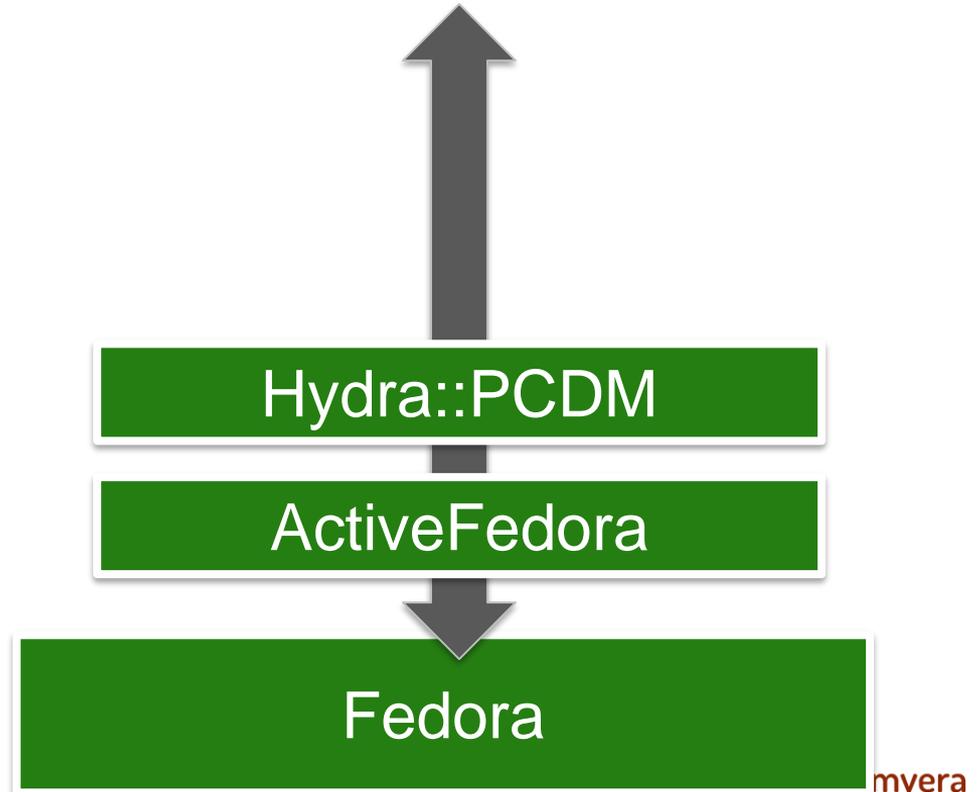
# Hydra's Core Gems

Interact with Fedora with a Rails/ActiveRecord pattern



# Hydra's Core Gems

Build digital objects using  
PCDM-based models

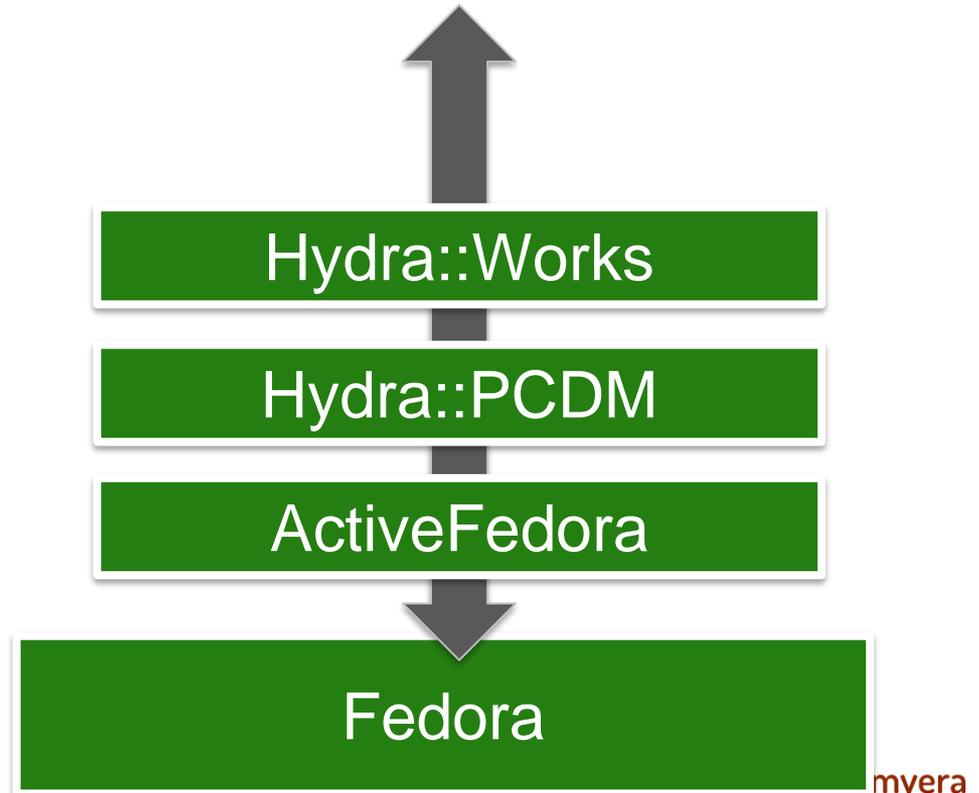


# Hydra's Core Gems

Work with files, like:

- Characterize
- Make derivatives
- Virus check
- Extract text

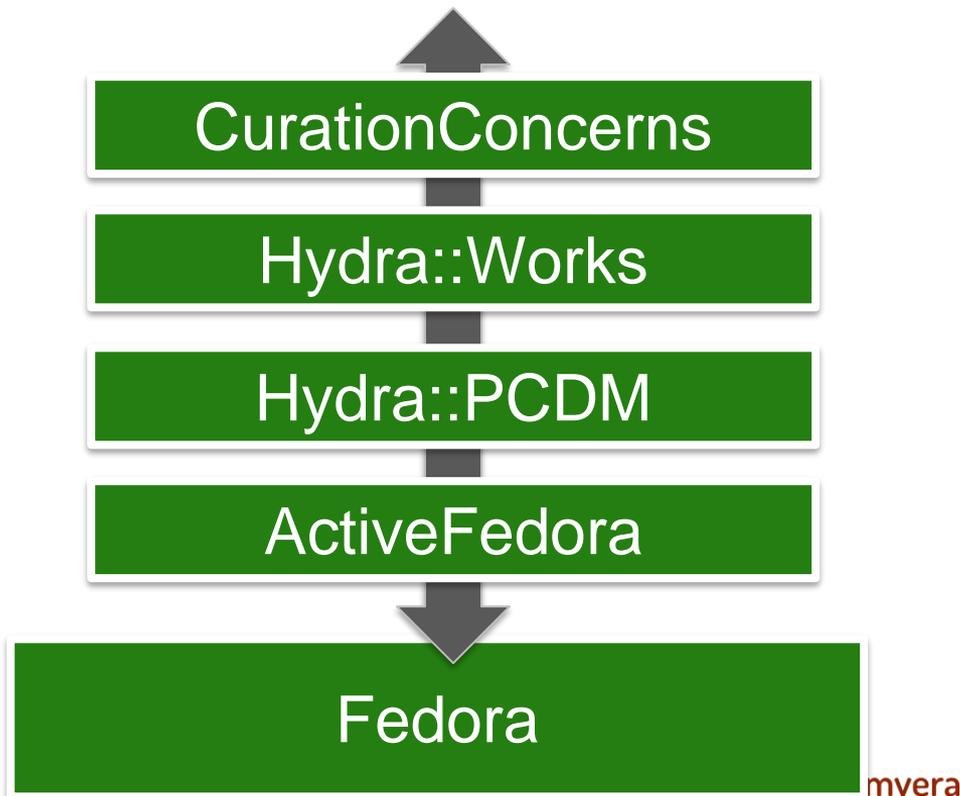
Group them to create  
“works” (or complex  
objects)



# Hydra's Core Gems

Get basic Ruby on Rails views, and more shared features:

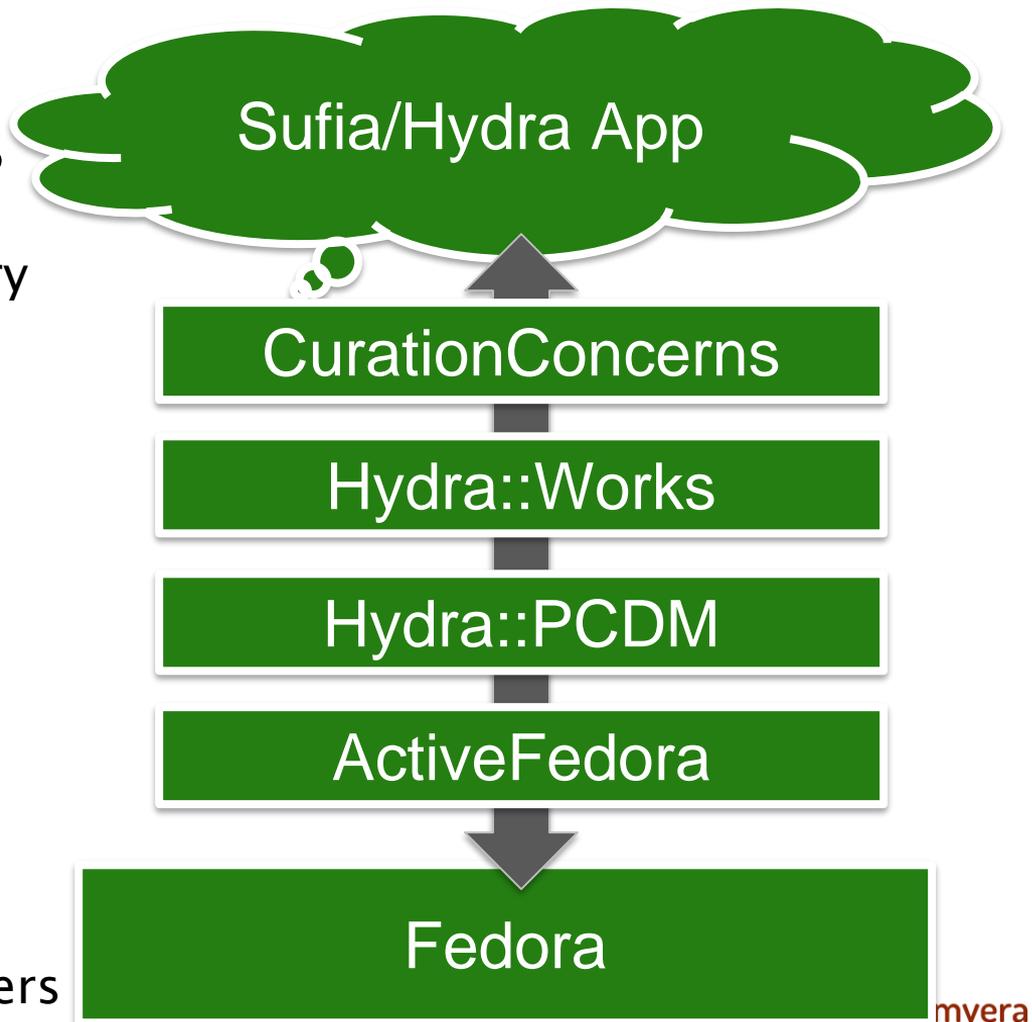
- Custom models
- Routing for CRUD
- File auditing
- Versioning
- Single-use links
- Upload file sets
- Access controls
- Leases and embargoes



# Hydra's Core Gems

A fully functional repository application, with:

- Proxy deposit
- User dashboard
- User profile
- Featured works & researchers
- Contact form
- WYSIWYG UI editing
- Activity streams
- Upload via cloud providers
- Integration with Zotero



# Hydra's Core Gems

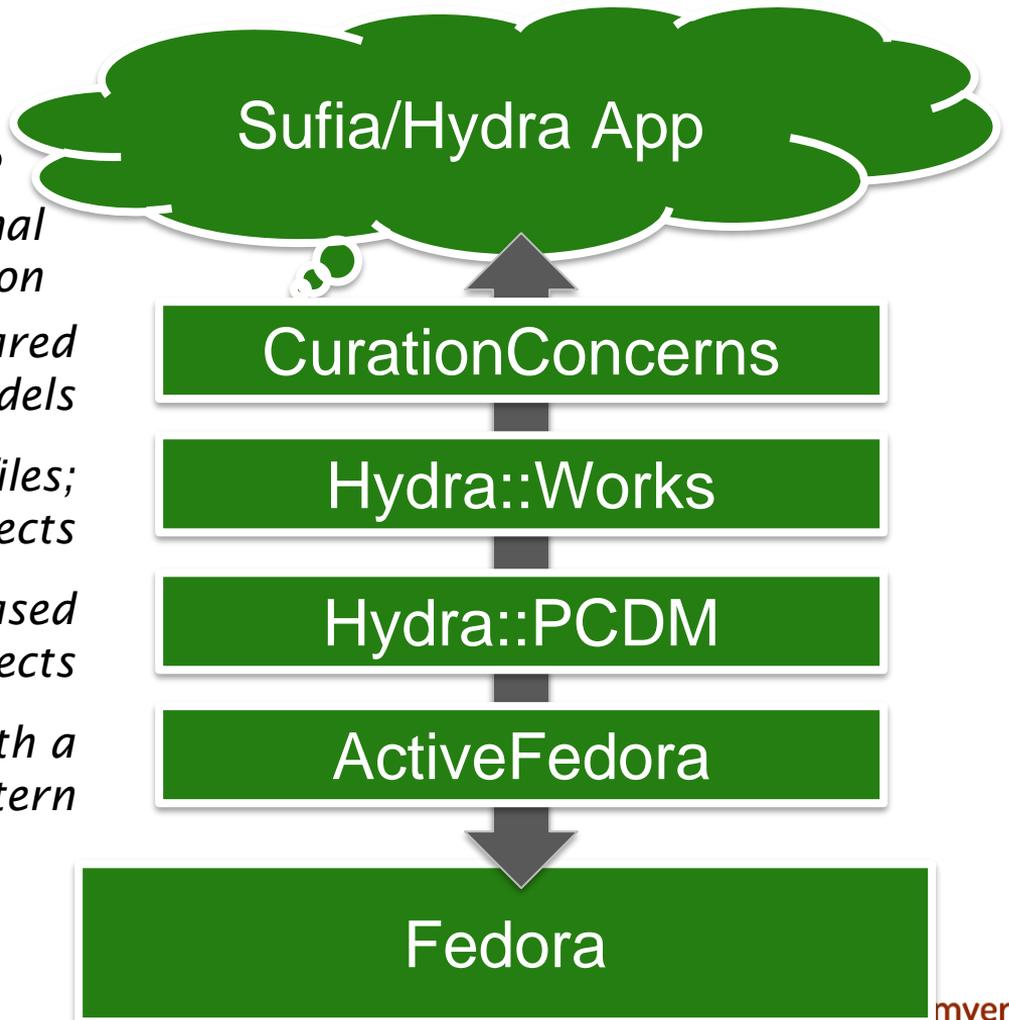
*A fully functional  
repository application*

*Build basic Rails views & shared  
features: custom models*

*Work with files;  
group them into digital objects*

*Build PCDM-based  
Models for Digital Objects*

*Interact with Fedora with a  
Rails/ActiveRecord pattern*





# Hyrax

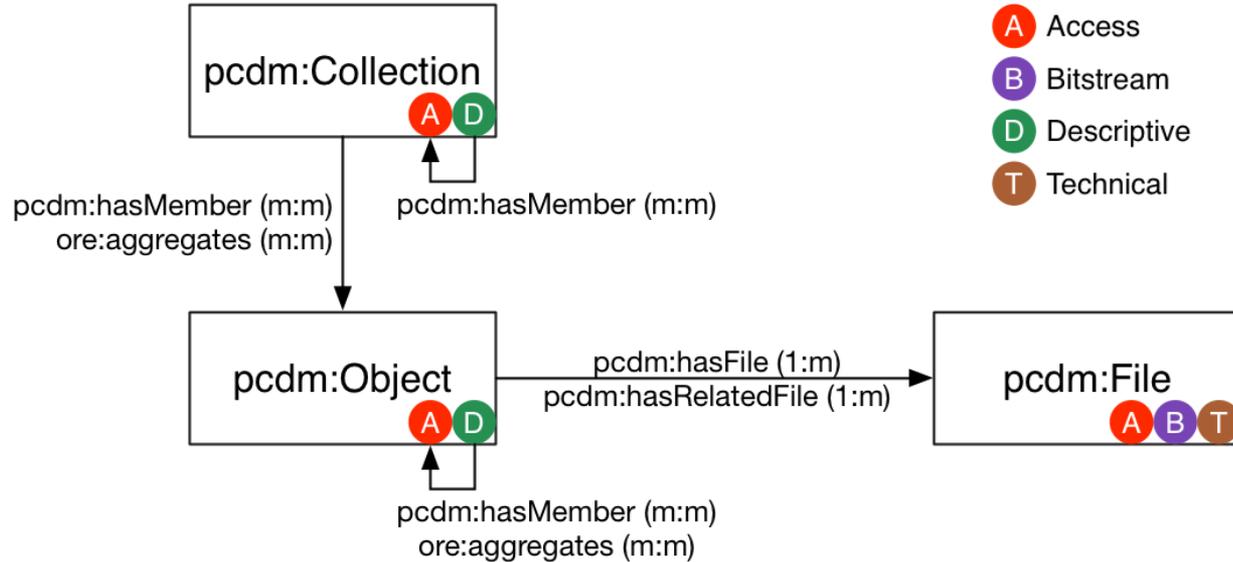
- Hyrax (formerly known as Sufia 7.x)
- Lots of features already done
- Complex objects based on PCDM in Fedora 4
- Building on Sufia, one of most actively maintained and widely used Hydra gems

# Agenda

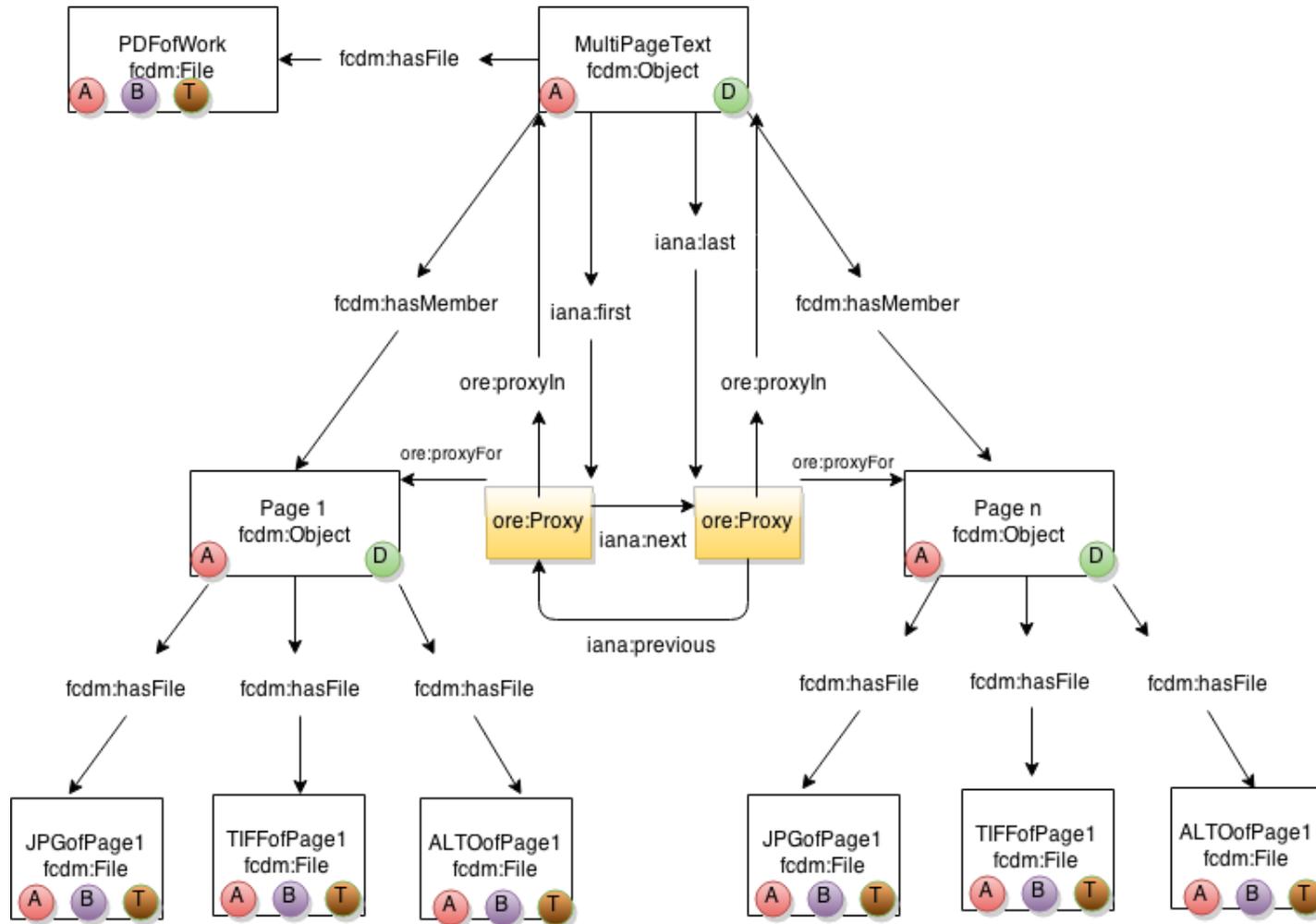
- What is Samvera?
- Community Framework
- Technical Framework
- *Data Modeling*
- Getting Started with Samvera
- Discussion



# PCDM = Portland Common Data Model



# Book



# Why PCDM?

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

# Agenda

- What is Samvera?
- Community Framework
- Technical Framework
- Data Modeling
- *Getting Started with Samvera*
- Discussion



# 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/hydra-head/wiki>



# Technical Engagement

- Software: <https://github.com/projecthydra>
  - Developers should start here:  
<https://github.com/projecthydra/hydra-head/wiki>
- Note: Most sites start from Sufia on Fedora 4 at this point
- [samvera-tech@googlegroups.com](mailto:samvera-tech@googlegroups.com), Slack and weekly tech phone calls = best channels for tech support

---

• Ask questions Seek help Commit back

# Typical Staffing

- Repository Manager
  - Typically serves as product owner & service manager. Essential to defining service goals, running service and working with users after set up
- Developers: numbers ranges from
  - 0: borrowed or contracted
  - 1: many sites have a single dev
  - 2-4: most Hydra teams don't get above 4 developers, usually on sprints rather than dedicated full time

---

- Optional:

# Skills & Training

- Technical Skills
    - Agile & social coding: Git, pull requests & code review, continuous integration are standard.
    - Ruby on Rails: eminently learnable. Rails Bridge is a good starting point if developers are new to Rails.
    - DevOps & Sys Admins: growing coterie of Hydra users leveraging each others' skills for deployment and operations
  - Hydra Camps
    - Typically organize a camp 2-3 times a year
    - Many sites host one as a way to bootstrap rapid
-

# A word of (polite) warning

- Hydra is open source – anyone can use it freely
- Working with Hydra is not, though, free
  - As in ‘puppies are not free’
  - It requires investment
    - Towards achieving a goal
    - Towards sustaining a solution
- Hydra recognises this and is seeking to provide a way to invest in a mutual direction of travel –  
come with us!



# Agenda

- What is Samvera?
- Community Framework
- Technical Framework
- Data Modeling
- Getting Started with Samvera
- *Discussion*



