An introduction to Samvera : Workshop

8th October 2018
Samvera Connect, University of Utah

Steve Van Tuyl,
Julie Allinson,
Robin Ruggaber

samvera

Let's Start with You

- Introductions
- What are your goals for this workshop?
- What interested you in Samvera?
- What are some things you would like to know more about?

What we'll cover

- What is Samvera?
- How We Work
- Example Sites
- Hands-On
- Technical Details
- The Samvera Pitch
- What Next?
- Q & A

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

- Collaboratively built "solution bundles" that can be adapted and modified to suit local needs.
  - ⇨ e.g. Hyrax, Avalon, etc.

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

samvera

Why Samvera?

- Need a flexible, rich, robust repository solution

- Be part of a community
  - Shared resources, shared risk, shared advances

- The whole (we) is greater than its parts (each of us)
  - Especially in this crowd

- Control your own destiny

- Build institutional capacity

- The way work is getting done now

samvera

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.

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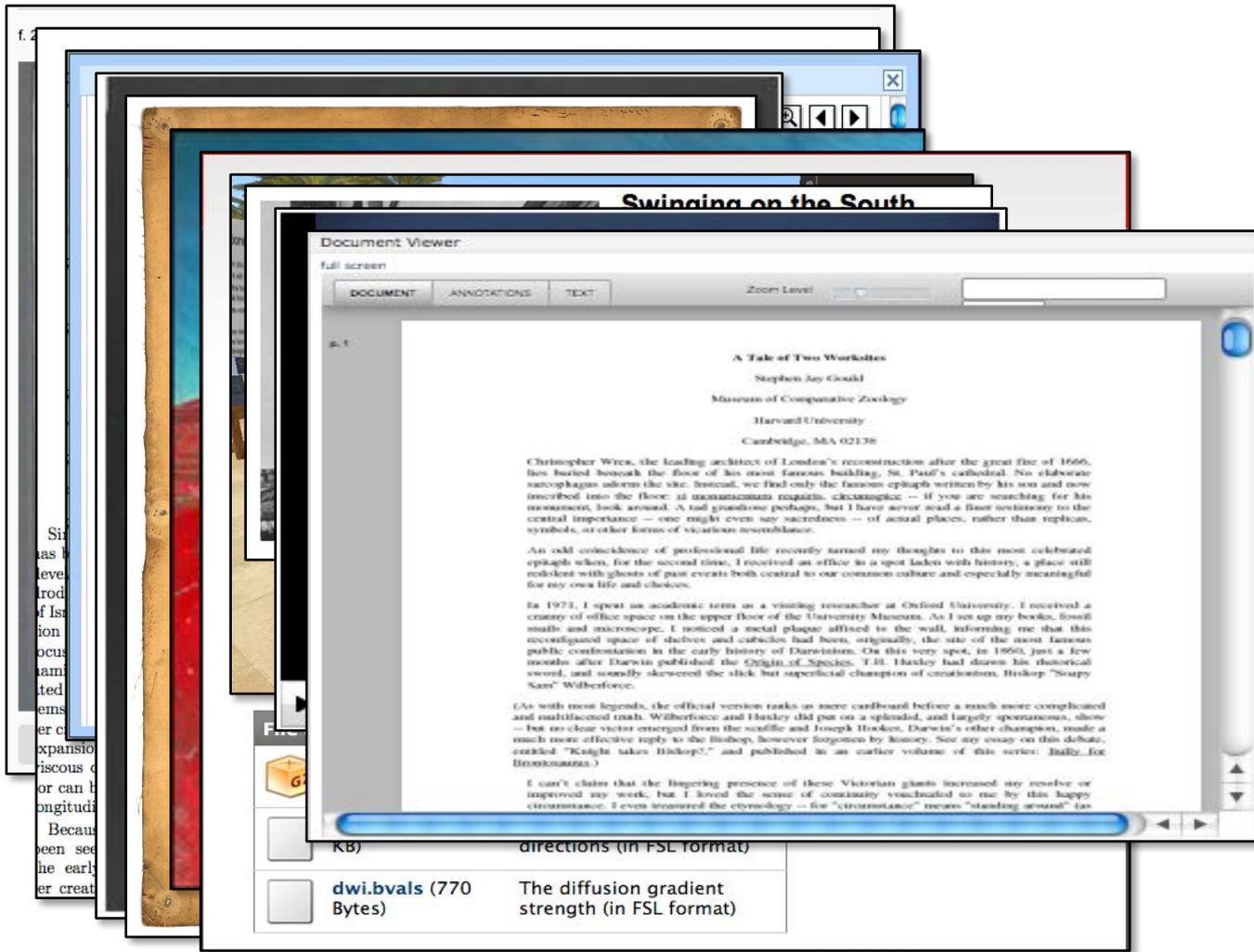
Books
Articles
Theses
Images
Maps
Data (Raster)
Data (Comp.)
Data (Observ.)
Audio
Video
Documents

Document Viewer

full screen

| DOCUMENT | ANNOTATIONS | TEXT |

Zoom Level

A Tale of Two Worksites

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: si monumentum requiris, circumspice -- if you are searching for his monument, look around. A tad grandiose 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 redolent with ghosts of past events both central to our common culture and especially meaningful for my own life and choices.

In 1971, 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 1860, 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 champion of creationism, Bishop "Soapy Sam" Wilberforce.

(As with most legends, the official version ranks as mere 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 any essay on this debate, entitled "Knight takes Bishop?," and published in an earlier volume of this series: Bully 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 vouchsafed to me by this happy circumstance. I even treasured the etymology -- for "circumstance" means "standing around" (as

Swinging on the South

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

Geospatial Inf. — Maps & GIS

Records Mgmt. — Documents

Management   Access   Preservation?   Effective?   Sustainable?

samvera

Repository Powered Approach
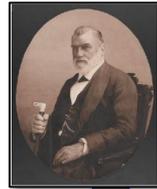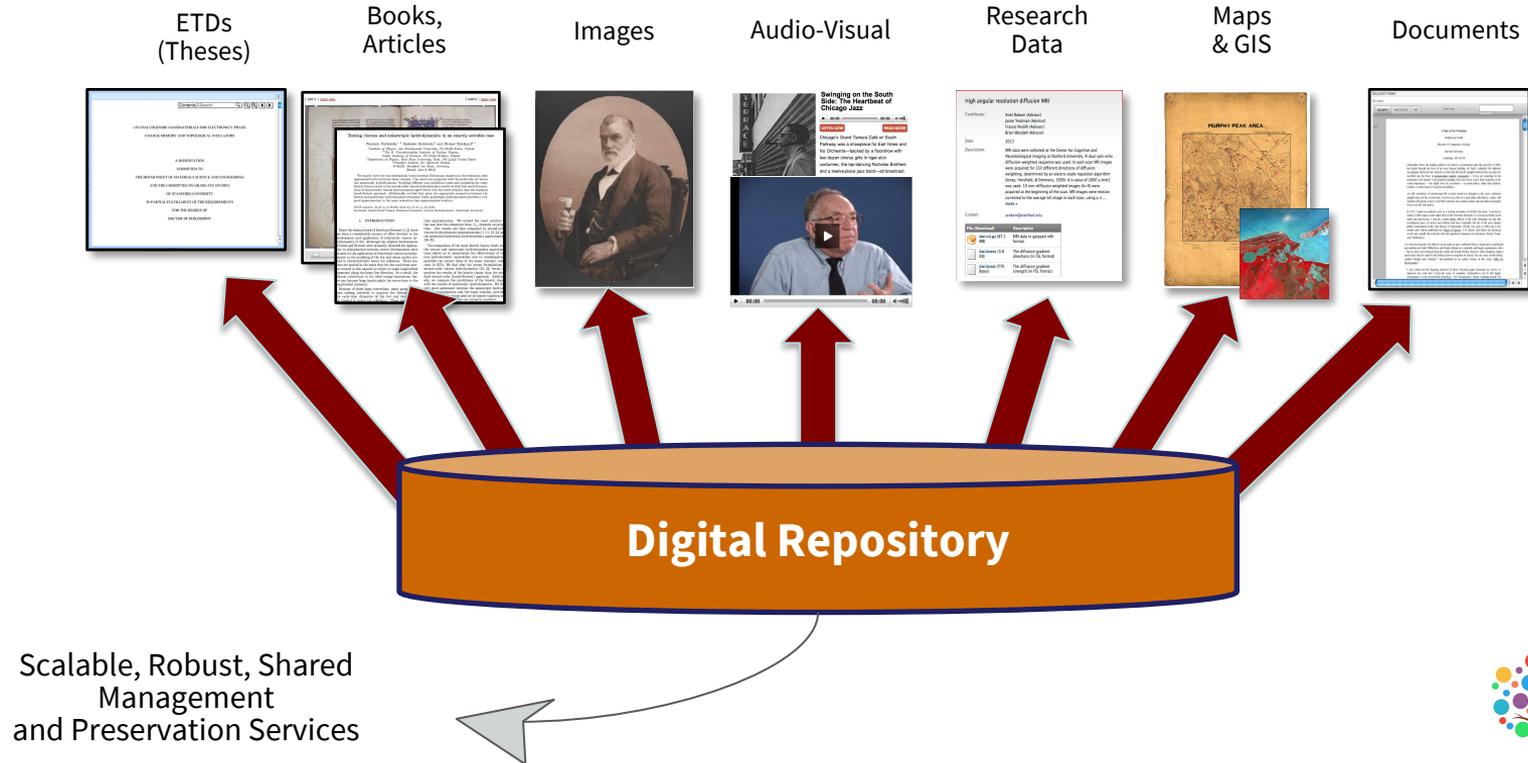
ETDs (Theses)

Books, Articles

Images

Audio-Visual

Research Data

Maps & GIS

Documents

Digital Repository

Scalable, Robust, Shared Management and Preservation Services

samvera

CRUD in Repositories

Create/Submit/Edit
(CUD)

Search/View
(R)

Repository/
Persistent Storage

samvera

CRUD in Repositories

Create/Submit/Edit
(CUD)

Search/View
(R)

Repository/
Persistent Storage

samvera
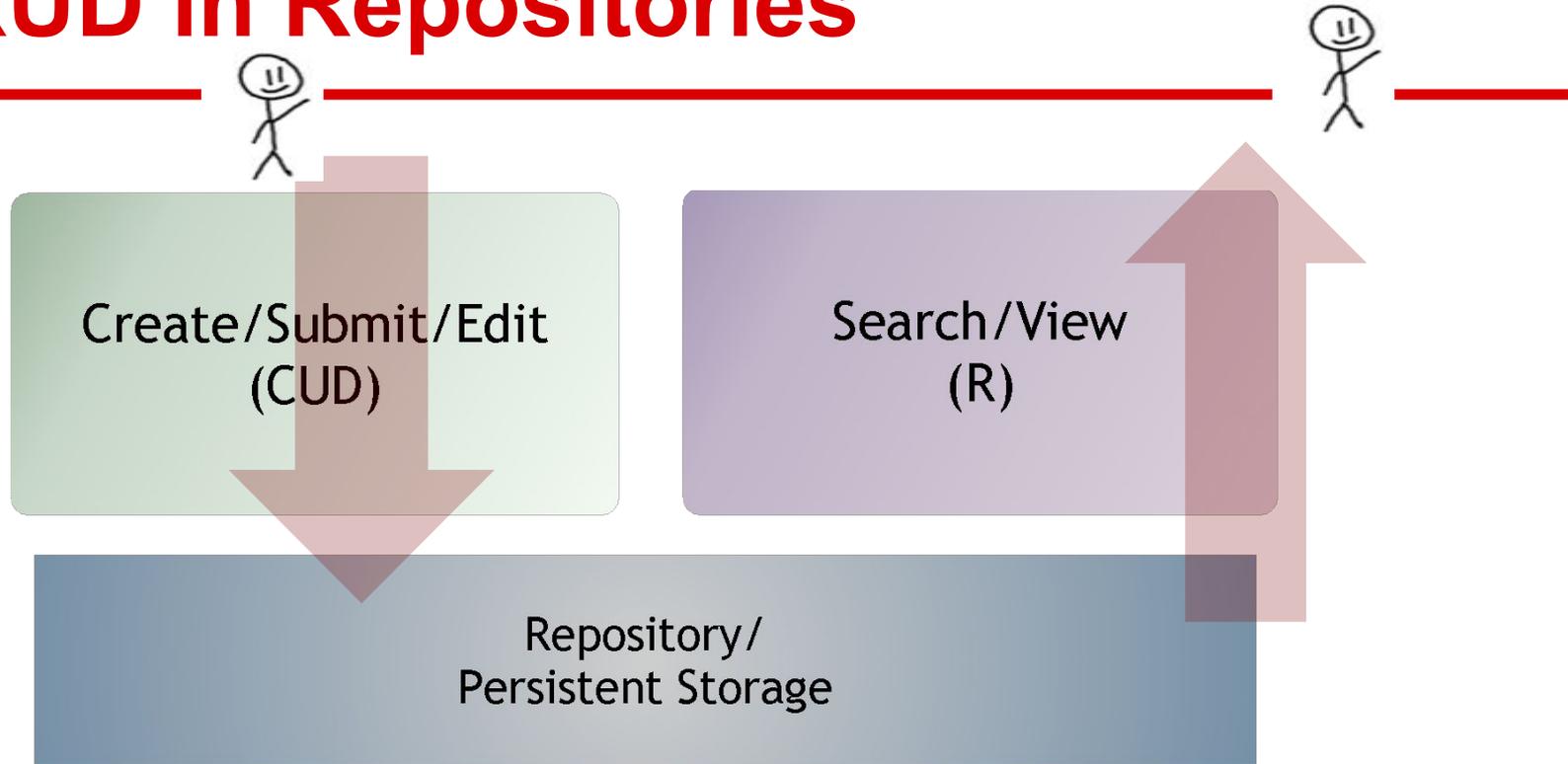
Used by:

- Large Universities
- Small Universities
- Colleges
- Public Broadcasting
- Government Ministry
- National Libraries
- National Lab
- Small Research Labs
- National Digital Repository
- Statewide Digital Libraries
- Science History Institute
- 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!

samvera

Next Up:

How we work
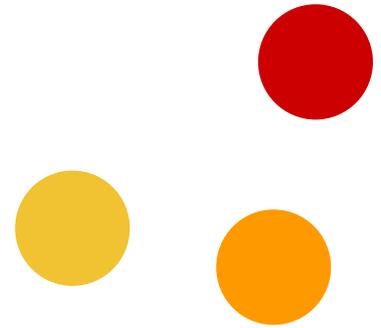
How do I get involved?

samvera

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

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."

samvera

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

samvera

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**
  Topically aligned groups that span organizations & roles

- **Committers**
  Developer community

- **Community adopters**

samvera

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

samvera

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://samvera.org/wiki)
  - 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!)

samvera

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 Export 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

samvera

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

samvera

Community Resources

- **Website -** Introduction to the community: https://samvera.org

- **Wiki -** Broad range of information types: https://samvera.org/wiki

- **Github -** Code, code wikis, technical details:

  https://github.com/samvera
  https://github.com/samvera-labs
  https://github.com/samvera-deprecated

samvera

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

samvera

First Up:

Example Sites ●
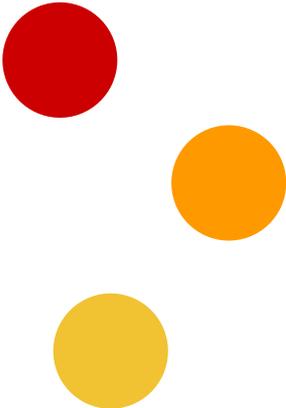
What are folks doing with it?

samvera

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

samvera

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

samvera

Earthworks

- Heterogeneous GIS Content

- Search and browse augmented by spatial search functions

- Includes content harvested from multiple sources

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

samvera

King's Fund Digital Archive

- Running on Hyku software

- Digitised publications by the King's Fund

- Heavy use of IIIF and Universal Viewer

Live Site: https://archive.kingsfund.org.uk/

samvera

Emory ETD Repository

- Hyrax Development

- Electronic Theses and Dissertations focus

- Student submission and workflow focused

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

samvera

Next Up:

Hands-On

Any questions so far?

samvera
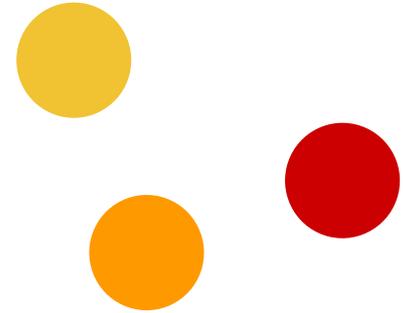
Create an account

1. Visit http://connect18.demos.leaf.ulcc.ac.uk

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"

samvera

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

samvera

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

samvera

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. Is your item there?
6. Log-out - how does your experience differ?

samvera

Deposit a new work

1. Reminder: Logo, Share, Select

2. Add required metadata, attach a file, choose visibility, accept the deposit agreement

3. Save your work

samvera

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. Everyone

    Gated discovery: Public, UoL, Private Collections

2. Admins

    Featured works

    Workflow / Review

    Admin dashboard

Next Up:

Tech Details
(totally optional)

Any questions so far?

samvera
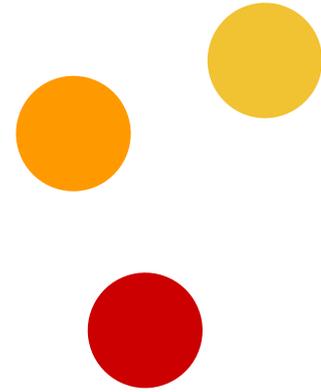
Major Components
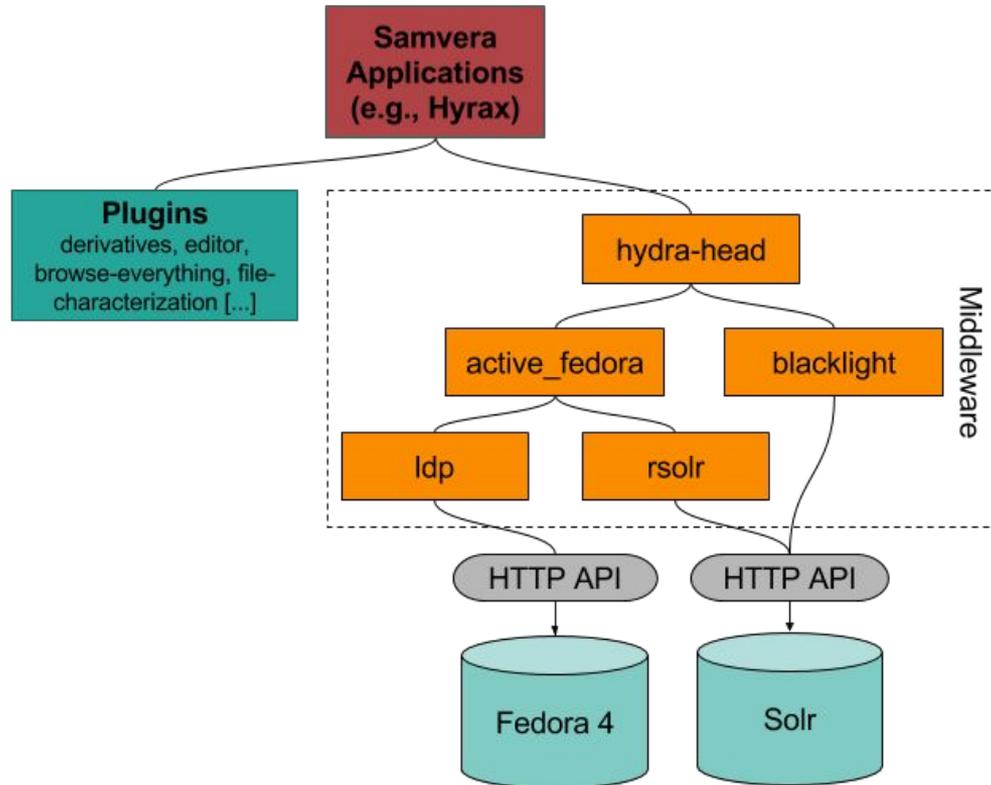
A Note on Fedora

- **F**lexible, **E**xtensible, **D**urable **O**bject **R**epository **A**rchitecture
  - 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

samvera

Current status of Fedora

- Fedora 4.0 released in Dec 2014

    - Major rewrite, incompatible with Fedora 3 and earlier

    - Current release: 4.7.5 as-of 12th February 2018

- Fedora 5, in development

    - Compatible with Fedora 4

    - API specification to allow multiple Fedora software implementations

samvera

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

samvera

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

Hyku

- Availability: DIY, Hosted (via several service providers)
- Goal: turnkey ease of installation and management
- Goal: Provide configured version of Hyrax with multitenancy

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.
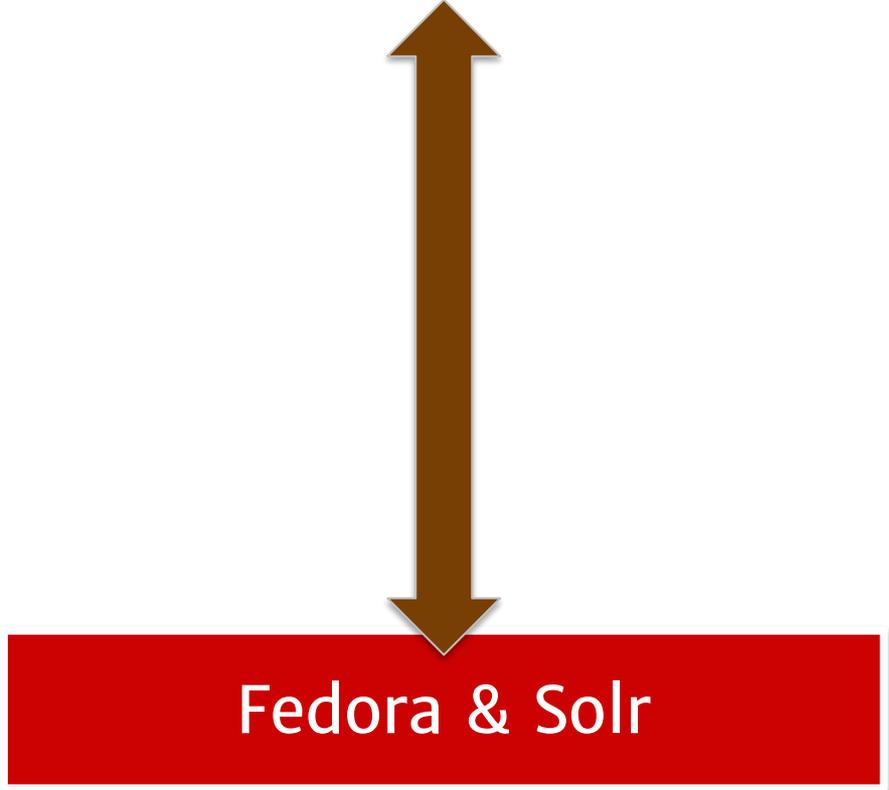
Fedora & Solr

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.

ActiveFedora

Fedora & Solr
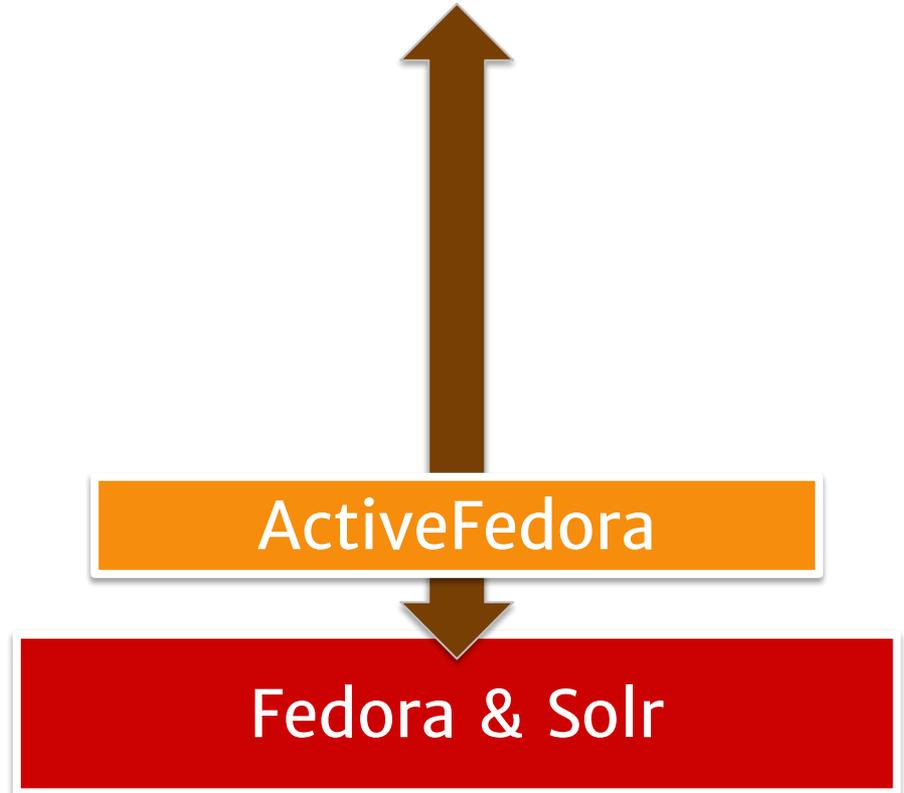
The Stack

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.

Hydra::PCDM

ActiveFedora

Fedora & Solr
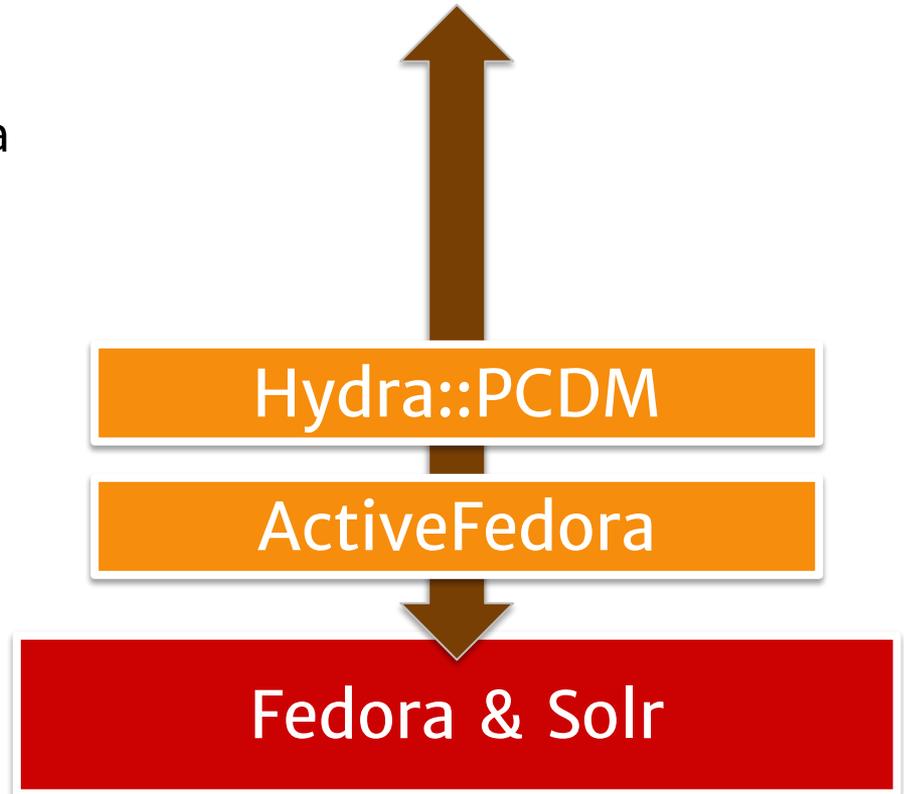
PCDM (interlude)
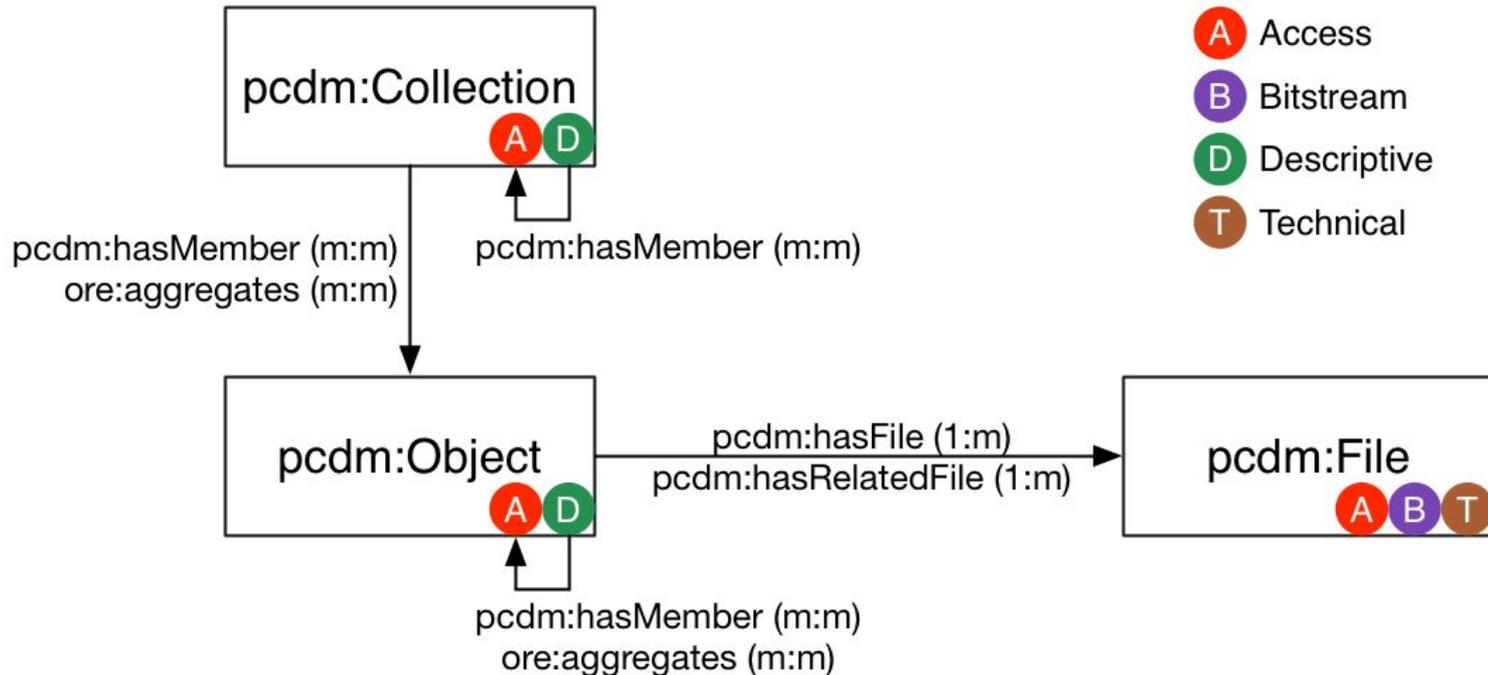
Book Modeled in PCDM
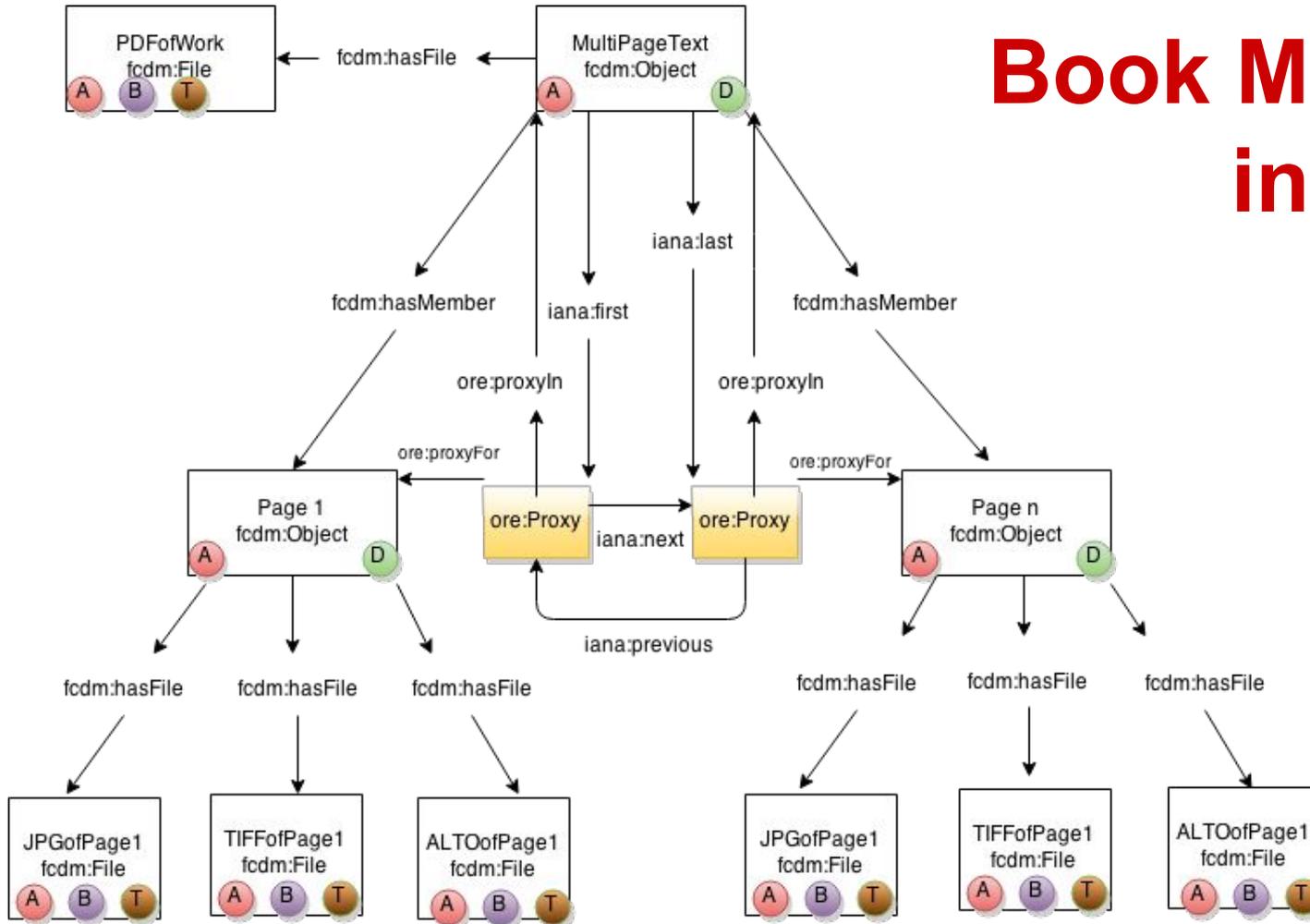
The Stack

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.

Hydra::Works

Hydra::PCDM

ActiveFedora

Fedora & Solr
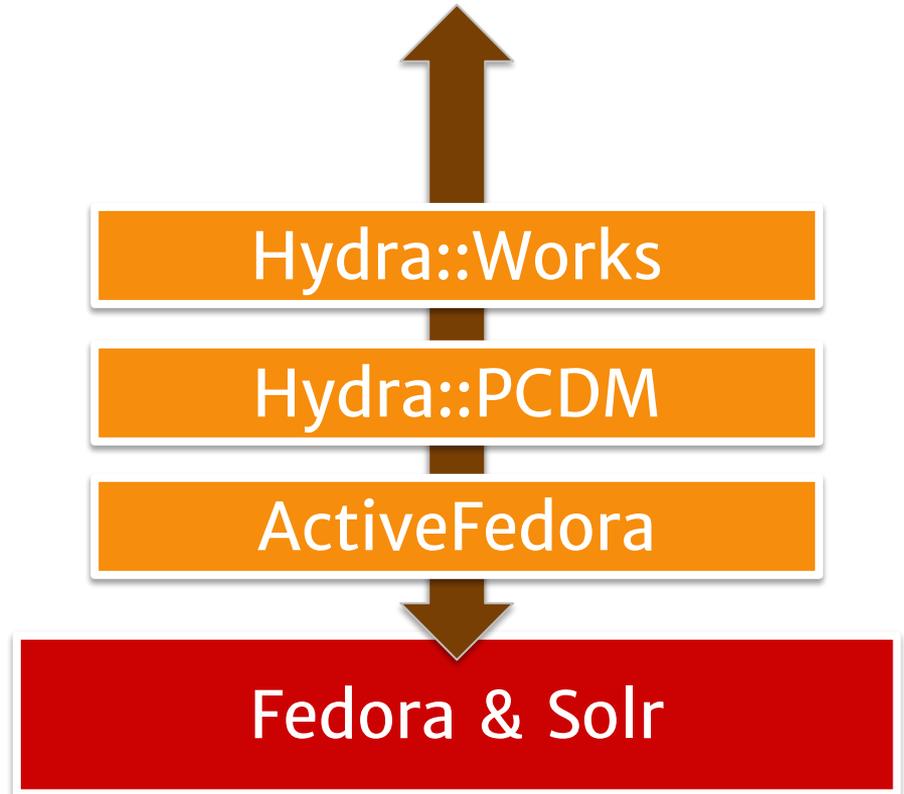
The Stack

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

Hyrax (optional)

Hydra::Works

Hydra::PCDM

ActiveFedora

Fedora & Solr
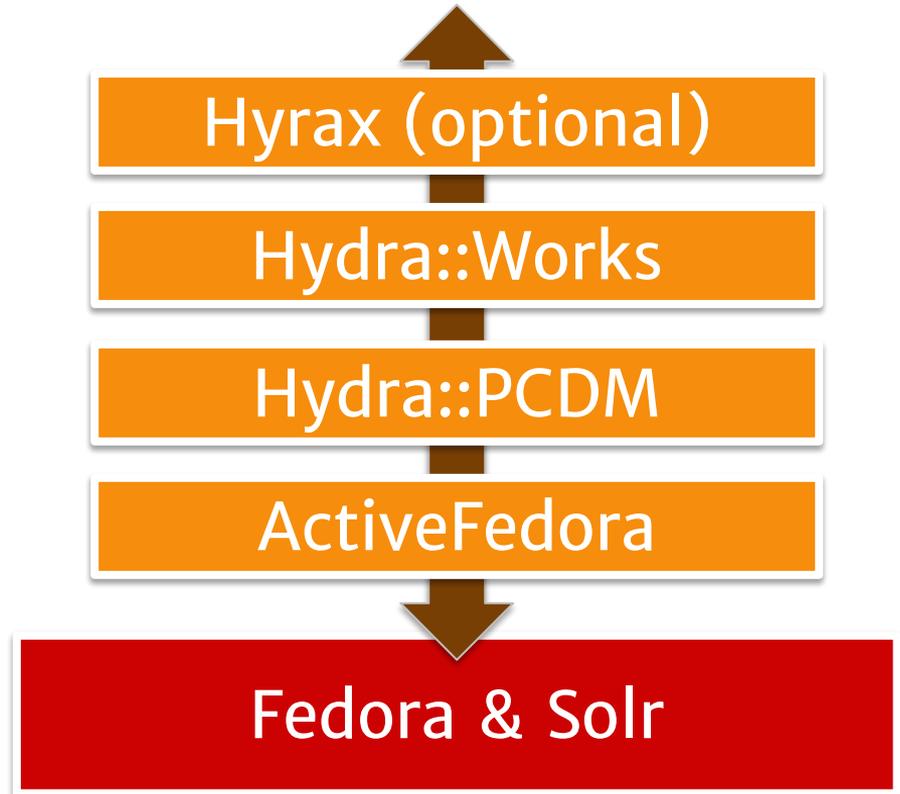
The Stack

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.

Your Application

Hyrax (optional)

Hydra::Works

Hydra::PCDM

ActiveFedora

Fedora & Solr
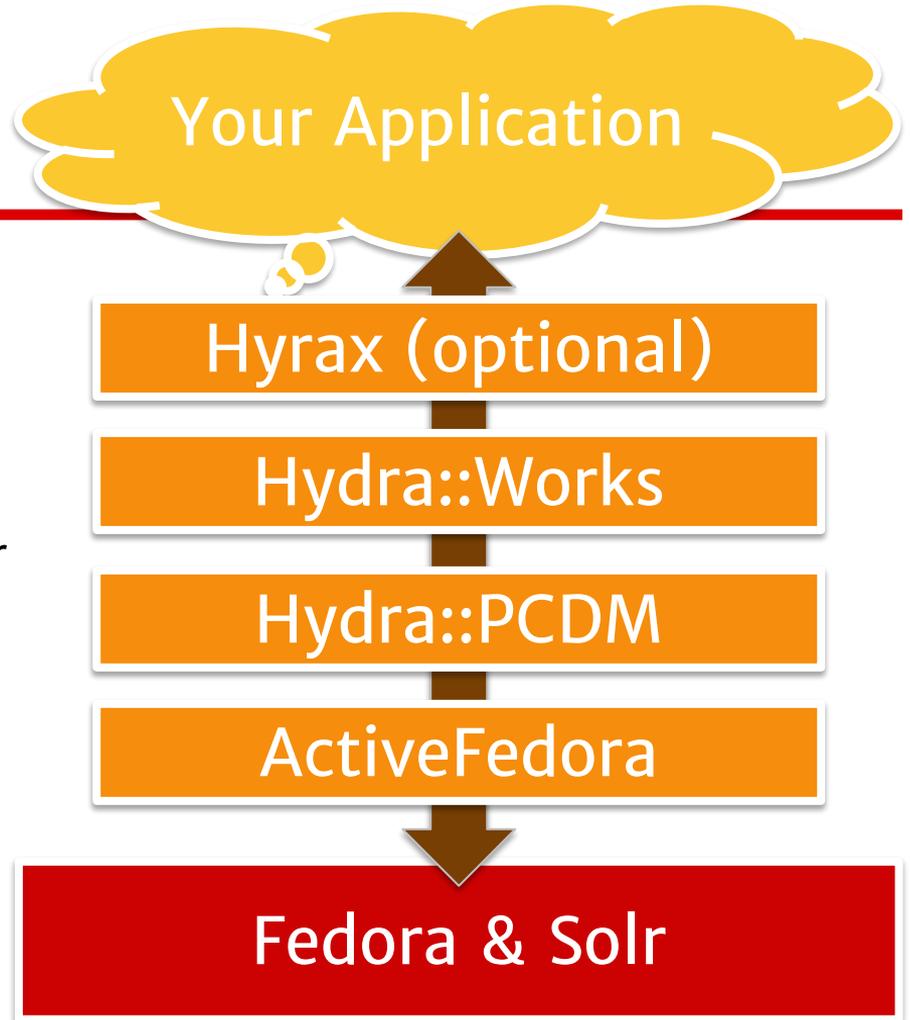
Next Up:

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

samvera

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

samvera

A Worldwide Presence

Samvera's Partners (red), adopters (blue) and solution bundle users (green) - 2017

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
- 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

samvera

Resource Sustainability

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

  …use Samvera

samvera

Solution Sustainability…

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

samvera

Samvera Code Snapshot

1.49M Lines of code

101 Current contributors

19 hours since the last commit

Activity Rating by Open Hub? Very High

Next Up:
What next?
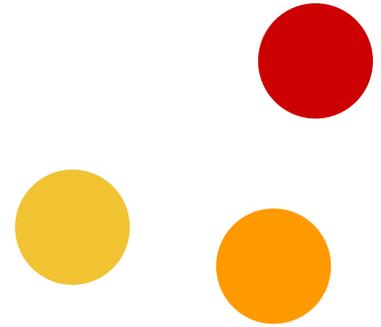
What do I do back home?

samvera

Community Resources

- **Website -** for the managers
  https://samvera.org

- **Wiki -** for the doers
  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

- https://github.com/samvera

- Demo Sites:

  https://tenejo.curationexperts.com/
  http://connect18.demos.leaf.ulcc.ac.uk

- Ask other partners for demos!

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

samvera

Questions

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

julie.allinson@cosector.com
robin.ruggaber@virginia.edu
steve.vantuyl@oregonstate.edu

Thank you

samvera