



cottage labs

Comparing three institutional software repository platforms

Anusha Ranganathan
Hrafn Malmquist
Steven Eardley

anusha@cottagelabs.com
hrafn@cottagelabs.com
steve@cottagelabs.com

us@cottagelabs.com

Three repository software platforms

DSpace 8	Website: https://or24-dspace.cottagelabs.com/ Code: https://github.com/cottagelabs/or24-dspace-angular
Invenio v12 (12.0.0b2.dev0)	Website: https://or24-invenio.cottagelabs.com Code: https://github.com/cottagelabs/or2024-invenio
Hyrax 5	Website: https://or24-hyrax.cottagelabs.com Code: https://github.com/cottagelabs/or2024-hyrax





OR2024 Demo Instance by Cottage Labs



Recent uploads

June 10, 2015 (v1)

Presentation

Open

Indexing repositories: pitfalls and best practices

Acharya, Anurag

Featured talk by Anurag Acharya, Distinguished Engineer, Google, at Open Repositories 2015 (OR2015), the 10th International Conference on Open Repositories, Indianapolis, Indiana.

Part of [Open Repositories 2024](#)

Uploaded on June 3, 2024

1 1

June 3, 2024 (v1)

Dataset

Open

Original dataset for ID 226 Y2O3 in Thermophysical Property Database

Ishikawa, Takehiko; Ishii, Masashi

This is the original dataset for ID 226 Y2O3 in Thermophysical Property Database (<https://thermophys.nims.go.jp/thermophysicalproperty/experiments/226>). The dataset was obtained at Japan Aerospace Exploration Agency (JAXA), and is a part of [Thermophysical Property Original Datasets](https://doi.org/10.48505/nims.3877) (<https://doi.org/10.48505/nims.3877>) as a collection of MDR.

InvenioRDM

InvenioRDM V12

- Packaged version of utilities powering Zenodo; from CERN + numerous partners. Zenodo ported in 2023
- Python & Flask web application with Semantic UI & React



Developing the OR24 InvenioRDM

- On a fresh VM, install CLI tool to bootstrap an instance
`invenio-cli init rdm`
- Minor changes to the included docker-compose; replaced self-signed SSL certificate with LetsEncrypt.
`invenio-cli containers setup --no-demo-data`
`invenio-cli containers build`
- Customisation via override files: sass, jinja template
`invenio-cli containers start`
- Deposit via admin user + webform

InvenioRDM USP

- Backed by CERN
- Developer-friendly tech stack
- Extensible metadata schema & controlled vocabularies
- Very scalable and performant
- The Invenio community

InvenioRDM Challenges

- A young platform; some features hotly anticipated
- Featureset heavily influenced by Zenodo's needs
- Something of a maze of interconnected packages can make development tricky

Hyrax

The screenshot shows the OR24 Hyrax website interface. At the top left, the text "OR24 Hyrax" is displayed. At the top right, there is a language dropdown menu set to "English" and a "Login" button. Below the header is a navigation bar with links for "Home", "About", "Help", and "Contact". A search bar is located on the right side of the navigation bar, with the text "Search OR24 Hyrax" and a "Go" button. Below the navigation bar is a large blue button labeled "Share Your Work" with an upload icon. Underneath this button is a link for "Terms of Use". The main content area is divided into two columns. The left column has two tabs: "Featured Works" and "Recently Uploaded". Below the "Featured Works" tab, it says "No works have been featured". The right column has two tabs: "Explore Collections" and "Featured Researcher". Below the "Explore Collections" tab, there are three collection items, each with a 3D cube icon and a title: "Videos", "Charles Dickens", and "Thermophysical Property Original Datasets". Below these items is a button labeled "View all collections". At the bottom of the page, there is a footer with the text "A service of Samvera." and "OR24 Hyrax v5.0.1" on the left, and "Copyright © 2024 Samvera Licensed under the Apache License, Version 2.0" on the right.

OR24 Hyrax

English Login

Home About Help Contact

Search OR24 Hyrax Enter search terms Go

Share Your Work

Terms of Use

Featured Works Recently Uploaded

No works have been featured

Explore Collections Featured Researcher

Videos

Charles Dickens

Thermophysical Property Original Datasets

View all collections

A service of Samvera.

OR24 Hyrax v5.0.1

Copyright © 2024 Samvera Licensed under the Apache License, Version 2.0

Hyrax

“Hyrax is a [Ruby on Rails Engine](#) built by the [Samvera community](#). Hyrax provides a foundation for creating many different digital repository applications.” ([doc](#))

Hyrax is not a web application.



Developing the OR24 Hyrax based application

- Reused existing docker-compose scripts
- Created a new Hyrax v5 application
- Created two new worktypes
 - Monograph (copied from Koppie)
 - Dataset (to showcase a nested metadata example)
- Commissioned a new machine and deployed the application
- Manually submitted the 15 records

Hyrax USP

- Develop any number of work types with
 - custom metadata model
 - custom workflow and administrative sets (admin sets - curated collections)
- Flexible adapters for metadata and storage with Valkyrie
- Custom workflows
- Integration with other services
- The Samvera community

Hyrax Challenges

- It's written in Ruby ;) and not a traditional RoR application
- It's got a long dependency chain of gems
- Need to strike a balance between feature richness and code complexity

DSpace



Communities & Collections All of DSpace ▾ Statistics

🔍 🌐 Log In ▾

COMPARING THREE INSTITUTIONAL REPOSITORY SOFTWARE PLATFORMS AT OR2024

DSPACE 8

OR2024 Demo Instance by [Cottage Labs](#)

Jens Hunt, Creative Commons Attribution-Share Alike 3.0 Unported license. Via Wikimedia Commons

DSpace

- Turn key institutional repository software
- Using the main branches for version 8



Developing the OR24 DSpace

- Commissioned a new machine and deployed the application
- Forked the DSpace Angular repository and the DSpace backend
- Created a community and a collection
 - Manually submitted the 15 records as an administrator
- Applied basic custom theming and set up certificates
- Ran filter media process to generate thumbnails and extract and index text

DSpace USP

- Highly customisable
- Modern architecture
- Strong community
- Integration with other services

DSpace Challenges

- Arbitrary two level container structure imposed
- Resource intensive
- Complex technologies
- No way to ingest files different directories

Differences 1/3

	DSpace	InvenioRDM	Hyrax
Setup effort	2 to 3 days to deploy a basic repository	~1 day to deploy a basic repository	3 to 4 days to deploy a basic repository
Versioning	Record level	File level	File level
Language	Java	Python	Ruby
Maturity	2001	2015 - Invenio3	2010
Customizing metadata	Administrator	Developer	Developer
Configurable submission workflows	Yes	No	Yes - greater control

Differences 2/3

	DSpace	InvenioRDM	Hyrax
PIDs	Handle, DOI	DOI	DOI
Communities/ Collections	Needs both	Communities	User and administrative collections
Usage statistics	Yes	(yes)	(yes)
Lease	No	No	Yes
Embargo management	Automatic	Automatic	Dashboard
Sharing, proxy deposit and ownership transfer	No	(yes)	Yes

Differences 3/3

	DSpace	InvenioRDM	Hyrax
Lease	No	No	Yes
Embargo management	Automatic	Automatic	Dashboard
Sharing, proxy deposit and ownership transfer	No	No	Yes
Meta tags for discoverability	Google scholar, Signposting	No	Google scholar, Schema.org
Multilingual support	Yes	(Yes)	Yes

Factors for choosing

- The object versioning model, types of metadata model and workflows you need to represent
- The skill set available to support the repository
- Community involvement