Comparing three institutional software repository platforms

Anusha Ranganathan anusha@cottagelabs.com
Hrafn Malmquist hrafn@cottagelabs.com
Steven Eardley steve@cottagelabs.com

us@cottagelabs.com

cottage labs

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

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
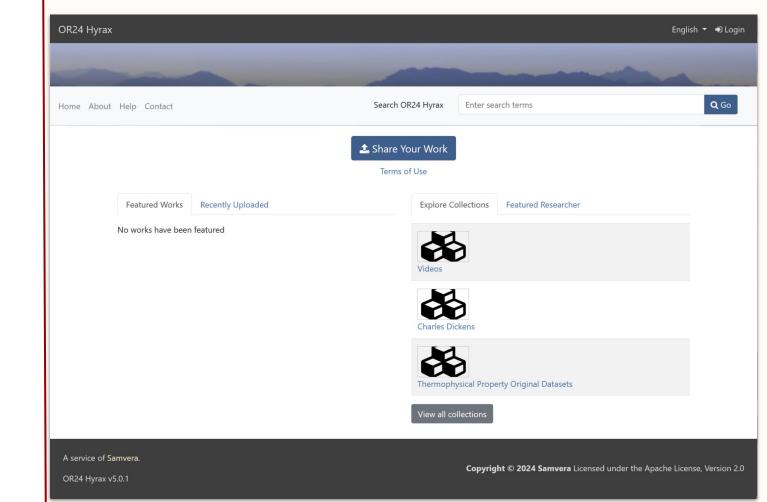
- On a fresh VM, install CLI tool to bootstrap an instance
  `invenio-cli init rdm`

- 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

Cottage Labs: make the most of your data

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

COMPARING THREE INSTITUTIONAL REPOSITORY SOFTWARE PLATFORMS AT OR2024

DSPACE 8

OR2024 Demo Instance by Cottage Labs

Communities & Collections   All of DSpace   Statistics

Q   Log In

OpenRepositories

DSPACE

Cottage Labs: make the most of your data

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 in 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