The Repository Rodeo

Dataverse, DSpace, EPrints, Fedora, Invenio, Islandora, Samvera Open Repositories 2024

Panelists

Dataverse: Gustavo Durand, Harvard University

DSpace: Kristi Park, Texas Digital Library

EPrints: Rory McNicholl, CoSector

Fedora: Kate Dohe, University of Maryland

Invenio: Alex Ioannidis, CERN

Islandora: Donald Moses, University of Prince Edward Island

Samvera: Heather Greer Klein, Samvera

Dataverse

Dataverse

Gustavo Durand
Technical Lead / Architect
IQSS, Harvard University









Introduction to Dataverse

Overview

- An open-source platform to publish, cite, and archive research data
- Built to support multiple types of data, users, and workflows
- Developed at Harvard's Institute for Quantitative Social Science (IQSS) since 2006
- Development funded by IQSS and with grants, in collaboration with institutions around the world
- Core team
 - @ IQSS developers, UX/UI, metadata specialists, curation team, leadership team
 - key contributors from the community with full privileges as IQSS team

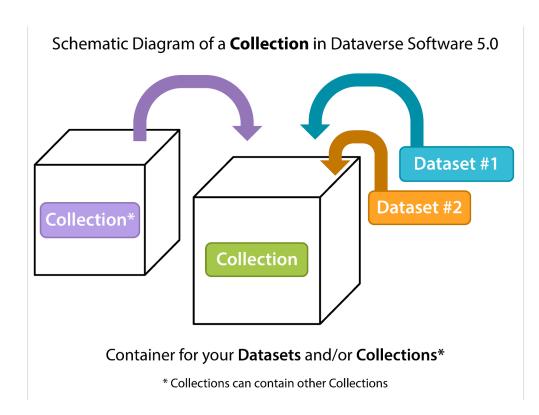
Dataverse Features

https://dataverse.org/software-features

- Main goal of core code is to focus on publishing (citing, sharing, versioning, etc.), FAIR Data principles
- Robust APIs to allow interoperability with "external tools" and other repositories / software

Dataverse Collections

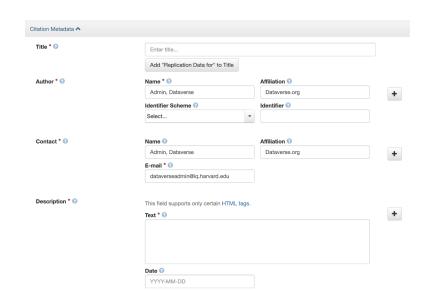
- Ability to create Dataverse collections to organize datasets according to your needs
- Dataverses collections can also contain other collections, enabling any hierarchical structure
- Different rules can be applied for different Dataverse collections,
 e.g. for Metadata, Permissions,
 etc.



Dynamic Metadata

- Metadata is defined dynamically at the database level, allowing for modularly adding new Metadata blocks
- Supports:
 - single or multiple values
 - simple or compound values
 - controlled vocabularies

Choose the metadata fields to use in dataset templates and when adding a dataset to this dataverse.			
☑ Citation Metadata (Required) [+] View fields + set as hidden, required, or optional			
Geospatial Metadata [+] View fields			
☐ Social Science and Humanities Metadata [+] View fields			
☐ Astronomy and Astrophysics Metadata [+] View fields			
☐ Life Sciences Metadata [+] View fields			
☐ Journal Metadata [+] View fields			



Flexible Permission System

- Supports multiple workflows by controlling who can add to your Dataverse collection, what they can, and what role they have on and created Datasets
- Roles are defined as a set of permissions to grant to users or to groups
- Groups can be defined statically or dynamically (e.g. users logging in from the same institution, via Shibboleth)

Edit Access

Who can add to this dataverse?

- Anyone adding to this dataverse needs to be given access
- Anyone with a Dataverse account can add sub dataverses
- Anyone with a Dataverse account can add datasets
- Anyone with a Dataverse account can add sub dataverses and datasets

When a user adds a new dataset to this dataverse, which role should be automatically assigned to them on that dataset?

- Contributor Edit metadata, upload files, and edit files, edit Terms, Guestbook, Submit datasets for review
- Curator Edit metadata, upload files, and edit files, edit Terms, Guestbook, File Restrictions (Files Access + Use), Edit Permissions/Assign Roles + Publish

Save Changes

Cancel

2 Users/Groups

User/Group Name (Affiliation) ≎	ID \$	Role ≎
Dataverse Admin (Dataverse.org)	@dataverseAdmin	Admin
Anyone with a Dataverse account	:authenticated-users	Dataverse + Dataset Creator

Robust APIs

- APIs for search, deposit, access, administration, metrics, etc.
- Additional APIs for harvesting (discovery) and interoperability with other systems
- External tools can be registered via APIs, so that Dataverse can provide links in the UI, then user is sent to tool to preview, explore, configure, and more

API Guide

Contents:

- Introduction
 - What is an API?
 - Types of Dataverse Software API Users
 - API Users Within a Single Dataverse Installation
 - Users of Integrations and Apps
 - Power Users
 - Support Teams and Superusers
 - Sysadmins
 - In House Developers
 - API Users Across the Dataverse Project
 - Developers of Integrations, External Tools, and Apps
 - Developers of Dataverse Software API Client Libraries
 - Developers of The Dataverse Software Itself
 - How This Guide is Organized
 - Getting Started
 - API Tokens and Authentication
 - Lists of Dataverse APIs
 - Client Libraries
 - Examples
 - Frequently Asked Questions
 - Getting Help
- · Getting Started with APIs
 - Servers You Can Test With
 - o Getting an API Token
 - curl Examples and Environment Variables
 - Depositing Data
 - Creating a Dataverse Collection
 - Creating a Dataset
 - Uploading Files
 - Publishing a Dataverse Collection
 - Publishing a Dataset

Dataverse Technology

Payara 6



Java 17 Java EE10

- Presentation: RESTful API, JSF (PrimeFaces)
- Business: EJB, Transactions, Asynchronous, Timers
- Storage: JPA (Entities), Bean Validation

Front end: SPA (React) - In Beta!

Storage: Postgres, Solr, File System / Swift / S3

Dataverse Community

Dataverse Community

- 190+ Github Contributors
- Hundreds of members of the Dataverse Community developers, researchers, librarians, data scientists
 - Workshops & Trainings
 - UX/UI Testing & Interviews
 - Global Dataverse Community Consortium
 - Dataverse Google Group / Matrix / Community Slack
 - Dataverse Community Calls
 - Dataverse Community Meeting

Dataverse Cup **Y**



Global Dataverse Community Consortium

Supporting Dataverse repositories around the world

The Global Dataverse Community Consortium (GDCC) is dedicated to providing international organization to existing Dataverse community efforts, and will provide a collaborative venue for institutions to leverage economies of scale in support of Dataverse repositories around the world.



http://DataverseCommunity.Global

Dataverse Community

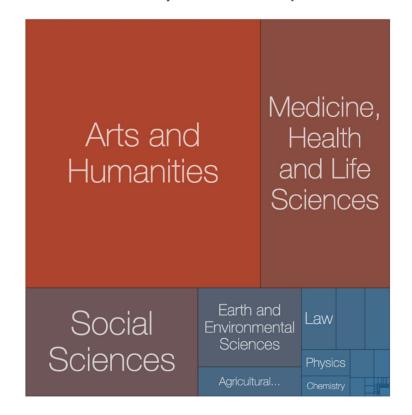
• 118 (self reporting) installations around the world



The Data (dataverse.org/metrics)

Datasets by Most Common Subject

- 118 installations
- 16,600 Dataverse Collections*
- 426,000 Datasets*
- 6,120,000 Files*
- 86,500,000 File Downloads*



^{*} metrics collected from 96 installations (running 4.9 and newer)

New Features

Releases since last year

Four releases:

- **5.14** August 2023
- **6.0** September 2023
- **6.1** December 2023
- **6.2** April 2024

Dataverse 5.14

- Developed in parallel with 6.0
- Highlights https://github.com/IQSS/dataverse/releases/tag/v5.14
 - Registering PIDs (DOIs or Handles) for files in select collections
 - Signposting for Dataverse
 - Permalinks support
 - Mechanism Added for Adding External Exporters
 - Creating datasets with incomplete metadata through API
 - Support for Grouping Dataset Files by Folder and Category Tag
 - Indexing performance improved
 - Changes to PID Provider JVM Settings
 - In preparation for a future feature to use multiple PID providers at the same time
 - New External Tool Type "query" type tool

Registering PIDs for files in select collections

- It is now possible to configure registering PIDs for files in individual collections.
 - registration of PIDs for files can be enabled in a specific collection when it is disabled instance-wide
 - Or it can be disabled in specific collections where it is enabled by default
- Available via the Admin API

:FilePIDsEnabled

Toggles publishing of file-level PIDs for the entire installation. By default this setting is absent and Dataverse Software assumes it to be false. If enabled, the registration will be performed asynchronously (in the background) during publishing of a dataset.

It is possible to override the installation-wide setting for specific collections, see :AllowEnablingFilePIDsPerCollection. For example, registration of PIDs for files can be enabled in a specific collection when it is disabled instance-wide. Or it can be disabled in specific collections where it is enabled by default. See Change Collection Attributes for details.

To enable file-level PIDs for the entire installation:

```
``curl -X PUT -d 'true' http://localhost:8080/api/admin/settings/:FilePIDsEnabled``
```

If you don't want to register file-based PIDs for your entire installation:

```
``curl -X PUT -d 'false' http://localhost:8080/api/admin/settings/:FilePIDsEnabled``
```

:AllowEnablingFilePIDsPerCollection

Toggles whether superusers can change the File PIDs policy per collection. By default this setting is absent and Dataverse Software assumes it to be false.

For example, if this setting is true, registration of PIDs for files can be enabled in a specific collection when it is disabled instance-wide. Or it can be disabled in specific collections where it is enabled by default. See Change Collection Attributes for details.

To enable setting file-level PIDs per collection:

```
``curl -X PUT -d 'true' http://localhost:8080/api/admin/settings/:AllowEnablingFilePIDsPerCollection`
```

When :AllowEnablingFilePIDsPerCollection is true, setting File PIDs to be enabled/disabled for a given collection can be done via the Native API - see Change Collection Attributes in the Native API Guide.

Dataverse 6.0

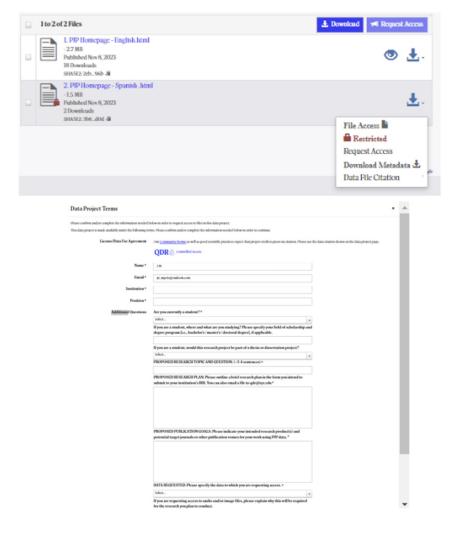
- Functionally the same as 5.14, but featuring major upgrades to core components:
 - The required Java version has been increased (from version 11) to version 17
 - Payara application server has been upgraded to version
 - Solr has been upgraded to version 9.3.0
 - PostgreSQL 13 remains the tested and supported version
- https://github.com/IQSS/dataverse/releases/tag/v6.0

Dataverse 6.1

- First "timed" release
- Highlights https://github.com/IQSS/dataverse/releases/tag/v6.1
 - Guestbook at request
 - Collection-level storage quotas
 - Globus support (experimental), continued
 - JSON Schema for datasets
 - OpenID Connect (OIDC) improvements
 - Solr improvements
 - New release of Dataverse Previewers (including a Markdown previewer)
 - External tools: configure tools now available at the dataset level

Guestbook at request

- Dataverse can now be configured to display any configured guestbook to users when they request restricted files
 - The default, showing guestbooks when files are downloaded, remains as it was in prior Dataverse versions
- The global default can be overridden at the collection level and at the individual dataset level



Collection-level storage quotas

- Instance admins can now define storage quota limits for specific collections
 - Storage quotas are inherited by subcollections
 - A storage quota defined on a child sub-collection overrides whatever quota that may be defined on the parent, or inherited from an ancestor
- the users will be informed of the remaining storage allocation on the file upload page
- Still an experimental feature; learn more at <u>https://guides.dataverse.org/en/6.1/admin/collectionquotas.html</u>

Search User Guide Admin Guide Dashboard External Tools Managing Harvesting Clients Managing Harvesting Server and Metadata Customization Metadata Export Dataverse Installation Application Make Data Count User Administration Managing Datasets and Dataverse Collections Solr Search Index IP Groups Mail Domain Groups Storage Quotas for Collections Monitoring Reporting Tools and Common Queries Maintenance Backups Troubleshooting

Storage Quotas for Collections

Please note that this is a new and still experimental feature (as of Dataverse v6.1 release).

Instance admins can now define storage quota limits for specific collections. These limits can be set, changed and/or deleted via the provided APIs (please see the Collection Storage Quotas section of the Native API guide). The Read version of the API is available to the individual collection admins (i.e., a collection owner can check on the quota configured for their collection), but only superusers can set, change or disable storage quotas.

Storage quotas are inherited by subcollections. In other words, when storage use limit is set for a specific collection, it applies to all the datasets immediately under it and in its sub-collections, unless different quotas are defined there and so on. Each file added to any dataset in that hierarchy counts for the purposes of the quota limit defined for the top collection. A storage quota defined on a child sub-collection overrides whatever quota that may be defined on the parent, or inherited from an ancestor.

For example, a collection A has the storage quota set to 10GB. It has 3 sub-collections, B, C and D. Users can keep uploading files into the datasets anywhere in this hierarchy until the combined size of 10GB is reached between them. However, if an admin has reasons to limit one of the sub-collections, B to 3GB only, that quota can be explicitly set there. This both limits the growth of B to 3GB, and also guarantees that allocation to it. Le. the contributors to collection B will be able to keep adding data until the 3GB limit is reached, even after the parent collection A reaches the combined 10GB limit (at which point A and all its subcollections except for B will become read-only).

We do not yet know whether this is going to be a popular, or needed use case - a child collection quota that is different from the quota it inherits from a parent. It is likely that for many instances it will be sufficient to be able to define quotas for collections and have them apply to all the child objects underneath. We will examine the response to this feature and consider making adjustments to this scheme based on it. We are already considering introducing other types of quotas, such as limits by users or specific storage volumes.

Please note that only the sizes of the main datafiles and the archival tab-delimited format versions, as produced by the ingest process are counted for the purposes of enforcing the limits. Automatically generated "auxiliary" files, such as rescaled image thumbnails and metadata exports for datasets are not.

When quotas are set and enforced, the users will be informed of the remaining storage allocation on the file upload page together with other upload and processing limits.

Part of the new and experimental nature of this feature is that we don't know for the fact yet how well it will function in real life on a very busy production system, despite our best efforts to test it prior to the release. One specific issue is having to update the recorded storage use for every parent collection of the given dataset whenever new files are added. This includes updating the combined size of the root, top collection - which will need to be updated after every file upload. In an unlikely case that this will start causing problems with race conditions and database update conflicts, it is possible to disable these updates (and thus disable the storage quotas feature), by setting the dataverse.storageuse.disable-storageuse-increments JVM setting to true.

Dataverse 6.2

- Highlights https://github.com/IQSS/dataverse/releases/tag/v6.2
 - Search and Facet by License
 - Support for Using Multiple PID Providers
 - Rate Limiting
 - Simplified SMTP Configuration
 - Harvesting Handle Missing Controlled Values
 - Ingested Tabular Data Files Can Be Stored Without the Variable Name Header
 - Uningest/Reingest Options Available in the File Page Edit Menu
 - New Accounts Metrics API
 - Several Existing APIs extended

Rate limiting

- The option to rate limit has been added to prevent users from over taxing the system either deliberately or by runaway automated processes
- Rate limiting can be configured on a tier level
 - tier 0 reserved for guest users
 - tiers 1+ for authenticated users
 - Superuser accounts are exempt from rate limiting
- Still an experimental feature; learn more at https://guides.dataverse.org/en/6.2/install-ation/config.html#rate-limiting

Search User Guide Admin Guide API Guide Installation Guide Introduction Preparation Prerequisites Configuration Upgrading OAuth Login Options OpenID Connect Login Options External Tools Advanced Installation Developer Guide Container Guide Style Guide QA Guide

Rate Limiting

Rate limiting has been added to prevent users from over taxing the system either deliberately or by runaway automated processes. Rate limiting can be configured on a tier level with tier of being reserved for guest users and tiers 1-any for authenticated users. Superuser accounts are exempt from rate limiting. Rate limits can be imposed on command APIs by configure the tier, the command, and the hourly limit in the database. Two database settings configure the rate limiting. Note: If either of these settings exist in the database rate limiting will be enabled (note that a Payara restart is required for the setting to take effect). If neither setting exists rate limiting is disabled.

:RateLimitingDefaultCapacityTiers is the number of calls allowed per hour if the specific command is not configured. The values
represent the number of calls per hour per user for tiers 0,1,... A value of -1 can be used to signify no rate limit. Tiers not
specified in this setting will default to - f (No Limit), l.e., of "10000" is equivalent to -d "10000,-1,-1,..."

curl http://localhost:8080/api/admin/settings/:RateLimitingDefaultCapacityTiers -X PUT -d '10000,20000'

:RateLimitingCapacityByTierAndAction is a JSON object specifying the rate by tier and a list of actions (commands). This allows
for more control over the rate limit of individual API command calls. In the following example, calls made by a guest user (tier 0)
for API GetLatestPublishedDatasetVersionCommand is further limited to only 10 calls per hour, while an authenticated user (tier
1) will be able to make 30 calls per hour to the same API.

rate-limit-actions.json Example JSON for RateLimitingCapacityByTierAndAction

curl http://localhost:808/api/admin/settings/:RateLimitingCapacityByTierAndAction — X PUT — d '[{"tier": 0, "limitPerHour": 10, "actions": ["GetLatestPublishedDatasetVersionCommand", "GetPrivateUrlCommand", "GetDatasetCommand", "GetLatestAccessibleDatasetVersionCommand"]} { "tier": 0, "limitPerHour": 1, "actions": ["CreateGuestbookResponseCommand", "UpdateDatasetVersionCommand", "DestroyDatasetCommand", "DeleteDataFil eCommand", "FinalizeDatasetPublicationCommand", "BetLatestPublishDatasetCommand", "GetPrivateUrlCommand", "GetPrivateUrlCommand", "GetDatasetCommand", "GetDatasetCommand", "GetDatasetCommand", "GetDatasetCommand", "DeleteDataFileCommand", "FinalizeDatasetUpslicationCommand", "PublishDatasetCommand", "DeleteDataFileCommand", "FinalizeDatasetPublicationCommand", "PublishDatasetCommand")]}

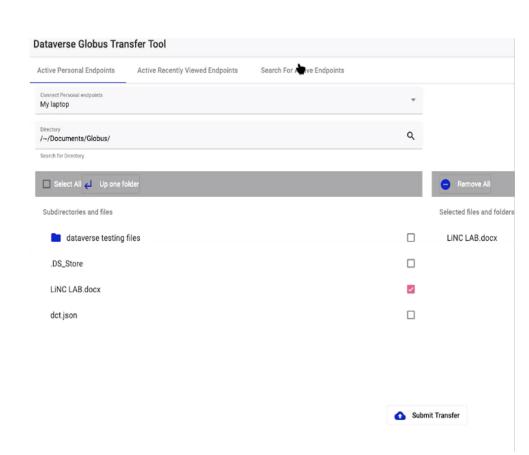
New Tools

Recent External Tools

- New File Previewers
 - Rich HTML previewer
 - RO-Crate
 - Shape files/geotiff
 - NcMI
 - o H5 Web
- Updated Globus Transfer Tool
- "Query" type tools for tabular data files
 - Ask the Data
- Dataset level "configure" tools
 - Turbo Curator

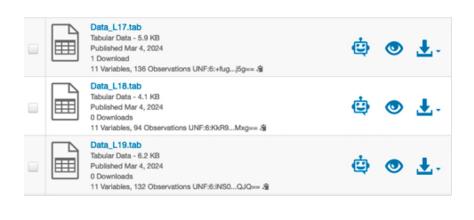
Globus Transfer Tool

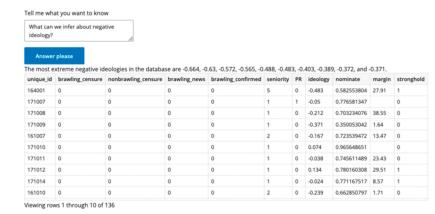
- Support upload and download in Dataverse beyond the limitations of using HTTPS & existing APIs
- Ability to add files or just file metadata (leaving file at source) to Dataverse
- Integration aims to support use cases for:
 - Medium to large data transfers into/out of Dataverse managed endpoints
 - Referencing data on trusted remote
 Globus endpoints
 - Integration with Computation
 - Sensitive Data



Ask the Data

- File level query tools allow the user to ask questions (e.g. natural language queries) of a data table's contents without having to download the file
- First experimental query tool developed at IQSS (by Stefano!),
 Ask the Data

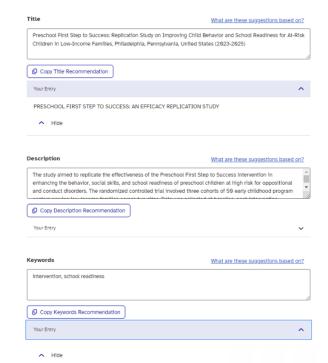




Turbo Curator

- First to use new Dataset level configure tools that allow (authorized) users to send metadata about the dataset back to Dataverse
- uses Open Al's ChatGPT & ICPSR best practices to provides recommendation to enhance metadata & generate meaningful titles, descriptions, and keywords
- Already in use with Harvard
 Dataverse





Future Plans

Activities for Dataverse Team @ Harvard

- Harvard Dataverse Support
- Community Development Facilitation
- NIH GREI (Generalist Repository Ecosystem Initiative)
 - Individual Proposal
 - "Coopetition" Activities
 - Dataverse, Dryad, Figshare, Mendeley Data,
 Open Science Framework, Vivli
- Other Dataverse related projects and partnerships

NIH GREI Program Activities

- Remote Large Storage Support
- Controlled Vocabularies for Biomedical
- Discovery for DDI-CDI
- Software and Biomedical Workflows
- Harvesting and Sharing Metadata Across Repositories
- Usage Metrics Make Data Count Support
- Revisiting of Sensitive Data Support
- Evaluation and Evolution of Architecture
- NIH Data Management Plans
- Training

Upcoming Release Plans

- new Timed Release Plan; every 3 months
- Next release 6.3 upcoming, mid-June 2024
 - Includes Globus and SPA updates
- Beyond
 - Continue SPA work for eventual retirement of current
 UI
 - Marketplace for easier configuration of modular components - exporters, metadata blocks, etc.
 - More Integrations and External tools!

Dataverse Roadmap

https://www.iq.harvard.edu/roadmap-dataverse-project

- Strategic Goals
- Implementation, Planning, Future

Thank you

Dataverse Community Meeting 2025 University of North Carolina at Chapel Hill





Open source research data repository software



Enjoy full control over your data. Receive web visibility, academic credit, and increased citation counts. A personal dataverse is easy to set up, allows you to display your data on your personal website, can be branded uniquely as your research program, makes your data more discoverable to the research community, and satisfies data management plans. Want to set up your personal dataverse?



Seamlessly manage the submission, review, and publication of data associated with published articles. Establish an *unbreakable link* between *articles in your journal* and *associated data*.

Participate in the open data movement by using Dataverse as part of your journal data policy or list of repository recommendations. Want to find out more about journal dataverses?



Establish a research data management solution for your community. Federate with a growing list of Dataverse repositories worldwide for increased discoverability of your community's data. Participate in the drive to set norms for sharing, preserving, citing, exploring, and analyzing research data. Want to install a Dataverse repository?



Participate in a vibrant and growing community that is helping to drive the norms for sharing, preserving, citing, exploring, and analyzing research data. Contribute code extensions, documentation, testing, and/or standards. *Integrate research analysis, visualization* and *exploration tools*, or other research and data archival systems with Dataverse. Want to contribute?

https://dataverse.org https://github.com/igss/dataverse

DSpace

A turnkey institutional repository application

DSPACE

Est. 2002

DSpace Overview





The platform – DSpace includes a core set of functionality out of the box that can be extended to, or integrated with, complementary services and tools in the larger scholarly ecosystem.



Its community – DSpace has a community of highly-skilled developers, engaged and active users, and supporting institutions committed to continuously expanding and improving the software.



Its Governance - based on a representative, community-based membership model incrementally rewarding members who have made the greatest commitments to the project.



>3,500 Installations*



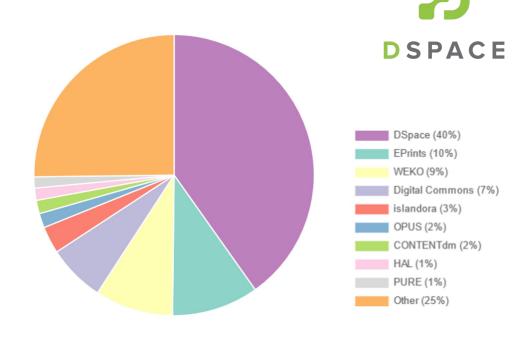
>120 Countries



109 Members



12 <u>Registered Service</u> Providers



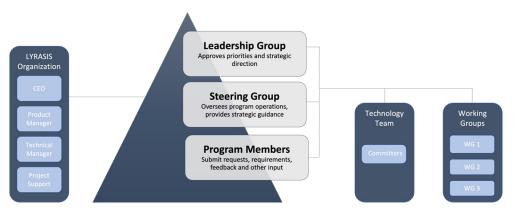
The most widely adopted platform for repositories

OpenDOAR Statistics

https://v2.sherpa.ac.uk/view/repository_visualisations/1.html

DSpace Community

- ★ Defined Governance
- ★ DSpace Community Advisory Team (DCAT)
- ★ 109 Organizational Members
- ★ 23 National User Groups
- ★ Lyrasis organizational home





★ Strategic Partnerships















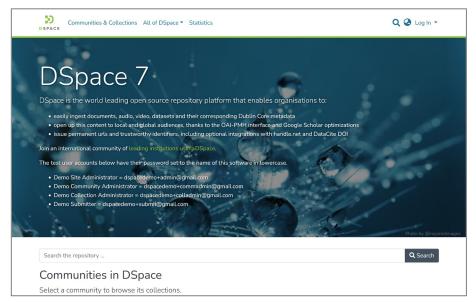
Largest release in DSpace history

- New Angular UI
- REST API

No new feature development for DSpace 7

Bug and security fixes only

Non-standard release schedule





Looking Ahead: DSpace 8

Release date: Approx. June 17-24

DSpace 8.0 had a shortened development cycle and DSpace Steering established four priorities

Include major features which just missed the 7.x series of releases

FEATURES INCLUDE

- 1
- ➤ COAR Notify
- ➤ OpenAIRE Correction service
- ➤ Port "REST-based Quality Control Reports"
- ➤ Basic Duplicate Detection in Submission form

2

New features which empower users (focus on the Admin User Interface)

FEATURES INCLUDE

3

Any other features which existed in 6.x but missed the 7.x series of releases

FEATURES INCLUDE

4

Necessary bug fixes, accessibility fixes, performance improvements & dependency upgrades

FEATURES INCLUDE



- Request withdrawal or reinstatement of Items
- > "Processes" page reorganization
- ➤ Edit Item supports authority control lookup & lookup via external sources
- > Item Submission forms can be configured for entire Community

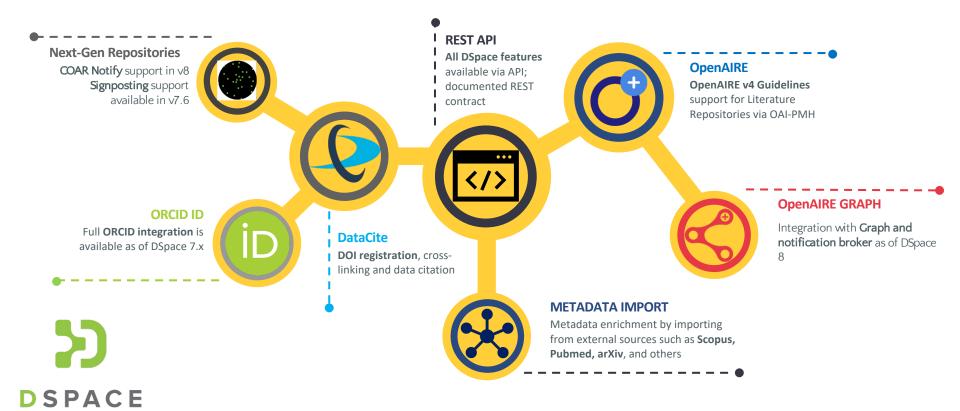


- Advanced Search
- Search facets on Homepage / Community / Collection
- ➤ Search within Community / Collection



- Large number of accessibility fixes (header, navbar, search, admin tools, MyDSpace, etc.)
- Accessibility fixes backported to 7.6.2 (due in July/Aug)
- Major dependency upgrades:
 - Backend upgraded to Spring v6
 - o Frontend upgraded to Angular 17

Towards an interoperable, standards-based ecosystem



Try out DSpace 8

https://wiki.lyrasis.org/display/DSPACE/Try+out+DSpace+8

- ★ DSpace 8 UI demo
 https://sandbox.dspace.org/home
 (uses the REST API demo as backend)
- ★ DSpace 8 REST API demo
 https://sandbox.dspace.org/server/



Stay up-to-date with DSpace



Want to stay up-to-date with latest development? Join us in Slack: https://wiki.lyrasis.org/display/DSPACE/Slack

Want general updates or have questions? See our Mailing Lists: https://wiki.lyrasis.org/display/DSPACE/Mailing+Lists

Other support options also available (StackOverflow, etc): https://wiki.lyrasis.org/display/DSPACE/Support

Funding & Support



DSpace is funded / developed / supported by its community

- **★** Contribute code
- ★ Become a member
- ★ Give to the DSpace Development Fund
- ★ Join <u>DCAT</u> discussions
- ★ Contribute to "Learning DSpace" documentation
- ★ Join (or create) your local DSpace User Group



Thank you to our 109 members!!!

★ https://www.lyrasis.org/programs/Pages/DSpace.aspx

EPrints



Rory McNichol (CoSector, University of London)

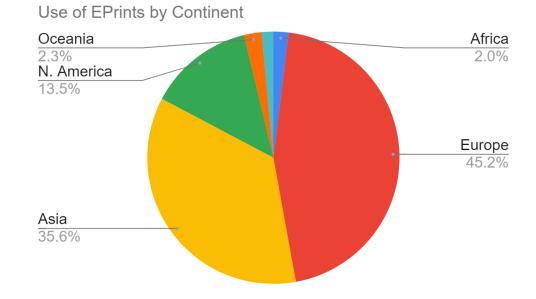
What does EPrints do?

- Open Access Repository Publications, Datasets, Open Education, Media
- Lightweight, very flexible data model(s)
- Open source, community driven, plugin based architecture
- Rodeo leaders in SEO
- Integration with 3rd party platforms and services
- Trusted solutions for compliance
- Fully customisable appearance and branding



The EPrints Repository Community

- 750+ EPrints repositories in ROAR
- Service providers
 - EPrints Services (Southampton, UK)
 - CoSector (London, UK)
 - Open Journal Theme (Indonesia)
- Many independently run repositories
- Active support community
- Regular software releases









































College of

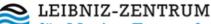
Ireland







U



für Marine Tropenforschung





























INSTITUTE FOR ADVANCED STUDIES





Sheffield University

















The Career University

ROYAL COLLEGE OF MUSIC



























Flexible Data Model

Fully customisable data model and workflow which allows the repository to...

- adapt to new data standards
- integrate with new PID infrastructure
- meet compliance requirements















Integrations

Repository standard interfaces allows EPrints to integrate with other parts of the research infrastructure...

- CRISes
- Archiving Platforms
- Other repositories!















Discoverability and Presentation of Research

EPrints also ensures that research is discoverable and presented in its best light....

- Optimised content, navigation and record metadata for SEO and Google Scholar
- Signposting to help machines navigate scholarly content
- Visualisations to highlight impact and prestige



















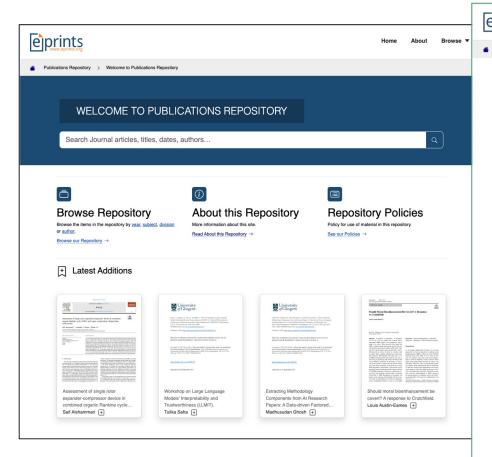






Look and Feel

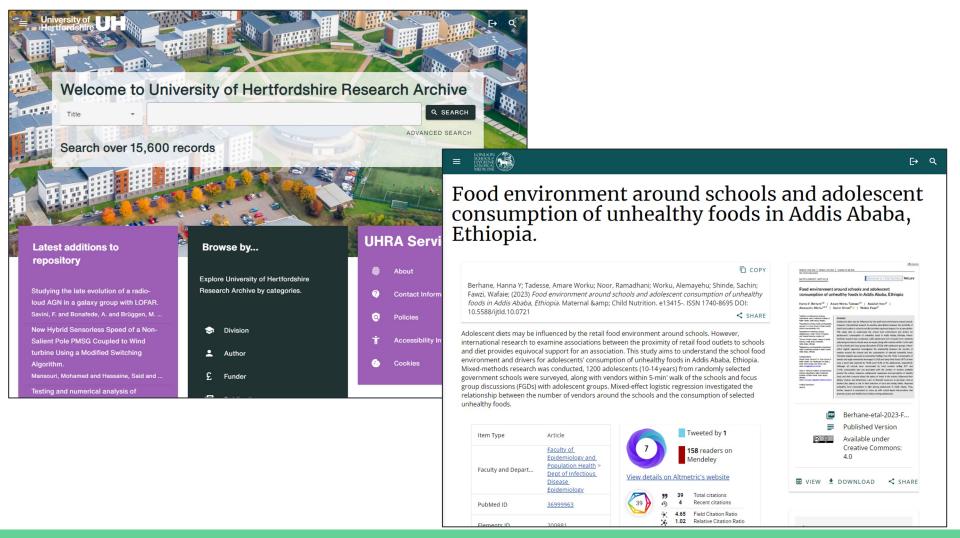
- Fully customisable theming and branding
- Compatible with range of CSS frameworks
- Feeds into institutional web estate staff profile pages
 - Even ones built in WordPress
- New out-of-the-box EPrints themes and templates!

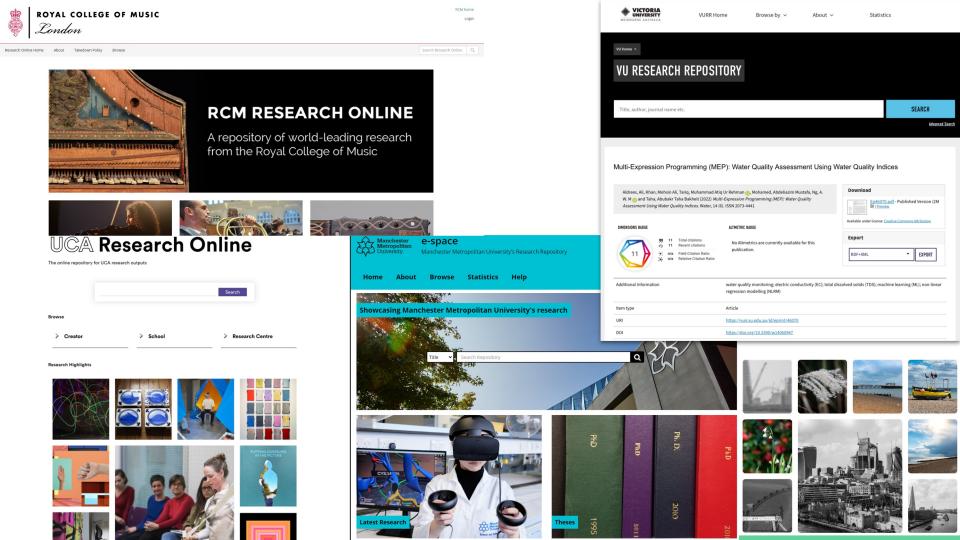


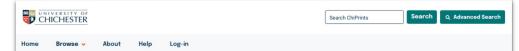


cooling efficiency and exergy efficiency are not affected by the rotor speed.

Learn more about EPrints 3.5 at https://eprints.org/3.5







> Academic Areas > Research Entities > Author

Our Research Areas

Explore our Research Areas



Humanities





Education & Teaching





Psychology & Coun

Performing Arts





Social Work & Practice



Login

Welcome to Royal Academy of Music PRESTO Archive

Search Repository

Search the repository using a full range of fields. Use the search field at the top of the

Browse Repository

Browse the items in the repository by year, subject or author.

Latest Additions

Howat, Roy



Debussy's Welte roll of La plus que lente.





An Opera about the 'Progress of Music': Charles Burney, Domenico Corri's The Travellers (1806) and the Macartney... Lee, Hayoung Heidi



Singing and music making: physiological responses across early to later stages of dementia.

Walker, Nina



"L'art d'évoquer les minutes heureuses" Miladic and Memory in the danie terrily

Emily Kilpatrick

"L'art d'évoquer les minutes heureuses"

Mélodie and Memory in the Année terrible.

Advanced Dementia.

Engagement (VASE) for people with

Debussy's Welte rolls.

@emeraldinsight

Development of a Video Analysis Scale of Co-creativity, well-being and agency: a case Co-creativity: possibilities for using the arts study analysis of a co-creative arts group for with people with a dementia. people with dementia.

Zeilig, Hannah

Zeilig, Hannah

Daniel Lai, L. L.

More..

Kilpatrick, Emily

Royal Academy of Music PRESTO Archive supports OAI 2.0 with a base URL of https://ram.eprints-hosting.org/cgi/oai2

Royal Academy of Music PRESTO Archive is powered by EPrints 3.4 which is developed by the School of Electronics and Computer Science at the University of Southampton, About EPrints I Accessibility









Current State and Direction of Technical Development

- 3.4.6 maintenance release
- 3.5 coming soon reworked user interfaces, slicker browsing and searching, updated out-of-the-box data model, and much more...!
- See more at https://eprints.org/roadmap



Ongoing developments...

- OpenAire
- RIOXX 3
- Hidden documents
- Modelling impact stories



Future Work...

- Further Arkivum v6 integration
- Continued ORCID and DataCite development ensuring reliable and thorough PID integration
- COAR Notify
- RAiDs
- Rules, Compliance & Validation Framework
- Enhanced compliance reporting across a range of framework
- Multi-language support for metadata



Sources of EPrints magic

- Main codebase: github.com/eprints
- User group: github.com/eprintsug
- Download: files.eprints.org
- Wiki: wiki.eprints.org
- Service Providers:
 - eprints.org
 - cosector.com/platform#eprints-hosting
 - openjournaltheme.com/eprints-repository
- Demos & Examples: eprints.org/tryme & eprints.org/examples



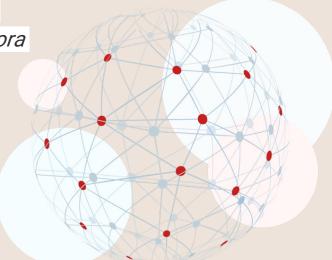


Fedora

What is **Fedora**™

Flexible, Extensible, Digital Object Repository (FEDORA)

"Fedora is a sustainable open-source repository system underpinning many types of digital libraries and archives. It's designed to store, manage, and preserve digital content. Fedora is particularly well-suited for institutions such as libraries, universities, research centres, and museums that need to preserve and provide access to their digital collections."

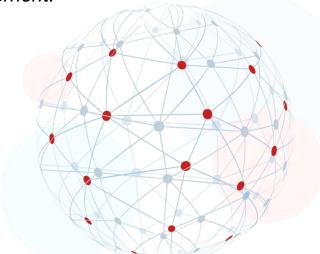


Why **Fedora**™

Features

- Built in Preservation Support
 - Oxford Common File Layout (OCFL)
- Standards-Based Services
- File Format Flexibility
- Diverse Storage Options
- Extensible Architecture
- Native Linked Data Support

"Fedora has a strong history of support and stability, and a vibrant community of loyal developers, implementers, and end users who work together on continuous improvement."





Our Community

Celebrating 20+ years since beta release of Fedora v1

Current version: Fedora 6.5.0 (released in Feb 2024)

State of the Community Today:

- ~300+ known installations around the world
- Wide range of institutional types, sizes & use cases
- Active Committer/Tech Team
- Technical and non-technical engagement opportunities
- Return to in-person conferences and events

Fedora users across the globe 389



Features & Benefits of 6.x

Features

- Enhanced digital preservation capabilities
 - Adoption of the Oxford Common File Layout as a persistence layer
- Extensive migration tooling & documentation thanks to an IMLS Grant
- Built-in simple search
- Real-time performance monitoring
 - Via Grafana & Prometheus integrations
- Metrics Collecting

Benefits

- Enhanced digital preservation & support
- Improved performance & scale
- Easier and fewer migrations



Meet the Team

Fedora Program Team

Arran Griffith

- Program Manager

Dan Field

Fedora Developer (PT)

Fedora Governance Group

Elected representatives from Fedora member institutions, plus 2 community representatives

Active Committers/Tech Team

Ben Pennell (UNC Chapel Hill)
Jared Whiklo (UManitoba)
Demian Katz (Villanova)
Doron Shalvi (NLM)
Calvin Xu (NLM)
Thomas Bernhart (DocuTeam)
James Alexander
(The Open University, UK)

→ Your name here...



Users on Fedora 6.x

Currently Using Fedora 6.x

- Acadia University
- Berklee College
- Berlin State Library
- Bodleian Libraries, Oxford University
- Cambridge University
- DocuTeam Clients
- Institute of Archaeology of the CAS Prague
- Ludwig-Maximilians-Universität München
- Saxon Academy of Science and Humanities
- Slovenian Public Archival Service
- The Ohio State University
- The University of California, Davis
- University of Las Vegas Nevada

- University of Texas at Austin
- Vasser College
- Villanova University
- Whitman College

Migrating to Fedora 6.x

- Emory University
- Érudit
- Lehigh University
- Texas A&M University
- The Open University, UK
- The National Library of Wales
- The National Library of Medicine



Year in Review

Virtual Fedora Showcase (Sept 2023)



- Over 100 registrants over the 3 days of programming
- Sparked engagement opportunities for User Groups
- Great learning and networking opportunity

Fedora Technology Survey (Dec 2023 - March 2024)



 Polled the community about how to prioritize efforts in the next 18-24 months

Cross Community Involvement

- Hyrax Fedora 6 User Group
- Islandora LAC-GLAM Group
- OCFL-java Implementers Group

User Group Gatherings

- Open Repositories (EU/UK Users)
- Texas Conference on Digital Libraries

Fedora Roadmap

 Built based on results of Tech Survey and with input from Tech/Committer Teams

Projects on the Go

Fedora 6.x Performance Testing

- Collaboration with Texas Advanced Computing Center at the University of Texas at Austin
- Understanding performance impacts in an HPC environment

Dependency Updates

Outdated dependencies within fcrepo core that will require updating

Documentation Upgrades

 Re-organizing current documentation to make navigation, search and accessibility more simplified

Fedora Community Roadmap

Program Priorities and Focus Areas

ON-GOING FFFORTS

- Migration Support Bringing all users to Fedora 6.x
- Hyrax Fedora 6 Working Group Investigating pathways to Fedora 6.x for Samvera repository platform users.

......



FEDORA 6.X PERFORMANCE IMPROVEMENTS

Establish performance benchmarking framework for on-going performance testing and improvements

- · Optimizations/bug fixes identified through testing
- Recommendations on infrastructure, architecture and hardware to optimize performance in High Performance Computing (HPC) environment
- · Performance infrastructure and baseline benchmarks for common use cases
- · Accurate performance and cost metrics on speed and scalability of Fedora 6.x



CRITICAL SECURITY AND STABILITY UPDATES

Security and stability improvements will be addressed for the following:

- SpringJava
- PostgresActiveMQ
- Velocity
 - Jettv



DOCUMENTATION UPDATES & IMPROVEMENTS

LEARN MORE AT: HTTPS://BIT.LY/43ZKMOU

Resources

Fedora Website: https://fedora.lyrasis.org/ (getting a facelift in the coming months)

Fedora Wiki: https://wiki.lyrasis.org/display/FF/Fedora+Repository+Home

• Fedora Community Roadmap: https://wiki.lyrasis.org/display/FF/Fedora+Community+Roadmap

Fedora YouTube Channel: https://www.youtube.com/channel/UCE wcxwlQCvVZkbKtC0ej3g

Download Fedora 6.x: https://github.com/fcrepo/fcrepo

Fcrepo Extras (migration tooling, camel toolbox etc): https://github.com/fcrepo-exts

Fedora Slack

- Recommended channels #General, #Tech, #Migration
- Join here: https://forms.gle/MjgCvNch7eaRoXRb8

Consider becoming a Fedora Member

Your contribution allows us to continue development efforts and to support our global community.

InvenioRDM

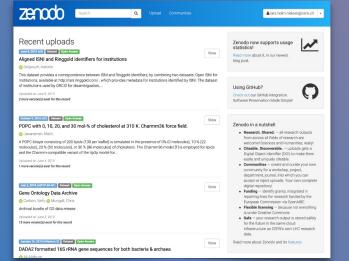
INVENIORDM

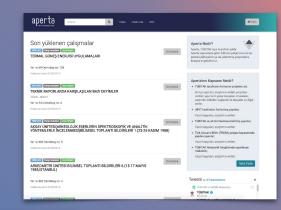
Alex Ioannidis, CERN Open Repositories 2024



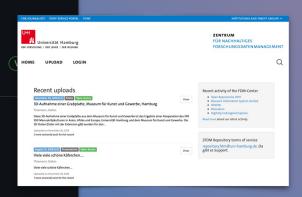
How it started...









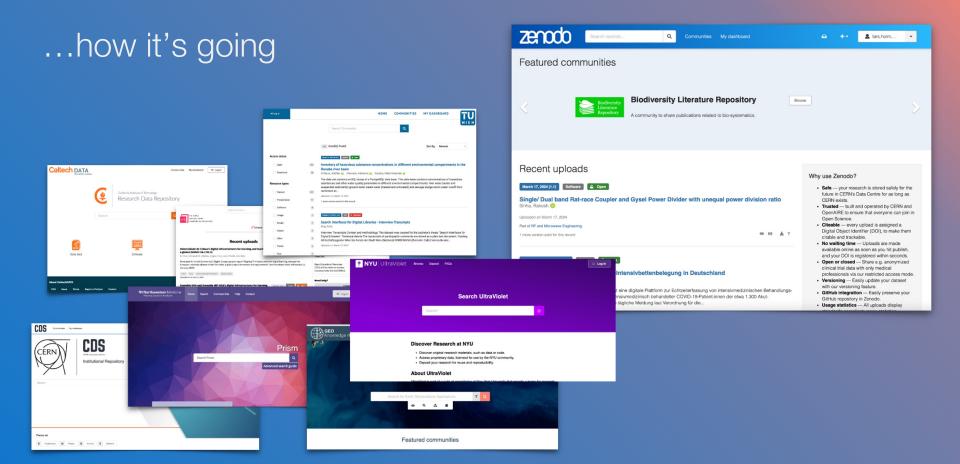




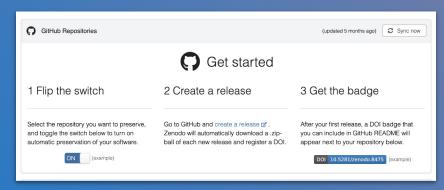
How it started...









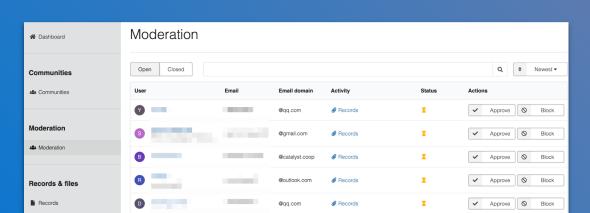


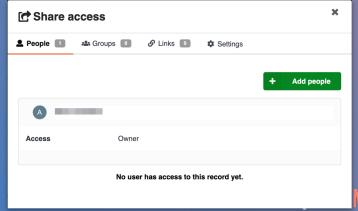




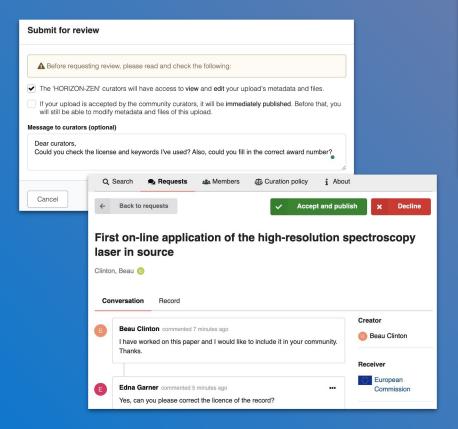
v12.0 Release Candidate

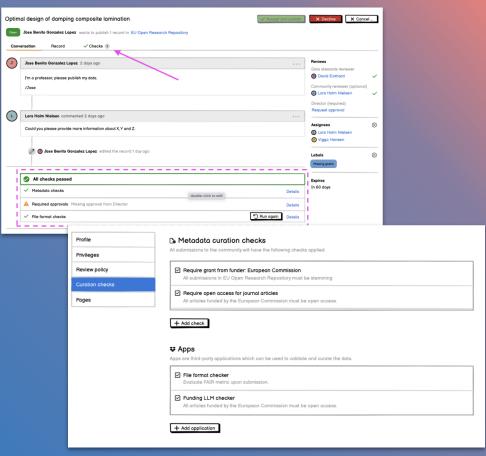






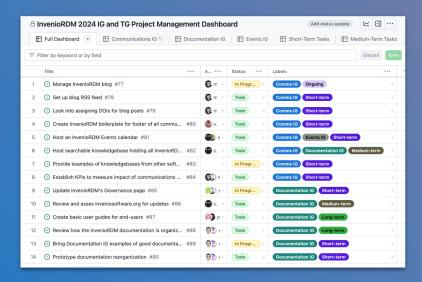
Collaboration. Made easy.

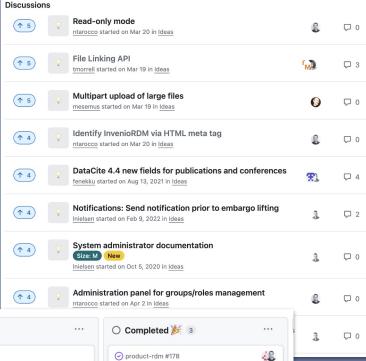


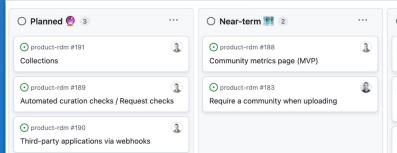


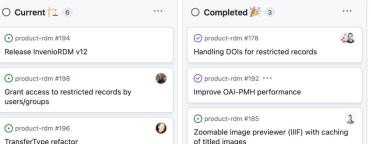


What's next?









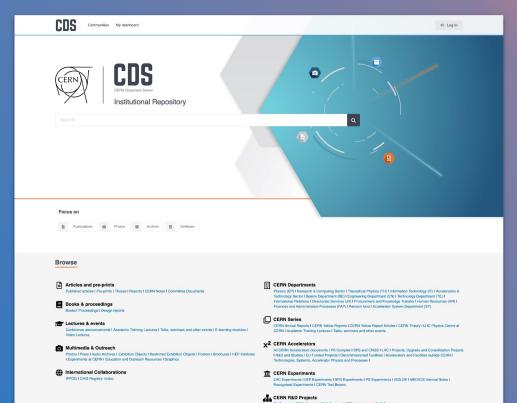


What's next?





INVENIORDM





How do we achieve that?







Join InvenioRDM!

- Website Governance Collaboration https://inveniosoftware.org
- Documentation
 https://inveniordm.docs.cern.ch
- Try it out! https://inveniordm.web.cern.ch
- Chat with us on Discord https://discord.gg/8qatqBC

















































Islandora

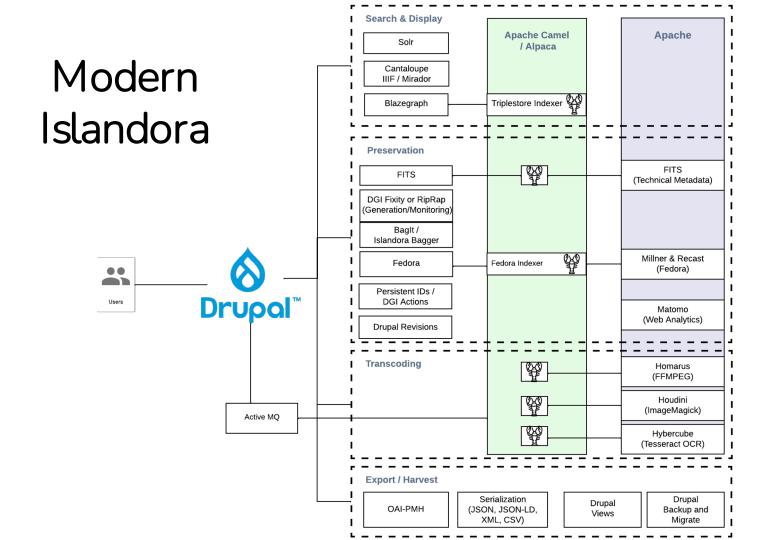


Open-source, Digital Asset Management

Collaboratively author, manage, discover, and steward digital collections across data types and knowledge domains.



https://www.islandora.ca/







Content Modeling and Serialization

- Leveraging Drupal's native capabilities for structured data (nodes, fields, taxonomies, and media)
- Model any metadata standard and the Starter Site comes configured with a rich collection metadata fields and taxonomies.
- Serialize your data in multiple formats including XML (eg. MODS, DC), JSON, JSON-LD, and CSV.

Administrative/System Fields ¶

Title		
The name given to the resource. Title is a system field and as such its configurations can not be adjusted. It is the only field that is required by the system.		
Machine Name	title	
Drupal Field Type	Text (plain)	
Required	yes	
Maximum Length	255 (character length is set by system but can be changed with a contrib module, see below)	
Repeatable	no	
CSL Citation Mapping	title	
RDF Mapping	dcterms:title	
XPath MODS	mods/titleInfo/title	
Transformation	To create a single string out of the subelements of <titleinfo>, we suggest to use the <titleinfo> section of the LOC MODS-DC transform https://www.loc.gov/standards/mods/v3/MODS3-5_DC_XSLT1-0.xsl. In words, it says to: • Take the value of <nonsort> add a crease.</nonsort></titleinfo></titleinfo>	

The Starter Site "<u>Data Dictionary</u>" describes the default metadata configuration for repository items.

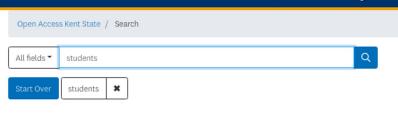


Open Access Kent State (OAKS)



10 items per page ▼

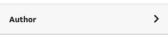
About - Browse - Login



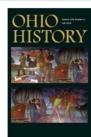
Limit your search











Ohio History Fall 2018

OHIO HISTORY

Contents for Volume 125, Number 2, Fall 2018

- · Contributors 4 Editor's Note 5
- · Pioneers and Land on the Ohio Frontier
- Mansel G. Blackford 7

Sort by Random ▼

- · Forgotten Frontiersman of the Ohio Valley: Simon Kenton's Early Years
- + Read More

Author: Mansel G. Blackford Arthur Andrew Savery

> Samuel R. Phillips David Simonelli Charles Lester Michael A. Beverly

Steven High Robert Colby Kevin F. Kern

Format:

Text



Ohio History Spring 2021

OHIO HISTORY

Home / Doris Collection





Genre

• photographs (15)

Date Created

From To Refine

Geographic Subject

• Scarborough (15)

Doris Collection

Displaying 1 - 15 of 15



Scarborough Bluffs at Bellamy Ravine



Scarborough Bluffs at Bellamy Ravine



Scarborough Bluffs at Bellamy Ravine





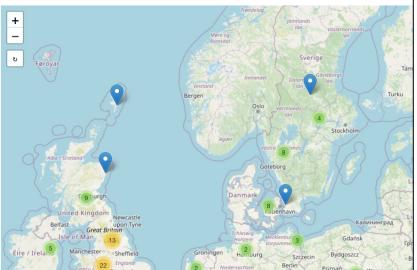


Harley Spiller Menu Collection

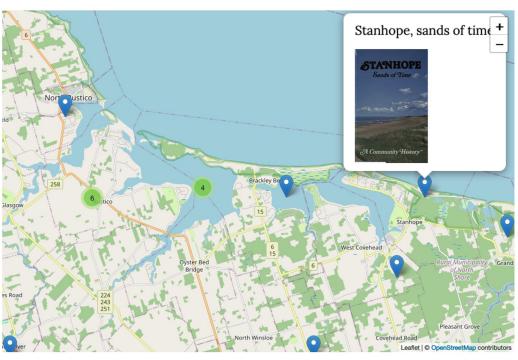
Home / Browse Menus by Geographical Location

Browse Menus by Geographical Location

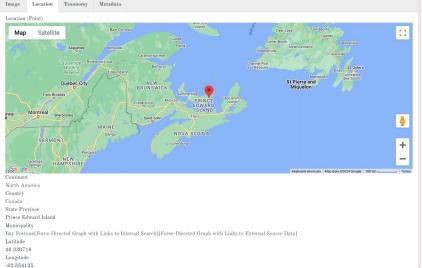
Use this map to browse menus by geographical location. Note that not all menus will have complete geographical represent the collection. Please provide any feedback to dsu.utsc@utoronto.ca.



IslandLives Map

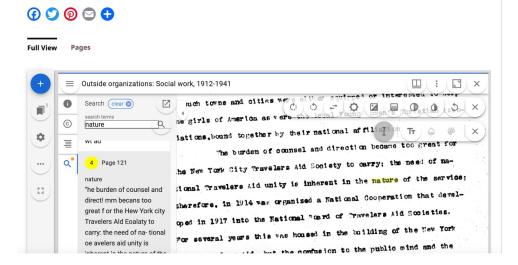






Home > Smith College Special Collections > Sophia Smith Collection of Women's History > YWCA of the U.S.A. microfilm records > Outside organizations: Social work. 1912-1941

Outside organizations: Social work, 1912-1941







U of T Scarborough Library Digital Collections

Home

Housing CSV for Scarborough 2016



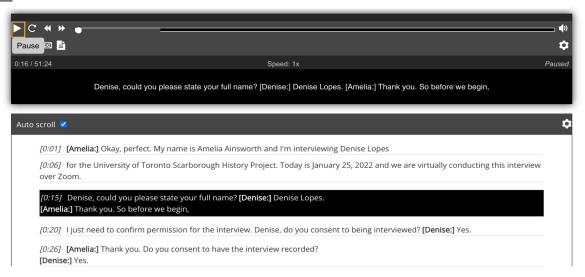
Media Display

GEO_UID CMA_code \$ CMA_name **♦** Census_Tract_name \$ Occupied_privat 535 0.0 2135910 5350000.0 Toronto 5350001.0 535 1.0 250 Toronto 5350002.0 535 2.0 270 Toronto 260 5350003.0 535 Toronto 3.0 535 3455 5350004.0 Toronto 5.0 3420 5350005.0 535 Toronto

Home / ... / Stories of UTSC

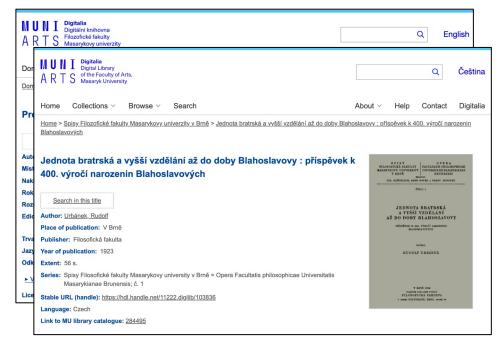
Interview with Denise Lopes





Multilingual Support

அடையாளம்காட்டி		61220/utsc8187
உள் அன இல் ஆக் மெ வன அள அள	Identifier	61220/utsc8187
	Local Identifier	8-2.27
	Member of	S., JV. Chelvanayakam Collection
	Description	Series 2. Professional material, Subseries 1. Professional files - 8-2 [ITAK branches - Upcountry]
	Date Created	1964-01-21
	Language	<u>Tamil</u>
	Genre	letters
	Resource Type	<u>text</u>
உரி	Extent	1 item
	Related Archival Item	Series 2. Professional material, Subseries 1. Professional files - 8-2 [ITAK branches - Upcountry]
	Rights	Digital content found in the UTSC Library's Digital Collections are meant for research and private study used in compliance with copyright legislation. Access to digital, and the technical capacity to download or copy it, does not imply permission to re-use. Prior written permission to publish, or otherwise use content, must be obtained from the copyright holder. Please contact the UTSC Library for further information.





Recent Community Projects

Starter Site

- Ready-to-customize Drupal site that shows off Islandora's features, hooked to module versions.
- Can be deployed using the Islandora Playbook, ISLE-DC, or ISLE Site Template.
- Available with demo content at https://sandbox.islandora.ca/

Islandora Workbench

- A command-line tool that allows creation, updating, and deletion of Islandora content from CSV data.
- Doesn't have to run on the Islandora server
- https://github.com/mjordan/island ora_workbench
- Documentation



Recent Community Achievements

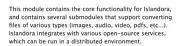
Download & Extend

Drupal Core Distributions Modules Themes General projects

Islandora

View Version control

Islandora is an open-source digital asset management framework designed to help institutions and organizations and their audiences collaboratively manage and discover digital assets using a best-practices framework. Islandora was originally developed by the University of Prince Edward Island's Robertson Library, but is now implemented and contributed to by an ever-growing international community.



Development for this module is done on Github.com so please contribute PRs there rather than MRs here.

To get a starter site that includes Islandora along with its external-to-Drupal ecosystem, visit our Download page.



islandora

These modules are used in the Islandora Ecosystem. This list is a combination of Islandora Foundation sponsored and stewarded tools, and popular third-party contributed modules often used by Islandora users.

Citation Features

Related Modules

- Citation Select: Adds a block that allows users to select and view citations of a node object from a list of citation styles.
- Islandora CSL: This Drupal module creates a psuedo field field_islandora_csl on a node of type islandora object

Media

- Advanced Queue Runner: This module provides a way to run Advance as daemon without manually using a Drush command or running a ground transfer of the provider of the pro
- Media Thumbnails Jp2: This module uses the Media Thumbnails fra entity thumbnails for ip2 files.
- Media Thumbnails TIFF: This module uses the Media Thumbnails fra entity thumbnails for tiff files.

Viewers

- Islandora Mirador: a Drupal module that wraps around the Mirador
- OpenSeadragon: Provides a field formatter to display images using viewer, which is compatible with a IIIF server such as Cantaloupe.
- List Archive Contents Field Formatter: Field formatter for archive file contents as a list.

Project information

Module categories: Content Editing Experience, Integrations, Media

Ecosystem: Islandora

122 sites report using this module

Created by kirsta on 21 September 2012, updated 2 May 2024

Drupal 10 is here!

The Islandora module is Drupal 10 compatible as of the 2.8.2 release.

Stable releases for this project are covered by the security advisory policy. Look for the shield icon below.

Releases

2.12.0 released 1 May 2024

Works with Drupal: ^9 || ^10

✓ Recommended by the project's maintainer.

hOCR feature added to IIIF Views plugin. Also various bug fixes and CI test improvements

Install: composer require 'drupal/islandora:^2.12'

https://www.drupal.org/project/islandora

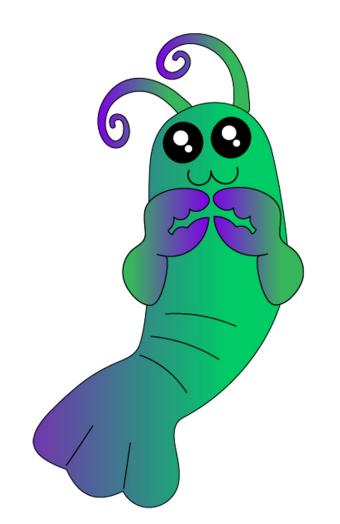
Community Channels

- Multi-channel Slack
- Google Group
- Weekly open tech call/Monthly open meetings
- Community Wiki
- Islandora GitHub organizations
- Learn more on islandora.ca



Stay tuned...

- Monthly Open Meetings
- Islandora Camp Fall 2024
- IslandoraCon 2025



Paislandora

Thank you!

Visit Us At: www.islandora.ca

Ask questions on Google Groups, in Slack, or email us at community@islandora.ca

Samvera



Samvera Community & Technologies

Open Repositories 2024



Samvera is a Community



Samvera is for collaboration

We're growing open technologies, together

Building what we need

The community uses and maintains multiple repository components, stacks and solutions





One Community, Many Repositories

Digital Collections Research Data Management Theses & Dissertations **Media Collections Self Hosted** Vendor Hosted Consortium Hosted





Flexible and Customizable

Large Universities

Small Colleges

Consortia

Governmental Agencies

Small Research Labs

National Digital Repositories

Statewide Digital Libraries









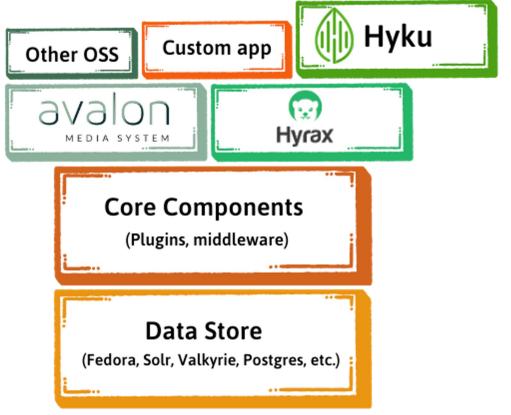
Serverless IIIF





CloverIIIF

The Stack





Whats new

Just shipped v7.7.2 >

MEDIA MANAGEMENT FOR TEACHING, LEARNING AND RESEARCH

Avalon Media System is a free, open source software system for managing and providing access to large collections of digital audio and video.

Documentation

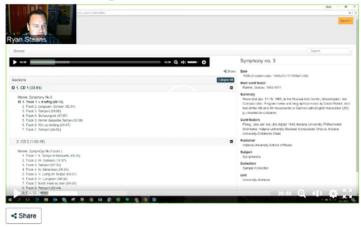
View on GitHub →



Search this site

Browse Collections

A Walkthrough of Avalon 6.4



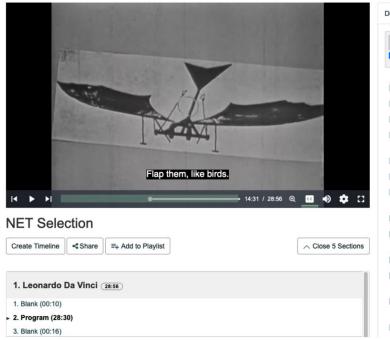
....

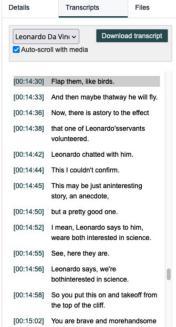
Avalon Media System Project Website

Contact Us

MEDIA SYSTEM

avalon The latest from Avalon





Ramp, a IIIF-based media component library

Demo site: ramp.avalonmediasystem.org

Multiple captions per file

Language can be specified for each caption file

Major accessibility work completed



Whats new

Just shipped hyrax-v5.0.1 >

AN OPEN-SOURCE, SAMVERA-POWERED REPOSITORY FRONT-END

Hyrax is a repository solution that allows deposit of content via configurable workflows; description with customizable metadata; and user-level control over that content.

Documentation

View on GitHub →





The latest from Hyrax

- Content migration example Metadata Only Works
- Before Valkyrie, on Fedora 4:
- 10,000 records per day
- After Valkyrie, on Postgres:
- 200,000 records per day

- Complete integration of the Valkyrie data mapper
- Flexibility to store and preserve files without compromising performance
- Up next: Hyrax 5.1 with migration adapters (Frigg and Freyja) for "lazy migration" to Postgres or Fedora 6 with no downtime





Whats new

Just shipped v5.1.0 >

THE NEXT-GENERATION REPOSITORY SOLUTION

Hyku is a digital repository that provides a robust and flexible platform for institutions to manage, preserve, and provide access to digital content.

Documentation

View on GitHub →



Featured items



RESEARCH REPORT State of the World's Plants and Fungi, 2023.

Antonelli, A.; Fry, C.; Smith, R. J.; Eden, J.; Govaerts, R. H. A. ...



JOURNAL ARTICLE

The global distribution of plants used by humans.

Pironon, S.; Ondo, I.; Diazgranados, M.; Allkin, R.; Baquero, A.



The latest from Hyku

Hyku 5.1:

- IIIF AV gem for audio and video support
- SSO Interface for SAML, OpenID Connect and CAS
- Accessibility improvements

Hyku 6.0 release candidate:

- Knapsack Encapsulates Hyku code and isolates overrides so that upgrades and feature updates are easier
- PDF.js viewer and other improvements from Hyku for Consortia

hyku.samvera.org



Features and Implementations

Getting Started

Presentations and Demos

News

Contact

Learn more →

The Next Generation Repository Solution

Hyku is a digital repository that provides a robust and flexible platform for institutions to manage, preserve, and provide access to digital content.

Get started

<u>Documentation</u> →

CloverIIIF

Whats new

Just shipped v2.8.3 >

SHOWCASE IIIF MANIFESTS AS INTEROPERABLE WEB CONTENT

Extensible IIIF front-end toolkit and Manifest viewer. Accessible. Composable. Open Source.

Documentation

View on GitHub →

Zagna "lunga"









Loft

addition mask, sir scarf to d

ALTERNAT Fava

CREATOR Fava, Ant

DATE 2012

Charles [

12

Serverless IIIF

Whats new

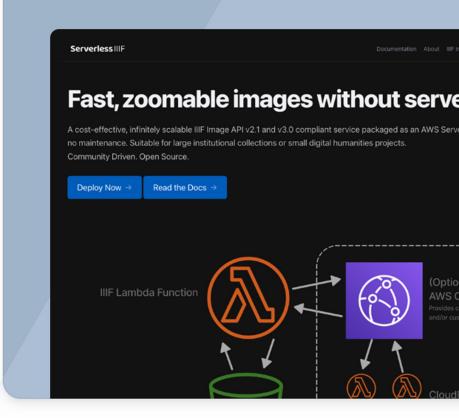
Just shipped v5.0.2 >

FAST, ZOOMABLE IMAGES WITHOUT SERVERS

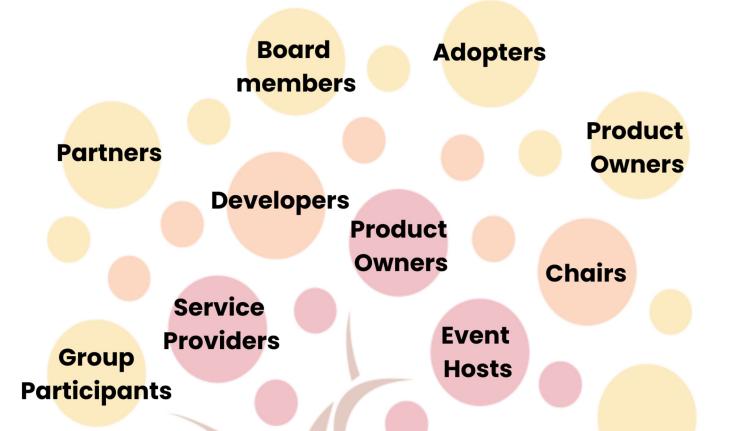
A cost-effective, infinitely scalable IIIF Image API v2.1 and v3.0 compliant service packaged as an AWS Serverless Application with minimum setup and no maintenance. Suitable for large institutional collections or small digital humanities projects.

Documentation

View on GitHub →



Samvera Community

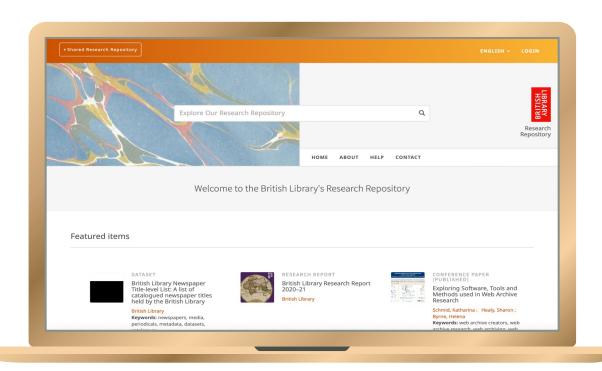




Welcome to our newest Partners



Welcome to our newest Partners





All Are Welcome!

Interest Groups:

Hyku, Hyrax, Avalon Metadata

Outreach & Engagement

Repository Management

Accessibility

Documentation



Goal: confident community contributors



Grow our community development leadership

Community training options

Community Sprints at regional meetings



Samvera Regional Meetings 2024

Midwest Regional Meeting

Indianapolis , IN September 25-26

Northeast Regional Meeting

Princeton, NJ October 14-15

Still to be scheduled:

- West Coast (early October)
- Europe



Samvera Connect **2025**



October 21 - 24 **2025**





Thank You!

heather@samvera.org



Questions