

# DATA CURATION NETWORK

## Sustaining Open Research Virtual Learning Series

Day 4: Developing and maturing IR technology platforms to  
support data sharing

# Research data repository ecosystem

Developing and maturing IR technology  
platforms to support data sharing

Data Curation Network, 2023-10-11

Amy  
Nurnberger

This work is licensed under a [Creative Commons  
Attribution 4.0 International License](#).

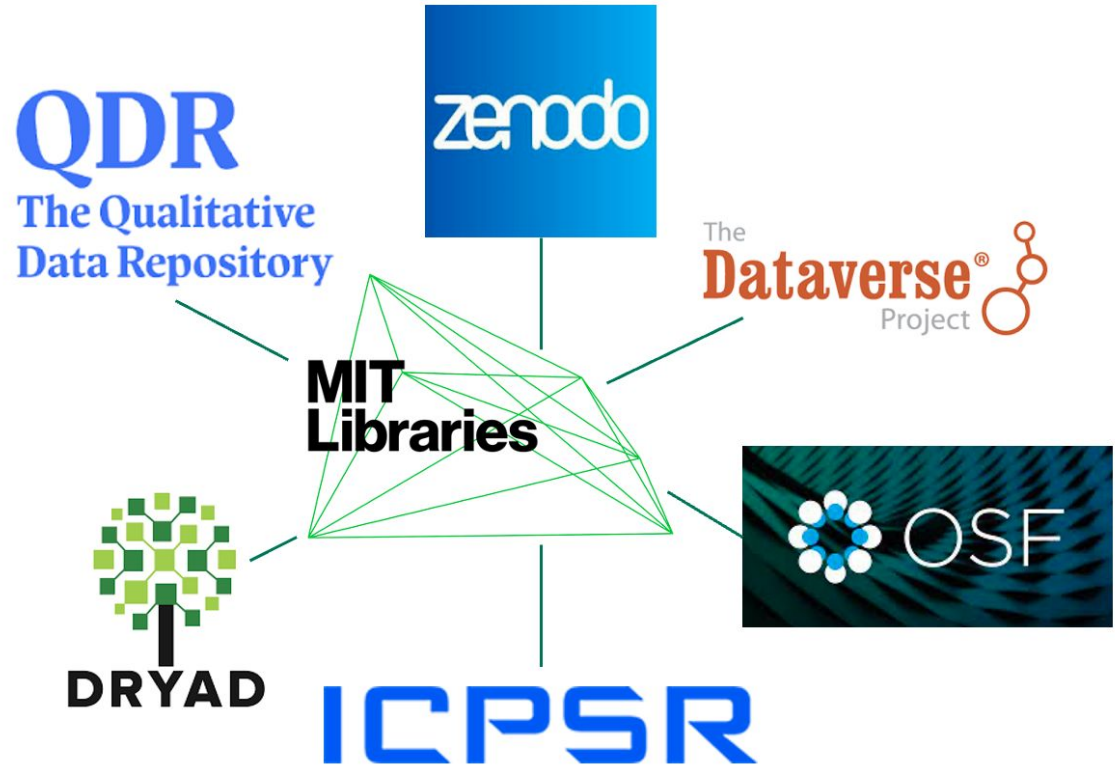


**MIT  
Libraries**

**What to do when  
you don't have  
an institutional  
data repository?**



**What to do when  
you don't have  
an institutional  
data repository?**



# Montana State University

Sara Mannheimer  
Associate Professor and Data Librarian  
Montana State University

## MSU Dataset Search

Access datasets from creators affiliated with Montana State University (MSU)

Q

Search

[Advanced Search](#)

<https://arc.lib.montana.edu/msu-dataset-search/>

### Recent Datasets

2023-01-04

[\*Interspecific trait variability and local soil conditions modulate grassland model community responses to climate\*](#)

Franklin Alongi, Jana R uthers, Justyna Giejsztowt, Katrina LaPaglia, and Anke Jentsch

2022-08-22

[\*Summer 2022 habitat selection in southwest Montana\*](#)

# Strengths and weaknesses of our current technology/strategy

- Strengths:
  - well-known and trusted, sustainable infrastructure, easy to use, requires Readme (includes guidance), great curation service and help service, good admin features (bulk dataset metadata easily available and up to date)
- Weaknesses
  - data size limits (300GB), lack of automated metadata entry, not optimized for search engines, no options for restricted data, cost

# Criteria or questions we would consider if launching a new repository technology or strategy today

- Took a long time to arrive at Dryad as our solution (MSU Data services launched in 2014, we became Dryad members in 2019)
- We would consider:
  - cost
  - ease of use (for both admins and data depositors)
  - data size limits
  - sustainability and community-mindedness
  - levels of access
  - curation services

## Areas of development to more closely align with *desirable characteristics*

Free and Easy Access ✓

Clear Use Guidance ✓

Risk Management ✓

Long-term Organizational Sustainability ✓

Unique Persistent Identifiers ✓

Metadata ✓

Curation and Quality Assurance - curation is lightweight

Broad and Measured Reuse - does provide metrics, could have more info - community-wide question

Common Format - recommended

Provenance - versioning isn't available

Authentication ✓ - submission with orcid

Long-term Technical Sustainability - technical partnership with CDL

Security and Integrity ✓

**Welcome to Prism!**

Upload scholarly work, create communities, get citable links and more. To get the most out of Prism, log in with your NetID and check out our guide.

# Prism

Northwestern University  
Feinberg School of Medicine

Kristi Holmes, PhD  
Director, Galter Health Sciences Library  
kristi.holmes@northwestern.edu

Prism

Search Prism



[Advanced search guide](#)

**INVENIO**RDM

zenodo



**USRN**



## Links!

Prism: <https://prism.northwestern.edu/>

InvenioRDM: <https://inveniosoftware.org/products/rdm/>

Invenio Blog: <https://inveniosoftware.org/blog/>

Invenio documentation: <https://inveniosoftware.org/documentation/>

Zenodo: <https://zenodo.org/>

OR23 workshop: <https://zenodo.org/communities/inveniordm-workshop-2023/>

GREI Community: <https://zenodo.org/communities/grei/>

## Thanks!

- Prism team
- InvenioRDM open source community

## Support: NIH – below and impact of others

- ODSS/Office of the NIH Director pursuant to OTA-21-009, “Generalist Repository Ecosystem Initiative (GREI)” through 10T2DB000013
- Northwestern Clinical and Translational Sciences Institute (NCATS UL1TR001422)

# Sample Use Cases



**Pre-Clinical:** *We're managing a large multi-site project, harmonizing data from numerous sources and managing research projects. We want to create communities of practice to integrate theories, data, techniques, and tools.*

**Population Health and Health Equity:** *Our multi-institution health equity project uses InvenioRDM to collaborate with our community-based partners and credit these partnerships. We can share materials from community health events, project materials, training materials, annual reports, and lay summaries of research. InvenioRDM helps us to be better partners, accountable to collaborators and the community. (ChicagoCHEC)*

**Clinical Trials:** *I am a clinical researcher. I need a way to pre-register protocols or research proposals, search on demographics of participants in similar studies, get insights into recruitment, and share portions of study for compliance. I also want an easy way to share materials such as recruitment protocols, and lay summaries of the trials with the community in an organized way.*

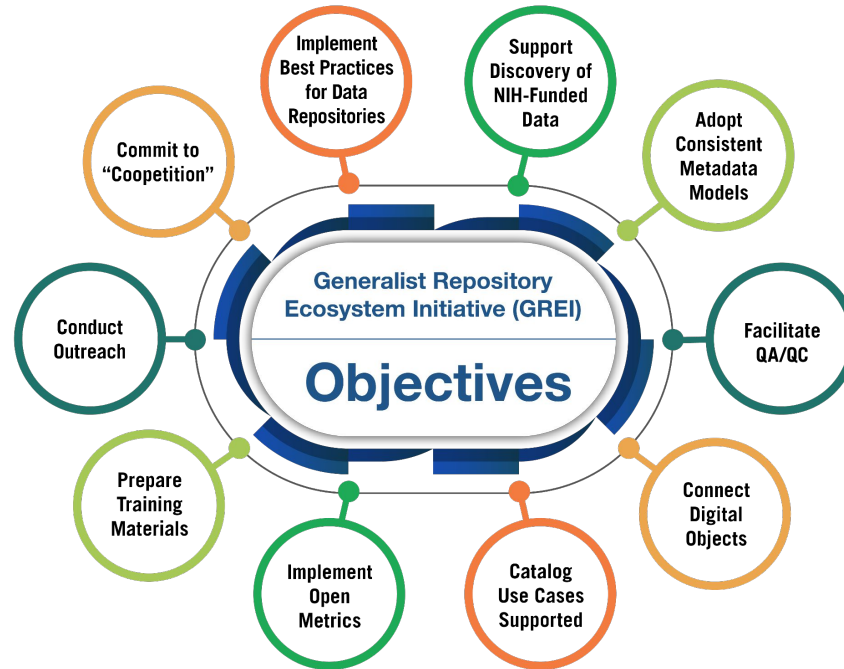
**Early-Stage Investigators:** *I'm an early career researcher just getting started on my research career. I want to "put my best foot forward" to showcase my work and demonstrate my expertise and collaborations. Our repository gives me a way to make all my research efforts findable and the metrics are helpful for reporting to leadership.*

**Basic Science:** *I lead a large basic science research group. We use the repository to support reproducible science by packaging our data and methods in a combined way. Everything gets a unique identifier and versioning is supported. Our lab prioritizes science communication, so the lab graduate students set up a collection of lay summaries of their research projects to enhance engagement and dissemination, tweeting back to the summaries.*

**Translation to Practice:** *My team wants to find out about clinical trial opportunities to better offer patients all options for treatment. It is important to us to openly share the latest research with patients. The InvenioRDM communities give us a way to make these materials openly, packaged in a cohesive and attractive manner. As resources are updated, we can upload the new versions and track access.*

**Dissemination:** *Our institute wants a way to publish and disseminate content like handbooks, lay summaries, and more. We want to credit all contributors and produce an attractive and interactive resource that can be easily updated.*

# Generalist Repository Ecosystem Initiative (GREI)



<https://datascience.nih.gov/data-ecosystem/exploring-a-generalist-repository-for-nih-funded-data>