# Research Data Management and Sharing and Research Visibility

Sara Mannheimer, Associate Professor, Data Librarian Jason Clark, Professor, Head of Digital Library Initiatives



### Outline

- Data management planning (Sara)
- Storing, archiving, and sharing data (Sara)
- Metadata basics (Sara)
- Benefits of metadata and identifiers (Jason)
- Research(er) visibility (Jason)



## Data are...



## Data management & sharing requirements

- 2013 OSTP memo required funding agencies with \$100 million in expenditures to require researchers to make data public
- 2022 OSTP memo expands requirement to all funding agencies
  - Notably, NIH will implement an updated data management and sharing policy, starting in January 2023.
    - NIH data sharing resources | NIH policy FAQs



### Expectations for Data and Research sharing and dissemination

National Endowment for the Humanities (NEH)

• <u>Dissemination of Project Results</u> [grant policy]

National Science Foundation (NSF)

<u>Dissemination and Sharing of Research Results</u> [grant policy]



### Expectations for Data and Research sharing and dissemination

Private funders

Gates Foundation, Mellon

Journals are requiring data sharing

PLOS, JAMA



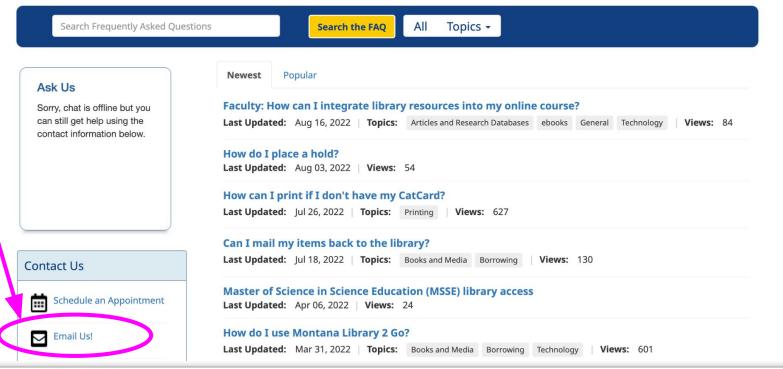
# Data management planning



### Ask the Library

Have a question? Find an answer!

#### ask.lib.montana.edu





#### **MSU Library**

Find

Request

Services

Spaces

People

Ask the Library

#### **Search Library**

GO

#### Quicklinks

Articles & Research Databases

Journal by Title

Archives and Special Collections

ScholarWorks

About Us

#### **MSU Library**

P.O. Box 173320

Montana State University Library / Services / Data Management Resources

### Data Management Resources <u>lib.montana.edu/services/data</u>









Data Management Planning Toolkit

Data Publication & Discovery

File Storage (UIT)

Data Transfer &
HighPerformance
Computing (RCi)

The Library can help you create and implement a data management strategy.

When ur research data is ready to be published, we can facilitate data publishing and promote discovery.

Contact the Data Librarian

Schedule a Data Consultation



### **FAIR Principles**

Findable

Accessible

Interoperable

Reusable

# www.go-fair.org/fair-principles

#### **MSU Library**

Find

Request

Services

Spaces

People

Ask the Library

#### **Search Library**

GO

#### Quicklinks

Articles & Research Databases Journal by Title

**Archives and Special Collections** 

ScholarWorks

About Us

#### **MSU Library**

P.O. Box 173320

Montana State University Library / Services / Data Management Resources

### Data Management Resources <u>lib.montana.edu/services/data</u>



Data Management **Planning Toolkit** 



**Data Publication** & Discovery



File Storage (UIT)



**Data Transfer &** High-Performance Computing (RCi)

#### The Library can help you create and implement a data management strategy.

When your research data is ready to be published, we can facilitate data publishing and promote discovery.

Contact the Data Librarian

Schedule a Data Consultation



### **Data Management Planning Toolkit**

The resources below can help you write data management plans to comply with funder requirements. For basic guidance on data management, we recommend reading Good Enough Practices for Scientific Computing (Wilson et al., 2016). Or contact us for one-on-one help.

★ Data Management on-demand workshops

### lib.montana.edu/services/data/toolkit



### Data Management **Plan Template**

Shape and structure your data management plan.



### Cut-and-Paste Language

Use in your data management plan.



#### Metadata Guidance

Help your data be machine readable and understood by others.



#### DMP Tool

Create data management plans that meet institutional and funder requirements. Review public DMPs. **DMP Tool Tutorial from** 

ASU



<b>DMP</b> Tool	Sara Mannheimer ▼ Language ▼			
Build your Data Management Plan My Dashboard Create Pla	n Funder Requirements Public DMPs Help			
Montana State University (MSU) (montana.edu)				
Create a new plan				
Before you get started, we need some information about your research project to set you up with the best DMP template for your needs.				
* What research project are you planning?				
☐ mock project for testing, practice, or educational purposes				
* Calaat the maintain and account agreemination				
* Select the primary research organization  Research organization				
Montana State University (MSU) (montana.edu)	<ul> <li>or -</li></ul>			
montana state oniversity (mso) (montana.eda)				
* Select the primary funding organization				
Funder	- or -   No funder associated with this plan or my funder is not listed			
Begin typing to see a list of suggestions.				
Create plan Cancel				





#### **MSU Library**

Find

Request

Services

Spaces

People

Ask the Library

#### **Search Library**

GO

#### Quicklinks

Articles & Research Databases

Journal by Title

**Archives and Special Collections** 

ScholarWorks

About Us

#### **MSU Library**

P.O. Box 173320 Centennial Mall Bozeman, MT 59717-3320 Montana State University Library / Services / Data Management Resources / Data Management Planning Toolkit / Cut-and-paste Language

### **Cut-and-Paste Language**

Provided by MSU Library Data Services

Last updated: 21 February 2021

Storage, backup, & security

Dryad

Zenodo

**ICPSR** 

Policies for Reuse

**MSU Data Policies** 

Storage, Backup, and Security 🔺

#### Microsoft OneDrive/SharePoint

MSU researchers and affiliates have access to the Microsoft OneDrive/SharePoint cloud storage system. This service is an institutional resource and fully compliant with all data policies, and it enables collaboration between teams within MSU and across institutions by allowing users to edit, download, and share files. It can also function as a cloud backup system by syncing content between local machines and the cloud. Data stored in OneDrive will be stored in the PI's account and shared as appropriate with other researchers.

Note: Any questions related to the storage of personally identifiable information and applicable security standards may be directed to MSU's Legal Counsel.

#### Knox

Knox is a server managed by the MSU IT Center that utilizes encryption to safeguard student data or other information that MSU has a legal obligation to protect.

#### **Data Transfer**

The MSU Research Cyberinfrastructure Storage Cloud platform offers a Globus Endpoint that researchers can use to



# Data Storage



### Data Storage

UIT provides support for OneDrive and High-Performance Computing.

The Library can connect you to solutions from UIT, or you can reach out to them directly.

### Data Storage - Research CyberInfrastructure (RCI) at MSU

### High Performance Computing

Access to computer processing

### **Storage Clusters**

Access to data storage and transfer

Research CyberInfrastructure (RCI) https://www.montana.edu/uit/rci/

# Data archiving and sharing



### Data Archiving and Sharing - Scenario

Potential ask from a partner:

"I have a series of video recordings that I'd like to share. How can I deidentify this type of material? How would I license and share it as data that can be reused?"

### Data curation

Help with human subjects considerations when sharing data, including deidentification guidance

Guidance on metadata and contextual information

Help with data repository selection and deposit

### **Data Publication and Discovery**

MSU Library recommends that datasets be published in trustworthy data repositories, especially those that are commonly used in your discipline. Our research suggests that publishing data in a disciplinary repository promotes discovery and reuse of research data.

The library provides a search tool for datasets published by MSU researchers. Click the button below to explore MSU research datasets.



### Recommended general repositories



Data publishing is provided to MSU affiliates for free via an MSU Library membership.

Log in with your NetID and password to deposit data

General purpose repository appropriate for any type of data

Non-profit organization

Provides lightweight data curation





#### National Snow and Ice Data Center

a part of CIRES at the University of Colorado Boulder



# The Qualitative Data Repository

Environmental Data Initiative





# Metadata basics





### datadryad.org

**Explore Data** 

About

Help 🔻

Login

#### **ORCID Account**

Dryad uses ORCID as its primary login method. Please use, or create, your ORCID login credentials to login to Dryad.



#### Dryad's Commitment to You

#### Curation

All datasets are curated to ensure they are Findable, Accessible, Interoperable, and Reusable

#### Compliance

Enabling adherence to funder and publisher open data mandates

#### **Community Supported**

Dryad is researcher-led and supported by our institutional and publisher members

Privacy Policy Accessibility Policy Terms of Service

Copyright (c) 2022 Dryad



Follow us on Twitter









#### **Dataset: Basic Information**



Author(s)

First Name *  Sara	Last Name *  Mannheimer	Institutional Affiliation: *  Montana State University Bozeman*	Author Email *
https://orcid.org/0000-	0002-1433-6782		
+ Add Author			

Research Domain \* Research Facility:

**Funding** 

Award Number Granting Organization: \* National Science Foundation N/A

remove



National Science Foundation

National Science Foundation of Sri Lanka

#### **Data Description DRYAD** Keywords: Adding keywords improves the findability of your dataset. E.g. scientific names, method type qualitative qualitative data collected? How has it been processed? qualitative study □ ✓ □ ✓ □ ✓ □ ✓ ✓ − qualitative studies qualitative mapping qualitative research qualitative interviews qualitative metasummary

#### **Related Works**

Are there any preprints, articles, datasets, software packages, or supplemental information that have resulted from or are related to this Data Publication?

**\$** 

Work Type Article

Identifier or external url

example: https://doi.org/10.1594/PANGAEA.726855

remove



### **Upload Your Files**



You may upload data via two mechanisms: directly from your computer, or from a URL on an external server (e.g., Box, Dropbox, AWS, lab server). We do not recommend using Google Drive.

We require that you include a <u>README.md</u> file based on our template in order to provide key information for understanding and reuse of your data.

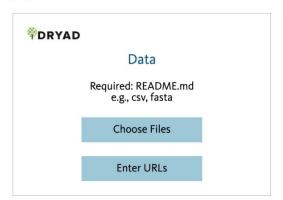
If you prefer, you can edit the Markdown online at hackmd.io.

Open the link (above)

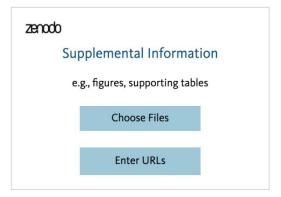
Copy and paste the text into a new note (create new by clicking the plus sign)

When you're done, save into a README.md and upload the file under the Data category

Software and Supplemental Information can be uploaded for publication at Zenodo. You will have the opportunity to choose a separate license for your software on the review page.









### Metadata

Metadata helps future users find and understand published datasets. Please contact the Library to discuss metadata standards in your field.

### Readme files and codebooks

A common way to provide metadata to your research dataset is through a readme file and/or a codebook.

- Readme checklist from Mozilla Science.
- Readme template, adapted from Cornell University.
- · Guide to codebooks from ICPSR.

#### Recommended descriptive elements for your dataset

When publishing data in a repository, we recommend that you provide the following descriptive elements:

- **Title** of the dataset (unique from associated articles)
- **Author** (the person/people who collected the data)

• Institution (Montana State University)

- · Date last modified
- Version, if applicable
- Short description of the dataset



### Metadata

Metadata helps future users find and understand published datasets. Please contact the Library to discuss metadata standards in your field.

### Readme files and codebooks

A common way to provide metadata to your research dataset is through a readme file and/or a codebook.

- Readme checklist from Mozilla Science.
- Readme template, Idapted from Cornell University.
- · Guide to codebooks from ICPSR.

#### Recommended descriptive elements for your dataset

When publishing data in a repository, we recommend that you provide the following descriptive elements:

- **Title** of the dataset (unique from associated articles)
- **Author** (the person/people who collected the data)

• Institution (Montana State University)

- · Date last modified
- Version, if applicable
- Short description of the dataset



### **README Template**

### lib.montana.edu/services/data/toolkit/readme/

Download as a TXT File

```
This DATSETNAMEreadme.txt file was generated on YYYY-MM-DD by NAME
<help text is included in angle brackets, and can be deleted before saving>
GENERAL INFORMATION
1. Title of Dataset:
2. Author Information
        A. Principal Investigator Contact Information
                Name:
                Institution:
                Address:
                Email:
        B. Associate or Co-investigator Contact Information
                Name:
                Institution:
                Address:
                Email:
```



# Metadata + Identifiers **Benefits**

a) Collect and make publicly available appropriate metadata associated with scholarly publications and data resulting from federally funded research, to the extent possible at the time of deposit in a public access repository. Such *metadata should* include at minimum:

- i) all author and co-author names, affiliations, and sources of funding, referencing digital persistent identifiers, as appropriate;
- ii) the date of publication; and, iii) a unique digital persistent identifier for the research output;

Office of Science and Technology Policy (OSTP) Memo on *Ensuring Free, Immediate, and Equitable* **Access to Federally Funded Research** (August 2022)



### Metadata + Identifier Standards

In standardized description and identification, we gain:

- Keywords
- Search/Browse
- Aggregation
- Indexability
- Analysis



### Identifiers

Creating permanent links for your research data.

- Useful in citation
- Useful as a permanent location

A <u>persistent identifier (PID)</u> is a long-lasting reference to a document, file, web page, or other object.



### Identifiers - Types

#### Article Identifiers

DOI, Digital Object Identifier

Researcher Identifiers

ORCID, Open Researcher and Contributor ID

Organizational Identifiers

Dissemination and Sharing of Research Results [grant policy]

#### Identifiers - Why?

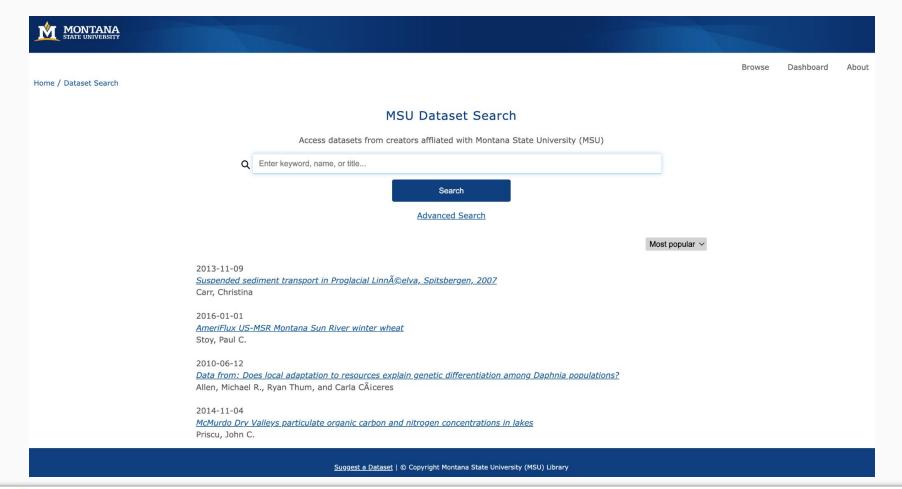
#### CUI BONO?

- Recognition
- New forms of impact and analysis

## Research data visibility -Dataset Search

arc.lib.montana.edu/msu-dataset-search







#### Metadata + Identifiers - Benefits

MSU Dataset Search as an example of this metadata and identifier work

- Index of research data shared by MSU researchers
- Refined metadata for broad findability
- Permanent identifiers built into the system



#### MSU Dataset Dashboard

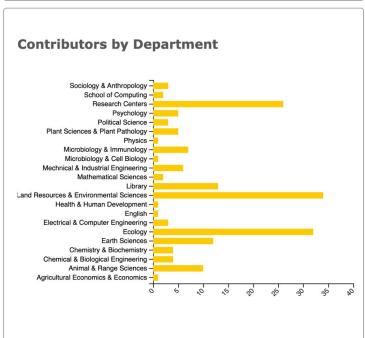
Last updated: 2022-09-27 11:09:52 am

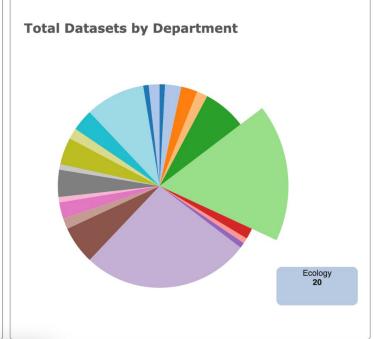
116

total datasets

176

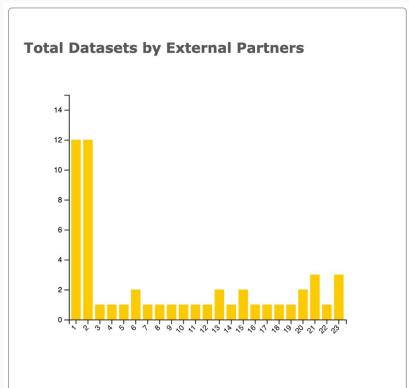
total contributors

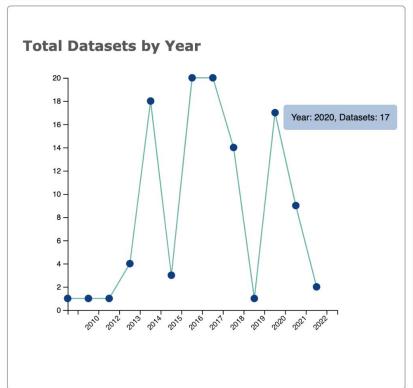






### Analytics







# Research(er) Visibility + Analytics

#### Visibility - Library Services and Expertise

- Consultation work
- Technical SEO and Semantic Web Optimization work
- Analysis and Data Science work

#### Visibility - Consultation

Research and data sharing strategies

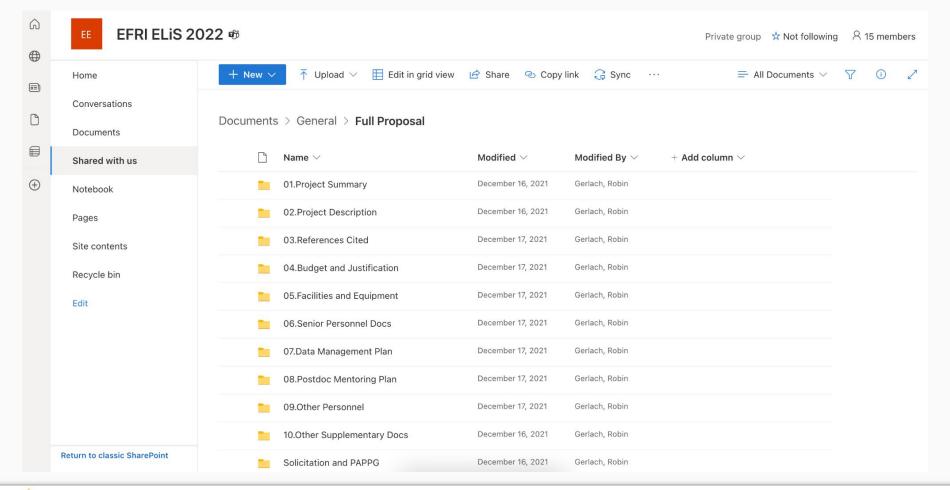
Metadata consults and automation

Data management planning



#### Visibility - Consultation Scenario

- Connect with Project Investigator.
- Work through goals for project data management, sharing, and research dissemination.
- Advise and draft sections of grant narrative.





Our goal for this project is to build a collaborative research team within a scientific focus area that is consistent with the NSF priorities of Engineering Living Systems (ELiS). We have defined four areas in our approach to sharing the project outcomes with the research community and the public at large. Our objectives are to distribute our research results by: Area 1- Disseminating research findings through peer-reviewed publications and presentations; Area 2- Communicating with researchers and the public at large; Area 3- Sharing designs/models of 3-D printed objects (biofilms and structural elements); and Area 4- Sharing method resources.

#### **Publications and scientific presentations**

The scientific community depends on peer reviewed publications to gauge the quality and impact of scientific research. As a team, we have a compelling commitment to maximize the use of our research dollars and share our findings through publications in peer-reviewed journals. Moreover, we have partnered with the research and data service units within the Montana State University Library to ensure best practices in data management, data archiving, and research findability using metadata are in place. Knowing that there is a fine line between balancing the completeness of datasets and analysis with the timeliness of reporting project outcomes, we have developed a series of target performance metrics (Table S1) to guide our production of high-quality, high-impact, and high-volume research throughout the duration of the proposed project. We intend to embrace open access journals (e.g., PLoS journals, Nature publications, Frontiers journals) when submitting our research for peer review publication, so that the scientific community can have immediate and free access to our findings. We will leverage institutional funding for open access journals, along with funds in the budget (included in Other Direct Costs category in budget justification). We will also strive to take advantage of options to publish supplementary materials to increase accessibility of datasets. In addition, our team will make use of other mechanisms to increase access to our publications (i.e., ResearchGate, Data in Brief, Zenodo - c.f., Data Management Plan). Similarly, our team will be very active in sharing our science with the broader research community by presenting our work at local, regional, national, and international meetings. Our expectation is that all

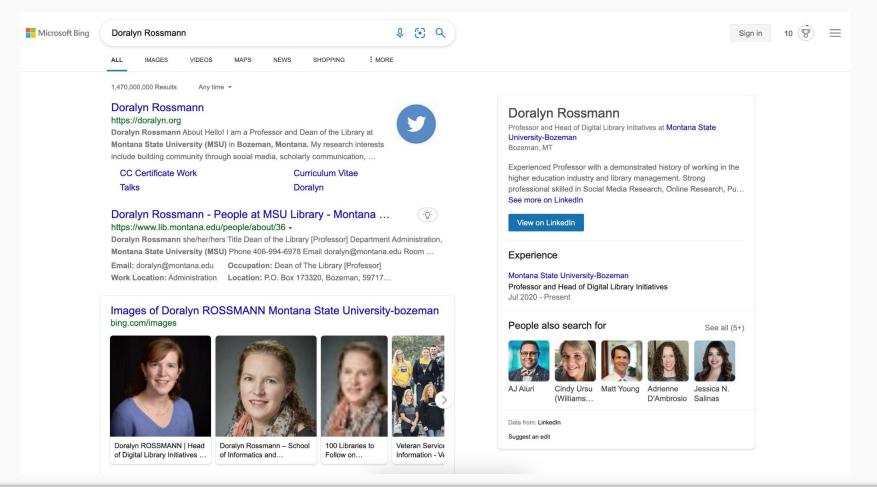


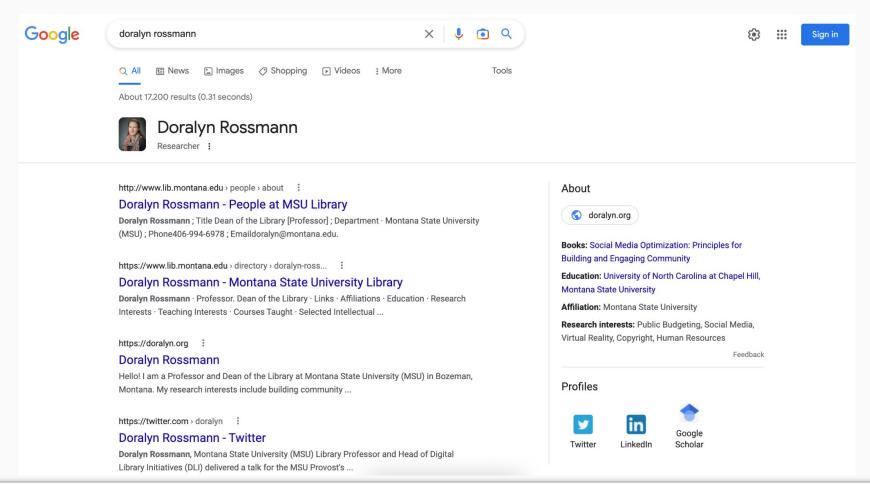
#### Visibility - Optimization

Technical Search Engine + Semantic Web Optimization

- Findability and Aggregation
- Creating your researcher profile and data profile
- Defining your research cohort or a center









#### Visibility - Optimization, how?

- Metadata and Identifier work
- Intentional Website Architectures
- Building Signals for Indexing Agents (crawlers)

#### Visibility - Optimization Techniques

### Technical Search Engine + Semantic Web Optimization

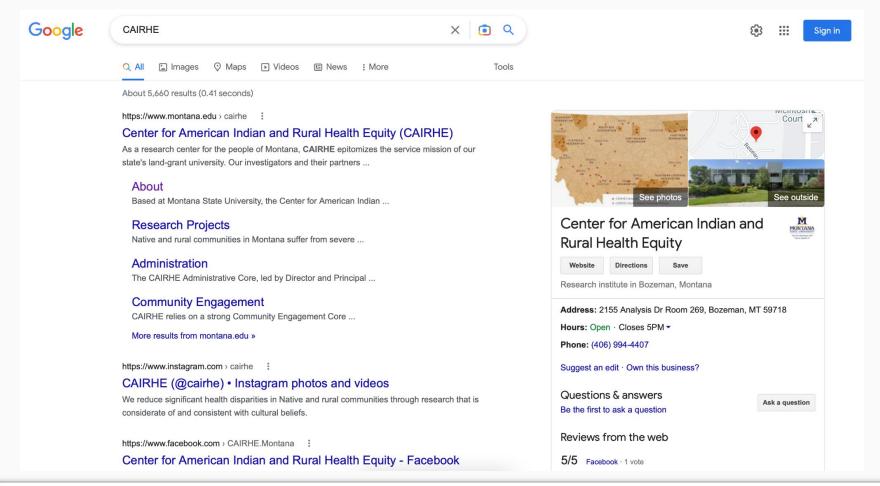
- Research Unit or Cohort
  - Define and Optimize for Findability
  - New Website Architectures and Tagging
  - Create Metadata sources for Indexing (Wikimedia)

#### Visibility - Optimization Scenario

Example: Center for American Indian and Rural Health Equity (CAIRHE)

- Work includes:
  - New Sitemap and Audit of Website Architecture
  - Google Business Registration
  - **Entity Creation** (Wikidata as a new data source)







#### Visibility - Analytics

Defining our institutions and university expertise

Increasing reach and impact of scholarship and data

Insights and data visualization

#### Analytics - Insights and Visualization

Analysis of research patterns and metadata

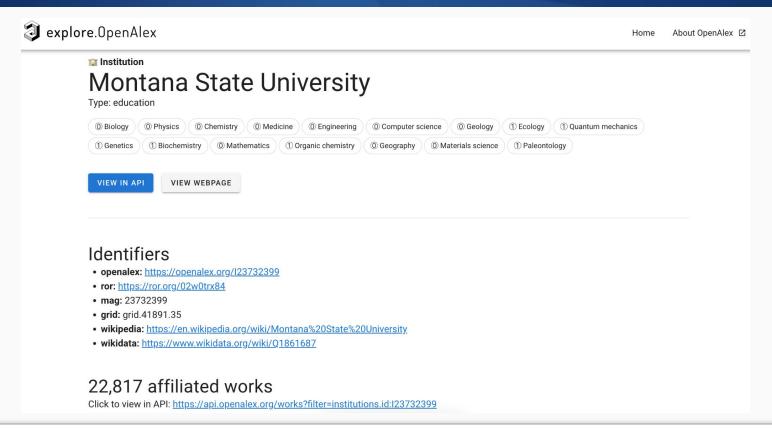
#### Could lead to:

- New insights about how your work is received.
- Community detection to help you find new collaborators.

# Expertise Search [beta]

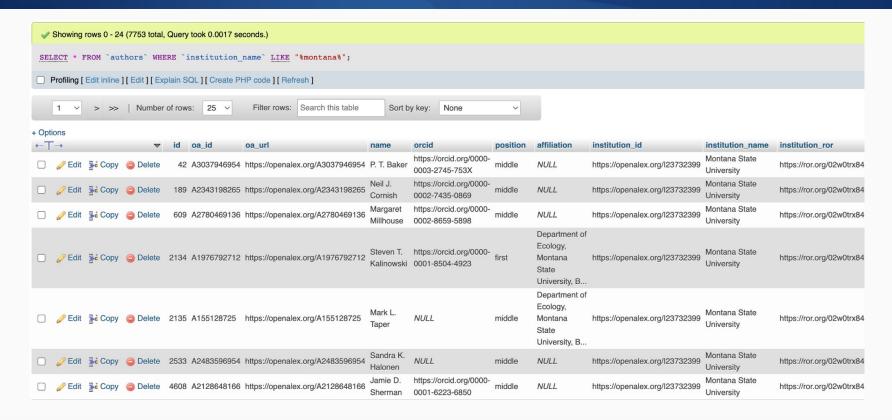


#### Analytics + Insights Scenario





#### Analytics + Insights - Benefits





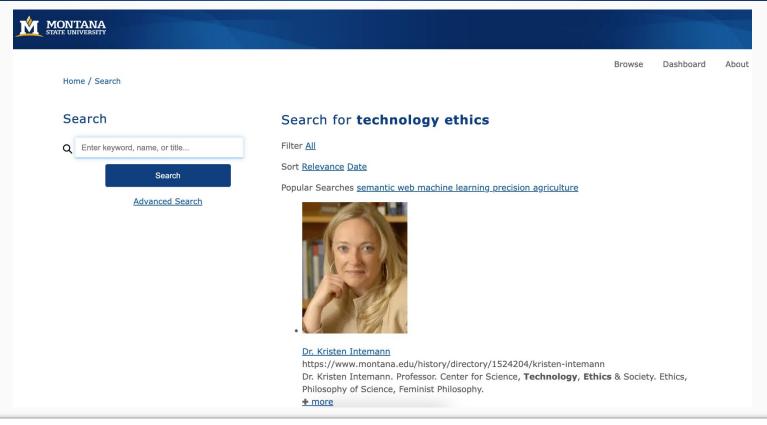
#### Analytics + Insights - Routine

Metadata + Identifier harvest ->

Aggregation and indexing ->

Expertise Search user interface

#### Analytics + Insights - Interface and Visualization





### Summary

The library can partner with you to support:

- Data management planning
- Data curation deidentification, metadata/description
- Data sharing
- Findability and visibility for research and research data
- Contact us: <u>ask.lib.montana.edu</u>

## Questions



#### Resources, Services, and Tools

- Data Management Resources MSU Library
- Research(er) Visibility Services MSU Library
- Data Storage and High Performance Computing RCI
- Office of Research Development
- Tools
  - MSU Dataset Search
  - <u>README template</u> (for research data)
  - README creator (for code)
  - **DMP Tool**



### Thanks!

