Montana State University Dataset Search

An open source tool to support data discovery, reuse, and analytics

Open Repositories 2021

Project Team

Sara Mannheimer, Project Lead

Jason A. Clark, Technical Lead

Jim Espeland, Software Engineer, Phase 1

Jakob Schultz, Software Engineer, Phase 2

Rhonda Borland, Metadata Curator, Phase 2

Kyle Hagerman, Student Research Assistant, Phase 1

Daniel Laden, Student Research Assistant, Phase 2

Prototype phase: OR2018

A **Prototype** for an Institutional Research Data Index

Sara Mannheimer, Jason A. Clark, and James Espeland

Montana State University

3

Dataset Search goals

1. Promote discovery for datasets from MSU*

2. Support community data repositories like <u>Dryad</u>

3. Support discovery for restricted data

4. Showcase research data as a legitimate research product that can be cited, reused, shared, and analyzed

* See Mannheimer, S., Sterman, L., & Borda, S. (2016). Discovery and reuse of open datasets: an exploratory study. Journal of eScience Librarianship, 5(1), e1091. https://doi.org/10.7191/jeslib.2016.1091

Features

Harvests metadata from DataCite and data repository APIs to automatically populate an index of datasets from MSU researchers

Human curators clean and edit metadata records

Visualization dashboards showcase data (in beta)

Optimized for search engines like Google

Supports discovery for restricted datasets

Outline

Public interface

Behind the scenes

Restricted data

Metadata + indexing

Continued challenges

Community collaboration

Public interface

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

MSU Dataset Search

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

🔍 wheat   ✕

Search

Advanced Search

Most popular ⌄

2018-12-10
Repository Analytics and Metrics Portal (RAMP) production snapshot dataset, 2018-11-01
Wheeler, Jonathan, Patrick O'Brien, Jeff Mixter, and Kenning Arlitsch

2014-01-01
Weddell Seal population Big Razorback TLS B-009 PS01 SV02
Rotella, Jay J., Robert A. Garrott, and Brendan Hodge

2017-07-17
A model predicting the relationships between seed size, yield, and actual yield in dry field peas
Smitchger, Jamin

Home / Search

Search

🔍 Enter keyword, name, or title...

Search

Advanced Search

Your search for **wheat** resulted in **5** records.

2017-12-07
Data from: Entomopathogenic nematodes combined with adjuvants presents a new potential biological control method for managing the wheat stem sawfly, Cephus cinctus (Hymenoptera: Cephidae)
Portman, Scott L., Krishnankutty, Sindhu M., Reddy, Gadi V. P.

2016-01-01
AmeriFlux US-Mj2 Montana Judith Basin summer fallow field
Stoy, Paul C.

2016-01-01
AmeriFlux US-Mj1 Montana Judith Basin wheat field
Stoy, Paul C.

2016-01-01
AmeriFlux US-MSR Montana Sun River winter wheat
Stoy, Paul C.

2016-01-01
AmeriFlux US-MC1 Creston, Montana pivot-irrigated spring wheat
Stoy, Paul C.

10

Home / Dataset View

2017-12-07

Data from: Entomopathogenic nematodes combined with adjuvants presents a new potential biological control method for managing the wheat stem sawfly, Cephus cinctus (Hymenoptera: Cephidae)

- Portman, Scott L. [Western Triangle Agricultural Research Center, US Department of Agriculture (USDA), International Society of Chemical Ecology, Entomological Society of America, Penn State University, University of Florida, Washington University School of Medicine, Southe]
- Krishnankutty, Sindhu M. [Invaio Sciences, USDA-PPQ/Xavier university, Montana State University, University of Notre Dame, Ivy Tech Community College, University of Illinois]
- Reddy, Gadi V. P. [Western Triangle Agricultural Research Center, Biological Control in Pest Management Systems of Plants: W3185, Ecological Society of America, Entomological Society of America, International Association for the Plant Protection Sciences]

Details

DOI:
10.5061/dryad.4p6pf

Keywords:
montana, cephus cinctus, wheat stem sawfly, agricultural pests, biological control

Categories:
Life Sciences & Earth Sciences

[Access dataset]

The wheat stem sawfly, (Cephus cinctus Norton) Hymenoptera: Cephidae, has been a major pest of winter wheat and barley in the northern Great Plains for more than 100 years. The insect's cryptic nature and lack of safe chemical control options make the wheat stem sawfly

11

Data from: Entomopathogenic Nematodes Combined with Adjuvants Presents a New Potential Biological Control Method for Managing the Wheat Stem Sawfly, Cephus cinctus (Hymenoptera: Cephidae)

Portman, Scott L., Montana State University
Krishnankutty, Sindhu M., Xavier University
Reddy, Gadi V.P.
Reddy, Gadi V. P., Montana State University
Publication date: December 7, 2017
Publisher: Dryad
https://doi.org/10.5061/dryad.4p6pf

Citation

Portman, Scott L.; Krishnankutty, Sindhu M.; Reddy, Gadi V.P.; Reddy, Gadi V. P. (2017), Data from: Entomopathogenic Nematodes Combined with Adjuvants Presents a New Potential Biological Control Method for Managing the Wheat Stem Sawfly, Cephus cinctus (Hymenoptera: Cephidae), Dryad, Dataset, https://doi.org/10.5061/dryad.4p6pf

Data Files

⬇ Download dataset

❯ December 7, 2017

Related Works

Article
https://doi.org/10.13...journal.pone.0169022

Metrics

👁 122 views

⬇ 12 downloads

🔖 1 citations

12

Dashboard
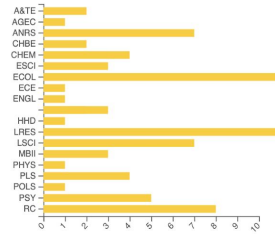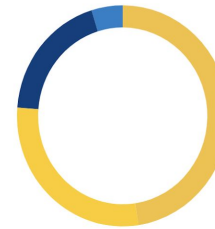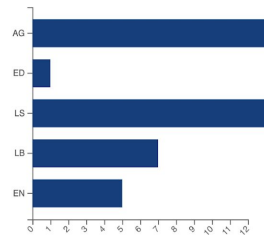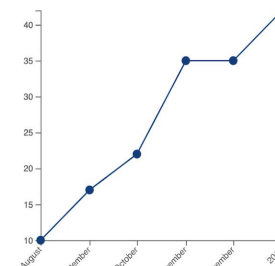
Behind the scenes

Harvesting

Dataset Candidates

DataCite_Jennifer_Thorson

Add Discard Thorson, Jennifer L. M. (02 Feb 2021). Additional file 1 of Epigenome-wide association study for pesticide (Permethrin and DEET) induced DNA methylation epimutation biomarkers for specific transgenerational disease

Add Discard Thorson, Jennifer L. M. (02 Feb 2021). Additional file 1 of Epigenome-wide association study for pesticide (Permethrin and DEET) induced DNA methylation epimutation biomarkers for specific transgenerational disease

DataCite_Peggy_Lamb

Add Discard Chavarro, Carolina (02 Feb 2021). Supplemental Material for Chavarro et al., 2020

Add Discard Chavarro, Carolina (02 Feb 2021). Supplemental Material for Chavarro et al., 2020

Add Discard Chavarro, Carolina (02 Feb 2021). Supplemental Material for Chavarro et al., 2020

Add Discard Chavarro, Carolina (02 Feb 2021). Supplemental Material for Chavarro et al., 2020

DataCite_Kenneth_Kephart

Add Discard Abe, F. (02 Feb 1993). Table 1

Add Discard Abe, F. (02 Feb 1983). Table 1

Add Discard Abe, F. (02 Feb 1991). Table 1

Add Discard Abe, F. (02 Feb 1998). Table 2

Select Feeds

☑ DataCite_montana_state
☑ DataCite_montana_state_university
☑ DataCite_Jackson_St_Clair
☑ DataCite_Miranda_Roy
☑ DataCite_Mitchell_Harris
☑ DataCite_Patrick_Carr
☑ DataCite_Shuang_Zhou
☑ DataCite_William_Franck
☑ DataCite_Thomas_Allen
☑ DataCite_Jennifer_Thorson
☑ DataCite_Peggy_Lamb
☑ DataCite_Kenneth_Kephart
☑ DataCite_Valerie_Smith
☑ DataCite_Qasim_Khan
☑ DataCite_Zachariah_Miller
☑ DataCite_Rachel_Leisso
☑ DataCite_Anamika_Sharma
☑ DataCite_John_Miller
☑ DataCite_Kate_Fuller
☑ DataCite_Brock_Smith
☑ DataCite_Mariana_Carrera
☑ DataCite_Justin_Gallagher
☑ DataCite_Daniel_Bigelow
🔄 retrieving data from DataCite_Diane_Charlton

16

MSU Dataset Search List by Status Admin

MSU Dataset Search List by Status

ID	STATUS	TITLE	OPTIONS
1	a	Data from: Entomopathogenic Nematodes Combined with Adjuvants Presents a New Potential Biological Control Method for Managing the Wheat Stem Sawfly, Cephus cinctus (Hymenoptera: Cephidae)	⊕ Edit | ⊖ Delete
2	a	Data from: Assessing the sustainability of African lion trophy hunting, with recommendations for policy	⊕ Edit | ⊖ Delete
3	a	Supporting Data: A large-scale dataset of solar event reports from automated feature recognition modules	⊕ Edit | ⊖ Delete
4	a	Molecular Models and Wave Function Definitions for Models A-E of the [2Fe]F Cluster in FeFe-hydrogenase Maturase Enzyme HydF	⊕ Edit | ⊖ Delete
5	a	Molecular Level Characterization of Dissolved Organic Carbon and Microbial Diversity in the WAIS Divide Replicate Core	⊕ Edit | ⊖ Delete

Metadata

MSU Dataset Search Edit Object

MSU Dataset Search Metadata

Dataset Name

Data from: Entomopathogenic Nematodes Combined with Adjuvants Presents a New Potential Biological Control Method fo

Creator 1

Creator 1 Name [Last, First Middle (or Middle Initial)]

Portman , Scott L.

Creator 1 ORCiD

Creator 1 Type

Creator 1 URL

Dataset Linked Data Category 2

Dataset Linked Data Category 3

Dataset Encoding Format (e.g. CSV)

Dataset Copyright Conditions

CC0

Dataset Version Number

Object Status

Object's current status is a. Edit object's status below. ⊡

○ a = Active   ○ r = Review   ○ p = Pending   ○ i = Inactive   ○ d = Discarded

○ u = Unprocessed

Submit Metadata Edits

20

Restricted Data

Pilot with Center for American Indian and Rural Health Equity (CAIRHE)

Browse   Dashboard   About

Home / Search

Search

🔍 | Enter keyword, name, or title |

Search

[Advanced Search](#)

Your search for **cairhe** resulted in **4** records.

2019-04-01
The healthy workplace survey, Gallatin County organizations, 2019 🔒
Knight, Kelly E., Ellis, Colter

2015-08-01/2016-07-01
Focus group discussions about drinking and driving with rural young adults 🔒
Greene, Kaylin

2017-09-01/2018-08-31
The Fort Peck substance abuse and recovery survey 🔒
Skewes, Monica

2014-09-01/2018-08-31
The Fort Peck sexual health project 🔒
Rink, Elizabeth

23

2019-04-01

The healthy workplace survey, Gallatin County organizations, 2019

- Knight, Kelly E. [Center for American Indian and Rural Health Equity (CAIRHE), University of Colorado, Boulder, San Francisco State University]
- Ellis, Colter [Center for American Indian and Rural Health Equity (CAIRHE), University of Colorado, Boulder, Idaho State University]

Request dataset

2018-01-01 to 2019-01-01(date collected)

This dataset contains responses from the Healthy Workplace Survey. Respondents were victim service providers (N = 412) working in organizations based primarily in Gallatin County, Montana. Measures assessed: social demographics; employment; occupational issues (burnout, employee work life mismatch, secondary traumatic stress, and

Details

🔒 **RESTRICTED ACCESS**

Keywords:
occupation-based secondary trauma, victim service providers, community-based participatory research, rural

Categories:
Social Sciences
Health & Medical Sciences

24

Browse   Dashboard   About

Home / Dataset View

2019-04-01

The healthy workplace survey, Gallatin County organizations, 2019

- Knight, Kelly E. [Center for American Indian and Rural Health Equity (CAIRHE), University of Colorado, Boulder, San Francisco State University]
- Ellis, Colter [Center for American Indian and Rural Health Equity (CAIRHE), University of Colorado, Boulder, Idaho State University]

Request dataset

2018-01-01 to 2019-01-01(date collected)

This dataset contains responses from the Healthy Workplace Survey. Respondents were victim service providers (N = 412) working in organizations based primarily in Gallatin County, Montana. Measures assessed: social demographics; employment; occupational issues (burnout, employee work-life mismatch, secondary traumatic stress, and

arc.lib.montana.edu says

To request access to deidentified data, contact CAIRHE (James Burroughs: jburroughs@montana.edu or 406-994-4407). All requests may be subject to prior review and approval by community partners, including tribal IRB and tribal council where applicable.

OK

Details

🔒 **RESTRICTED ACCESS**

Keywords:
occupation-based secondary trauma, victim service providers, community-based participatory research, rural

Categories:
Social Sciences
Health & Medical Sciences

25

Metadata + Indexing

Enriched metadata from researcher profiles

Search engine optimization

Google Structured Data Testing Tool

```html
1   <!doctype html>
2   <html lang="en" vocab="http://schema.org/" typeof="WebPage">
3   <head>
4   <title>MSU Data Catalog Record - Montana State University (MSU)
    Library</title>
5   <meta http-equiv="Content-Type" content="text/html; charset=utf-8" />
6   <meta name="description" content="Dataset metadata" />
7   <meta name="keywords" content="" />
8   <meta name="robots" content="noindex, nofollow" />
9   <link rel="canonical" href="http://shelf.lib.montana.edu/msu-data-
    catalog/item/1" />
10  <link rel="stylesheet" type="text/css" media="screen, projection"
    href="/msu-data-catalog/meta/styles/master.css" />
11  </head>
12  <body class="fullWidth default">
13  <div id="contain">
14  <div id="mastHead">
15    <h1><a href="/msu-data-catalog/index.php">Montana State University
    Data Catalog - Montana State University (MSU) Library</a></h1>
16    <p class="hidden"><a href="#mainContent" title="main content">skip
    navigation</a></p>
17  </div><!-- end masthead div -->
```
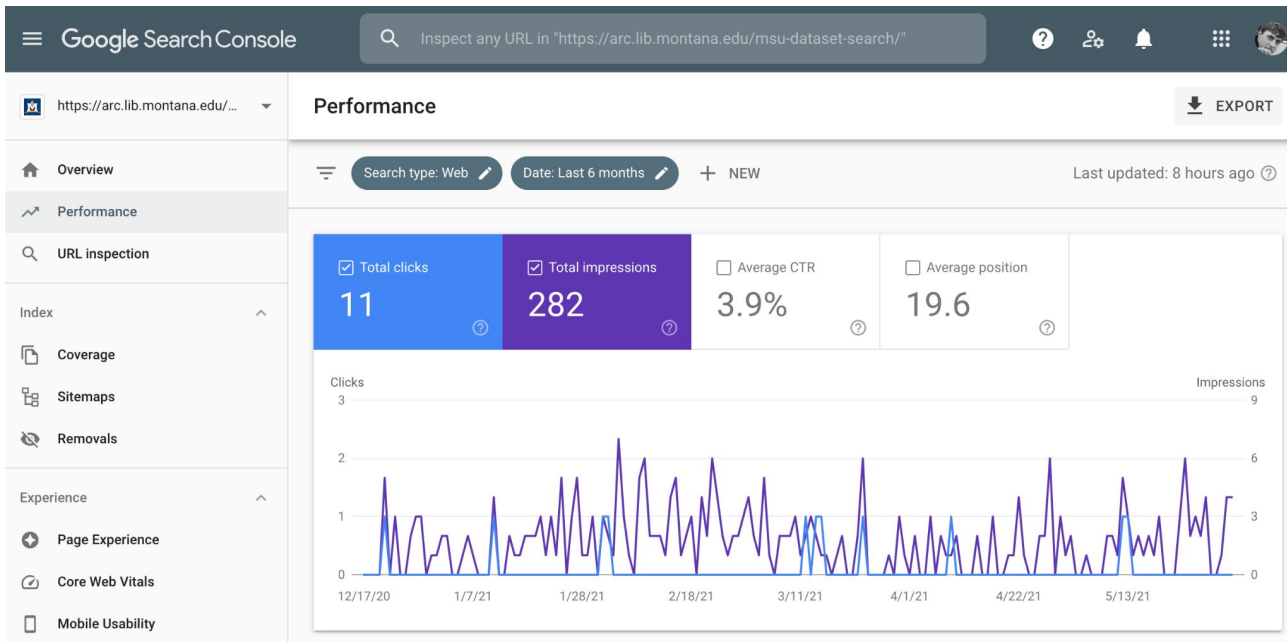
Google

🔍 montana state university wheat data   🎤

Google Search   I'm Feeling Lucky

Search engine indexing

Code available in Github, and archived in Zenodo

Continued challenges

Continued challenges

Accuracy of automated search
- Scholar disambiguation
- Manual metadata

Ideas for addressing challenges

- Scholar disambiguation
  - Curate list of faculty and graduate students from Office of Planning and Analysis
  - Add Google Scholar keywords

Ideas for addressing challenges

- Manual metadata
  - More/most data repositories provide DOIs
    - May narrow our API call to only DataCite
  - More data repositories allow ORCIDs
  - DataCite inclusion of Research Organization Registry IDs

Community collaboration

Data Discovery Collaboration

Members

- NYU Langone Health
- Northwestern University
- Memorial Sloan Kettering Cancer Center
- Montana State University
- University of Maryland, Baltimore
- University of Pittsburgh
- Wayne State University
- Weill Cornell Medicine
- Zucker School of Medicine at Hofstra/Northwell

https://datadiscoverycollaboration.org/members

Data Discovery Collaboration

Mission

- To enhance discovery of data and other research products in order to maximize their value.
- Governance
  - Steering Committee
  - Core Groups:
    - Metadata Development & Cataloging
    - Software Development
    - Advocacy and Communications

https://datadiscoverycollaboration.org/about-us

Funding and advisors

Funding for Dataset Search has been provided by IMLS, 2018-2019 and by Network of National Libraries of Medicine - Pacific Northwest Region, 2019-2020

2018 Dataset Search Advisory Board:

- Jeffrey Grethe (BioCaddie/DataMed)
- Cynthia Hudson-Vitale (SHARE)
- Kevin Read and Nicole Contaxis (Data Discovery Collaboration Project)

Thank you!

Sara Mannheimer

sara.mannheimer@montana.edu

[@saramannheimer](https://twitter.com/saramannheimer)

Jason A. Clark

jaclark@montana.edu

[@jaclark](https://twitter.com/jaclark)