

# Introduction to Metadata in Dataverse Repositories

## For Researchers and Support Staff

November 15, 2023

Julian Gautier

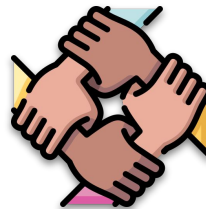
IQSS, Harvard University





Introduce yourself in chat

Post questions in Q&A





Review [Contributor Covenant Code of Conduct](#)

- Use welcoming and inclusive language
- Respect differing viewpoints and experiences
- Gracefully accept constructive criticism
- Focus on what is best for the community
- Show empathy towards other community members



# What I'll be covering

-  1 | Introduction
-  2 | About metadata in Dataverse repositories
-  3 | Searching and browsing
-  4 | What is “good” metadata?
-  5 | Questions, survey, upcoming webinars



# Introduction

By the end of the webinar...



You should be able to use  
Dataverse's metadata features  
to make data more FAIR



I should be able to help  
improve our webinars, like  
this one, based on your  
feedback

# How this webinar was designed



Conversations with you  
(mostly with support staff)

Conversations with  
curation team at Harvard  
Dataverse



Reviews of support  
emails and others  
recorded conversations  
about metadata



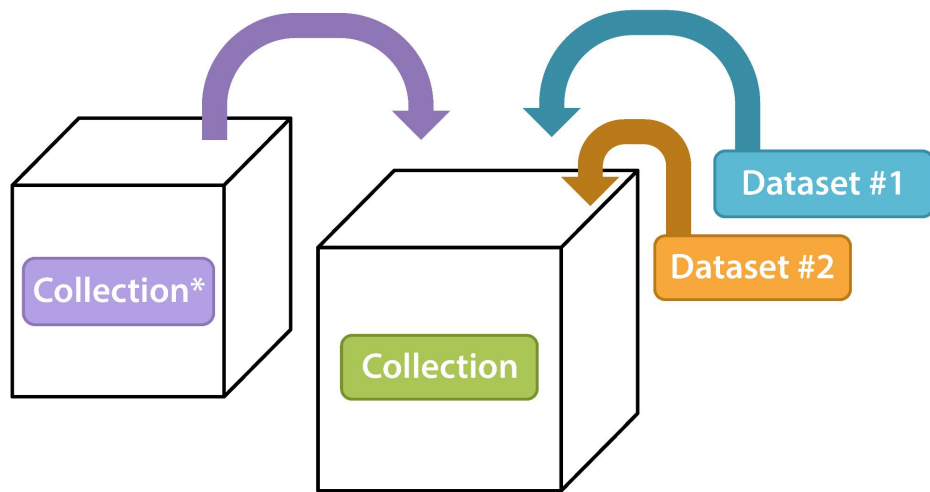
A web survey

# About me

- Working on Dataverse for ~7 years
- User experience researcher, informing design with qualitative research (e.g. interviews, usability testing, surveys) and quantitative research (e.g. analytics)
- Working on metadata design
- Working on curation and support team at Harvard Dataverse

# About metadata in Dataverse repositories

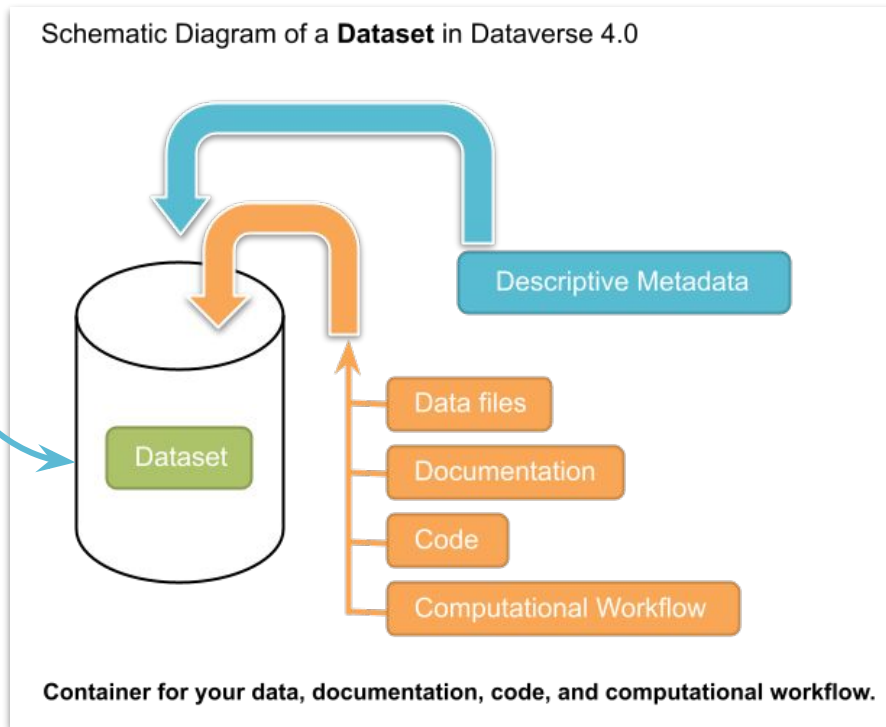
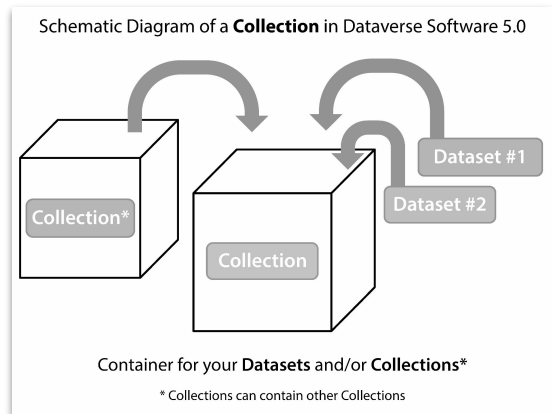
Schematic Diagram of a **Collection** in Dataverse Software 5.0



Container for your **Datasets** and/or **Collections\***

\* Collections can contain other Collections

## About metadata in Dataverse repositories





## Metadata you can use to describe

- Collections
- Datasets
- Files

## Metadata that Dataverse creates automatically

- Such as persistent IDs, publication dates, version numbers
- Metadata saved from "ingested" tabular data

# Metadata you can use to describe collections

- Names
- Affiliations
- Identifiers of collection URLs
- Categories
- Emails
- Descriptions

New Dataverse

[Demo Dataverse](#) >

\*Asterisks indicate required fields

**Host Dataverse** ?

Enter Dataverse Name

**Dataverse Name** \* ?

Julian Gautier Dataverse

**Affiliation** ?

Harvard University

**Identifier** \* ?

<https://dataverse.harvard.edu/dataverse/>

**Category** \* ?

Select one...

**Email** \* ?

juliangautier@g.harvard.edu

**Description** ?

This field supports only certain [HTML tags](#).

# Metadata you can use to describe datasets

Using many fields shown  
in collapsible panels,  
such as “Citation  
Metadata” and “Terms”

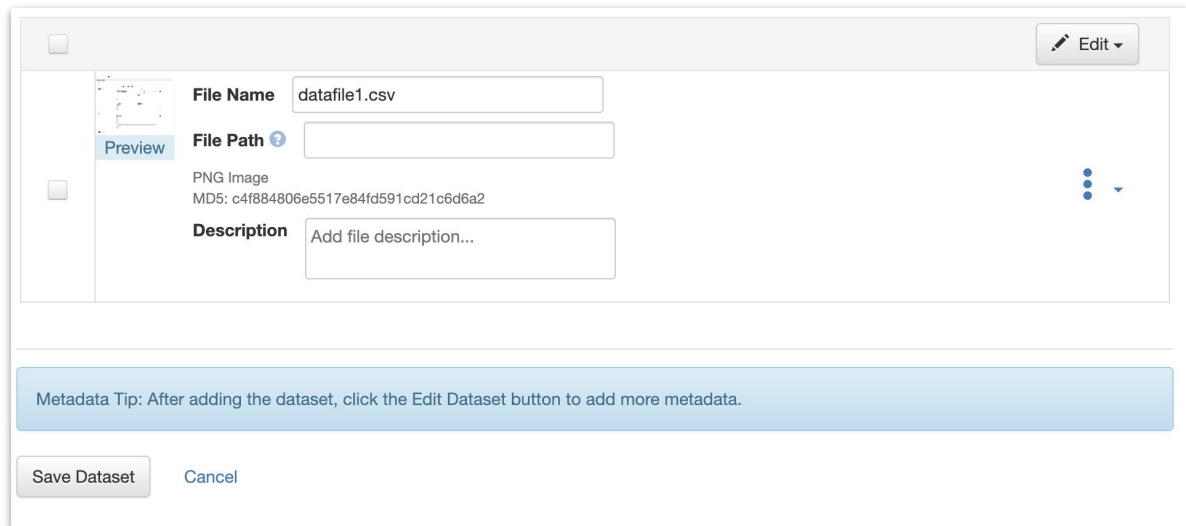
\*Asterisks indicate required fields

### Citation Metadata ^

<b>Title *</b> ?	<input type="text"/>		
	<button>Add "Replication Data for" to Title</button>		
<b>Author *</b> ?	<b>Name *</b> ?	<b>Affiliation</b> ?	
	<input type="text" value="Gautier (SU), Julian"/>	<input style="border: 1px solid #ccc;" type="text" value="Harvard University"/>	<input style="border: 1px solid #ccc;" type="button" value="+"/>
		<b>Identifier Type</b> ?	
		<input style="border: 1px solid #ccc;" type="text" value="Select..."/>	
	<b>Identifier</b> ?		
	<input type="text"/>		
<b>Point of Contact *</b> ?	<b>Name</b> ?	<b>Affiliation</b> ?	
	<input type="text" value="Gautier (SU), Julian"/>	<input style="border: 1px solid #ccc;" type="text" value="Harvard University"/>	<input style="border: 1px solid #ccc;" type="button" value="+"/>
	<b>E-mail *</b> ?		
	<input type="text" value="juliangautier@g.harvard.edu"/>		
<b>Description *</b> ?	This field supports only certain <a href="#">HTML tags</a> .		
	<b>Text</b> ?	<input type="text"/>	

# Metadata you can use to describe files

- File names
- File paths
- Descriptions



The screenshot shows the Dataverse file metadata editor interface. It features a sidebar on the left with a 'Preview' button. The main area contains fields for 'File Name' (datafile1.csv), 'File Path' (with a help icon), 'Description' (Add file description...), and a 'PNG Image' section with an MD5 hash. An 'Edit' button is in the top right. A blue tip box at the bottom states: 'Metadata Tip: After adding the dataset, click the Edit Dataset button to add more metadata.' At the very bottom are 'Save Dataset' and 'Cancel' buttons.

File Name: datafile1.csv

File Path:

Preview

PNG Image

MD5: c4f884806e5517e84fd591cd21c6d6a2

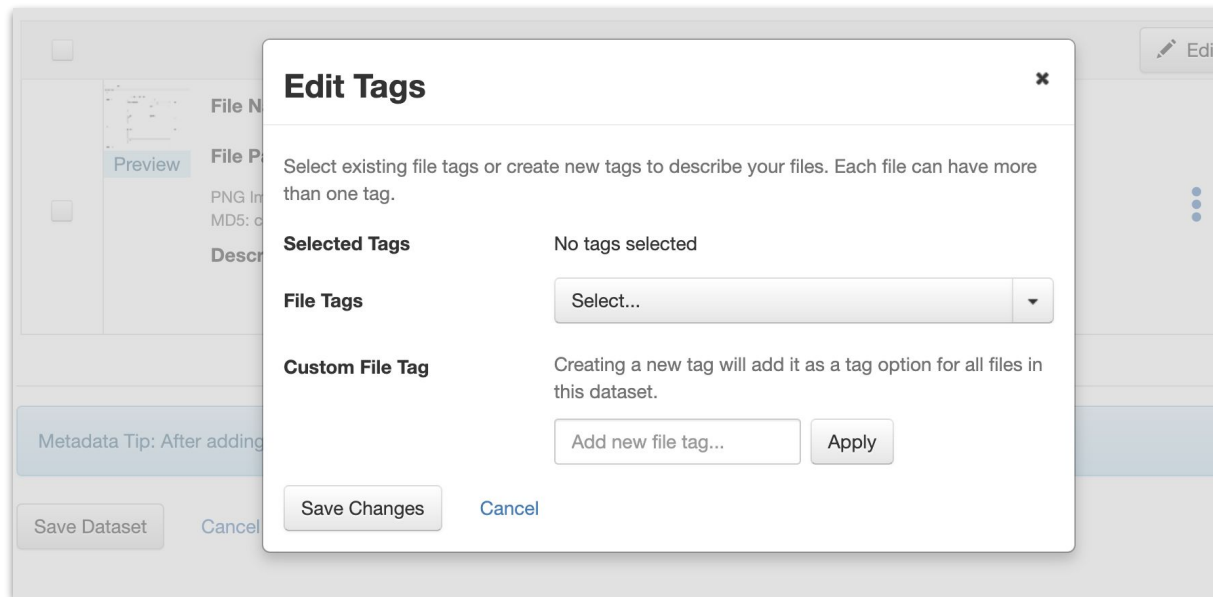
Description: Add file description...

Metadata Tip: After adding the dataset, click the Edit Dataset button to add more metadata.

Save Dataset Cancel

# Metadata you can use to describe files

- Tags
- Restrictions
- Embargo dates and reasons




# Metadata you can use to describe files


- Tags
- Restrictions
- Embargo dates and reasons

### Restrict Access

Restricting limits access to published files. People who want to use the restricted files can request access by default. **If you disable request access, you must add information about access to the Terms of Access field.**

Learn about restricting files and dataset access in the [User Guide](#).

**Request Access**  ☒ Enable access request

**Terms of Access for Restricted Files** 

Save Changes [Cancel](#)

# Metadata you can use to describe files

- Tags
- Restrictions
- Embargo dates and reasons

## Edit Embargo

Edit the planned embargo for the selected file or files. Once this dataset version is published, you will need to contact an administrator to change the embargo date or reason of the file or files.

**Add or Change**      **Select the embargo end-date \***

2023-11-14

**November 2023**

S	M	T	W	T	F	S
			1	2	3	4
5	6	7	8	9	10	11
12	13	14	15	16	17	18

Save Changes

# Metadata Dataverse creates for datasets

- Persistent IDs
- Creation dates
- Publication dates
- Version numbers
- Sum of file downloads
- Sum of page views  
(in repositories with Make Data Count enabled)

The screenshot displays the 'Metadata' tab of a Dataverse dataset page. At the top, there are four tabs: 'Files', 'Metadata', 'Terms', and 'Versions'. Below the tabs, there are two buttons: 'Add + Edit Metadata' (with a pencil icon) and 'Export Metadata' (with a download icon and a dropdown arrow). The main content area is titled 'Citation Metadata' with an upward arrow. It contains three rows of metadata:

<b>Persistent Identifier</b> ?	doi:10.7910/DVN/VZ5JFG
<b>Publication Date</b> ?	2023-11-13
<b>Title</b> ?	Replication Data for: "Prevalence of Neuromyths Among Students and Pre-



# Metadata Dataverse creates for datasets

- Differences among versions

	Dataset Version	Summary	Contributors	
<input type="checkbox"/>	12.0	<b>Citation Metadata:</b> Description (1 Changed); Notes (Changed); <b>Additional Citation Metadata:</b> (3 Changed); <b>Files (Added: 98; Removed: 84);</b> <a href="#">View Details</a>	Julian Gautier	202
<input type="checkbox"/>	11.0	<b>Citation Metadata:</b> Description (1 Changed); Notes (Changed); <b>Additional Citation Metadata:</b> (4 Changed); <b>Files (Added: 84; Removed: 57);</b> <a href="#">View Details</a>	Julian Gautier	202
<input type="checkbox"/>	10.0	<b>Citation Metadata:</b> Notes (Changed); Description (1 Changed); <b>Additional Citation Metadata:</b> (2 Removed, 2 Changed); <b>Files (Added: 57; Removed: 50);</b> <a href="#">View Details</a>	Julian Gautier	202
<input type="checkbox"/>	9.0	<b>Citation Metadata:</b> Description (1 Changed); <b>Additional Citation Metadata:</b> (1 Added); <b>Files (Added: 50; Removed: 37);</b> <a href="#">View Details</a>	Julian Gautier	202
<input type="checkbox"/>	8.1	<b>Citation Metadata:</b> Description (1 Changed); <a href="#">View Details</a>	Julian Gautier	202
<input type="checkbox"/>	8.0	<b>Citation Metadata:</b> Description (1 Changed); <b>Additional Citation Metadata:</b> (1 Added, 1 Changed); <b>Files (Added: 1; Replaced: 1);</b> <a href="#">View Details</a>	Julian Gautier	202
<input type="checkbox"/>	7.0	<b>Citation Metadata:</b> Description (1 Changed); <b>Files (Added: 35; Removed: 36);</b> <a href="#">View Details</a>	Julian Gautier	202

# Metadata Dataverse creates for files

- File types
- Byte sizes
- Creation dates
- Publication dates
- Checksums (MD5s)
- File paths
- Counts of downloads



[DP\\_LIVE\\_11072022201648574.tab](#)

Tabular Data - 62.3 KB

Published Nov 13, 2023

1 Download

8 Variables, 1121 Observations UNF:6:HSeM...Qow== 

This .csv file contains information on the labour force participation rate (source: OECD database).

Data

# Metadata Dataverse creates for ingested tabular files

- Counts of variables and observations
- Variable statistics
- Checksums (UNFs)

File Metadata ^	
File UNF	UNF:6:HSeMI8h1XrpAPAOzE50Qow==
Original File MD5	05ed7a1c7330df3e88e57629bf15b5ff
Type	Tab-Delimited
Variables	8
Observations	1121

## TIME

f7570319 Location: Summary Statistics: Min. 2000.0; Valid 1121.0; Mean 2010.6030330062445; Max. 2021.0; StDev 6.318365357642705;

Variable Format: numeric

# Metadata blocks

- Enabling and configuring
- Editing and creating

## Metadata Fields

Choose the metadata fields to use in dataset templates and when adding a dataset to this dataverse.

- ☒ Use metadata fields from Harvard Dataverse
- ☒ Citation Metadata (Required) [\[+\] View fields](#)
- ☐ 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](#)
- ☐ Computational Workflow Metadata [\[+\] View fields](#)

# Metadata blocks

- Enabling and configuring
- Editing and creating

## Metadata Fields

Choose the metadata fields to use in dataset templates and when adding a dataset to this dataverse.

☐ Use metadata fields from Harvard Dataverse

☒ Citation Metadata (Required)

<input checked="" type="checkbox"/> Title	Required by Dataverse
<input checked="" type="checkbox"/> Subtitle	<input type="radio"/> Required <input checked="" type="radio"/> Optional
<input checked="" type="checkbox"/> Alternative Title	<input type="radio"/> Required <input checked="" type="radio"/> Optional
<input checked="" type="checkbox"/> Alternative URL	<input type="radio"/> Required <input checked="" type="radio"/> Optional
<input checked="" type="checkbox"/> Other Identifier	
Other Identifier Agency	<input type="radio"/> Required <input checked="" type="radio"/> Optional

About metadata in Dataverse repositories

# Metadata blocks

- Enabling and configuring
- Editing and creating



# Dataset templates

- Copying metadata to multiple datasets
- Creating instructions

The screenshot shows the 'Template Name' field at the top, followed by two explanatory notes: '\*Asterisks indicate metadata fields that users will be required to fill out while adding a dataset to this dataverse.' and '\*Asterisks indicate required fields'. Below these is a section titled 'Citation Metadata' with a chevron icon. This section contains five rows of metadata fields, each with a label, a required field indicator (asterisk), a help icon (question mark), and a 'Custom Instructions' link. The fields are: Title, Subtitle, Alternative Title, Alternative URL, and Other Identifier. Each field has a corresponding text input box. The 'Alternative URL' box contains the text 'https://'. The 'Custom Instructions' for each field are currently set to '(None - click to add)'.

Template Name \* ?

\*Asterisks indicate metadata fields that users will be required to fill out while adding a dataset to this dataverse.

\*Asterisks indicate required fields

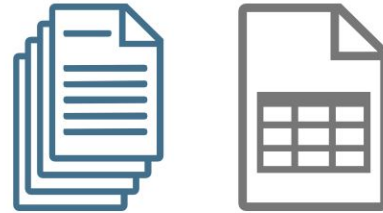
Citation Metadata ^

Title * ?	Custom Instructions: (None - click to add)
Subtitle ?	Custom Instructions: (None - click to add)
Alternative Title ?	Custom Instructions: (None - click to add)
Alternative URL ?	Custom Instructions: (None - click to add)
Other Identifier ?	Custom Instructions: (None - click to add)

# Searching and browsing



Searching in  
collections

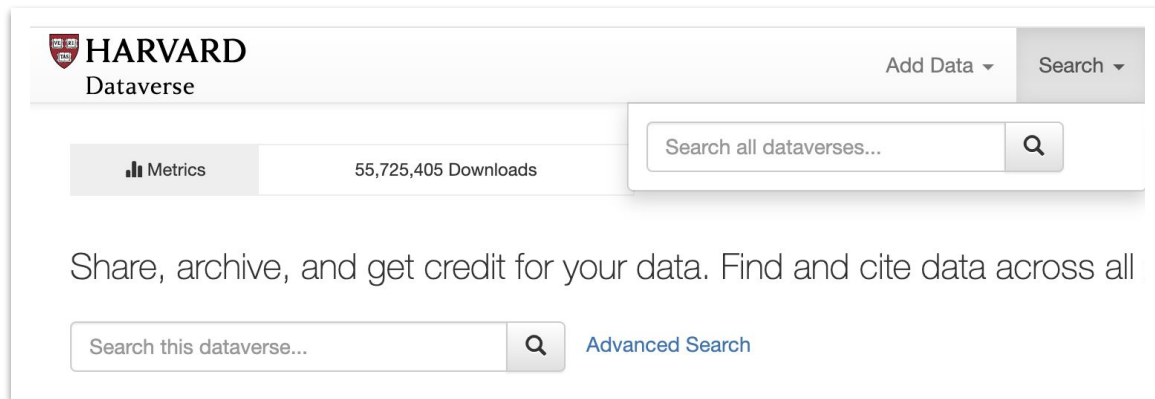


Searching in  
datasets for files



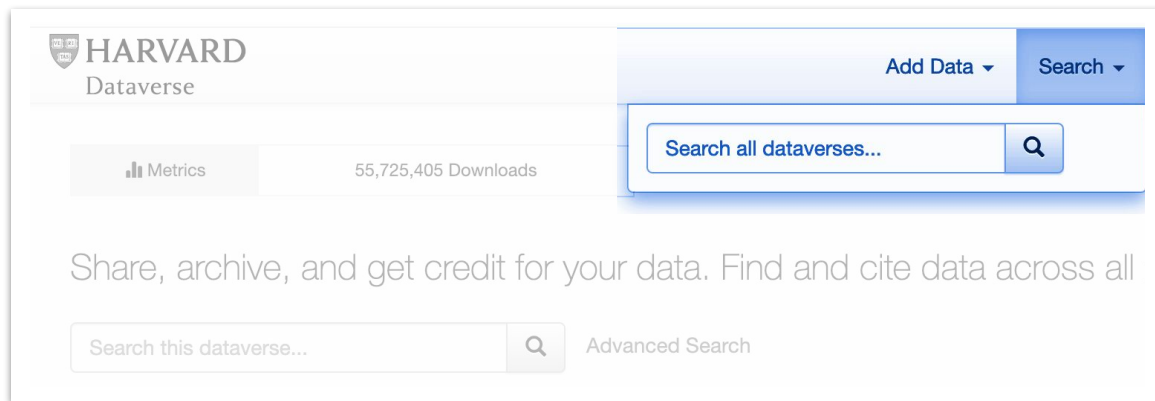
# Searching in collections

- Using the general search boxes
- Using advanced search
- Filtering with search facets
- Sorting



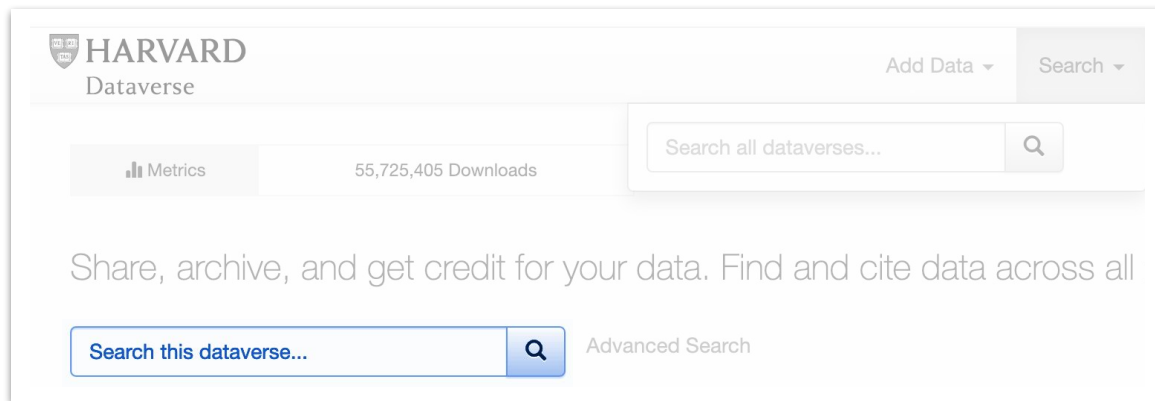
# Searching in collections

- Using the general search boxes
- Using advanced search
- Filtering with search facets
- Sorting



# Searching in collections

- Using the general search boxes
- Using advanced search
- Filtering with search facets
- Sorting



# Searching in collections

- Using the general search boxes
- Using advanced search
- Filtering with search facets
- Sorting

The screenshot shows a search interface for 'Dataverses'. At the top is a 'Find' button. Below it is a header 'Dataverses' with an upward arrow. The main search area contains five input fields: 'Name', 'Identifier', 'Affiliation', 'Description', and 'Subject'. Each field has a blue question mark icon to its left. The 'Subject' field is a dropdown menu showing a list of categories: Agricultural Sciences, Arts and Humanities, Astronomy and Astrophysics, Business and Management, and Chemistry. At the bottom of the interface is a bar labeled 'Datasets: Citation Metadata' with a downward arrow.

Find

Dataverses ^

Name ?

Identifier ?

Affiliation ?

Description ?

Subject ?

- ☐ Agricultural Sciences
- ☐ Arts and Humanities
- ☐ Astronomy and Astrophysics
- ☐ Business and Management
- ☐ Chemistry

Datasets: Citation Metadata v

## Searching and browsing

# Searching in collections

- Using the general search boxes
- Using advanced search
- Filtering with search facets
- Sorting

The screenshot displays a search interface with a left sidebar for filtering and a main area for search results.

**Left Sidebar Filters:**

- Dataverses (6,345)** (checked)
- Datasets (152,193)** (checked)
- Files (2,061,989)** (unchecked)
- Dataverse Category**
  - Research Project (2,356)
  - Researcher (1,981)
  - Organization or Institution (504)
  - Research Group (444)
  - Journal (130)
- Metadata Source**
  - Harvard Dataverse (95,091)
  - Harvested (63,447)
- Publication Year**
  - 2023 (7,513)
  - 2022 (19,832)
  - 2021 (23,904)
  - 2020 (9,905)
  - 2019 (4,510)
- Subject**
  - Social Sciences (61,213)
  - Arts and Humanities (36,231)
  - Medicine, Health and Life Sciences (9,478)

**Main Search Results:**

1 to 10 of 158,538 Results

Sort ▾

**Replication Data for "You Had Better Mention All of Them: Race and Gender Effects in Election Loss Narratives"**

Nov 13, 2023

Haines, Pavielle, 2023, "Replication Data for "You Had Better Mention All of Them: Race and Gender Effects in Election Loss Narratives"", <https://doi.org/10.7910/DVN/FDFMZT>, Harvard Dataverse, V1, UNF:6:8RX4GeZTwMdhjmZSMadP0w== [fileUNF]

The replication files posted here correspond to the article Replication Data for "You Had Better Mention All of Them: Race and Gender Effects in Election Loss Narratives." This article is available through Political Research Quarterly.

**First-choice assay - 12 replicates**

Nov 13, 2023 - 1-Time-lapsed First-choice assay

Costa-da-Silva, AL, 2023, "First-choice assay - 12 replicates", <https://doi.org/10.7910/DVN/69COMS>, Harvard Dataverse, V1

These are the time-lapse recorders for 12 trials in the 2-choice water trap assay. Each trial has multiple files in order. The files are named by the number of the experiment (ExpX), replicates on the same day (lowercase letters, when there is more than one trial on the same day...

**1-Time-lapsed First-choice assay (Florida International University)**

Nov 13, 2023 Andre Silva Dataverse

**Oviseekmeter assay (choice dynamic) - 30 gravid females - 9 replicates**

Nov 13, 2023 - 2-Time-lapsed Oviseekmeter assay

Costa-da-Silva, AL, 2023, "Oviseekmeter assay (choice dynamic) - 30 gravid females - 9 replicates", <https://doi.org/10.7910/DVN/BZY3YJ>, Harvard Dataverse, V1

## Searching and browsing

# Searching in collections

- Using the general search boxes
- Using advanced search
- Filtering with search facets
- Sorting

The screenshot displays a search interface with the following components:

- Filters (Left Sidebar):**
  - Dataverse Category:** Research Project (2,356), Researcher (1,981), Organization or Institution (504), Research Group (444), Journal (130). A "More..." link is present.
  - Metadata Source:** Harvard Dataverse (95,091), Harvested (63,447).
  - Publication Year:** 2023 (7,513), 2022 (19,832), 2021 (23,904), 2020 (9,905), 2019 (4,510). A "More..." link is present.
  - Subject:** Social Sciences (61,213), Arts and Humanities (36,231), Medicine, Health and Life Sciences (9,478).
- Search Results (Main Content):**
  - 1 to 10 of 158,538 Results
  - Sort Menu:** A dropdown menu with options: Name (A-Z), Name (Z-A), Newest, and Oldest.
  - Result 1:** "Replication Data for 'You Had Better Mention All of Them: Race and Gender Effects in Election Loss Narratives'" by Haines, Pavielle, 2023. Includes a DOI link and a description of the replication files.
  - Result 2:** "First-choice assay - 12 replicates" by Costa-da-Silva, AL, 2023. Includes a DOI link and a description of the time-lapse recorders.
  - Result 3:** "1-Time-lapsed First-choice assay (Florida International University)" by Andre Silva, 2023.
  - Result 4:** "Oviseekmeter assay (choice dynamic) - 30 gravid females - 9 replicates" by Costa-da-Silva, AL, 2023. Includes a DOI link and a description of the time-lapse recorders.

# Searching in datasets for files

- Using the general search box
- Filtering with search facets
- Sorting
- Viewing files in folders

The screenshot displays a web interface for browsing dataset files. At the top, there are tabs for 'Files', 'Metadata', 'Terms', and 'Versions'. Below these is a 'Change View' section with 'Table' and 'Tree' options. A search bar labeled 'Search this dataset...' is positioned next to a magnifying glass icon. To the right of the search bar is a '+ Upload Files' button. Below the search bar, there are filters for 'File Type: All' and 'Access: All', along with a 'Sort' button with an upward arrow icon. The main content area shows a list of files, starting with '1 to 10 of 88 Files'. Each file entry includes a file icon, a filename, a description, file size, publication date, download count, and MD5 hash. The files listed are:

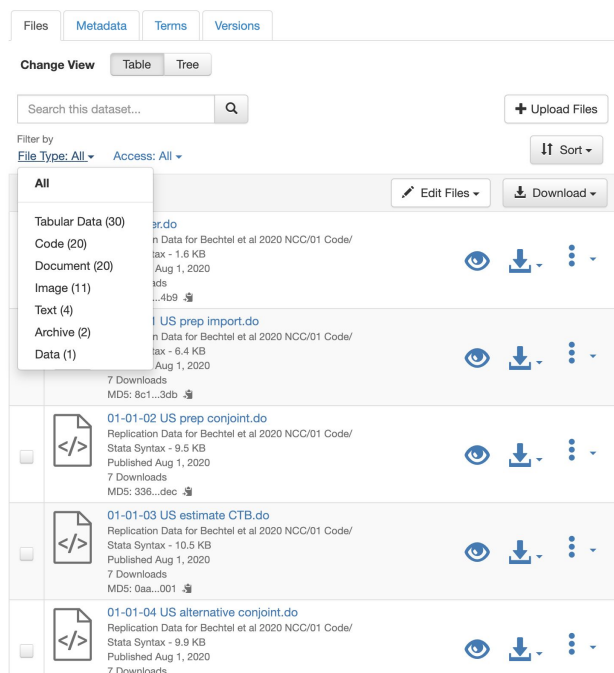
- 00 master.do**: Replication Data for Bechtel et al 2020 NCC/01 Code/ Stata Syntax - 1.6 KB, Published Aug 1, 2020, 8 Downloads, MD5: b71...4b9
- 01-01-01 US prep import.do**: Replication Data for Bechtel et al 2020 NCC/01 Code/ Stata Syntax - 6.4 KB, Published Aug 1, 2020, 7 Downloads, MD5: 8c1...3db
- 01-01-02 US prep conjoint.do**: Replication Data for Bechtel et al 2020 NCC/01 Code/ Stata Syntax - 9.5 KB, Published Aug 1, 2020, 7 Downloads, MD5: 336...dec
- 01-01-03 US estimate CTB.do**: Replication Data for Bechtel et al 2020 NCC/01 Code/ Stata Syntax - 10.5 KB, Published Aug 1, 2020

Each file entry also has icons for viewing, downloading, and a menu of additional actions.

## Searching and browsing

# Searching in datasets for files

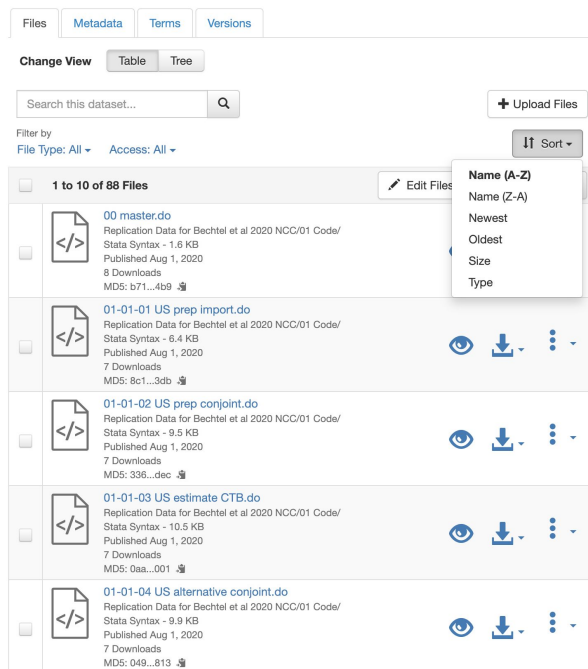
- Using the general search box
- Filtering with search facets
- Sorting
- Viewing files in folders





# Searching in datasets for files

- Using the general search box
- Filtering with search facets
- **Sorting**
- Viewing files in folders



# Searching in datasets for files

- Using the general search box
- Filtering with search facets
- Sorting
- Viewing files in folders



# What is “good” metadata?

Metadata that helps others find datasets

Metadata that helps others re-use datasets

Examples of well-described datasets

What is “good” metadata?

# Metadata that helps others find datasets

Metadata fields that discovery systems use when returning search results

Metadata that leverages how people look for data

\*Asterisks indicate required fields

### Citation Metadata ^

<b>Title *</b> ?	<input type="text"/>		
	<input type="button" value="Add 'Replication Data for' to Title"/>		
<b>Author *</b> ?	<b>Name *</b> ?	<b>Affiliation *</b> ?	
	<input type="text" value="Gautier (SU), Julian"/>	<input style="border: 1px solid #ccc;" type="text" value="Harvard University"/>	<input style="border: 1px solid #ccc;" type="button" value="+"/>
		<b>Identifier Type *</b> ?	
		<input style="border: 1px solid #ccc;" type="text" value="Select..."/>	
	<b>Identifier *</b> ?		
	<input type="text"/>		
<b>Point of Contact *</b> ?	<b>Name *</b> ?	<b>Affiliation *</b> ?	
	<input type="text" value="Gautier (SU), Julian"/>	<input style="border: 1px solid #ccc;" type="text" value="Harvard University"/>	<input style="border: 1px solid #ccc;" type="button" value="+"/>
	<b>E-mail *</b> ?		
	<input type="text" value="juliangautier@g.harvard.edu"/>		
<b>Description *</b> ?	This field supports only certain <a href="#">HTML tags</a> .		
	<b>Text *</b> ?		
	<input type="text"/>	<input style="border: 1px solid #ccc;" type="button" value="+"/>	

What is “good” metadata?

# Metadata that helps others find datasets

Metadata fields that discovery systems use when returning search results

Metadata that leverages how people look for data

Metadata

\*Asterisks indicate required fields

Citation Metadata ▼

Geospatial Metadata ▼

Social Science and Humanities Metadata ▼

Astronomy and Astrophysics Metadata ▼

Life Sciences Metadata ▼

Journal Metadata ▼

What is “good” metadata?

# Metadata that helps others find datasets

Metadata fields that discovery systems use when returning search results



Metadata that leverages how people look for data



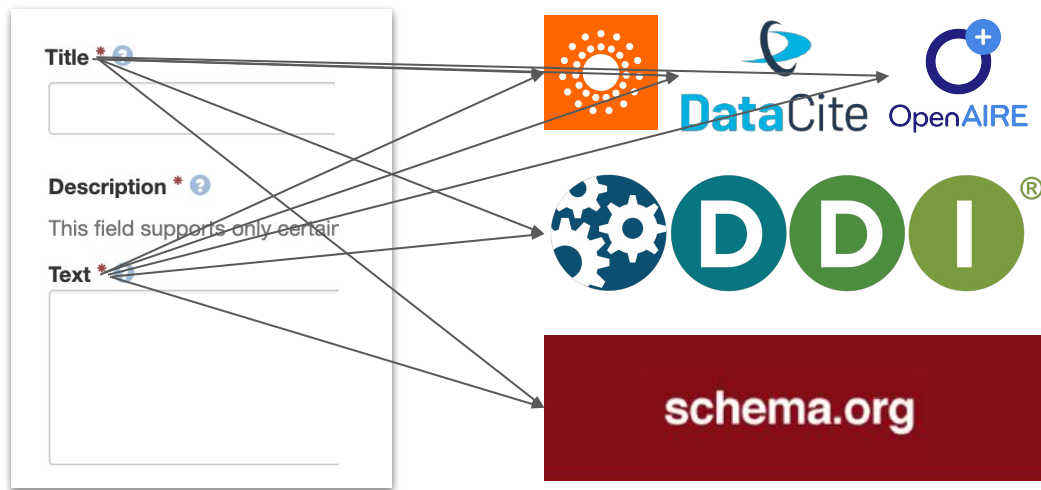
[schema.org](https://schema.org)

What is “good” metadata?

# Metadata that helps others find datasets

Metadata fields that discovery systems use when returning search results

Metadata that leverages how people look for data



What is “good” metadata?

# Metadata that helps others find datasets

Metadata fields that discovery systems use when returning search results

Metadata that leverages how people look for data

“Everything you always wanted to know about a dataset: Studies in data summarisation”

<https://doi.org/10.1016/j.ijhcs.2019.10.004>

Laura Koesten, Elena Simperl, Tom Blount, Emilia Kacprzak, Jeni Tennison. Everything you always wanted to know about a dataset: Studies in data summarisation, International Journal of Human-Computer Studies, Volume 135, 2020, <https://doi.org/10.1016/j.ijhcs.2019.10.004>



<sup>a</sup> The corresponding author at: University of Southampton, University Rd, Southampton SO17 1BJ, UK

<sup>b</sup> Email address: l.koesten@soton.ac.uk (L. Koesten).

<sup>c</sup> In this paper, a "dataset" refers to structured or semi-structured information collected by an individual or organisation, which is distributed in a standard format for instance as CSV files. In the context of search, it refers to the artifacts returned by a search algorithm in response to a user query.

<sup>d</sup> <https://scholar.google.com/>

© 2019 Elsevier B.V. All rights reserved.



What is “good” metadata?

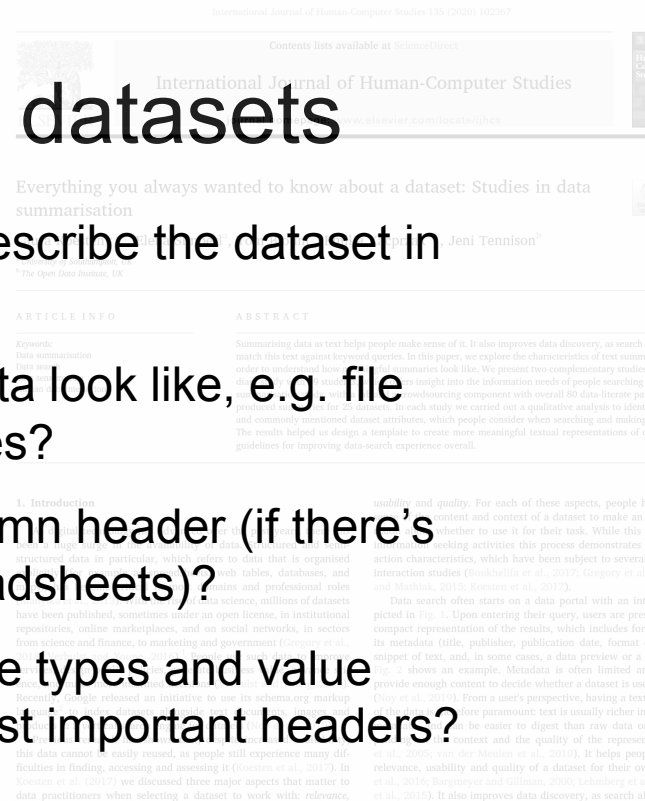
# Metadata that helps others find datasets

Metadata fields that discovery systems use when returning search results

Metadata that leverages how people look for data

1. How would you describe the dataset in one sentence?
2. What does the data look like, e.g. file formats, data types?
3. What are the column header (if there's tabular data/spreadsheets)?
4. What are the value types and value ranges for the most important headers?

<https://doi.org/10.1016/j.ijhcs.2019.10.004>



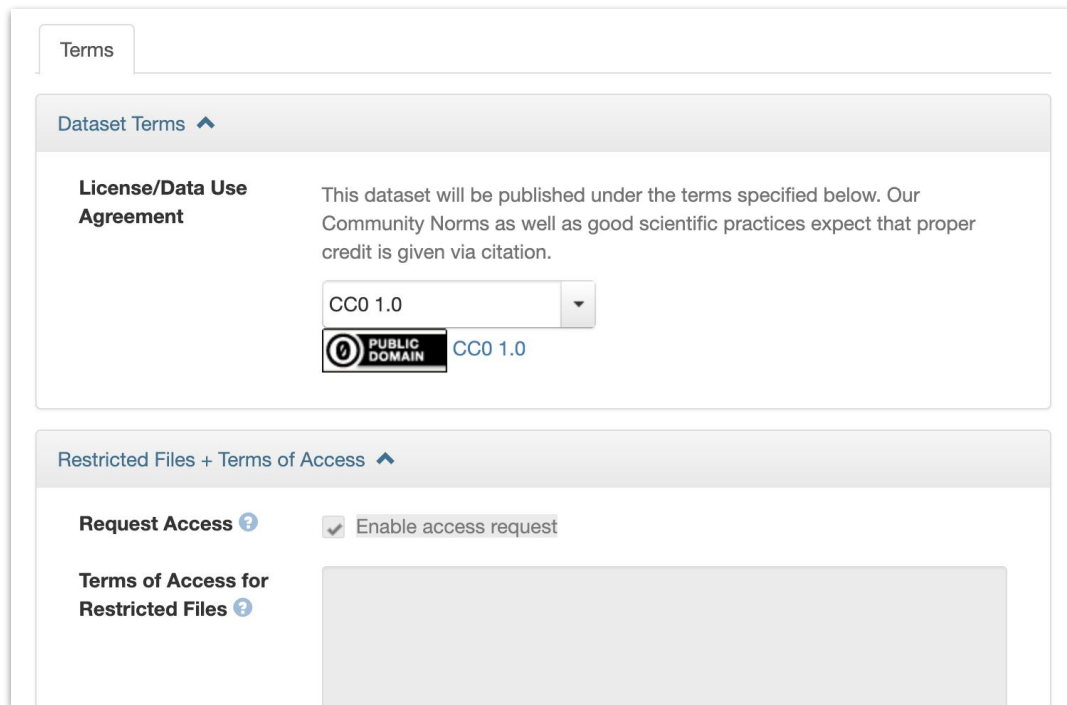
What is “good” metadata?

# Metadata that helps others re-use datasets

Terms of use and  
access metadata, e.g.  
CC0

Links to other resources,  
e.g. “Related  
Publications” metadata

Readme files



The screenshot shows a web form for dataset metadata. At the top, there is a tab labeled 'Terms'. Below it, a section titled 'Dataset Terms' is expanded, showing a 'License/Data Use Agreement' dropdown menu set to 'CC0 1.0'. Below the dropdown is a 'PUBLIC DOMAIN' icon and a 'CC0 1.0' link. The text above the dropdown states: 'This dataset will be published under the terms specified below. Our Community Norms as well as good scientific practices expect that proper credit is given via citation.' Below this, another section titled 'Restricted Files + Terms of Access' is expanded. It contains a 'Request Access' section with a checked checkbox for 'Enable access request' and a 'Terms of Access for Restricted Files' section with a question mark icon.


Terms

Dataset Terms ^

**License/Data Use Agreement**

This dataset will be published under the terms specified below. Our Community Norms as well as good scientific practices expect that proper credit is given via citation.

CC0 1.0

 CC0 1.0

Restricted Files + Terms of Access ^

**Request Access ?** ☒ Enable access request

**Terms of Access for Restricted Files ?**

What is “good” metadata?

# Metadata that helps others re-use datasets

Terms of use and  
access metadata, e.g.  
CC0

Links to other resources,  
e.g. “Related  
Publications” metadata

Readme files

**Related Publication ?**

**Citation ?**

Wilkinson, M., Dumontier, M., Aalbersberg, I. et al. The FAIR Guiding Principles for scientific data management and stewardship. Sci Data 3, 160018 (2016). <https://doi.org/10.1038/sdata.2016.18>

**Identifier Type ?**

doi

**Identifier ?**

doi:10.1038/sdata.2016.18

**URL ?**

<https://doi.org/10.1038/sdata.2016.18>



























What is “good” metadata?

# Metadata that helps others re-use datasets

Terms of use and  
access metadata, e.g.  
CC0

Links to other resources,  
e.g. “Related  
Publications” metadata

Readme files

1 to 10 of 48 Files		 Edit Files ▾	 Download ▾	
	<div><div></div><div><a href="#">README.txt</a> Plain Text - 4.1 KB Published Nov 14, 2023 2 Downloads MD5: 5a1...79e </div></div>		 ▾	 ▾
	<div><div></div><div><a href="#">DP18608.pdf</a> Adobe PDF - 7.1 MB Published Nov 14, 2023 2 Downloads MD5: 871...f17 </div></div>		 ▾	 ▾
	<div><div></div><div><a href="#">9. Généralité Limoges (right).jpg</a> Atlas des Bailliages/ JPEG Image - 36.0 MB Published Nov 14, 2023 2 Downloads MD5: bf0...7fe </div></div>		 ▾	 ▾
	<div><div></div><div><a href="#">9. Généralité Limoges (left).jpg</a> Atlas des Bailliages/ JPEG Image - 35.7 MB Published Nov 14, 2023 2 Downloads MD5: c97...1cc </div></div>		 ▾	 ▾

What is “good” metadata?

# Examples of well-described datasets

International Food  
Policy Research  
Institute

## Citation Metadata ^

### Title ?

Survey on Gendered Constraints to Employment and Entrepreneurship, Lesotho

### Description ?

The dataset comprises a sample of small and medium-sized enterprises in Lesotho interviewed in 2021. The sectors represented are farms; agro-processors; other agricultural businesses; manufacturers; tourism industries & creative industries. The survey instrument was designed to capture information on the experiences of and constraints to employment and entrepreneurship among under-represented groups, specifically: women, youth & persons with disabilities.

What is “good” metadata?

# Examples of well-described datasets

## Murray Research Archive

### Citation Metadata ^

#### Title ?

Oral History of the Tenured Women in the Faculty of Arts and Sciences at Harvard University, 1981

#### Description ?

The purpose of this 1981 study was to document the history of women faculty at Harvard University. Of the then 13 women tenured in the Faculty of Arts and Sciences at Harvard in 1981, 11 professors and the husband of one deceased professor participated in interviews. Nine of these women were of foreign background and one was African American. The interviews focused on the development of their careers, the challenges they faced, their educational and occupational histories, life histories, and demographic information.

The 12 participants took part in intensive, open-ended interviews. The interview followed a rough chronological outline and produced a biographical sketch of each participant, in addition to information about academic work, career, and view of Harvard. The interviews were transcribed and approved by the participants.

The Murray Research Archive has the original tapes and the interview transcripts for all participants. Each professor made her own confidentiality agreements with Dr. Walzer. Because it is not possible to disguise the participants' identities, access to some transcripts is restricted and requires special permission.

**Audio Data Availability Note:** This study contains audio data that have been digitized. There are 78 audio files available.

What is “good” metadata?

# Examples of well-described datasets

Nord University's  
collection in  
DataverseNO

[https://doi.org/10.18710/  
D1HDHZ](https://doi.org/10.18710/D1HDHZ)

## Citation Metadata ^

### Title ?

Video recordings of spawning behavior of Arctic charr; date: 2016-09-19; spawning ground: 3; time: morning; camera no: 6

### Author ?

Egeland, Torvald B. (Nord University) - ORCID: [0000-0003-4768-1678](https://orcid.org/0000-0003-4768-1678)  
Folstad, Ivar (UiT The Arctic University of Norway) - ORCID: [0000-0003-1472-5128](https://orcid.org/0000-0003-1472-5128)  
Nordeide, Jarle Tryti (Nord University) - ORCID: [0000-0003-2315-3635](https://orcid.org/0000-0003-2315-3635)

### Description ?

This dataset contains video recordings of spawning behavior of Arctic charr (*Salvelinus alpinus*) in Lake Fjellfrøsvatnet (69°08'N 19°34'E), Troms, Northern Norway. The recordings were made with camera no. 6 at spawning ground 3 in the morning of 19 September 2016.

A description of the data structure and format is gathered in the documentation dataset: Egeland, Torvald B.; Folstad, Ivar; Nordeide, Jarle Tryti, 2021, "Documentation of the Spawning behavior of Arctic charr video collection", DataverseNO, <https://doi.org/10.18710/VIQB0B>

**This dataset is part of a larger collection: Egeland, Torvald B.; Folstad, Ivar; Nordeide, Jarle Tryti, 2021, "The Spawning behavior of Arctic charr video collection". DataverseNO. <https://doi.org/10.18710/HTM6-F146>.**

What is “good” metadata?

# Examples of well-described datasets

Nord University's  
collection in  
DataverseNO

[https://doi.org/10.18710/  
D1HDHZ](https://doi.org/10.18710/D1HDHZ)

## Citation Metadata ^

### Keyword ?

Arctic charr  
Salvelinus alpinus (Global Biodiversity Information Facility (GBIF)) <https://www.gbif.org/species/4284021>  
spawning behavior  
vibrational communication  
mate guarding  
sperm competition  
gamete synchrony  
courtship  
sexual selection  
female choice  
egg predation  
filial cannibalism  
protective behaviour  
male-male competition

### Related Publication ?

Brattli, MB, Egeland, TB, Nordeide, JT, Folstad, I. Spawning behavior of Arctic charr (*Salvelinus alpinus*): Spawning synchrony, vibrational communication, and mate guarding. *Ecol Evol.* 2018; 8: 8076– 8087. doi: 10.1002/ece3.4277 <https://doi.org/10.1002/ece3.4277>

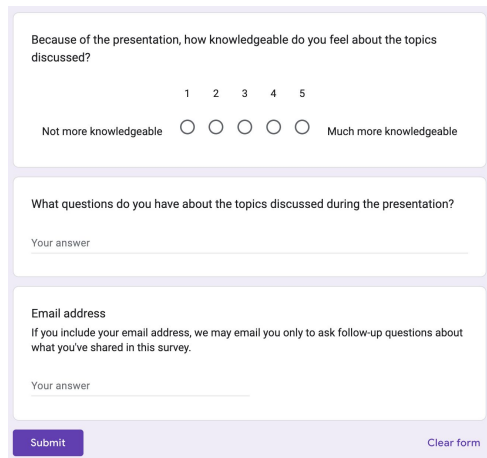


# Questions, survey, upcoming webinars

## Q&A

### 3-question webinar survey

<https://forms.gle/Hxyy76fYFJGjFJG78>



Because of the presentation, how knowledgeable do you feel about the topics discussed?

1 2 3 4 5

Not more knowledgeable ☐ ☐ ☐ ☐ ☐ Much more knowledgeable

What questions do you have about the topics discussed during the presentation?

Your answer

Email address

If you include your email address, we may email you only to ask follow-up questions about what you've shared in this survey.

Your answer

Submit Clear form

## Upcoming webinars

December 4, 2023

Intermediate Curation on the Dataverse Platform

January 31, 2024

Introduction to Dataverse APIs

2024

Introduction to Metadata in Dataverse Repositories: For Installation Administrators