

Dataverse for data sharing: An introduction

**Chantal Ripp & Jane Fry
June 16, 2021**

Agenda

- What is Dataverse
- Benefits of Dataverse
- Using Dataverse for your research data
- Resources – many valuable ones!



Information about Dataverse in the following slides is taken from:

<https://learn.scholarsportal.info/all-guides/dataverse/faq/>

What is Dataverse?

- A research data platform for sharing, discovering and preserving research data
- A repository for datasets
- For University of Ottawa and Carleton University
 - Students, researchers, faculty and staff
- In partnership with [ScholarsPortal](#)
 - a service of the Ontario Council of University Libraries ([OCUL](#))

Key benefits

- Robust Canadian storage network
 - Hosted in University of Toronto
 - On Canadian servers
 - Long-term access and preservation
- Data management
 - Meets one of the requirements of the [Tri-Agency Research Data Management Policy](#)
 - Data deposit
 - Data dissemination
 - Open access (when permitted)
- Access control
 - You determine who has access to your data
 - IP controlled

More key benefits

- DOIs and citation standards support
 - DOIs - digital object identifiers
 - An international standard for simple, persistent identification for access to your data on the web
 - Automatically provided
 - Some journals are requiring DOIs for datasets associated with a journal article
 - This is often just a replication dataset
 - Mandatory for major Economics journals
 - Other journals will be following soon
 - Standard and rich metadata fields
 - Allow for maximum discoverability for your datasets
 - Dataverse automatically generates a data citation
 - Can be used in any research publication
 - Citation can be downloaded in various formats
 - Simply click the “Cite Data” button

More ...

- Discoverability, recognition and increased visibility
 - Share your data with a global research community
 - Researchers will get recognition and proper academic credit for their scholarly work
 - Indexed in
 - re3data.org
 - Registry of research data repositories
 - More than 2450 research data repositories
 - The most comprehensive source of reference for research data infrastructures globally
 - [DataCite's Global Search Service](https://datacite.org)
 - Web of Sciences' Data Citation index
 - Dataverse is regularly crawled by Google
 - *Note: only published and public files have their metadata harvested*

What File Size?

- Dataverse supports file sizes less than 3GB
 - If larger, you can zip the files
 - If it is a tabular file, you can break it into smaller files (with an accompanying *Readme file*)
- Recommended to keep files to less than 500MB for tabular data files that require more time to upload
 - For file sizes larger than 500MB, contact Dataverse support of UoO/Carleton
- No limit on total size of all files you can upload for a single dataset
 - If more than 10GB in total file size, contact Dataverse support of UoO/Carleton
- See [Scholars Portal Dataverse Guide](#) for more info
- If you are working with big data
 - Check out [FRDR](#) (Federated Research Data Repository)

What types of files?

- Any file type can be published
 - Shapefiles
 - Images
 - Compressed files
 - Stata
 - SPSS
 - Rdata
 - .wav
 - .jpeg
 - .png
 - Audio files
 - GIS data
 - Csv
 - Tab
 - Ascii
 - ...

Collaboration

- Can I collaborate with colleagues in a Dataverse?
 - Yes! Has to be hosted with a subscribing institution.
 - Dataverse supports groups of researchers working on the same research project.
 - Very useful when working with researchers at other institutions.
 - Can assign different group permissions.
 - Can assign different account roles.
 - Can also share your dataset without publishing it.



Source: <https://blog.jostle.me/blog/whv-collaboration-is-important>

Safety of your data

- Scholars Portal has backup copies
- Can be restricted to authorized users only
 - Individual IPs
 - Specific IP groups, eg., your institution only
- Can keep a copy of your research data in Dataverse
 - As a backup copy that you are keeping elsewhere

Dataverse vs Dataset

- A Dataverse
 - Structure is much like the structure of Windows Explorer
 - main folders/dataverses
 - sub-folders/sub-Dataverses
 - It is a folder/bucket that can have other Dataverses in it
 - can also have datasets in it
- A dataset
 - must be in a Dataverse
 - contains data files and metadata about those data files

Dataverse vs a Website

	<i>Dataverse</i>	<i>Website</i>
<i>Sustainability</i>	No worries	Keep an eye on it (not 'one and done')
<i>Discoverability</i>	Maximum	Need proper SEO (Search engine optimization)
<i>Security</i>	Can't be hacked	Hopefully won't be hacked, but ...

- Can use both together
 - Put your dataset in Dataverse and the DOI on your website
 - Then your research project can have its own branding

More FAQs

- Can I test it out to see how it works
 - Sure!
 - <https://demodv.scholarsportal.info/dataverse/sp?q=demo+dataverse>
 - Try it out to see if you like its features
- What is the cost to use Dataverse?
 - None!
 - University of Ottawa and Carleton belong to the consortium that supports it (OCUL)
 - Therefore, no direct cost to researchers
 - Must use your institutional email address

Still more!

- How much data can I download from Dataverse?
 - No restrictions
 - As much openly available data as you want
- Can I upload .zip bundles?
 - Yes!
 - Dataverse will automatically unpack your single zip package and present individual files
- How many files can be in my .zip bundle?
 - Less than 500 because of additional processing that is required if there are more

Dataverse limitations

- Dataverse may not be suitable if your data
 - Requires
 - high volume computational processing infrastructure to share or access
 - an automatic embargoed period
 - You would have to manually embargo it
 - Is sensitive data
 - Dataverse cannot accept any data that contains confidential or sensitive information

Making your data FAIR

- Good research data practice
 - data management and stewardship
- **F**indable
 - discoverable
- **A**ccessible
 - Data can be downloaded
- **I**nteroperable
 - Data can be integrated with other data and can be analyzed
- **R**eusable
 - Metadata and data are well-described
 - They can be replicated and/or combined in different settings
- Here's a checklist to follow:
 - [FAIR-Aware](#)
 - [FAIR-Aware Tool video](#) (1.21 min)



Creating Your Account

Options 1

Institutional Single Sign on

- sign-in using the same username and password you use to access your institution's online services

The screenshot shows the 'Log In' section of the Scholars Portal Dataverse. Under 'Your Institution', a dropdown menu is open for 'University of Ottawa', listing various institutions. Below this, there are 'Username/Email' and 'Sign' buttons. To the right, the 'Account Information' form is displayed, showing the user's name (Chantal Ripp), email (cripp@uottawa.ca), and affiliation (University of Ottawa). Below the form, there is a 'General Terms of Use' section with a checkbox that is checked, indicating acceptance of the terms. At the bottom, there is a 'Login' button and a password field with a 'Forgot your password?' link.

Search ▾ User

Log In

Log in or sign up with your institutional account — more information about account creation. Leaving your institution? Please contact Dataverse Support for assistance.

Your Institution

University of Ottawa

Please select... Continue

MacEwan University
McGill University
McMaster University
OCAD University
Polytechnique Montréal
Queen's University
Université du Québec à Chicoutimi
University of Alberta
University of British Columbia
University of Calgary
University of Guelph
University of Ottawa
University of Regina
University of Toronto
University of Victoria
University of Waterloo
University of Western Ontario
University of Windsor
UQAM | Université du Québec à Montréal
Wilfrid Laurier University

Institution not listed?

Username/Email Sign

Scholars Portal Dataverse

Search ▾ User Guide Support English ▾ Log

Account Information – This information is provided by your institution and will be used to create your Dataverse account.

Account Information

Name Chantal Ripp

Email cripp@uottawa.ca

Affiliation University of Ottawa

General Terms of Use ⓘ

Service.

By way of example, and not as a limitation, you shall not (or permit others to) either (a) take any action or (b) upload, download, post, submit or otherwise distribute or facilitate distribution of any content using any communications service or other service available on or through the Service, that:

- infringes any patent, trademark, trade secret, copyright, right of publicity or other right of any other person or entity;
- is unlawful, threatening, abusive, harassing, defamatory, libelous, deceptive, fraudulent, invasive of another's privacy, tortious, obscene, offensive, or profane;
- constitutes unauthorized or unsolicited advertising, junk or bulk e-mail ("spamming");
- contains software viruses or any other computer codes, files, or programs that are designed or intended to disrupt,

I have read and accept the Dataverse General Terms of Use as outlined above.

This page displays in the default language of your browser and device. To see this page in another language, [change your language settings](#).

cripp@uottawa.ca

.....

[Forgot your password?](#)

Login

Creating Your Account

Options 2

Username/password

- If you are affiliated with an institution, please use your institutional e-mail address. Users that are not affiliated with a participating institution will have limited access. Refer to [SP FAQ](#) for more information.

The screenshot shows the 'Scholars Portal Dataverse' account creation page. At the top, there is a navigation bar with 'Search', 'User Guide', 'Support', 'English', and 'Log In'. Below this is a blue banner with the text: 'Create Account - Why have a Dataverse account? To create your own dataverse and customize it, add datasets, or request access to restricted files.' The main form is titled 'Account Information' and includes a note: 'You can also create a Dataverse account with one of our other log in options.' The form fields are: 'Email' (chantal.ripp@uottawa.ca), 'Affiliation' (University of Ottawa), 'Username' (with instructions: 'Create a valid username of 2 to 60 characters in length containing letters (a-z), numbers (0-9), dashes (-), underscores (_), and periods (.).'), 'Password' (with instructions: 'Your password must contain: At least 6 characters (passwords of at least 20 characters are exempt from all other requirements); At least 1 character from each of the following types: letter, numeral'), 'Retype Password', 'Given Name' (Chantal), 'Family Name' (Ripp), and 'Position'. At the bottom, there is a 'General Terms of Use' section with a scrollable area containing the 'Acceptance of Terms' text and a checkbox for 'I have read and accept the Dataverse General Terms of Use as outlined above.' The 'Create Account' button is at the bottom left.

Publishing Datasets in Dataverse

The screenshot displays the Scholar Portal Dataverse interface. At the top, the header includes the site name "Scholar Portal Dataverse", a search bar, and navigation links for "About", "User Guide", "Support", and "Sign Out". Below the header, the user's profile is visible, showing "Scholar Portal Dataverse - Account" and a "Log Out" button. The user's name "A. B. B." is displayed in a circular profile picture. The "My Data" section includes tabs for "Publications", "Associated Citations", and "API Feeds". A message states: "Here are all the datasets, datasets, and files you have access to. You can filter through them by publication status and more." Below this is a search bar for "Search my data" and a "Q, Find" button. The left sidebar contains filters for "Dataset ID", "File", "Publication Status", "Role", and "Contributor". The main content area shows "1 to 4 of 4 Results" with a list of datasets. The first dataset is "Sample Survey Data" (Feb 14, 2016) with buttons for "View", "Download", "Share", and "Download". The second is "Scholar Portal 2016, 'Scholar Portal Data'" (Feb 10, 2016). The third is "SP Test Dataset" (Jan 10, 2016). The fourth is "SP Test Dataverse" (Jan 10, 2016). At the bottom left, there are logos for "SCREENSHOT" and "MATIC".

Modifying Terms

Files Metadata Terms Versions

Edit Terms Requirements

Terms of Use ^

Terms of Use ^

Waiver ?

Datasets will default to a [CC0 public domain dedication](#) . CC0 facilitates reuse and extensibility of research data. Our [Community Norms](#) as well as good scientific practices expect that proper credit is given via citation. If you are unable to give datasets a CC0 waiver you may enter custom Terms of Use for datasets.

Yes, apply CC0 - "Public Domain Dedication"

No, do not apply CC0 - "Public Domain Dedication"

Terms of Use ?

If you are unable to use CC0 for datasets you are able to set custom terms of use. Here is an example of a [Data Usage Agreement](#) for datasets that have de-identified human subject data.

Want to apply a different licence? Click "Edit Terms Requirements"

Add another open licence under "Terms of Use"

Dataverse and FAIR

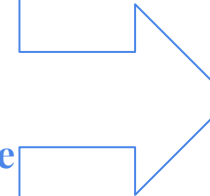
Data citation with persistent identifier (DOI)

Data files

Metadata

Terms/Data Licence

Dataset version



Scholars Portal **Dataverse** Search - User Guide Support English - Log In

COVID-19 Destigmatizing Chinese Communities Project (York University)

Scholars Portal Dataverse > York University Dataverse > Aaida Mamuji Dataverse > COVID-19 Destigmatizing Chinese Communities Project >

COVID-19 Destigmatizing Chinese Communities Project: Greater Toronto Area Interview Data

Version 1.0

Mamuji, Aaida; Rozdilsky, Jack L.; Lee, Charlotte T.; Mwarumba, Ngiki; Tubula, Martin; Chu, Terri. 2020. "COVID-19 Destigmatizing Chinese Communities Project: Greater Toronto Area Interview Data", <https://doi.org/10.5603/SP2/GC3QZM>, Scholars Portal Dataverse, V1

Cite Dataset - Learn about Data Citation Standards.

Access Dataset - Contact Owner Share

Dataset Metrics 99 Downloads

Description

The purpose of this phase of the research was to explore how members of Chinese diaspora communities in large global urban centers, namely the Greater Toronto Area (GTA) and the Greater Nairobi Metropolitan Region, are coping and reacting to COVID-19. To be included, participants must have identified as being ethnically Chinese, be over 18 years of age, and have lived in the GTA or Nairobi Metropolitan for at least 1 year prior to the start of the pandemic. A team of researchers was assembled from York University (Canada), Ryerson University (Canada), the University of Nebraska - Omaha (United States), and Atoscope Consulting (Kenya) to conduct a series of interviews in either English, Cantonese or Mandarin. Audio recordings of the interviews were transcribed, and where necessary translated into English. This dataset contains the de-identified interview data collected in the GTA.

Subject Social Sciences

Keyword COVID-19 (Disease), Sigma (Social psychology), Chinese diaspora, Emergency management

Related Publication Mamuji, A., Rozdilsky, J., Lee, C., Mwarumba, N., Tubula, M., and Chu, T. (2020). Expanding the Narrative on Anti-Chinese Sigma During COVID-19: Lessons on Complexity & Capacity in Toronto and Nairobi. (York University Disaster and Emergency Management Technical Report #009). doi: 10.6084/m9.fgshare.12759512

Files Metadata Terms Versions

Search this dataset. Find

Filter by File Type: All Access: All Sort

1 to 10 of 86 Files Download

9. Toronto Key Informant Interview Guide.pdf	Adobe PDF - 55.4 kB - Sep 22, 2020 - 10 Downloads	MD5: db8fae703951e8905f15a2c20527e01	Download
Interview 01 - Public Transcript.rtf	Rich Text Format - 1.1 MB - Sep 22, 2020 - 3 Downloads	MD5: 0b7e1e729c30e0f039a10e2ab0a9d6	Download

Resources

- [Guide to depositing data in uOttawa Dataverse](#)
- [Carleton University Dataverse](#)
- [Scholars Portal Dataverse](#)
 - [Dataverse overview: Finding and exploring data](#) (7 min video)
 - [Creating, editing and Publishing a Dataverse](#) (6 min video)
 - [Publishing datasets in Dataverse](#) (6 min video)
- [Portage](#)
 - [Can I share my data?](#)
 - [Documentation and supporting material required for deposit](#)
 - [Recommended repositories for COVID-19 research data](#)

More resources

- [Dataverse North metadata best practices guide](#)
- [Introduction to the Harvard Dataverse Repository](#) (1 hr 36 min video)
- [FAIR](#)
 - [FAIR-Aware Tool video](#) (1.21 min)
 - [FAIR-Aware Tool](#)
 - [How FAIR are your data?](#)

Thank You!



Source: <https://workingnation.com/2020-presidential-debate-future-of-work-questions/>

Contact Information

Chantal Ripp

Research Data Management Librarian (interim) |

Bibliothécaire de gestion des données de recherche (intérim)

309E, Bibliothèque Morisset Library

University of Ottawa | Université d'Ottawa

chantal.ripp@uottawa.ca

Jane Fry

Data Services Librarian

MacOdrum Library, Carleton University

jane.fry@carleton.ca

<https://library.carleton.ca/find/data>



Carleton University
MACODRUM LIBRARY