


# Best Practices for Managing Data in your Research

 **Sherman  
Centre**  
for Digital Scholarship

Wednesday, October 1, 2025

10:30am – 11:30am **(Online)**





## Land Acknowledgement

McMaster University is located on the traditional territories of the Mississauga and Haudenosaunee Nations. Settlers have responsibilities under the Silver Covenant Chain Wampum, part of the 1764 Treaty of Niagara.

Georgia Kirkos, "Cootes Trail," October 29, 2021, McMaster University, Hamilton, Ontario, Canada  
<https://brand-resources.mcmaster.ca/asset-bank/action/viewAsset?id=40841&index=14&total=34&view=viewSearchItem>





## Code of Conduct, Session Recording, & Privacy

**Code of Conduct:** The Sherman Centre and the McMaster University Library are committed to fostering a supportive and inclusive environment for its presenters and participants. Please refer to our code of conduct webpage for more information: [scds.ca/events/code-of-conduct](https://scds.ca/events/code-of-conduct)

**Session Recording:** This session is being recorded with the intention of being shared publicly via the web for future audiences. In respect of your privacy, participant lists will not be shared outside of this session, nor will question or chat transcripts. Questions asked via the chat box will be read by the facilitator without identifying you. Note that you may be identifiable when asking a question during the session in an audio or visual format.



## Certificate Programs

### The Sherman Centre for Digital Scholarship Certificate of Attendance

The Sherman Centre's certificate program recognizes attendance at our workshops. It complements degree training, supports the development of critical competencies in data analysis, research data management, and digital scholarship, and formalizes core skills fostered by our workshops.

Participants are invited to collect **seven** workshop points to receive a certificate of attendance. To verify your participation in today's workshop, we will provide a code and additional instructions at the end of the session.

You can learn more about the certificate program at [scds.ca/certificate-program](https://scds.ca/certificate-program)

### The Canadian Certificate for Digital Humanities

This workshop is also eligible for the Canadian Certificate for Digital Humanities. To learn more about the certificate, visit [ccdhhn.ca](https://ccdhhn.ca). You can also contact local liaison Alexis-Carlota Cochrane at [scds@mcmaster.ca](mailto:scds@mcmaster.ca).

## Fall 2025: Upcoming Workshops

### Data Analysis Support Hub

**September 24:** Create Simple Maps using Google Maps and Microsoft Excel

**October 9:** Getting Started with Linear Regression in R

### Digital Research

**October 21:** Establishing and Maintaining Researcher Profiles

**November 19:** Creating Compelling Research Impact Visualizations

### Research Data Management

**November 19:** Data Management Plan Bootcamp (Virtual)

**November 25:** Data Management Plan Bootcamp (In-Person)

**January 27:** Streamline Your Research Materials Photos with Tropy

### Do More with Digital Scholarship

**October 8:** Visions of Generative AI: Historical and Ethical Dimensions of Visual Media Literacy in the Era of AI

**October 23:** Creating High-Quality Documents with LaTeX

**November 26:** Making and Querying Databases in SQL with DuckDB

Register for Upcoming Workshops: <https://u.mcmaster.ca/scds-workshops>

Library



# Research Data Management Services

PhD in Anatomy  
and Cell Biology,  
member of  
MREB - can help  
with data  
management for  
sensitive data



Isaac Pratt, PhD

MA in Media  
Studies, excited  
about data justice,  
community research,  
+ connecting with  
curious disciplines!



Danica Evering, MA

# Let us know where you're from and one thing you're hoping to learn!

*Shout it out! Get in the comments!*



# Outline



## Best Good Enough Practices for Managing Data in Your Research

- Everyone has data (even if you don't think you do!)

- Don't be like Dave or Alice

### 1. Make a plan

### 2. Organize & document what you're doing consistently

### 3. Store & back up data securely

### 4. Get ready for archiving & sharing



**Everyone has data  
(even if you don't  
think you do!)**

# What counts as “research data”?

Information or materials that are used as evidence in research and scholarly work.

scds.ca

scds@mcmaster.ca

3	2	2023-05-09 10:05	MCG_001	0.20001	0.07100	0
4	3	2023-05-09 10:05	MCG_001	0.394381	7.452087	0
5	4	2023-05-09 10:06	MCG_001	0.444642	0.858403	0
6	5	2023-05-09 10:06	MCG_001	0.684542	6.038661	0
7	6	2023-05-09 10:07	MCG_001	1.332544	14.29202	0
8	7	2023-05-09 10:07	MCG_001	1.431062	14.00633	0
9	8	2023-05-09 10:08	MCG_001	1.114414	7.702386	0
10	9	2023-05-09 10:08	MCG_001	0.701831	13.06353	0
11	10	2023-05-09 10:09	MCG_001	0.97043	38.50039	0
12	11	2023-05-09 10:09	MCG_001	0.691576	4.583931	0
13	12	2023-05-09 10:10	MCG_001	0.483963	13.99204	1
14	13	2023-05-09 10:10	MCG_001	0.354242	10.55881	1
15	14	2023-05-09 10:11	MCG_001	0.199756	0.725879	1
16	15	2023-05-09 10:11	MCG_001	0.102357	0.658201	1
17	16	2023-05-09 10:12	MCG_001	0.16149	2.02036	1
18	17	2023-05-09 10:12	MCG_001	0.085709	0.846466	1
19	18	2023-05-09 10:13	MCG_001	0.024841	0.225617	1
20	19	2023-05-09 10:13	MCG_001	0.098602	0.79793	1
21	20	2023-05-09 10:14	MCG_001	0.1185	0.140091	1
22	21	2023-05-09 10:14	MCG_001	0.18612	0.10714	1
23	22	2023-05-09 10:15	MCG_001	0.093189	0.716232	1
24	23	2023-05-09 10:15	MCG_001	0.176302	0.826936	1
25	24	2023-05-09 10:16	MCG_001	0.098325	0.416565	1
26	25	20				
27	26	20				
28	27	20				
29	28	2023-05-09 10:18	MCG_001	0.14202	0.000004	1
30	29	2023-05-09 10:18	MCG_001	0.000000	1.510000	1

Data for Blanding's turtle (*Emydoidea blandingii*)  
behavioural states from multi-sensor biollogger data  
McMaster Dataverse

# What counts as “research data”?

Information or materials that are used as evidence in research and scholarly work.

Clockwise: Christina WOC in Tech via Unsplash; Rachel Glaves, "Collage party at Million Fishes," CC-BY 2.0, Microsoft Stock Images.



**surveys, interviews,  
and focus groups**



**research-creation  
and community art**



**battery charging data  
for training ML models**



**photos and metadata  
from archival records**



## I lost 2 years' worth

The title says it all..

I don't know how to process this rig

I've been doing astronomy research  
on the computer in my office, until :

At the time I thought it was just a m  
and I actually used this time to take

Yesterday they got back to me, and  
campus, and somehow something v  
scared" mode, and they told me the  
do a clean restart.

They haven't done it yet, because th  
do it, because I'd lose 2 years' worth

I already know people are gonna sa  
but it just never crossed my mind (fi  
scheduled for next Wednesday). I ha  
large chunk of my work will definite

The projects that I was working on v

[Read more](#) ▾



475



80



s

## Lost a TON of data for my PhD and my PI is pissed. Does anyone have advice?

Need Advice

### HELP! LOST DATA FOR MY PHD

Alright, so, my Google Drive wasn't syncing 1,296 files. I saw that you can re-sync by deleting the DriveFS folder in the local Google folder. Did that and files started to upload, great. HOWEVER, all the previous files that didn't sync are now gone. I basically lost 8 months of work. I can't simply redo it because I work with mice and tumors that take a long time to work with.

I cut and pasted the old cache file to a new folder but the file is so small it can't possibly have all my data.

There's also the fact that I cut and pasted instead of using the "recover" tool in the recycle bin. Could that be an issue?

And no, there's nothing in my drive trash bin or anything like that. The location address just simply does not exist.

I found that someone downloaded their new Google drive then placed in the old cache then downloaded everything that was lost, then got in the new cache and simply added the lost files in. Again, the old cache is so tiny though.

I'm using drill disk right now. Do you think that will work?

oy not checking if the

nce I just defended my PhD today

ata, wrote the code, did all the

I so they had to take my PC and wipe it  
I the files back.

d send to me quickly.

cup is done, and then I give it to them.

), I just saw the tick mark, and in my  
handful of small files did get backed up,

had a mini heart attack once I realized  
blished works, they were lost too.

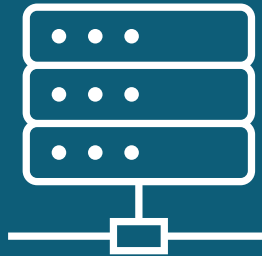
t wasn't published yet. Also, I actually  
rom memory, and thankfully I  
y work, but managed to replicate the

it.

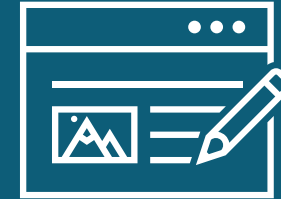
# *Is your data vulnerable?*



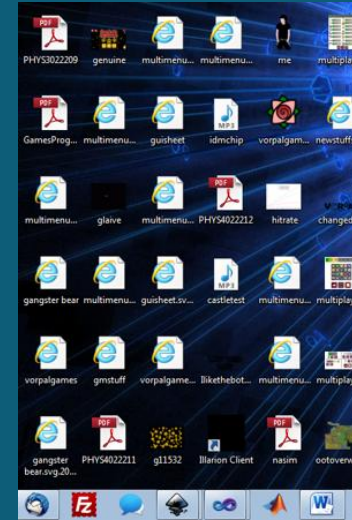
researcher/  
research team



research data



publication/thesis



# Research Data Management activities



## Planning

- Creating a **data management plan**
- **Finding** existing data
- Obtaining systems and software for managing data



## Data collection and analysis★

- **De-identification** of information
- **Organizing** data efficiently
- **Documenting** data effectively
- **Storing** data securely



## Knowledge Sharing

- Preparing data for **archival** and sharing
- **Publishing** data and providing access to participants and other researchers



Home • Services • Data Analysis Support Hub (DASH)

## Data Analysis Support Hub (DASH)

<https://library.mcmaster.ca/services/dash>

Data Analysis Support Hub (DASH)

Dash Resources

DASH GitHub Repository

Workshops

Contact DASH

### Data Analysis Support Hub (DASH)

Software and data analysis?

### AskResearch

AskResearch is a collaborative network of support units providing digital research services to the McMaster research community.

AskResearch is coordinated by the Digital Research Commons.

About Us

# What Research Data Management isn't

<https://askresearch.mcmaster.ca>



What is digital research?

An intellectual process or creative



Digital Research Kickstarter

The online Digital Research



Digital Research Resources

Explore our growing set of digital



Consults

Have questions about what digital research support is available?

Libraries

**Don't be like “Dave”  
and “Alice”**



## Let's look at an example:

**“Dave”** is a graduate student working in Biomedical Science, using x-ray imaging. Dave's data has 3 parts:

- **Image files** – x-ray scan images, microscope images.
- **Experiment files** – instrument metadata, records, scripts, etc
- **Measurement data files** – spreadsheet files

Dave's data is stored in a few places:

- Image files are large (~10 TB) and stored wherever there is space - on lab computers and miscellaneous external hard drives
- The other files are smaller (~10 GB) and stored on a personal laptop backed up in the cloud
- Data is not consistently documented.



# What went wrong

One of the external drives fails and is unrecoverable. A segment of the data that had yet to be analysed was only stored on that drive.

- The data loss is not discovered for several weeks.
- **There is no back up of this data** because there was no documentation on where data was stored

This leaves Dave with two choices:

- Recollect that data which would add several months of delay to the degree
- Publish what he can, even though the power of the study has been reduced



# Let's look at an example:

**“Alice”** is a graduate student working in Media Studies studying the practices of artist-researchers.

Alice's data is made up of:

- **Audio Files** – Recorded interviews with artists.
  - Some interviews are in a studio, one is in a **café**. Interviews are recorded on a Zoom device with notes added into a notebook. These are eventually moved from the recorder to her **hard drive**.
- **Handwritten Notes** – Written during work at a socially engaged arts organization.
  - On a desk in Alice's apartment.
- **References** – Scholarly writings for context
  - PDFs on hard drive, books are on bookshelf.

# What went wrong for Alice

- When Alice finally sits down to listen to her interviews, one of them is unusable. Instead of the responses, she instead hears the chaos of the coffee shop, the steamer hissing, the portafilter slamming, people catching up.
- This leaves Alice with two choices:
  - Reach out to the artist to rearrange an interview, knowing the time involved and their responses will be trying to approximate something they've already said.
  - Publish the other interviews without that participant's contribution.





# What could you do to be better?

<b>Make a plan for data</b>	Prepare for future challenges and problems.
<b>Organize and document data consistently</b>	Save time and resources in the future.
<b>Store and back-up data securely</b>	Avoid loss of data from theft, corruption, or failure of storage devices.
<b>Ensure data is ready for archival and sharing</b>	Increase the accessibility of research and allow others to reproduce and use results.

# 1. Make a plan

- A **Data Management Plan (DMP)** is your plan for how you will create, store, organize, document, secure, preserve, and share your research data.
- A document which speaks to the management of data both **during** the active phases of your research and **after** the completion of the research project.







→ Gates  
Portes  
→ Oversized Baggage  
Bagages hors format

Check in  
Enregistrement  
**3**

**4**

OS7  
Oversized Baggage  
Bagages hors format



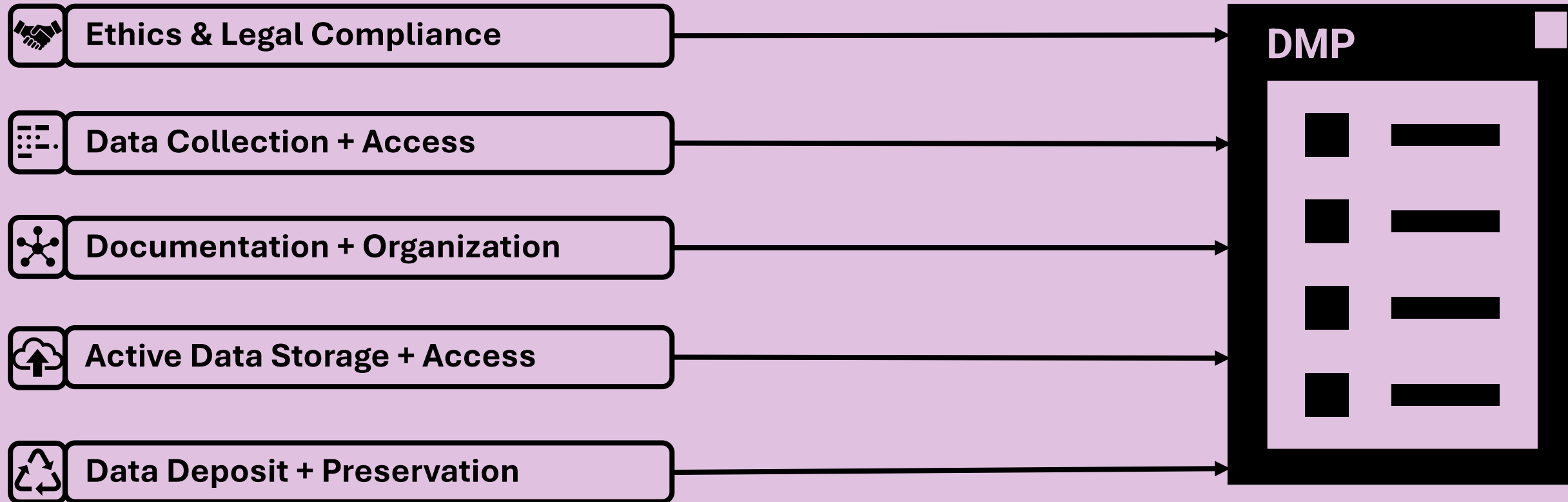
# Why create a DMP?

- A **living document** – something you'll work with, adapt, and change through your research.
- Create it at the **start** of your research - avoid pitfalls and problems before they occur.
- Prepare for **future** stages of research including potential data sharing (if desired).
- Research is a **team** effort – collaborate on your DMP.
- Many research **funders** require grant applicants to submit a DMP – including the Tri-Agencies (NSERC, CIHR, SSHRC – started 2022), NIH, and others.





# What goes in a Data Management Plan?

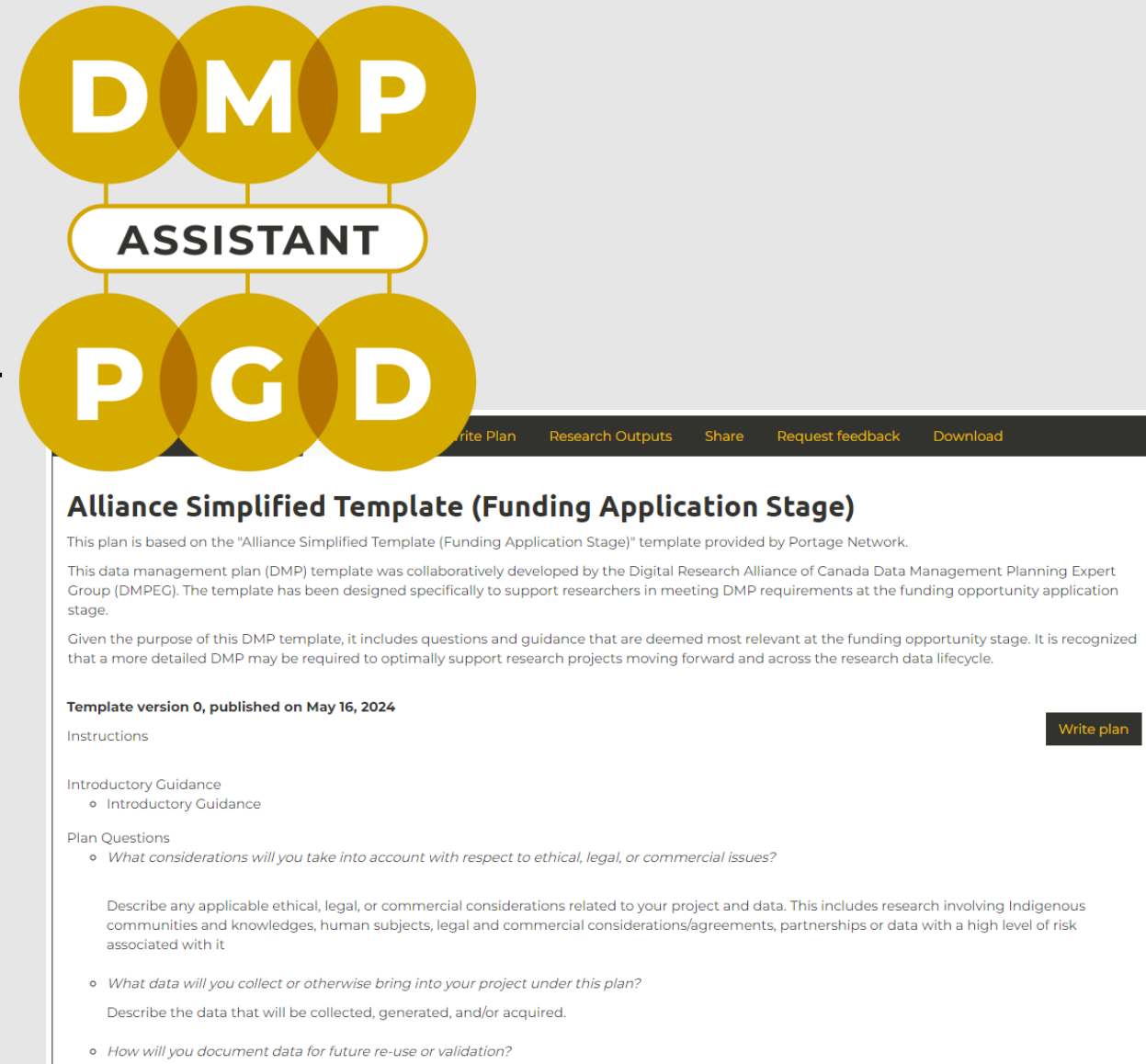


[scds.ca](https://scds.ca)

[scds@mcmaster.ca](mailto:scds@mcmaster.ca)

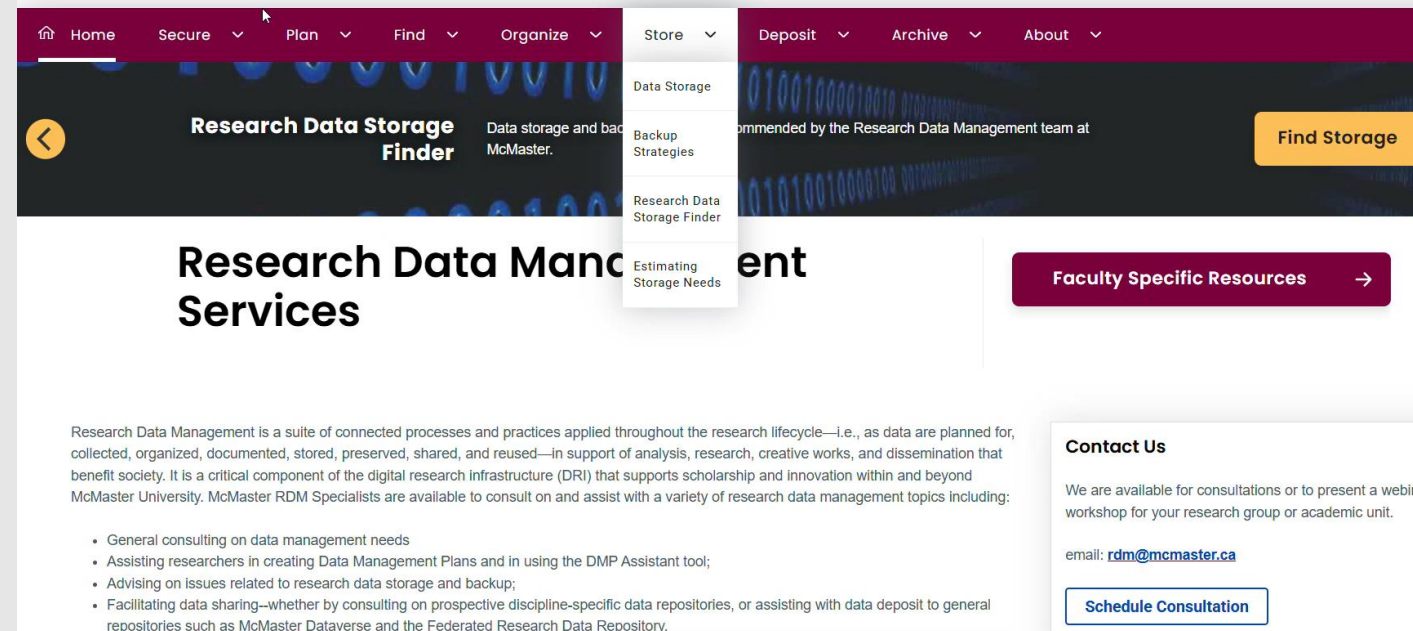
# DMP Assistant

- A web-based, bilingual data management planning tool
- Walks you through relevant questions for data management
- McMaster-specific guidance
- Exportable data management plans
- Send to RDM Services through the “Request feedback” tab!
- Access at [dmp-pgd.ca/plans](https://dmp-pgd.ca/plans)



# DMP Database

- DMPs can vary across disciplines, methodologies, and data types.
- 450+ example DMPs from resources across the world.
- Filters by McMaster Department for selected plans
- Search by field, location, funder, keyword, and more.
- Link to our database at [rdm.mcmaster.ca/dmps](https://rdm.mcmaster.ca/dmps)



## How and where will your data be stored and backed up during your research project?

It is expected that the study will generate, at most, a terabyte of digital data for one complete copy of the dataset. This digital data volume is well within the existing storage capacities of this research group. Similarly, the paper documents generated (consent forms, CRFs, PHINs) are expected to fill 3 large, locked filing cabinets, which the PI already has...Approximately 210 boxes, or less than 13 cubic feet, will be required to store the biological samples in the freezer (-80°C).

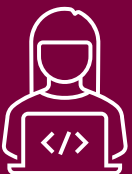
Hay, J. L., Ducas, J., & Duhamel, T. (2023). Women's Advanced Risk-Assessment in Manitoba (WARM) Hearts. Zenodo.

<https://doi.org/10.5281/zenodo.11074529>

# What help is available

- **Before Writing:** Review our Data Management Plan Database for examples in your discipline.
- **Writing:** Data Management Plan Bootcamps
  - November 19: Data Management Plan Bootcamp (Virtual – no snacks)
  - November 25: Data Management Plan Bootcamp (In-Person – snacks!)
- **Writing:** Data Management Plan Assistant is always available!
- **After Writing:** Submit for feedback through the Assistant or reach out for a consultation with us: [rdm@mcmaster.ca](mailto:rdm@mcmaster.ca) or [u.mcmaster.ca/rdm-appointments](https://u.mcmaster.ca/rdm-appointments)

**Learn more:**



**Building a Data Management Plan for your research project.**

**<https://learn.scds.ca/intro-rdm/dmp.html>**

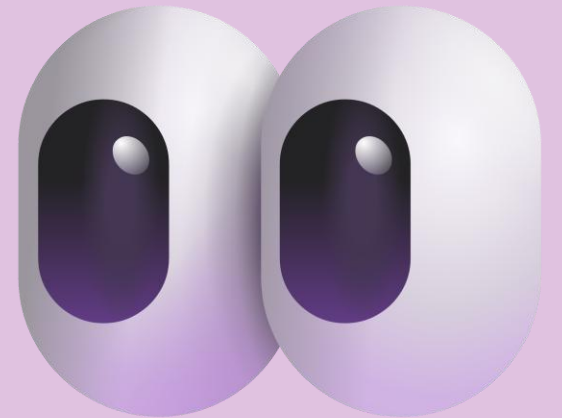
Libraries



**Sherman  
Centre**  
for Digital Scholarship



**Who do you think could  
potentially be involved in  
reviewing a DMP?**



## 2. Organize & document what you're doing consistently

- **Raw data are complicated.**
- Documentation and organization make data easier to understand and reuse—for you and others
- Whether you're working solo or collaborating across teams, documenting your data will reduce confusion and increase efficiency

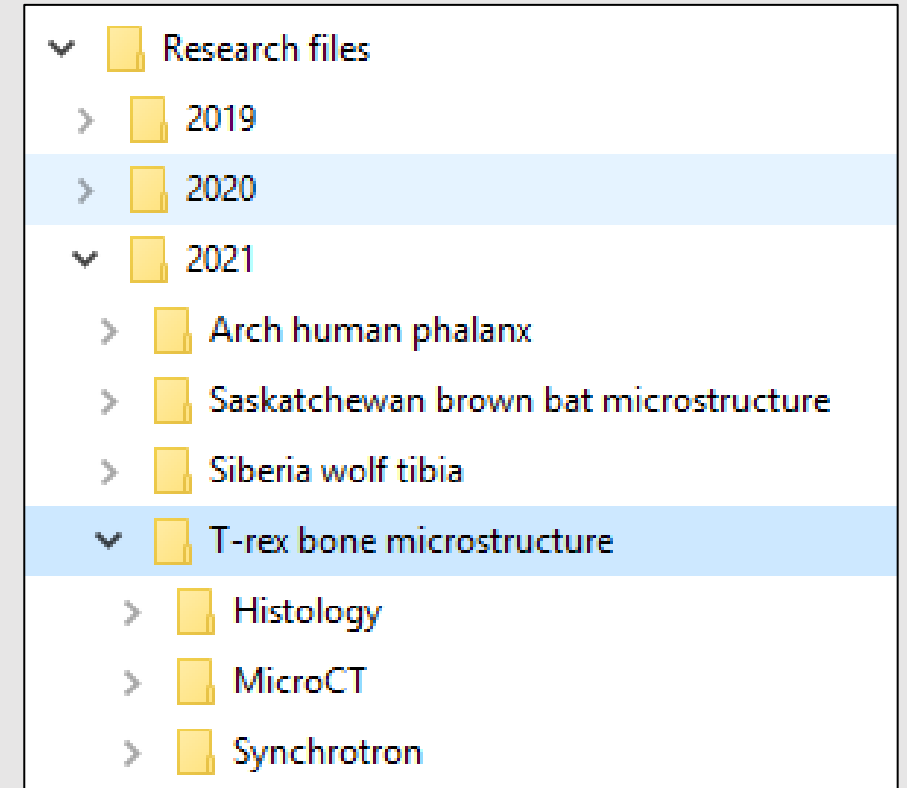
# Keeping folders organized makes it easier to find things.

Make organizing files and folders a habit – that way, it's easy to know where things go.

File organization schemes can include:

- By project
- By researcher
- By experiment type
- By date (often year)
- By some combination of the above

*(i.e. a two level structure of year -> project)*



# Research Data Organizing and Documenting

- Think about file and folder names and organization – *i.e.*  
`2025_10_22_LakeMercury_TestData1_EP.csv`

**Date:** 2025\_10\_22 (collection date)  
**Project Name:** LakeMercury  
**Short Description:** TestData1  
**Name:** EP (Elizabeth Phillips)

- Think about file types and formatting
- Document your data so anyone can understand it (including future you!)
  - Proper **metadata** about who produced it, when, how, and where
  - Explanations of file and variable names and relationships
  - Add a **README** file – [we have a template!](#)

```
#####  
[DATASET TITLE] - This readme file was generated on [YYYY-MM-DD] by [NAME]  
#####  
<A readme is a simple text documents that describe the files and organ  
others and your future self to understand and interpret what you've do  
to your dataset may be removed or altered where appropriate.>  
  
<This readme template is provided by McMaster University RDM Services  
and is adapted from Cornell's readme template (https://data.research.c  
Francesco Varrato, Alain Borel and Chiara Gabella's "README file for D  
template" (https://infoscience.epfl.ch/record/298249)>  
  
<Help text in angle brackets should be deleted before finalizing your  
brackets should be changed for your specific dataset.]>
```

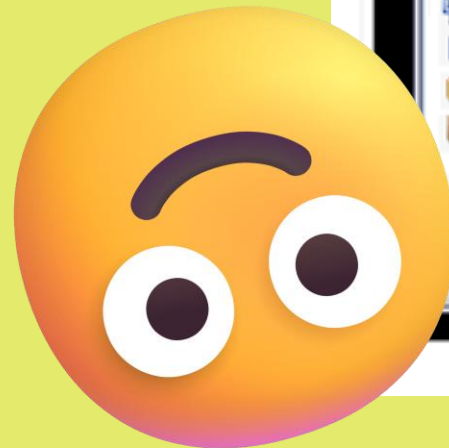


# Build a documentation scheme you will **actually use!**

- The most important aspect of documentation is doing it.
- Whatever file naming and organization scheme you choose, make sure it's **descriptive**, use it **consistently** and **document** it (in a readme.txt file).
- **Take advantage of the software that is out there**, including note-taking software, reference management software, and collaboration software.



**What issues can we see with these files? What did this person do well? What could they have done better?**



A STORY TOLD IN FILE NAMES:

Location: C:\user\research\data

Filename	Date Modified	Size	Type
data_2010.05.28_test.dat	3:37 PM 5/28/2010	420 KB	DAT file
data_2010.05.28_re-test.dat	4:29 PM 5/28/2010	421 KB	DAT file
data_2010.05.28_re-re-test.dat	5:43 PM 5/28/2010	420 KB	DAT file
data_2010.05.28_calibrate.dat	7:17 PM 5/28/2010	1,256 KB	DAT file
data_2010.05.28_huh??.dat	7:20 PM 5/28/2010	30 KB	DAT file
data_2010.05.28_WTF.dat	9:58 PM 5/28/2010	30 KB	DAT file
data_2010.05.29_aaarrgh.dat	12:37 AM 5/29/2010	30 KB	DAT file
data_2010.05.29_#\$\$@*&!!!.dat	2:40 AM 5/29/2010	0 KB	DAT file
data_2010.05.29_crap.dat	3:22 AM 5/29/2010	437 KB	DAT file
data_2010.05.29_notbad.dat	4:16 AM 5/29/2010	670 KB	DAT file
data_2010.05.29_woohoo!!!.dat	4:47 AM 5/29/2010	1,349 KB	DAT file
data_2010.05.29_USETHISONE.dat	5:08 AM 5/29/2010	2,894 KB	DAT file
analysis_graphs.xls	7:13 AM 5/29/2010	455 KB	XLS file
ThesisOutline!.doc	7:26 AM 5/29/2010	38 KB	DOC file
Notes_Meeting_with_ProfSmith.txt	11:38 AM 5/29/2010	1,673 KB	TXT file
JUNK...	2:45 PM 5/29/2010		Folder
data_2010.05.30_startingover.dat	8:37 AM 5/30/2010	420 KB	DAT file

type: Ph.D Thesis Modified: too many times Copyright: Jorge Cham www.phdcomics.com

# 3. Store & back up data securely

- **Data Loss**
  - Theft of devices
  - Loss of devices
  - Accidental damage or destruction
  - A USB drive is not an archive!
- **IT Security:**
  - Computer viruses, malware, ransomware.





# Where should I store data?

## Local



- On your computer or an external drive
- **You** need to make sure that storage is backed up
- Relatively cheap
- Easy to use
- Available offline
- **Higher chance of loss**

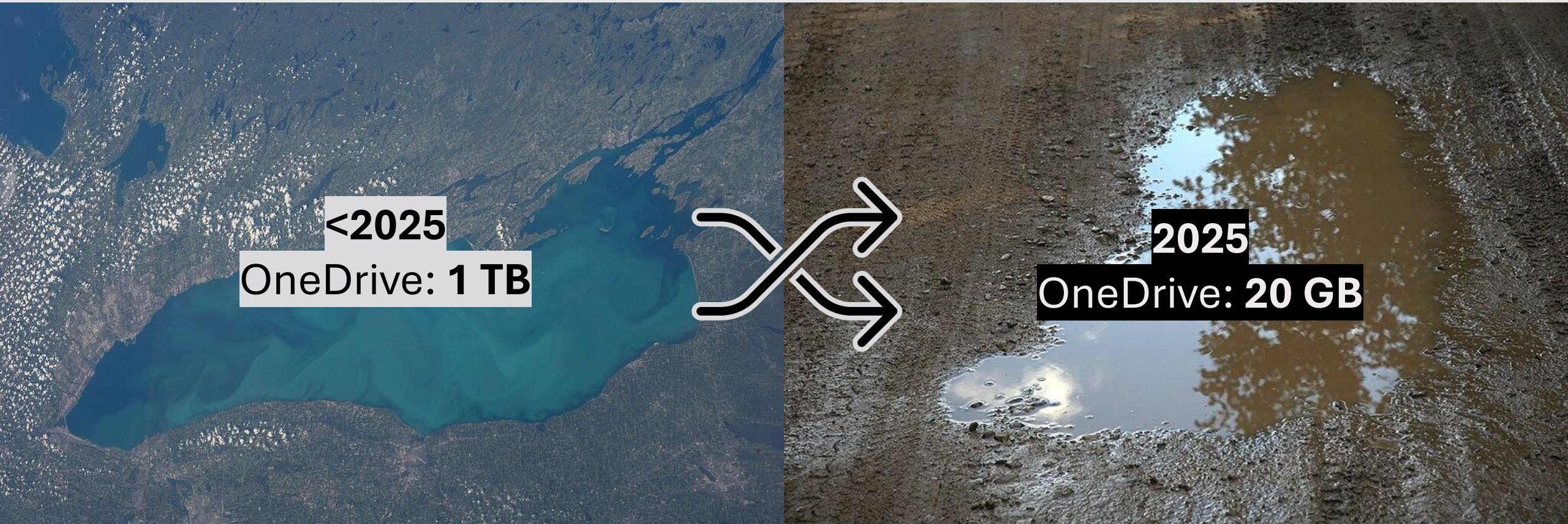
## Cloud



- On someone else's computer
- Automated backups, file versioning, and recovery
- File synchronization to your computer
- **Easy** to access and share files
- Subscription costs
- **Privacy can be opaque**



# Microsoft storage has shrunk for students



Above: NASA, "ISS-36 Lake Ontario (horizontal)," 2013-08-24, Wikimedia Commons, Public Domain, [https://commons.wikimedia.org/wiki/File:ISS-36\\_Lake\\_Ontario\\_\(horizontal\).jpg](https://commons.wikimedia.org/wiki/File:ISS-36_Lake_Ontario_(horizontal).jpg)

Below: Dmitry Makeev, "Reflections in Puddles," 2006-06-30, Wikimedia Commons, CC-BY-SA 4.0 [https://commons.wikimedia.org/wiki/File:Reflections\\_in\\_puddles.\\_img\\_012.jpg](https://commons.wikimedia.org/wiki/File:Reflections_in_puddles._img_012.jpg)

Libraries



# Research Data Storage

- If your data fits in the cloud, try OneDrive
  - Available to all campus users for free – **until you leave the institution**
  - Storage is shrinking
- If you need data to be portable, use local storage. Make sure local storage drives are encrypted with passwords.
- Use both if you can! Whatever you do, **make sure you make backups**

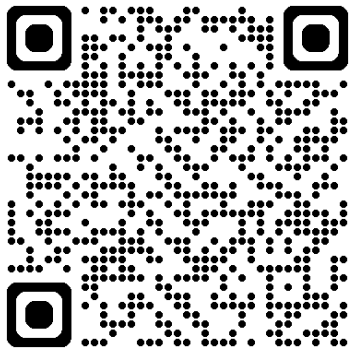
## *Example:*

1 copy stored locally on **hard drive** for analysis  
1 copy stored on **cloud storage** platform  
1 copy stored in a **secure campus drive**



# Research Data Storage Finder Tool

Find vetted storage providers depending on **risk, volume, collaboration**, and other needs.



scds.ca  
scds@mcmaster.ca

<https://rdm.mcmaster.ca/finder>

Webinar on data storage: <https://scds.github.io/intro-rdm/storage.html>

Step 1: Answer these questions to narrow down storage provider options.

[Clear Answers](#)

1. What risk level is your data? ⓘ

☐ Low  
☐ Medium  
☐ High

2. What type of data storage are you looking for? ⓘ

☐ Active research  
☐ Backup  
☐ Archival & Open data sharing

3. Are you collaborating with other researchers? ⓘ

☐ Other McMaster researchers  
☐ Specific researchers external to McMaster

Step 2: Select data storage providers you would like to compare

[Select All](#) [Clear Selections](#)

<b>Advanced Research Computing</b> Digital Research Alliance of Canada systems, storage and software	<b>Compute Canada Nextcloud</b> Digital Research Alliance of Canada online file hosting service	<b>FRDR</b> Find and Share Canadian Research Data	<b>Github</b> Distributed version control system for software code	<b>MacDrive</b> File Synchronization and Sharing solution
<b>MacDrive with Encrypted Data</b> Store sensitive data in encrypted files on MacDrive	<b>MacDrop</b> Web service to store and transfer files	<b>McMaster Dataverse</b> Store, share, publish and discover research data	<b>McMaster based custom solution</b> Contact us directly for help with complex projects	<b>Microsoft Sharepoint</b> Communication and document storage for large groups
<b>Microsoft Teams (institutional)</b> Create a group space for your research team	<b>OSF</b> Open platform for collaborative research	<b>OneDrive (institutional)</b> Save all your work and files to OneDrive and get them from any device, anywhere	<b>OneDrive (institutional) with Encrypted Data</b> Store sensitive data	<b>RHPCS Backup</b> Automated backup of your research computers

Libraries



**Sherman  
Centre**  
for Digital Scholarship

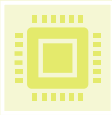




## Data Security (first steps!)



**Passwords:** Use password managers with a different complicated password for each service and device.



**Update Software and Hardware:** Don't put off updates—there are often important security settings.

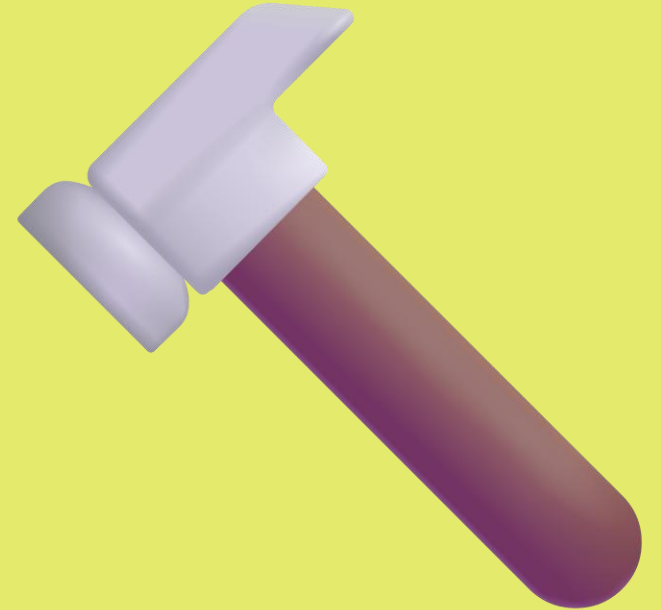


**Multi-Factor Authentication (MFA):** This added layer of security should be enabled on all platforms.



**Sensitive Data:** Limit collection and access. De-identify data, use encryption on storage and devices

**Extra credit: What is the worst thing that has happened to your data?**



# What help is available

- Check our **Research Data Storage Finder** for options
- Look at the existing **training modules** on the RDM website
- **Set up a consultation** if you have questions or particular storage or documentation considerations

**[rdm@mcmaster.ca](mailto:rdm@mcmaster.ca) or [u.mcmaster.ca/rdm-appointments](https://u.mcmaster.ca/rdm-appointments)**

**Learn more:**



**Strategies for research data storage and backup.**

**<https://learn.scds.ca/intro-rdm/storage.html>**

**Sensitive Data Management**

**<https://learn.scds.ca/intro-rdm/sensitive.html>**

Libraries



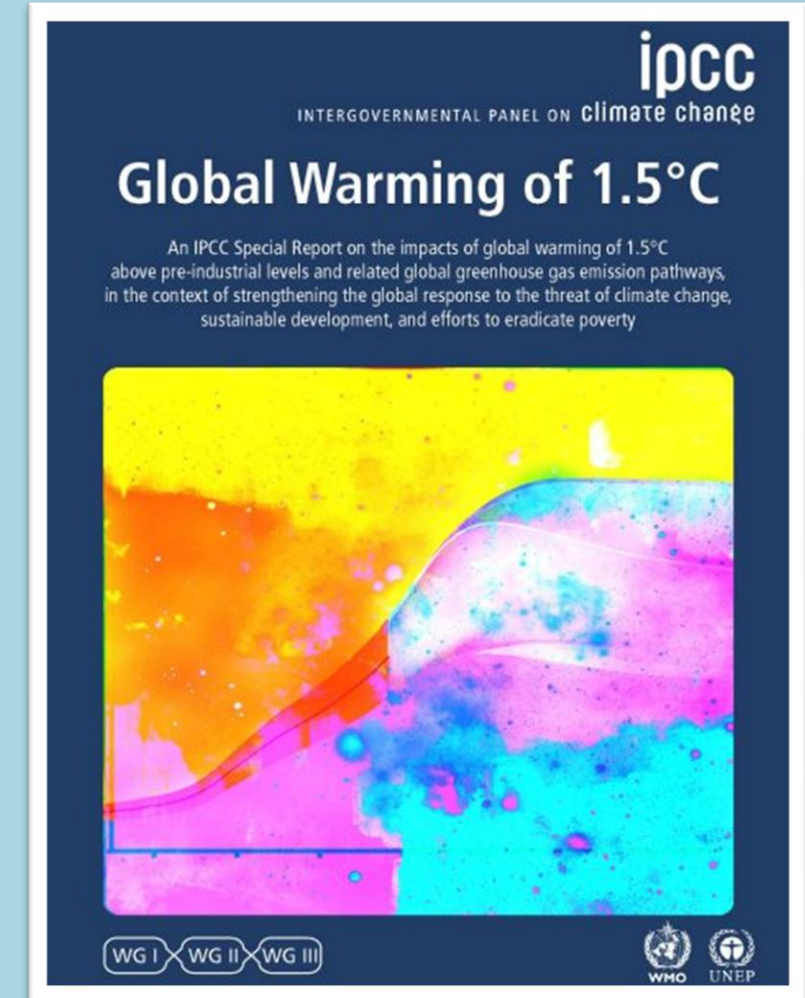
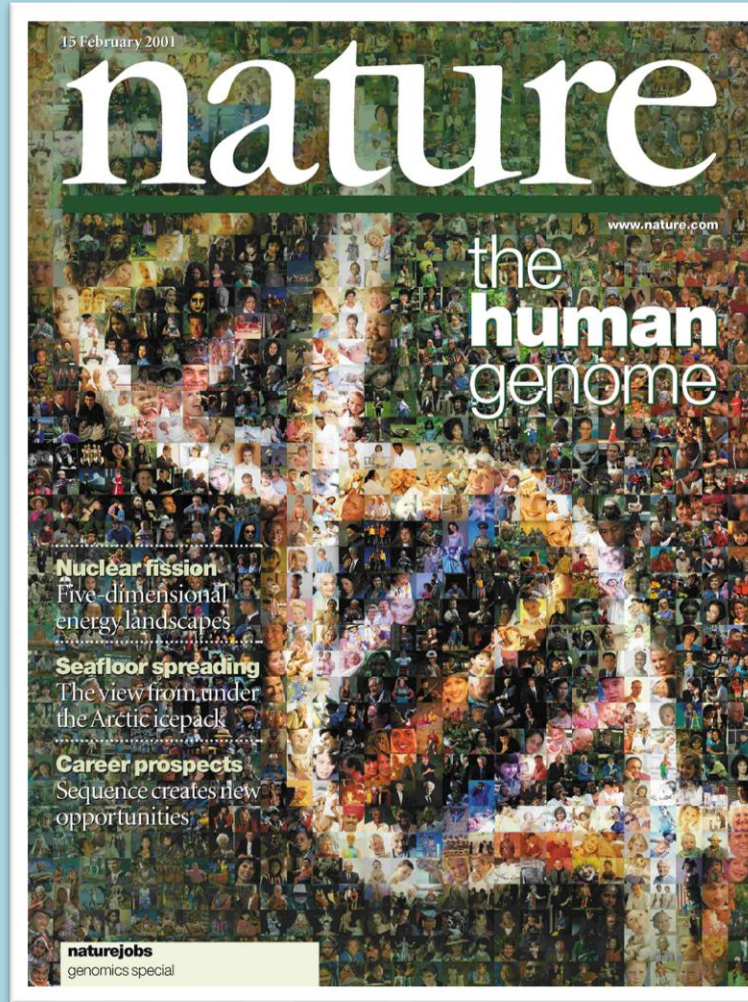
**Sherman  
Centre**  
for Digital Scholarship

# 4. Get ready for archiving & sharing

- **Data Sharing:** Open and free data sharing supports research ideals like transparency, accessibility, reproducibility, collaboration, and maximizes the impact and visibility of research.
- **Data Archiving:** What do you plan to do with your data once it's been published? How will you ensure that your data remains accessible (to you and others) long-term?



# Science built on the power of open data:



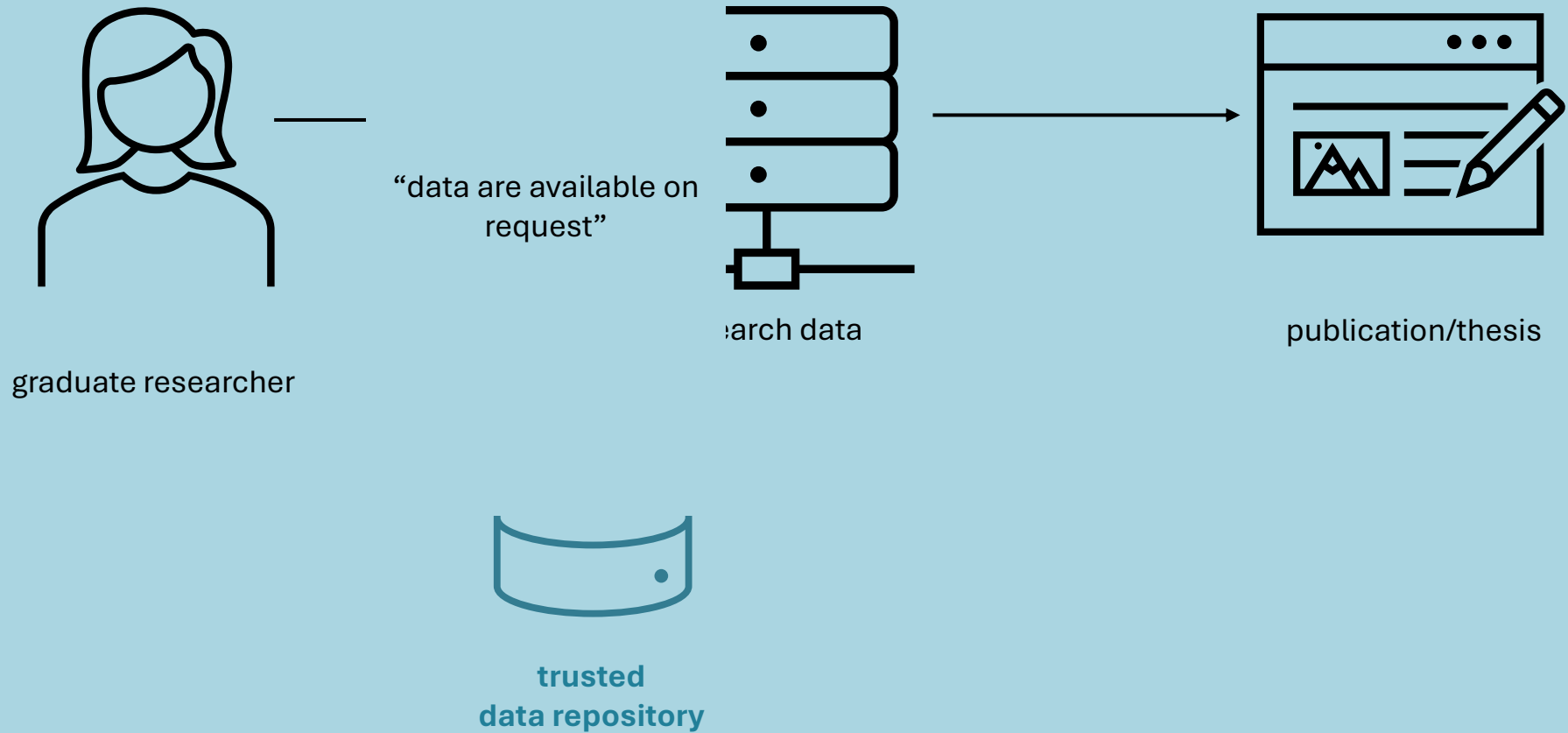
# Open Research



- Sharing data openly is a critical element in pushing academia towards openness, transparency, reproducibility, and research integrity.
- Research is often publicly funded – data as a “public good” that is accessible by journalists, other researchers, etc.
- Limits over-researching communities for qualitative studies. Share datasets with community partners – value added, support long-lasting research relationships.

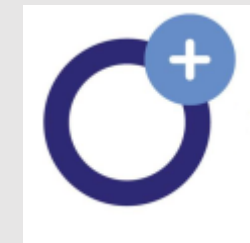


# What does data deposit look like?



# Finding existing Research Data

- Sometimes the data you want already exists – it's only a matter of knowing about it and getting access to it.
- **Data discovery platforms:**
  - Lunaris (national, Canada-based)
  - OpenAIRE Explore (European, AI-driven)
  - DataCite Commons (European)
  - Google Dataset Search (Global)
- Secondary data (even from a data sharing agreement or industry partner) has its own difficulties
  - May need reorganizing and extra documenting
  - Need to consider data ownership vs. data stewardship, especially for [sensitive](#), [Indigenous](#), and/or proprietary data





# Why share data?

**Citation Impact**

**Journal Requirements**

**Funder Requirements**

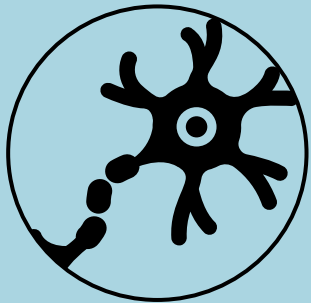
**Data Preservation**





# Data Repositories

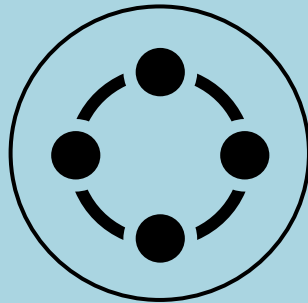
Publishing data in a recognized data repository is the best way to share data. There are thousands of data repositories.



## Domain Specific Repositories

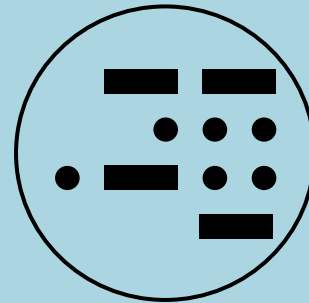
Focus on certain types of data such as genomic information or astronomical information.

<https://www.springernature.com/gp/authors/research-data-policy/recommended-repositories>



## General Repositories

Accept broader types of research data. ex. *McMaster Dataverse (part of Borealis)* and *Canada's Federated Research Data Repository (FRDR)*, *Open Science Framework (OSF)*.



## Code Repositories

There are also code-specific repositories like Github, Gitlab, BitBucket, SourceForge



## Controlled Access

For sensitive or qualitative data look at Vivli (Clinical Data), Qualitative Data Repository, Inter-university Consortium for Political and Social Research (ICPSR)

# McMaster Dataverse

<https://borealisdata.ca/dataverse/mcmaster>

- McMaster's Institutional Data Repository is a home for research data created by McMaster researchers. (*Not for sensitive data*)
- Provides basic data curation services
- Data is stewarded by professionals at McMaster
- Contains tools for tabular data exploration and analysis
- Depositor "Pre-Flight" Checklist helps you prepare for deposit

The image shows a screenshot of the McMaster University Dataverse website. The top navigation bar includes the Borealis logo, search, user guide, support, language, and login options. The main header features the McMaster University logo and links to the Dataverse and RDM services. A section titled "About Dataverse" explains its purpose as a secure research data repository. Below this, a "McMaster Dataverse Depositor Checklist" is displayed, listing mandatory requirements for data submission. The checklist includes items such as developing a README, gathering documentation, anonymizing sensitive data, obtaining ethics approval, and ensuring proper citation and file formats. A "Pre-Flight" checklist section is also visible, providing a final set of instructions before depositing data.

**McMaster University Dataverse** (McMaster University) [McMaster RDM Services](#)

[Borealis](#) >

[Contact](#) [Share](#)

### About Dataverse

The McMaster University Dataverse is a research data repository for our faculty, students, and staff. Files are held in a secure environment on Canadian servers. Researchers can choose to make content available publicly, to specific individuals, or to keep it private.

For information on research data management, including best practices for sharing and preserving research data, consult McMaster's [Research Data Management](#) website.

The Library can assist with using the repository service and will help you prepare your data for sharing and preservation.

- Before depositing a dataset in the McMaster Dataverse, please look through the [Data Deposit Guidelines](#) for information on what is required to deposit data.
- For a full walkthrough of depositing data to McMaster Dataverse, please [click here](#).

If you need assistance please contact: [rdm@mc.mcgill.ca](mailto:rdm@mc.mcgill.ca)

### McMaster Dataverse Depositor Checklist

This checklist augments our more comprehensive [McMaster Dataverse Data Deposit Guidelines](#) and [McMaster Dataverse Deposit Walkthrough](#) documents. Items marked with an asterisk \* are mandatory for submissions.

#### Before Deposit

- ☐ \***README:** Develop a README file to submit alongside dataset [template, [more info](#)]. You do not need to use the template but information equivalent to sections 1-3 from the template is required.
- ☐ **Documentation:** Gather all the documentation for your dataset, potentially including: data dictionary, data model, code book, interview guide, etc. [[more info](#)]
- ☐ \***Anonymization/De-identification:** McMaster Dataverse does NOT accept datasets containing confidential or sensitive information. Remove, replace, or redact data until they are de-identified and non-confidential. [[more info](#)]
- ☐ \***Ethics Approval:** Confirm you have MREB or HiREB approval to share data (if applicable).
- ☐ \***Citation and Credits:** Confirm you have credited, linked, and cited third-party sources, including data, code, or software (if applicable)
- ☐ **Dataset organization:** Use consistent file naming and folder organization. A well-structured dataset is easier to understand and share. [[more info](#)]
- ☐ **Sustainable file formats:** Make your data files accessible long-term and for re-use. [[more info](#)]

#### During Deposit

Visit [McMaster Dataverse](#) and set up your account by clicking "Log In"

**What are some opportunities you  
can think about with sharing data?  
What are some hesitations you  
might have?**







**Learn more:**



**Essentials of open data sharing.**  
<https://learn.scds.ca/intro-rdm/sharing.html>

## What help is available

- **Before Deposit:**
  - Look at McMaster Dataverse for existing examples in your field!
  - Look at our Dataverse Deposit checklist
  - Look at our README template
  - Set up an individual consultation with us!
- **During Deposit:** Data Deposit Bootcamp – May 19, 1-4 PM (snacks)
- **After Deposit:** We will review your dataset to make sure the data are appropriate and have the right metadata and documentation to be reusable

**[rdm@mcmaster.ca](mailto:rdm@mcmaster.ca) or**  
**[u.mcmaster.ca/rdm-](https://u.mcmaster.ca/rdm-appointments)**  
**[appointments](https://u.mcmaster.ca/rdm-appointments)**

Libraries



# So, where should you start?

Store your data with enough information that you will still understand it in 2 years!

1. **Clean up** your files – filenames, folders, etc.!
2. **Save** your data in multiple safe places – but **choose** what is worth saving
3. Keep documents that describe your data
4. **Whatever you do, do it intentionally!**



A Data Management Plan can help you think through options!

## Research Data Management Links

Send RDM Services an email:

[rdm@mcmaster.ca](mailto:rdm@mcmaster.ca)

Review resources on our webpage:

<https://rdm.mcmaster.ca>

Join our Community of Practice:

<https://u.mcmaster.ca/rdm-community>

Make an appointment:

<https://u.mcmaster.ca/rdm-appointments>

---

## SCDS Links

Send SCDS an Email:

[scds@mcmaster.ca](mailto:scds@mcmaster.ca)

Subscribe to our Newsletter:

<https://u.mcmaster.ca/sign-up>

Register for a Workshop:

<https://u.mcmaster.ca/scds-workshops>

Schedule a Consultation:

<https://libcal.mcmaster.ca/appointments>

