September 20, 2023 | 10:30am - 12pm
Virtual Workshop + Sandbox Session

Crash Course: Digital Tools to Ace Grad School

u.mcmaster.ca/scds-events

RDM

SCDS

Library

McMaster University

Crash Course: Digital Tools to Ace Grad School

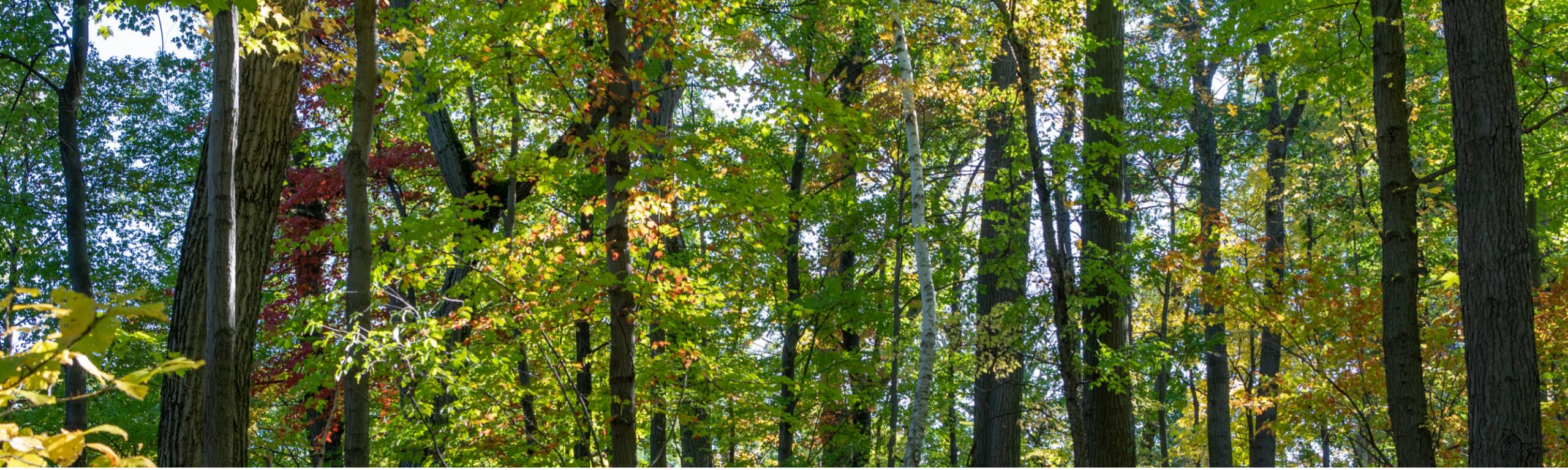
Jeff Demaine
Bibliometrics Librarian
demainj@mcmaster.ca

Danica Evering, MA
Research Data Management Specialist
rdm@mcmaster.ca

Research Data Management Workshop Series
September 20, 2023

Lewis & Ruth Sherman Centre for Digital Scholarship
scds.ca

McMaster University Library



McMaster University is located on the traditional territories of the Mississauga and Haudenosaunee Nations, and within the lands protected by the “Dish With One Spoon” wampum agreement.

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

The Sherman Centre and the McMaster University Library are committed to fostering a supportive and inclusive environment for its presenters and participants.

As a participant in this session, you agree to support and help cultivate an experience that is collaborative, respectful, and inclusive, as well as free of harassment, discrimination, and oppression. We reserve the right to remove participants who exhibit harassing, malicious, or persistently disruptive behaviour.

Please refer to our code of conduct webpage for more information: scds.ca/events/code-of-conduct/

Certificate Program

The Sherman Centre offers a Certificate of Attendance that rewards synchronous participation at 7 workshops. We also offer concentrations in Data Analysis and Visualization, Digital Scholarship, and Research Data Management.

Learn more about the Certificate Program: https://scds.ca/certificate-program

Verify your participation at a session: https://u.mcmaster.ca/verification

At an unspecified point during the workshop, a code will be read aloud. This is the answer to the third question of the form.

Research Data Management Workshops

Register for upcoming RDM events: https://rdm.mcmaster.ca/events

Sept. 27: “Best Practices for Managing Data in Your Research"

Sept. 28: “RDM Community of Practice - Data Management Plan Roundtable”

Oct. 25: “Data Management Plans and Intro to DMP Assistant”

Nov. 29: “Depositing & Sharing Data Online with McMaster Dataverse”

Feb. 14: “Storage Scores: Store & Back Up Data at McMaster”

Mar. 20: “How to Implement Encryption to Protect Your Research Data”

Apr. 17: “Sensitive Data Management”

May 14: “Data Management Plan (DMP) Bootcamp”

Jun. 18: “Data Deposit Bootcamp”

Session Recording and Privacy

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.

Outline

TL;DR - What to do with graduate research data

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

1. Make a plan to manage it
2. See if someone already collected some you can use
3. Keep it safe and secure
4. Organizing (satisfying)
5. Share (if you can)!

TL;DR – Be a distinguished researcher

- Get an ORCID
- Distinguish yourself from doppelgangers
- Get credit for what you write
- Connect your research

Sandbox Session

McMaster Research Data Management Services

We can help you with any questions about:

- Data security, storage and backup
- Data Management Plans
- Data organization and documentation
- Data sharing, archival, and preservation
  - Check out RDM Services online: rdm.mcmaster.ca
  - **Email us at** rdm@mcmaster.ca or book an appointment with us.

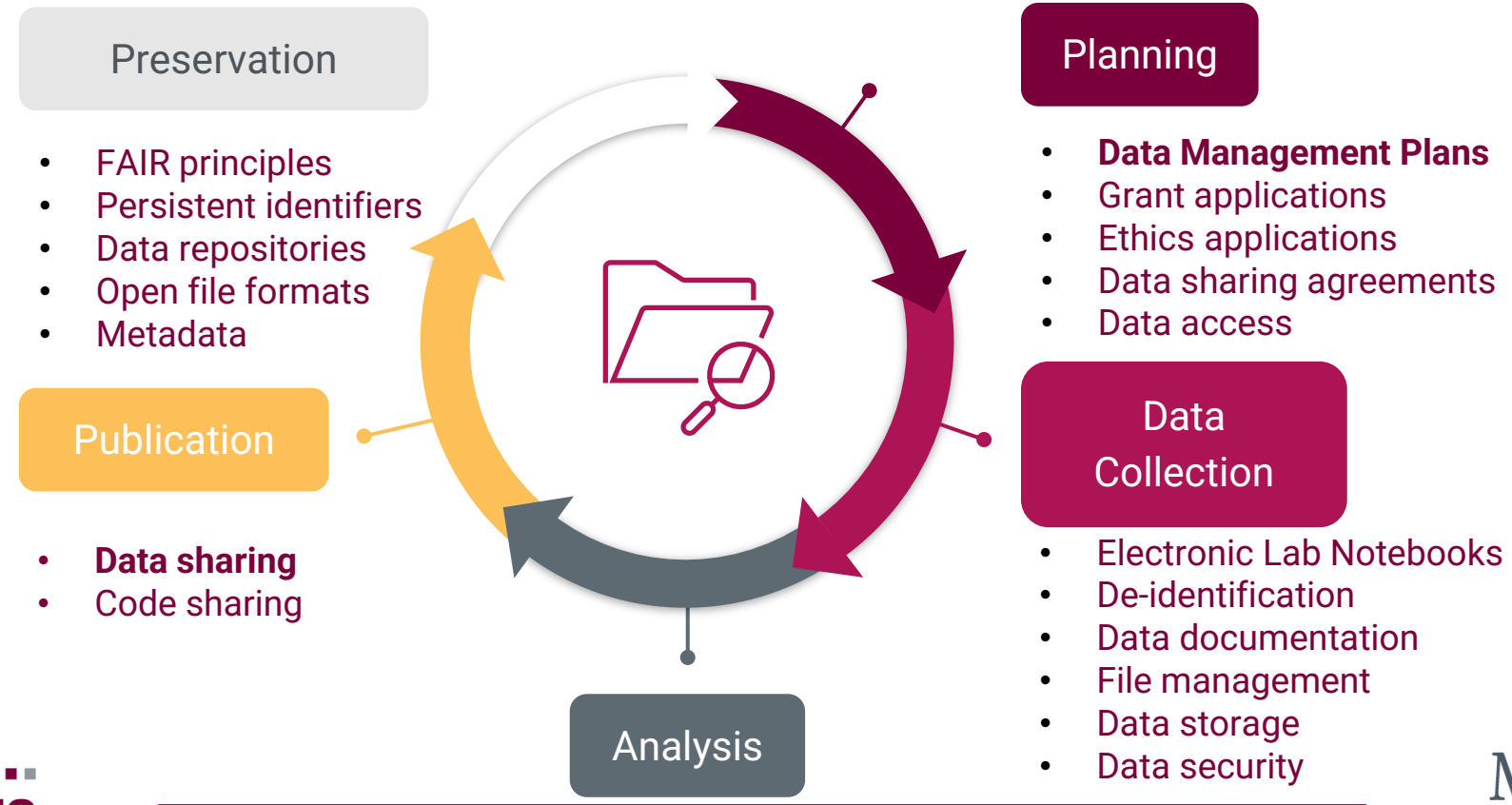
Do I even have research data to manage? (You do!)

"Data may be in any format or medium taking the form of text, numbers, symbols, images, films, video, sound recordings, pictorial reproductions, drawings, designs or other graphical representations, procedural manuals, forms, diagrams, workflows, equipment descriptions, data files, data processing algorithms, software, programming languages, code, or statistical records."

Innovation, Science and Economic Development Canada. "Frequently Asked Questions -Tri-Agency Research Data Management Policy." Government of Canada. Innovation, Science and Economic Development Canad, October 19, 2021. https://science.gc.ca/eic/site/063.nsf/eng/h_97609.html#1a

What is **Research Data Management**?

Research Data Management is a suite of connected processes and practices applied throughout the research lifecycle, as data are **planned** for, **collected**, **organized**, **documented**, **stored**, **preserved**, **shared**, and **reused**, in support of analysis, research, creative works, and dissemination that benefit society.



RDM practices	make research better.
Data organization and planning	saves time and resources.
Good data storage and backup practices	help avoid loss of data from theft, corruption, or failure of storage devices.
Secure data management	protects research participants and sensitive data.
Depositing and sharing data openly	increases the accessibility of research and allows others to reproduce and verify results.

Photo by Louis Reed on Unsplash.

1. Make a plan to manage it

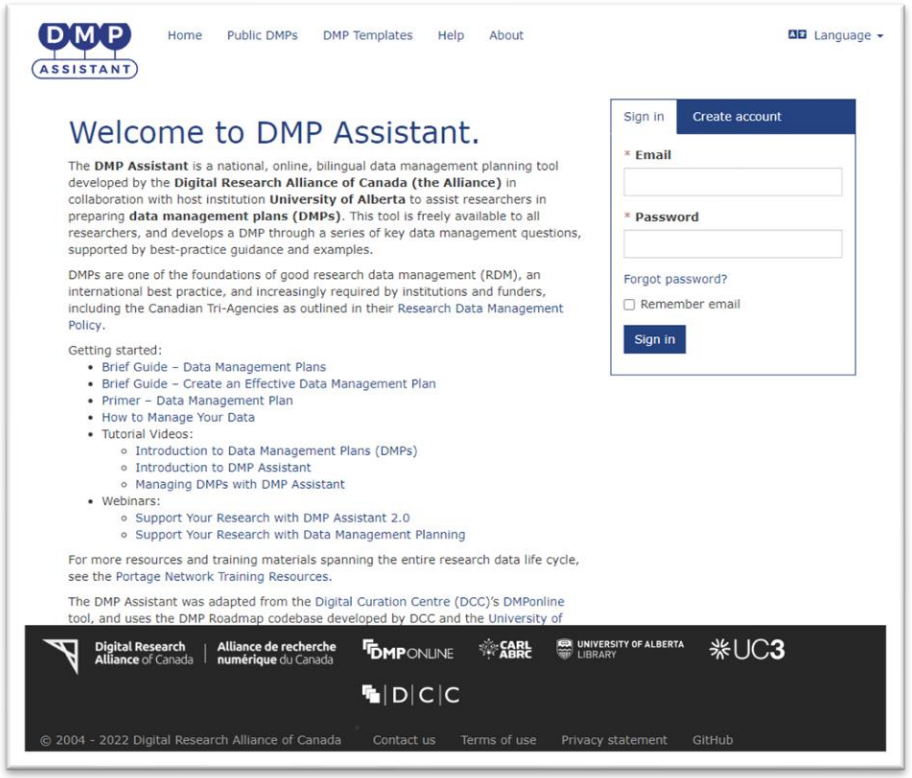
- A **Data Management Plan (DMP)** describes how you will create, store, organize, document, secure, preserve, and share your research data.
- A document which speaks to the management of data both **during** the active phases of your research and **after** the completion of the research project.

Why create a DMP?

- A blueprint for how you will be working with your data during your project
- Avoid potential pitfalls and problems before they occur
- Prepare for future stages of research including potential data sharing
- Many research funders require grant applicants to submit a DMP – Including Tri-Agency (starting 2022)

What goes in a Data Management Plan?

Digital Research Alliance of Canada | Alliance de recherche numérique du Canada



Tool: DMP Assistant

- A web-based, bilingual data management planning tool
- Available to all researchers in Canada
- Walks you through relevant questions for data management
- Exportable data management plans

assistant.portagenetwork.ca/

2. See if someone already collected some you can use

Tool: Research Data Discovery Platforms and Search Engines

DataCite Search https://search.datacite.org/

Google Dataset Search https://datasetsearch.research.google.com/

McMaster Library Data Services https://library.mcmaster.ca/services/data-services

Inter-university Consortium for Political and Social Research (ICPSR) https://www.icpsr.umich.edu/

Statistics Canada Microdata https://www.statcan.gc.ca/en/microdata

Nature Scientific Data: Data Repository Guidance https://www.nature.com/sdata/policies/repositories

3. Keep it safe and secure

- Avoiding data loss
  - Theft or loss of devices, accidental damage or destruction
- IT Security
  - Computer viruses, malware, ransomware

Data Loss

University of Manitoba Psychology

Lewis & Ruth Sherman Centre for Digital Scholarship

scds.ca

Library

Tool: Back-Up Strategies

Make sure you have more than one copy of your data. Follow the 3-2-1 rule:

3 Copies of your data (at least!)

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**

2 Copies are on-hand (easily accessible) on different systems (internal hard drive, cloud storage, etc.)

- a "**production**" (working) copy
- a "**production backup**" copy

1 Copy is in another location ("off-site") from the others with a ***trusted*** service provider

Where should I store my data?

Features to look for when deciding on a storage platform:

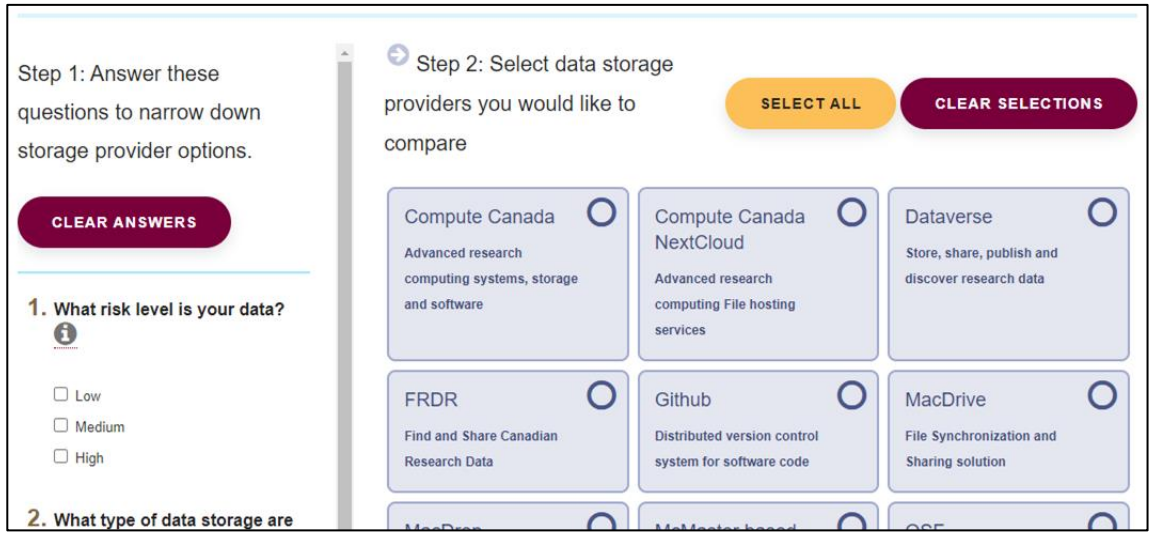
- Version control
- File recovery
- Security features (2FA, encryption)
- Collaboration features
- Storage provided
- Cost
- Storage location

Special considerations: Sensitive data, indigenous data, computational needs, code

Tool: Research Data Storage Finder

If you're looking for a good place to store your research data, try our **Research Data Storage Finder** webtool, where you can get personalized recommendations.

u.mcmaster.ca/storagefinder



Check out our recorded webinar on data storage:

https://scds.github.io/intro-rdm/storage.html

Tool: MFA and Password Management

- Enable **Multi-Factor Authentication (MFA)** wherever you can

- Follow good password practices everywhere
  - Use unique passwords for each important service/website
  - Use a good password for your computer **and your phone**
  - Consider using a password manager
  - **Don't share your password with anyone**

IT Tips from UTS:

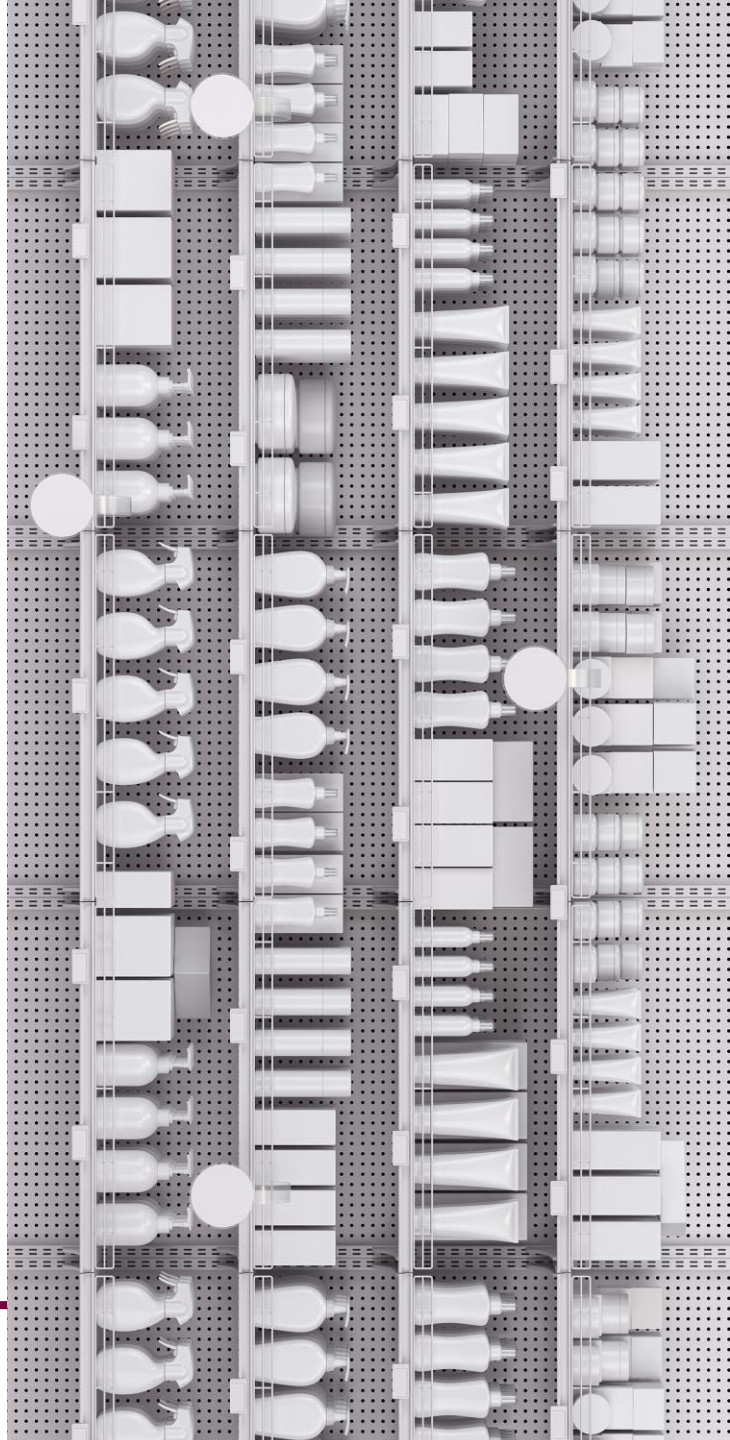
- https://uts.mcmaster.ca/ready-for-the-fall-term-heres-what-you-need-to-know/

4. Organizing (satisfying)

Your files should be **easy to find** and **easy to recognize**

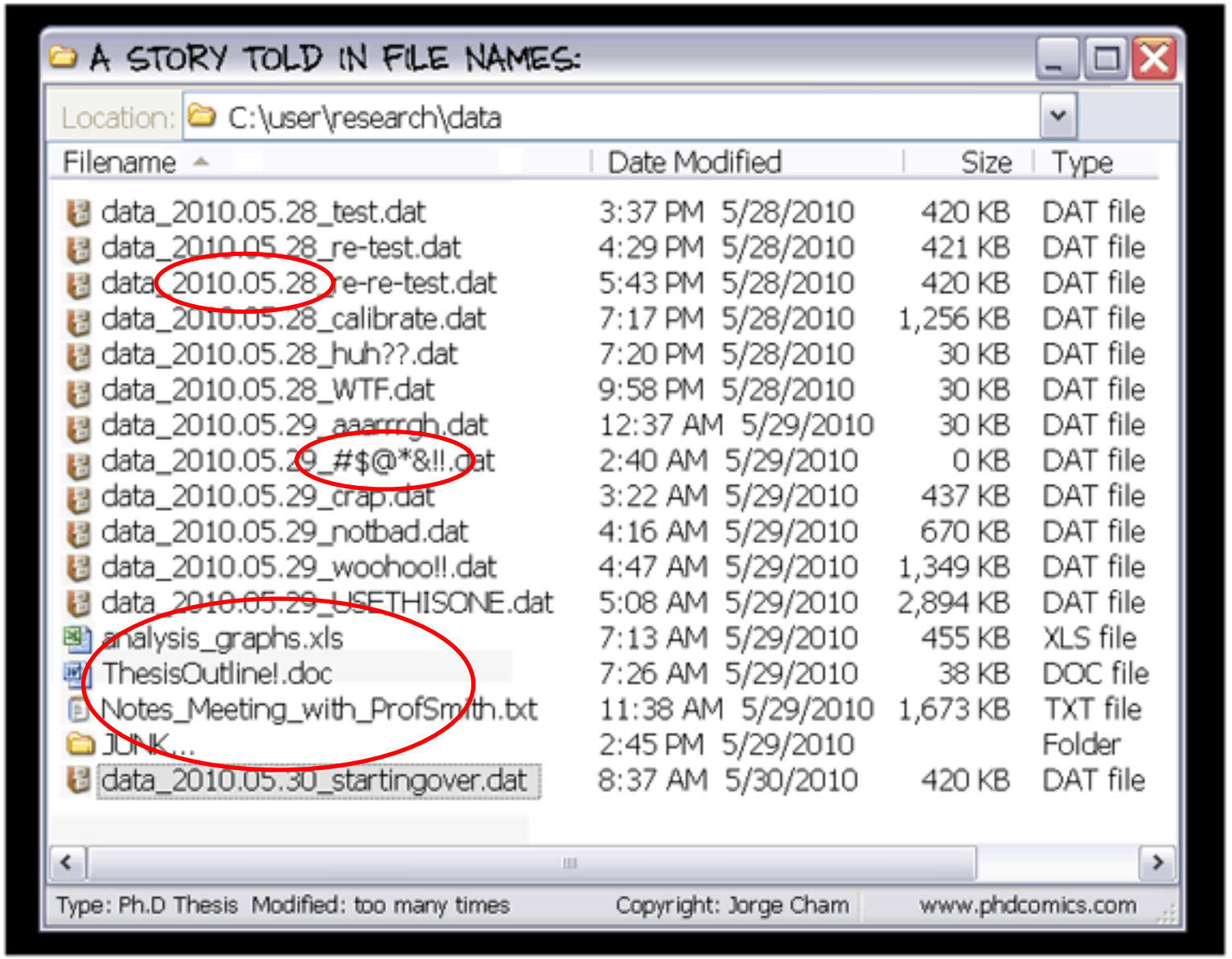
Research projects create a lot of files, often with similar names:

- Your work – including **data files**, coursework, publications, conference presentations, and your thesis
- Other people's work – including research literature, citations, presentation, etc.



Do you have files named like this?

Is this a good file name system?



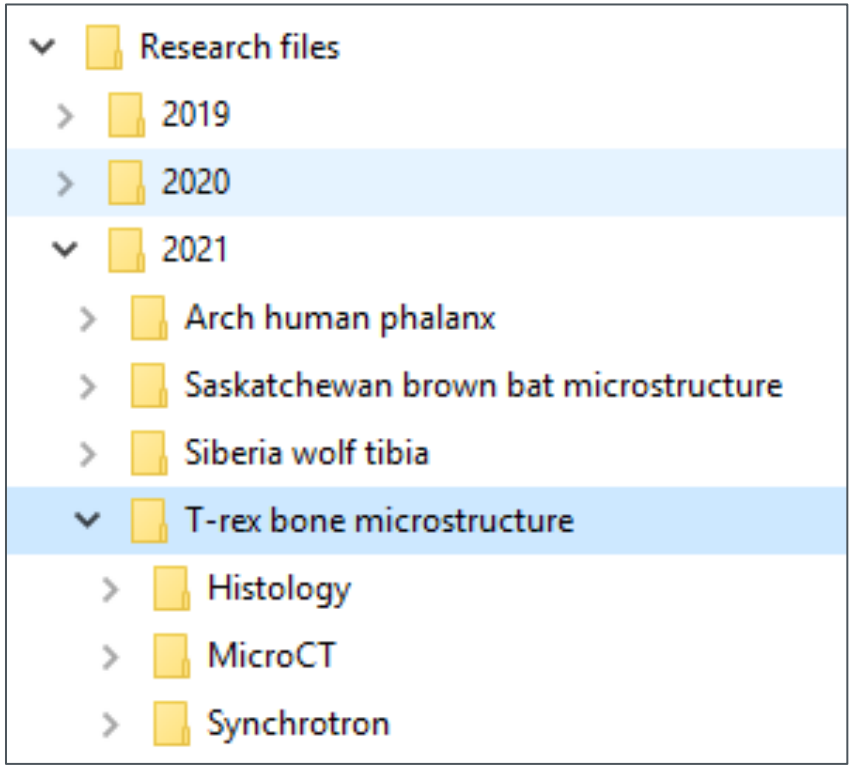

Does your desktop look like this?

Tool: File organization scheme

The key to organizing files is to make it a **habit**. Make it easy to know where files go.

Files can be organized:

- 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)



Tool: File naming scheme

A good file name makes it **easy to recognize** a file's contents:

 dataset.csv vs  2020_12_01_MercuryTestData.csv

File naming schemes can include:

- A short description of file contents
- The date the file was created (try using YYYY_MM_DD)
- File version (if applicable)
- Initials of researcher (if working on a collaborative file)

Try to keep names **short** and avoid special characters such as:

& , * % # * ( ) ! @$ ^ ~ ‘ { } [ ] ? < > –

Keep your files organized.

- 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).
- Collaboration software like Electronic Lab Notebooks, Reference Management software, or the Research Project Management Software can help.

Tool: Documentation

Notebooks: Write down your methodology as well as details around what decisions you make and why as you go. You can refer to this when you're writing your paper later.

Codebooks: For survey or statistics data, but also could be used for documents. Keep track of your tagging system to find things more easily and notice patterns in qualitative research.

Data Dictionaries: Usually for tabular/spreadsheet data – write out your variables, units, labels, etc. to keep data clean.

5. Share (if you can)!

Sharing your data is a strong part of open research!

Researchers share data to support reproducibility, trust in research, contribution/impact of your research, benefits to participants and funders, and to meet journal and funder requirements.

Tool: Research Data Repositories

Try out data sharing as part of your graduate school experience to learn this research best practice. You will publish your graduate thesis in MacSphere. When you do, consider sharing your data in McMaster Dataverse - **https://borealisdata.ca/dataverse/mcmaster**

If you're interested in learning more:

RDM Webinar Series

rdm.mcmaster.ca/events

Next event September 27th 10:30 AM

RDM Services:

web: rdm.mcmaster.ca

email: rdm@mcmaster.ca