Best Practices for Managing Data in your Research

Oct. 21, 2026

11:00 - 12:00 p.m.

Online

Sherman Centre for Digital Scholarship



Stream

Research Data Management

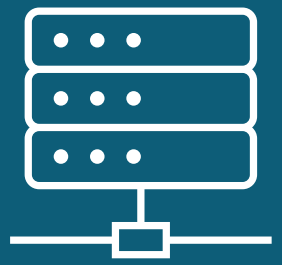


What is research data management (RDM)?

Are your data vulnerable?



researcher/
research team



research data



publication/thesis











?

“Dave” is a graduate student.

Dave’s data:

- **X-ray images** (~10 TB total)
- **Other misc files** (~10 GB) instrument metadata, records, scripts, spreadsheet files

Dave’s data management:

- Image files are too **large** for cloud and stored wherever there is space - on lab computers and hard drives
- The other files are smaller and stored on a personal laptop and cloud
- Data is not consistently documented.

scds.ca
scds@mcmaster.ca

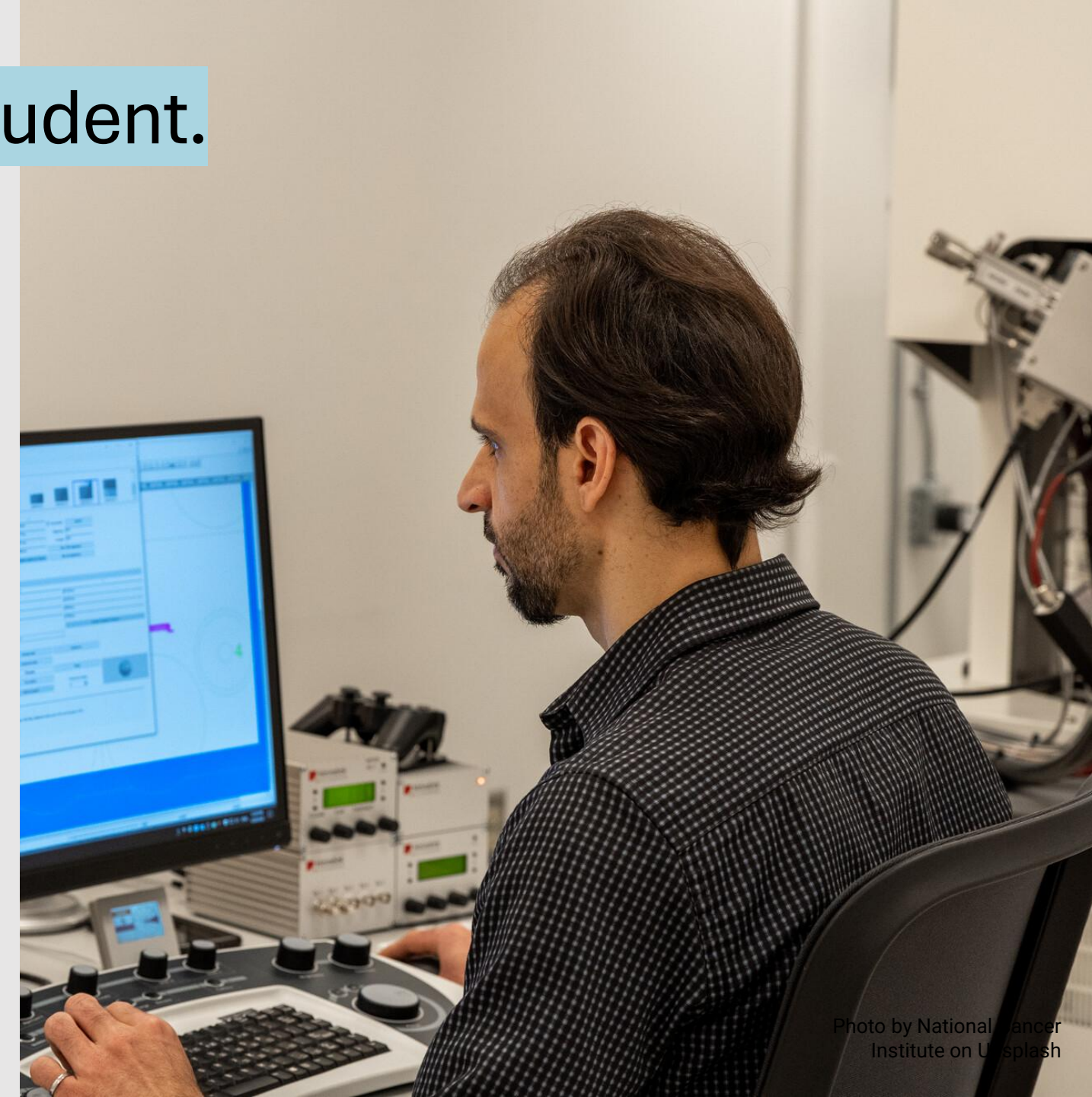


Photo by National Cancer Institute on Unsplash

Disaster strikes:

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.

- **There is no back up of this data** because there was no documentation on where data was stored

What would you do?

- **Try to recollect that data?** This could add months of delay redoing your experiment
- **Keep going?** Is your work still viable with the loss of part of the dataset?



This can happen in **qualitative** research too:

Note to self: don't record your interviews in a coffee shop!

- There were also things that could have gone wrong but didn't:
  - No **plan** for the data
  - Bad **documentation** and **organization**
    - Took notes but on both paper and laptop
    - PDF articles; typed out each reference
  - Risky **storage** practices
    - Only one recording device
    - Files stored on just my hard drive
  - **Archived** files on portable hard drives
    - No notes left on what the files were



Photo by Benjamin Riot on Unsplash

scds.ca
scds@mcmaster.ca

To Do:

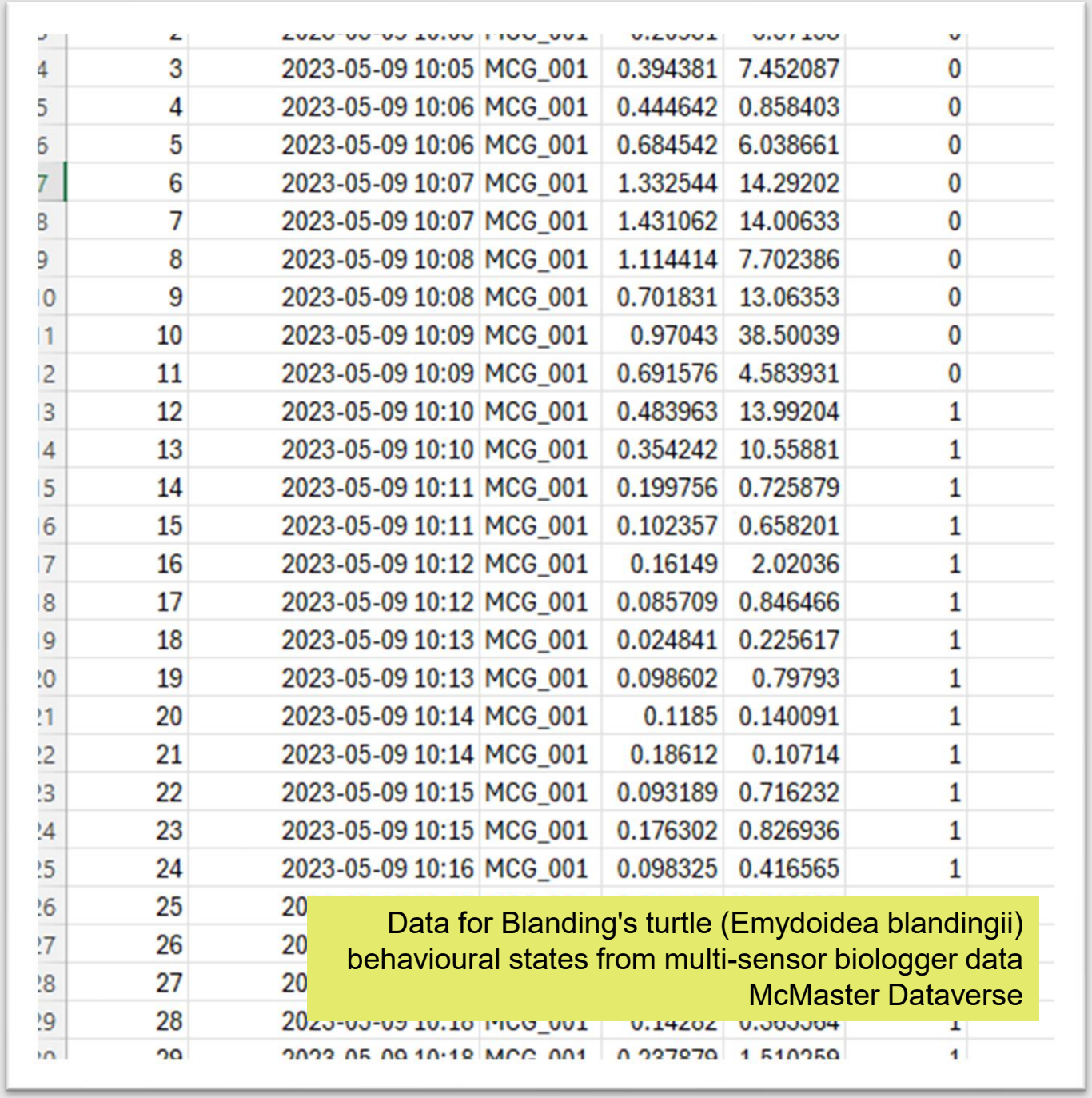
- ☐ 1.Make a plan
- ☐ 2.Organize & document data consistently
- ☐ 3.Store & back up data securely
- ☐ 4.Get ready for archiving & sharing



What counts as “research data”?

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

scds.ca
scds@mcmaster.ca



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	20				
30	29	2023-05-09 10:18	MCG_001	0.237879	1.510259	1

Data for Blanding's turtle (Emydoidea blandingii) behavioural states from multi-sensor biologger 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.

scds.ca
scds@mcmaster.ca

A joint initiative of McMaster's Faculty of Humanities and McMaster University Libraries
McMaster University
Sherman Centre for Digital Scholarship

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

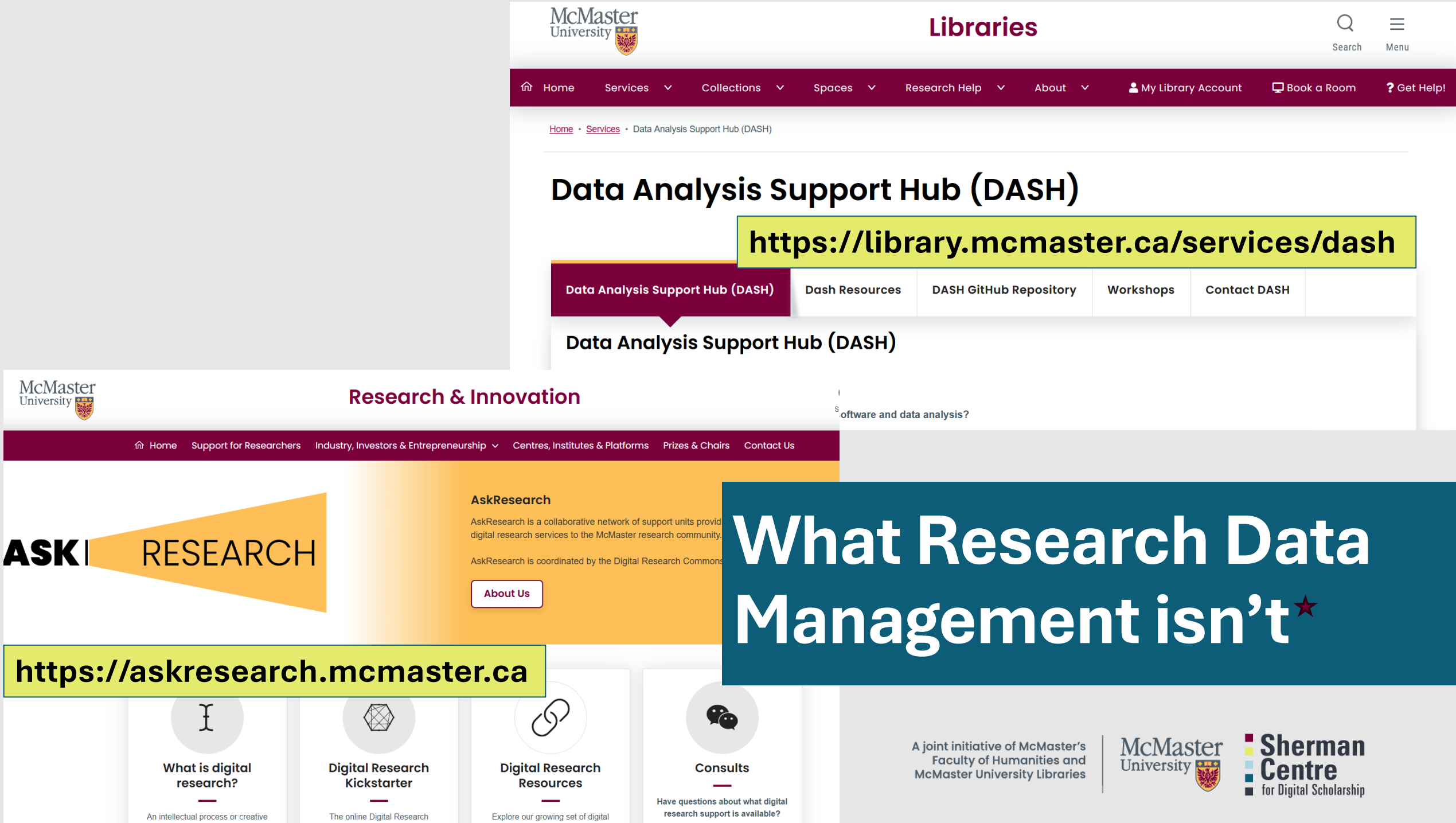


Knowledge Sharing

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

scds.ca
scds@mcmaster.ca

A joint initiative of McMaster's Faculty of Humanities and McMaster University Libraries
McMaster University
Sherman Centre for Digital Scholarship


McMaster University
Libraries
Search
Menu
Home
Services
Collections
Spaces
Research Help
About
My Library Account
Book a Room
Get Help!
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)
oftware and data analysis?
McMaster University
Research & Innovation
Home
Support for Researchers
Industry, Investors & Entrepreneurship
Centres, Institutes & Platforms
Prizes & Chairs
Contact Us
ASK RESEARCH
AskResearch
AskResearch is a collaborative network of support units provid
digital research services to the McMaster research community.
AskResearch is coordinated by the Digital Research Commons
About Us
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?
What Research Data Management isn't*
A joint initiative of McMaster's Faculty of Humanities and McMaster University Libraries
McMaster University
Sherman Centre for Digital Scholarship

- [x] **1.Make a plan**
- [ ] *2.Organize & document data consistently*
- [ ] *3.Store & back up data securely*
- [ ] *4.Get ready for archiving & sharing*




Gates
Portes
Oversized Baggage
Bagages hors format
Check-in
Enregistrement
3
4
OS7

Data Management Plans

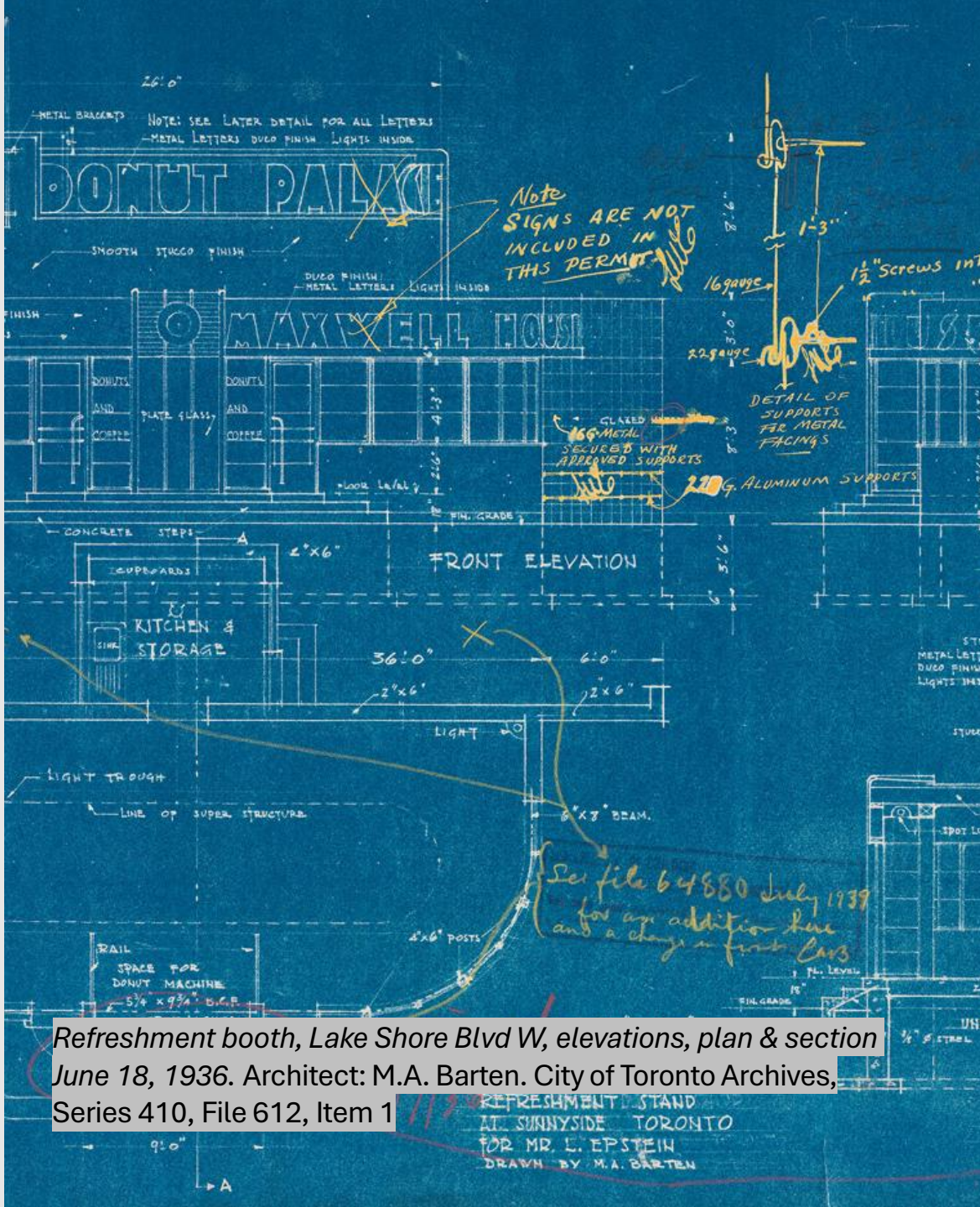
- **Plan for your work**: Describes how you will create, store, organize, document, secure, preserve, and share your research data—both during and after your project.

- **Ongoing conversation**: An opportunity to engage your research team, partners, and collaborators in an ongoing conversation about how you will be managing research data.

- **Funder requirement:** Required for several CIHR, NSERC, and SSHRC grants, Canadian Foundation for Innovation (CFI) – Sustainability Section, Ontario Research Fund (ORF), Horizon Europe

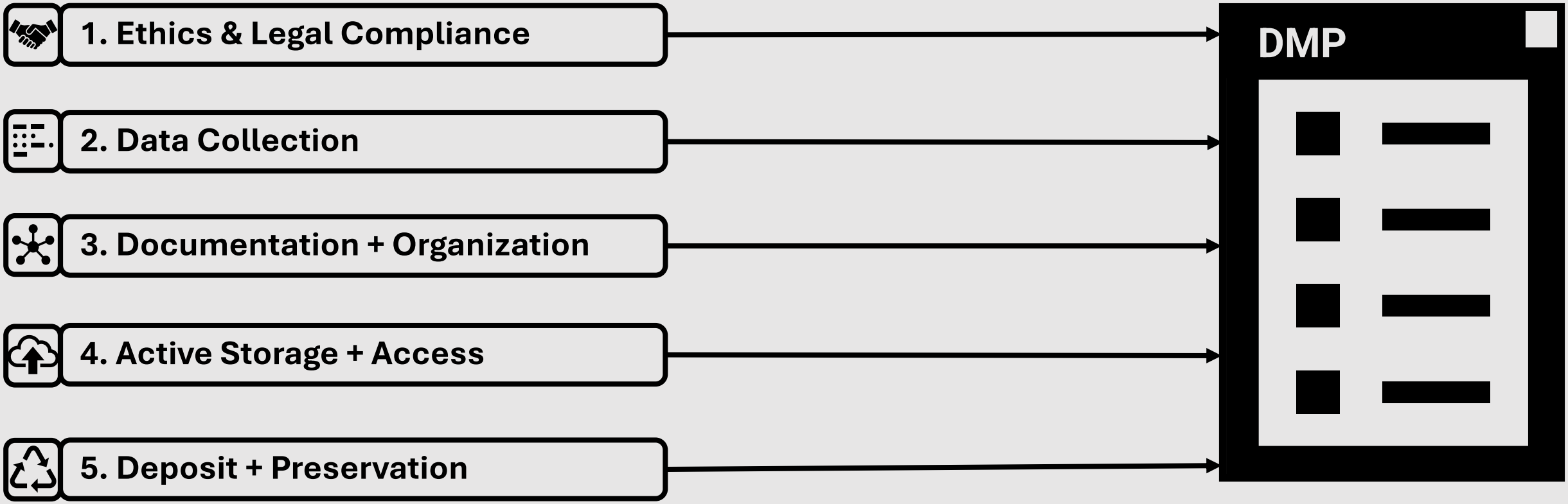
- *RDM Services reviews DMPs!*

scds.ca
scds@mcmaster.ca



Refreshment booth, Lake Shore Blvd W, elevations, plan & section June 18, 1936. Architect: M.A. Barten. City of Toronto Archives, Series 410, File 612, Item 1

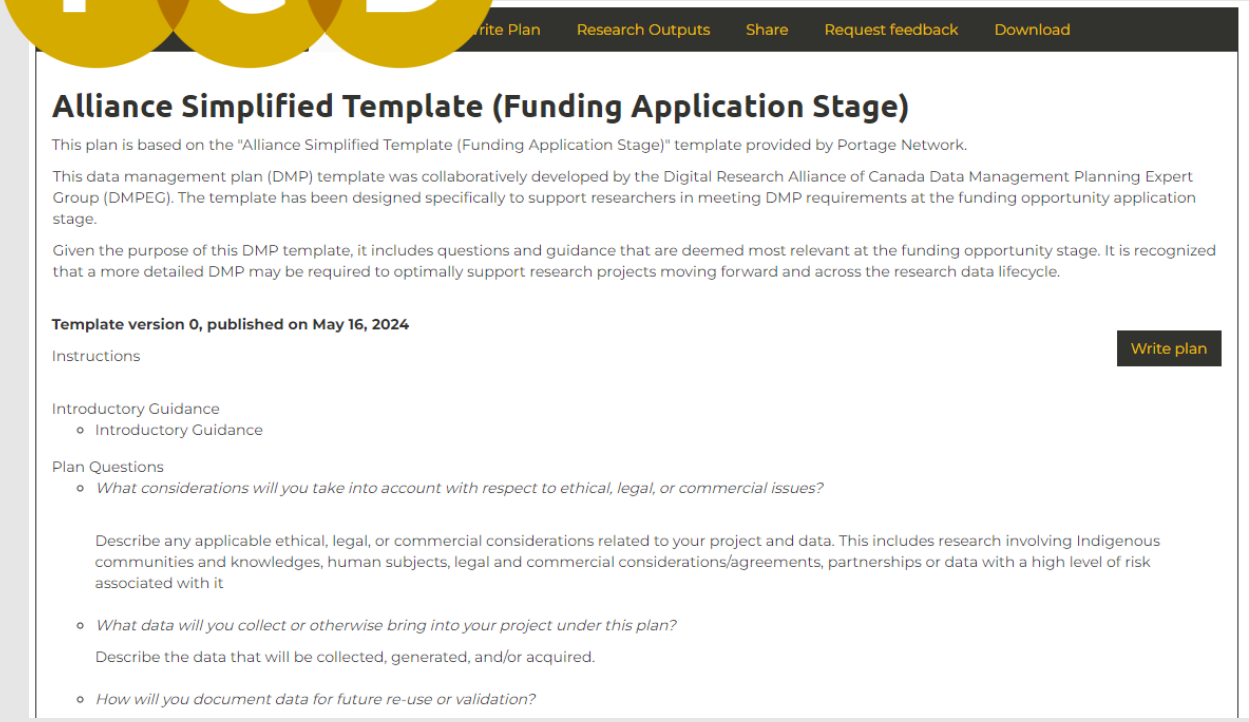
What goes in a Data Management Plan?



scds.ca
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**

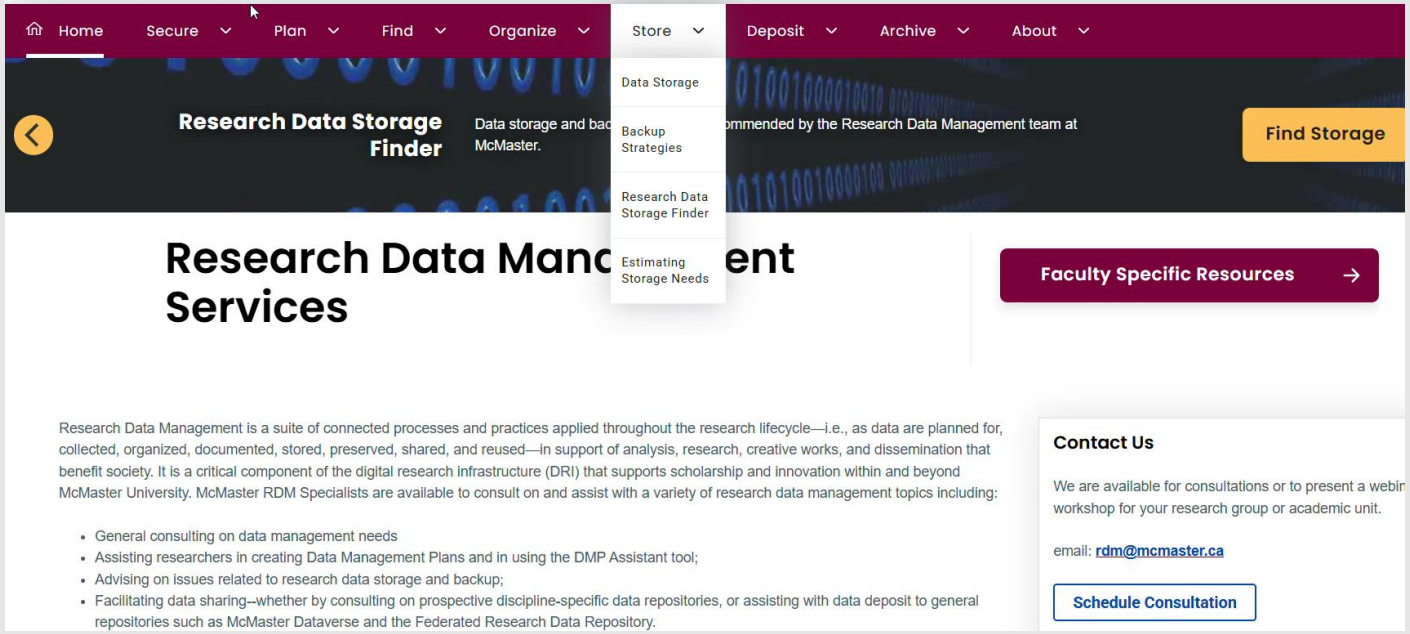


scds.ca
scds@mcmaster.ca

A joint initiative of McMaster's Faculty of Humanities and McMaster University Libraries
McMaster University
Sherman Centre for Digital Scholarship

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**



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 (-80ºC).

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

scds.ca
scds@mcmaster.ca

Plan: Find existing Research Data

- Sometimes the data you want already exists – it's only a matter of knowing about it and getting access to it.







- **Where to look:**
  - Lunaris (national, Canada-based) - https://www.lunaris.ca/en
  - OpenAIRE Explore (European, AI-driven) - https://explore.openaire.eu/
  - Google Dataset Search (Global) - https://datasetsearch.research.google.com/
  - Statistics Canada Microdata through McMaster's RDC https://rdc.mcmaster.ca/
  - McMaster Library Data Services - https://library.mcmaster.ca/services/data-services



Learn more:

Building a Data Management Plan for your research project.
https://learn.scds.ca/intro-rdm/dmp.html

What help is available?

- **Before Writing:** Review our Data Management Plan Database for examples in your discipline.
- **Writing:** Brand new asynchronous Data Management Plan module!
- **Writing:** Data Management Plan Bootcamps
  - **May 11**: Data Management Plan Bootcamp (in-person = snacks!)
- **After Writing:** Submit for feedback through the Assistant or reach out for a consultation with us: **rdm@mcmaster.ca or u.mcmaster.ca/rdm-appointments**

A joint initiative of McMaster's Faculty of Humanities and McMaster University Libraries

McMaster University

Sherman Centre for Digital Scholarship

- [x] *1.Make a plan*
- [x] **2.Organize & document data consistently**
- [ ] *3.Store & back up data securely*
- [ ] *4.Get ready for archiving & sharing*



Basics: Organizing and Documenting

- **Organizing:** set up a folder structure for your research
- **Documenting:** take notes, and keep contextual information about your data with your data
- **Take advantage of the software that is out there**, including note-taking software, reference management software, and collaboration software.
- Most importantly, build a documentation scheme you will actually use! Don’t make it too complex.



scds.ca
scds@mcmaster.ca

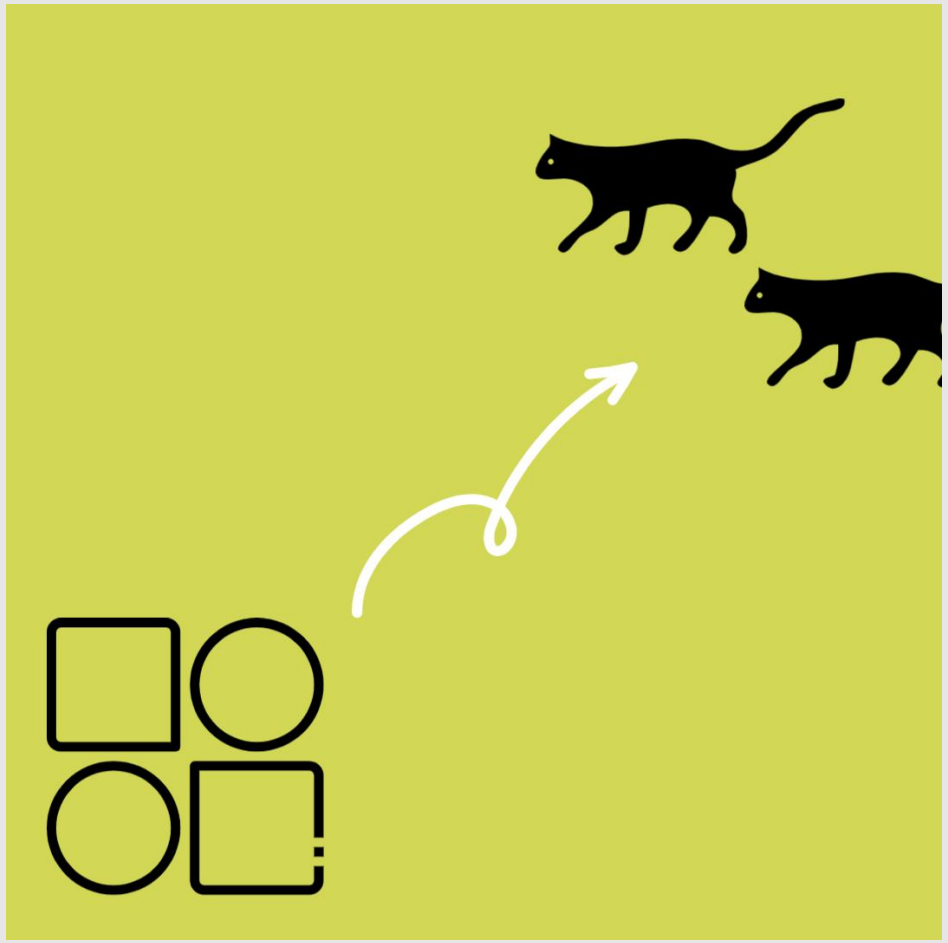


Trnava University on Unsplash

Documenting + Notetaking

- Take notes as you do things - **Don’t rely on your memory**
- Write down any decisions during data collection or analysis that weren’t in your plan
- Write down things that went wrong
- Keep track of your progress - when/where/who collected data and where is it kept
- Try taking digital notes - we use OneNote, Andy Roddick gave a great talk on using Obsidian.
- Try to keep notes in *one place,* and make sure they are backed up with your data.

What were you thinking?



What did you mean when you made this variable "cat2"? Did it mean cat? category? caterpillar? Something else?).

Raw data is usually nuanced.

Documentation and organization make data easier to understand and reuse—for you and others!

Example from Samantha Ming, "Bad Variable Names to Avoid," Medium, October 8 2018. https://medium.com/@samanthaming/bad-variable-names-to-avoid--bbab6614f36

scds.ca
scds@mcmaster.ca

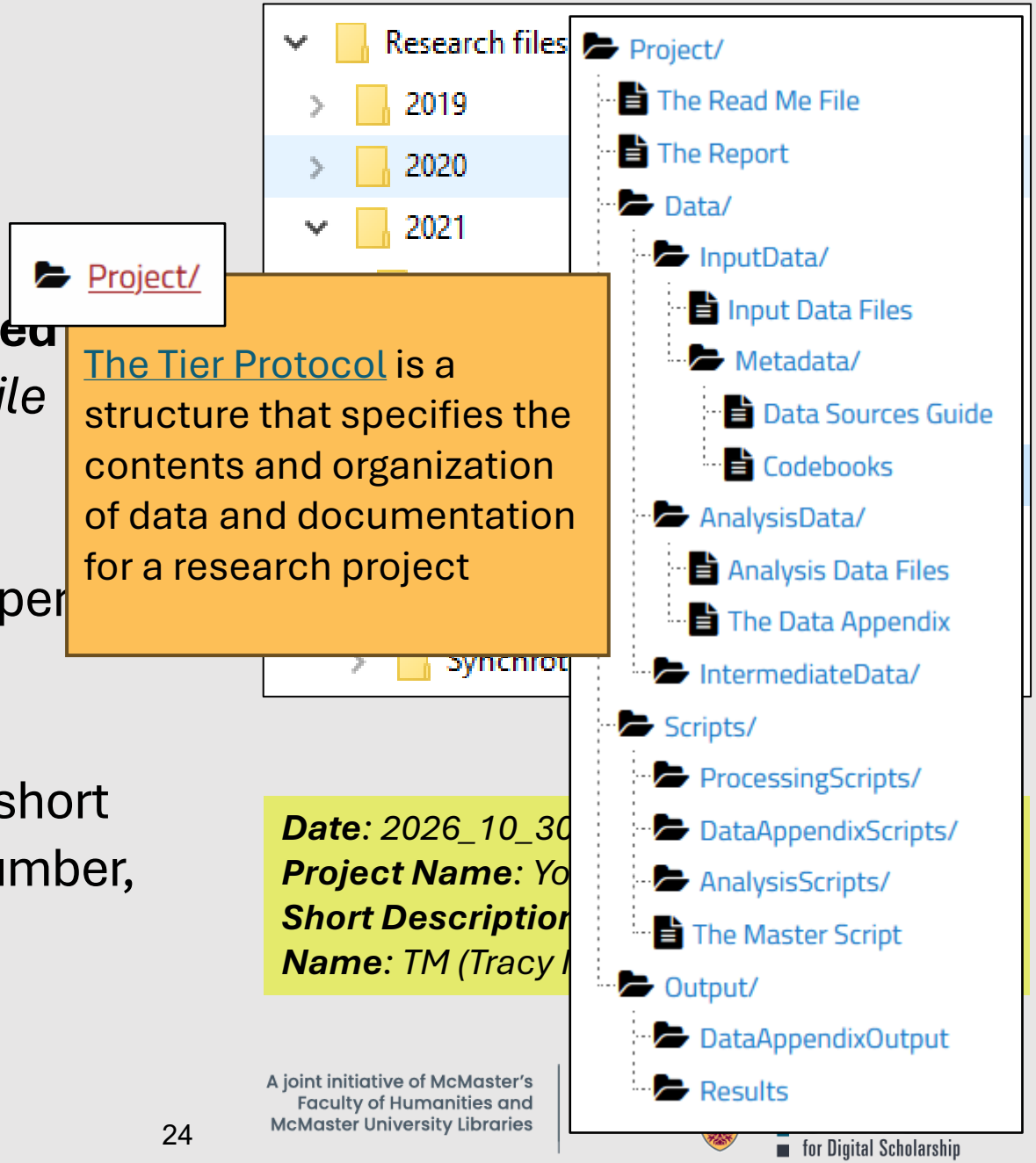
A joint initiative of McMaster's Faculty of Humanities and McMaster University Libraries
McMaster University
Sherman Centre for Digital Scholarship

Files + Folders

A good file organization system should be **descriptive**, **standardized**, and **implemented consistently**. *You should know what's in a file without opening it up!*

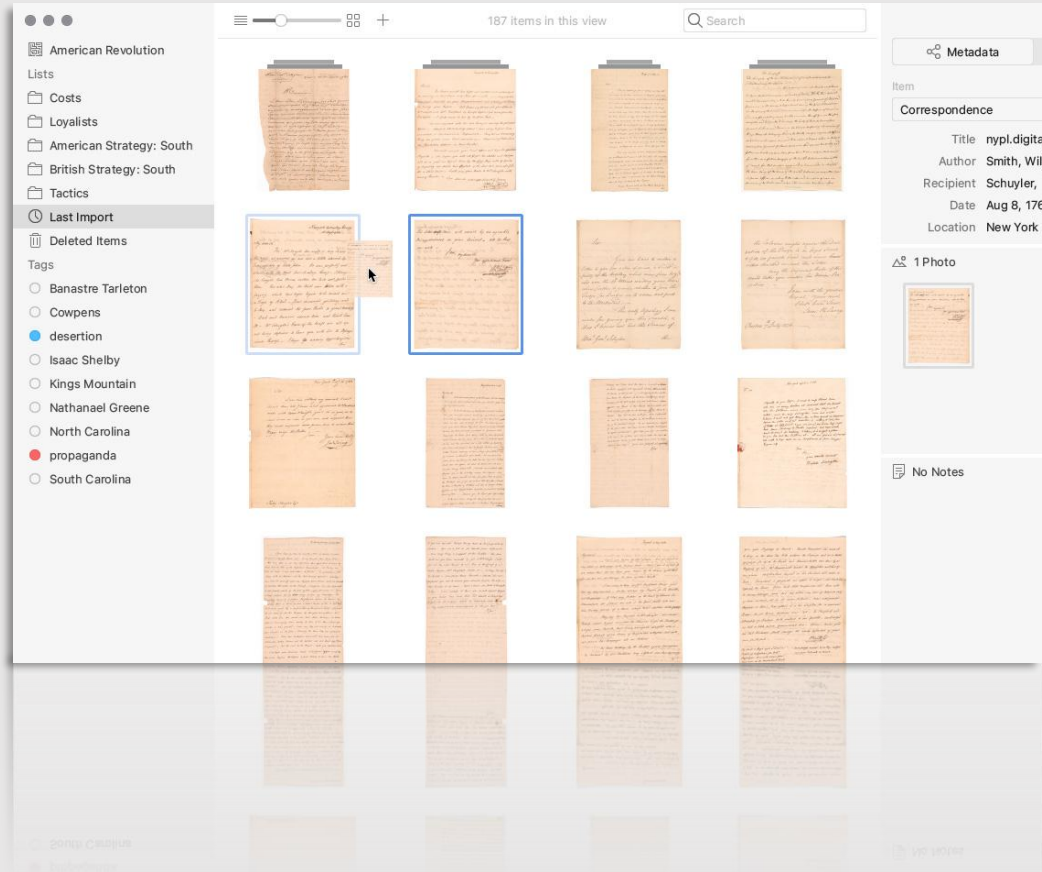
Folders: Organize by project, researcher, exper type, date, some combo (*year -> project)*

Files: Names can have date, project name, short description, initials of researcher, version number, and other metadata (like location).
2026-10-30_YOGuelph_StaffFG_TM.wav



Organizing: **Qualitative Data**

- **Reference Managers:** Zotero lets you collect, organize and cite publications. Annotate PDFs and add notes and tags. Plugin lets you click a button when you've found an article to add it as a citation. Master Libraries overview of reference managers.
- **Research Photos**: Add tags, notes, invert microfiche, organize into project sections, and more. Download and learn how to use Tropy.
- **Qualitative Analysis Tools**: QualCoder is an open-source web-based qualitative analysis tools. Import PDFs and Word documents to highlight themes, significant quotes you want to highlight, key understandings, and more. Our colleague Subhanya can help with this!

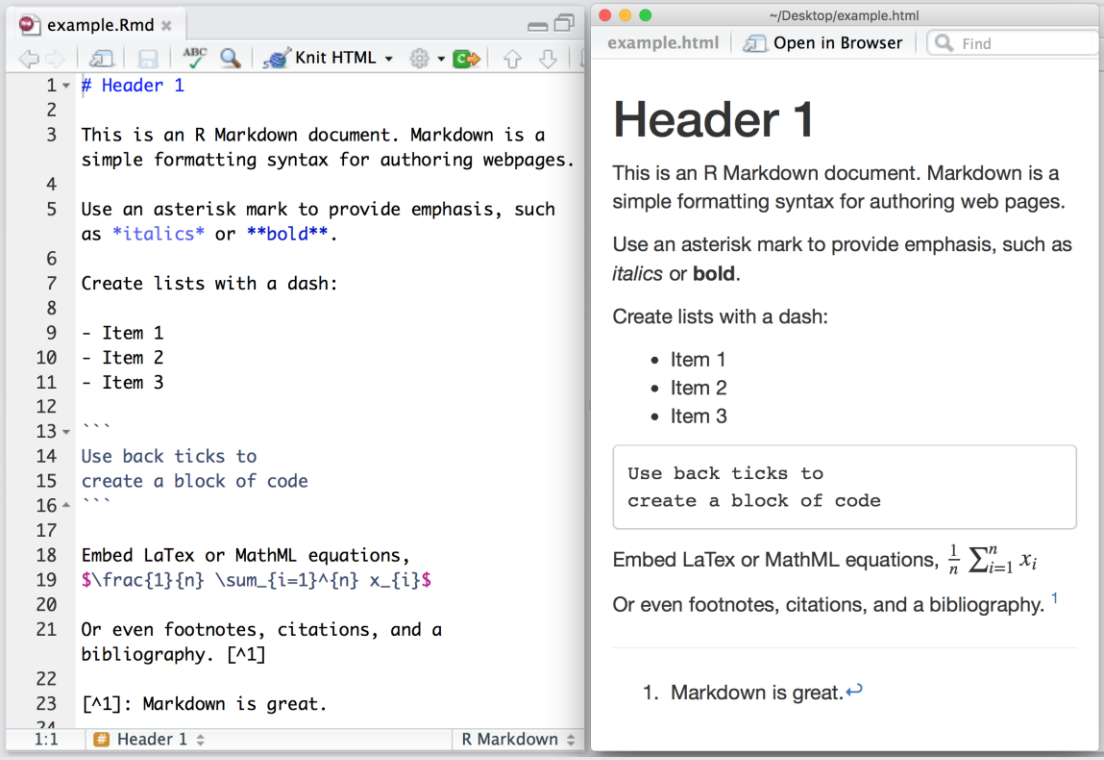


scds.ca
scds@mcmaster.ca

A joint initiative of McMaster's Faculty of Humanities and McMaster University Libraries
McMaster University
Sherman Centre for Digital Scholarship

Organizing: **Quantitative Data**

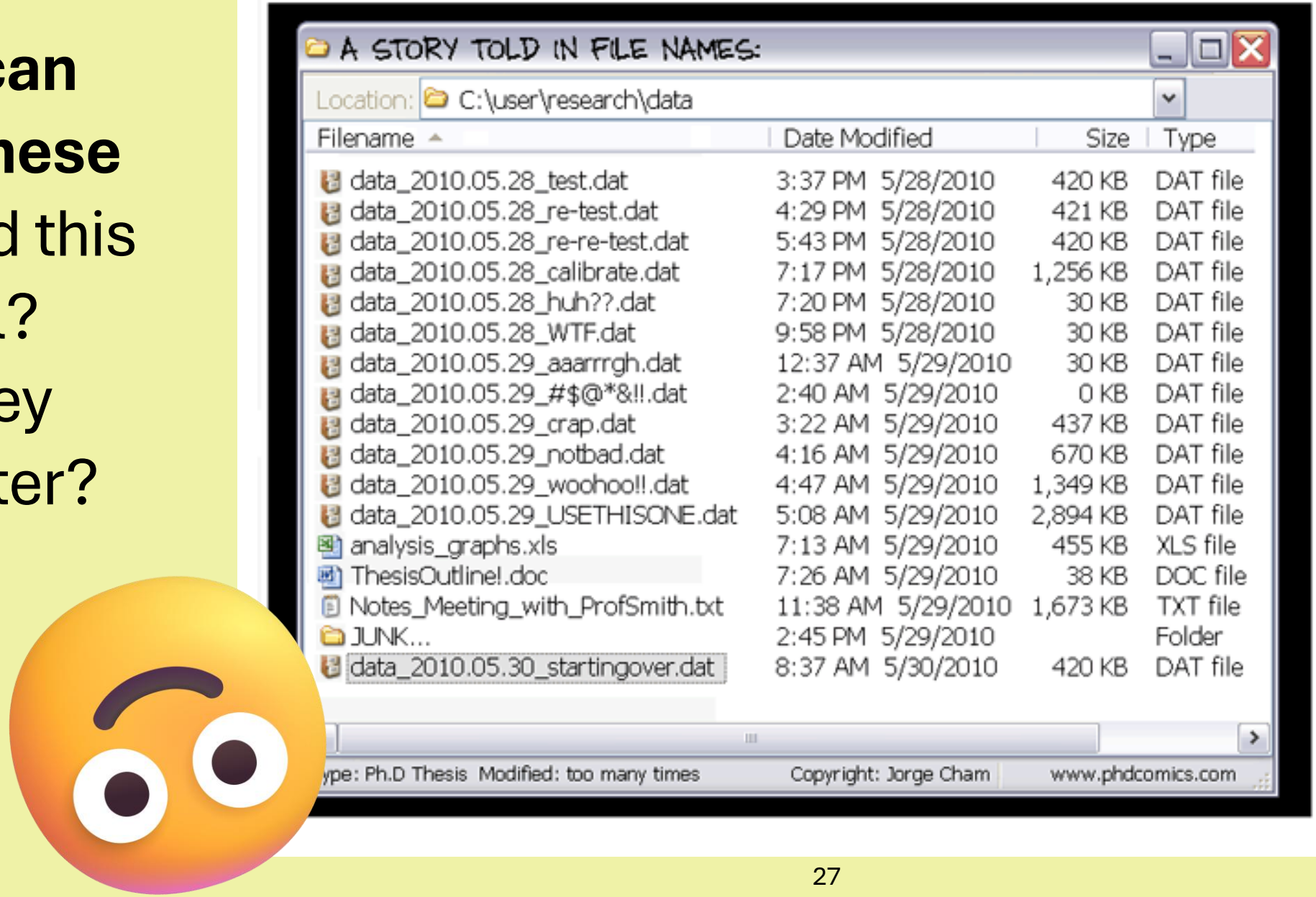
- **Codebooks + Data Dictionaries**: Explain layout and structure, define variables and elements, define ranges for units of measurement, acceptable ranges of values.
- **Code annotations + commit tags:** Describe what you're doing in your code and commits. Comments should be informative and current.
- *Dr. Antonio Paez – Reproducible Research Workflow with GitHub and R: https://github.com/paezha/Reproducible-Research-Workflow*
- *Dr. Jeremy Freese – Key principles for transparent social science code: https://www.boydetective.net/workflow/*



scds.ca
scds@mcmaster.ca

A joint initiative of McMaster's Faculty of Humanities and McMaster University Libraries
McMaster University
Sherman Centre for Digital Scholarship

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



- ☑ *1.Make a plan*
- ☑ *2.Organize & document data consistently*
- ☑ **3.Store & back up data securely**
- ☐ *4.Get ready for archiving & sharing*



r/GradSchool • 4 yr. ago
[deleted]

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 w
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 (fu
scheduled for next Wednesday). I ha
large chunk of my work will definite

The projects that I was working on w

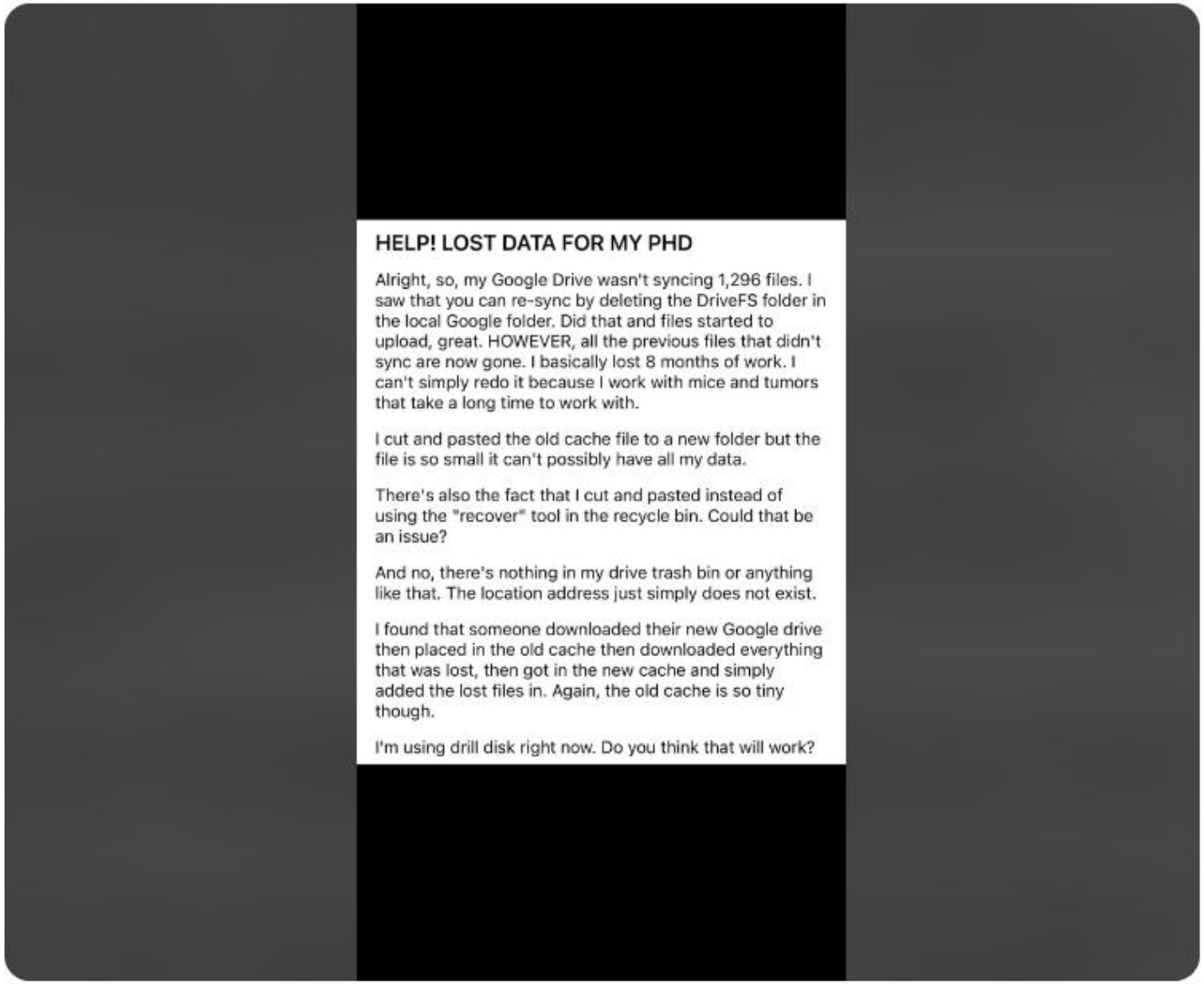
Read more

475 80

r/PhD • 1 yr. ago
JuriPlz

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

Need Advice



oy not checking if the

nce I just defended my PhD today

data, wrote the code, did all the

l so they had to take my PC and wipe it
l 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.

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

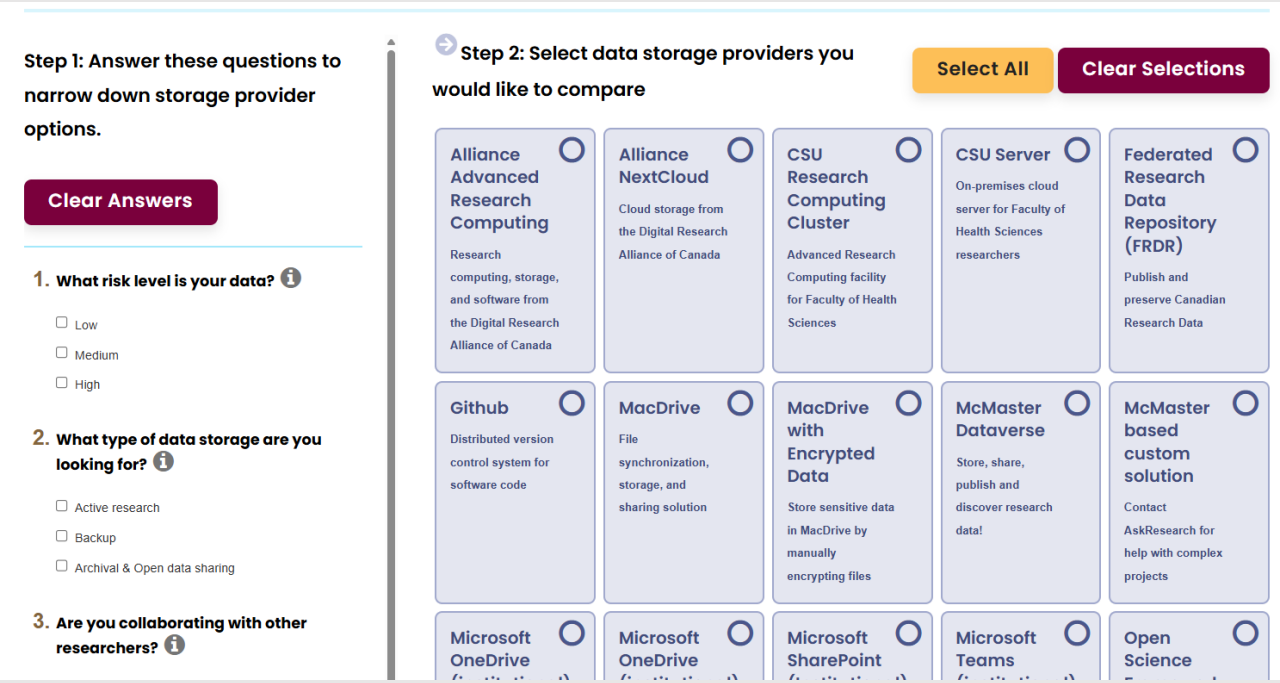


- On someone else's computer
- Automated backups, file versioning, and recovery
- File synchronization to your computer
- **Easy** to access and share files
- Subscription costs
- **Privacy can be opaque**

Basics: Data Storage

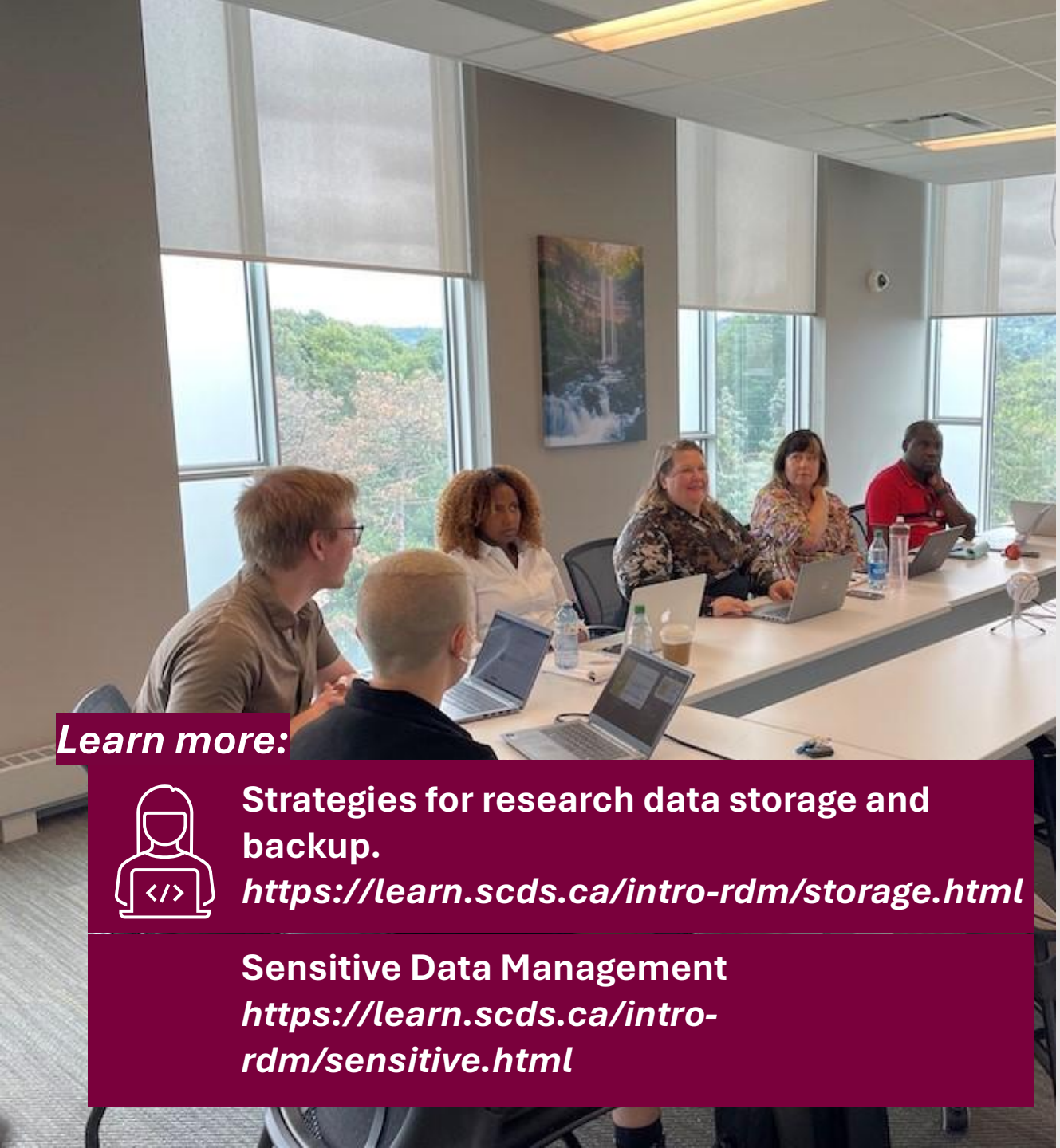
- If your data fits in the cloud, try **OneDrive**
  - Available to all campus users for free – **until you leave the institution**
  - Storage is shrinking
- If you need data to be portable, use **local storage**. Make sure local storage drives are encrypted with passwords.
- Use both if you can! Whatever you do, **make sure you make backups.**
- Research Data Storage Finder lets you compare options: **https://rdm.mcmaster.ca/finder**

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**



scds.ca
scds@mcmaster.ca

A joint initiative of McMaster's Faculty of Humanities and McMaster University Libraries
McMaster University
Sherman Centre for Digital Scholarship



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

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 or u.mcmaster.ca/rdm-appointments

A joint initiative of McMaster's Faculty of Humanities and McMaster University Libraries
McMaster University
Sherman Centre for Digital Scholarship

- ☑ *1.Make a plan*
- ☑ *2.Organize & document data consistently*
- ☑ *3.Store & back up data securely*
- ☑ **4.Get ready for archiving & sharing**



Basics: Data Archiving

Has anyone had their parents give them the entire contents of their childhood bedroom in a box?

1. **Clean up** your files – filenames, folders, file types!
2. **Save** your data in multiple safe places – but **choose** what is worth saving
3. Keep **documents** that describe your data

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

A Data Management Plan can help you prepare in advance!



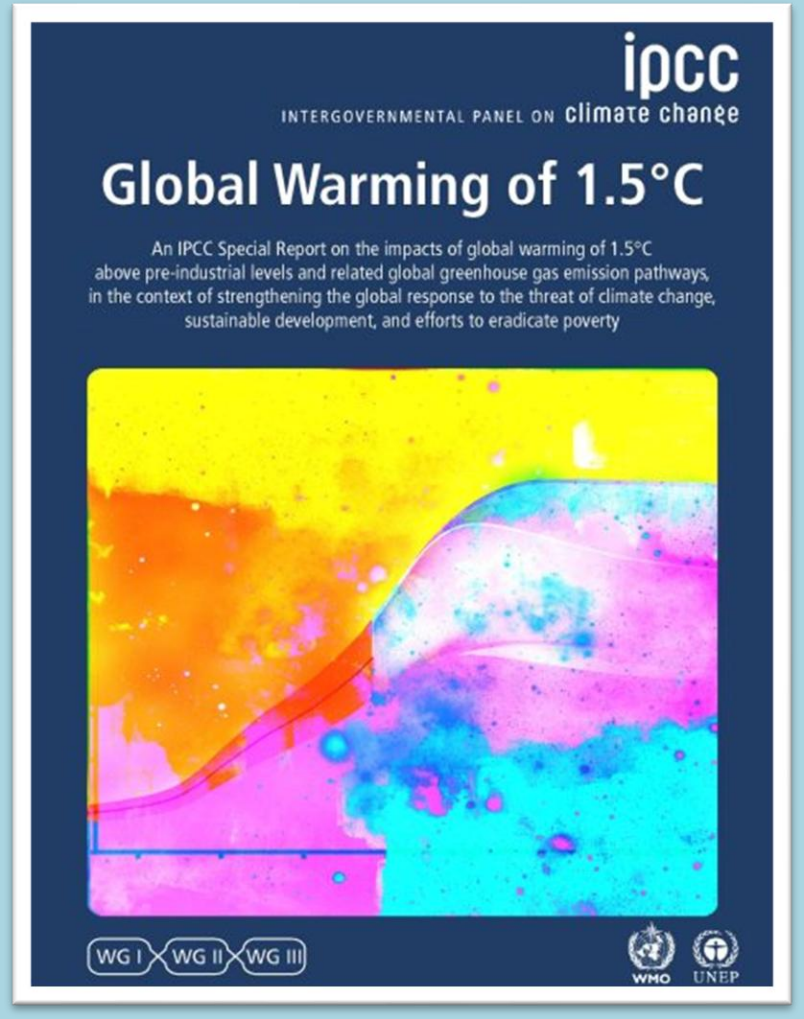
From u/SunshineEnjoyer on r/nostalgia on Reddit.

scds.ca
scds@mcmaster.ca

Science built on the power of open data:







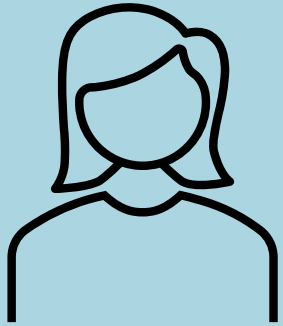
Why share data?

- Citation Impact
- Journal Requirements
- Funder Requirements
- Data Preservation



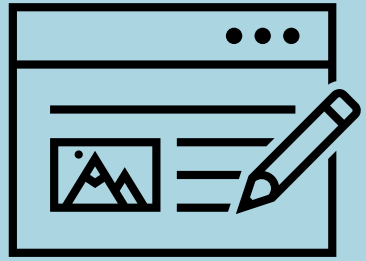
Denise Elbs on unsplash

What does data deposit look like?



researcher

"data are available on request"



publication/thesis

McMaster Dataverse
Our institutional data repository, a place for researchers to deposit and share research data

McMaster University



trusted data repository

Can be general (all data types) or domain specific (astronomy, genomics, political and social research).



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 asynchronous module - *coming soon.*
- **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 or u.mcmaster.ca/rdm-appointments

McMaster University | Sherman Centre for Digital Scholarship