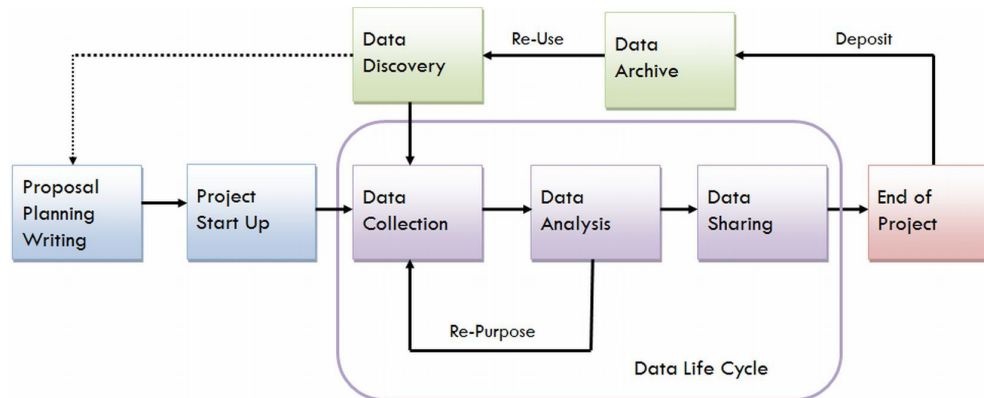


# Research Data Management in the Atlantic Provinces

Sarah Davis, Alison Farrell, Jamie MacKenzie, Erin MacPherson  
June 7, 2019

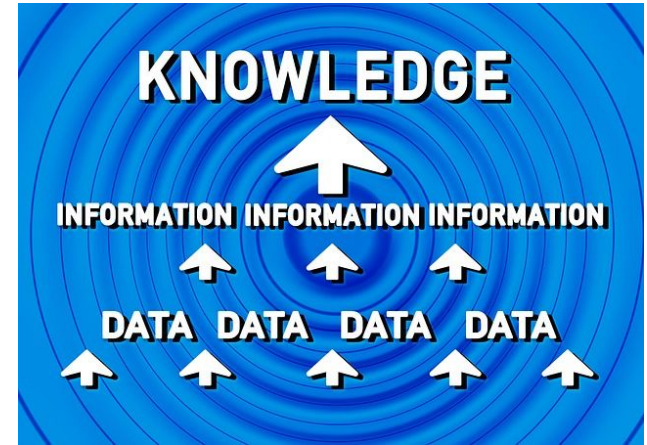
# Research Data Management - What is it?

- The everyday management of data produced during the lifecycle of a research project
- Good data management should describe the organization, storage, preservation and sharing of data



# RDM - Why is it important?

- Can help to turn data into knowledge
- Impact and visibility of research
- Encourage the discovery and reuse of data
- Ensure the accuracy and validity of data
- Long term preservation
- Compliance with ethics and privacy policies/regulations
- Compliance with funding authorities



# RDM in Canada - Portage

The Portage Network is dedicated to the shared stewardship of research data in Canada through:

- Developing a national research data culture
- Fostering a community of practice for research data
- Building national research data services and infrastructure

Portage was launched in 2015 by the Canadian Association of Research Libraries (CARL)

# The Tri-Agencies: NSERC, SSHRC, CIHR

## Statement of Principles on Digital Data Management

- Released in 2016, the Tri-Agencies provided a statement to “promote excellence in digital data management practices and data stewardship”
- Its aim was to provide researchers and research communities with guidance for adhering to current and future requirements of Tri-Agency grants.

# The Tri-Agencies: NSERC, SSHRC, CIHR

## Research Data Management Policy

- The Tri-Agencies released a draft Research Data Management Policy for consultation in spring 2018.
- Universities and other organizations (CARL, Portage, MUN, UofT, etc.)
- Expected to be released iteratively, starting Fall 2019
- Expected to include:
  - Institutional Strategies
  - Data Management Plan Requirements
  - Data Sharing Requirements

# RDM in the Atlantic Provinces



# RDM @ MUN

- Memorial Libraries Research Data Management Strategic Planning Project Team
  - Developed as one of 7 Strategic Planning Projects
    - Define and develop high quality, measurable research data curation tools and services
  - Created a Project Charter in Fall 2018
  - Team consists of Library staff and Librarians, as well as one member from AceNet.
  - Goal is to improve RDM practices, supports and resources





# RDM @ MUN



- Memorial University Research Data Management Institutional Strategy Working Group
  - Consists of members from the library, office of research, research ethics board, various academic departments, and has representation from all campuses.
  - Tasked with developing an institutional strategy to:
    - ensure Memorial is ready to meet the needs of our researchers
    - meet Tri-Agency funding requirements
  - Will provide recommendations for an operational process for the proposed strategy
  - Will develop and oversee a communications plan to increase awareness in the university community about the strategy

# Current RDM landscape @ MUN



- Lib Guide
- MUN Branded version of the DMP assistant
- Building a community of expertise
- No data repository - yet!
- Confusing options for active data storage
- No research data policies
- Inventory of data assets.



# RDM @ Dal



Image by [rawpixel](#) from [Pixabay](#)

- Started building a team back in 2015 (and earlier) and began to meet monthly
- In 2016, the team developed a **project charter** for research data management.
- The team plans new initiatives and share updates and news.

# RDM @ Dal



Image by [rawpixel](#) from [Pixabay](#)

The team consists of representatives relevant departments including:

- ❖ Office of Research Services
- ❖ Academic Technology Services
- ❖ Dalhousie Libraries

It is led by the Associate University Librarian, Research.

# Creating an Institutional Strategy for RDM

- A *Research Data Advisory Committee* was formed in late 2017 by the University Librarian.
- Committee consists of: Library representatives, AVPR, reps from Office of Research Services, Director of Academic Technology Services, reps from various Faculty, University Archivist.





# Creating an Institutional Strategy for RDM

- A draft of the RDM Institutional strategy, using the Portage template and guidance document, was completed in late 2018 by a small working group from this committee.
- Currently working on RDM survey, results will be incorporated in the strategy. This will provide information on current practices and attitudes help inform the strategy.



# Current RDM landscape @ Dal



- We have a data repository (Dataverse)
- DMP Assistant branded for Dalhousie
- LibGuide with links to educational resources and services available at Dalhousie and beyond (ie. Portage, etc)
- No research data policies, working on our strategy which may inform and help develop further policies.
- Several options at Dalhousie available for data management and storage.
- Have good and existing relationships with ITS, Information Security, Research Ethics and Office of Research Services.

# RDM @ UNB



## Research Data Management Services Working Group

- Director, Scholarly Technologies
- Functional Specialists
  - STEM Librarian (T. Zaraiskaya)
  - Data/GIS Librarian (S. Hanratty)
  - Scholarly Communications Librarian (M. Nason)
- Library Systems Group (J.Carter, B. Cassidy)
- Centre for Digital Scholarship (E. Moore)
- No formal faculty involvement... yet



# Activity and Engagement @ UNB



- RDM Outreach
  - Faculty events, meetings
  - ORS workshops
- GEAR: Graduate Essentials for Academic Research (2018-2019)
  - Series of workshops introducing core research skills and issues
  - Offered in collaboration with School of Graduate Studies
  - RDM-focused module
- Research Data Day (Fall 2018)
  - Offered in Partnership with Office of Research Services
  - Half-day Event
    - Led by Jeff Moon (Portage Director)
    - Presentation, faculty panel, discussion

# Current RDM Landscape @ UNB



- UNB Dataverse
- Support for RDM Assistant
- RDM Services Web Page & Subject Guide
- Next:
  - Policy Development
  - Institutional Strategy
- Challenges
  - Interest ≠ Deposit
  - Data (services) Deluge
    - Multiple “research data” entities on campus
    - Clarity, differentiation and/or partnership

*Image by James Mann via Flickr*

# Canadian RDM Survey Consortium

“Several Canadian universities have committed to working together, using a common survey instrument to gather information about their respective researcher communities and generate a richer understanding of their users’ RDM practices and attitudes.”

- Dalhousie University
- McGill University
- McMaster University
- Queen’s University
- Ryerson University
- University of Alberta
- University of British Columbia
- University of Ontario Institute of Technology
- University of Ottawa
- University of Toronto
- University of Victoria
- University of Waterloo
- University of Windsor
- Western University
- Memorial University
- University of New Brunswick

# Purpose

- To better understand local and national RDM needs
- To generate a richer understanding of disciplinary practices
- To inform the development of Research Data Management services
- To make researchers aware of RDM issues

# Content

- Data Storage
  - Size
  - Formats
  - Location
  - Backups
  - Documentation
- Data Sharing
  - Practices
  - Attitudes
  - Repositories
  - Restrictions
- Data Services
  - Data Management Plans
  - Needed/wanted supports
- Demographics



# A few results from our colleagues

In an analysis of survey responses from Health Sciences researchers at 8 Institutions, Barnes et al report that:

- 20% of researchers do not know how much data they create
- 42.3% create less than 10GB of data per project
- 28.7% create more than 50GB of data per project
- Many store their data on hard drives, flash drives or external hard drives
- Some store their data on cloud servers or Institution shared drives
- Few store their data in repositories
- 89% would like or would need assistance when writing DMPs

# More results from our colleagues

In a report of Science and Engineering researchers across five Canadian institutions, it was found that:

- 46% of respondents use less than 50GB per project
- A majority of respondents use external data repositories for data deposit
- 83% of Tri-Agency funded researchers need or want assistance in preparing DMPs
- There is high interest in RDM services

# Data Sharing results

Biggest reason for sharing data:

**Collaborative Scholarship/Encourage collaborative science!!!**

Reasons for not sharing:

**For Health Researchers:** Privacy/legal/security reasons

**For Science Researchers:** Incomplete data, still wish to derive value from data,  
lack of willingness to share

Barnes, L., Brown, H., Cheung, M., Cooper, A., Dorgan, M., Ellis, U., Gertler, M., Miller, K., Morris, M., Sikora, L. (2018, June). Telling Canadian Research Data Management (RDM) Stories in the Health Sciences. Poster presented at the Canadian Health Libraries Conference 2018 in St. John's, Newfoundland

Sewerin, C., Barsky, E., Dearborn, D., Henshilwood, A., Hwang, C., Keys, S., Mitchell, M., Spence, M., Szigeti, K., & Zaraiskaya, T. (2016, June). **Research Data Management (RDM) Needs of Science and Engineering Researchers: A View from Canada**. Poster presented at IASSIST 2016 Conference in Bergen, Norway.

# Survey process and results



# Survey Experience @ Dal



- Initial survey conducted in 2016 (Science, Engineering, Computer Science, Agriculture)
- Extremely low response rate
- Tried again in 2019 (all Faculty @ Dal). Submitted to Ethics in March.
- Ethics said this was considered “assessment” (in 2016 we were approved as research).



Image by [Wokandapix](#) from [Pixabay](#)



Image by [jmw02824](#) from [Pixabay](#)

*What's the difference???*

# Survey Experience @ Dal



- We re-submitted as research and are awaiting feedback.
- Lessons learned:
  - Large institution = things take time
  - Ethics Boards change
  - Getting in the “survey line-up” can cause delays depending on priorities/timelines
- Stay tuned!

# Survey Experience @ MUN

- No ethics approval needed - assessment (even if we want to present/publish)
- Center for Institutional Analysis and Planning
  - Helped to time and distribute survey
- Institutional Survey Oversight Committee
  - Needed to approve survey before dissemination
  - Suggested the survey was too long!
- Three surveys
  - Data Storage
  - Data Sharing
  - Data Services

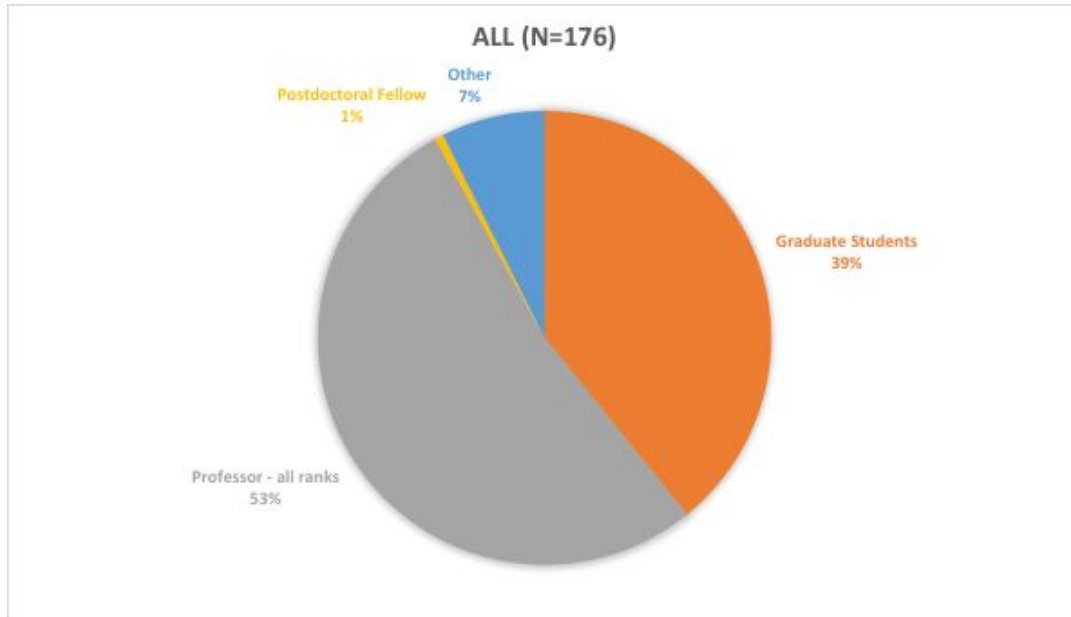
# Survey Experience @ MUN

- Distribution
  - We planned to send to all Faculty, Graduate Students, Postdoctoral Students and Medical Residents.
- Problems with distribution
- Problems with demographics
- Too many surveys!

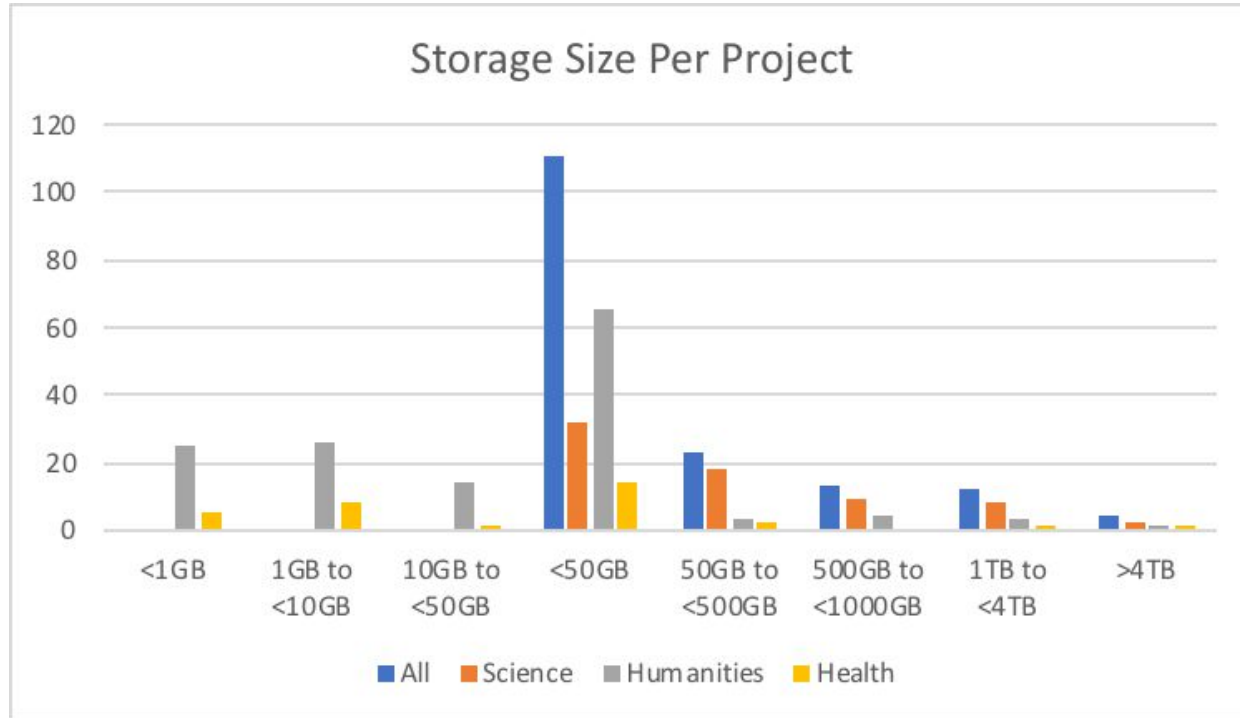


# MUN - Results (Very Preliminary!)

255 Total Responses

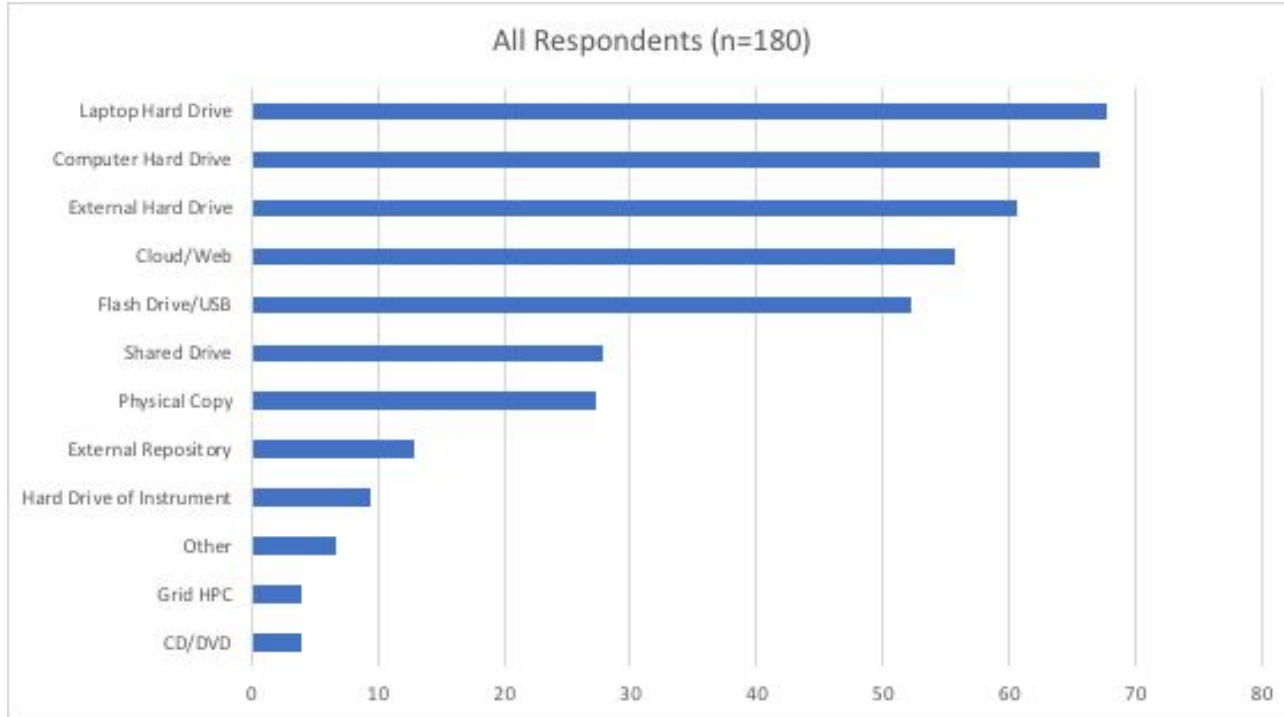


# How much data storage do you use?



- 32% of Humanities respondents use less than 1GB data per project (n=79)
- 28% of Health respondents use less than 1GB data per project (n=18)
- 57% of Science respondents say they either do not, or are unsure if they have enough documentation for others to understand and use their data (n=81)

# Where do you store your data?



- Almost 20% of Science respondents use external data repositories, while only 6% and 11% of Humanities and Health respondents (respectively) do.
- Only 15% of Humanities respondents use institutional shared drives (37% - Science, 44% - Health)



# Number of Locations that Data is stored

- Promisingly, 77% of respondents store their data in 3 or more locations.
- However, we do not know if they have 3 backups or if they have their data spread out over multiple locations



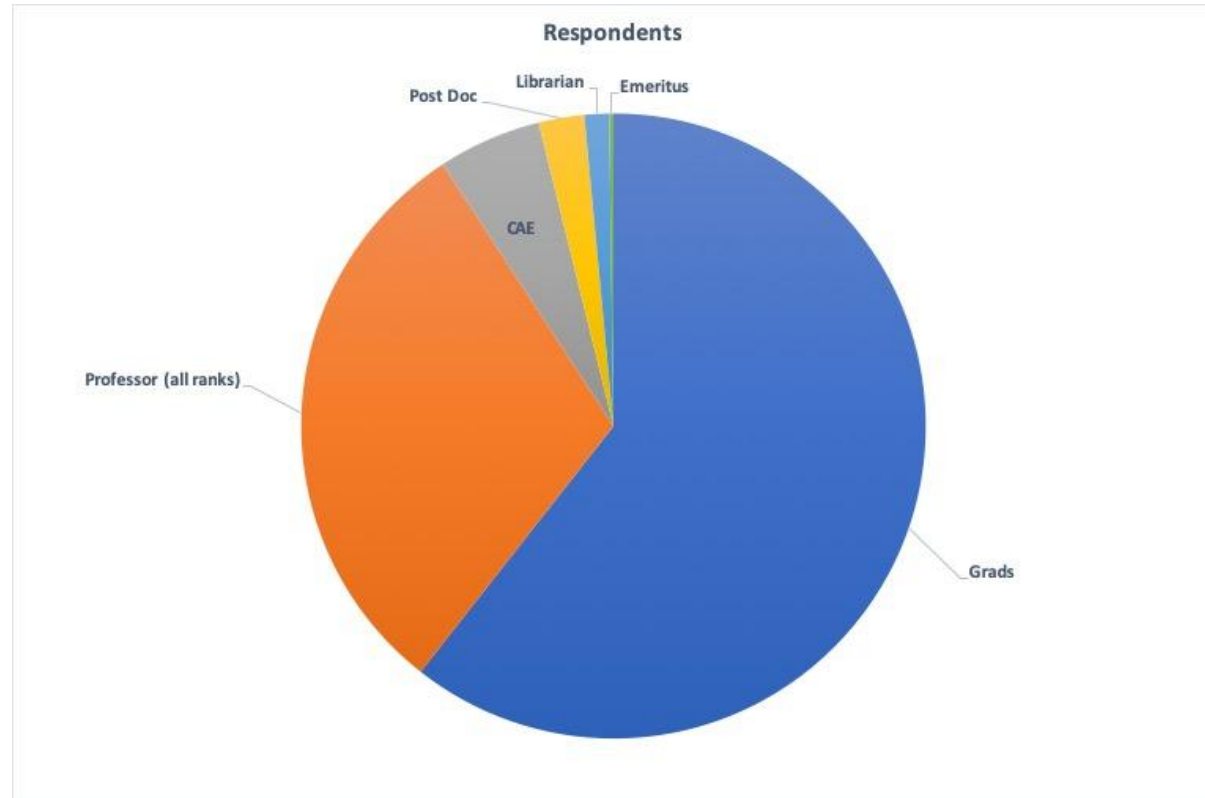
# Survey Experience @ UNB



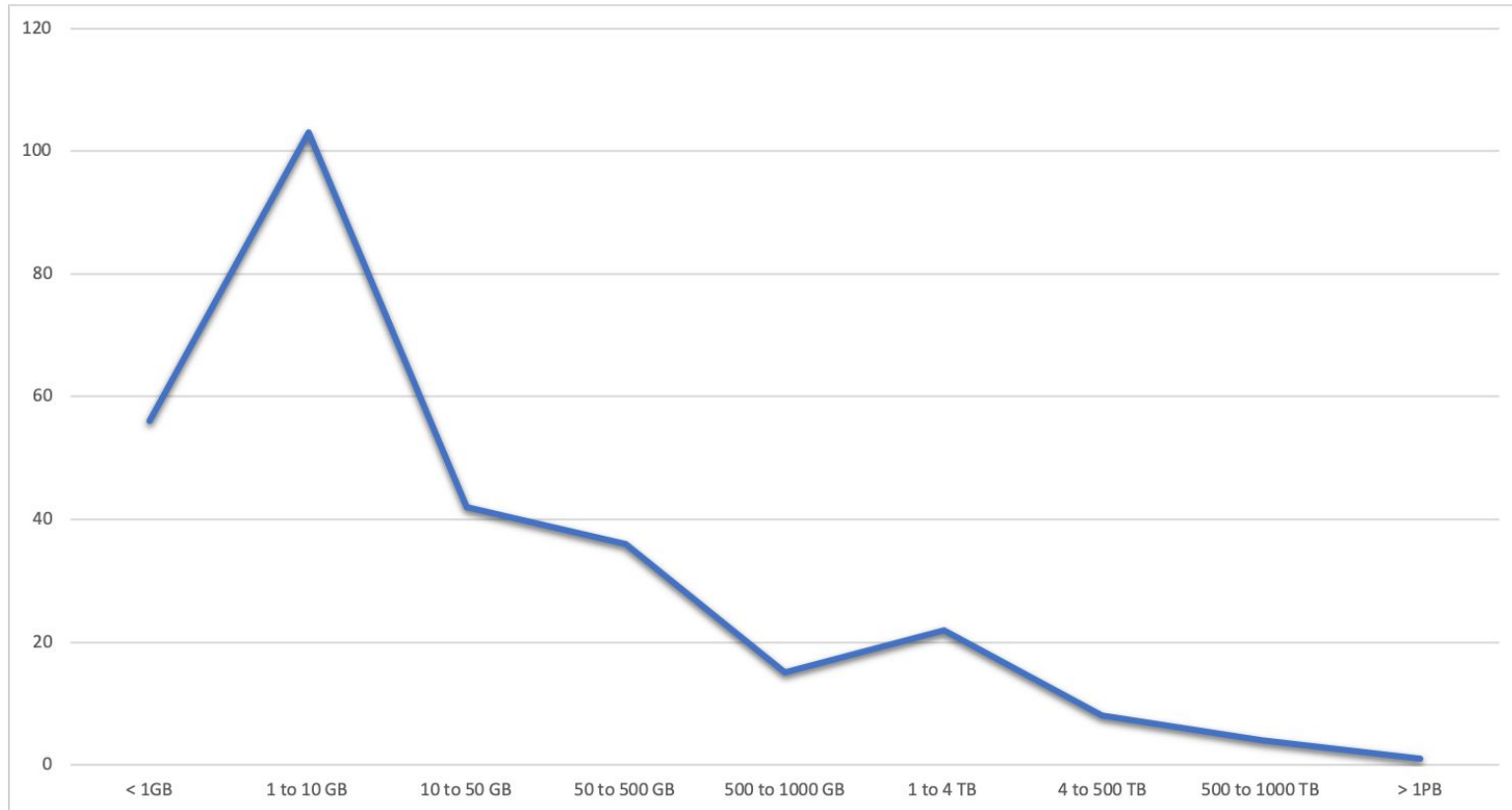
- Prior work: Ithaka S&R (2016)
- REB Application Approved: Early February
- Acquire / Prep Contact Data: January - February
- Pre-survey announcement (from Dean): Mid-February
- Survey Administered: Feb. 21 - Mar. 15
  - Lime Survey (local)
  - Direct email invitation
  - Faculty, researchers, librarians, graduate students... ca. 2600
  - Two campuses
- Interest and support!!
  - REB, ORS, Secretariat, School of Graduate Studies

# Survey Response

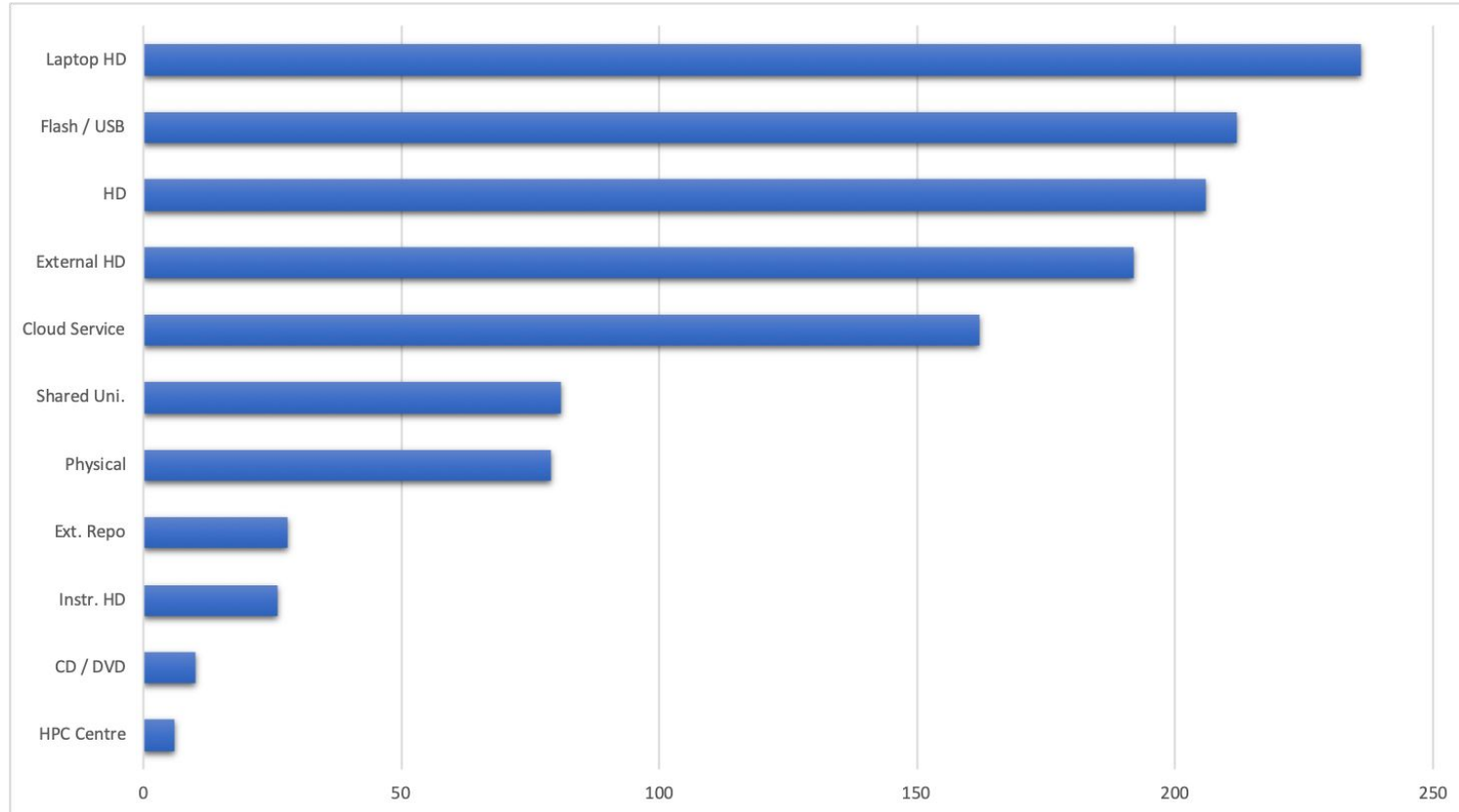
- 467 valid surveys
- 17.6% response (n=2639)
  - Grad: 16.2%
  - Prof: 23%
- Top Depts/Faculties:
  - Education
  - Business
  - Biology
  - Computer Science
  - Forestry / Env. Mgmt
  - Psychology
  - Electrical / Com. Eng'g
  - Nursing
  - Kinesiology



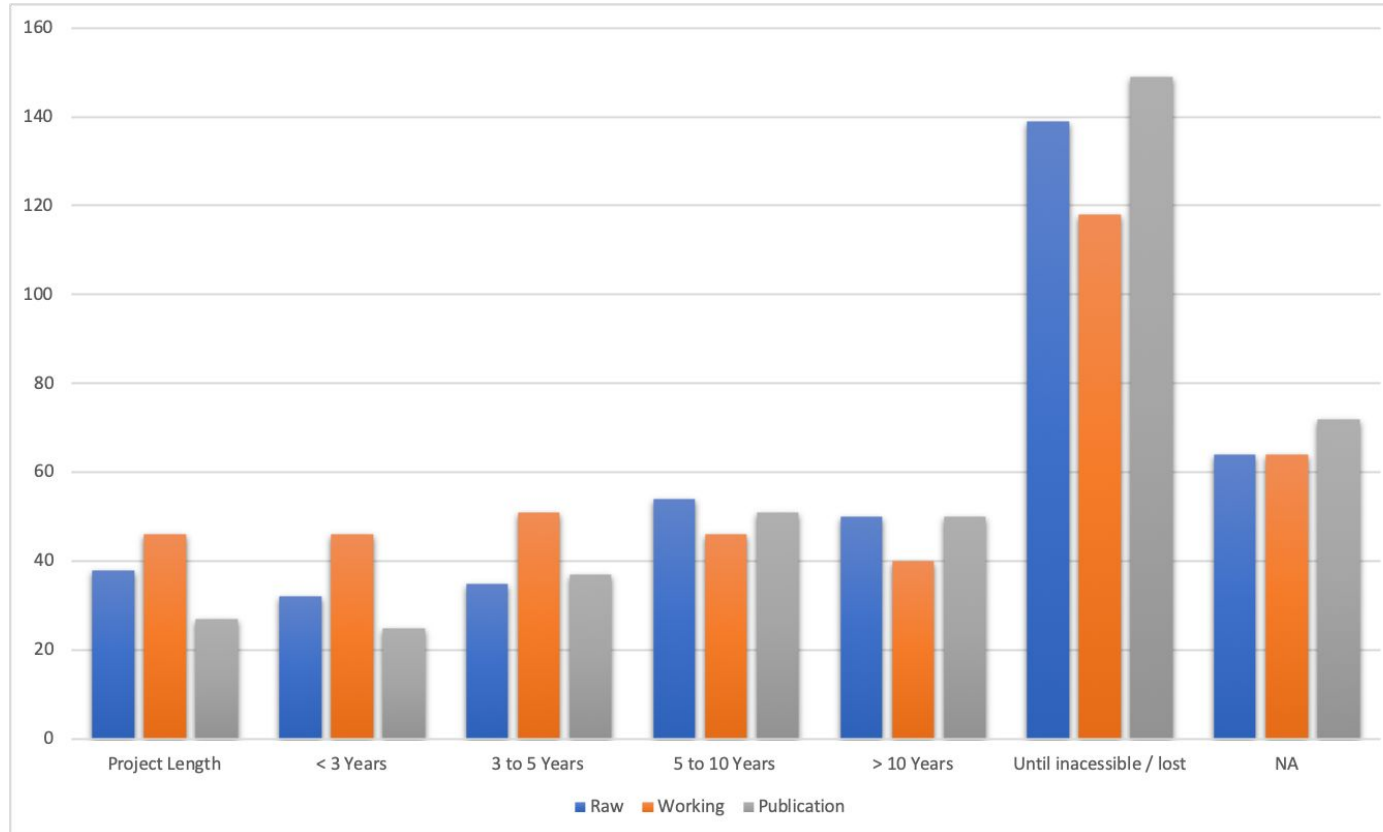
# How much data storage do you use?



# Where do you store your research data?



# How long do you keep your data?



# Data Sharing

## Do you share your research data?

- No: 49%
- Yes: 51%
  - Personal requests: 74%
  - Supplements to publication: 21%
  - Restricted online access: 15%
  - Repositories and websites: 8-9%

## Do restrictions prevent sharing?

- No: 65%
- Yes: 35%
  - Privacy and confidentiality
  - Need to publish first
  - Contractual obligations
  - May jeopardize rights

77%

*I am willing to share my research data and associated methods/tools/algorithms!*



# Comments

- Concerns and perceptions
  - Skepticism and concern about usefulness to others
    - Messy, incomplete data can't be useful
    - Data “minus our theoretical lens”
  - “Researchers are [already] dealing with data management all the time”
  - Sharing doesn't always work
    - Ethics
    - Confidentiality
  - Does this really apply to me / my discipline?
    - Humanities

# Comments

- Positive perceptions of RDM services... (mostly)
- HELP!!!
  - "I need help... my lab needs help"
  - "some of the data storage has been rather messy... my lab and our collaborators would be interested in learning how to rectify"
  - "... the general researcher is too unfamiliar with the process"
  - "The question of how and when best to share it is very relevant... I have a lot of uncertainty"
  - "I am perfectly willing [...] but have no time or funding to deal with the details"

# Takeaways



# Next Steps

# Lessons learned

- Survey fatigue is real!
- Delivering three surveys is awkward for dissemination
- Timing is everything
- REBs are not all created equal

# Key Points

- Researchers are interested in RDM!
- Researchers need support for RDM!
- There is great interest in sharing research data, but there are many barriers too.



- Relationship building is key!

# Where we go from here

## Data Analysis

- More analysis on our individual and combined survey data.
- Comparison across disciplines

## Data Sharing

- Deposit to respective dataverses to add to growing body of RDM survey data

## Service building

- Continue to develop and expand RDM services based on best practices and survey data

# Thank You

Questions?