Research Data Management
What is Research Data Management

- How you organize, structure, store, and care for data used in or generated by a research project

- Planning:
  - How data will be used during a project
  - How data will be looked after following the end of the project
  - A formal data management plan is often part of this
Why Plan for Data Management

- Accuracy
- Data & Metadata Definition
- Security
- Preservation
- Accessibility
- Funding Agency Requirements
Accuracy

- Data can be fragile
- Ensuring data is not compromised
- Possible issues to consider:
  - User error
  - Data corruption
  - Hardware failure
Data & Metadata Definition

- What data are being collected?
- What formats are the data in:
  - Digital Information
    - What formats?
  - Hard Copy
  - Physical samples
- How are the data described
  - Info describing content – text description, keywords, etc.
  - Data definition – column names, SQL database definition, etc.
Security

- Some data may have considerations that lead to limiting access:
  - Personal medical data
  - Ethical issues
  - Intellectual Property
    - Copyright
    - Trade secrets
    - Licensing Agreements
  - National Security
Preservation

- Data can be useful beyond its intended purpose
- Might the data be useful to someone in 10 years, 20 years, 100+ years?
- File formats change
  - Can you open a Word Perfect 5.1 file?
- Hardware technologies change
  - Can you read an 8-inch floppy disk?
- Software exists to help with this
Accessibility

- Availability of data to others
  - External Collaborators
  - Other Researchers
  - Public

- Open Data movement
  - Associated with Open Access movement
  - Public availability of data
  - Popular with governments, more funding agencies requiring open data
Funding Agency Requirements

- More funding agencies are requiring data management planning
  - More international agencies than Canadian ones
  - For now!

Funders usually have specific requirements as part of application process.

- Plans required at application
- Data deposit requirements
- Data preservation periods (e.g. 3 years after project ends)
- Access/Data sharing requirements
NSF Data Management Plan


- Data Management Plans required for all proposals
  - Describe types of data to be produced
  - Describe standards for data and metadata
  - Access policies
  - Data re-use policies
  - Plans for archiving and sharing data

- California Digital Libraries Tool for US Funding Agencies:
  - [https://dmptool.org/](https://dmptool.org/)
UK DCC Data Management Plans

- Digital Curation Centre:
  - [http://www.dcc.ac.uk/resources/data-management-plans](http://www.dcc.ac.uk/resources/data-management-plans)

- Variety of resources for data management planning
  - Guide for development of DMPs
  - List of UK Funding agency requirements

- DMPOnline tool: [http://dmponline.dcc.ac.uk/](http://dmponline.dcc.ac.uk/)
Canadian DMP Initiatives

- **CARL Data Management Page**
  - Guidance for Canadian Researchers
    - Project will become Portage

- **U of A Library DMP Builder**
  - [https://dmp.library.ualberta.ca/](https://dmp.library.ualberta.ca/)
University of Saskatchewan

▪ Research Data Management Support
  • http://www.usask.ca/ict/services/research-technologies/data-management/index.php
  • research_computing@usask.ca

▪ Procedures for Stewardship of Research Records at the University of Saskatchewan
  • http://research.usask.ca/documents/ethics/research_records_stewardship.pdf

▪ We’re here to help!
Storage Options

- Where to put the data? Where to back it up?
- Desktop/Laptop
- External Disks
- Network Attached Storage
  - Cabinet
  - Datashare
  - Departmental server
- Compute Canada/WestGrid – Silo (U of S storage site)
Sharing Options

- Getting data to others who can use it
- How will data be shared?
  - Small amounts (several GB):
    - OwnCloud (U of S, WestGrid)
    - Dropbox or other cloud services
    - FTP/SFTP
    - Mailed disks
  - Large Amounts (10s of GB to TB+)
    - Globus (Compute Canada)
  - Repositories
    - Over to Jessica!