

Create a Data Management Plan (DMP) for Your Research Project

March 4th, 2024

LILY REN, SCHOLARLY COMMUNICATIONS & METRICS LITERACY LIBRARIAN

STEPHANIE QUAIL, DIRECTOR

OPEN SCHOLARSHIP DEPARTMENT, YORK UNIVERSITY LIBRARIES

How to cite this presentation:

Ren, L., & Quail, S. (2024). Create A Data Management Plan (DMP) For Your Project. York University Libraries.

Original content in this slide deck is licensed under the [Creative Commons Attribution 4.0 International license](https://creativecommons.org/licenses/by/4.0/).



Land Acknowledgement

York University recognizes that many Indigenous Nations have longstanding relationships with the territories upon which York University campuses are located that precede the establishment of York University. York University acknowledges its presence on the traditional territory of many Indigenous Nations. The area known as Tkaronto has been care taken by the Anishinabek Nation, the Haudenosaunee Confederacy, and the Huron-Wendat. It is now home to many First Nation, Inuit and Métis communities. We acknowledge the current treaty holders, the Mississaugas of the Credit First Nation. This territory is subject of the Dish with One Spoon Wampum Belt Covenant, an agreement to peaceably share and care for the Great Lakes region.

About Us



STEPHANIE QUAIL

Director, Open Scholarship
department, York University
Libraries



LILY REN

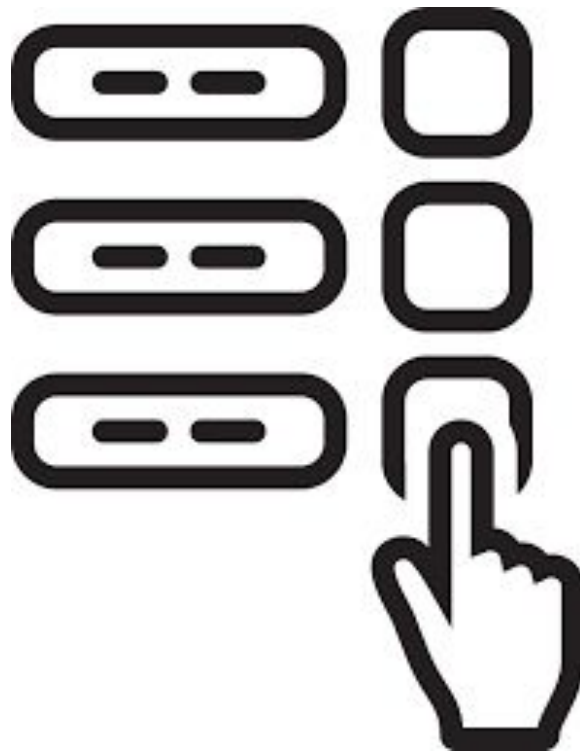
Scholarly Communications &
Metrics Literacy Librarian, Open
Scholarship department, York
University Libraries



MINGLU WANG
***On Sabbatical**

Research Data Librarian, Open
Scholarship department, York
University Libraries

Poll 1



Workshop Objectives

- Understand the importance of research data management (RDM)
- Learn key data management plan (DMP) considerations and best practices
- Explore the DMP Assistant tool and gain familiarity with the York U template

Agenda

1. The Tri-Agency RDM policy and funder requirements
 - Guest Speaker from the Faculty of Science (NSERC)
2. Library support and services for RDM, DMP, and grants
3. What is a DMP and why do I need one?
4. DMP tasks and the research data lifecycle
5. Crafting a DMP using the DMP Assistant tool and York template

Commonly used terms

6. RDM = Research Data Management
7. DMP = Data Management Plan

Guest Speaker

Randy Lewis, Professor, Department of Physics & Astronomy, Faculty of Science

- Principal Investigator for NSERC Subatomic Physics grant

Randy will cover:

- A brief description of his research
- Nature of the data for this project
- Considerations for developing your Data Management Plan for a Tri-Agency grant
- Recommendations for researchers

DMP Requirements for



• **CIHR**

- Network Grants in Skin Health, Bone Health and Muscular Dystrophy (October 2022)
- Team Grants: Strengthening the Health Workforce for System Transformation (June 2023)
Formerly "Virtual Care/Digital Health Team Grants"
- Team Grants: Lung Health (July 2023)
- Operating Grant: Clinical Trials Projects (July 2023)
- Team Grants: Improving Health and Administrative Data and Monitoring for Rare Diseases (August 2023)
- Infectious and Congenital Syphilis in Canada: Implementation and Intervention Research Response (September 2023)
- Canadian Consortium on Neurodegeneration in Aging (September 2023)
- Team Grants: Embracing Diversity to Achieve Precision and Health Equity (November 2023)
- Team Grants: HIV/AIDS and Sexually Transmitted and Blood-Borne Infections Community-Based Research (TBD)
- Directed Grant: Canadian Immunization Network Renewal (TBD)

• **NSERC**

- Subatomic Physics Discovery Grants - Individual and Project (November 2023)

• **SSHRC**

- Partnership Grants Stage 2 (October 2023)

RDM and the Tri-Agency

• Tri-Agency RDM Policy

- “The agencies expect the research they fund to be conducted to the highest professional and disciplinary standards, domestically and internationally. These standards support research excellence by ensuring that research is performed ethically and makes good use of public funds, experiments and studies are replicable, and research results are as accessible as possible. **Research data management (RDM) is a necessary part of research excellence.**”
- “The agencies believe that research data collected through the use of public funds should be responsibly and securely managed and be, where ethical, legal and commercial obligations allow, available for reuse by others. To this end, the agencies **support the FAIR (Findable, Accessible, Interoperable, and Reusable) guiding principles for research data management and stewardship.**”
- “The agencies therefore recognize that there are legitimate **differences in the standards for RDM** among the disciplines, areas of research, and modes of inquiry that the agencies support.”

• Tri-Agency grants with DMP requirements

- “All grant proposals submitted to the agencies should include methodologies that reflect best practices in RDM. For certain funding opportunities, the agencies will require data management plans (DMPs) to be submitted to the appropriate agency at the time of application, as outlined in the call for proposals; in these cases, the DMPs will be considered in the adjudication process.”

Research Data Management Support @ York University Libraries

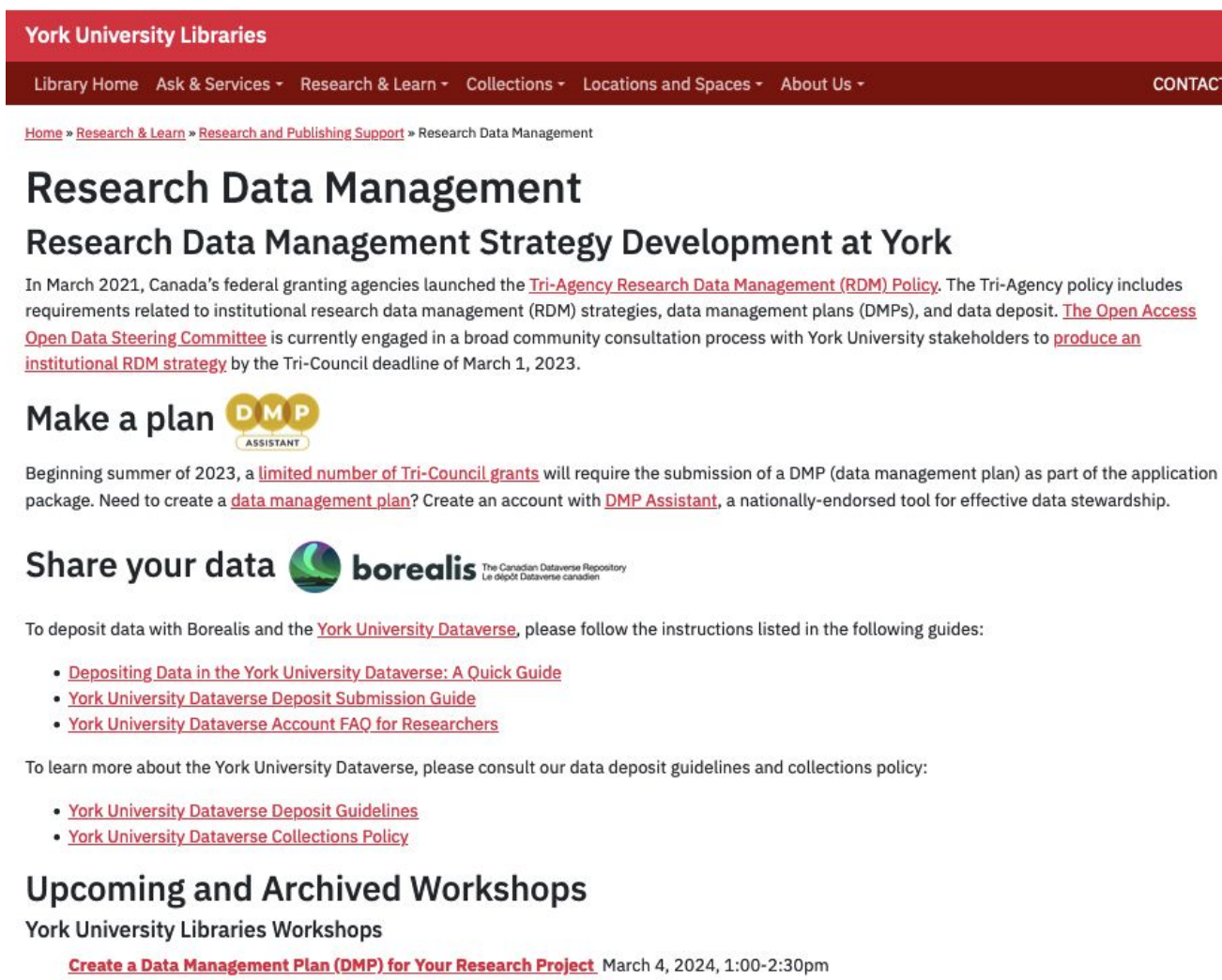
• Workshops

- Best Practices for Managing Research Data
- Create a Data Management Plan (DMP) for Your Research Project
- [Publishing Your Research Data in a Data Repository](#) (March 18th | 10:30-12:00)
- Safeguarding Your Research Data

• [Web guide](#)

• Individual consultations and support

- Support the deposit and curation of data in the institution data repository York U @ Dataverse
- Support in creating and reviewing data management plans (DMPs)
- Support in all areas of research data management
- RDM services and support as part of the [Libraries' Grant Support Services](#)



The screenshot shows the York University Libraries website. The header includes navigation links: Library Home, Ask & Services, Research & Learn, Collections, Locations and Spaces, About Us, and CONTACT. The main content area is titled "Research Data Management" and "Research Data Management Strategy Development at York". It contains text about the Tri-Agency Research Data Management (RDM) Policy, a link to the Open Access Open Data Steering Committee, and a section titled "Make a plan" with a DMP ASSISTANT logo. Below this is a section "Share your data" with the Borealis logo and a list of guides for depositing data in the York University Dataverse. The page also mentions upcoming and archived workshops, including one on "Create a Data Management Plan (DMP) for Your Research Project" on March 4, 2024.

York University Libraries

Library Home Ask & Services Research & Learn Collections Locations and Spaces About Us CONTACT

Home » Research & Learn » Research and Publishing Support » Research Data Management

Research Data Management

Research Data Management Strategy Development at York

In March 2021, Canada's federal granting agencies launched the [Tri-Agency Research Data Management \(RDM\) Policy](#). The Tri-Agency policy includes requirements related to institutional research data management (RDM) strategies, data management plans (DMPs), and data deposit. [The Open Access Open Data Steering Committee](#) is currently engaged in a broad community consultation process with York University stakeholders to [produce an institutional RDM strategy](#) by the Tri-Council deadline of March 1, 2023.

Make a plan

Beginning summer of 2023, a [limited number of Tri-Council grants](#) will require the submission of a DMP (data management plan) as part of the application package. Need to create a [data management plan](#)? Create an account with [DMP Assistant](#), a nationally-endorsed tool for effective data stewardship.

Share your data **borealis** The Canadian Dataverse Repository Le dépôt Dataverse canadien

To deposit data with Borealis and the [York University Dataverse](#), please follow the instructions listed in the following guides:

- [Depositing Data in the York University Dataverse: A Quick Guide](#)
- [York University Dataverse Deposit Submission Guide](#)
- [York University Dataverse Account FAQ for Researchers](#)

To learn more about the York University Dataverse, please consult our data deposit guidelines and collections policy:

- [York University Dataverse Deposit Guidelines](#)
- [York University Dataverse Collections Policy](#)

Upcoming and Archived Workshops

York University Libraries Workshops

[Create a Data Management Plan \(DMP\) for Your Research Project](#) March 4, 2024, 1:00-2:30pm

Library Support for Grant Funded Research Projects

- **Step 1:** Email libgrants@yorku.ca at **least 2 weeks** before your **faculty's internal review deadline** to arrange a consultation
- **Step 2:** After the consultation, a formal letter will be prepared (*dependent on library capacity*) and signed by the Dean of Libraries.
- **The following in-kind services are available:**
 - Computational research methods & tools consultations
 - Data visualization consultations & workshops
 - Digital preservation consultations
 - Exhibits and discovery consultations & workshops
 - Media Creation Lab spaces, equipment, & workshops
 - Metadata consultations
 - Open access publishing consultations & workshops
 - Research data management consultations & workshops
 - Research visibility and impact consultations & workshops
- [Review Library Support for Grant Funded Research Projects webpage](#)

Scenario

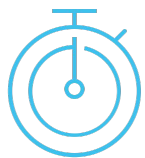
- You publish your manuscript



INSTEAD



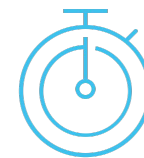
- Two years later ...
 - Someone reads your paper and asks for your data and code to replicate your work and build on your idea
 - You don't have access to the data



- You publish your research
- Publish and share the data in a generalist or disciplinary data repository with a README file and other information to understand the data



- Two years later ...
 - Someone reads your paper and downloads your data and code from the repository and uses the README file to understand it
 - They cite your data in their work



What is a Data Management Plan (DMP)?

- A DMP is a document that describes how research data is to be managed during a research project and after it has been completed
- Why a DMP?
 - Saves time and prevents data loss
 - Ensure correct attribution, rigor, reproducibility, and accountability
 - Increase the visibility and impact of your research
 - Meet institution, funder and grant guidelines
 - Adhere to journal data sharing requirements
 - Supports adherence to the FAIR (findable, accessible, interoperable, and reusable) Principles to maximize the research potential of the data



FAIR Principles for research data

Examples

OMB No. 0925-0001 and 0925-0002 (Rev. 07/2022 Approved Through TBD)
Sample DMS Plan – Survey/Interview Data Project

DATA MANAGEMENT AND SHARING PLAN

If any of the proposed research in the application involves the generation of scientific data, this application is subject to the NIH Policy for Data Management and Sharing and requires submission of a Data Management and Sharing Plan. If the proposed research in the application will generate large-scale genomic data, the Genomic Data Sharing Policy also applies and should be addressed in this Plan. Refer to the detailed instructions in the application guide for developing this plan as well as to additional guidance on [sharing.nih.gov](https://www.nih.gov/data-management). The Plan is recommended not to exceed two pages. Text in italics should be deleted. There is no "form page" for the Data Management and Sharing Plan. The DMS Plan may be provided in the format shown below.

Element 1: Data Type

A. Types and amount of scientific data expected to be generated in the project:

- Survey** This study will collect quantitative and qualitative survey data from 3,000 U.S. adults. The survey instrument will include 40-50 fixed-scale and open-ended text items including novel measures, scales identified in the PhenX toolkit, and two proprietary measures from [insert entity]. Raw and recoded data including open-ended text responses and resulting recoded variables will be produced. Open-ended text items will not request personally identifying or sensitive information but will be reviewed for privacy disclosure risks and redacted accordingly.
- Interviews** This study will also conduct digitally recorded semi-structured interviews with patients (n=40), healthcare providers (n=40) and industry professionals (n=20). Deidentified, raw transcripts will be generated and coded using methods described in Aim 2. Codebooks will be developed and used for coding transcripts, as described in Aim 2. Codebooks and coding summary files will be shared as detailed below.

B. Scientific data that will be preserved and shared, and the rationale for doing so:

Survey Data: Except where mentioned in Section 5 below, de-identified individual and aggregate survey data (including raw and recoded data) will be shared. The de-identification process will remove direct and indirect respondent identifiers. Once data are confirmed final, respondent identifiers will be deleted.

Interview Data: Following generation and quality check of raw transcripts from interviews, digital voice recordings will be permanently deleted to protect participant privacy. Respondent identifiers will not be shared. Raw transcripts will be maintained but not shared. Transcripts from interviews with patients and healthcare providers will be de-identified and sensitive content redacted where identification is plausible. These de-identified and redacted transcripts and coding summaries will be shared. Transcripts from industry stakeholder interviews will not be shared to protect participant privacy (see 5A). All shared data sets mentioned above, and metadata (see below) will be made publicly available through the Analysis, Visualization, and Informatics Lab-space (AnVIL).

C. Metadata, other relevant data, and associated documentation:

Documentation to be made publicly available to the research community will include PDF documents containing:

- Survey instruments with proprietary measures redacted (Note we will not be at liberty to share proprietary instruments used in the survey but will provide citations and contact information for proper licensure)
- Interview guides
- All data collection protocols including sample and subject selection methods
- Copies of blank, dated, stamped consent forms and IRB approvals, and resulting limitations of data usage
- Survey codebook including question number, question text, variable name, variable label, value labels, codes for missing, non-applicable, "don't know," and refusal values
- Methods used to code open-text survey responses
- Codebook for analyses of interviews, including a list and definition of all codes used, and coding examples
- Steps taken to remove direct and indirect identifiers in the data
- Description of software and analytical methods used in survey and interview data analyses
- R code used in survey data analyses.
- A standard citation and unique identifier to facilitate attribution of data use.

These will be shared in AnVIL. To the extent the context of data collection can be revealed without compromising privacy and identity of research participants, it will be included in study protocols.

Element 2: Related Tools, Software and/or Code

Novel tools and software will not be generated. Proprietary data analysis software such as [insert name] may be needed to analyze transcripts data and must be licensed independently by data users. Other data analyses will be conducted using

Data Collection

Describe the type(s) of data that you will collect, including all survey, interview and/or focus group data. If there are any additional types of data that will be collected or generated describe these as well.

When using acronyms within your DMP (Data Management Plan) it is always recommended to spell them out at least once within any given category so that readers understand what the acronym is referring to.

The data will consist of project data, survey data, semi-structured interviews and focus groups.

Project data includes identity lists of participants, working documents and other project work that will also be collected. Any other potential data not mentioned here that is collected throughout the project will be incorporated into the DMP as it becomes available. The research data collected will produce tabular data, video audio, and text based data (including transcriptions for all interviews and focus groups).

Are there any existing data that you can re-use and that will provide insight or answer any of your research questions? If so, please explain how you will obtain these data and integrate them into your research project.

We have searched a wide range of disciplinary-specific, institutional ([Borealis Dataverse](https://borealisdata.ca/dataverse/ualberta): <https://borealisdata.ca/dataverse/ualberta>), and national repositories ([Federated Research Data Repository \(FRDR\)](https://www.frd-rdfr.ca/repo/): <https://www.frd-rdfr.ca/repo/> and Statistics Canada public use microdata files) for any existing data that may be re-purposed or integrated into our project but were unable to find any relevant data. The project team also consulted with Anna Bombak, Digital Content Specialist within the University of Alberta Library, to confirm that there are currently no existing data that the team can reuse or repurpose.

Is it important to identify and understand as early as possible the methods which you will employ in collecting your data to ensure that they will support your needs, including supporting the secure collection of sensitive data if applicable. Describe the method(s) that you will use to collect your data.

Survey data will be collected using the REDCap software platform which is hosted by the Women and Children's Health Research Institute (WCHRI) located at the University of Alberta and is available to University of Alberta researchers. Survey participants will use a secure and unique web-based link to access the survey, and all information they provide will be entered and saved directly into the REDCap system.



**Data Management Plan Exemplar #2:
Digital Humanities and Secondary Data**

Historical Canadian Census Data

Data Collection

What types of data will you collect, create, link to, acquire and/or record?

Text-based data are collected from Statistics Canada, falling under the Statistics Canada Open License.

What file formats will your data be collected in? Will these formats allow for data re-use, sharing and long-term access to the data?

The data will be collected in an Inmagic v.15 database from which non-proprietary file-formats, including ASCII, HTML and/or XHTML text files, may be extracted. Inmagic is a content management system, a type of database used to manage large amounts of content, such as documents, images, and more.

What conventions and procedures will you use to structure, name and version- control your files to help you and others better understand how your data are organized?

As records are created and/or edited, they will be time-stamped with changes. The name of the person making the changes is also requested. A log file is kept to record information about database changes.

Documentation and Metadata

What documentation will be needed for the data to be read and interpreted correctly in the future?

Four different pieces of documentation will be needed.

- A description of the project, which will include a description of the process undertaken to identify the various historical census data files;
- A description of the field structure in Inmagic (e.g., whether a field is required, uses a controlled vocabulary, is repeatable, etc.);
- The data entry instructions to be followed in populating the database; and

[Survey/Interview Data](#) (NIH)

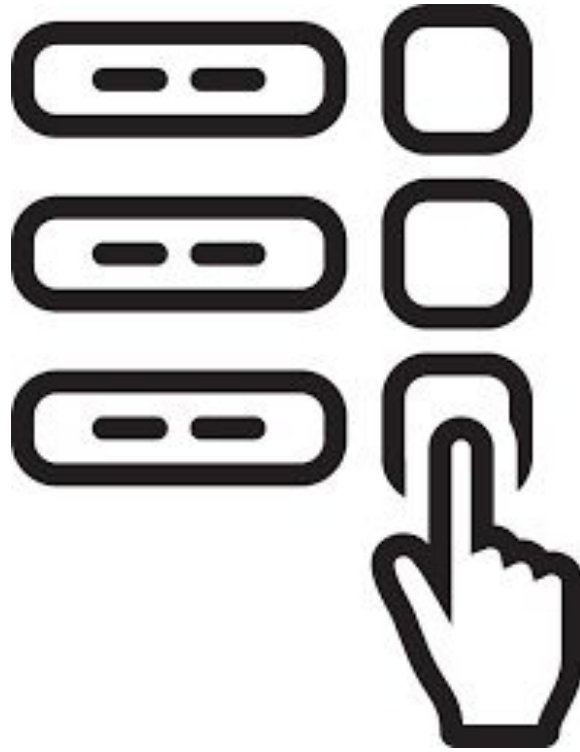
[Mixed Methods Research](#) (Digital Research Alliance of Canada)

[Digital Humanities and Secondary Data](#)
(Portage)

YORK U



Poll 2

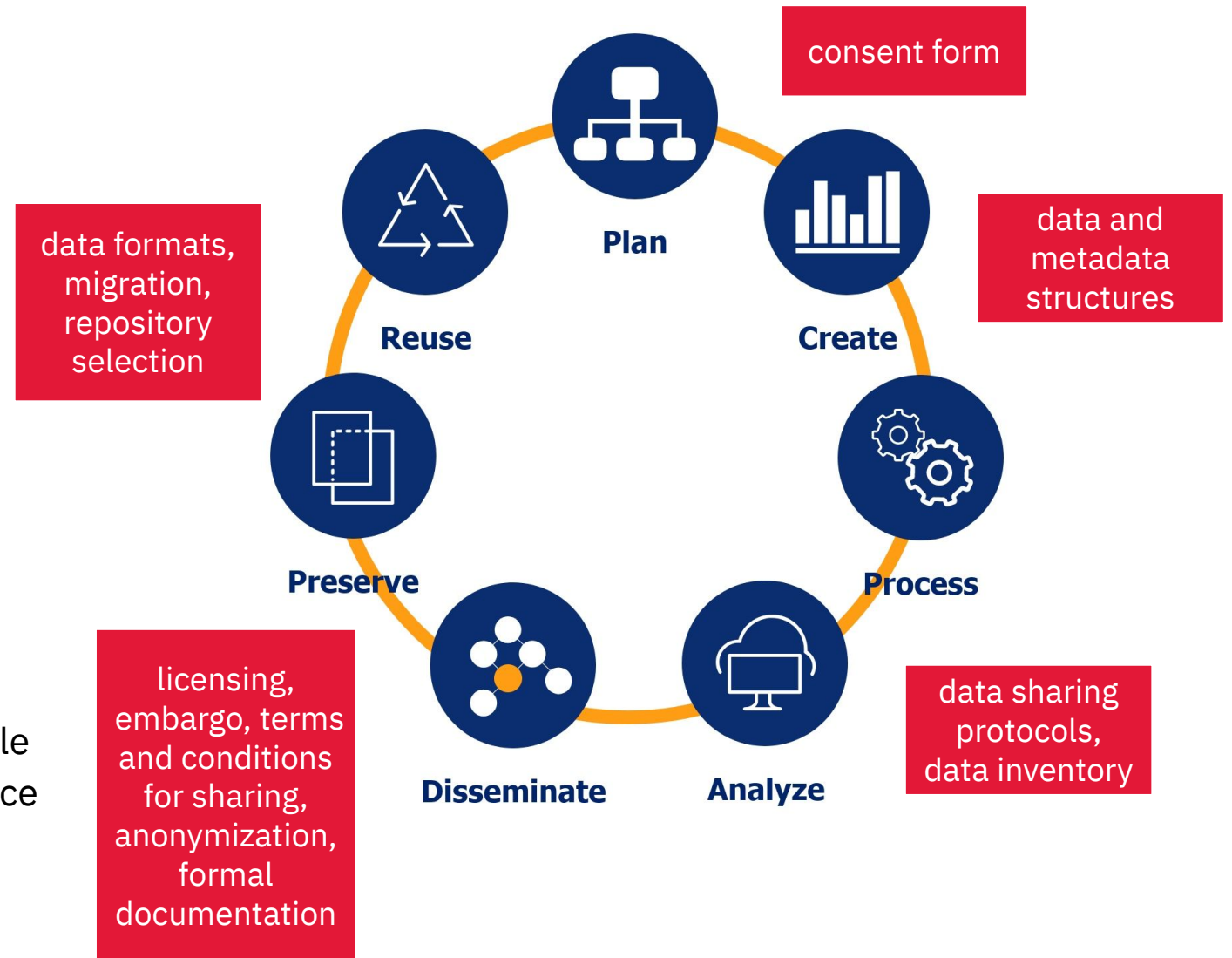


Types of Research Data

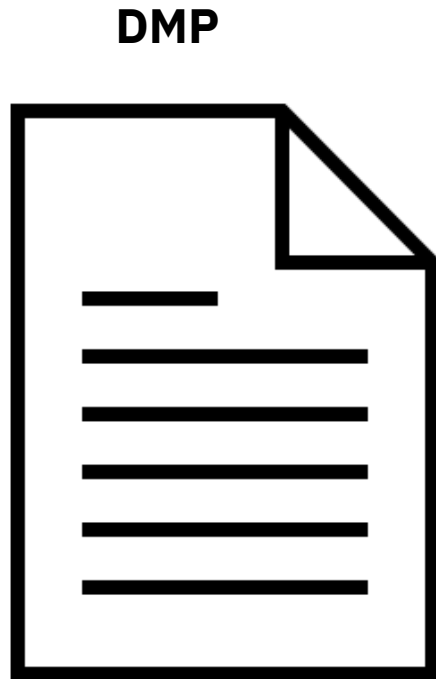
General	Social Sciences	Natural/Physical Sciences
<ul style="list-style-type: none">• Audio• Video• Images• Numerical measurements• Software and code• Secondary data	<ul style="list-style-type: none">• Survey• Interview transcripts• Economic indicators• Demographics• Opinion polling• Field notes• Drawings	<ul style="list-style-type: none">• Measurements• Experiments• Computer modeling• Simulations• Observations• Field studies• Specimen

DMP in the Research Data Lifecycle

- A dataset has a longer lifespan than the research project that creates it
- Data should be “as open as possible and as closed as necessary”
- Data can be used and reused if
 - Shared
 - Managed well
 - Properly preserved
 - Made available
- Data principles
 - FAIR Principles
 - Findable, Accessible, Interoperable, Reusable
 - CARE Principles for Indigenous Data Governance
 - Collective Benefit, Authority to Control, Responsibility, Ethics
 - Sensitive Data



Elements of a DMP



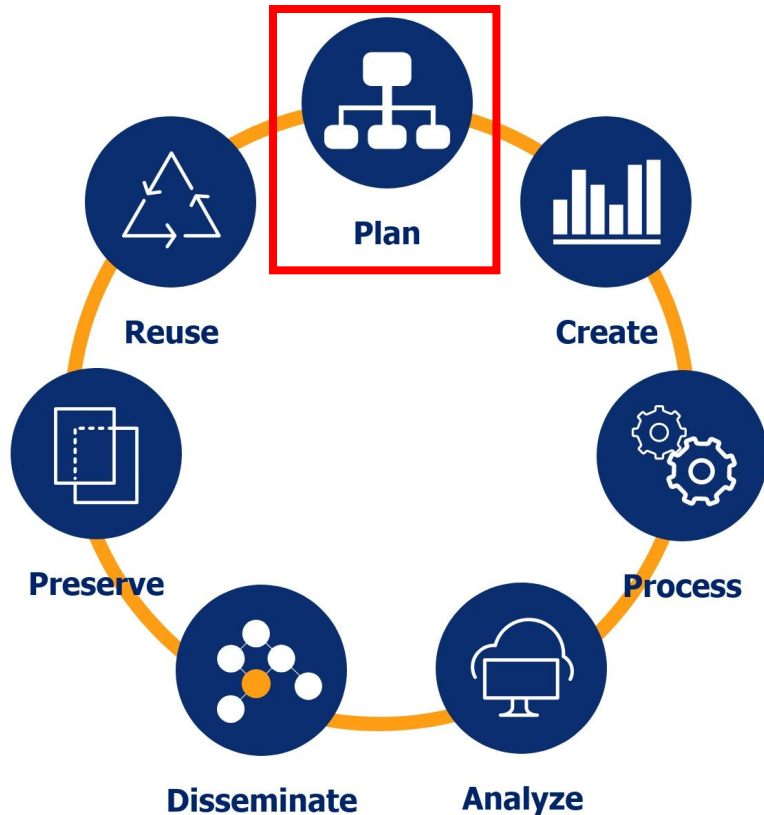
- ❑ Responsibilities & resourcing
- ❑ Project description
- ❑ Data collection
- ❑ Documentation & metadata
- ❑ Storage & backup
- ❑ Access, sharing & reuse
- ❑ Preservation
- ❑ Ethics & legal compliance

DMP Responsibilities and Resourcing

- Who?
 - PI or corresponding researcher
 - Collecting, creating, processing, and analyzing data
 - Transcribing and translating
 - Managing and administering research
 - Supporting data storage, security, back-up
 - Sharing the data
- Roles and responsibilities
 - Who is responsible for collecting and managing the data?
 - Who are the producers and owners of the data?
- Resourcing and cost
 - People (ex. training, time)
 - Data management (ex. software, data tools)
 - Sharing and archiving resourcing (ex. storage, preservation)

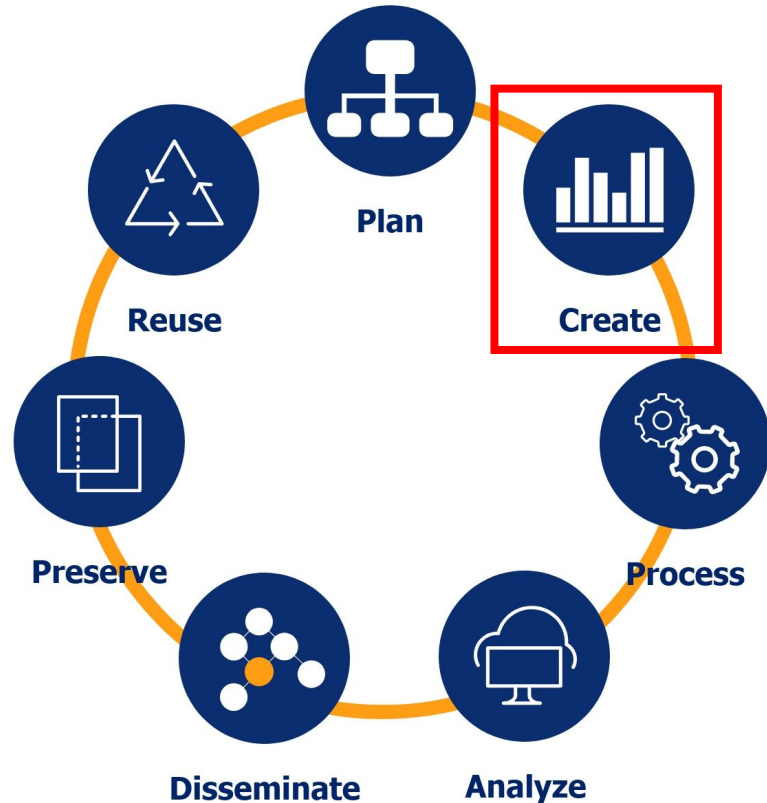


Project Description



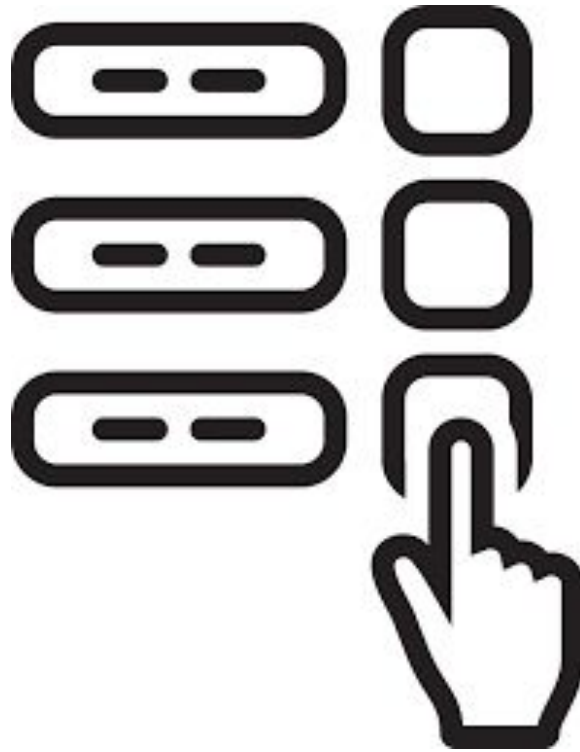
- Project and data description
 - What is it?
 - How is the data collected and produced?
 - How much data is generated?
 - Is there any personally identifiable information or confidential data?
- Secondary data
 - Is data produced by someone else used? If so, where is it from?
 - Do you have a data use agreement? Are there restrictions on accessing and sharing the data?
- Consider project-specific policy requirements including ethical and legal restrictions, and funder and publisher requirements
- Identify stakeholders (ex. funder, journal, data producer, data owners, and research participants)

Data Collection



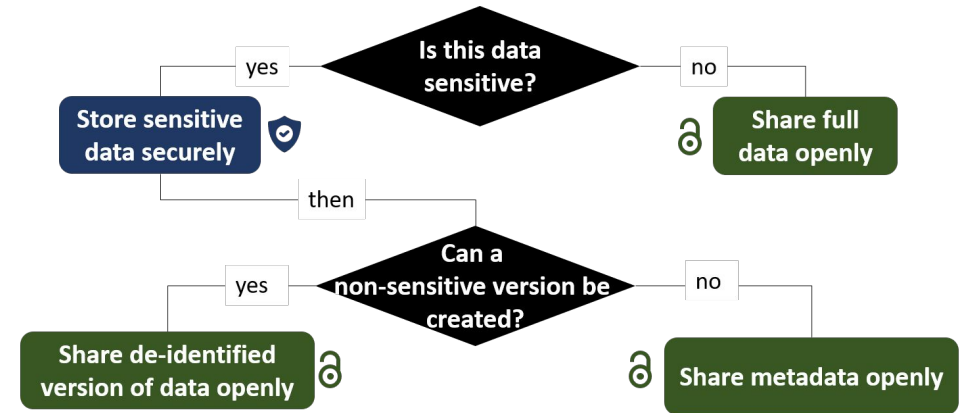
- Provide an overview of data type, format, and size
 - Ex. text documents, spreadsheets, artefacts, sample analyses, survey data, drawings, photographs, recorded interviews, sensor reading, patient records, results from theoretical calculations etc.
- Are there any privacy or sensitive considerations?
- Have an inventory of all your data files

Poll 3



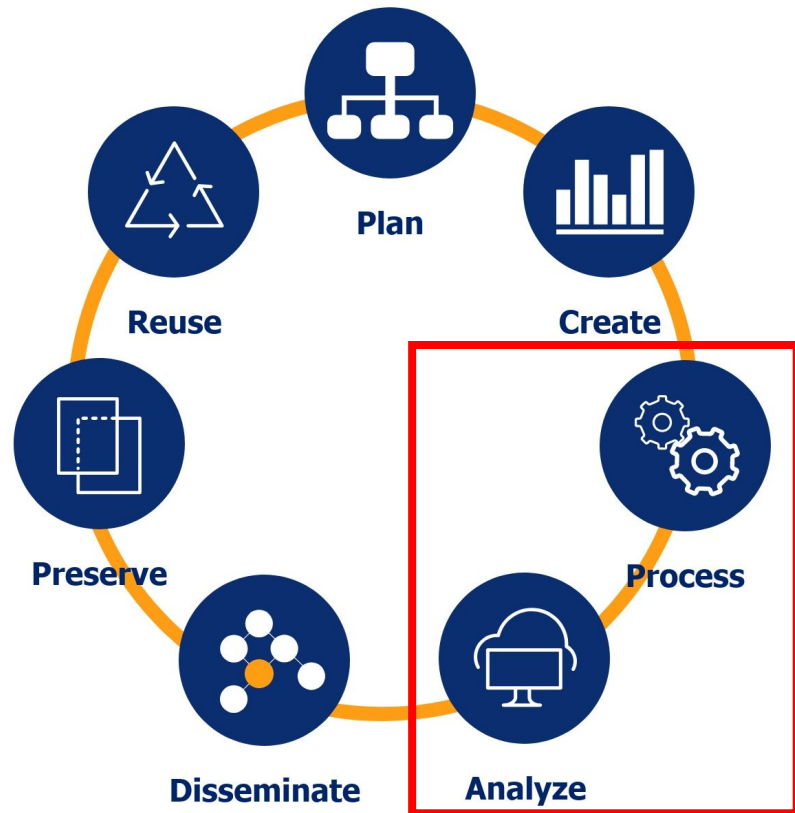
Managing Sensitive and Confidential Data

- Decision about sharing
 - Is sharing data required or prohibited?
 - Is sharing data supported by your research communities and/or publishers?
- Data anonymization
 - Direct identifiers (name, address, phone number, membership number, etc.)
 - Indirect identifiers (zip code, education level, medical diagnosis, race/ethnicity, occupation, etc.)
- Indigenous data
- Consent forms for data sharing
- Resources
 - [York's Office of Research Ethics](#)
 - [York's Data Security Guideline](#)



["What is sensitive data?". La Trobe University](#)

Documentation (aka metadata)



- Plan for versioning, naming, and structuring your variables, files, and folders
- What documentation do you need to make the data understandable to your team and other researchers?
 - What types of information are important for describing, discovering, and using the data?
 - File structure
 - File naming
 - Format
 - Software
 - Which mechanisms for metadata capture make sense given your research workflow?
 - Any restrictions on your data?
- Are you using documentation standards and controlled vocabulary in your field?
 - What directory and file naming conventions will you use?
- Be consistent

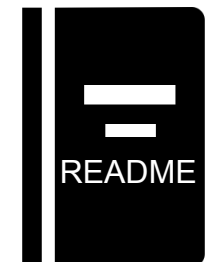
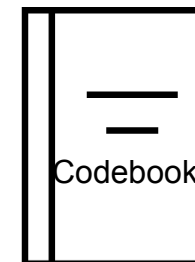
Accompaniments to the Data

- README files, codebooks, data dictionaries, data vocabularies, user guides should accompany the data to ensure others can understand and use your data
- Data require
 - Accurate documentation describing the steps taken to process and code data
 - Meaningful labels and descriptions
 - Avoid the use of acronyms or ambiguous language
 - Use consistent file and document structures

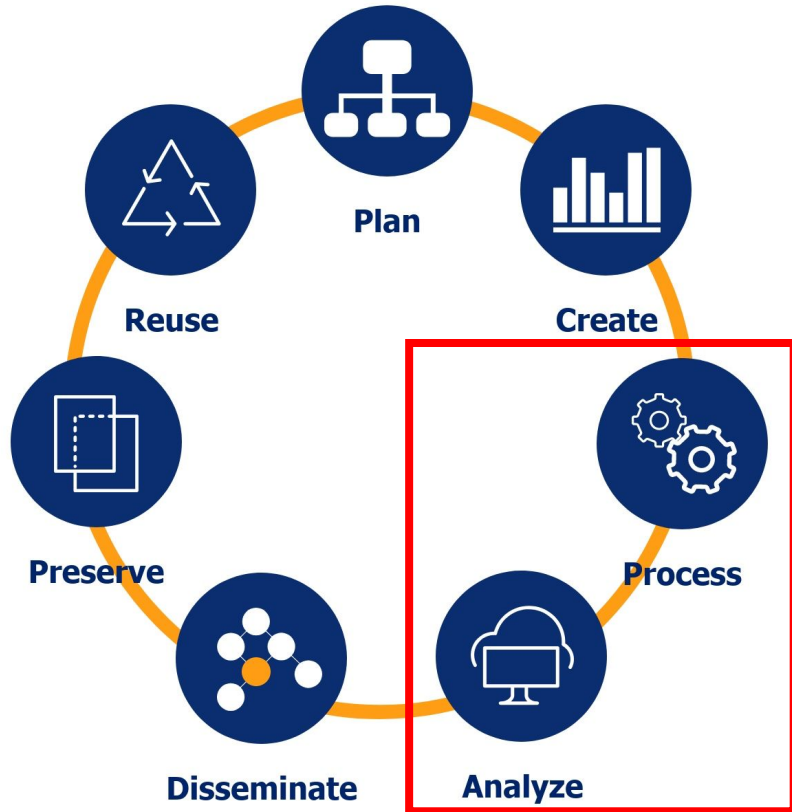
File naming convention

Place a single component of the name in each box (eg. date). Think about how the file names will sort when you determine the order.

ON001_Audio_20190530_MS_Original.wav
ON001_Image_20190624_MW_Cropped.png



Storage and Security

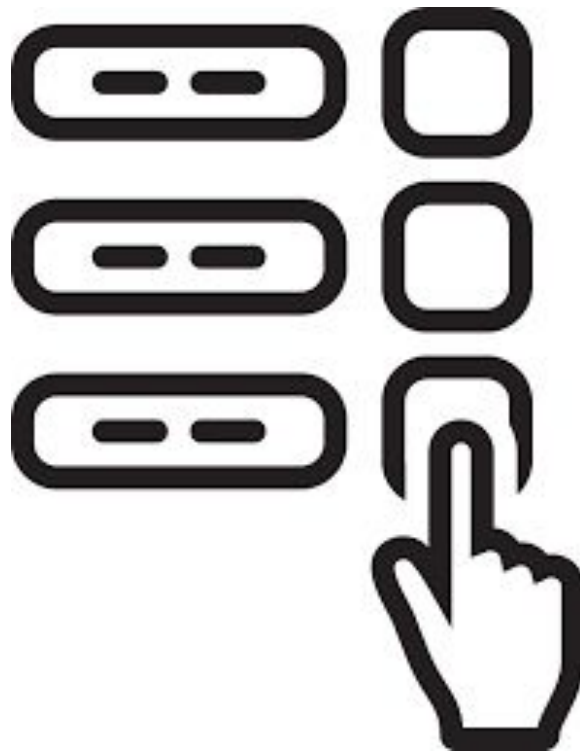


- Access
 - How will you access, modify, and contribute data throughout the project
 - What tools or software are required to read or view the data?
- Storage
 - Will this data require secure storage?
 - Short vs long term storage for active and/or completed research
 - Backup and data retention schedule and procedures
- Security
 - Cloud based vs physical backup
 - Any personally identifiable or confidential information?

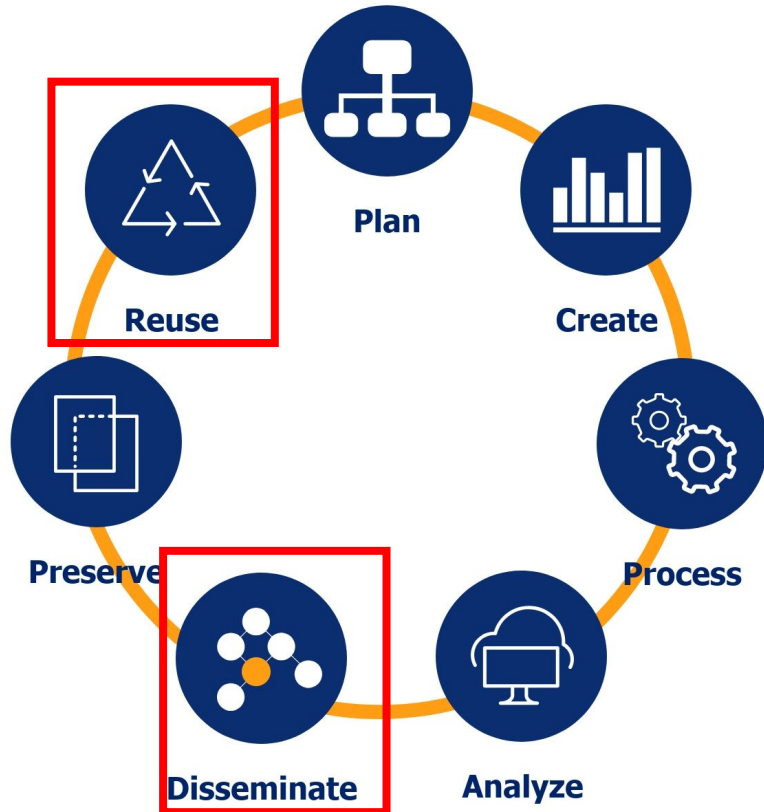
Securing Sensitive or Confidential Data

- Are you working with or sharing only de-identified or non-confidential data?
- Tips for securing sensitive or confidential data
 - Host sensitive data on infrastructure physically located at the University or on a secure Canadian network or server
 - Avoid working directly with sensitive or confidential data
 - Utilize secure data anonymization and encryption methods
 - Always ensure storage options chosen support the safe handling of data

Poll 4

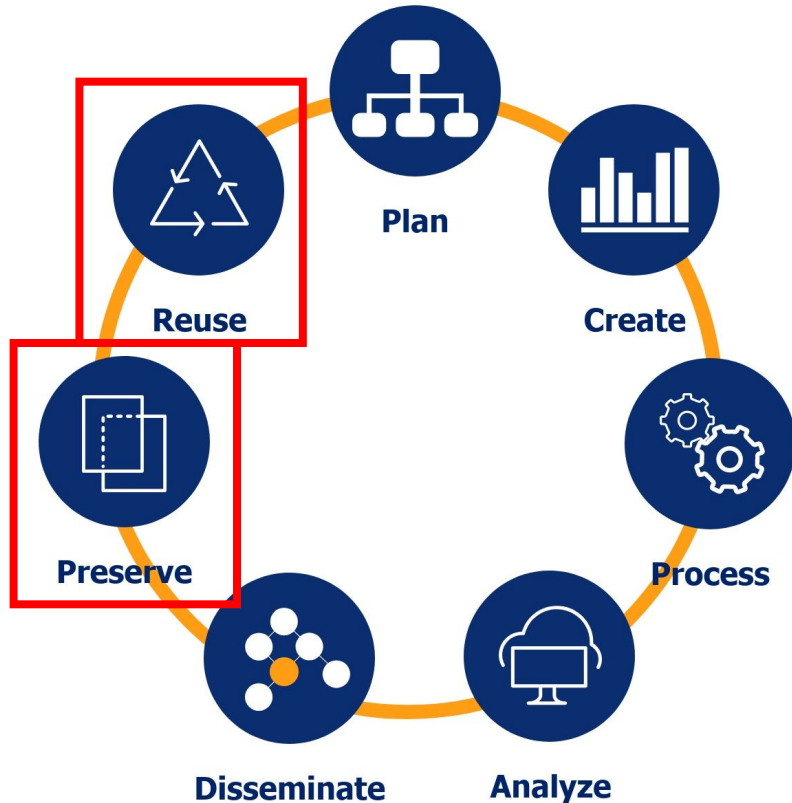


Sharing and Reuse



- What data can be shared?
 - When? How? Where? For how long?
 - Research data repositories
 - Will the research be published in a journal that requires research data to accompany articles?
 - Will there be any restrictions or embargoes on the data?
 - Security, sustainability, and discoverability of the data
- Data licensing
 - Who can view and use the data, for what purpose?
 - Creative Commons licensing
- Does sharing the data raise privacy, ethical, or confidentiality concerns?
 - Do you have a plan to protect or anonymize data?

Retention and Preservation



- What data will be kept / discarded?
 - How long will the data be preserved?
 - Select open data formats for longevity and interoperability, and common to your field
 - Check data repository requirements
- Who will be interested in re-using the data?
- Is there sufficient information to allow easy re-use of the data?
- Restrictions

Open Data Formats

- FAIR file formats for preservation
 - Containers: TAR, GZIP, ZIP
 - Databases: XML, CSV, JSON
 - Geospatial: SHP, DBF, GeoTIFF, NetCDF
 - Video: MPEG, AVI, MXF, MKV
 - Sounds: WAVE, AIFF, MP3, MXF, FLAC
 - Statistics: DTA, POR, SAS, SAV
 - Images: TIFF, JPEG 2000, PDF, PNG, GIF, BMP, SVG
 - Tabular data: CSV, TXT
 - Text: XML, PDF/A, HTML, JSON, TXT, RTF

DMP Ethical and Legal Compliance

- Consider ethical and legal requirements (of participant community, research field, grants, institution, journal publisher, etc.) and develop strategies for compliance
- Licensing, embargo, restrictions
- Consult with local IP and copyright offices and ethic review boards
 - [Copyright at York](#)
 - [Resources and Services for Researchers at York](#)
 - [Research Ethics at York](#)

Sensitive Data Toolkit for Researchers

Part 2: Human Participant Research
Data Risk Matrix

Part 3: Research Data Management
Language for Informed Consent

DMP Exemplars and Templates

- [York University template on the DMP Assistant](#)
- [DMP Assistant templates](#)
- [DMP templates from Digital Research Alliance of Canada](#)
 - [Arts-Based Research](#)
 - [History and the Humanities](#)
 - [Mixed Methods \(Surveys & Qualitative Research\)](#)
 - [Neuroimaging in the Neurosciences](#)
 - [Qualitative Health Sciences Research](#)
 - [Systematic Reviews](#)
- [DMPs shared by Canadian researchers](#)

York University

This plan is based on the "York University" template provided by York University.

Template version 4, published on March 13, 2023

Instructions

Write plan

Project Information

- 1.1 Name of project or sub-project and project number (if applicable).
- 1.2 Please provide an outline description of the research project and the intended products of the research.
- 1.3 Please provide details of the project timelines.
- 1.4 Principal Investigator(s)
- 1.5 Research data management policies

Data Collection

- 2.1 What types of data will you collect, create, link to, acquire and/or record?
- 2.2 What file formats will your data be collected in? Will these formats allow for data re-use, sharing and long-term access to the data?
- 2.3 How much data do you anticipate collecting? Include an estimate of how much storage space you will require (in megabytes, gigabytes, terabytes). This estimate should also take into account storage space required for file versioning, backups, and the growth rate over time.
- 2.4 Are there any existing data that you can re-use? If so, please explain how you will obtain that data and integrate it into your research project.
- 2.5 What conventions and procedures will you use to structure, name and version-control your files to help you and others better understand how your data are organized?

Documentation and Metadata

- 3.1 What documentation will be needed for the data to be read and interpreted correctly in the future? This includes study-level documentation, data-level description, and any other contextual information required to make the data usable by other researchers.
- 3.2 If you are using a metadata standard and/or tools to document and describe your data, please list here.
- 3.3 How will you make sure that documentation is created or captured consistently throughout your project?

Storage, Access, and Backup

- 4.1 What are the anticipated storage requirements for your project, in terms of storage space (in megabytes, gigabytes, terabytes, etc.) and the length of time you will be storing it? If applicable, where are hardcopy notebooks and physical samples going to be physically stored?
- 4.2 How and where will your data be stored and backed up during your research project?
- 4.3 How will the research team and other collaborators access, modify, and contribute data throughout the project? If applicable, how will you ensure that sensitive data is stored securely and only accessible to the research team and other collaborators throughout the project?

Sharing and Reuse

- 5.1 What data will you be sharing and in what form? (e.g. raw, processed, analyzed, final).
- 5.2 Where will you deposit your data and provide access at the end of your research project? Is there specific software needed in order to use or interpret your research data collection?

DMP Assistant Tool for Creating a DMP



[Go to the DMP Assistant](#)

- A free, bilingual, open-source Canadian tool for preparing data management plans (DMPs). The tool follows best practices in data stewardship and walks researchers step-by-step through key questions about data management.
- York University guidance
- Disciplinary templates

Accessing York's DMP Template (Part 1)

The screenshot shows the 'Create account' form with the following fields and options:

- Sign in** / **Create account** (highlighted)
- * First Name** (text input)
- * Last Name** (text input)
- * Email** (text input)
- * Organization** (text input with placeholder: "Begin typing to see a list of suggestions.")
- * Password** (text input)
- Show password
- I accept the [terms and conditions](#)
- Security check**
 - I'm not a robot (reCAPTCHA)
- Create account** (button)

The screenshot shows the 'My Dashboard' page with the following content:

- My Dashboard**
- Welcome**
You are now ready to create your first DMP. Click the 'Create plan' button below to begin.
- There are no records associated
- Create plan** (button)

Step 1: Go to the DMP Assistant Website. On the right side of the page, click **Create Account**. Fill out the relevant information. You can use a non-YorkU email

Step 2: After creating your account, you'll be redirected to a new page. Click the **Create plan button**.

The screenshot shows the 'Create a new plan' form with the following sections:

- Create a new plan**
Before you get started, we need some information about your research project to set you up with the best DMP template for your needs.
- * What research project are you planning?**
 - My Test Data Management Plan (text input)
 - mock project for testing, practice, or educational purposes (checkbox)
- * Indicate the primary research organization**
 - Organization**
 - York University (dropdown menu)
 - or - No research organization associated with this plan or my research organization is not listed (checkbox)
- Which DMP template would you like to use?**
 - York University (dropdown menu)
 - We found multiple DMP templates corresponding to your primary research organization
 - Create plan** / **Cancel** (buttons)

Step 3: On the next page, enter a title for your research project. You can also click the box beside, 'Mock project for testing, practice or educational purpose' for test DMPs. Use the primary research organization dropdown menu to select **York University**. Use the DMP template dropdown to select **York University**. Click **Create Plan** to proceed.

Accessing York's DMP Template (Part 2)

My Test Data Management Plan

Project Details Contributors Plan overview **Write Plan** Research Outputs Share Download

* Project title
My Test Data Management Plan

mock project for testing, practice, or educational purposes

Project abstract

B I [list icon] [link icon] [grid icon]

Research domain
- Please select one -

Project Start
yyyy - mm - dd

Project End
yyyy - mm - dd

Select Guidance

To help you write your plan, DMP Assistant can show you guidance from a variety of organizations.

Select Up To 6 Organizations To See Their Guidance.

Find guidance from additional organizations below

See the full list

Save

Step 4: You will now see York's DMP template. Use the **menu at the top of the page** to access and fill out each section of the DMP template (i.e., Project Details, Contributors, Write Plan, etc.).

My Test Data Management Plan

Project Details Contributors Plan overview **Write Plan** Research Outputs Share Download

expand all | collapse all 0/30

Project Information (0 / 5) +

Data Collection (0 / 5) +

Documentation and Metadata (0 / 3) -

Because data are rarely self-explanatory, all research data should be accompanied by metadata (information that describes the data according to community best practices). Metadata standards vary across disciplines, but generally state who created the data and when, how the data were created, their quality, accuracy, and precision, as well as other features necessary to facilitate data discovery, understanding and reuse. Any restrictions on use of the data must be explained in the metadata, along with information on how to obtain approved access to the data, where possible.

3.1 What documentation will be needed for the data to be read and interpreted correctly in the future? This includes study-level documentation, data-level description, and any other contextual information required to make the data usable by other researchers.

B I [list icon] [link icon] [grid icon]

Save

Guidance

YorkU

Your documentation may include study-level information about:

- who created/collected the data
- when it was created
- any relevant study documents
- conditions of use
- contextual details about data collection methods and procedural documentation about how data files are stored, structured, and modified.

A complete description of the data files may include:

- naming and labelling conventions
- explanations of codes and variables
- any information or files require to reproduce derived data.

More information about data documentation is available at the UK Data Archive

Step 5: When you get to the **Write Plan section** of the template, you will notice that on the **right side of the page**, York specific advice has been embedded to help guide you in writing your DMP.

Accessing York's DMP Template (Part 3)

My Test Data Management Plan

Project Details Contributors Plan overview Write Plan Research Outputs Share **Download**

Format

pdf

Download settings

Optional Plan Components

- project details coversheet
- question text and section headings
- unanswered questions

PDF formatting

Font

Face: "Times New Roman", Times, Serif Size (pt): 10

Margin (mm)

Top	Bottom	Left	Right
25	20	12	12

Download Plan

Step 6: When you complete all sections, you can go to the last section, **Download**, and click **Download plan**. Use the file type dropdown menu to select the file type your prefer (i.e., PDF, docx, CSV, etc.).

Questions?

Research Data Management Services

- Contact:
 - Email: yul_rdm@yorku.ca
 - [RDM Services Website](#)

Library Support For Grant Funded Research Projects

- Contact
 - Email: libgrants@yorku.ca
 - [Library Support for Grant Funded Projects Website](#)