Creating Research Data Management Plans Using DMPTool

Dr. Xuan Zhou
Research Data Services
University Libraries | Texas State University
GOALS FOR TODAY

1. Understand the basic principles of research data management;
2. Be aware of data management planning tools, support and guidance which are available to academic researchers;
3. Be able to use DMPTool to develop a data management plan, and maintain it through the course of your research.
OVERVIEW

1. Why is data management important?
2. What is a data management plan?
3. Use the DMPTool
OVERVIEW

1. Why is data management important?
2. What is a data management plan (DMP)?
3. The DMPTool
WHAT IS RESEARCH DATA?

- Research Data is recorded, factual material commonly accepted in the scientific community as necessary to validate research findings (Awasthi & Tripathi, 2019)

- Lab notebooks
- Surveys
- Questionnaires
- Interview notes
- Code books
- Models and algorithms
- Content analyses
- Audio and video files
You have achieved global recognition with your groundbreaking study on genetic mutations and neurological disorders, resulting in two widely-cited papers. However, your life in the scientific community was disrupted three years later when another researcher accused you of data falsification, alleging statistical improbabilities and missing critical details in your research.
WHAT IS RESEARCH DATA MANAGEMENT

- Research data management (RDM) is about handling research data effectively and appropriately throughout the life of a research project and beyond.
- RDM refers to all aspects of creating, storing, sharing and archiving data and is an essential aspect of conducting responsible research.
WHY IS DATA MANAGEMENT IMPORTANT?

- Data management is a set of practices across the research lifecycle

<table>
<thead>
<tr>
<th>Mandate</th>
<th>Facilitate</th>
<th>Reuse</th>
<th>Impact</th>
</tr>
</thead>
<tbody>
<tr>
<td>Meet requirements and expectations set by funding agencies, publishers and domain associations</td>
<td>Ensure that your data is complete, documented, and accessible to you and to future researchers</td>
<td>Encourage the discovery and reuse of your data to further discoveries in your field of research</td>
<td>Receive credit for your data and increase its impact and visibility</td>
</tr>
</tbody>
</table>
RESEARCH DATA LIFE CYCLE
OVERVIEW

1. Why is data management important?

2. What is a data management plan?

3. The DMPTool
DATA MANAGEMENT & SHARING MANDATES

- Journals – PLOS, Nature, JDAP partners
- Funders – NSF, NIH...
- The White House OSTP memo (2003) – Federal agencies with over $100 million/year in R&D must develop a plan to support public access to research
START WITH A DMP

- Funder requirements
- Institutional policy
- Mitigate error and loss
- Avoiding unforeseen costs
- Be able to return the data
- Getting a handle on the complexity of data
COMPONENTS OF A BASIC DMP

1. Describing the Research Data
2. Data Standards and Metadata
3. Ethic Issue and legal Compliance
4. Intellectual Property and Re-use
5. Data Storage and Backup
DESCRIBING THE RESEARCH DATA

- Give a brief description of the data. Outline and justify your choice of format and consider the implications of data format and data volumes in terms of storage, backup and access.

  **What data will you collect or create?**
  - What type, format and volume of data?
  - Do your chosen formats and software enable sharing and long-term access to the data?
  - Are there any existing data that you can reuse?

  **How will the data be collected or created?**
  - What standards or methodologies will you use?
  - How will you structure and name your folders and files?
  - How will you handle versioning?
  - What quality assurance processes will you adopt?
DATA STANDARDS AND METADATA

- Describe the types of documentation that will accompany the data to help other users to understand and reuse it.
- Metadata is the information that describes and documents research data. Metadata will make your datasets searchable in an archive or repository, easily located from a citation, and easily understood by people who might want to use your data.

What documentation and metadata will accompany the data?
- What information is needed for the data to be to be read and interpreted in the future?
- How will you capture / create this documentation and metadata?
- What metadata standards will you use and why?
Ethical issues affect how you store data, who can see/use it and how long it is kept. Managing ethical concerns may include: anonymization of data; referral to departmental or institutional ethics committees; and formal consent agreements.

How will you manage any ethical issues?
- Have you gained consent for data preservation and sharing?
- How will you protect the identity of participants if required? e.g. via anonymization
- How will sensitive data be handled to ensure it is stored and transferred securely?
INTELLECTUAL PROPERTY AND RE-USE

- State who will own the copyright and IPR of any data that you will collect or create, along with the licence(s) for its use and reuse.
- Consider any relevant funder, institutional, departmental or group policies on copyright or IPR.
- Consider permissions to reuse third-party data and any restrictions needed on data sharing.
DATA STORAGE, ACCESS AND BACKUP

- Consider where data will be stored and backup during the research and how you will control access to the data.
- Consider how will you manage access and security
DATA PUBLISHING, SHARING AND PRESERVING

- Consider how datasets that have long-term value will be preserved and curated beyond the lifetime of the project.
- Consider who will be responsible for data management.
- Consider how people might acknowledge the reuse of your data.
- Outline the plans for preparing and documenting data for sharing and archiving.
Data will be published in Texas State University Dataverse Repository under a Creative Commons CC0 public domain dedication, so that others may freely access, use, and build upon the work.
OVERVIEW

1. Why is data management important?
2. What is a data management plan?
3. The DMPTool
The DMPTool is an online platform guiding DMP development according to the requirements of specific funding agencies.

Texas State University researchers log in with their NetID and passwords.
DMP TOOL FOR DATA MANAGEMENT PLANS

- Helps researchers meet requirements of NSF, NIH and other U.S. funding agencies.
- Guides researchers through the process of creating a data management plan.
- Is available to everyone.
- Provides additional help for researchers at DMPTool partner institutions.
Sign In

To get started, go to DMPTool.org
Sign In

• Use your Texas State email address
• Click Continue
• Use your Texas State NetID and password to login
Create a Plan
Start developing DMP

1. Enter the project title
2. Enter the funder information
3. Click Create Plan

Lock here if you would like to create a test DMP
1. Enter the project title
2. Enter the project abstract
3. Enter the project and funding information
4. Click the + icon to add a related work if there is one
5. Click Save
Collaborators

Add the contributors of the project as needed

invite your collaborators via email
Write Plan

Expandable sections for each topic covered in a DMP.

Click the + icon to see detailed instructions for each section and enter text.
Click here to list your research output.

Fill in (anticipated) research outputs information accordingly if needed.
You can download your plan for view once you complete all the sections.
You can register your plan for a unique DMP ID and link it to your ORCID

When you have completed all sections of your plan, you can publish your DMP when you are ready and decide the visibility of the plan.
Finalize / Publish

the unique DMP ID and reference for citing
Given a DMP ID, you can access the ID to view a DMP landing page that includes details about the plan.

[View an example of a DMP ID landing page]
Templates For TXST Researchers

- Template Text: Data Sharing and Access
- Template Text: Metadata for Data Management
OPENxt: TXST Dataverse Data Repository: Preserve, Publish, and Share Your Research Data

Upcoming Sessions

Tuesday, November 14, 2023, 2:00PM - 3:00PM { 1 / 25 registered } Online

Description:
The Texas State University Dataverse Repository is hosted on the Dataverse platform, developed and used by Harvard University. It offers researchers a trusted repository to deposit, share, manage, and publish their research datasets. Researchers can also find and cite data across all research fields. In the workshop, you will be able to recognize appropriate data sharing practices in order to minimize data loss and maximize efficiency. You will also be able to use the TXST Dataverse Repository in order to manage, preserve, publish, and share data in an open access repository.

Tags:
data-management, dataverse, faculty, faculty-commons, graduate, library, open-access, open-txt, principle-investigators, repository, research, staff, workshops

Duration:
60 minutes

Department:
University Library

Additional Info:
https://www.library.txst.edu/services/research-services/research-data-management.html
THANK YOU!

Xuan Zhou, PHD
Research Data Services
University Libraries
UL-RDS@txstate.edu
References
