DATA MANAGEMENT PLANNING FOR SOCIAL SCIENCE RESEARCH

PATRICIA B. CONDON, PHD
Research Data Services Librarian
patricia.condon@unh.edu

SOCIAL SCIENCE LIBRARIAN BOOT CAMP
Tisch Library, Tufts University
June 1, 2018
OUTLINE AND LEARNING OBJECTIVES

Outline

- Presentation
  - Exploring topics to consider when creating DMPs for social science research
- Interactive Breakout Session
  - Case study and DMP discussions
- Wrap-up and questions

Learning Objectives

- Describe
  - the components of a data management plan
- Articulate
  - key data management/sharing challenges in research involving human subjects
- Reflect
  - on approaches to addressing those challenges
GOOD RESEARCH PRACTICE

Research Data Management

RDM is caring for, facilitating access to, and adding value to research data throughout its lifecycle and involves activities that contribute to effective storage, preservation, and reuse of data and documentation.

Image Source: The University of Sheffield
http://library.nuigalway.ie/digitalscholarship/researchdata/researchdatamanagement/
GOOD RESEARCH PRACTICE

Challenges Managing & Sharing Data

- Time & resources
- Knowledge and skills
- Large amounts of noisy data
- Researcher concerns about sharing data
- Changing notions of data as scholarly output
- Finding secure storage and long-term access repositories
- Lack of planning around data management and data sharing

Image Source: The University of Sheffield
http://library.nuigalway.ie/digitalscholarship/researchdata/researchdatamanagement/
Good Data Sharing Techniques

- Plan for sharing/reuse from the start
- Practice good data management throughout the research lifecycle
- Provide enough metadata and documentation
- Evaluate the sensitivity of data
- Submit data to a suitable repository
- Give appropriate recognition of reused data
DMP helps researchers consider, during the research design and planning stages, how the data will be managed during the research process itself and potentially shared afterwards with the wider research community.

DMPs are often required by funders and take the form of a brief description of how the researcher will comply with funder’s data sharing policy.

Image source: https://data.uni-bielefeld.de/en/data-management-plan
Five General Areas of a DMP

- Data types & sources
- Documentation & metadata
- Policies for re-use
- Policies for sharing
- Long-term preservation & access

And more…

Image source: https://data.uni-bielefeld.de/en/data-management-plan
<table>
<thead>
<tr>
<th><strong>DATA MANAGEMENT PLAN</strong></th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Data types &amp; sources</strong></td>
</tr>
<tr>
<td>What types of data will be created, captured, or generated (experimental measures, observational or qualitative, model simulation, existing)? How will the data be captured, created, and/or processed (Identify instruments, software, imaging, etc. used)? What quality assurance &amp; quality control measures will you employ?</td>
</tr>
<tr>
<td><strong>Documentation &amp; metadata</strong></td>
</tr>
<tr>
<td>What contextual details or documentation are needed to make data meaningful to others? What standards will be used for data and metadata/documentation? What file formats and naming conventions will be used?</td>
</tr>
<tr>
<td><strong>Policies for re-use</strong></td>
</tr>
<tr>
<td>How will you make the data available? When will you make the data available? What steps will be taken to protect privacy, confidentiality, security, intellectual property, or other rights or requirements (IRB, anonymization of data)? What restrictions need to be placed on re-use of the data?</td>
</tr>
<tr>
<td><strong>Policies for sharing</strong></td>
</tr>
<tr>
<td>What is the process for gaining access to the data? Will any permission restrictions need to be placed on the data? Who is likely to be interested in the data? What and who are the intended or foreseeable uses the data?</td>
</tr>
<tr>
<td><strong>Long-term preservation &amp; access</strong></td>
</tr>
<tr>
<td>What is the researchers’ long-term plan for preservation and maintenance of the data? Which archive/repository/database have you identified as a place to deposit data? How long will/should data be kept beyond the life of the project? What metadata/documentation will be submitted alongside the data or created on deposit/ transformation in order to make the data reusable?</td>
</tr>
</tbody>
</table>
Search for research funders' open access and data archiving policies
http://v2.sherpa.ac.uk/juliet/

Browse data sharing requirements by federal agency
http://datasharing.sparcopen.org/data

Always check with the funding agency for up-to-date requirements.
FUNDER’S REQUIREMENTS FOR DATA MANAGEMENT

DMPTool.org
US Funding Institutions

DMPonline.dcc.ac.uk
European Funding Institutions
PROTECTING DATA

Digital preservation

Data security

Privacy, confidentiality & disclosure

Image Sources:
- Digital Curation at the University of Maine: http://digitalcuration.umaine.edu/media/var_confusing_strategies parti_vga.png
- New York City College of Technology: https://www.gc.cuny.edu/getmedia/615af9fd-454f-4ac0-baee-92435cb46e18/HRPP-Texts1?width=470&height=260
SHARING DATA RESPONSIBLY

- Regulations/Protected Data
  - EU General Data Protection Regulation (GDPR)
    https://www.eugdpr.org/
  - HIPAA and Protected Health Information (PHI)
- Informed Consent
  - Permission to share
  - Whether data will be shared; if so, in what form (e.g., deidentified) and with whom (e.g., other researchers, placed in repository)
- Please Share Responsibly
  - Just because you can, doesn’t mean you should

Image Source: https://www.cuahsi.org/Files/Pages/images/27195/data_publication.jpg
FINDING A DATA REPOSITORY

- Social Science Repositories
  - ICPSR - https://www.icpsr.umich.edu/icpsrweb/
  - QDR Qualitative Data Repository - https://qdr.syr.edu/

- Open Access repositories
  - Figshare - https://figshare.com/
  - Havard Dataverse - https://dataverse.harvard.edu/
  - Zenodo - https://zenodo.org/
  - GitHub for code - https://github.com/

- Resources for locating repositories
  - Registry of Research Data Repositories - https://www.re3data.org/
RESEARCH DATA MANAGEMENT TEACHING CASE

Activity

1. Split into pairs or teams of three
2. In your team, read and discuss the short case study and data management plan (DMP) associated with the case.
3. Make edits/comments on the data management plan with the objective to more thoroughly address the key recommended components and improve the content. Questions to consider for each section of the DMP:
   - What are your key data management concerns for the researcher’s current practice?
   - What additional details could be added to help the researcher more effectively address the prompts?
   - Are there any issues related to managing & sharing data that the researcher needs to consider when preparing his IRB materials?
4. Teams report-out to the larger group for general discussion.

Purpose

- Small groups are asked to review and provide edits/comments to improve a Data Management Plan.

Skills & Knowledge

- Describe the components of a data management plan.
- Articulate key data management/sharing challenges in research involving human subjects.
- Reflect on approaches to addressing those challenges.

Outcome

- Team report-outs are intended to generate a larger group discussion about questions to consider when creating DMPs for social science research.
THANK YOU!

PATRICIA B. CONDON, PHD
Research Data Services Librarian
patricia.condon@unh.edu

SOCIAL SCIENCE LIBRARIAN BOOT CAMP
Tisch Library, Tufts University
June 1, 2018