Data Management Strategies Graduate Self-Assessment

This document is intended to help you identify areas of need in your data management strategies. This document does not proscribe or suggest courses of action. (Indeed, some areas may not be applicable to your research at all.) Instead, this is a tool to explore your current data management practices and identify areas of strength and weakness. For the purpose of this document the term “research collaborators” includes both co-investigators and graduate student researchers. For further information on any of these topics or for a data strategies consultation, contact msn@purdue.edu.

DATABASES AND DATA FORMATS
1. Are you building databases to house your data?
2. Are those databases documented in a written format that is complete and accessible by your research team members?
3. Is the database structured in such a way that the data content is clearly labelled and human readable?
4. Do you have a plan to maintain forward compatibility, access to data, and avoid data loss with updates to software?

DATA CREATION
5. Do you have a specified workflow such as a plan for the steps of analysis for your dataset?
6. Do you have a specified database, analysis and visualization software, coding protocols, and other tools that should be used on your dataset?
7. Have you identified the existing level of skill among research collaborators in using these tools?

DATA MANAGEMENT AND ORGANIZATION
8. Are there standard operating procedures (SOP) to be used by all individuals as part of the research team?
9. Are all individuals made aware of the SOP verbally and in writing?
10. Do you have a standardized file structure for your research project that makes clear where raw, analyzed, and final data files belong for each aspect of your research project?
11. Do all project members have access to the storage tool where data files are stored?
12. Do you have a standardized naming convention that enables easy retrieval of specific files?
13. Are all research members required to deposit their latest data files to this location?
DATA QUALITY AND DOCUMENTATION
14. Is the data set examined for inconsistencies and errors?
15. Is the data set documented in an ongoing manner such that final
documentation can easily be compiled from research records?
16. Have files been tagged so that errors or changes are documented on
files?
17. Do you have a version control system in place for your research data?
18. Have there been a conversation among the research team to identify
when new versions should be created and how those decisions should be
documented?

METADATA AND DATA DESCRIPTION
19. Is the amount of documentation and description that your research
team members provide sufficient for you to be able to understand and
make use of data?
20. Have you developed a tagging schema for your files so that you team or
others can use and find the documents?

DATA CURATION AND RE-USE
21. Are you creating your data with the intention of making part of all of your
data publicly available?
22. Have you identified which portions of the data have value to the
discipline as a whole?
23. Have you learned how to make data citable?
24. Have you deposited data into a data repository?

DATA PRESERVATION
25. Have you built data preservation into your project lifecycle?
26. Have you developed a preservation policy for your data that includes
how long the data will be kept, where it will be housed, how it will be
accessible, and how forward compatibility will be managed?
27. Have you talked with a preservation professional about what is entailed
by preserving a data set?