

DataCommons@PSU: Data Plan Questionnaire

datacommons@psu: <http://www.datacommons.psu.edu>

The following questions were created provided by Penn State's "DataCommons@PSU". They are intended to assist in the development of data plans for management of final (archival) databases.

1. **Project Timeline**—What is the proposed length of the project and expected completion date?
2. **Data Type(s)**—What data types will be produced by this project? For example, tabular, spatial, large scale databases, images, videos, etc.
3. **Software Platform**—What software platform(s) will you and your project partners be using to enter, create, and analyze the data?
4. **Data Format(s)**—What will be the resulting data format(s)? For example, xls, shp, pdf, tiff, jpg, netcdf, etc.
5. **Data Standards**—Do you have any scientific or research community standards by which you must create your data or metadata? Have you determined standard platforms and formats for all project members who will be creating data? Who will be creating the metadata? Do you need support for this?
6. **Data Users**—Who will be the primary users of this data during the project life cycle and where will they be located—at PSU or will they be remote users? Will these users require login access?
7. **Data Input**—Who will be responsible for the data? For example, this person would be tracking the latest version of the database, data set, etc. Who will be inputting data? Will they be located at PSU or be entering data remotely?
8. **Expected Storage Needs**—How much data do you expect to generate during the life cycle of this project? What do you expect to be your yearly storage needs?
9. **Additional Data & Models**—What additional data might you need for this project? Are you using any models or analysis techniques that could be documented? Will users be accessing these results or running analyses remotely.
10. **Expected Access Needs**—What access needs will be required at the completion of this project?

DataCommons@PSU: Data Questionnaire

datacommons@psu: <http://www.datacommons.psu.edu>

This questionnaire is designed to assist in the creation of metadata for your data set. It should capture the basics of “Who, What, When, Where, How” for the data.

1. **Primary Contact**--Please provide contact information for the primary data contact if possible. This would be the person who knows the most about the data.
2. **Title of Data Set**--Please list your preferred title for your data set.
3. **Date of final version of data set**--Please list the date of the final version of the data set. For example, if the last update to the data was in 2012, you would list 2012. For a data set that spans multiple years, you can list this as a range of dates, for example 2002-2012. If this is data that is updated, you can list the month and year of the latest update.
4. **Abstract of purpose and method**--Please describe the data as well as the purpose of the data. You can use the description or abstract from your project or research proposal for this.
5. **Researcher(s)**--Please list the researchers on the project. You can include researcher role, Principal Investigator, Co Principal Investigator, Senior Personnel if you would like. Please also include your preferred top level affiliation. For example, if you are with a center such as the Center for Informatics, or an institute, college or department, you can list that as your top level affiliation. You may also include your collaborators at other institutions.
6. **Data Access**--If your data will be downloadable via FTP, just include the words FTP in this section. If you have other needs such as data that can be geoenabled (visualized in a GIS), please let us know.
7. **Additional Information**—Is there any additional information about your data? A website you want us to link to?
8. **References**—Include the titles and journal information for any publications related to this data that you would like us to list.
9. **DOI**—If you would like a DOI for your data or if you have a DOI for a publication and would like this included in the metadata, please let us know.
10. **Supplementary Materials**—Do you have any additional material you would like to include with this data? For example, a final report , presentation, or data dictionary.
11. **Additional Points of Access**—Are there any existing or required access points for your data such as a specialized data repository for your field?