**Directorate for Mathematical and Physical Sciences – Division of Chemistry (NSF-MPS/CHE)**

**Products of Research**

Describe the types of data and products that will be generated in the research, for example numerical data on chemical systems such as spectra, diffraction patterns, physical properties, time-dependent information on chemical and physical processes, theoretical formalisms, computational strategies, final or intermediate numerical results from theoretical calculations, software, and curriculum materials.

Give a short description of what "data" and "data collection" will mean in your research—explain what the contents of each dataset will be, including size if known. Name the type of data, the instrument or collection approach, and how the data will be sampled. *Consider these questions*:

* What data will be generated in the research?
* What data types will you be creating or capturing?
* How will you capture or create the data? (e.g. experimental measures, observational or qualitative, model simulation, processed etc.)
* If you will be using existing data, state that fact and include where you got it.
* What is the relationship between the data you are collecting and the existing data?

**Data Format**

Describe the format in which the data or products are stored (e.g., hardcopy notebook and/or instrument outputs, ASCII, html, jpeg or other formats). Where data are stored in unusual or not generally accessible formats, explain how the data may be converted to a more accessible format or otherwise made available to interested parties. You may also comment on the current or anticipated need for interested parties outside of your laboratory to access your primary data.

Explain the format of your data, and how that format will remain useable and accessible over time. While you aren’t specifically asked to describe what metadata will be necessary for accessing your data, such a description will demonstrate anticipation of the needs of future users of the data. Ensuring that metadata standards exist will also help you secure the longevity of your data. Think about what information “interested parties outside of your laboratory” would need to use your data. *Consider these questions*:

* Which file formats will you use for your data, and why?
* What transformations (to more shareable formats) will be necessary to prepare data for sharing?
* What form will the metadata describing/documenting your data take?
* How will you create or capture these details?
* Which metadata standards will you use and why have you chosen them? (e.g. accepted domain-local standards, widespread usage).
* What contextual details (metadata) are needed to make the data you capture or collect meaningful?
* What and who are the intended or foreseeable uses / users of the data?

**Access to Data and Data Sharing Practices and Policies**

"Access to data" refers to data made accessible without explicit request from the interested party, for example those posted on a website or made available to a public database. Describe your plans, if any, for providing such general access to data, including websites maintained by your research group, and direct contributions to public databases (e.g., the Protein Data Bank, Cambridge Crystallographic Data Centre, Inorganic Crystal Structure Database in Karlsruhe, Zeolite Structure Database). Also note if you submit your data in the form of tables, graphs, computer code or other format to the supplementary materials sections of peer-reviewed journals. Describe your practice or policies regarding the release of data for access, for example whether data are posted before or after formal publication. Finally, note as well any anticipated inclusion of your data into databases that mine the published literature (e.g., PubChem, NIST Chemistry WebBook). "Data sharing" refers to the release of data in response to a specific request from an interested party. Describe your policies for data sharing, including where applicable provisions for protection of privacy, confidentiality, intellectual property, national security, or other rights or requirements.

* The differentiation between "Access to data" and "Data sharing" is a key distinction made by the CHE data management plan guidelines. Please keep in mind that you are expected to adequately provide responses for both how you plan on making your data accessible without a specific request from a researcher, and how you will be able to provide data upon demand. Use this section to also explain issues of confidentiality and intellectual property as their impact on the dissemination of your data. *Consider these questions*:
  + How and when will you make the data available? (Resources needed: equipment, systems, expertise, etc.)
  + What is the process for gaining access to your data? Where will you provide general access to data (websites, public databases, or other databases)
  + Do you plan on publishing findings which rely on the data? If so, do your prospective publishers place any restrictions on other avenues of publication?
  + Explain details of any embargo periods for political/commercial/patent or publisher reasons.
  + How long will the original data collector/creator/principal investigator retain the right to use the data before opening it up to wider use?
  + Are there ethical and privacy issues? If so, how will these be resolved?
  + Who will hold the intellectual property rights to the data and how might this affect data access?

**Policies and Provision for Re-Use, Re-Distribution and Products of Derivatives**

Describe your policies regarding the use of data provided via general access or sharing. For example, if you plan to provide data and images on your website, will the website contain disclaimers, or conditions regarding the use of the data in other publications or products? Describe these disclaimers and/or terms of use.

Explain how the policies you outline in the section above can be applied to the re-use and re-distribution of your data. Identify who will be allowed to use your data, how they will be allowed to use your data and whether or not they will be allowed to disseminate your data. **If you will be restricting access, use or dissemination of the data, you must explain how you will codify and communicate these terms.** *Consider these questions*:

* Will any permission restrictions need to be placed on the data?
* What and who are the intended or foreseeable uses / users of the data?
* How will the dataset be licensed if rights exist? (e.g. any restrictions or delays on data sharing needed to protect intellectual property, copyright or patentable data.)

**Archiving of Data**

Describe how data will be archived and how preservation of access will be handled. For example, will hardcopy notebooks, instrument outputs, and physical samples be stored in a location where there are safeguards against fire or water damage? Is there a plan to transfer digitized information to new storage media or devices as technological standards or practices change? Will there be an easily accessible index that documents where all archived data are stored and how they can be accessed?

This portion of the Data Management Plan asks the researcher to provide a long-term strategy for archiving and preserving the data from the research described in the proposal. *Consider these questions*:

* What is the long-term strategy for maintaining, curating and archiving the data?
* Which archive/repository/database have you identified as a place to deposit data?
* What procedures does your intended long-term data storage facility have in place for preservation and backup?
* How long will/should data be kept beyond the life of the project?

Also consider these questions about the data and associated information that will be deposited:

* What data will be preserved for the long-term?
* What transformations will be necessary to prepare data for preservation?
* What metadata/documentation will be submitted alongside the data or created on deposit/transformation in order to make the data reusable?
* What related information will be deposited (e.g. references, reports, research papers, fonts, the original bid proposal, etc.)?

For **non-digital data** consider this question:

* Where will hardcopy notebooks, instrument outputs, and physical samples be stored (consider locations with safeguards against fire or water damage)?