Plan Overview

A Data Management Plan created using DMPTool

Title: Science and Engineering Network for Solar Energy Innovations

Creator: michael heagy

Affiliation: New Mexico Institute of Mining and Technology (nmt.edu)

Principal Investigator: michael heagy

Data Manager: michael heagy

Funder: National Science Foundation (nsf.gov)

Funding opportunity number: NSF 155517

Template: NSF-CHE: Chemistry Division

Last modified: 02-17-2015

Copyright information:

Science and Engineering Network for Solar Energy Innovations

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.

The EPSCoR Track II consortium will generate a broad array of data including: experimental data tabular data; spectral data, analytical data from ion chromatography, mass spectral data, fluorescence lifetime data, transient absorption and quantum yields of photoproducts. With the exception of long-term instrumentation data, most data sets collected by EPSCoR will be small in size (< 10 MB). Currently, we anticipate that no more than 25 GB of data will be collected as part of the project with the larger data volumes associated with rmedia data products, videos, and learning modules.

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.

Most of the data collected (by volume) will be spectral by nature; i.e. NMR, optical absorption, IR, transient absorption. These data will be documented using tools that support the ISO 19115 metadata standard. Some of the data collected will be textual in nature and will be saved as text, MS Word, and pdf documents (e.g., supplementary info, experimental notes). Any tabular data collected will be captured in spreadsheets or data tables and saved in .csv files for long-term accessibility; we plan to make use of the DataUp tool available through DataONE that converts Excel spreadsheets into preservation-ready products and that enables the creation of associated metadata. In addition, presentation data such as pptx, or project files will be archived within the Track I EPSCoR Owncloud.

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.

Data will be made available through several mechanisms. First, during the data and information gathering portion of the project, most data will be available to all project participants via a password-controlled website that houses the virtual lab notebook, copies of non-copyrighted materials, drafts of working white papers and publications, and other data and information generated during the course of the project. The exception includes data and information related to inventions that are to be patented.

Data collected through this project may be embargoed for a period of up to one year to allow time for publication by students and researchers; any exceptions to this embargo period must be approved by the Project Director in writing. There will be no charge for data and information and they will be easily discoverable and acquired via the EPSCoR data portal. The research will be conducted in full compliance with both federal and University regulations.

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.

Our policy towards accessing the data will be structured so that it's available upon request. There are no charges for this data and in as much as this proposal was primarily constructed within the SENSEI team via Dropbox, we envision using a similar intitial repository for team data.

ntended or foreseeable users of the data include other researchers, business and industry, educators, governmental and nongovernmental organizations, and educators and students. The data can be made available upon request via email to those interested in any of the uploaded documents. We wish to invoke a one year data embargo on any uploaded material until the manuscript or patent has been approved for publication, etc. Intellectual property rights belong to the originating lab/university where results were obtained.

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?

In addition to preliminary data sharing via Dropbox, our consortium may make use of the Track I repository system known as Owncloud. This system has been set up as part of the Track I EPSCoR award and contains several policies of data preservation, term of project and sharing that will be invoked in our Track II proposal.

Key data such as published spectral data and associated ISO 19115 compliant metadata will be preserved for the long-term after transformation to generic, preservation-ready formats.