

Plan Overview

A Data Management Plan created using DMP Tool

Title: Center: Great Lakes Center for Fresh Waters and Human Health

Creator: Tracey Wiley

Affiliation: University of North Carolina at Chapel Hill (UNC-CH) (unc.edu)

Funder: National Science Foundation (nsf.gov)

Funding opportunity number: NSF 23-1

Grant: 161753

Template: NSF-CHE: Chemistry Division

Project abstract:

https://ramses2.research.unc.edu/ramses/de_abstract.cfm?prop_id=183475

Start date: 02-01-2024

End date: 01-31-2029

Last modified: 07-08-2024

Copyright information:

The above plan creator(s) have agreed that others may use as much of the text of this plan as they would like in their own plans, and customize it as necessary. You do not need to credit the creator(s) as the source of the language used, but using any of the plan's text does not imply that the creator(s) endorse, or have any relationship to, your project or proposal

Center: Great Lakes Center for Fresh Waters and Human Health

Products of the Research

Describe the types of data (including metadata and annotations, primary or analyzed) and products that will be generated by the research, for example description of samples, numerical data on chemical systems such as spectra, chemical and physical properties, time-dependent information on chemical and physical processes, theoretical formalisms, experimental protocols, algorithm specifications, database schemas and data tables, data produced by simulations and software. Data and products generated from Broader Impact activities, such as educational materials, participant information, tutorials and other web-based materials, as well as assessment results, should also be included in the DMP.

Observational (primary)

Chemical and Biological (primary)

Microbial (analyzed)

Raw Samples (primary)

Data Format

Describe the format and media 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. In general, solutions and remedies to providing data in an accessible format should be provided with minimal added cost.

Raw Data (.dat and ASCII)

Excel Spreadsheet Files (.xlsx) converted to ASCII for analysis (Matlab, R)

Field and sample data to include date, time, location, and depth

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 or software repositories (e.g., NMRShiftDB, the Protein Data Bank, Cambridge Crystallographic Data Centre, Inorganic Crystal Structure Database in Karlsruhe, Zeolite Structure Database and Github). For software or code developed as part of the project, include a description of how users can access the code (e.g. licensing, open source) and specific details of the hosting, distribution and dissemination plans. Also describe your practice or policies regarding the release of data for access, for example whether data are posted before or after formal publication. Note as well any anticipated inclusion of your data in databases that mine the published literature (e.g. PubChem, NIST Chemistry WebBook). Consider using the Digital Object Identifiers (DOI) assignment mechanism not just for journal articles, but for suitably-archived, publishable data sets.

"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. Discussion on the compliance with the NSF's *Public Access Policy* is also encouraged.

All files will be backed up weekly and stored both onsite on personal computers with redundant backup on shared network drive and offsite on portable media.

Data will be shared among project participants using email and password-protected cloud storage.

Embargo Period - a maximum of two years will be taken to ensure adequate time for original research collaborators to publish results.

Access - after embargo period, complete data set made available thru NCBI and JGI thru local web servers and through inclusion in DRYAD Digital Repository.

Policies for Re-Use, Re-Distribution, and Production of Derivatives

Describe your policies regarding the use of data provided via general access or sharing. Practices for appropriate protection of privacy, confidentiality, security, intellectual property, and other rights should be communicated. The rights and obligations of those who access, use, and share your data with others should be defined. 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?

Project collaborator users will be able to contact the PIs with questions regarding access and use of the data.

Once the embargo period ends the project data will be public domain.

Two requirements: Publications presenting project data will be acknowledged by all users. Footnotes and addenda to the database will be maintained regularly to track publications.

Archiving of Data

Describe when the data should be archived, how data will be archived, and how preservation of access will be handled. Are there provisions for data backup? 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? What are the physical and cyber resources and facilities that will be used for data preservation and storage? Will there be an easily accessible index that documents where all archived data are stored and how they can be accessed? What are the roles and responsibilities of all parties with respect to the management and archiving of the data after the grant ends? How long will the data be maintained after the grant ends?

CHE-supported large research centers or other programs may specify more stringent data storage, sharing and archiving procedures for research conducted under their awards. Such requirements will be specified in the program solicitation and award conditions.

All data will be retained for at least 3 years beyond the award period, as required by NSF. PI will work closely with NCBI and JGI curators related to data preservation. Redundant storage in this clearing house will guarantee the widest dissemination possible to future access for the supported data and the core data set.

Planned Research Outputs

Dataset - "Observational, Chemical and Biological data"

New data will be available to the P.I. and collaborators and can be retrieved from data repositories for future publications.

Planned research output details

Title	Type	Anticipated release date	Initial access level	Intended repository(ies)	Anticipated file size	License	Metadata standard(s)	May contain sensitive data?	May contain PII?
Observational, Chemical and Biological data	Dataset	2024-02-29	Open	NCBI, JGI		Creative Commons Attribution Share Alike 4.0 International	e Metabolomics Standards Initiative-Environmental Context Working Sub-Group, MSI-ECWSG ABCD (Access to Biological Collection Data)	No	No