Plan Overview

A Data Management Plan created using DMPTool

Title: LTREB: Drivers of temperate forest carbon storage from canopy closure through successional time

Creator: Jason Tallant - **ORCID:** 0000-0003-4790-8772

Affiliation: University of Michigan–Ann Arbor (umich.edu)

Principal Investigator: Luke Nave, Knute Nadelhoffer, Christopher Gough

Data Manager: Luke Nave

Funder: National Science Foundation (nsf.gov)

Funding opportunity number: 35572

Grant: https://www.nsf.gov/awardsearch/showAward?AWD_ID=1353908

Template: NSF-BIO: Biological Sciences

Project abstract:

This document details plans for the management and dissemination of scientific materials produced and synthesized as part of the project: “LTREB: Drivers of temperate forest carbon storage from canopy closure through successional time.”

Last modified: 09-18-2018

Copyright information:

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

---

LTREB: Drivers of temperate forest carbon storage from canopy closure through successional time

Data and Materials Produced

Describe the types of data, physical samples or collections, software, curriculum materials, and other materials to be produced in the course of the project. (For collaborative proposals, the DMP must cover all the various data types being collected by each collaborator.)

This project generates 7 types of data and scientific materials:

1. Physical samples: These are the samples that we collect from the long-term plots comprising our experimental design. They include samples of soil, roots, tree cores, leaves, and leaf litterfall. Some samples are stored in bulk and processed forms, e.g., soils, roots, and leaf litterfall, we store samples of bulk dried material and also milled subsamples for laboratory analysis. An inventory and sample tracking system is integrated with the UMBS Information Management System (IMS) in order that sample documentation is related to published data sets, research sites, and related projects.
2. Field observations: These are hard-copy data sheets containing observations made in the field on our sampling infrastructure (e.g., plots). Examples include soil temperature and respiration rates, tree species and diameter for each tree in each permanent plot, and deployment notes for root ingrowth cores. These are subsequently transcribed for use in compiled datasets. The raw data tables are kept in their original condition and are referenced by compiled datasets.
3. Laboratory data: These consist of hard-copy data sheets and digital files that contain information recorded during lab processing (e.g., masses of physical samples), archiving (e.g., accession numbers), and chemical analysis (e.g., percent carbon by elemental analyzer). Transcribed and instrument outputs are kept in raw form and referenced for the creation of compiled data sets.
4. Compiled datasets: These are minimum units of data that we prepare for publication via institutional and outward-facing data sharing websites. They are quality-assured products that typically require some compilation (e.g., merging lab processing and chemical analysis data) and computation (e.g., calculation of averages across multiple analytical replicates). These datasets would be considered by most researchers and data users to be processed data, they have been quality controlled and are ready for analysis.
5. Historical records, samples, and datasets: These artifacts include types 1, 2, 3, and 4 above. They exist in multiple (redundant) formats including physical (e.g., original samples in vials,

photocopied lab notebooks or datasheets) and digital (e.g., .CSV and TXT files).

6. Published datasets: These are the finalized, documented, fully attributed, published datasets consisting of packages of digital files needed by contemporary researchers to make use of our project's information resources. The files typically include the compiled datasets, along with relevant metadata and links to peer-reviewed or other publications.
7. Publications, presentations, and reports: These are the most fully developed and detailed resources produced by our project. Publications include peer-reviewed journal articles and other products (e.g., teaching modules, theses); presentations may be indexed abstracts or video or audio recordings from professional conferences; reports include annual reports for NSF that document progress towards project objectives.

Standards, Formats and Metadata

Describe the standards to be used for all the data types anticipated, including data or file format and metadata.

Basic standards (e.g., sample vials with labels that can be tracked across analytical and compiled datasets, clear handwriting on data sheets) and written protocols exist and are applied to the physical artifacts of our research, but as the future users of our scientific resources will most likely be using resources from types 4-7 above, we are most concerned with proper attribution of metadata and file formats related to compiled and published datasets, scanned documents, and publications, presentations and reports. Metadata for the project conform to the Ecological Metadata Language (EML) standard. Associated researchers and the UMBS Data Manager will curate metadata within the UMBS IMS. The IMS subsequently generates an EML compliant XML document to be shared with the data set and published to EDI for data that are public. EDI generated DOI’s and code come back to the UMBS IMS for use by the local research community. Our published datasets and publications form the core of these resources because these are the most detailed and universally accessible in terms of viewing project data and understanding their origins, methods, and locations.

Roles and Responsibilities

Describe the roles and responsibilities of all parties with respect to the management of the data (including contingency plans for the departure of key personnel from the project).

Final responsibility for data management (including future inquiries regarding access and use not already specified by licensing of publicly released data through EDI) rests with the project PI (Nadelhoffer), or upon retirement, his delegates (e.g., project co-PI's Nave and Gough, UMBS Information Manager Tallant). Nave is responsible for preparation of data and other scientific resources for archival and curation during the course of the project, working in close collaboration with Gough, UMBS Research Specialist Den Uyl, Tallant, and UM staff and interns (e.g., EDI) involved in data management.

Dissemination Methods

Describe the dissemination methods that will be used to make data and metadata available to others during the period of the award, and any modifications or additional technical information regarding data access after the grant ends.

Datasets are compiled, versioned, and shared among project team members using multiple redundant copies on a local file server and an institutional supported cloud drive. Data in this phase of development and use (i.e., pre-publication) are stored and shared in a variety of formats that are geared towards the project team rather than by the wider research community. To disseminate final data sets to the broader community, the UMBS IMS serves as the local data and research information repository. Metadata and public data sets are pushed to EDI. EDI publicly publishes metadata and data sets through their data portal. As a member node of DataONE, EDI pushes metadata from the project to DataONE making the deposited data discoverable at multiple scales.

Policies for Data Sharing and Public Access

Describe the PI's policies for data sharing, public access and re-use, including re-distribution by others and the production of derivatives. Where appropriate, include provisions for protection of privacy, confidentiality, security, intellectual property rights and other rights.

Per NSF requirements, processed data sets will be published to EDI two years after completion of data collection. All published data sets will carry Creative Commons By Attribution (CCBY; https://creativecommons.org/licenses/by/4.0/). Prior to publication, the PI and co-PIs may decide to share data with collaborators or other researchers upon request. This will be done by granting file

access to collaborators through institutionally supported cloud services. Downloading and versioning will be tracked.

Archiving, Storage and Preservation

Where relevant, describe plans for archiving data, samples, software, and other research products, and for on-going access to these products through their lifecycle of usefulness to research and education. Consider which data (or research products) will be deposited for long-term access and where. (What physical and/or cyber resources and facilities (including third party resources) will be used to store and preserve the data after the grant ends?)

The discoverability and usability of our final, publicly available processed data is ensured through the next 20 years with the institutional support that exists for the UMBS IMS, with redundant archival on DataONE. These types of curation services are widely viewed as best practices in our research community; while they do not guarantee “permanent” access, affirmation of accessibility and interoperability for two decades is significant. Raw data and physical data sheets will be archived, as has long been done at UMBS, in rodent-proof filing cabinets in a designated, windowless room in a climate-controlled facility. Provenance of the physical samples in the archive building (renovated with FSML support) is tracked in the IMS sample database. Their accessibility is ensured through the duration of the careers of the collective project team at UMBS; beyond that, they will remain secure and stored at stable conditions as long as institutional support exists for infrastructure at UMBS, which has been operating for 110 years.
