

Plan Overview

A Data Management Plan created using DMPTool

Title: Effects of mine waste materials in the north-central Mojave desert

Creator: Hong Xu

Affiliation: Texas A&M University – Corpus Christi (tamucc.edu)

Principal Investigator: Hong Xu

Data Manager: Hong Xu

Funder: National Science Foundation (nsf.gov)

Funding opportunity number: 23144

Template: NSF-EAR: Earth Sciences

Last modified: 09-16-2016

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

Effects of mine waste materials in the north-central Mojave desert

Types of data

Preservation of all data, samples, physical collections and other supporting materials needed for long-term earth science research and education is required of all EAR-supported researchers.

- Physical samples
 - Bulk samples of different media including mine tailings, waste rock, streambed sediments, and background materials. All sample stored in borosilicate glass jars with Teflon lids.
 - Streambed water samples stored in HDPE bottles.
 - Field notes about information of bulk samples and water samples
- Data of samples
 - Data about the bulk samples in XLS/CSV
 - Data about the water samples in XLS/CSV
 - Scanned images of field notes in TIFF
- Lab data
 - Lab reports in DOC/PDF(PDF/A)
 - Experiment data in XLS/CSV
 - Images in TIFF
- Geochemistry modeling software tools
 - Software program code in EXE
- Scholarly works: electronic copies of conference presentations and pre-prints of published articles in PDF(PDF/A)
- Lesson materials:
 - Lesson plans developed by public school teachers in PDF(PDF/A)
 - Audio/video in WAV

Data and metadata standards

Data archives must include easily accessible information about data holdings, including quality assessments, supporting ancillary information, and guidance and aids for locating and obtaining data.

We will register all our physical samples with the System for Earth Sample Registration (SESAR) [<http://www.geosamples.org/>]. Then, every sample will have an International GEO Sample Number (IGSN) [<http://www.geosamples.org/aboutigsn>]. The documentation of IGSN metadata can be found at [<http://trac.gfz-potsdam.de/igsn/wiki/WikiStart>]. Data of samples will be prepared according to the IGSN metadata fields to make the sample registration straightforward. Field notes template will be developed at the beginning of the project to include all necessary background information fields that would help develop metadata for samples during the project.

Lab data spreadsheets will be developed according to the Geochemistry Data Templates designed by Integrated Earth Data Application (IDEA)'s EarthChem Library [<http://www.earthchem.org/data/templates>] to include the recommended sheet structure design and to use the standard EarthChem controlled vocabularies measured for parameter, experiment method, mineral, and variable unit. X-Ray mapping images will be prepared with required and optional metadata fields as specified by the IDEA's MediaBank [<http://media.marine-geo.org/contribute>]. Documentation of geochemistry modeling tools and codes will be prepared with necessary notes and comments to make the replication of the modeling possible. When deposited into the EarthChem Library [<http://www.earthchem.org/library/submit>] at the end of our project, the lab data and the software code will be submitted with the library's cataloging metadata, as defined by the DataCite metadata kernel [<http://schema.datacite.org/>].

Policies for access and sharing

It is the responsibility of researchers and organizations to make results, data, derived data products, and collections available to the research community in a timely manner and at a reasonable cost. In the interest of full and open access, data should be provided at the lowest possible cost to researchers and educators. This cost should, as a first principle, be no more than the marginal cost of filling a specific user request. Data may be made available for secondary use through submission to a national data center, publication in a widely available scientific journal, book or website, through the institutional archives that

are standard for a particular discipline (e.g. IRIS for seismological data, UNAVCO for GP data), or through other EAR-specified repositories. Data inventories should be published or entered into a public database periodically and when there is a significant change in type, location or frequency of such observations. Principal Investigators working in coordinated programs may establish (in consultation with other funding agencies and NSF) more stringent data submission procedures.

The research data and software code will be accessible to the PI and other researchers involved during the project. Three copies will be made for the data and software code; the first copy will be stored on the lab computer, the second copy will be stored on the network server which the university's IT department manages, and the third copy will be stored on an external hard drive which is kept at the PI's residence (away from the office and lab). Once the project is finished, the datasets and software code will be deposited into the IEDA EarthChem Library and will be made accessible publicly and freely.

Policies and provisions for re-use, re-distribution

For those programs in which selected principle investigators have initial periods of exclusive data use, data should be made openly available as soon as possible, but no later than two (2) years after the data were collected. This period may be extended under exceptional circumstances, but only by agreement between the Principal Investigator and the National Science Foundation. For continuing observations or for long-term (multi-year) projects, data are to be made public annually.

The datasets and software code will be published by IEDA under the Creative Common License BY-NC-SA 3.0. [<http://creativecommons.org/licenses/by-nc-sa/3.0/>]. The research data should be cited in the format of Creator (PublicationYear): Title. Publisher. Identifier. IEDA will have Digital Object Identifiers (DOIs) assigned to the datasets.

Plans for archiving and preservation of access

Remember - Data may be made available for secondary use through submission to a national data center, publication in a widely available scientific journal, book or website, through the institutional archives that are standard for a particular discipline (e.g. IRIS for seismological data, UNAVCO for GP data), or through other EAR-specified repositories.

To increase access to the research outcome from this project, the PI will deposit the data and related software code into the IEDA EarthChem Library [<http://www.earthchem.org/library>], an online data

repository that archives and publishes data and other related digital artefacts from geosciences research. The physical samples will be stored in a secure, locked cabinet and made available for examination for at least 7 years after the project completion. The file formats of the datasets will be converted to preservation formats (CSV, PDF(PDF/A), TIFF, WAV).

Conference materials, lesson plan and publications will be deposited in an institutional repository for open access.

Software code will be deposited in github.
