Plan Overview

A Data Management Plan created using DMPTool-Stage

Title: DMPRoadmap: Making Data Management Plans Actionable

Creator: Stephanie Simms - **ORCID:** 0000-0001-9129-3790

Affiliation: University of California, Irvine

Principal Investigator: Stephanie Simms, Günter Waibel

Data Manager: Stephanie Simms

Funder: National Science Foundation (nsf.gov)

Funding opportunity number: 27195

Grant: https://www.nsf.gov/awardsearch/showAward?AWD_ID=1745675

Template: NSF-GEN: Generic

Last modified: 07-01-2024

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

DMPRoadmap: Making Data Management Plans Actionable

Types of data produced

Types of data, samples, physical collections, software, curriculum materials, and other materials to be produced in the course of the project.

Three types of data will be generated during this project:

1) Qualitative and quantitative survey results collected as part of Unit One will be captured in paper forms and via electronic methods (e.g., Qualtrics). Paper surveys will be digitized by hand and and originals will be kept for the duration of the project. All survey data will be compiled as csv files. We will comply with Institutional Review Board policies and secure approval before conducting the surveys with human subjects. Simms and Jones will be responsible for all survey data.

2) Software produced or modified as part of the project will be maintained in the DMPRoadmap GitHub repository during and after the project (https://github.com/DMPRoadmap/roadmap), where it is available under an open source MIT License. Riley will be responsible for managing the repository and related documentation.

3) Miscellaneous research products such as additional use cases developed as part of Unit Three, community feedback, project status communications, etc. will be disseminated publicly throughout the project via the project website, blogs, peer-reviewed articles, and data repositories as appropriate. Communication channels and resulting products will be managed by Simms and Jones.

Data and metadata standards

Standards to be used for data and metadata format and content (where existing standards are absent or deemed inadequate, this should be documented along with any proposed solutions or remedies).

The DMPRoadmap software produced by the project will conform to accepted community best practices, including version control (using Git) with tagging of major releases, a permissive open-source license (MIT License), public availability in a community repository (GitHub), inline comments, reference and tutorial documentation with download and installation instructions that is available from within the software and from a community website, and extensive test coverage.

Policies for access and sharing

Policies for access and sharing; Provisions for appropriate protection of privacy, confidentiality, security, intellectual property, or other rights or requirements.

All three types of data produced in the course of this project will be publicly available for review, evaluation, and use as they are generated during the project and after its completion. Announcements about software and data availability will be made using a variety of channels (e.g., blogs, Twitter, email lists) targeting all interested stakeholder communities. All survey data will be anonymized to remove personally identifiable information in compliance with IRB protocols for human subjects research.

Policies for re-use, redistribution

Policies and provisions for re-use, re-distribution, and the production of derivatives.

All software products resulting from this project will be reusable and redistributable during the project and after its completion. The only restriction placed on redistribution of the software is that the copyright and license statement be kept intact as required by the MIT open source license. The software is expected to be of interest to national and international data infrastructure providers, data centers and repositories, institutional administrators, and individual researchers.

Plans for archiving and preservation

Plans for archiving data, samples, and other research products, and for preservation of access to them.

DMPRoadmap is an active software product and will continue to be managed in the community GitHub repository. We will document current and future releases with release notes and the code for each version will be publicly available using Git history.

All survey data will be deposited to Dash, a data publication platform, and preserved in the Merritt Repository Service to provide public access and long-term storage upon completion of the project. All data will be publicly available upon deposit and findable through a DataCite DOI granted by Dash. The datasets will be maintained for as long as they are of continuing value to researchers and project collaborators, for a minimum of five years.

Dash and Merritt are both managed by the University of California Curation Center (UC3). Merritt relies on a highly fault tolerant microservices architecture with significant redundancy of all computational and storage components and has not experienced any data loss over its seven years of production operation.