

Plan Overview

A Data Management Plan created using DMPTool

Title: CAPERS - Cloud Applications and Platforms Enhancing Research Scalability

Creator: Lew Lefton

Affiliation: Georgia Institute of Technology

Principal Investigator: Lew Lefton

Data Manager: Lew Lefton

Funder: National Science Foundation (nsf.gov)

Funding opportunity number: 39576

Template: NSF-CISE: Computer and Information Science and Engineering

Last modified: 02-15-2019

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

CAPERS - Cloud Applications and Platforms Enhancing Research Scalability

Types of data

The Data Management Plan should describe the types of data, metadata, scripts used to generate the data or metadata, experimental results, samples, physical collections, software, curriculum materials, or other materials to be produced in the course of the project.

CAPERS will be the primary organization serving as a principal interface between the CISE research and education community and public cloud computing providers. The types of data produced by this organization will be one of the following

- a) internal operational data (e.g. budget and finances, workflow documentation, cloud provider partnership agreements, etc.)
- b) external training manuals, strategic planning documents, community updates

Data and metadata standards

The Data Management Plan should address the 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). It should also cover any other types of information that would be maintained and shared regarding data, e.g. the means by which it was generated, detailed analytical and procedural information required to reproduce experimental results, and other metadata.

Only relevant metadata will be basic file metadata (time and date of last modification, size, etc).

Policies for access, sharing, and privacy

The Data Management Plan should address the policies for access and sharing including provisions for appropriate protection of privacy, confidentiality, security, intellectual property, or other rights or requirements. It should cover any factors that limit the ability to manage and share data, e.g. legal and ethical restrictions on access to human subject data.

As an NSF funded organization with no sensitive data, we anticipate that all CAPERS data will be publicly available, with the exception of internal personnel records.

Policies for re-use, re-distribution, derivatives

The Data Management Plan should address the policies and provision for re-use, re-distribution, and the production of derivatives.

As an NSF funded organization with no sensitive data, we anticipate that all CAPERS data will be open and freely sharable. Things like training manuals, strategic planning documents, community updates will be open to anyway via websites. Internal operational data will be available to anyone upon request.

Plans for archiving and preservation

The Data Management Plan should address the plans for archiving data, samples, and other research products, and for the preservation of access to them. It should cover the period of time the data will be retained and shared; how data are to be managed, maintained, and disseminated; and mechanisms and formats for storing data and making them accessible to others, which may include third party facilities and repositories.

All CAPERS data, both operational and external-facing will be backed up using enterprise-grade services and following standard industry practices. We will leverage version control systems like github wherever possible to include version control for documentation, etc.

Roles and responsibilities

The Data Management Plan should clearly articulate how the PI and co-PIs plan to manage and disseminate data generated by the project. The plan should outline the rights and obligations of all parties as to their roles and responsibilities in the management and retention of research data, and consider changes that would occur should a PI or co-PI leave the institution or project. It should describe how the research team plans to deposit data into any relevant and appropriate disciplinary repositories that are appropriately managed and that are likely to maintain the metadata necessary for future use and discovery. Any costs associated with implementing the DMP should be explained in the Budget Justification.

The PIs will be responsible for initial data management as described above. The CAPERS budget includes a full time director of the organization whom, once she or he is hired, will assume these responsibilities.

