

CSET UTC Data Management Plan

Updated: May 16, 2017

This document provides Data Management Plan (DMP) information related to Center for Safety Equity (CSET) UTC projects. All research proposals submitted to CSET will be required to contain a project-specific DMP as outlined here. This CSET DMP is intended to inform researchers on how to handle digital data both during and after the completion of CSET research projects and how research proposals will conform to DOT policy on the dissemination and sharing of research results. All CSET-funded research projects are expected to be conducted pursuant to their approved DMPs. A DMP may evolve with the research project and should be reviewed for possible revision whenever a data management procedure is changed. Modifications and changes to a DMP are subject to approval by CSET.

PIs should adhere to the project-specific DMP timeline as follows:

1. Preliminary DMP – submitted with research project proposal.
2. Finalized DMP – submitted at the end of the first quarterly review period after project award announcement.
3. Modified DMP – submitted at the end of the quarterly review period in which a significant change to the DMP occurred.

In general, project-specific DMPs will include the following elements:

1. Data Descriptions – the anticipated final research data to be produced during the entirety of the active project period and may include
 - a. Surveys and interviews
 - b. Cataloged and inventoried infrastructure elements
 - c. Trauma and injury data from hospitals and emergency response
 - d. Statewide- and village-based crash records
 - e. Historic weather
 - f. Traffic characteristics and traveler behavior
 - g. Spatial data related to transportation infrastructure
2. Data Standards – the standards that will be used for data and metadata format and content.
3. Data Access Policies – the policies for access and sharing the final research data, including:
 - a. provisions for appropriate protection of privacy
 - b. confidentiality
 - c. security
 - d. intellectual property
 - e. other rights or requirements.
4. Data Derivatives - the policies and provisions for re-use, re-distribution, and the production of derivatives.
5. Data Archiving and Preservation – the plans for archiving and preserving access to final research data and other research products.

Guidance on these DMP elements are discussed herein.

Data Description

The project-specific DMP must include a section that provides a description of the data that will be gathered by the PI or affiliated researchers during the course of the project. This description must address the nature, scope, and scale of the data that will be collected. Describe the characteristics of the data, their relationship to other data, and provide sufficient detail so that reviewers will understand any disclosure risks that may apply. Discuss value of the data over the long-term.

A DMP data description section will typically provide the following level of detail:

1. Name the data, data collection project, or data producing program.
2. Describe the purpose of the research.
3. Describe the data that will be generated in terms of nature and scale (e.g., numerical data, image data, text sequences, video, audio, database, modeling data, source code, etc.).
4. Describe methods for creating the data (e.g., simulated, observed, experimental, software, physical collections, sensors, satellite, enforcement activities, researcher-generated databases, instrument generated digital data output such as images and video, etc.).
5. Discuss the period of time data will be collected and frequency of update.
6. If using existing data, describe the relationship between the data you are collecting and existing data.
7. List potential users of the data.
8. Discuss the potential value of the data have over the long-term for not only your institution, but also for the public.
9. If you request permission not to make data publicly accessible, explain rationale for lack of public access.
10. Indicate the party responsible for managing the data.
11. Describe how you will check for and ensure adherence to this data management plan.

Standards Used

Describe the anticipated formats that your data and related files will use. To the maximum extent practicable, and in accordance with generally accepted practices in your field, your DMP should address how you will use platform-independent and non-proprietary formats to ensure maximum utility of the data in the future. If you are unable to use platform-independent and non-proprietary formats, you should specify the standards and formats that will be used and the rationale for using those standards and formats. Identify the metadata standards you will use to describe the data. Project-specific DMPs should address the following:

1. List in what format(s) the data will be collected. Indicate if they are open or proprietary.
2. If you are using proprietary data formats, discuss your rationale for using those standards and formats.
3. Describe how versions of data be signified and/or controlled.
4. If the file format(s) you are using is (are) not standard to your field, describe how you will document the alternative you are using.
5. List what documentation you will be creating in order to make the data understandable by other researchers.



6. Indicate what metadata schema you are using to describe the data. If the metadata schema is not one standard for your field, discuss your rationale for using that scheme.
7. Describe how will the metadata be managed and stored.
8. Indicate what tools or software is required to read or view the data.
9. Describe your quality control measures.

Access Policies

Describe any access restrictions that may apply to your data. In general, data from research projects funded wholly or in part by CSET must be made publicly accessible. Exceptions to this policy are data that contain personally identifiable information, confidential business information, or classified information.

Protecting research participants and guarding against the disclosure of identities and/or confidential business information is an essential norm in scientific research. Your project-specific DMP should address these issues and outline the efforts you will take to provide informed consent statements to participants, the steps you will take to protect privacy and confidentiality prior to archiving your data, and any additional concerns (e.g., embargo periods for your data). If necessary, describe any division of responsibilities for stewarding and protecting the data among PIs or other project staff.

If you will not be able to anonymize the data in a manner that protects privacy and confidentiality while maintaining the utility of the dataset, you should describe the necessary restrictions on access and use. If a project should include human subject research, the project must go through the University of Alaska Fairbanks IRB or the PIs home institution IRB if they have one. The project-specific DMP should describe the IRB process and how informed consent forms will permit sharing with the research community and whether additional steps may be required to protect privacy and confidentiality. . Project-specific DMPs should address the following information in the Access Policies section:

1. Describe what data will be publicly shared, how data files will be shared, and how others will access them.
2. Indicate whether the data contain private or confidential information and if so:
 - a. Discuss how will you guard against disclosure of identities and/or confidential business information;
 - b. List what processes you will follow to provide informed consent to participants; and
 - c. State the party responsible for protecting the data.
3. Describe what, if any, privacy, ethical, or confidentiality concerns are raised due to data sharing.
4. If applicable, describe how you will anonymize your data before sharing. If not:
 - a. Identify what restrictions on access and use you will place on the data.
 - b. Discuss additional steps, if any you will use to protect privacy and confidentiality.

Re-Use, Redistribution, and Derivative Products Policies

The University of Alaska Fairbanks or the home institution of the research project PI holds the intellectual property rights for the data created by the project. The project-specific DMP must address, if appropriate, whether or not rights will be transferred to a data archive. If transfer rights are not described, rights will be maintained by the University of Alaska Fairbanks or the home institution. Identify whether any



copyrights apply to derivatives of the data or cases when data may be collected using copyrighted instruments. If you will be enforcing terms of use or a requirement for data citation through a license, this must be indicated in the project-specific DMP. Describe any other legal requirements that might need to be addressed. Also include the following information as appropriate:

1. Name who has the right to manage the data.
2. Indicate who holds the intellectual property rights to the data.
3. List any copyrights associated with data derivatives. If so, indicate who owns them.
4. Discuss any rights be transferred to a data archive.
5. Describe how your data will be licensed for reuse, redistribution, and derivative products.

Archiving and Preservation Plans

The CSET UTC will archive all data on the Harvard Dataverse (HD), <https://dataverse.harvard.edu/>, which a data archiving site currently considered conformant by the USDOT, meets the Data Seal of Approval, and allows for the recording of persistent identifiers linked to the data. When a final report is submitted for a project, the associated PI will have 60 days to archive the project data on HD and provide confirmation to CSET that the data has been uploaded. Upon upload, the PI must describe the dataset that is being archived with a minimum amount of metadata that ensures its discoverability including the generation of a data-specific DOI and must provide for maintenance of all identifiers throughout the 10-year preservation lifecycle of the data. The project-specific DMP plan must address how the archiving and preservation will meet these requirements.

In addition, the project-specific DMP must address and include the following archiving information:

1. Identify where the PI will store the data prior to being sent to an archive.
2. Describe how back-up, disaster recovery, off-site data storage, and other redundant storage strategies will be used to ensure the data's security and integrity.
3. Describe how data will be protected from accidental or malicious modification or deletion prior to receipt by the archive.
4. Discuss your chosen data archive's policies and practices for back-up, disaster recovery, off-site data storage, and other redundant storage strategies to ensure the data's security and integrity for the long-term.
5. Indicate why the 10-year preservation life is not appropriate if proposing an alternative usable life and data preservation period.

