



Center for Pedestrian and Bicyclist Safety

Data Management Plan

8/4/2023

Nicholas N. Ferenczak, PhD, PE
CPBS Director
University of New Mexico
cpbs@unm.edu
www.pedbikesafety.org

The Center for Pedestrian and Bicyclist Safety (CPBS) is a Tier-1 University Transportation Center (UTC) funded by the United States Department of Transportation (USDOT) through the Bipartisan Infrastructure Law (BIL). It is composed of University of New Mexico (UNM), San Diego State University (SDSU), University of California, Berkeley (UCB), University of Tennessee, Knoxville (UTK), and the University of Wisconsin-Milwaukee (UWM).

The Center will ensure that data will be managed for security, consistency, and public (i.e., open-access and non-proprietary) dissemination when appropriate. Data can be primary data collected by each project principal investigator (PI) or secondary data aggregated and analyzed for new information. This document details the general requirements for all activities funded by CPBS while it is conducting research, education, workforce development, and technology transfer.

This Program-Level Data Management Plan (DMP) generally describes how CPBS researchers will handle digital data both during and after a research project is completed. As per USDOT guidelines for public data access, PIs for each CPBS research project will be responsible for submitting a Project-Specific DMP and will have overall responsibility for tracking and managing progress and adherence. PIs can use this Program-Level DMP to reference or quote when developing their Project-Specific DMP. Future CPBS funding for the PI is contingent on successful implementation of their Project-Specific DMP. For guidance on creating a Project-Specific DMP, refer to the US DOT Public Access Guidance webpages, especially “Creating Data Management Plans for Extramural Researchers” at <https://doi.org/10.21949/1520571>.

1. Data Description

The Center expects to generate a variety of data including but not limited to the following types of data: [PIs choose applicable data for your specific project]:

- Research data:
 - Data generated from experiments, interviews, or surveys
 - Crash data
 - Infrastructure data
 - Demographic information
 - Traffic operations data
- Educational activity data:
 - Course syllabi
 - Video/webcasts of presentations and lectures
 - Student surveys
- Technology transfer and workforce development activity data:
 - Video/webcasts of presentations/seminars/invited talks
 - Summary of event attendees and outcomes
 - Conference proceedings

Each PI must provide the following information regarding the data identified above. Amendments may be made throughout the project. If amendments are made, the PI must promptly notify CPBS of those amendments and update the Project-Specific DMP as needed.

1. 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.).
2. Describe methods for creating the data (e.g., simulated; observed; experimental; software; sensors; satellite; enforcement activities; etc.).
3. Discuss when data will be collected and frequency of update.

2. Data Standards

PIs will store data in platform-independent, non-proprietary formats whenever possible. Examples of such data formats include CSV, TXT, etc. If PIs are using proprietary data formats, they will discuss their rationale and detail the software required to read the data format.

The PI will describe their project's data quality control measures (both internal measures and any measures taken by original data owners). Metadata describing datasets and the terms of access will be made available in a schema appropriate to the type of data collected, and with a DCAT-US Schema v1.1 .JSON file, as described at <https://resources.data.gov/resources/dcat-us/>.

Each PI must detail:

1. The format of their data including whether the data is in a proprietary form. If so, please justify and explain what software is needed to read the data.
2. Any quality control measures the PI or original data owner utilizes on the data including verification and cleaning methods.
3. What metadata will be provided with the datasets.

3. Data Access and Sharing

All projects involving human subjects will abide by the requirements of the Institutional Review Board (IRB) of the PI's institution. Such projects must have an IRB-approved protocol for ensuring informed consent of participants and protecting privacy and confidentiality. Data should be shared only after redaction of all individual identifiers, including names, residential addresses, personal geo-coordinates, and email addresses.

Projects using proprietary data from commercial or public sources will abide by all conditions and requirements imposed on the use of the data. If the source organization, license agreement, or terms of sale prohibit the public sharing of the data, the project will be granted an exception from data sharing requirements. When working with Tribal communities' data, PIs will adhere to CARE Principles for Indigenous Data Governance (<https://www.gida-global.org/care>).

In situations where it is not feasible for Researchers to de-identify data in a manner that sufficiently protects privacy and confidentiality while preserving the dataset's usefulness, researchers will provide a description of the necessary restrictions on accessing and using the data (e.g., when working in Tribal communities where data sovereignty is necessary).

Each PI must detail:

1. Whether IRB approval is required and if so, when in the project schedule IRB approval will be obtained.
2. Whether there will be any issues (e.g., ownership or privacy issues) with sharing the data publicly and steps taken to overcome those issues.

4. Policies for Re-Use, Redistribution, and Derivatives

CPBS will upload the data to Zenodo for access and data sharing. Each PI's institution will hold the intellectual property (IP) and copyright for data and other materials created by the project, but the PI will agree with the non-exclusive licensing agreement of Zenodo. PI will be required to release their data in an open license for reuse, redistribution and derivative products which will be based upon the open licenses and provided by Zenodo. Researchers are obligated to cite the data source and specify the license under which they accessed and utilized third-party data in their project's DMP.

All inventions and patents developed with CPBS funding must be reported through iEdison electronic system. In accordance with the requirements included in the General Provisions of UTC Grants and the flow-down provisions included in the subawards issued by UNM, all subrecipients will notify UNM of any inventions and patents developed with UTC funding. UNM shall be responsible for submitting all required invention reports through iEdison.

Each PI must detail:

1. Who has the right to manage the data, including any intellectual property rights or copyrights.
2. Describe how your data will be licensed for reuse, redistribution, and derivative products.
3. Whether they anticipate any inventions or patents to be generated through their project.

Data packages deposited into Zenodo must include the following documentation files:

1. The final dataset (including preservation-friendly non-proprietary formats) needed to support and replicate the findings of the final report or paper;
2. A robust Data Dictionary defining each variable;
3. A README.txt describing the data, the methods, the other files;
4. A DCAT-US v1.1 .JSON metadata file;
5. Scripts, Code tables, or any other files used to analyze the data, so that DOT can have a complete picture of how research results were achieved, and so other researchers may replicate or reproduce the results.

5. Archiving and Preservation

Upon completion of a research project, the PI will submit all data and metadata created by the project to the CPBS Director (cpbs@unm.edu). CPBS will upload the data to Zenodo for access and data sharing. Zenodo will be the primary repository for CPBS. The data will be assigned a digital object identifier (DOI) upon uploading to Zenodo. Zenodo is an approved data repository by USDOT. PIs may opt for other repositories with the approval of the CPBS Director; other repositories must also meet federal requirements and support the creation and maintenance of persistent identifiers (e.g., DOIs, handles, etc.) and must provide for maintenance of those identifiers throughout the preservation lifecycle of the data.

All PIs will be required to provide sufficient metadata when submitting to CPBS to ensure its discoverability. Data and metadata will be retained for the lifetime of the repository, which is as long as CERN (the European Organization for Nuclear Research, the host of Zenodo) exists.