SMARTER Center Data Management Plan

Contact: Joseph Niehaus, Editor, joseph.niehaus@morgan.edu
Date Effective: June 1, 2023
Grant Period: June 1, 2023 – May 31, 2029
Grant Number: 69A3552348303

Listed below are the standards and procedures that govern data management for the Sustainable Mobility and Accessibility Regional Transportation Equity Research (SMARTER) Center. Keeping in line with the Department of Transportation guidelines for public data access, each principal investigator for SMARTER research projects will be responsible for submitting a data management plan during the Request for Proposals process and archiving any relevant data from the project in a publicly accessible data repository upon completion. Appropriate measures will be taken to ensure that any personal data collected during research will be secure.

Data Description

The SMARTER Center will collect a wide range of data from testing facilities at its associated universities. This includes but is not limited to behavioral data collected from surveys and field observations, data collected from driving simulators, emissions data collected from the field and simulators, crash statistics, demographic information, video footage, traffic operations data, and geolocation data. This data will be either collected directly by the principal investigators carrying out the projects in question (primary data collection) or aggregated and analyzed to produce new insights (secondary data).

In many cases, SMARTER research team members will analyze data using multivariate statistical analysis packages, such as R, SPSS, and SAS. We anticipate that some researchers will also take advantage of other data tabulation software as more open-source analytical tools (such as those provided by Google) become more available. All of these figures will be stored in open-access data repositories upon project completion.

The SMARTER Center will require each principal investigator to submit detailed data descriptions for their research projects as part of their data management plan. Each project's data management plan will be completed during the Request for Proposals process. As DMPs are living documents, amendments should be made as needed throughout the research project, for example, each time the project parameters or staff changes, and when the project is being finalized and published.

1. Describe 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, tables, and/or spreadsheets; 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 the transportation field and the public.

Data Standards

To ensure that data generated from SMARTER research projects can be utilized in future transportation research, principal investigators are encouraged to store data in platform-independent, non-proprietary formats. Principal investigators can notify the Center regarding the use alternative formats in their project applications during the Request for Proposals process, specifying the software and version, standards, proprietary file formats, and rationale for their use.

Data from models and simulations will be saved in batch file format, either ASCII or RTF (Rich Text Format), which can be directly read by the modeling/simulation code. Reports, papers, graphs, figures, dissertations, and other publications will be saved as Adobe PDF files. All researchers must possess final datasets that are not proprietary and are in the standard data formats commonly used in the field, such as CSV, TXT, and so on.

Finally, researchers will be responsible for data accessibility and relevant metadata. Principal investigators should be prepared to describe their project's quality control measures, tools, and software necessary to read their datasets, and any supplemental documentation needed to understand the collected data. This can be done when the data is uploaded to an open-access data repository by providing a brief description of the data collection process.

Data Access Policies

In general, data collected in SMARTER research projects must be publicly accessible (see Archiving and Preservation). Exceptions to this policy are data that contain personally identifiable information, confidential business information, or classified information.

Researchers will articulate their plans for obtaining informed consent from participants and outline the measures they will implement to ensure privacy and confidentiality before archiving the data. Additionally, they will address any additional concerns, such as embargo periods for data, and if applicable, describe how the responsibilities for safeguarding and protecting the data will be divided among project staff.

In situations where it is not feasible for Researchers to de-identify the data in a manner that sufficiently protects privacy and confidentiality while preserving the dataset's usefulness, they will provide a description of the necessary restrictions on accessing and using the data. In matters of human subject research, your DMP should describe how your informed consent forms will permit sharing with the research community and whether additional steps, such as an Institutional Review Board, may be used to protect privacy and confidentiality.

As general guidance you may consider addressing the following:

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 contains private or confidential information. If so:
    - Discuss how you will guard against disclosure of identities and/or confidential business information.
    - List what processes you will follow to provide informed consent to participants.
    - 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 deidentify your data before sharing. If not:
    - Identify what restrictions on access and use you will place on the data.
    - Discuss additional steps, if any, you will use to protect privacy and confidentiality.

Re-Use, Redistribution, and Derivative Products Policies

The National Transportation Center at Morgan State University holds the IP data for all SMARTER Projects. Researchers must explicitly state whether they are transferring the rights to the data archive. If they do not provide this information, the rights to the data will be retained by their home institution. The USDOT also reserves a royalty-free, nonexclusive and irrevocable license to reproduce, publish, or otherwise use and to authorize others to use the work for government purposes.

Additionally, if researchers are utilizing third-party data in their project, they are obligated to cite the data source (using the dataset's recommended citation) and specify the license under which they accessed and utilized the data in their project's data management plan. With respect to data ownership, researchers will address the following in their project DMPs:

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 to the data. 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

The principal investigators of each SMARTER research project will be responsible for archiving any final datasets produced during their study. Each project that receives approval will require its own data management plan in which the PIs specify where their data is stored. The PIs must provide the persistent identifier links (DOI, Handle, other persistent id) to their stored datasets to the SMARTER center upon completion of the study.

Data from SMARTER projects must be stored on a safe, open-access data repository that meets all the requirements of the DOT Public Access Plan. We suggest the following pre-approved data repositories:

1. MD-SOAR is an open-access data repository platform funded by the University System of Maryland and Affiliated Institutions (USMAI)
2. Zenodo is a free, international data repository operated by CERN. Researchers can store a variety of different data types on this site, which they get full access to using their ORCID identifier.

Researchers who wish to use an alternative data archive method must notify SMARTER that they plan to do so and explain their rationale.

Change Log

August 11th, 2023 – Added contact information and grant period information, specified USDOT rights to reproduce and publish data generated by SMARTER Center PIs