



FERSC Data Management Plan

Contact: Mingzhou Jin, jin@utk.edu

Date Effective: June 1, 2023

Grant Period: June 1, 2023 – May 31, 2029

This data management plan (DMP) governs data management for the Center for Freight Transportation for Efficient and Resilient Supply Chain (FERSC), a Tier 1 University Transportation Center (UTC) consortium led by the University of Tennessee, Knoxville. To comply with the U.S. Department of Transportation (DOT) Public Access Plan (<https://www.transportation.gov/mission/open/official-dot-public-access-plan-v11>), the FERSC UTC will require each principal investigator of each FERSC project to submit a detailed DMP during the proposal submission process, following DOT guidelines at <https://ntl.bts.gov/ntl/public-access/guidelines-evaluating-repositories>. Projects will not be funded until the DMP is received, approved by the FERSC Associate Director, and added to the project record within the UTC Projects Database. Each project is expected to archive any relevant data from the project in a publicly accessible data repository, detailed in the section of “Plans to Archive, Preserve, and Deposit the Research Data” upon completion. As living documents, project DMPs should be updated as needed, and the updated DMPs should be submitted and approved by the FERSC Associate Director, who will upload the revised DMPs into the UTC Projects Database.

Data Description

The FERSC investigators should submit detailed data descriptions for their individual research projects, as outlined in the guidance.

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, 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 the frequency of updates.
6. If using existing data, describe the relationship between the data you being collected and existing data.
7. List potential users of the data.
8. Indicate whether the data should be preserved for long-term access (e.g., longer than five years) and how long the data will be preserved.
9. Discuss the potential value of the data over the long term (e.g., longer than five years) for not only the investigator’s institution but also for the public.
10. If requesting permission not to make data publicly accessible, explain the rationale for the lack of public access. For example, there may be restrictions if the data are sensitive.
11. Indicate the party responsible for managing the data.
12. Describe how adherence to this DMP will be monitored.

Data Standards and Machine-readable Formats

To the maximum extent practicable, FERSC investigators will use DCAT-US Schema v1.1 (Project Open Data Metadata Schema) described at <https://resources.data.gov/resources/dcat-us/> to ensure maximum utility of the data in the future and share data across agents.

1. All investigators will be required to have all non-proprietary final datasets in the standard data format of the field, such as CSV and TXT.
2. If this is not possible, investigators will be required to describe how they will document the

alternative formats they are using and why the file format(s) they are using is (are) not able to be in the standard data format such as CSV.

3. If investigators are using proprietary data formats, they will be required to discuss their rationale for using those standards and formats and receive prior approval from the FERSC Associate Director.
4. Investigators will be required to describe the data process log to clarify the final version of data shared with the public.

Investigators will list what documentation they will be creating in order to make the data understandable to other researchers.

1. Investigators will indicate what metadata schema they are using to describe the data. The metadata schema may include READMEs, Data Dictionaries, and Codebooks. If the metadata schema is not one standard for their field, discuss their rationale for using that scheme.
2. Investigators will have to describe how the metadata will be managed and stored.
3. Investigators will indicate what tools or software are required to read or view the data.
4. Investigators will describe their quality control measures.

Data Access and Sharing Policies

The general policy of FERSC is that data collected in research projects must be publicly accessible. Investigators will be required to address any access restrictions and rationales for such restrictions in the project DMP that they submit to the FERSC. For project DMPs, investigators will address issues and outline the efforts they will take to provide informed consent statements to participants, the steps they will take to protect privacy and confidentiality prior to archiving their data, and any additional concerns (e.g., embargo periods for their data). If necessary, they will describe any division of responsibilities for stewarding and protecting the data among other project staff. If investigators are not able to deidentify the data in a manner that protects privacy and confidentiality while maintaining the utility of the dataset, investigators will describe the necessary restrictions on access and use. If an individual research project includes human subject research, investigators will be required to go through the University of Tennessee Institutional Research Board (IRB) or their home institutions' IRB, if they have one. Researchers working with Indigenous populations or Tribal Nations should also collect data in alignment with the CARE Principals for Indigenous Data Governance at <https://www.gida-global.org/care>.

In general, investigators will be required to address the following in their project DMPs:

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

Policies for Re-use, Redistribution, Derivatives

The University of Tennessee or the home institution of the investigators holds the intellectual property for data created by the project. Investigators must describe if they are transferring rights to the data archive. If they do not describe this, their home institution maintains the rights. Investigators will be required to cite the data source and license under which they used the data in their project DMPs.

In general, investigators 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 to be transferred to a data archive.
5. Describe how their data will be licensed for reuse, redistribution, and derivative products.

Plans to Archive, Preserve, and Deposit the Research Data

The principal investigators of each FERSC research project will be responsible for archiving any final datasets produced during their study. Each project DMP is required to specify where the data will be stored. Plans for archiving should support the capture and provision of the U.S. Federal Government Project Open Data Metadata Schema. In addition, the archive will support the creation and maintenance of persistent identifiers (e.g., DOIs, handles, etc.) and will provide for the maintenance of those identifiers throughout the preservation lifecycle of the data. Data from FERSC projects must be stored on a safe, open-access data repository that meets all the requirements of the DOT Public Access Plan. FERSC prefers the use of CERN's Zenodo at <https://zenodo.org/>, which is a free, international data repository conformant with US DOT guidelines as described at <https://ntl.bts.gov/publicaccess/evaluatingrepositories.html>.

1. Investigators are required to archive all publications and data on CERN's Zenodo, <https://zenodo.org/communities/FERSC>, which is conformant with US DOT guidelines as described at <https://ntl.bts.gov/publicaccess/evaluatingrepositories.html>.
2. When a project submits a final report, the investigators will have 60 days to archive their data on Zenodo.
3. Investigators will maintain and back-up data until it is uploaded to Zenodo.
4. Zenodo's procedures and policies for back-up, data recovery, retention, security and integrity are outlined in <https://zenodo.org/policies>.
5. Zenodo provides 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.
6. Zenodo will retain data for the lifetime of the repository. This is currently the lifetime of the host laboratory CERN, which currently has an experimental program defined for the next 20 years at least.
7. Each data upload in Zenodo gets a Digital Object Identifier (DOI) to make them easily and uniquely citable.
8. The DOI to data deposited in Zenodo, or any other repository, must be reported to USDOT and the National Transportation Library at the time final reports are submitted, or within 60 days of that submission.