Coastal Research and Education Actions for Transportation Equity
Tier-1 University Transportation Center
Center Data Management Plan

Point of Contact: Stacey Kulesza, sekulesza@txstate.edu
Effective Date: 09/01/2023
Grant Period: 06/01/2023 – 05/31/2029
Grant Number: 69A3552348330

Lead Institution
Texas State University

Partner Institutions
Oregon State University
Texas A&M University
University of Miami
University of Puerto Rico, Mayagüez

Overview

This document serves as the CREATE UTC Center Data Management Plan (DMP); it describes our policies for handling data during and after our research program's grant period. CREATE is a Tier-1 UTC funded by the USDOT that focuses on statutory research priority area D, Improving the durability and extending the life of transportation infrastructure. We view this DMP as a living document that will be updated as our project evolves. Updates to this Center DMP will be distributed to all CREATE researchers and to USDOT.

In this document, the term Faculty refers to individual project Principal Investigators. The Center refers to the CREATE UTC. Faculty will follow the guidelines and policies in this Center DMP. All Faculty must submit individual project DMPs following the guidelines and format of this Center DMP, and then provide details that are unique to their specific research project. Project DMPs will be reviewed and approved by CREATE program leadership, and made available to USDOT grant managers. Each Project DMPs is also a living research team knowledge management document that will be reviewed and updated 1) as the research project, research personnel, or data preservation infrastructure changes, and/or, 2) at each project review stage (at least twice per year). Changes to Project DMPs will be forwarded to USDOT. Project DMPs will be stored in the shared Center location.

Faculty seeking more guidance on the US DOT Public Access Plan can find guidance at https://doi.org/10.21949/1520571

Data Description

CREATE UTC research projects will focus on four research thrusts: transformational coastal infrastructure design and construction; coastal transportation infrastructure evaluation, prediction, and prevention; equitable response to unprecedented hazards; and pathways to blue economy transportation careers. Research project approaches will include computational methods,

experimental (laboratory and field) methods, and research involving human subjects. The four research thrusts will address challenges of coastal transportation infrastructure durability.

Initial Projects DMPs for CREATE UTC research should be submitted at the time of proposal and should answer each prompt below. Responding to these prompts help make data and data handling clearer and more understandable to the research team, USDOT, and other researchers.

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).
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).
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 Faculty 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 the Faculty's academic institution, but also for the public.
9. If Faculty 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 Faculty will check for adherence to this data management plan.

Data format and metadata standards

We expect that data will be collected and generated in a range of formats (e.g., .ascii, .mov, .xlsx). To ensure open and accessible data, Faculty will preserve and submit final datasets in open, non-propriety formats, such as .csv or .txt, when possible. Datasets may also be preserved in industry standard formats, such as shapefiles for GPS data. Preserving data using open formats will allow data to be utilized in future transportation research.

1. Faculty will list the types of data they expect will be collected, used, or re-used for this research project. Discuss which file formats the data will be collected in or being used.
2. If Faculty are using proprietary data formats, they will be required to discuss the rationale for using those standards and formats and why the file format(s) they are using is(are) not able to be in non-propriety formats.
3. If the file format(s) is(are) not standard to the their field, Faculty will describe how they will document the alternative formats they are using and why.
4. Faculty will describe their data process log to clarify the final shared version to the public.
5. Faculty will list what documentation they will be creating in order to make the data understandable by other researchers. Data documentation package elements include: the DMP; a data dictionary; a README.txt that describes the data and methodology; a DCAT-US .json metadata file; source code used and formulas used for data analysis; any other supporting tables or files needed to contextualize the data.

6. Faculty will indicate what metadata schema they are using to describe the data. If the metadata schema is not one standard for the field, discuss the rationale for using that scheme. The minimum metadata is author, author identifier (such as an ORCID), contributors and their identifiers (such as ORCID), title, abstract, and keywords.
7. Faculty will describe how the metadata will be managed and stored.
8. Faculty will indicate what tools or software is required to read or view the data, especially if proprietary software is used and proprietary data formats are generated.
9. Faculty will describe quality control measures.

Access and sharing policies

Faculty will be required to address any access restrictions that may apply to their data in their project DMP they submit to the Center. Faculty 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., data embargo periods) in their Project DMP. If necessary, they will describe any division of responsibilities for stewarding and protecting the data among other project staff. If Faculty will not be able to deidentify the data in a manner that protects privacy and confidentiality while maintaining the utility of the dataset, Faculty will describe the necessary restrictions on access and use. If an individual research project includes human subject research, Faculty will be required to go through Texas State University's Institutional Review Board (IRB) or their academic institution's IRB approval process, if they have one. Any research involving First Nations or tribal members should adhere to the CARE Principles of Indigenous Data Governance https://www.gida-global.org/care

Restrictions on data sharing should be identified in the Project DMP. Should Faculty believe that they are unable to share any of the data, submit a justification for consideration by USDOT program staff. USDOT program staff may wish to have a discussion about creating a sharable subset of data or how to allow restricted access. Refer to USDOT Q&A at
 https://doi.org/10.21949/1520567

Faculty will be required to address the following in their Project DMP when they submit to the Center:

1. Describe what data will be publicly shared, how and at which repository the data files will be shared, and how others will access them.
2. Indicate whether the data contain private or confidential information. If so:
   o Discuss how they will guard against disclosure of identities and/or confidential business information.
   o List what processes they will follow to provide informed consent to participants.
   o 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 data will be deidentified before sharing. If not:
   o Identify what restrictions on access and use they will place on the data.
   o Discuss additional steps, if any they will use to protect privacy and confidentiality.

Re-use, redistribution, and derivative products policies

Faculty will be required to submit a copy of all final data created by the project to the Center to be deposited to Zenodo for access and data sharing (see plans for archiving and preservation

below). Texas State University or the home institution of the Faculty will hold the intellectual property rights for the data and other materials created by their project. 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.

Faculty will be required to submit the recommended citation for the data source and the data license under which they used the data of others in their project DMPs. Faculty 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 the archive.

The DMP should specifically:
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 plans

Preserving research data is an important aspect of scientific replicability, reproducibility, and open science. All final datasets that support research conclusions, or do not support the original research hypothesis, must be archived in an appropriate repository. Faculty or the project data management staff will be responsible for ensuring data submission to an appropriate repository.

1. The CREATE Center will archive all data on Zenodo, https://zenodo.org/, which is conformant with the USDOT Public Access Plan. Faculty may also use their institutional data archives such as the Texas Data Repository (TDR) for Texas State University and Texas A&M University.
2. When Faculty submit a final report, they must also provide the Zenodo link to dataset(s).
3. Faculty will maintain their data until it is uploaded to Zenodo.
4. Faculty will 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.
5. Faculty will describe how data will be protected from accidental or malicious modification or deletion prior to receipt by the archive.
6. 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 for the long-term.
7. Zenodo uses CERN's Data Centre and will be available for as long as CERN exists. CERN currently has research scheduled through 2041.
8. Every Zenodo upload is assigned a Digital Objective Identifier (DOI), to make them citable and trackable.

References

United States. Department of Transportation. (2022) *DOT Public Access* [Home page]. https://doi.org/10.21949/1503647

United States. Department of Transportation. (2022). *Creating Data Management Plans for Extramural Research*. https://doi.org/10.21949/1520571