Mobility21 University Transportation Center
Data Management Plan (DMP) Requirements for Research Projects

- *The use of the word Faculty in this document shall mean the project Principal Investigator.*

The Mobilty21 UTC requires each Faculty to submit a detailed project data management plan for individual research projects per this center-wide plan as outlined below and the US Department of Transportation Guidance.*

Data management plans should contain each of the sections described below. Each section should contain a narrative addressing the section prompts, as applicable. The DMP may be brief, but should be as detailed and robust as needed. The DMP will also serve as a useful knowledge management document for the research team.

Faculty must submit the project data management plan for review to the lead for their institution for review and approval before any work can begin on the project.[1] The lead will review and approve DMPs within 5 business days of receipt. The approved DMP will then be uploaded into the relevant project record in the Mobility21 UTC Research Database.

As DMPs are living documents, the plans should be updated as needed to reflect changes in the research project, and an updated DMP provided to the lead Principal Investigator. The lead Principal Investigator will again review the DMP and upload the revised DMP into the relevant project record in the Mobility21 UTC Research Database within 5 business days.

Projects will not be funded until the DMP is received, approved, and added to the project record within the Mobility21 UTC Research Database.

Data description

1. Name the data, data collection process, or data producing program. A template will be provided for the information to be entered to satisfy both repository and Center reporting requirements.

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 being collected and existing data.

7. List potential users and re-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 requesting 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 adherence to this data management plan will be monitored.

Data format and metadata standards

1. For preservation purposes, the Center will require all Faculty to submit links to final datasets. Datasets should be available in open, non-proprietary formats, such as .csv or .txt, when possible. Or datasets may be preserved in industry standard formats, such as shapefiles for GPS data.

2. If Faculty are using proprietary data formats, they will be required to discuss the rationale and provide open data source if possible. They will be required to discuss rationale and software to read the data format.

3. Faculty be required to describe the data process log to clarify the final version of data shared to the public. Refer to the "README FILE" for additional guidance.

4. Faculty will be required to describe how they will document the alternative formats they are using and why.

5. Faculty will list what documentation they will be creating in order to make the data understandable by other researchers. Faculty will use standard document template to create a data dictionary in order to make the data understandable by other researchers.

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, contributors, title, abstract, keywords. For reporting purposes, researchers will also be required to supply. Name the data, data collection project, or data producing program. A template will be provided for the information to be entered to satisfy both repository and Center reporting requirements

7. Faculty will have to describe how the metadata be managed and stored.

8. Faculty will indicate what tools or software is required to read or view the data.

9. Faculty will describe the quality control measures.

Policies for access and sharing

Faculty will address issues and outline the efforts they will take to provide informed consent statements to participants, the steps they will take the protect privacy and confidentiality prior to archiving the data, and any additional concerns (e.g., embargo periods for data). 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 de-identify or anonymize 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 Carnegie Mellon University's IRB or the academic partner's IRB approval process, if they have one.

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

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
   1. Discuss how disclosure of identities and/or confidential business information will be guarded against.
   2. State the party responsible for protecting the data.
   3. List what processes will be followed to provide informed consent to participants.

3. Describe what, if any, privacy, ethical, or confidentiality concerns are raised due to data sharing.

4. If applicable, describe how data will be de-identified before sharing. If not:
   - a. Identify what restrictions on access and use will be placed on the data.
   - b. Discuss additional steps, if any, will be used to protect privacy and confidentiality.

Policies for re-use, redistribution, derivatives

Carnegie Mellon University or the home institution of the Faculty or staff will hold the IP and copyright for data and other materials created by the project. The USDOT hereby 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.[2]

Faculty will be required to cite the data source and license under which they used the data of others in the project DMPs. Faculty will be required to release the data in an open license for reuse, redistribution and derivative products which will be based upon the open licenses and provided by the data archive.

Note the Copyrights per the GENERAL PROVISIONS OF GRANTS FOR 2016, 2018 and 2019, UNIVERSITY TRANSPORTATION CENTERS (copied below) (https://www.transportation.gov/sites/dot.gov/files/docs/utc/350636/fast-act-generalprovisions-nov-2016-rev-july-2019-ver-14clean.pdf) *or the prevailing version of the General Provisions of Grants for UTCs document.*

b) Copyrights

(1) The author or the Grantee's organization may copyright any books, publications, or other copyrightable materials developed in the course of or under this Grant, but USDOT hereby 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.

(2) The Grantee must not incorporate material copyrighted by others into any work product delivered under this Grant unless it has acquired for USDOT a royalty-free, nonexclusive and irrevocable license to reproduce, publish, or otherwise use and to authorize others to use the work for government purposes.

(3) The Grantee may arrange for publication of initial reports of original research, supported in whole or in part by USDOT funds, in primary scientific journals and copyright by the journal unless the journal's copyright policy would preclude an individual from making or having made by any means available, without regard to the copyright of the journal and without royalty, a single copy of any such article for the individual's own use.

(4) The Grantee must be responsible for any losses that result from or arise out of the negligent use of or breach of provisions by its employees or agents under this Grant regarding the publication, translation, reproduction, delivery, use, or disposition of any data or protected privacy information furnished under this Grant provided that this provision must not be deemed a waiver by the Grantee of any immunities to which it may be entitled under applicable Federal, State, or Tribal law.

Plans for archiving and preservation

1. The Faculty can archive data on Zenodo, zenodo.org, and KiltHub, kilthub.cmu.edu, the Carnegie Mellon University data repository which are known to be conformant with the USDOT Public Access Plan. Faculty may also choose another publicly available platform as outlined in the DMP.

2. When Faculty submits a project final report, the Faculty must also provide the link to the data set(s).

3. Faculty will maintain the data until it is uploaded to the platform outlined in the DMP.

4. 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. Describe how data will be protected from accidental or malicious modification or deletion prior to receipt by the archive.

6. Discuss the 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. Include a link to the archive's online documentation.

7. Indicate how long the chosen archive will retain the data.

8. Indicate if the chosen archive employs, or allows for the recording of, persistent identifiers linked to the data. Include a link to the archive's online documentation.

9. Discuss how the chosen data repository meets the criteria outlined on the **Guidelines for Evaluating Repositories for Conformance with the DOT Public Access Plan** page (https://ntl.bts.gov/public-access/guidelines-evaluating-repositories)**.**

---

[1] Contacts for Each Academic Partner: Carnegie Mellon University, Lisa Kay Schweyer, University of Pennsylvania, Rahul Mangharam, and The Ohio State University, Umit Ozguner.
[2] https://ntl.bts.gov/ntl/public-access/creating-data-management-plans-extramural-research