SOUTHERN PLAINS
TRANSPORTATION CENTER

The Southern Plains Transportation Center (SPTC)

UTC Region 6

SPTC Data Management Plan

Contact: Musharraf Zaman

Phone: 405-325-4682

Email: Zaman@ou.edu

Date Effective: June 1, 2023

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

Grant No. 69A3552348306

Lead Institution: The University of Oklahoma

Partner Institutions:

University of Texas at El Paso

University of New Mexico

University of Arkansas

Texas Southern University

Texas A&M University Transportation Institute

Oklahoma State University

Navajo Technical University

Louisiana Tech University

Louisiana State University

El Paso Community College

1. Introduction:

The Southern Plains Transportation Center (SPTC) recognizes the importance of effective data management to support its research, education, and outreach activities. This Data Management Plan (DMP) outlines the strategies, policies, and procedures that the SPTC will employ to ensure the proper collection, storage, analysis, and sharing of data generated from projects that are awarded through the center. This DMP will serve as a guide for all SPTC projects.

2. Data Description:

SPTC may handle various types of data, including but not limited to:

a. Research Data: Collected through surveys, experiments, simulations, and field observations.

b. Geospatial Data: Maps, GPS coordinates, spatial datasets, and aerial imagery.

c. Traffic and Transportation Data: Traffic flow, congestion, transportation modes, infrastructure, and related datasets.

d. Administrative Data: Budgetary, personnel, and project management data.

e. Education Data: Course materials, assessments, and student information.

f. Outreach Data: Public engagement records, stakeholder feedback, and communication materials.

SPTC will require each PI to submit detailed data descriptions for their research projects that will include:

- Name of the data, data collection project, or data producing program.
- Description of the purpose of the research.
- Description of 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.).
- Description of the 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.)
- The period of time data will be collected and frequency of update.

3. Data Standards and Formats

Formats for SPTC data will include tabular data formats (e.g., CSV, Excel), standard formats for textual data (e.g., DOC, TXT, XML, JSON), image and video formats (e.g., JPEG, PNG, MP4, PPT), and audio formats (e.g., WAV, MP3). SPTC will ensure that data will be provided in an open and accessible format and include applicable software and versioning documentation in the event that a proprietary format

is used. Researchers will ensure that whenever possible, non-proprietary formats will be used. If proprietary formats are used, the PI will discuss the rationale and explain the software and version needed for a user to open and view the data files. SPTC will require that projects create a DCAT-US v.1.1 metadata file (https://resources.data.gov/resources/dcat-us/) for all funded projects.

4. Data Life Cycle, Preservation, and Archiving:

The SPTC will follow a comprehensive data life cycle management approach to ensure data integrity, accessibility, and usability. The PI should include the following elements in their DMP to achieve this objective:

a. Data Collection: Establish protocols and standardized data collection instruments including identifying research objectives, determining data variables, identifying sampling methods, data collection methods, data collection procedures, and data quality assurance. Continuous monitoring of the data collection process to identify and address any issues or deviations from protocols will occur, and version control procedures will track any updates or revisions made to protocols over time.

b. Data Organization: Researchers will implement a structured data organization system, including metadata. See Metadata elements in Appendix A. Metadata will be created at the time of data acquisition and will be updated as needed by database managers. Metadata will be in compliance with DCAT-US Metadata schema, found here.

c. Data Storage and Preservation: When a researcher submits a final report, they will have 45 days to archive their data with SPTC. Researchers will submit their data to SPTC for storage (in addition to their own organizational storage locations if needed). SPTC will use its institutional repository with the University of Oklahoma for storage and will maintain backups for retrieval in case of catastrophic events. This will involve the use of cloud storage solutions, on-premises storage infrastructure such as storage area networks, scheduled backups of data to external hard drives and other storage locations and implementing data redundancy and replication to ensure data is on multiple devices and locations. In addition to using their own repositories at their institutions, data will be stored on the SPTC shared drive, a cloud-based storage system maintained by the University of Oklahoma.

d. Data Documentation: Researchers will provide documentation detailing data processing methods, assumptions, and limitations. This will involve providing an overview of documentation, explaining its purpose and scope, describing the methods used to collect and acquire the data (including sources, instruments, and sampling techniques), detail specific methods used to process and analyze the data, document any assumptions made during the data processing and analysis, and identify the limitations of the data processing methods used. An easily identifiable file naming convention will be used, using names such as "README," "DMP," "codes/scripts," and others.

e. Data Sharing: SPTC will encourage open data practices using appropriate data sharing mechanisms while respecting privacy and confidentiality. Mechanisms include posting (when legally allowed) to the SPTC SharePoint site for staff and PI's, and to other repositories such as the National Transportation Library, TRB RiP, FHWA, and others. Data will also be shared during presentations, workshops, seminars, webinars, and other outreach activities related to presentation of research information. Data will be marked with a recommended citation and persistent identifier (PID) to track data sharing, including email address and contact information for individual/s involved in creating and sharing the data.

5. Data Access and Protecting Sensitive Data:
SPTC will ensure that researchers will execute a plan to preserve data access while protecting privacy and confidentiality when needed. To ensure data security and privacy, the SPTC will implement the following measures:

a. Access Control: Restrict data access based on role-based permissions, ensuring authorized personnel are the only ones able to edit or view restricted data.

b. Data Anonymization: Remove personally identifiable information (PII) and sensitive data whenever possible as needed.

c. Data Encryption: Apply encryption techniques to protect data during transmission and storage as needed. Additionally, university partners in SPTC will use their IRB boards to review projects involving people.

d. Ethical Considerations: When working with or conducting research that includes Indigenous populations or Tribal communities, SPTC researchers will adhere to the CARE Principles for Indigenous Data Governance: https://www.gida-global.org/care.

e. Data Retention: Establish data retention policies compliant with relevant legal and regulatory requirements as needed.

f. Data Handling Training: Provide training on data security, privacy, and ethical considerations to personnel involved in data management as needed.

g. Data Destruction: When restricted data is found on old hard drives, disks, or electronics that are due to be taken out of service, SPTC will implement data erasure services to ensure no data or information will be compromised.

6. Data Sharing and Re-use, and Redistribution:
The SPTC will actively promote data sharing and open science practices, striving to maximize the impact of its research. The following steps will be taken:

a. Data Publication: Publish research data in open repositories or institutional data repositories with appropriate licenses and embargoes. These will include but are not limited to the SPTC website (http://www.sptc.org/),the National Transportation Library's Repository & Open Science Access Portal (ROSA P) (https://rosap.ntl.bts.gov/),

Transportation Research Board's Research in Progress (RiP) Database (https://rip.trb.org/), Transportation Research International Documentation (https:/trid.trb.org/), and USDOT Research Hub ((https:/researchhub.bts.gov/search).

b. Data Access: Facilitate data access through standardized formats and clear access procedures.

c. Data Citations: Encourage proper citation of datasets to acknowledge the contributors and promote reproducibility.

d. Collaboration: Foster collaborations with all SPTC partner institutions, stakeholders, and other interested institutions to enhance data sharing and collective knowledge.

Additionally, 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.

7. Data Management Roles and Responsibilities:
The SPTC will assign the following roles and responsibilities to ensure effective data management:

a. Principal Investigators (PIs): Oversee data management activities within their respective projects, including data quality assurance.

b. Data Manager: Facilitate data organization, documentation, and sharing, ensuring adherence to data management policies.

c. IT Support: Provide technical assistance in data storage, security, backup, and recovery.

d. Compliance Officer: Ensure compliance with data protection regulations and policies.

e. Center Director: Center Director will submit data management plan to USDOT within 90 days of receiving funding. If there are any changes to the data management plan, the Center Director will notify the grant manager. The Center Director must ensure that each project is in compliance with US DOT public access plan and data management plans.

8. Data Management Plan Review:
Given that DMP's change over time and based on various project levels the SPTC will periodically review and update this DMP to align with evolving technological advancements, regulatory changes, and best practices in data management.

By implementing this Data Management Plan, the SPTC aims to ensure the efficient and responsible handling of data, facilitating high-quality research, collaboration, and societal impact.

Change Log:

09-20-2023: Original draft
10-26-2023: Revisions and additions based on feedback
01-24-2024: Revisions made by NTL staff.

Appendix A: Metadata Elements

The SPTC requires that the following information is included in each project's metadata file. We require that projects create a DCAT US metadata file (https://resources.data.gov/resources/dcat-us/), but if another schema is desired, the project team can create an additional file. This additional file must include the following:

a. Dataset Title: A descriptive title for the dataset.

b. Description: A summary of the dataset's content, purpose, and scope.

c. Creator/Contributor: The individuals or organizations responsible for creating or contributing to the dataset.

d. Date Created/Modified: The dates when the dataset was created or last modified.

e. Data Collection Methods: Information about the methods, instruments, and protocols used to collect the data.

f. Variables/Fields: A list of variables or fields present in the dataset, along with their definitions, units, and allowed values.

g. Data Format: The file format or structure of the dataset.