

Center for Social and Economic Mobility for People And Communities through Transportation (SEMPACT)

A USDOT Region 2 University Transportation Center under BIL

Center Data Management Plan

Version 1.1

Effective Date: 9/1/2023

Grant Period: 2023 – 2028

Grant Number: 69A3552348302

Lead Institution: **The City College of the City University of New York**

Partner Institutions:

CUNY Bronx Community College
New Jersey Institute of Technology
New York University
Princeton University
Rensselaer Polytechnic Institute

Rutgers University
SUNY University at Albany
SUNY Polytechnic Institute
SUNY Stony Brook University
University of Puerto Rico Mayaguez

Contacts:

Dr. Camille Kamga
SEMPACT Director
ckamga@utrc2.org

Jerry He
Assistant Director, Research
jhe@utrc2.org

1. About SEMPACT

The “Center for Social and Economic Mobility for People And Communities through Transportation” (**SEMPACT**) is a USDOT Region 2 University Transportation Research Center, covering regions of New York, New Jersey, Puerto Rico and US Virgin Islands.

Led by University Transportation Research Center (UTRC) at the CUNY City College of New York, and building from the foundation of our existing legacy transportation systems, **SEMPACT** seeks to leverage recent advances in vehicle and infrastructure technologies; data; and modeling, analytics, and visualization to plan, design, implement, and operate a more equitable, resilient, and sustainable transportation system for Region 2 that supports mobility and access to opportunity for all the region’s residents and businesses. To advance toward these goals, the **SEMPACT** team will focus research in four priority areas that address both USDOT & regional needs. Work will address the statutory priority area **Improving mobility of people and goods** and the USDOT research priorities Economic Strength & Global Competitiveness (primary), Equity, and Climate & Sustainability.

SEMPACT prioritize research of:

- (1) accessible tools for informed transportation decision making
- (2) inclusive advanced technology applications
- (3) climate resilient infrastructure
- (4) green and resilient supply chains

This document serves as the **SEMPACT** Center Data Management Plan (“Center DMP”). Researchers should reference the Center DMP to reduce the amount of copy-and-pasted language in their project-specific DMPs.

The Project DMP should describe the unique and specific aspects of the research project for each section outlined in the Center DMP. Project DMPs are dynamic knowledge management tools that should be reviewed and updated regularly, especially when there is a significant change in the research project, the data collected, or the project personnel.

2. Data Description, Activities & Sources

SEMPACT expects to collect or create different types of data, such as trip Origin-Destination data, geographic information system (GIS) data, freight data, demographics data, depending on the activities it funds. This data can be about transportation systems infrastructure & users, or the environment, from both open-source, public sector, and private companies. The data is collected or created through:

2.1 Research Projects

Research projects, funded by **SEMPACT** and led by Principal Investigators from each institution, generate a variety of scientific outputs, including:

- Publications: working papers, project reports, and open access articles
- Code that simulate, modify, or analyze project-specific data
- Data generated from experiments, surveys & interviews
- Sensitive, confidential or protected data shared by public agencies, private companies, or third parties

2.2 Educational Activities & Career Training Initiatives

These educational activities are funded by **SEMPACT** to provide overall Center priority to “improve the mobility of people and goods”, with a focus on STEM education, job readiness, university education, and professional development. Data include but not limited to:

- Project final presentation and reports
- Video/webcasts/recordings of webinars, presentations, and lectures
- Course outlines and curriculums

2.3 Knowledge Transfer and Outreach Activities

SEMPACT conducts knowledge transfer activities through symposia, conferences, community engagement to implement assistance to stakeholder groups, outreach underrepresented communities, and disseminate and report research results. Data collected include but not limited to:

- Slides and recordings of presentations/seminars/invited talks
- Conference proceedings
- De-identified attendance summary

3. Data format and standards

Principal investigators are encouraged to use accessible, interoperable, open, and non-proprietary data formats when possible, to ensure that their data is accessible and reusable to future users:

- The PI shall describe how the data will be managed and stored, and indicate what tools or software is required to read, view, or modify the data.
- If a PI chooses to use proprietary data formats (e.g. certain industry standard data formats), they should explicitly document why decisions were made to use proprietary formats. They should also explicitly document procedures the public can take to access, use, and modify the chosen proprietary data formats.

USDOT views data as first-class research objects, on par with research papers. Data should have their own searchable and citable repository pages with metadata to make it easy for future researchers to reuse data and credit the original researchers:

- The PI shall describe the metadata. [Data Catalog Vocabulary \(DCAT-US\) metadata schema v1.1](#) is a recommended option to use. At the minimum, metadata should include author, contributors, title, abstract, and keywords.
- For preservation purposes, **SEMPACT** will require all PIs to submit links to final datasets, to be attached when project final reports are due.

4. Data Access and Protecting Sensitive Data

All research data and results shall be made publicly accessible, except for data that contain Personally Identifiable Information (“PII”), confidential business information, or classified information. For research projects that involve PII and/or confidential information, a de-identified dataset that supports research conclusions shall be released, if at all safely possible, rather than no dataset at all.

The level of sharing will depend on the nature of the data and project. Principal investigators shall indicate in their project proposals if specific sets of research data are not accessible to the public and explain why.

When working with human subjects, researchers shall follow their Institutional Review Board (“IRB”) policies and seek IRB approval before beginning the study. DMPs involving human subjects should document the consent process, such as the steps that will be taken to obtain informed consent from participants. Proper documents must be prepared to address privacy and confidentiality concerns.

Further, when working with, or conducting research that includes Indigenous populations or Tribal communities, **SEMPACT** researchers will adhere to the [CARE Principles for Indigenous Data Governance](#)¹.

5. Data Sharing, Re-use, and Redistribution

All **SEMPACT** projects are “Public Access” as defined by the Official [USDOT Public Access Plan v1.1](#)². The PIs shall make the public aware of research outputs, including data, and sharing data that supports research conclusions whenever possible, except to protect privacy and security.

Upon project initiation, the PIs shall:

- Describe what data and to what extent will be shared, how data files will be shared, and how other researchers and the public will access them.
- Identify whether the data contain private or confidential information.
- Identify if the public needs proprietary software to access data.
- If applicable, describe whether and how data will be de-identified before sharing.
- Identify the source, copyright, and attribution of data, for the redistribution of future users.

Note: The USDOT also reserves a royalty-free, non-exclusive and irrevocable license to reproduce, publish, or otherwise use and to authorize others to use the work for government purposes.

6. Data Preservation and Archiving

Upon project completion, PIs shall:

- **Store data in a publicly accessible place.** This could be Zenodo (www.zenodo.org) or another platform that is approved by the PI’s institution. The dataset should also be given a Persistent Identifier.

¹ <https://www.gida-global.org/care>

² <https://doi.org/10.21949/1520559>

- **Include a link to the data set(s) in the final project report.** This will make it easy for others to find the data and reuse it in their own research.
- **Keep the data safe until it has been uploaded to the archiving platform.** This means backing it up and storing it in a secure location.
- **Explain how the data will be protected from being accidentally or maliciously changed or deleted before it is archived.** This could involve using encryption, password protection, or other security measures.
- **Learn about the data archiving platform's policies and practices for long-term data preservation.** This will help to ensure that the data will be safe and accessible for many years to come.

7. Change Log

2023-10-03: Updated draft based on USDOT suggestions

2023-09-30: Original draft