



Mineta Consortium for Equitable, Efficient, and Sustainable Transportation (MCEEST)

Master Data Management Plan

Version 1.0
(August 23, 2023)

Lead Institution

San José State University

Director: Dr. Karen Philbrick, karen.philbrick@sjsu.edu

Partner Institutions

Howard University

Navajo Technical University

University of South Florida



About MCEEST

The Mineta Consortium for Equitable, Efficient, and Sustainable Transportation is a USDOT Tier-1 University Transportation Center led by the Mineta Transportation Institute (MTI) at San José State University. Consortium partners include: Howard University, a Historical Black College & University (HBCU) located in Washington, D.C.; Navajo Technical University, a Tribal College & University (TCU), located in Crownpoint, New Mexico; and the University of South Florida, located in Tampa, Florida.

MCEEST conducts research, education, workforce development, and technology transfer activities to improve mobility of people and goods. MCEEST's work advances equitable, efficient, and sustainable transportation so people in every community can connect to the people, places, and opportunities that make their lives meaningful.

MCEEST's research agenda places equity, safety, and sustainability at the forefront and will proceed with a primary focus on advancing the USDOT Strategic Plan goals of Equity, with Climate and Sustainability and Economic Strength and Global Competitiveness as secondary strategic goals. Research projects focus on efforts to expand access and broaden participation in our nation's transportation systems and workforce by improving understanding and facilitating connections, especially with traditionally underserved communities. Projects will use data and public opinion to inform policy, infrastructure, and technology benefiting diverse communities, including bicyclists and pedestrians, urban and rural regions, women and gender-nonconforming people, and others, to address equity and barriers to opportunity.

As Consortium lead, MTI will ensure that data of all types will be managed and organized for security, consistency, and public dissemination in compliance with the U.S. Department of Transportation's Publication Access Plan. MCEEST researchers will follow the guidelines and policies in this Master Data Management Plan (DMP).

1. Data Description

The Mineta Consortium for Equitable, Efficient, and Sustainable Transportation anticipates a variety of data to be gathered or generated as a part of program activities. MCEEST's research will rely upon a wide range of transportation data including existing data sets, collection of new data, as well as simulations.

Data developed as part of MCEEST's research activities will include the following:

- Surveys of individuals, agencies and/or organizations;
- Geographic/spatial data, including geographic information systems (GIS) data;
- Data developed as a result of laboratory/fields experiments. These data may include, for example: photos, videos, spreadsheets of experiment/laboratory results, etc.



- Simulated data produced through simulation tools.
- Source code for development algorithms, simulations, and other tools/products
- Demographic, socioeconomic, and census data.
- Data describing transportation network/supply characteristics.
- Travel survey data.
- Transit passenger data.

All data collection efforts involving human subjects will follow each university's Institutional Review Board (IRB) policies. San José State University information can be found at: <https://www.sjsu.edu/research/research-compliance/irb/index.php>.

Data will be collected throughout the duration of the five-year grant. The long-term value of the data will be to provide transportation decision-makers, the general public, and others interested in transportation research the ability to replicate research findings or utilize the data (with attribution) in subsequent studies.

Researchers funded through this grant will be required to indicate compliance through the creation of the data management plan prior to approval of any individual research project, and update the DMP as needed during the research project. Researchers funded through the grant will be expected to manage the data according to any approved IRB protocol. In limited cases, it is possible that data providers or project partners may place legal restrictions on openly sharing data. In these cases, principal investigators (PIs) will provide the details and rationale for limited public access.

2. Data Formats and Standards

MCEEST's policy is to use data in open, standard formats and to provide appropriate metadata and data dictionaries so that others may access and understand the data.

Data from MCEEST activities will be primarily collected and generated in a range of formats, including but not limited to:

- .docx, .xlsx, .pptx
- Xml
- Pdf/a
- Ascii
- .txt
- .csv
- .jpg, .tiff, .npg
- .avl
- .gps
- .mxd, .yr, .gdp, .shp, .dbf
- .html
- .mov, .mpeg, .mp4



While every effort will be made to archive data in open formats, proprietary formats will be utilized when those are the best way to present the data to retain the information or when the proprietary format (e.g. .doc) is commonplace. In the event that proprietary formats are used the specific software, version, and other relevant details that a user might need to view the data will be included.

Data will be anonymized and stripped of personally identifiable information in accordance with San José State's University policies, specifically the Policy for Protection of Human Research Subjects (<http://www.sjsu.edu/senate/docs/S08-7.pdf>). SJSU's Institutional Review Board has a federal-wide assurance through the U.S. Department of Health and Human services and complies with all federal regulations for research involving human subjects.

Metadata for each dataset will be provided following the standards set forth in the DCAT-US Schema v1.1 (Project Open Data Metadata Schema) (<https://resources.data.gov/resources/dcat-us/>). Criteria included will focus on authorship, subject, scope, as well as spatial or temporal extent of the data. Metadata about the collected project data will be represented in either a JSON file or as a README.TXT file. The JSON or README.TXT file will accompany the dataset to provide maximum context.

Data, including all associated metadata, will be stored in a publicly accessible archive (San José State University's ScholarWorks repository: <http://scholarworks.sjsu.edu/>) and users will be able to download the data to their own computers for use. Direct viewing of the data within the repository will not be available.

All funded research will undergo peer review.

3. Policies for Access and Sharing

MCEEST's policy is to openly share data generated/used in research and activities funded by the U.S. Department of Transportation and our partners, subject to an confidentiality, privacy, security or human subjects restrictions.

All MCEEST research that involves human subjects will be required to obtain approval from San José State University's Institutional Review Board (or equivalent entity from partner institutions). The University's Institutional Review Board ensures that the research is in compliance with university policy and federal regulations established to ensure the safety of research participants and the ethical and responsible conduct of investigators. Research that involves human subjects and collects data with any personally identifiable information must document how individual privacy is protected. Researchers are expected to use protocols such as pseudonyms or coding systems to maintain confidentiality. Researchers must specify in their IRB application how data is collected and kept safe (e.g. encrypted files, password-protected computer, locked cabinet, etc.), who has access to data, and a retention plan for the data. In addition, researchers will be expected to document how that data will protect privacy and confidentiality while maintaining the use of the dataset for final distribution.



Policies for re-use, redistribution, derivatives

Intellectual property rights for the data will comply with San José State policy (<http://www.sjsu.edu/senate/docs/F98-3.pdf>). As such, ownership of copyrightable works are owned by the authors (e.g. SJSU faculty, staff, or students). Data produced under the auspices of this grant and subject to the requirement to post the data to a publicly accessible repository will utilize a CC-BY-NC license. This Creative Commons license requires attribution to the original author and allows the distribution, remix, re-use, and derivative works, as long as it is not for commercial purposes. The CC-BY-NC license will be noted in the record of the data posted to the online repository along with contact information for the corresponding author who can field inquiries regarding the data and its ownership.

Plans for archiving and preservation

MCTM data will be archived through San José State University's ScholarWorks repository. ScholarWorks is a data repository conformant with the U.S. Department of Transportation's Public Access Plan. SJSU ScholarWorks provides reliable, secure, and sustainable access to the data.

Prior to submission to ScholarWorks, all data will be stored by individual principal investigators in a manner compliant with the approved IRB (for human subjects research). For submission to ScholarWorks, the data will be described and identified according to this data management plan.

- The University Library is committed to archiving and preserving SJSU community materials through SJSU ScholarWorks. There may be instances where the University Library may convert submissions to other mediums or formats for the purpose of preservation, or keep more than one copy for purposes of security, back-up, and preservation. In all cases, content will not be altered.
- In the event of a platform change, the University Library is committed to securely and reliably migrating records and metadata.
- Every entry in SJSU ScholarWorks is provided a permalink or permanent link to ensure reliable access to content. The University Library has further enhanced selective records with digital object identifiers (DOI), enabling interoperability with CrossRef and its database of scholarly citations. All MCEEST reports and data products will be assigned a DOI to facilitate preservation and citation networking.
- Digital Commons is hosted by Amazon Web Services and maintained in a high availability, redundant configuration across multiple Availability Zones. Full nightly backups of SJSU ScholarWorks are performed to ensure the integrity of the content. Further information on technical specifications is available from bepress: https://bepress.com/reference_guide_dc/safeguarding-content-digital-commons/.