

CENTER FOR INTEGRATED ASSET MANAGEMENT FOR MULTI-MODAL TRANSPORTATION INFRASTRUCTURE SYSTEMS (CIAMTIS)



U.S. DOT Region 3 University Transportation Center

Data Management Plan

BACKGROUND

The Center for Integrated Asset Management for Multi-modal Transportation Infrastructure Systems (CIAMTIS) will work collaboratively with their university partners and regional transportation stakeholders to improve the ability of agencies to deliver safe and cost-effective transportation infrastructure by targeting high-return aspects of infrastructure asset management. Enhanced infrastructure asset management has the capability to extend the life of existing infrastructure, provide improved performance while controlling the costs to both agencies and users, and provide targeted data for enabling the fruition of life-long design. CIAMTIS will focus on the following three themes:

1. Application of innovative materials and technologies.
2. Condition assessment and health monitoring.
3. Infrastructure management and innovative financing.

CIAMTIS will undertake research, education, and outreach activities to support these themes. The Center has developed a strategy for processing and archiving digital data sets, which are outlined in this data management plan. The plan is organized into five sections. The first section provides a high-level description of the types of data that the Center anticipates gathering in the course of conducting research activities. In the second section, an outline of the proposed standards and machine-readable formats that will be used for research activities is provided. The third section describes the data access policies that will govern disclosure of identities, confidential business information, national security information, etc., and whether public use files may be generated from the data. In the fourth section, a general discussion of policies for re-use and re-distribution of the research data is offered. The fifth section describes how, when, and where CIAMTIS plans to archive, preserve, and deposit the research data.

DESCRIPTION OF DATA

Research activities within CIAMTIS will result in the generation of four categories of data:

(a) **Hard copy experimental data:**

Written descriptions of sample preparations and test setups (e.g., date/time, temperature, reference to specific ASTM or other test protocols, material formulations and proportions, casting and curing practices, sample dimensions), hardcopy experimental observations and results;

(b) **Digital experimental data:**

Digital descriptions of sample preparations and test methods, images and photographs, videos, details pertaining to the assembly of laboratory specimens, raw data from computer-controlled testing equipment, processed test data including statistical analysis of test results (e.g., using Excel, MatLab, R), and software reports (e.g., image analysis);

(c) **Analytical or numerical data:**

Details of the derivation of analytical models, and input files and results of numerical modeling and simulations;

(d) **Observational data:**

Details of data collected from transportation agency sources, such as pavement ride quality, pavement surface friction, bridge rating, etc., which will be used to develop statistical models of infrastructure deterioration or performance;

(e) **Data generated for the purpose of professional and broader dissemination of project outcomes:** Figures, tables, and other graphics generated for the purpose of peer-reviewed journal publications/reports/theses, professional and public presentations (e.g., Powerpoint files), CIAMTIS website as well as workshop and curriculum materials, including the classroom notes for the classes, seminars, webinars, and other educational and technology transfer activities.

The experimental/physical measurements will be conducted according to strict testing protocols (e.g., relevant ASTM and AASHTO standards) in order to generate statistically defensible results that are acceptable by transportation agencies in the region and by peer-reviewed journals. It is anticipated that all data files assembled from laboratory experiments, or data assembled for the purposes of analytical/numerical/statistical modeling, will be preserved for long-term access.

OUTLINE OF STANDARDS AND MACHINE-READABLE FORMATS

The formats for the data generated in each category are as follows:

(a) **Hard copy experimental data:**

Hard copy data will be collected in standard laboratory notebooks that will include the date/time, description of experiment or activity, ambient conditions (e.g., temperature, humidity, etc.), observations, and results obtained. All graduate students and researchers will be responsible to record this information in their individual lab notebooks. The PIs of each

research project will be responsible for making sure project data are properly documented and for maintaining the laboratory notebooks for five years after conclusion of each project.

(b) Digital experimental data:

- Sample preparation processes in MS Excel spreadsheets (*.xls or *.xlsx) or Word documents (*.doc or *.docx)
- Test data (LabVIEW project and results files in *.lbv format)
- Signal or image processing data (MATLAB format using *.m or *.mat file format)
- Multimedia (*.jpg, *.gif, *.tiff, *.mp4 files)
- Data codes and variable descriptions (documentation) will be saved for each data file in an ASCII text file (*.txt), MS Excel (*.xls or *.xlsx), or MS Word format (*.doc or *.docx)

(c) Analytical, numerical, and statistical analysis data:

- Details of the derivation of analytical models (MS Word or LaTeX documents in *.doc, *.docx or *.pdf)
- Simulation software (MATLAB *.m files)
- Results of numerical modeling and simulations (database with selected outputs, e.g. in *.db)
- Statistical analysis software and pertaining results (MATLAB *.m files, *.mat files, *.fig figures, and binary output files with accompanying ASCII data header files and Readme instructions)

(d) Observational data:

- Analysis files (MS Excel *.xls or *.xlsx)
- Statistical modeling files (SPSS, Stata, Minitab, R)

(e) Data generated for the purpose of professional and broader dissemination of project outcomes:

- Reports and technical briefs (MS Word *.doc or *.docx, and *.pdf)
- Presentations (MS Powerpoint *.ppt or *.pptx, and *.pdf)
- Journal manuscripts (*.pdf)

DATA ACCESS POLICIES

It is not anticipated that research projects undertaken by faculty or students in CIAMTIS will contain any personally-identifiable information (PII), confidential business information, or national security information. However, if PII data are collected, the research project Principal Investigator will seek study approval from their university Institutional Review Board (IRB) prior to collecting this information, which will include informed consent and data collection protocols. If these data are collected for a research project funded using CIAMTIS funds, PII information will be retracted from all data files, and replaced with a participant identification number.

Similarly, if confidential business or national security information is collected as part of any CIAMTIS research project, this information must be considered by the project Principal Investigator. The PI will be asked to abide by all conditions imposed on these data, including withholding information, or altering the data, so that the privacy information is not compromised.

Some research projects conducted by faculty or students affiliated with CIAMTIS may collect data related to patented research. So that these data are not shared prematurely, CIAMTIS will place an embargo on the data for a limited period of time (e.g., one year after project completion) to protect the intellectual property of the researchers. The researchers developing the data will be encouraged to seek additional guidance from their university regarding the ownership and management of intellectual property.

Research project Principal Investigators must document IRB approvals, as well as planned collection of confidential business or national security information, in their proposals and research progress reports. Any concerns related to data privacy must be reported early in the project so that these issues may be mitigated (or exceptions granted) to the data sharing policies of CIAMTIS.

POLICIES FOR RE-USE AND RE-DISTRIBUTION OF DATA

Project Principal Investigators will maintain intellectual property rights of all original data that are generated for research projects. In the event that the data have copyright restrictions, these will be identified in the documentation associated with the data, and users of the data will be asked to abide by the conditions identified in the documentation. Further, all data will contain a reference, which recipients of the data will be asked to acknowledge in their resulting research projects.

It is anticipated that research reports conducted under CIAMTIS will be made available to the public no later than 30 days after the conclusion of the project period of performance. The data corresponding to the report will also be made available during this same window.

Data, samples, and derivatives resulting from CIAMTIS activities will be recognized if/when they are re-used or re-distributed for future research. Proper acknowledgement of UTC support will be made .

PLANS TO ARCHIVE, PRESERVE, AND DEPOSIT DATA

Sharing data has many benefits in the scientific community, including: (1) promoting alternative research methods, (2) testing alternative hypotheses, (3) supporting professional development of new researchers, (4) producing new or enhanced datasets, or (5) exploring new research topics not addressed by the original research. Research project principal investigators in CIAMTIS will be asked to store their “final research data,” including appropriate documentation to explain the codification of data, in a data repository (see below for dissemination information). CIAMTIS will also encourage research PIs preserve raw and processed data files in their own archives.

CIAMTIS will encourage project Principal Investigator’s to make research results widely known, accessible, and available to interested parties via publication of research results in archival journals. In addition, the Center’s website will contain electronic copies of research reports, with instructions where to find final research data. To this end, data generated through CIAMTIS will be made publicly available using a generalist data repository, such as Dryad Digital Repository or figshare. There is a nominal cost to use the Dryad repository (\$120 for the first 20 GB and \$50 for each additional 10 GB), while the figshare repository is free for the first 100 GB (additional

fees apply for larger datasets). The project PI will transfer research outcomes to the repository within 30 days after the final report is published, provided that there is not copyright or intellectual property concerns with the data.

In addition to making final research data available via a public access data repository, Penn State will store data (including documentation of the data) in the cloud using box.psu.edu for at least three years after the CIAMTIS period of performance has concluded. Principal Investigators will be asked to share their data with Penn State for archiving purposes at the same time as the data are transferred to a public data repository. In the event that the repository data are unexpectedly lost, researchers wishing to access the data via the Penn State Box account will be able to do so by contacting the Center Director. The Center Director will then offer a link to the data for download.