

U.S. Department of Transportation  
Office of the Assistant Secretary for Research and Technology  
University Transportation Centers Program  
Grant No. 69A3551947136



**NATIONAL INSTITUTE FOR  
CONGESTION REDUCTION**

## **DATA MANAGEMENT PLAN**

University of South Florida  
in partnership with  
University of California, Berkeley  
Texas A&M University and its affiliated Texas A&M Transportation Institute  
University of Puerto Rico at Mayagüez

v1.0 December 30, 2019  
Approved January 28, 2020

---

## INTRODUCTION

This document serves as the National Institute for Congestion Reduction (**NICR**) **Data Management Plan (DMP)**, describing our policies for handling digital data both during and after our research program. NICR aims to optimize the efficiency and reliability of travel for all transportation system users, supported by data modeling and analytical tools to evaluate the effects of shifting transit incentive structure and ridesharing and alternative forms of transportation. We view the DMP as a living document that will be updated as our programs evolve. The DMP includes discussions of handling, storage, dissemination, and sharing of research results. OMB Circular A-110 defines research data as: “the recorded factual material commonly accepted in the scientific community as necessary to validate research findings, but not any of the following: preliminary analyses, drafts of scientific papers, plans for future research, peer reviews, or communications with colleagues. This “recorded” material excludes physical objects (e.g., laboratory samples).” All of our efforts will comply with the U.S. Department of Transportation (DOT) [Public Access Plan](#), and this DMP will be available via the NICR website ([www.cutr.usf.edu/nicr](http://www.cutr.usf.edu/nicr)). NICR benefits from helpful [Data Management Planning Resources](#) organized by the USF Library. The NICR Director and Executive Committee will be responsible for implementation of the DMP with input from all project principal investigators (PIs).

We also note that we will be using Scholar Commons, a service of the University of South Florida (USF) Library as a permanent repository for all NICR products, including reports, and U.S. DOT reporting, with digital object identifiers (DOIs) to permanently identify a document and link to it on the NICR website. Scholar Commons is part of the Digital Commons Network™ that includes more than 3.4 million original scholarly works from 566 institutions. We are exploring the use of Digital Commons for posting datasets, that includes on-demand metrics, different levels of authorization settings, licensing and copyright tools, customizable metadata, different display options and persistent URLs.

### 1. Data Description

**NICR’s policy is that data used/generated at the project level will be available in open access format**, free from proprietary formats and barriers. NICR’s research aims to produce tools (methods, guidelines and policies) for improving congested, multimodal systems in the immediate future. By operationalizing theories developed by our consortium, using recent advancements in big data science and adaptations to existing technologies, our tools will be available in the short run, without having to wait for future technological advances to understand congestion impacts. By generalizing our theories, our tools will be ready for immediate use in complex, real-world environments more

---

efficiently. Further, the tools we develop will be suitable for application to multimodal systems. Harnessing the consortium's intellectual horsepower and deep domain knowledge in these ways will ensure NICR's success.

We have identified a set of 17 projects that will be led by consortium faculty; each project will be collaborative and will include at least one faculty member from two different universities. We envision that the funded research projects will rely upon a wide range of quantitative and qualitative transportation data, including use of existing data sets, collection of new data, and simulation. These include but are not limited to:

- Tolling data including rates, speeds, volumes and automatic vehicle identification (AVI) data from sensors.
- Databases and spreadsheets including metadata.
- Vehicle/traveler location data at various temporal resolutions, including Global Positioning Systems (GPS) and mobile phone data.
- Data describing transportation network/supply characteristics.
- Images including maps.
- Simulated data produced via simulation tools.
- Travel surveys, interviews, expert panels, and focus groups.
- Video surveillance from roadside cameras and/or unmanned aircraft (UA).
- Traveler activity and transport demand data including origins and destinations.
- Traffic count, occupancy and speed data at various temporal resolutions.
- Geographic information systems (GIS) spatial data.
- Demographic, socioeconomic, and census data.
- Transit vehicle, taxi, and transportation network company (TNC) fleet position, passenger ons/off, and transit schedule data.
- Source code for developing algorithms, simulations and other tools/products.

All data collection efforts involving human subjects at NICR will follow each university's Institutional Review Board (IRB) and Human Research Protection Program (HRPP) policies. The University of South Florida (USF) policies are listed here:

<https://www.usf.edu/research-innovation/research-integrity-compliance/ric-programs/irb/irb-hrpp/index.aspx>.

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 lack of public access. Each project will include a work plan that will be peer-reviewed and the PIs will be required to submit a project

---

DMP as part of the project scope of work that will address project-specific data dissemination and sharing issues, including:

1. Purpose of research project.
2. Project data to be collected or used, and any details of the data-producing program.
3. Description of project data that will be generated, including nature and scale.
4. Methods and procedures for creating the data.
5. Time period over which data will be collected, and update frequency.
6. Description of relationship between data being collected for project and existing/previously collected data, if applicable.
7. Potential data users.
8. Potential long-term data value for project, NICR, U.S. DOT and the public.
9. Rationale for limiting public access (in limited cases, as applicable)
10. Party responsible for managing data at project level.

Based on the project-level input, this DMP will be updated throughout the life of the NICR program as warranted.

## 2. Standards Used

**NICR's policy is to use data in open, standard formats and to provide appropriate metadata and data dictionaries so that others may understand our data.** At the NICR program level, we anticipate that data will be collected and generated in a range of formats including but not limited to:

- .docx, .xlsx, .pptx
- .xml
- .pdf/a
- .ascii
- .txt
- .csv
- .ascii
- .tiff
- .jpeg
- .png
- .avl
- .gps
- .mxd, .yr, .gdp, .shp, .dbf
- .html
- .php
- .mov
- .mpeg
- .mp4

As described above, any data used in NICR projects will be converted to open access formats so that the public may easily access the data. Data tables and databases will include necessary metadata and/or data dictionaries so that it is understandable to future users. Some data may be offered to the Federal Highway Administration [ITS DataHub](#), and some may be compliant with the [CARMA Platform](#). Each project's work plan will request PIs to provide the following, as applicable:

- 
1. Project open data formats. In limited cases PI will justify/explain rationale for use of proprietary data format and/or convert to open format.
  2. Data version control/indication.
  3. Metadata schema or data dictionary to make the data understandable by other researchers, and management/storage of same.
  4. Are any tools or software required to read or view the data.
  5. Data quality control measures.

### 3. Access Policies

**NICR's policy is to openly share data generated/used under our projects and programs funded by the U.S. DOT and our partners**, subject to any confidentiality, privacy, security or human subjects restrictions. In this spirit we welcome other researchers to pursue the principles of Reproducible Research whereby they may verify findings and build upon them. University campus IRB and HRPP policies will be followed and no private, confidential or personally identifiable information (PII) will be revealed.

One of NICR's research projects (3-1 Influencing Travel Behavior via an Open Source Platform) proposes to create a *OneBusAway* data center on the USF campus that will allow for access and collaboration. Other projects may also generate online web-based access points as part of a larger USF Transportation Data Archive. For individual NICR research projects, each PI will be responsible for protecting privacy and confidentiality according to their campus IRB and HRPP policies. Each research project work plan will request the following information as applicable:

1. What data will be publicly shared and how it will be accessed.
2. Plans to follow campus IRB and HRPP policies, and methods for protecting PII, confidentiality and security.
3. Any restrictions on data access.
4. Any additional steps beyond campus IRB and HRPP policies to protect privacy and confidentiality.

### 4. Re-Use, Redistribution, and Derivative Products Policies

**NICR's policy is to make project data available for open sharing** under the Creative Commons Zero (CC0) universal public domain dedication. Under CC0, data and derivative products will be available for reuse and redistribution without restriction. In rare cases, as described above, PIs for individual research projects may face

---

restrictions due to use of proprietary data, and this will be disclosed in individual project work plans.

## 5. Archiving and Preservation Plans

**NICR's policy is to archive and preserve project data, with appropriate disaster recovery procedures to ensure that data will remain openly available and recoverable.** We intend to use ScholarCommons at USF as the repository for NICR project data. We intend to include metadata in conformance with the U.S. Federal Government [Project Open Data Metadata Schema](#). ScholarCommons supports the generation of persistent generators, in particular we intend to use permanent DOIs for all of our reports and datasets. ScholarCommons files are duplicated and backed up on- and off-site. DOIs are deposited with CrossRef and content is submitted to LOCKSS for inclusion in their archive. Individual PIs may propose to use institutional data archives such as the Texas Data Repository (TDR) at Texas A&M University or the Merritt Repository at U.C. Berkeley.

## REFERENCES

- [Guidelines for Evaluating Repositories for Conformance with the DOT Public Access Plan](#)
- [Plan to Increase Public Access to the Results of Federally Funded Scientific Research Results](#)
- [Creating Data Management Plans for Extramural Research](#)
- National Science Foundation [Dissemination and Sharing of Research Results](#)
- Office of Management and Budget. Uniform Administrative Requirements for Grants and Agreements with Institutions of Higher Education, Hospitals and Other Non-Profit Organizations ([OMB Circular A-110](#))