DATA RESOURCES TESTBED 101 OVERVIEW

WHAT IS THE DATA RESOURCES TESTBED?

The Data Resources Testbed (DRT) is a Federal Highway Administration (FHWA)-sponsored initiative to support project data storing and sharing of transportation operations related data for the Office of Operations, the Office of Operations Research and Development, and other FHWA offices.

WHAT DATA AND RESOURCES WILL RESIDE IN THE DRT?

The DRT collects and provides access to operations-related data that are not suitable for storing on one of the U.S. Department of Transportation (USDOT) public facing data portals. Depending on the research project, the generated or collected data may include foundational primary data (e.g., traffic flow data, video data, and connected vehicle data) or derived data (e.g., traveler information, traffic operations analysis). The DRT maintains an inventory of available primary date, derived data, and other resources, including: description of the data and resources, metadata document, folder hierarchy, data management plans (DMPs), contract scope of works, and data use agreements.

WHAT PROJECT DATA SHOULD BE SAVED IN THE DRT?

The type, characteristics, and expected use of data determines what are appropriate to save and archive in the DRT. Data generated by operations research projects are collected, archived, and made available in the DRT if the project managers deem it unfeasible to make the data accessible via a public facing USDOT data portal. A project's DMP assists project managers with deciding what data should be collected, saved, and archived, and what data should be available for others to access and use.

HOW DO YOU DECIDE IF DATA SHOULD BE MADE PUBLIC OR REMAIN STORED?

The project manager must first determine if the data are valuable to the public, if there is any personally identifiable information (PII) in the data, if they contain any other sensitive information, if there's a data use agreement, and lastly, if it's a large dataset. Answers to these questions will help project managers determine if the data belong in the DRT (figure 1).
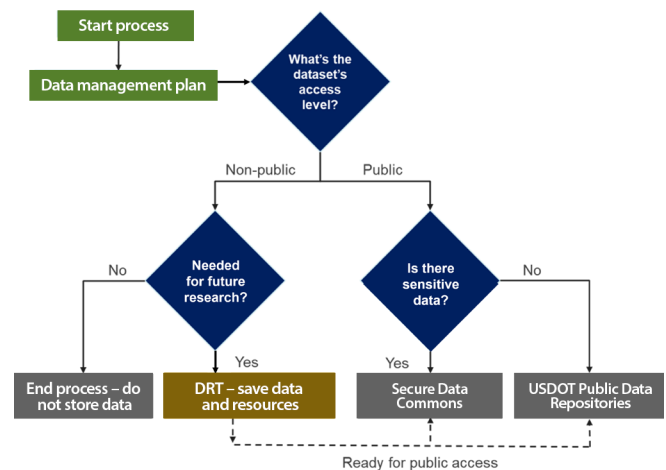
Source: FHWA.

Figure 1. Diagram. Data management decision process.

WHAT DOCUMENTS ASSIST IN DETERMINING WHAT DATA SHOULD BE SAVED AND WHERE?

Resources have been developed to assist FHWA researchers with determining what data may be appropriate to save, where they could be archived, and who may be allowed to access and use the data.

These resources include:

- Detailed procedural guideline on what data to save, archive, and make available.
- Project checklist to assist with identifying the lifecycle of data throughout the duration of the research project.
- Complete scope of work requirements and issues to consider during procurement of documents for research projects involving the collection, saving, and making of data available to use.
- DMP template.
- Metadata document (e.g., data structure and documentation).
- Data sharing agreement template.

WHAT ARE METADATA AND DATA MANAGEMENT PLAN DOCUMENTS AND WHY ARE THEY IMPORTANT TO DRT?

The metadata document will assist users with understanding the data origins, how data can be assessed, what elements the data contains, details (e.g., spatial, temporal), and formats. The goal is to help users find relevant information, discover resources, and understand the significance of the data. It will assist users to find, in one document, information on the data collection procedures, types of data, and different formats and definitions. It describes any post processing performed on the data, defines the folder and file naming conventions used, and identifies the number and size of files. Metadata are structured/standardized information that describes, clarifies, and supports the use or management of an information resource. Additional guidance on these metadata requirements for data submitted to the DRT is available at: https://www.its.dot.gov/data/.

A DMP should be developed for each project. The document describes how the data was collected, saved, and managed throughout the project. Should there be any reason that the data collected or generated by this project will not be suitable for use by the public (e.g., sensitive, or proprietary, PII), the justification will be documented in this plan and in the data use agreement governing the allowed use of the project's data.

WHERE CAN YOU FIND DATA AVAILABLE FOR PUBLIC SHARING?

The following is the list of USDOT and FHWA websites where specific types of data can be accessed publicly:

- Intelligent Transportation Systems (ITS) Public Data Hub: Point of entry to search and access all ITS and Office of Operations related data publicly available at https://www.its.dot.gov/data/.
- National Transportation Library: Portal for technical reports and derived transportation data available at https://ntl.bts.gov/.
- Data.transportation.gov: Portal for primary research data available at https://data.transportation.gov/.
- Secure Data Commons (SDC): A cloud-based analytics platform that enables traffic engineers, researchers, and data scientists to access transportation-related datasets available at https://www.transportation.gov/data/secure.

However, because some data are sensitive, large, or otherwise cannot be made publicly available, that data are temporarily stored and managed in the Saxton Transportation Operations Laboratory (STOL) DRT and later may be moved to the SDC.

WHO DECIDES WHEN DATA FROM PROJECTS SHOULD BE SAVED AND MADE AVAILABLE FOR FUTURE USE?

When it's determined reasonable and practical to make data from FHWA-funded operations projects available to the public, the STOL DRT team works with the FHWA project managers to determine on which of the USDOT-sponsored portals the data should be archived and made available.