PLAN OVERVIEW

A Data Management Plan created using DMP Tool

DMP ID: https://doi.org/10.48321/D1R67H
Title: Inclusive Wayfinding Through NaviLens
Creator: Mira Philipson – ORCID: 0009-0007-9551-538X
Affiliation: United States Department of Transportation (DOT) (transportation.gov)
Principal Investigator: Mira Philipson – ORCID: 0009-0007-9551-538X
Data Manager: Mira Philipson, Lisa Mae Fiedler, Christine Singh, Preksha Agarwal, Rahnuma Tarannum
Project Administrator: Mira Philipson
Contributor: Yuvraj Ramsaywack - **ORCID:**0009-0002-9619-6409
Funder: United States Department of Transportation (DOT) (transportation.gov)
Funding opportunity number: DOT-SMART-FY23-01
Grant: SMARTFY22N1P1G45

Template: SMART Grants Stage 1 Data Management Plan (DMP)

Project abstract:

The Project is a Systems Integration solution in partnership with NaviLens. NaviLens is a mobile phone application that utilizes customized colorful QR-like codes (“codes”) to provide users with context-specific indoor and outdoor wayfinding and service information.

NaviLens can translate visual signage into audio, and to solve the “first mile/last mile” problem, in which GPS data only guides a user to the general vicinity of a location, as opposed to directly to the location itself. The app can detect codes at significant distances and angle, informing the user of what destination has been detected (i.e., “stairs down to R platform” or “bus stop 23 St, Broadway”) and their distance and angle (i.e., “50 feet away, to the right), to help guide the customer directly to the exact location of the code. Once the user has arrived at the location & scans the code, additional information can be presented including live service information (e.g., train/bus arrivals, elevator status). NaviLens can also translate this information into dozens of languages, a boon in a diverse city like New York.

The Project demonstrates the feasibility of NaviLens as a cost-effective technology solution that provides context-specific wayfinding and real-time service information, in corresponding audio and visual formats, throughout an end-to-end transit journey. This information can assist all transit riders, particularly those with visual disabilities and non-native English speakers. Stage 1 allowed the MTA to pilot the NaviLens’ wayfinding solution with both real-time and static NYCT information for both the subway and bus systems, understand the internal and external resources needed for a larger deployment in a possible systemwide rollout, proactively identifying and tackling roadblocks. These lessons can be applied to other agencies and that NaviLens can help provide a cost-effective, accessible solution for systems looking to provide better communication to its customers.

Start date: 09-01-2023
End date: 09-01-2025
Last modified: 01-22-2026

INCLUSIVE WAYFINDING THROUGH NAVILENS

DATASET AND CONTACT INFORMATION

Please provide as much of the the following information as possible:

1. **Name of the project;**
2. **Grant number;**
3. **Name of the person submitting this DMP;**
4. **ORCiD of the person submitting this DMP (need an ORCID? Register here: https://orcid.org/);**
5. **Email and phone number of the person submitting this DMP;**
6. **Name of the organization for which the person submitting this DMP is working;**
7. **Email and phone number for the organization;**
8. **Link to organization or project** website, if applicable; and,
9. **Date the DMP was written.**

1. Inclusive Wayfinding Through NaviLens
2. Grant number: SMARTFY22N1P1G45
3. Miriam Philipson
4. ORCiD: 0009-0007-9551-538X
5. miriam.philipson@mtahq.org; 917-890-0515
6. Metropolitan Transportation Authority (NY)
7. miriam.philipson@mtahq.org; 917-890-0515
8. mta.info
9. 1/8/2026

DATA DESCRIPTION

Please provide as much information as possible:

1. **Provide a description of the data that you will be gathering in the course of your project or data from a third party that you will re-use, if any;**
    1. **If there will be no data collected or re-used from another source, state that this is case;**
2. **Address the expected nature, scope, and scale of the data that will be collected, as best as you can at this stage;**
3. **As best as you can, describe the characteristics of the data, their relationship to other data, and provide sufficient detail so that reviewers will understand any disclosure risks that may apply;**
    1. **If data might be sensitive, please describe how you will protect privacy and security, if you know that now;**
    2. **You may need to update your DMP later to add more detail;**
4. **Discuss the expected value of the data over the long-term.**

1. We are interested in the codes that our customers find useful. NaviLens allows us to collect anonymized count data to help us track how often customers are scanning individual codes. This will help inform how codes can be placed in a systemwide rollout to ensure seamless navigation. This will also allow us to figure out which codes are most helpful, which can be determined by the number of scans. We are working with NaviLens to determine if we can track the languages used when scanning codes so we can determine how well we are reaching those with limited English proficiency.

2. All of the NaviLens codes and their respective content are stored in a large database, functionally like an Excel spreadsheet. Content and codes can be updated in real-time depending on customer use, particularly to make them more useful. We are looking to collect count data per code for now.

3. The data are all anonymized from NaviLens, so we will not collect/store/analyze data that could infringe on our customers' privacy. This stage builds on our past proof of concept/pilot of NaviLens, and the company has adjusted their data storage policies in accordance with New York State cybersecurity regulations. We will work off past cybersecurity agreements, which will help protect our data servers and our customers.

4. We are interested in tracking use over time as this will inform a larger/longer-term rollout. We are particularly interested in tracking usage across 2 subgroups: those who are blind and those who have limited English proficiency. Depending on the program's success with non-English speakers, perhaps a future stage can be targeted at intermodal or "gateway" facilities (e.g., JFK/EWR airports, Port Authority Bus Terminal) or at stations with high non-English speaking users (e.g., Flushing - Main St, 181 St, Jackson Heights - Roosevelt Av). We are waiting for the data to be collected before drawing further conclusions, but trends in usage will be helpful in strategizing what an expansion, if any, could look like. We do not expect decreased usage over time to be a project failure, however.

DATA FORMAT AND METADATA STANDARDS EMPLOYED

Please provide as much information as you can:

1. **Describe the anticipated file formats of your data and related files;**
2. **To the maximum extent practicable, your DMP should address how you will use platform-independent and non-proprietary formats to ensure maximum utility of the data in the future;**
    1. **If you are unable to use platform-independent and non-proprietary formats, you should specify the standards and formats that will be used and the rationale for using those standards and formats.**
3. **Identify the metadata standards you will use to describe the data.**
    1. **At least one metadata file should be a DCAT-US v1.1 (https://resources.data.gov/resources/dcat-us/) .JSON file, the federal standard for data search and discovery.**

NaviLens data exports are in .csv formats. This data is easily exportable, and data on the individual scan level will be transparent. There is no data in proprietary formats.
This data will be public and open-access, in .csv format. We are exploring metadata standards and will ensure the data has at least a DCAT-US v1.1 .JSON file. No data will be released in proprietary forms.

ACCESS POLICIES

In general, data from DOT-funded projects must be made publicly accessible. Exceptions to this policy are: data that contain personally identifiable information (PII) that cannot be anonymized; confidential business information; or classified information. Protecting research participants and guarding against the disclosure of identities and/or confidential business information is an essential norm in scientific research. Your DMP should address these issues and outline the efforts you will take to provide informed consent statements to participants, the steps you will take the protect privacy and confidentiality prior to archiving your data, and any additional concerns. In general, in matters of human subject research, your DMP should describe how your informed consent forms will permit sharing with the research community and whether additional steps, such as an Institutional Review Board (IRB), may be used to protect privacy and confidentiality. Additionally, when working with, or conducting research that includes Indigenous populations or Tribal communities, researcher will adhere to the CARE Principles for Indigenous Data Governance https://www.gida-global.org/care and make an explicit statement to that effect in this portion of the DMP.

Please provide as much information as possible:

1. **Describe any sensitive data that may be collected or used;**
2. **Describe how you will protect PII or other sensitive data, including IRB review, application of CARE Principles guidelines, or other ethical norms and practices;**
    1. **If you will not be able to deidentify the data in a manner that protects privacy and confidentiality while maintaining the utility of the dataset, you should describe the necessary restrictions on access and use;**
3. **Describe any access restrictions that may apply to your data;**
4. **If necessary, describe any division of responsibilities for stewarding and protecting the data among Principal Investigators or other project staff.**

All data collected are anonymized. For determining whether or not a user is blind, we use the wayfinding-specific feature of the application as a proxy as this feature is also available to sighted users. We do not foresee any access restrictions with this data and will work with our internal data teams to provide this publicly. The MTA already shares most of its performance data through the New York State open data portal, and we believe this would also be a suitable place to store this information to maximize public research benefit.

RE-USE, REDISTRIBUTION, AND DERIVATIVES PRODUCTS POLICIES

Recipients are reminded:

1. Data, as a collection of facts, cannot be copyrighted under US copyright law;
2. Projects carried out under a US DOT SMART Grants is federally funded; therefore, as stated in grant language:
    1. Recipients must comply with the US DOT Public Access Plan, meaning, among other requirements, project data must be shared with the public, either by the researchers or by US DOT;
    2. That by accepting US DOT funding through this grant, recipients have granted to US DOT a comprehensive non-exclusive, paid-up, royalty-free copyright license for all project outputs (publications, datasets, software, code, etc.). This includes all rights under copyright, including, but not limited to the rights to copy, distribute, prepare derivative works, and the right to display and/or perform a work in public; and,
    3. In accordance with Chapter 18 of Title 35 of the United States Code, also known as the Bayh-Dole Act, where grant recipients elect to retain title to any invention developed under this grant, US DOT retains a statutory nonexclusive, nontransferrable, irrevocable, paid-up license to practice or have practiced for or on behalf of the United States any such invention throughout the world.

Please provide as much information as possible:

1. **Describe who will hold the intellectual property rights for the data created or used during the project;**
2. **Describe whether you will transfer those rights to a data archive, if appropriate;**
3. **Identify whether any licenses apply to the data;**
    1. **If you will be enforcing terms of use or a requirement for data citation through a license, indicate as much in your DMP;**
4. **Describe any other legal requirements that might need to be addressed.**

MTA's contracts with NaviLens require that they share all usage data with the MTA. While NaviLens ultimately owns the codes, MTA owns all related data. All information to be fed into the NaviLens codes are publicly available (e.g., real-time train/bus arrival, bus number, service status). All data is publicly available on data.ny.gov.
Since this data is not sensitive information, a license is not needed. Anyone can access information through the New York State Open Data Portal (data.ny.gov) without a login or license, and the data is free of charge to access.

ARCHIVING AND PRESERVATION PLAN

Please provide as much information as possible:

1. **State where you intend to archive your data and why you have chosen that particular option;**
2. **Provide a link to the repository;**
3. **You must describe the dataset that is being archived with a minimum amount of metadata that ensures its discoverability;**
    1. **Whatever archive option you choose, that archive should support the capture and provision of the US Federal Government DCAT-US Metadata Schema https://resources.data.gov/resources/dcat-us/**
4. **In addition, the archive you choose should support the creation and maintenance of persistent identifiers (e.g., DOIs, handles, etc.) and must provide for maintenance of those identifiers throughout the preservation lifecycle of the data;**
5. **Your plan should address how your archiving and preservation choices meet these requirements.**

Data and files will be kept in MTA's Github account (https://github.com/nymta). Usage and survey feedback data is publicly posted on the existing New York State open data portal (data.ny.gov), where all of MTA's other public data feeds are housed. Data is completely anonymized and any identifying information is not retained. All data is collected at the scan level (each record is an individual scan event).

PLANNED RESEARCH OUTPUTS

DATASET - "MTA_NAVILENS_USAGE"

Navilens usage data contains records of scan events (each row represents a time a NaviLens code was scanned). Scan records include the code identifier, the type of asset scanned (e.g., bus stop, bus vehicle, subway station, train car, or vehicle feature), the app used (NaviLens vs. NaviLens GO), the device language setting at time of scan, and timestamps. For certain asset types, scan events are also enriched with location fields such as bus stop identifiers and names or subway station identifiers and names. This data supports evaluation of the NaviLens deployment and related project reporting, including work funded through the USDOT SMART Grant.

DATASET - "MTA_NAVILENS_FEEDBACK_SURVEY"

Survey responses include overall satisfaction, agreement ratings about ease of use and information quality, recommendations, interest in expanding code coverage, comparisons to other wayfinding/travel apps, and open-ended comments. The survey also includes optional demographic questions (e.g., age range, typical transit use frequency, Access-A-Ride customer status, and disability identification) to help contextualize feedback. This survey supports evaluation of the NaviLens deployment and related project reporting, including work funded through the USDOT SMART Grant.

PLANNED RESEARCH OUTPUT DETAILS

Title	Type	Anticipated release date	Initial access level	Intended repository(ies)	Anticipated file size	License	Metadata standard(s)	May contain sensitive data?	May contain PII?
MTA_NaviLens_Usage_	Dataset	2026-01-08	Open	Github		Creative Commons Zero v1.0 Universal	DCAT-US	No	No
MTA_NaviLens_Feedback_Survey	Dataset	2026-01-08	Open	Github		Creative Commons Zero v1.0 Universal	DCAT-US	No	No