

Advancing Rural Mobility: Michigan Public Transit Open Data Standards Program

Project 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. Name of the project: Advancing Rural Mobility: Michigan Public Transit Open Data Standards Program
2. Grant number: SMARTFYNN21P1G33
3. Name of the person submitting this DMP: Jeremy Papuga
4. ORCID of the person submitting this DMP: 0009-0007-2292-1440
5. Email and phone number of the person submitting this DMP: jpapuga@hntb.com / 313.316.3723
6. Name of the organization for which the person submitting this DMP is working: Michigan Department of Transportation
7. Email and phone number for the organization: GeisslerJ1@michigan.gov, 517.930.6582
8. Link to organization or project website, if applicable; and: <https://www.michigan.gov/mdot/travel/mobility/pub-transit>, <https://miruralmobility.org/>
9. Date the DMP was written. 12/15/2024 - updated 6/3/2025

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;
 1. If you answered "No data" above, then you are finished and may skip the most of the steps and sections below, after you:
 1. Save your DMP as it exists;
 2. Submit it to your Grant Manager or the NTL staff for review.
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.

The Advancing Rural Mobility project (ARMP) has generated a diverse and valuable collection of data, including GTFS-Flex data feeds, MiTripPlanner code, media assets from the Charlevoix press conference, site visit presentations, rider survey data, the project website and newsletter, marketing kits for transit agencies, and project presentations. This collection of data plays a crucial role in advancing rural mobility by improving transit accessibility, fostering industry-wide adoption of GTFS-Flex, and supporting informed decision-making for transit agencies.

The expected long-term value of the data from the ARMP project is substantial, especially in shaping the future of rural transit accessibility and efficiency. The adoption of GTFS-Flex as a standard data specification in April 2024 marked a

significant step forward for the transit industry, providing a structured format for flexible transit services. This not only encourages more agencies to adopt GTFS-Flex but also enhances interoperability and consistency across transit systems, making it easier for riders and agencies to access and share critical transit data.

As the ARMP pilot continues to expand across the state, the accessibility of transit information will improve, benefiting both riders and transit operators. Lessons learned from the pilot will aid in refining best practices and ensuring smoother implementation for additional rural transit agencies. Over time, transit agencies will be better equipped to make data-driven decisions, optimizing services, planning routes, and allocating resources effectively based on actual travel patterns, peak demand periods, and underserved areas.

By making transit data publicly available, riders will have greater access to real-time and scheduled transit information, leading to improved trip planning and convenience. Transparency in transit services fosters trust among users and encourages broader utilization of rural mobility options. Additionally, the project's commitment to ethical data practices ensures that sensitive information is protected while still allowing for meaningful analysis. Since no personal data is recorded or shared, the initiative aligns with privacy standards while maintaining the integrity of transit data.

The availability of structured transit data has the potential to influence policy decisions and funding allocations in the long run. Legislators and transportation authorities can leverage insights from this project to advocate for increased investment in rural mobility solutions, supporting sustainable growth and innovation in the sector. Overall, the ARMP project is laying the groundwork for a more connected, data-driven, and accessible rural transit system, with long-term benefits extending to improved service efficiency, better-informed transit policies, and enhanced rider experiences.

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.**

The ARMP project data will be available in various file formats to ensure accessibility and usability across different platforms. The GTFS-Flex data feeds will primarily be stored in CSV and TXT formats, following the General Transit Feed Specification (GTFS) standard, which is widely used in the transit industry for structured transit data. The MiTripPlanner front-end web application code is written primarily in typescript - an open source programming language developed by Microsoft. Other project-related materials, such as the news release, photos, and videos from the Charlevoix media conference, will be provided in PDF, JPEG, PNG, and MP4 formats. Project presentations will be available in PDF formats. The rider survey data will be stored in PDF formats for easy analysis. The project website and newsletter will be accessible in HTML and PDF, while the marketing kits for transit agencies will be distributed in PDF and PNG formats.

To maximize the utility of the data in the future, the project will prioritize platform-independent and non-proprietary formats wherever possible. GTFS-Flex data is already structured in an open format that is widely accepted across transit systems. GeoJSON is preferred for GIS-related files due to its compatibility with web-based applications and open-source mapping tools. CSV is used for tabular data to ensure broad accessibility across different software platforms. PDF is chosen for documents to maintain formatting consistency across devices. However, some proprietary formats, such as Shapefiles (SHP), may be necessary due to industry standards and software compatibility. In cases where proprietary formats are used, efforts will be made to provide alternative open formats, such as GeoJSON for spatial data and PDF versions of presentations.

For metadata standards, the project will adhere to DCAT-US v1.1, the federal standard for data search and discovery. At least one metadata file will be provided in JSON format, ensuring compliance with federal open data requirements.

Additionally, metadata for transit-related datasets will follow GTFS metadata conventions, ensuring consistency and interoperability with existing transit data platforms. These metadata standards will enhance discoverability, usability, and long-term accessibility of the project's data.

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 to 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.**

The ARMP project is committed to ensuring ethical data practices, privacy protection, and compliance with all relevant standards for data governance. The project does not collect personally identifiable information (PII) or sensitive data, as all data gathered remains anonymous. No personal data was recorded or shared, and it is not anticipated that any sensitive data will be collected. The primary datasets include GTFS-Flex transit data feeds, MiTripPlanner code, media assets, site visit presentations, rider survey data, the project website and newsletter, marketing kits for transit agencies, and industry presentations. These datasets are structured to be publicly accessible while adhering to ethical guidelines.

To protect privacy and confidentiality, the project follows all local, state, and federal standards for ethical data handling, including the Privacy Act of 1974 and DOT privacy risk management policies. Since no PII is collected, an Institutional Review Board (IRB) review is not required. However, the project team ensures that all transit agencies participating in the initiative provide informed consent for data use and sharing. The project also aligns with CARE Principles for Indigenous Data Governance when working with Indigenous populations or Tribal communities, ensuring that data is handled responsibly and with respect for community interests.

Given that all data is anonymized, there are no restrictions on access beyond ensuring that non-sensitive data is made publicly available following open standards. Transit agencies participating in the project agree to share their data in compliance with these principles. The project team is responsible for stewarding and protecting the data, ensuring that it is stored, transferred, and archived securely. Project staff oversee data governance, ensuring compliance with ethical norms and best practices.

Overall, the ARMP project prioritizes transparency, accessibility, and ethical data management, ensuring that all collected information contributes to improving rural mobility while safeguarding privacy and confidentiality.

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.**

The intellectual property rights for data created in the ARMP project will be held by the project team and participating transit agencies. However, as a US DOT SMART Grant-funded project, all outputs—including datasets, publications, and software—comply with the US DOT Public Access Plan and will be made publicly available.

Where appropriate, intellectual property rights may be transferred to a data archive to ensure long-term accessibility. The data will generally follow open data principles, with GTFS-Flex data feeds, MiTripPlanner code, rider survey data, and other non-sensitive datasets made publicly available under an open license. The project also follows local, state, and federal data protection standards to ensure all shared data remains anonymized and non-sensitive.

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.**

The dataset will be archived in the National Transportation Library Repository and Open Science Access Portal (ROSA P) at <https://doi.org/10.21949/z23z-bb22>. Prior to archiving, the data are stored on the secured BTS networks and

drives, which are backed up nightly. The US DOT systems are secured from outside users and backed up daily.

Files in ROSA P are backed up in NTL drives at US DOT, daily; at the Centers for Disease Control, the repository managing facility, daily; and in Amazon Web Service Cloud servers in Virginia and Oregon daily.

The dataset will be retained in perpetuity.

NTL staff will create persistent Digital Object Identifiers (DOIs) for each dataset stored in ROSA P. These DOIs will be associated with dataset documentation as soon as they become available for use.

Planned Research Outputs

Data paper - "Advancing Rural Mobility Implementation Report"

The Implementation report will document lessons learned, key project findings and next steps.

Dataset - "Project Dataset "

The ARMP project data resulted in various file formats to ensure accessibility and usability across different platforms. The GTFS-Flex data feeds is primarily stored in CSV and TXT formats, following the General Transit Feed Specification (GTFS) standard, which is widely used in the transit industry for structured transit data. The MiTripPlanner front-end web application code is written primarily in typescript - an open source programming language developed by Microsoft. Other project-related materials, such as the news release, photos, and videos from the Charlevoix media conference, will be provided in PDF, JPEG, PNG, and MP4 formats. Project presentations are available in PDF formats. The rider survey data will be stored in PDF formats for easy analysis. The project website and newsletter will be accessible in HTML and PDF, while the marketing kits for transit agencies will be distributed in PDF and PNG formats.

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?
Advancing Rural Mobility Implementation Report	Data paper	2024-08-15	Open	None specified		Creative Commons Attribution Share Alike 4.0 International	None specified	No	No
Project Dataset	Dataset	2025-06-12	Open	National Transportation Library Repository and Open Science Access Portal		Creative Commons Attribution 4.0 International	DCAT-US	No	No