DATA MANAGEMENT PLAN AND METADATA SCHEMA TEMPLATE

Data Management Plan

Name of Contractor: University of Wyoming
Name of the Project: Review of the Current Transportation Asset Management Practices in Wyoming
Project Duration: Start Date: March 14, 2022 **End Date**: March 14, 2025
DMP Version:
Date Amended, if any:
Name and ORCID (Open Researcher and Contributor Identifier) Number for each author:
WYDOT Project Number:
Waleed Aleadelat, ORCID: 0000-0003-2127-5057
Khaled Ksaibati, ORCID: 0000-0002-9241-1792
Shahbaz Khan, ORCID: 0000-0002-3484-4530
Bernard Boake, ORCID: 0009-0005-9223-4430
Lucille (Luci) Cawley, ORCID: 0000-0002-8528-2021
Marwan Hafez, ORCID: 0000-0002-0303-7867

Introduction

What constitutes data will be determined by the Principal Investigator (PI), Project Champion, and the Research Manager. In general, your plan should address the data in the final research project.

The following forms of data/datasets should be reviewed when determining what data should be archived and listed in this Data Management Plan (DMP):

a) Primary data used in the production of the report: Raw, verified data that has been obtained directly from a source. It can be captured through experiments, surveys, interviews, focus groups, or other direct interactions with individuals in the field. Does not include analysis data.
b) Unpublished datasets: Materials and methods; clear description of the variables presented; supported by unpublished reports; and any other relevant material.
c) Secondary Data: Pre-existing data not gathered or collected by the authors. Usually collected by another organizations or source.
d) Metadata: Set of data that describes and gives information about the dataset – cataloging information.
e) Dataset description document: Describes all variables in the dataset and the measurement units used.
f) Codebook: A list of variable names, variable labels, and label values. Should specify the data position of each variable, describe the contents of each variable, and identify the range of possible codes and the meanings of the codes.
g) Questionnaires: An unused copy of the questionnaire.
h) Handbooks, guides, and manuals derived from research.

Determination of what counts as data and what should be archived will depend on the Principal

Investigator's knowledge of the data and what he/she believes is valuable. As part of the research project, your DMP should address unique data that may arise from your research.

Data that does not need to be archived or saved includes preliminary analyses of a project, drafts, plans for future research, peer reviews, interoffice communications, emails, letters, or other forms of correspondence. The Principal Investigator and the Project Champion will have the opportunity to discuss what data and/or other digital material should be excluded prior to finalizing the project.

WYDOT expects the timely release and sharing of data to be no later than the acceptance for publication of the main findings from the final dataset, and there will be no embargo period approved for this project without prior approval from the Research Center.

Definitions

a) Code or scripts include code used in the collection, manipulation, processing, analysis or visualization of data, but may also include software developed for other purposes.
b) Copyright is a set of legal rights extended to copyright owners that govern such activities as reproducing, distributing, adapting, or exhibiting original works fixed in tangible forms.
c) Data means 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, communications with colleagues. Recorded material excludes physical objects (e.g. laboratory samples). Research data also does not include trade secrets, commercial information, materials necessary to be held confidential, and personnel and medical information, including information that falls under the HIPAA and PII confidentiality impact levels, and similar information the disclosure of which would constitute a clearly unwarranted invasion of personal privacy. Any information that falls under this definition shall not be considered open source and shall not be publically available, Data Archive is a site where machine-readable materials are stored, preserved or possibly redistributed to individuals interested in the materials.
d) Data Management Plan is a document that specifies your plans for managing your data and files for a research project.
e) Dataset means collection of data.
f) Metadata refers to structured data about data that helps define administrative, technical, or structural characteristics of the digital content.

I. For peer reviewed publications provide the following for each:

A. Name of all peer reviewed publications that have been generated using data from this project:
Paper accepted and presented in TRB titled "Optimizing Pavement Maintenance and Rehabilitation Strategies for Large-Scale Pavement Networks: A Case Study of Wyoming". Another paper is published in a journal titled "Network-level Pavement Maintenance and Rehabilitation Planning Using Genetic Algorithm".

B. Any Digital Object Identifier (DOI) assigned to any peer reviewed publication or data generated by this project:
https://doi.org/10.1007/s41062-024-01534-1

C. All persistent uniform resource locators (URLs) for all peer reviewed publications that have been generated using data from this project:
https://link.springer.com/article/10.1007/s41062-024-01534-1#citeas

D. Dataset URLs, if available:

II. The purpose of this research project is to:

The main objectives of this research are to:

- Nationwide survey data collection on the best management practices of transportation assets by state DOTs.
- Analysis of central-funding and decision-making for pavement managed by WYDOT.
- Analysis of central-funding and decision-making for bridges managed by WYDOT.
- Trade-off analysis of the different transportation assets managed by WYDOT.
- Evaluating the historical funding records for pavement assets managed by WYDOT.

III. Data Types and Storage

The types of data and/or datasets generated and/or used in this project include:

The procedure of creating a comprehensive transportation asset management plan involves four distinct steps. These steps are as follows:
Pavement network optimization. The data collected includes pavement conditions, traffic conditions, cost of maintenance and budget of each maintenance for a pavement network optimization.
Bridge network optimization. The data collected includes bridge conditions, traffic conditions, cost of maintenance and budget of each maintenance for a bridge network optimization.
Cross asset management optimization. The data collected includes different assets together for a cross asset management optimization
Historical funding distribution. The data includes the historical pavement maintenance funding by WYDOT.

NOTE: Provide a description of the data that you will gather in the course of your project. Address the nature, scope, and scale of the data that will be collected. 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. Discuss value of the data over the long-term. Provide the name of all repositories where the data will be housed during the

lifetime of the project.

Checklist

- What type of data will be produced?
- How will data be collected? In what formats?
- How will the data collection be documented?
- Will it be reproducible? What would happen if it got lost or became unusable later?
- How much data will it be, and at what growth rate? How often will it change?
- Are there tools or software needed to create/process/visualize the data?
- Will you use pre-existing data? From where?
- Storage and backup strategy?

IV. Data Organization, Documentation, and Metadata

The plan for organizing, documenting, and using descriptive metadata to assure quality control and reproducibility of these data includes:

The data formats are in Microsoft Excel spreadsheets which can be readily imported by different software packages. The data is prepared and stored in accordance with the SQL standard. Data validation is done by applying identifiable rules that will ensure that the data is accurate. Data that does not meet minimum levels of acceptable quality will be flagged and removed from the database.

NOTE: Your DMP should describe the anticipated formats for data and related files. To the maximum extent practicable, and in accordance with generally accepted practices in your field, your DMP should address how you will use platform-independent and non-proprietary formats to ensure maximum utility of the data in the future. If you are unable to use platform-independent and non-proprietary formats, specify the standards and formats that will be used and the rationale for using those standards and formats.

NOTE: Attach the Metadata Schema URL for data generated, and all peer reviewed publications from this project.

Checklist

- What standards will be used for documentation and metadata?
- Is there good project and data documentation format/standard?
- What directory and file naming convention will be used?

- What project and data identifiers will be assigned?
- Is there a community standard for metadata sharing/integration?

V. Data and/or Database Access and Intellectual Property

What access and ownership concerns are there?

No intellectual property issues are expected to exist with the data collected for this project. Wyoming highway datasets that will be collected are owned by WYDOT, who will restrict personal and business details from the database. WYDOT will be responsible for allowing access to the data for processing and modification, and for access restrictions. Access to the data will involve seeking authorization from WYDOT, who may then allow access to a part or the whole dataset. There will be no embargo periods to uphold.

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. If necessary, describe any division of responsibilities for stewarding and protecting the data among Principal Investigators.

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. 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.

Checklist

- What steps will be taken to protect privacy, security, confidentiality, intellectual property or other rights?
- Does your data have any access concerns? Describe the process someone would take to access your data.
- Who controls it (e.g., PI, student, lab, University, funder)?
- Any special privacy or security requirements (e.g., personal data, high-security data)?
- Any embargo periods to uphold?

VI. Data Sharing and Reuse

The data will be released for sharing in the following way.

Property rights for the data created will be held by WYDOT. The data will be collected and stored in a manner that allows for the efficient sharing of the data with other interested parties. The data will be submitted to WYDOT after the project and may be shared through the WYDOT Data Warehouse. Interested parties will seek permission for access to the data from WYDOT. The data will be available for access and modification using standard data and statistical software such as Microsoft Excel and R.

Describe who will hold the intellectual property rights for the data created by your project. Describe whether you will transfer those rights to a data archive, if appropriate. Identify whether any copyrights apply to the data, as might be the case when using copyrighted instruments. If you will be enforcing terms of use or a requirement for data citation through a license, indicate as much in your DMP. Describe any other legal requirements that might need to be addressed.

Checklist

- If you allow others to reuse your data, how will the data be discovered and shared?
- Any sharing requirements (e.g., funder data sharing policy)?
- Audience for reuse? Who will use it now? Who will use it later?
- When will I publish it and where?
- Tools/software needed to work with data?

VI. Data Preservation and Archiving

The data will be preserved and archived in the following way(s).

After the completion of the project, the data is saved on the Warehouse for 3 years at the University of Wyoming.

Describe how you intend to archive your data and why you have chosen that particular option. You may select from a variety of options including, but not limited to:

- Use of an institutional repository.
- Use of an archive or other community-accepted data storage facility.
- Self-dissemination.

You must describe the dataset that is being archived with a minimum amount of metadata that ensures its discoverability. Whatever archive option you choose, that archive must support the capture and provision of the National Transportation Library metadata requirements. In addition, the archive you choose must support the creation and maintenance of persistent identifiers and must provide for maintenance of those identifiers throughout the preservation lifecycle of the data. Your plan should address how your archiving and preservation choices

meet these requirements.

Checklist

- How will the data be archived for preservation and long-term access?
- How long should it be retained (e.g., 3-5 years, 10-20 years, permanently)?
- What file formats? Are they long-lived?
- Are there data archives that my data is appropriate for (subject-based? Or institutional)?
- Who will maintain my data for the long-term?

NOTE:
Researchers evaluating data repositories as the option(s) for storing and preserving their data should examine evidence demonstrating that the repository:

a. Promotes an explicit mission of digital data archiving.
b. Ensures compliance with legal regulations, and maintains all applicable licenses covering data access and use, including, if applicable, mechanisms to protect privacy rights and maintain the confidentiality of respondents.
c. Has a documented plan for long-term preservation of its holdings.
d. Applies documented processes and procedures in managing data storage.
e. Performs archiving according to explicit workflows across the data life cycle.
f. Enables the users to discover and use the data, and refer to them in a persistent way through proper citation.
g. Enables reuse of data, ensuring appropriate formats and application of metadata.
h. Ensures the integrity and authenticity of the data.
i. Is adequately funded and staffed, and has a system of governance in place to support its mission.
j. Possesses a technical infrastructure that explicitly supports the tasks and functions described in internationally accepted archival standards like Open Archival Information System (OAIS).

VII. Generative Artificial Intelligence (AI) Tools

Authors who use AI tools in the writing of a manuscript, production of images or graphical elements of any report, or in the collection and analysis of data, must be transparent in disclosing in the materials and methods (or similar section) of the report how the AI tool was used. Authors are fully responsible for the content of their manuscript, even those parts produced by an AI tool, and are thus liable for any breach of publication ethics.

Specifically, you must set out the following:

- **Clearly indicate the use of language models in the manuscript,** including which model was used and for what purpose. Please use the methods or acknowledgements section, as appropriate.

  Generative language models such as ChatGPT and google Gemini to assist in writing the manuscript.

- **Verify the accuracy, validity, and appropriateness of the content** and any citations generated by language models and correct any errors or inconsistencies.

  The content was written by the author and refined by the language models and all citations were generated manually.

- **Provide a list of sources used to generate content** and citations, including those generated by language models. Double-check citations to ensure they are accurate and properly referenced.

  Content and citation were developed manually using existing literature.

- **Be conscious of the potential for plagiarism** where the LLM may have reproduced substantial text from other sources. Check the original sources to be sure you are not plagiarizing someone else's work.

  LLM was not used to generate text.

- **Acknowledge the limitations of language models in the manuscript,** including the potential for bias, errors, and gaps in knowledge.

  This is not applicable for the present study.

Please note that AI bots, such as ChatGPT should not be listed as an author on your submission.

Metadata Schema

Fill out all information in the below table.

Elements	Example of what is expected for each element
Title[1]	Review of the Current Transportation Asset Management Practices in Wyoming
Creator/contact point	Khaled Ksaibati, ORCID: 0000-0002-9241-1792
Publication Date(s)	July 10th, 2025
Description/Abstract	WYDOT maintains a vast network consisting of Interstate, NHS and non-NHS roads. The existing infrastructure needs continues maintenance to keep in good condition and safe for users. To achieve this, WYDOT needs a considerable amount of funding every year and anticipates the demand for future maintenance through performance models. This future maintenance requires dedicated funds and strategy for the long-term performance for WYDOT infrastructures. This report seeks to achieve this by proposing various supportive techniques to managing WYDOTs asset management programs using various analysis procedures. Firstly, a nationwide survey is conducted amongst all the state DOTs in the country to gather the best asset management practices from industry experts. Separate management analysis is then done for pavement and bridges using both multi-criteria and single-objective techniques based on asset data provided by WYDOT. A combined cross asset management of different transportation assets including ancillary items is also further examined. Finally, the historical funding distributions of pavement assets over a 10-year period are also analyzed to show the spending allocation trends of WYDOT. This report seeks to provide WYDOT and other relevant state agencies with supportive future expenditure management practices that would allow them to dedicate funding to the maintenance, rehabilitation, and reconstruction of projects using the existing performance models.
Subject and Keywords	Subject: Transportation Asset Management Keywords: Cross Transport Asset Management; Pavements; Bridge; Optimization
Identifier[2] **and/or source**	
Collection and Related Documents	
Edition	July 10th, 2025
Related Documents	
Coverage	Wyoming
Language	English
Publisher/Distributor	FHWA and Wyoming Department of Transportation
Funding agency	FHWA and Wyoming Department of Transportation
Access Restrictions	Public data

Intellectual Property and Other Rights	No restriction
License	Public Domains
Code and software needs	Microsoft Excel
Format	.xlsx
Choice of Repository	Public data
Reading Level Success Criterion 3.1.5 for Section 508 compliance	This research report is written in a highly technical language; the Table of Contents is written in plain language, and the tables are figures are 508 compliant.

NOTE: Each separate report, dataset, collection, existing collection, and software developed must have its own table. All fields in this Schema must be completed at the time of the final report.
NOTE: This Metadata Schema is created as a derivative from the Common Core required fields that can be found at https://project-open-data.cio.gov/schema/

[1] *To include alternate title; conference title; and journal title, if they are different.*
[2] *To include record numbers; report numbers; NTIS number; TRIS Accession Number; OCLC Number; ISBN; ISSN; contract number; and DOI if available.*