Data Management Plan

Name of Contractor	**University of Wyoming**
Name of project	**Driver Performance and Behavior in Adverse Weather Conditions: An Investigation Using the SHRP2 Naturalistic Driving Study Data—Phase 2**
Project Duration	**Start date : March 2016** **End: March 2018**
DMP Version	
Date Amended, if any	
Name of all authors, and ORCID number for each author	Mohamed Ahmed, Ph.D., P.E. (orcid.org/0000-0002-1921-0724) Ali Ghasemzadeh, Ph.D. (orcid.org/0000-0003-1232-251X) Britton Hammit, Ph.D. Candidate (orcid.org/0000-0002-8859-9075) Nasim Khan, MS Student (orcid.org/0000-0001-5996-091X) Anik Das, MS Student (orcid.org/0000-0003-4674-5334) Elhashemi Ali, Ph.D. Candidate (orcid.org/0000-0001-7643-7143) Rhonda K. Young, Ph.D., P.E. (orcid.org/0000-0001-6745-5008) Hesham Eldeeb, Ph.D. (HIS) (orcid.org/0000-0001-9669-5991)
WYDOT Project Number	WY- 18/05F
Name of all peer reviewed publications which have been generated using data from this project	1. Ghasemzadeh, A., and M. M. Ahmed. Drivers' Lane-Keeping Ability in Heavy Rain: Preliminary Investigation Using SHRP 2 Naturalistic Driving Study Data. *Transportation Research Record: Journal of the Transportation Research Board*, No. 2663, 2017, pp. 99–108. 2. Ahmed, M. M., R. K. Young, A. Ghasemzadeh, B. Hammit, A. Elhashemi, N. Khan, A. Das, and H. Eldeeb. *Implementation of SHRP2 results within the Wyoming connected vehicle variable speed limit system: Phase 2 early findings*

report and phase 3 proposal. 2017.
3. Ghasemzadeh, A., and M. M. Ahmed. A Probit-Decision Tree Approach to Analyze Effects of Adverse Weather Conditions on Work Zone Crash Severity Using Second Strategic Highway Research Program Roadway Information Dataset. 96th Transportation Research Board Annual Meeting. Washington, D.C. 2017.
4. Ahmed, M. M., A. Ghasemzadeh, H. Eldeeb, S. Gaweesh, J. Clapp, K. Ksaibati, and R. Young. *Driver Performance and Behavior in Adverse Weather Conditions: An Investigation Using the SHRP2 Naturalistic Driving Study Data—Phase 1. DOI: 10.13140/RG.2.2.24061.05602*. 2015.
5. Ghasemzadeh, A., B. Hammit, M. M. Ahmed, and H. Eldeeb. Complementary Methodologies to Identify Weather Conditions in Naturalistic Driving Study Trips: Lessons Learned from the SHRP2 Naturalistic Driving Study & Roadway Information Database. 97th Transportation Research Board Annual Meeting. Washington, D.C. 2018.
6. Ghasemzadeh, A., and M. Ahmed. Multivariate Adaptive Regression Splines and Logistic Regression Models to Identify the Impact of Rainy Weather on Driver Lane-keeping Performance Considering Driver Demographics and Roadway Characteristics Using SHRP2 Naturalistic Driving Data. 2018.
7. Ghasemzadeh, A., B. Hammit, M. Ahmed, and R. Young. Using Parametric Ordinal Logistic Regression and Non-Parametric Decision Tree Approaches for Assessing the Impact of Weather Conditions on Driver Speed Selection Using Naturalistic Driving Data. *Transportation Research Record: Journal of the Transportation Research Board*, Vol. In press, 2018.
8. Ghasemzadeh, A., and M. M. Ahmed. Crash Characteristics and Injury Severity

at Work Zones Considering Adverse Weather Conditions in Washington Using SHRP 2 Roadway Information Database. 95th Transportation Research Board Annual Meeting. Washington, D.C. 2016.
9. Hammit, B., A. Ghasemzadeh, M. M. Ahmed, and R. K. Young. Evaluation of Weather-Related Freeway Car-Following Behavior using the SHRP2 Naturalistic Driving Study. 97th Transportation Research Board Annual Meeting. Washington, D.C. 2018.
10. Ghasemzadeh, A., and M. M. Ahmed. A Tree-Based Ordered Probit Approach to Identify Factors Affecting Work Zone Weather-Related Crashes Severity in North Carolina Using the Highway Safety Information System Dataset. 96th Transportation Research Board Annual Meeting. Washington, D.C. 2017.
11. Khan, N. M., A. Ghasemzadeh, and M. M. Ahmed. Investigating the Impact of Fog on Freeway Speed Selection Using the SHRP2 Naturalistic Driving Study Data. *Transportation Research Record: Journal of the Transportation Research Board*, Vol. In press, 2018.
12. Das, A., A. Ghasemzadeh, and M. M. Ahmed. A Comprehensive Analysis of Driver Lane-Keeping Performance in Fog Weather Conditions Using the SHRP2 Naturalistic Driving Study Data. 2018.
13. Ahmed, M. M., and A. Ghasemzadeh. Exploring the Impacts of Adverse Weather Conditions on Speed and Headway Behaviors Using the SHRP2 Naturalistic Driving Study Data. 96th Transportation Research Board Annual Meeting, 2017. 2017.
14. Ghasemzadeh, A., and M. M. Ahmed. Utilizing naturalistic driving data for in-depth analysis of driver lane-keeping behavior in rain: Non-parametric MARS and parametric logistic regression modeling approaches. *Transportation*

	Research Part C, Vol. In press, 2018. 1.
Any Digital Object Identifier (DOI), including any CROSSREF number, which has been assigned to any peer reviewed publication or data generated by this project	http://dx.doi.org/10.3141/2663-13 http://dx.doi.org/10.1177/0361198118758035 http://dx.doi.org/10.1016/j.trc.2018.03.018
URLs for all peer reviewed publications which have been generated using data from this project	http://trrjournalonline.trb.org/doi/abs/10.3141/2663-13 https://trid.trb.org/view/1437513 https://trid.trb.org/view/1495703 https://trid.trb.org/view/1497181 https://trid.trb.org/view/1494970 https://trid.trb.org/view/1494546 https://trid.trb.org/view/1497201 https://trid.trb.org/view/1439711 https://trid.trb.org/view/1439630 https://trid.trb.org/view/1392336 https://trid.trb.org/view/1415699 http://shrp2.transportation.org/Documents/Safety/05-Phase2EarlyResearchFindings-WY-Weather.pdf
RiP RH Display ID Number	
Dataset URL, if available	

What constitutes data will be determined by the Principle Investigator, Project Champion, and the Research Manager. In general, your plan should address final research data. This includes recorded factual material commonly accepted in the scientific community as necessary to validate research findings. Final research data do not include laboratory notebooks, partial datasets, preliminary analyses, drafts of scientific papers, plans for future research, peer review reports, communications with colleagues, or physical objects, such as gels or laboratory specimens. See Chapter 12, subsections 12.2 and 12.3 for full details on what data to retain. As part of your research, you may also generate unique data, which are data that cannot be readily replicated. Your DMP should also address unique data that may arise from your research.

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, unless the Principle Investigator will be embargoing the data. In such a case, the data cannot be embargoed for a period longer than 12 months. See Chapter 12, subsection 12.13 and 12.14 for information on retention and embargos.

1. Introduction

The purpose of this research project is to:

The primary goal of the Wyoming SHRP2 IAP project is to leverage the SHRP2 NDS and RID databases to enhance the understanding of how drivers respond to adverse weather and road conditions. The ultimate objective of this research is to use the findings to develop feasible countermeasures that WYDOT can implement on state interstates and highways to improve the reliability of the transportation network VSL systems during adverse weather conditions.
The original project proposal presented the following research questions:
1. Can NDS trips occurring in inclement weather be identified efficiently and effectively using available NDS and RID data?
2. Can driver behavior (e.g., speed selection, car-following, and lane wandering) during inclement weather conditions be characterized efficiently from the NDS data?
3. What are the best surrogate measures for weather-related crashes that can be identified using the NDS data?
4. What type of analysis can be performed and conclusions drawn from the resulting dataset?

2. 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 and similar information the disclosure of which would constitute a clearly unwarranted invasion of personal privacy.
d. Data Archive is a site where machine-readable materials are stored, preserved or possibly redistributed to individuals interested in the materials.
e. Data Management Plan is a document that specifies your plans for managing your data and files for a research project.
f. Dataset means collection of data.
g. Metadata refers to structured data about data that helps define administrative, technical, or structural characteristics of the digital content.

3. Data Types and Storage

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

This study utilized data from the SHRP2 safety data, which is comprised of two main datasets including Naturalistic Driving Study (NDS) data and the Roadway Information Database (RID). NDS refers to the real-time behavioral data that was collected from more than 3,400 volunteer drivers in six site locations including New York, Florida, Washington, Indiana, Pennsylvania, and North Carolina. Each driver received an annual incentive of $500 and was asked to provide access to his/her vehicle in order to remove and replace the hard drive with the collected data. Some participants remained in the study for several months and then replaced with new drivers and vehicles, while others remained for the full duration of the study. The project duration was about 39 months (October 2010-December 2013) and during this time drivers' behavior and their vehicle-use were recorded continuously, which made this project one of the largest naturalistic driving studies worldwide.

The Data Acquisition System (DAS) was developed by VTTI (Virginia Tech Transportation Institute), which includes four cameras, forward radar, GPS, vehicle network information, on-board computer vision lane tracking system, as well as other computer vision systems and capability of storing data. The collected NDS data includes but not limited to vehicle kinematics and network information recorded at 10 HZ frequency (10 observations per second) such as speed, acceleration, and steering wheel position; video views including forward and rear views, and driver's face and hands view, in addition to information obtained from a variety of sensors such as forward radar, alcohol sensor, and accelerometers.

Roadway Information Database (RID) was also utilized in this study, which was developed by the Center for Transportation Research and Education (CTRE) of Iowa State University. The RID contains a comprehensive description of roadway characteristics for the six NDS states. The roadway data was collected not only from the state highway department, but also from other sources including equipped vans that collected roadway characteristics while traveling at posted speed limits on most traversed routes by SHRP2 vehicles. The NDS and RID data were linked in this study to associate driver lane-keeping behavior with roadway characteristics.

Weather data extracted from the National Climate Data Center (NCDC) were also utilized in this study to initially identify NDS trips occurred in adverse weather conditions. It is worth noting that SHRP2 do not have weather conditions reported for the NDS normal driving trips, weather conditions are only provided for SHRP2 event data (i.e., crashes and near crashes). The NCDC archives weather information from different weather stations at a national level and provide public access to the historical weather data.

Provide a description of the data that you will be gathering in the course of your project. You should 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. Please 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?

4. Data Organization, Documentation and Metadata

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

The data acquired in Phase 1 and 2 is stored on secured UW computers in keyed graduate student offices. The data are encrypted while not in use.

Your DMP should describe the anticipated formats that your data and related files will use. 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, you should 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?

5. Data and/or Database Access and Intellectual Property

What access and ownership concerns are there…

All data are housed at secured computers that only accessible by the co-PIs and graduate students working on this project. UW implement a rigorous security measures. No identifiable personal information was acquired/collected in this project.

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?

6. Data Sharing and Reuse

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

The University of Wyoming will hold the intellectual property rights for the data. The PI and graduate students, from the civil engineering department, whom are responsible for the study are currently using the different datasets. Data will be used under supervision of the PI. Data will not be accessed or used by other parties other than university of Wyoming.

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?

7. Data Preservation and Archiving

The data will be preserved and archived in the following ways …

Collected data will be retained for 5 years after the end date of the project (Phase 3).

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).

NOTE: This DMP is created as a derivative from the DMP belonging to the University of Minnesota and can be found at https://www.lib.umn.edu/datamanagement/DMP

REV: 8-2017

Metadata Schema

Title[1]	Human-readable name of the asset. Should be in plain English and include sufficient detail to facilitate search and discovery. A name given to the publication or data element. All substitute or alternative titles must have a different Metadata Transmittal Schema.
Creator/contact point	An entity/person(s) primarily responsible for making the content of the resource. Contact person's name, ORCID number, and email for the asset.
Publication Date(s)	The date associated with the final report/dataset.
Description/Abstract	Human-readable description (e.g., an abstract) with sufficient detail to enable a user to quickly understand whether the asset is of interest. May include abstract, table of contents, reference to a graphical representation of content or a free text account of the content.
Subject and Keywords	The topic of the content of the resource. Tags (or keywords) help users discover your dataset; please include terms that would be used by technical and non-technical users.
Identifier[2] **and/or source**	A unique identifier for the dataset/publication. Examples: URI, URL, DOI, ISNB, ISSN.
Collection and Related Documents	If there is a secondary dataset, cite source. The collection of which the dataset is a subset should be listed. Include all identifiers and/or sources.
Edition	Most recent date on which the dataset was changed, updated or modified.
Related Documents	Related documents such as technical information about a dataset, developer documentation, etc.
Coverage	Spatial location, temporal period, jurisdiction.
Language	The language of the dataset/publication.
Publisher/Distributor	FHWA and Wyoming Department of Transportation List all other publishing companies that this publication has been sent to.

[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.*

Funding agency	FHWA and Wyoming Department of Transportation
Access Restrictions	The degree to which this dataset could be made publicly available, *regardless of whether it has been made available*. Choices: public (Data asset is or could be made publicly available to all without restrictions), restricted public (Data asset is available under certain use restrictions), or non-public (Data asset is not available to members of the public).
Intellectual Property and Other Rights	This may include information regarding access or restrictions based on privacy, security, or other policies. This should also serve as an explanation for the selected "accessLevel" including instructions for how to access a restricted file, if applicable, or explanation for why a "non-public" or "restricted public" data asset is not "public," if applicable.
License	The license or non-license (i.e. Public Domain) status with which the dataset or API has been published.
Code and software needs	List all code specific information. Is there specific software needed to run the database or data.
Format	The machine-readable file format. May include media type or dimensions. Used to determine the software, hardware or other equipment needed to display or operate the resources.
Choice of Repository	If you have a preference, list the repository where you will archive your data/datasets.

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 which can be found at https://project-open-data.cio.gov/schema/.

REV: 8-2017