Data Management Plan

Name of Contractor: University of Wyoming
Name of the Project: Connected Autonomous Traffic Signal Control Algorithms and Fleet Vehicles
Project Duration: 32.5 months Start Date: 12/31/2018 End Date: 9/17/2021
DMP Version: 2
Date Amended, if any: 12/03/2021
Name of all authors, and ORCID number for each:

Milan Zlatkovic, ORCID: 0000-0002-6777-230X
Mohamed Ahmed, ORCID: 0000-0002-1921-0724
Zorica Cvijovic, ORCID: 0000-0003-3558-7967
Sara Bashir, ORCID: 0000-0003-4328-4552

WYDOT Project Number: RS03219

- **Name of all peer reviewed publications, which have been generated using data from this project to include:**

  Assessment of Queue Warning Applications on Signalized Intersections for Connected Freight Vehicles
  DOI: 10.1177/03611981211015247
  URL: https://journals.sagepub.com/doi/10.1177/03611981211015247

  Multi-Level Conditional Transit Signal Priority in Connected Vehicle Environments
  DOI: 10.31075/PIS.67.02.01
  URL: https://www.putisaobracaj.rs/index.php/PiS/article/view/182

  Conditional Transit Signal Priority for Connected Transit Vehicles
  DOI: 10.1177/03611981211044459
  URL: https://journals.sagepub.com/doi/10.1177/03611981211044459

- **Dataset URL, if available:**
  N/A

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. 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 11 for information on retention and embargos.

1. Introduction

The purpose of this research project is to:

The purpose of the research project is to create field-ready, Connected-Autonomous Vehicles (CAV)-based traffic control programs that will improve operations and safety of trucks and fleet vehicles. The study reviews the current protocols and recommends options applicable to Wyoming conditions. It analyzes, assesses and develops traffic control algorithms that use CAV technologies to improve operations and safety of signalized intersections. The focus is on optimizing operations of freight and fleet vehicles through signalized intersections. The analyzed strategies include freight priority, queue warning and speed harmonization.

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 …

The data for this project include geometrical characteristics of the six selected interchanges on I-80 (in Evanston, Rock Springs, Rawlins, Laramie and Cheyenne), such as horizontal and vertical alignments, cross section elements, roadside elements, ramp characteristics, auxiliary lane dimensions; operational characteristics (traffic volumes, traffic composition, speeds); and signal timing data.
The geometry data was collected from the WYDOT's Roadway database, GIS maps, Google Earth and field visits. These data are used to create and exist in traffic simulation models. Geometry data is quite stable over time with minimal changes.
Operational data were collected through available WYDOT databases and field observations. They are summarized in spreadsheet, and exist in the developed models. The dataset includes historical operational data, which is used to assess the changes over time and perform projections for the future.
Signal timing data have been obtained from WYDOT in a form of Synchro models. These data are replicated in the simulation models. Since the team used existing databases and field data collection, the datasets are reproducible.
Spreadsheet software (such as Excel and Access) are be the main tools needed to create, process and visualize the data. PTV VISSIM with Econolite ASC/3 external controller functionality and Python installation is required to run the simulation models.
The data are stored on UW team computers and UW's OneDrive, which have excellent protection. Backup data are stored to external devices (USBs) which are only accessible to the research team.

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 majority of data collected, redacted and analyzed in this research are in non-proprietary forms (spreadsheets, GIS and KML types). Some data are duplicated and analyzed in proprietary software (PTV VISSIM and Synchro) for simulation and analysis. The simulation results are presented in the same spreadsheet and software forms for easy access by wider audiences.
The metadata scheme contains all information about the data collection procedures, storage, file types and naming conventions. It is in a text format with descriptions and explanations.
The naming convention is designed to be self-explanatory, so that the names of the directories and files correspond to the contents in the data file.
The UW team was guided by our experience and lessons learned about data organization from previous research projects.

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…

This research project did not use nor create data which contain confidential information. Nevertheless, access to the data is limited to the UW team (PIs and the students working on the project). The data are stored on the UW team computers and UW OneDrive, which are accessible only to the individual that owns the account, or to whom the permission has been given. Personal and high-security data were not used in this research. The UW team and WYDOT maintain the intellectual rights to data analysis results. There are no embargo periods for the data 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?

6. Data Sharing and Reuse

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

The complete data sets collected in the study will only be shared with WYDOT. WYDOT can then reuse the data for subsequent analysis or share with third parties for future studies. The UW team may also use the collected data for future research. There is not a special data archive for this project.
Reduced data sets and analysis results have been included in journal publications and conference presentations. Research papers were submitted to transportation-oriented journals throughout the research.
We do not foresee any special requirements for data sharing. We also do not expect to have any copyrights or legal requirements for the data. The software needed to work with the data will be spreadsheets and map tools, VISSIM simulation and Python programming language.

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 …

The complete data are kept on UW researchers' computers and the UW OneDrive with protected access, and shared with WYDOT. Redacted data are stored on external storage devices in the UW team research labs. We expect to keep copies of the datasets for about 10 years upon completion of the project. The formats are be in .xlsx type (spreadsheet), GIS, KML, VISSIM and Python.
The team will share reduced and analyzed data through journal and other types of publications, which will be available to wider audiences.

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

Metadata Schema

Elements	Example of what is expected for each element
Title[1]	**Connected Autonomous Traffic Signal Control Algorithms and Fleet Vehicles**
Creator/contact point	**Milan Zlatkovic, ORCID: 0000-0002-6777-230X** **Mohamed Ahmed, ORCID: 0000-0002-1921-0724** **Zorica Cvijovic, ORCID: 0000-0003-3558-7967** **Sara Bashir, ORCID: 0000-0003-4328-4552**
Publication Date(s)	December 2021
Description/Abstract	Connected and Autonomous Vehicle (CAV) technologies enable communication among vehicles, and vehicles and infrastructure, paving the way for multiple safety and operational applications. This research developed and tested traffic signal control algorithms and control programs which utilized CAV-equipped heavy trucks and traffic signals. The focus of the study was on Intelligent Traffic Signals (ISIG), Freight Signal Priority (FSP), Queue Warning (Q-WARN), and Speed Harmonization (SPD-HARM) applications. The application, testing and analysis were performed through Traffic In Cities Simulation Model (VISSIM) microsimulation software, coupled with real-world traffic control software (Econolite ASC/3). Communication and information sharing was modeled using Python programming language, while signal control applications was programmed directly in the signal control software. The test-case networks included six signalized intersections adjacent to I-80, in Wyoming. Model scenarios included different rates of CAV-equipped trucks. Additionally, the study tested CAV-based Transit Signal Priority using a test-network in Salt Lake City, Utah. First, the study developed a communication process which uses latitude/longitude coordinates of CAVs and intersections to define the detection zone and enabled information sharing. Once the communication was established, CAV application can be called and implemented as needed. FSP provides extra time for trucks approaching a signalized intersection, therefore, minimizing delays. The implementation of FSP has the potential to reduce truck intersection delays between 10 and 70 percent, with minimal impacts on other traffic. Q-WARN can warn drivers on a downstream queue of the conditions

[1] *To include alternate title; conference title; and journal title, if they are different.*

Elements	Example of what is expected for each element
	and allow extra time to slow down. This application can reduce CAV truck delays 2 to 6 percent, with a significant increase in spacing between vehicles. SPD-HARM optimizes the speed of vehicles as they approach an intersection so that the delay is minimized. The results show that SPD-HARM can reduce intersection delays for trucks between 4 and 82 percent. TSP implementation allows extra time to transit vehicles at signalized intersections, and the developed algorithm has the potential to reduce transit delays up to six percent. The developed control programs can be directly implemented in the field, as they use the information which exists within CAV technology and the field-based signal control software.
Subject and Keywords	Connected and Autonomous Vehicles; Traffic Signal Control; Intelligent Traffic Signals; Freight Signal Priority; Queue Warning; Speed Harmonization; Transit Signal Priority; Wyoming
Identifier[2] and/or source	
Collection and Related Documents	
Edition	
Related Documents	
Coverage	2018-2021
Language	English
Publisher/Distributor	Wyoming Department of Transportation
Funding agency	Wyoming Department of Transportation
Access Restrictions	public (Data asset is or could be made publicly available to all without restrictions)
Intellectual Property and Other Rights	This document is available through the National Transportation Library and the Wyoming State Library. Copyright © 2018. All rights reserved. State of Wyoming, Wyoming Department of Transportation, and the University of Wyoming.

[2] *To include record numbers; report numbers; NTIS number; TRIS Accession Number; OCLC Number; ISBN; ISSN; contract number; and DOI if available.*

Elements	Example of what is expected for each element
License	
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	University of Wyoming 1000 E University Ave Laramie, WY 82070 and Wyoming Department of Transportation Programming Research Unit 5300 Bishop Blvd. Bldg. 6100 Cheyenne, WY 82009

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/.