

Estimating the Effects of Vehicle Automation and Vehicle Weight and Size on Crash Frequency and Severity: Phase 1

Data Collection

What data will you collect or create?

The main objective of this project is to develop an open, replicable, deployable tool for assessing the effects of vehicle automation and heavier vehicles on crash frequency and severity by utilizing national crash datasets such as the Fatality Analysis Reporting System (FARS) and Crash Reporting Sampling System (CRSS). Data on the existing road network and simulation results will be stored as CSV files and in separate folders. These files will be relatively small in size (<1 GB) and therefore easily stored in the CSV format.

How will the data be collected or created?

The data will be collected by downloading files from online and through application programming interfaces (API). Files will be stored on my hard drive as well as cloud storage service such as Box. All data files will be stored in the same folder for ease of access. In order to ensure quality of data I'll have other members of the project team, including myself review the validity of the data.

Documentation and Metadata

What documentation and metadata will accompany the data?

A documentation file will be created outlining how the methodology for accessing each data source. The documentation will include who created or contributed to the data, its title, date of creation and under what conditions it can be accessed.

Ethics and Legal Compliance

How will you manage any ethical issues?

The data planning to be collected for the project is publicly available and will be used for research purposes only. The data will be shared amongst the research team for the duration of the project. This research does not involve human participants or the collection of any personally identifiable information.

How will you manage copyright and Intellectual Property Rights (IP/IPR) issues?

The data is publicly available and can be used by third-parties for research purposes. The data created from this project can be shared once the project is completed

Storage and Backup

How will the data be stored and backed up during the research?

Files will be stored on my hard drive as well as cloud storage service such as Box for backup. If extra space is needed an external hard drive will be used to create additional storage space. In the event of an incident, the data can be re-downloaded from online.

How will you manage access and security?

Only the individuals on the project team will be given permission to access data. Data will be stored online and only shared with members of the research group.

Selection and Preservation

Which data are of long-term value and should be retained, shared, and/or preserved?

The data is publicly available, so can be stored and reused by other researchers as long as its still valuable.

What is the long-term preservation plan for the dataset?

The data will be stored on a cloud storage service, which is a free service offered by the institution. This will ensure the data can be used effectively beyond the lifetime of the grant.

Data Sharing

How will you share the data?

The data will be shared with other researchers and through a public box folder so that the data can be downloaded and accessed by others. All code written by the research team used to format or process any of the data in this project will be made publicly available on a Github repository.

Are any restrictions on data sharing required?

There are no restrictions to sharing the data.

Responsibilities and Resources

Who will be responsible for data management?

- Data Capture- Students
- Metadata production- Students
- Data quality- Corey Harper
- Storage and backup-Corey Harper

The principal investigator and data manager will be responsible for implementing the DMP, and ensuring its reviewed and revised.

What resources will you require to deliver your plan?

No additional resources are required to deliver plan.

Planned Research Outputs

Data paper - "Estimating the Number of Lives Saved by Crash Avoidance Technologies"

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?
Estimating the Number of Lives Saved by Crash Avoi ...	Data paper	2024-06-29	Open	None specified		None specified	None specified	No	No