Plan Overview

A Data Management Plan created using DMPTool

Title: Bus on the edge: Passengers

Creator: Christoph Mertz

Affiliation: Carnegie Mellon University (CMU)

Funder: United States Department of Transportation (DOT)

Template: Digital Curation Centre

Project abstract:

In this project we will use cameras mounted on the outside and inside of a transit bus to observe passenger behavior. We are interested to determine the full trip of the passengers, from where they enter to where they exit the bus. We want to develop algorithms that can automatically detect unusual behavior. Of special interest are the trips of people with disabilities, we want to determine which parts of the trip or interactions with other people create barriers to access or equitable ridership. This information will help the transit agencies to make the trips more efficient and equitable.

Start date: 07-01-2022

End date: 06-30-2023

Last modified: 11-18-2021

Bus on the edge: Passengers

Data Collection

What data will you collect or create?

The data will consist of videos and images inside and outside of a transit bus tagged with time and GPS localization. All the data will be taken in public areas and will contain traffic (passengers, vehicles, pedestrian, etc.), infrastructure (road, signs, etc.) and the environment (vegetation, etc.).

How will the data be collected or created?

Five cameras are mounted on the outside of a transit bus and up to two in the inside. The video streams go into a computer where interesting data are selected. The selected data will then be transmitted to our central data server.

Documentation and Metadata

What documentation and metadata will accompany the data?

The images are in standard JPEG or equivalent format. The metadata (GPS, time stamp, camera info) will be in plain text format. The metadata will be stored with the images.
Most of the raw data will be analyzed on the edge computer inside the bus. Only relevant observations will be sent to the central computer on a regular basis. Some samples of complete trips of the raw data will be saved for development of the edge computing software.
The final results will be in .csv, GIS or similar format and can be viewed with standard mapping software (ArcGIS, Google Earth, etc.)

Ethics and Legal Compliance

How will you manage any ethical issues?

All the data is taken in public areas. We do not expect any ethical issues. We already have IRB approval that there are no restrictions collecting data on the outside of the bus. We will confirm with IRB that the same is the case with data from inside the bus.

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

The data is collected by CMU and therefore it is the IP of CMU. The CMU IP rules apply, which we will follow.

Storage and Backup

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

The raw data will be stored in a raid server with automatic backup. Researchers will copy parts of the raw data onto their work computers where they will analyze the data and develop algorithms. Automatic backup will be installed on all work computers.

How will you manage access and security?

The data server is password protected. Researchers of the project will get an account on the server.

Selection and Preservation

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

All the raw data should be preserved for the long term. If storage space becomes an issue, we will delete data which is similar to other data we collected.

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

The data server will be available for the long term.

Data Sharing

How will you share the data?

We will share the data within CMU and our UTC partner university OSU.

Are any restrictions on data sharing required?

We do not expect that any restrictions will be required, but we will confirm with IRB.

Responsibilities and Resources

Who will be responsible for data management?

Christoph Mertz will be responsible for the data management. He might delegate some parts of the management if necessary.

What resources will you require to deliver your plan?

A raid server and automatic backup. Both are already available. We will make use of standard SCS hardware and software support.