**Project Title: Latest Generation Data Portal for the Intelligent Mobility Meter**

Principal Investigator: Bernardo Pires, PhD

**Data Management Plan**

**1. Data Description**

The purpose of the research is to advance the state of art on the automatic detection, classification and tracking of pedestrians, bicyclists, and motor vehicles as well as to collect traffic modeling statistics at specific locations identified by selected partners. To this end, the project shall receive video recordings from selected partners and shall produce data in the form of periodic counts, of vehicles, bicycles, and pedestrian. The video data shall be collected by the partners and be outside the scope of this data management plan. The annotations shall be generated automatically and verified manually. No audio information shall be received.

The annotations shall be compiled during the last six months of the period of performance. (One year project.) No updates to the data are expected. Project will not use pre-existing data.

Potential users of the data include local governments, who may use the data to make infrastructure decisions. This use case shows the value of the data in the near term. Additionally the data will have significant impact in the long term as it is expected that computer vision researchers will use the data to create and/or test methodologies for the automatic detection of motor vehicles, bicycles, and pedestrians.

All data shall be made public. The data will be managed by the PI, Dr. Pires, who will check for adherence to this Data Management Plan.

**2. Data format and metadata standards**

All data submitted will use open standards. The annotations shall be submitted in txt files. No special software shall be required to view the annotation data, as any text editor shall be able to read the data. Because it is not expected that the data will be updated, there will be a single version of each file, which shall be made public.

The video files shall require no documentation beyond the metadata. The annotation file shall include a data dictionary to ensure the data can be understood by other researchers in the field. Annotation quality shall be ensured by manual verification of some or all annotations.

Metadata shall be recorded following the schema of the Federal Highway Administration’s Traffic Monitoring Guide, which is one of the standards for the description of vehicle, bicycle and pedestrian traffic. The metadata shall be stored with the associated data.

**3. Policies for Access and Sharing**

Data shall be freely available to the public and no restrictions shall be placed on data or metadata distribution. Sharing of the data is not expected to raise any ethical, privacy or confidentiality concerns.

**4. Policies for re-use, redistribution, derivatives**

A copy of all data created by the project, as well as its corresponding metadata, shall be submitted to Mobility21 University Transportation Center to be deposited to Zenodo for access and data sharing (see more details below). Carnegie Mellon University will hold the IP and copyright for data and other materials created by the project. The PI of the project agrees, upon deposit of their data, to the non-exclusive licensing agreement of Zenodo and the CMU data archive. The data shall be released in an open license for reuse, redistribution and derivative products which will be based upon the open licenses and provided by Zenodo and the CMU data archive.

**5. Plans for Archiving and Preservation**

The Mobility21 UTC will archive all data on Zenodo, https://zenodo.org/, which is an approved data repository by USDOT. The PI agrees to archive all project data with Zenodo within 60 days of the submission of the project final report. The PI will maintain the data until it is uploaded to Zenodo.

Data security (back-up, redundancy, disaster recovery) shall be maintained by recording all data and metadata to two external hard drives, kept at separate locations, as well as a private cloud copy on Amazon Web Services. Integrity of the data (protection from accidental or malicious modification or deletion prior to receipt by the archive) shall be insured by compilation of cryptographic hashes of all project files, which shall be maintained and verified periodically by the PI.