

# Safe Intersection Crossing for Pedestrians with Disabilities

---

## Data Collection

### What data will you collect or create?

This project aims to further develop and deploy PedPal, a mobile app designed to assist pedestrians with disabilities in safely crossing signalized intersections. As such, data collection is not an explicit end goal of the project. However, to evaluate the efficacy of some of the PedPal extensions that have been proposed, experiments will be performed and data will be collected for purposes of quantifying performance. In particular, an attempt will be made to measure the overall accuracy of mechanisms developed both for monitoring pedestrian progress through the intersection and for detecting when a pedestrian moves outside of the crosswalk.

In the case of progress monitoring, a series of crossing trials will be performed at various surtrac controlled intersections in the east end. Intersections will be chosen so as to vary the crossing distance and incorporate intersections with different size characteristics. For each crossing trial we will collect the following information:

- crossing trial number
- crossing distance
- start time of the cross
- end-time of the cross
- detected position (location) of the pedestrian after each second of the cross as provided by the app.
- actual (ground truth) position of the pedestrian after each second of the cross (as derived from a timed video of the cross with street markings)
- accuracy of each detected location relative to ground truth.

In the case of detecting veering outside of the crosswalk, a second series of crossing trials will be performed at various surtrac controlled intersections in the east end. Intersections will be chosen so as to vary the crossing distance and incorporate intersections with different crosswalk characteristics (perpendicular, slanted). For each crossing trial we will collect the following information:

- crossing trial number
- crossing distance
- for each weave during the cross that moves the pedestrian outside of the crosswalk, (1) the detected time and location of the movement outside of the crosswalk, (2) the actual time and location of the movement outside of the crosswalk (as derived from a timed video of the cross with street markings), and (3) the accuracy of each detected crosswalk violation relative to ground truth.

### How will the data be collected or created?

To collect the information related to the detected locations over time, both for the progress monitoring and the crosswalk violation crossing trials, we will instrument the PedPal app to log this information and subsequently retrieve it from the logs.

To obtain the actual (ground truth) locations over time for both types of crossing trials, we will record a timed video of each trial, and derive the actual locations from this information. To facilitate this process, we will introduce temporary markings in the intersection (e.g., marking tape) at regular intervals.

From this baseline information, accuracy information will be derived.

The entire data set will be maintained in an accessible format, most likely a comma separated values file with appropriate header information.

## Documentation and Metadata

### What documentation and metadata will accompany the data?

A data dictionary will be provided with the file containing the data in comma separated form and header information, to explain the semantics of the values in each field (column).

## **Ethics and Legal Compliance**

### **How will you manage any ethical issues?**

Our intent is that all trials will be performed by project personnel, strictly for purposes of quantifying the accuracy of the mechanisms for progress monitoring and crosswalk violation detection. It is not our intent to conduct a study of prospective users of the extended PedPal system. Hence, there are no ethical issues to manage to our understanding.

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

CMU will retain IP rights to all data collected and created over the project lifetime. It is not anticipated that there will be any 3rd party interest in this data. However, we expect to publish the results of these experimental analyses, and if there is an interest in the data for purposes of comparison with other tracking technologies, then a mechanism for making the data available will be developed.

## **Storage and Backup**

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

The data will be stored in the Intelligent Coordination and Logistics Laboratory's GitHub repository and version control system. This repository is regularly backed up by GitHub.

### **How will you manage access and security?**

Access to this data is provided only to project personnel through controlled accounts which are issued by the PI. GitHub encrypts all data in transit, all login information and credentials are always protected. GitHub stores a one-way hash of all user passwords using *bcrypt*. Your account login is protected from brute force attack with rate limiting.

Data collected in the field will be downloaded from the app and extracted from the video using local project member laptops and then entered into GitHub.

## **Selection and Preservation**

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

All of the data mentioned above will be considered sharable once the project has completed. As mentioned earlier, our intent is to publish results following from the experimental trials and the data will be preserved to support this activity. We expect to retain the data generated for a period of at least 2 years after the end of the project.

Although we don't anticipate it, we will make this data available in as is condition, should another third party desire access to the data for future research purposes.

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

As indicated above, the data will remain in the ICLL GitHub repository for at least 2 years after the completion of the project. If a reason materializes for preserving the data indefinitely, we will seek an appropriate public repository service. One natural candidate will be the data repository currently being configured by the FHWA Accessible Transportation Technology Initiative (ATTRI) program for management of data generated by ATTRI research projects (of which PedPal was one).

## **Data Sharing**

### **How will you share the data?**

We do not anticipate any external party being interested in this data. But we are willing to share it for various educational and/or research purposes. We will take a low overhead approach to sharing, where an interested party contacts the PI and the PI then responds by forwarding the data by email in its planned comma separated values format.

### **Are any restrictions on data sharing required?**

No.

## **Responsibilities and Resources**

### **Who will be responsible for data management?**

The Intelligent Coordination and Logistics Laboratory will be responsible for implementing the data management plan, until which time it is deemed necessary to enlist an independent repository service for indefinite preservation (assumed to be unlikely).

### **What resources will you require to deliver your plan?**

No additional software, hardware or staff training will be necessary.