
DATA MANAGEMENT PLAN

1 Types of Data to Be Collected

We will collect data on human motion within a public space using sensor data. We expect to collect video and depth data. This data will be converted to measurements on human motion. All sensors will be human-safe.

Where applicable, we will use appropriate standards or standard practices for the data collected. For example, data collected from robots will be stored by ROS and in forms that other ROS systems can read. Proper documentation will be utilized for any new data structures created.

2 Institutional Review Board

All human participant data is subject to IRB regulations and we will file protocols and modifications for IRBs as needed. The team is well versed in these procedures, and Steinfeld serves on the Carnegie Mellon Minimal Risk IRB Board.

3 Data Retention, Access, Archiving, and Storage Facilities

The team typically retains data for at least 5 years beyond the end of each study in order to support longer publication cycles and students who are still working on their PhD theses.

Access to study data is limited to team members and all servers are kept in locked rooms. Data retention is monitored by the IRB.

The team has secure lab servers with independent, industry-grade back-up regimens.

4 Data Analyses

The faculty team members have extensive training and experience on study design and analysis. They will ensure proper statistical techniques are employed. Most analyses will be conducted using custom written methods in either industry standard packages (e.g., MatLab, etc.) or new code (e.g., Python). The team is also well versed in a variety of statistical software packages.

5 Data Sharing and Re-Use

When possible, we will seek to configure data collection efforts in a way that allow sharing and re-use. Due to the presence of personally identifiable information (e.g., video), we may be unable to share study data beyond our project team due to IRB review requirements. Data collected within buildings maintained by 3rd parties (e.g., transit stations) may also have limitations on sharing and re-use, depending on negotiated agreements.

The team will share analyzed results through traditional publication and dissemination outlets. We will place a priority on publications in archived, easily searched venues (e.g., ACM, IEEE, etc.).

6 Types of Software to Be Produced and Dissemination Plan

We utilize a code repository with version control for our own development. Some code developed under the proposed work may be open sourced. Some aspects of this work will be developed with the open source ROS (Robot Operating System), which has quickly become a common standard in robotics. We will contribute our ROS-specific packages to the community

once stable and performing well. These will likely be general robot and user interaction components.

7 IP Policies and Commercialization of Hardware

The team has no plans for commercialization of software or hardware developed under this work. If these plans change, the team will work with the university technology transfer and conflict of interest offices to ensure best practices are followed. Steinfeld has experience in this process.

8 Roles and Responsibilities with Respect to the Management of the Data

Steinfeld and Admoni will lead all efforts and be responsible, in conjunction with IRB, for monitoring data management activities. Student members of the team will be responsible for day-to-day activities.