

## **Data Management Plan:**

This data management plan covers the research results and project reports. We plan to keep all the documents related to the project on the project website (e.g., project plan, and results). All articles and reports published will be online. The website will be secured and protected from any attacks. The above items will include open access data on openly available website at the University of North Texas (UNT).

**De-identification:** Based on our prior experience, we do not need IRB study. If required, we plan to apply to the Institutional Review Board (IRB) for this integration and study.

**Privacy concerns** We are not expecting any privacy concerns due to deidentification

**Secured Storage:** All the data to the portal includes a secured transmission and secured storage. Data will be stored securely on the portal at the University of North Texas (UNT).

**Security, Storage, Management and Back-up of Data:** The simulated data during the experiments will be archived within UNT's Network Security Laboratory. The laboratory has a lab manager. We will be strictly following the UNT guidelines for storage and security of data (<http://security.unt.edu/policy>). During the development of the project, data will be physically stored on a password-protected, VPN-based secured server maintained by UNT IT staff. We plan to backup on a daily basis. The researchers are currently responsible for storage during the project development. The IT staff will be responsible for maintenance and backup of data.

**Policies for collecting data:** Currently we are not anticipating any data collection. If required, we plan to follow ALL IRB requirements/guidelines for collecting the data (currently, no plan but if required). Comprehensive UNT computing facility policies (e.g., <http://policy.unt.edu/policy-by-chapter>) will be strictly applied for collecting the data.

**Legal guidelines and requirements:** We do not anticipate any copyright, intellectual property, or licensing issues associated with the data.

**Access, Sharing and Re-use of Data:** We are not aware of any reasons that might prohibit the sharing and re-use of the data (mostly simulated). We will remove direct identifiers, aggregate or reduce the parameters, generalize the meaning, and restrict the lower and upper-case ranges of the variable. Results will be published in open journals/conferences by the PIs and students.

**Data Standards and Capture:** Data will be stored (wherever required) in the European Data Format (EDF) file, a simple and flexible file format well-known for storing multichannel biological and physical data. In EDF Files usually each file is generated for uninterrupted digitized recording. A data file consists of header record holds the Metadata about the recording signal and followed by the data records. The header stores the identification information like the subject name, time information, the experiment name, the type of data collected and other technical information about the recorded signal Also it is more appropriate if the file format is simpler since it will be processed in real-time. Therefore, the data will be stored in CSV files where each file will be given a unique id for the simulated data

**Preservation, Review and Long-term Management of Data:** The data (e.g., software code in the github) will be stored in a specific virtual archive and will be made available through UNT network to the PIs and the students working on the project. The UNT networks are well established, trusted (no major incidence reported for several years), managed by the computing center. We will be archiving the login information, downloads, and any data access. All the formats and physical storage media will ultimately become obsolete. As a long-term measure, every 3-5 years, we will migrate the data to sustainable formats and medium. After software conversions, data will be checked for errors or changes that may be caused by the conversion process. We will revise the Metadata accordingly. Also, we will explore public repositories for long-term management of data.

**Dissemination of data and results:** We are planning an active dissemination program to complement the activities of this project. We plan to publish the results of the student research and dissertations.

The first dissemination mechanism will be the standard archival channels of scholarly communication, involving peer-reviewed publications and technical reports (ACM/IEEE intelligent transportation society). As a second, more dynamic means of dissemination, we plan to develop a website that will be used to communicate the project aims, data types, and other initiatives related to issues. We will also explore the

public data repositories. We follow standard procedures for documenting the software code, and this includes detailed description of hardware and software.