

## Traffic 21 Smart Mobility Challenge 2017 - Data Management Plan

### *Building a pilot peer-to-peer platform for ride-sharing in Lawrence County*

All models, data and outputs of this project will reside on workstations within Carnegie Mellon University. They will be periodically and systematically backed up onto multiple external hard drives and centralized backup systems provided by the institution, to ensure full data recovery in the event of equipment failure. The following steps will be undertaken for data management.

**Models** This project will develop data visualization, data analyses and an online platform for promoting ride-sharing in the Lawrence County. Ultimately, we aim to demonstrate the potential effectiveness of a technological solution to enhance mobility in rural areas, and assess its adoption dynamics among its users. The data analysis will be performed using the Microsoft Excel software and the R programming language (<https://www.r-project.org>). Optimization models will be built using the GAMS ([www.gams.com](http://www.gams.com)) and MATLAB ([www.mathworks.com](http://www.mathworks.com)) programming languages, and the IBM ILOG CPLEX Optimizer solver ([www-01.ibm.com/software/commerce/optimization/cplex-optimizer/](http://www-01.ibm.com/software/commerce/optimization/cplex-optimizer/)), all commonly used in scientific and engineering research. All software and model codes, along with simple demonstration applications, will be made available to interested parties upon request, and will be transmitted electronically via e-mail.

**Data** Data inputs will be obtained from the Lawrence County, partner organizations, and public databases for demographic and geographic data. The largest databases will be stored in the mySQL format, and all others will be stored in the .csv format, in computer servers hosted at Carnegie Mellon University. Data will be made available upon request, whenever this does not violate any agreement with the third parties. When sharing the data, we commit to protect privacy, confidentiality, and security.

**Outputs** The output data files will characterize potential travel outcomes with a technological ride-sharing solution, under different user adoption and operating scenarios. They will be made available as mySQL databases or in standard .csv format upon request. If the research and discoveries of the project might be secured with intellectual property, we will work with our respective Technology Transfer Offices to protect potential proprietary data.

**Presentations and Publications** All significant research insights obtained in this project will be promptly shared with the practitioners and the scientific community through presentations in seminars, conferences and symposia, and through publications in scientific journals and reports. All presentations will be made by the researchers involved in the project, and all publications will be co-authored by all researchers who have actively participated in that particular part of the project.

**Educational Materials** Software, sample applications, visual interfaces and teaching materials will be packaged into modules, and potentially used in courses taught by the PIs.