

VEHICLE DATA MODELING: A TOOL TO AID RESEARCHERS IN TRANSPORTATION POLICY ANALYSIS AND DECISION-MAKING.

Virginia Allen, McNair's Scholars Program, Academic Affairs, vla8446@rit.edu; Eugene Marcotte, Undergraduate Student, Department of Software Engineering, exm3763@rit.edu; J. Scott Hawker, Department of Software Engineering, hawker@mail.rit.edu

The National Science Foundation is currently funding the Materials Use: Science, Engineering, and Society (MUSES) project to create the Computational Automotive Policy Analysis (CAPA) system which will contain linked models. CAPA will show how a greenhouse gas (GHG) policy on passenger cars and light trucks might affect vehicle manufacturers' and consumers' decisions along with material flows, costs, and any environmental impacts. The data for passenger cars and light trucks built and sold in the past will be used to create some of the models needed for CAPA. This historical vehicle data provides modelers with the information on what vehicles are in operation, the vehicle characteristics that consumers preferred, and other information needed by those creating the models needed for CAPA. However, this is a problem because information resides in different data sources or is created in different models, meaning modelers end up spending much of their time collecting the data they need. To solve this problem, data management tools are being developed to help with the process of collecting the data. The first tool, Data Describe will allow the modeler to define what a data source is and what the attributes in that data source mean. The second tool, Data Query, will allow the modeler to run queries on a data source. This will allow modelers to obtain the specific data they need from a data source in a time efficient manner and save it for later modeling and analysis. The research done this summer will consist of continuing the development of Data Describe and to begin the development of the Data Query tool. These two tools are being developed in Eclipse 3.3, Java 5, and the queries in Data Query will be completed by using Java Object SQL (JoSQL). These two tools lead up to the final one, Data Match. This will allow a modeler to match attributes from different data sources when the data sources differ on the definition of a specific vehicle make, model, and style. This tool will be a challenge since the data sources do not follow the same method of identifying vehicles. For example, one source may have the manufacturer name as Ford, where other may use FRD or Ford Motors Company, etc. In the end, all of these tools will be used to create the models needed for CAPA and will allow for CAPA to access the data needed for these models.