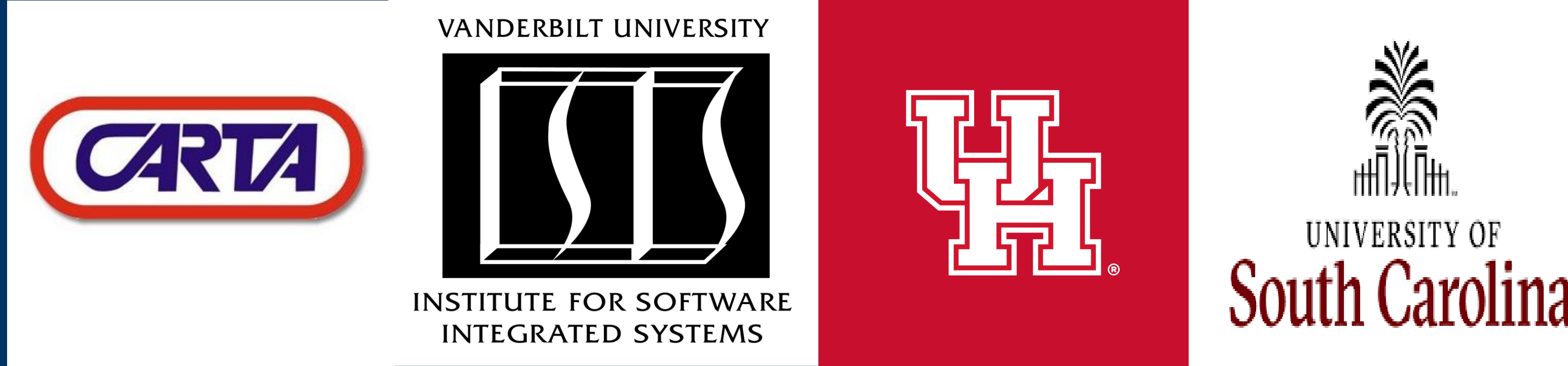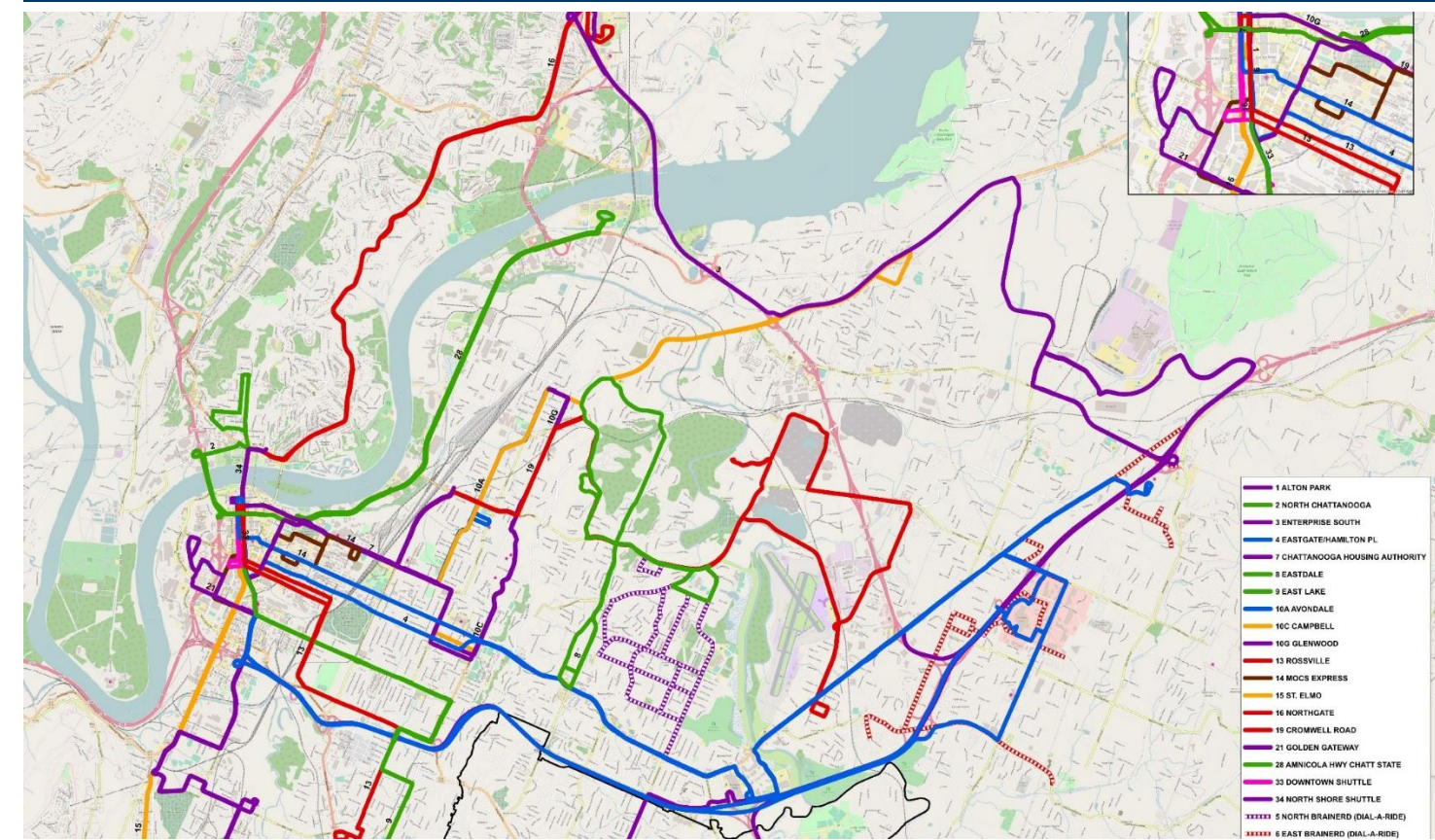# High-dimensional Data-driven Energy optimization for Multi-Modal Transit Agencies (HD-EMMA)

Philip Pugliese, Aron Laszka, Afiya Ayman, Abhishek Dubey, Michael Wilbur, Fred Eisele, Yuche Chen, Yunteng Zhang

VANDERBILT UNIVERSITY
INSTITUTE FOR SOFTWARE INTEGRATED SYSTEMS

CARTA

UNIVERSITY OF South Carolina

Background

- Energy costs are a major concern for transit agencies like Chattanooga Area Regional Transportation Authority
- This project is building a high-resolution system-level data capture and analysis for the transit operations to provide CARTA the capability to identify energy bottlenecks and accurately predict energy costs of all operations.
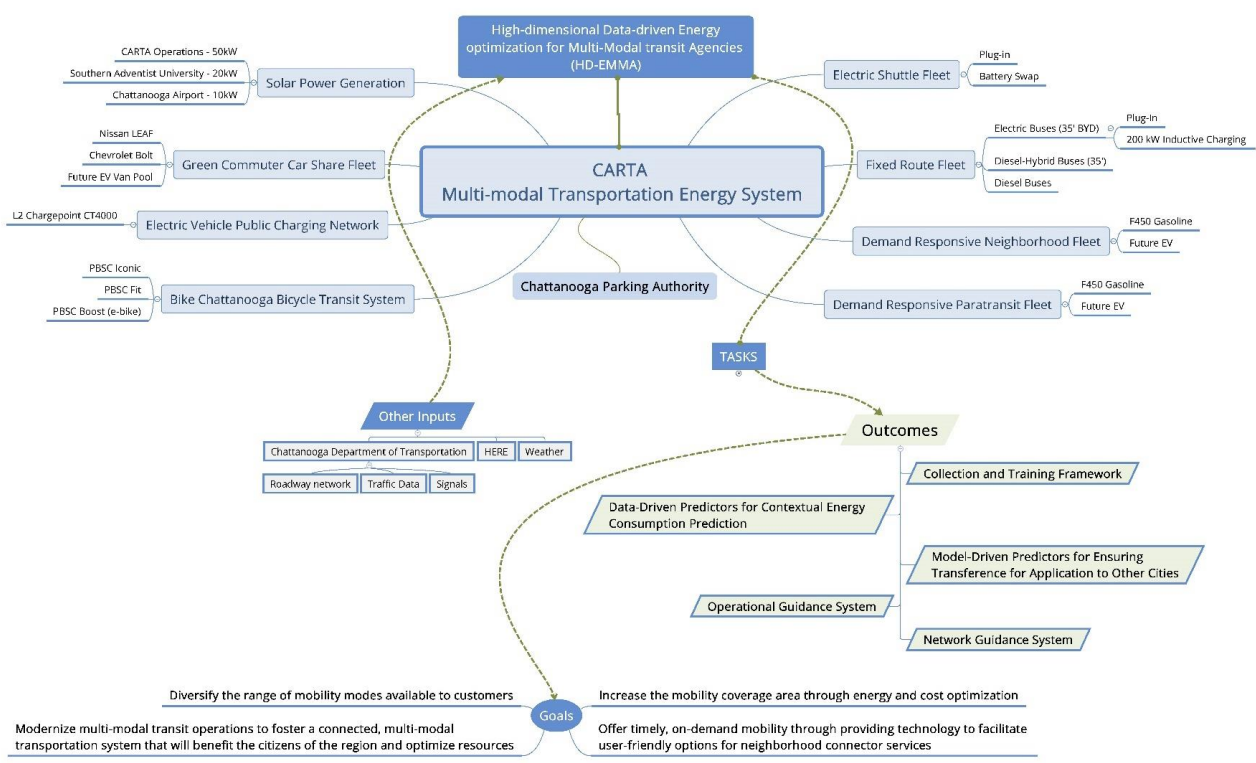
HD-EMMA Approach
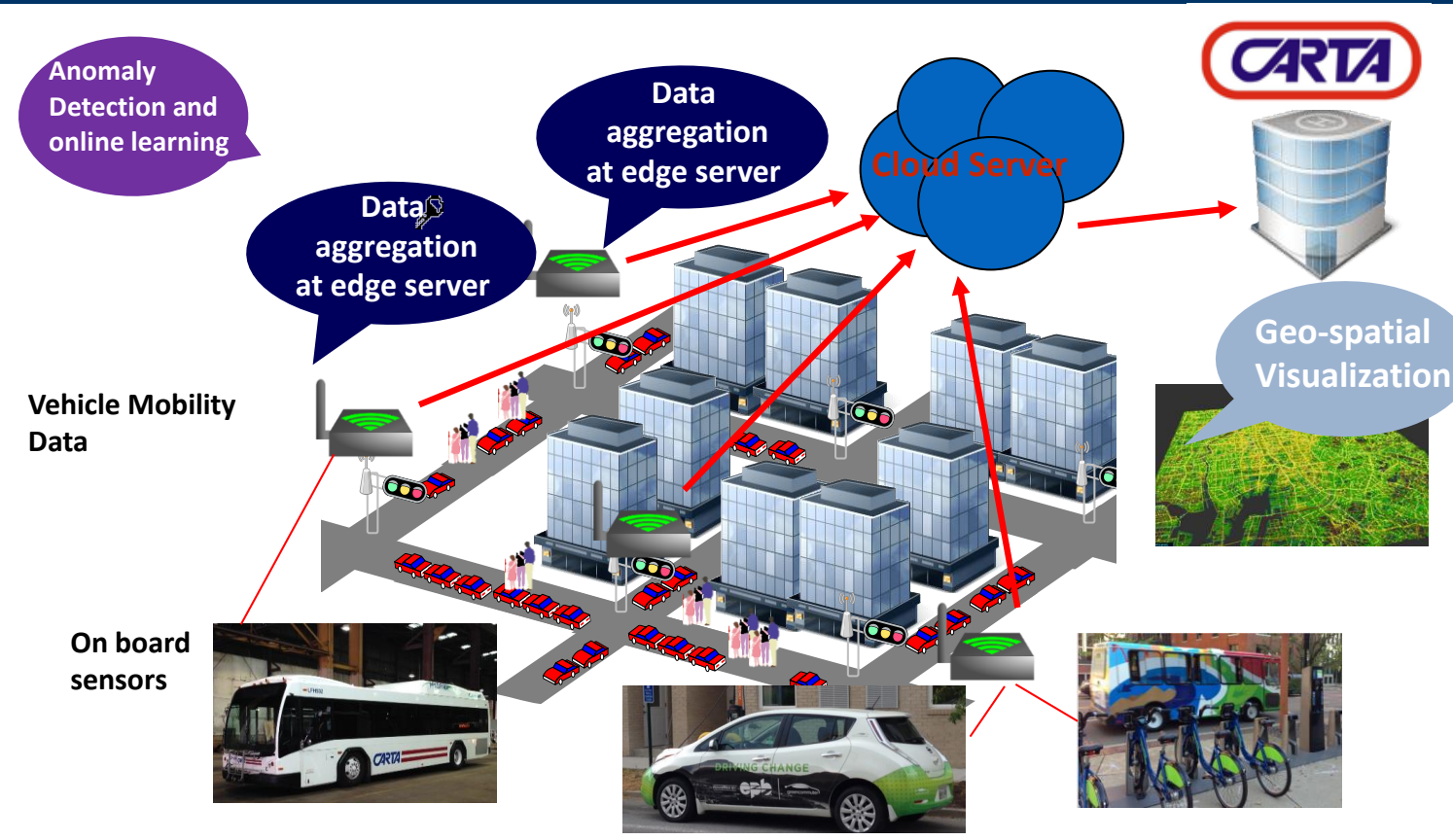
1. Efficient sensor data aggregation
2. Anomaly detection and data store for supporting high integrity, velocity, and volume of operational data
3. Micro and macro energy predictors
4. Operational guidance and city-wide geo-spatial visualization
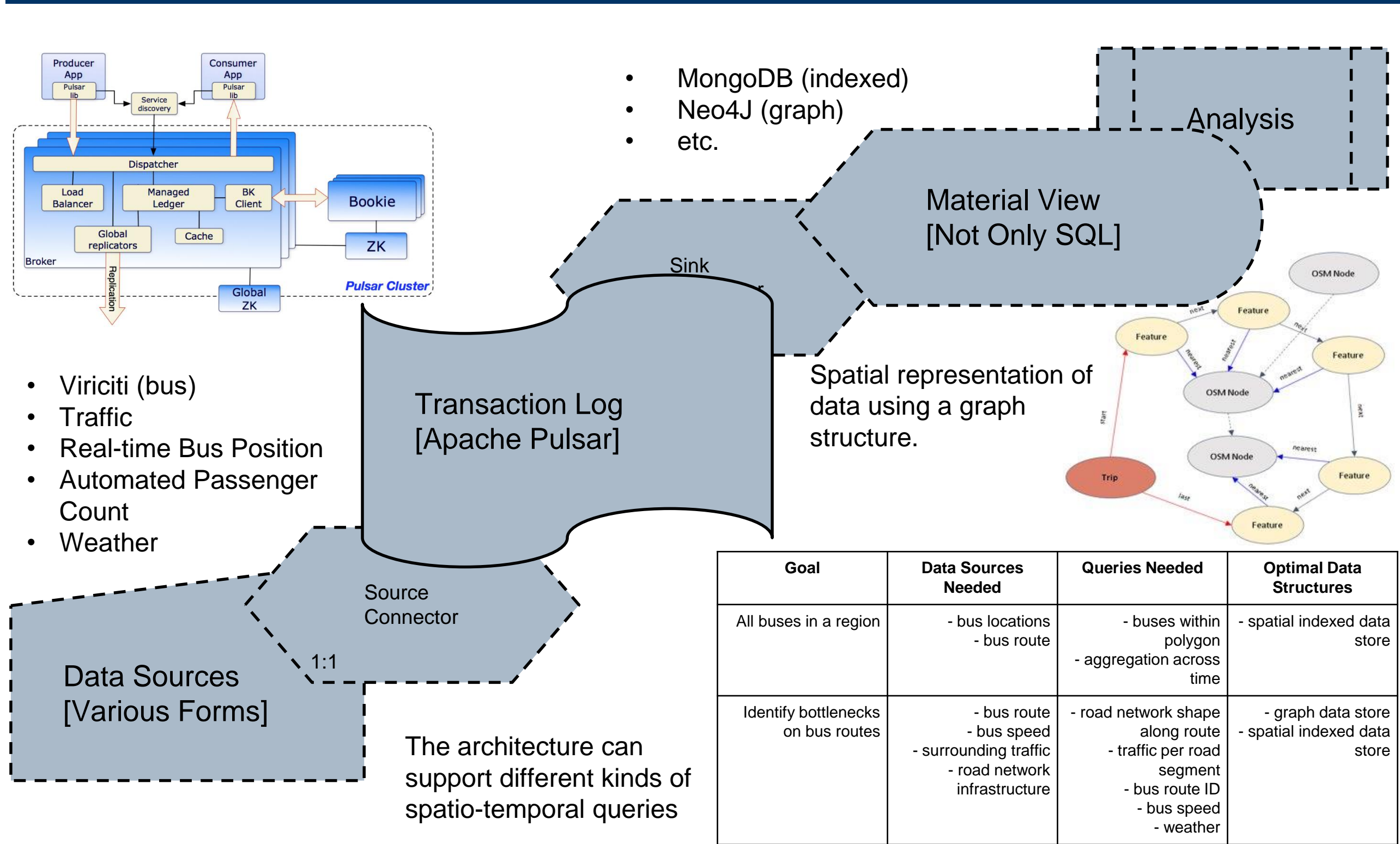
Sensor Data Collection
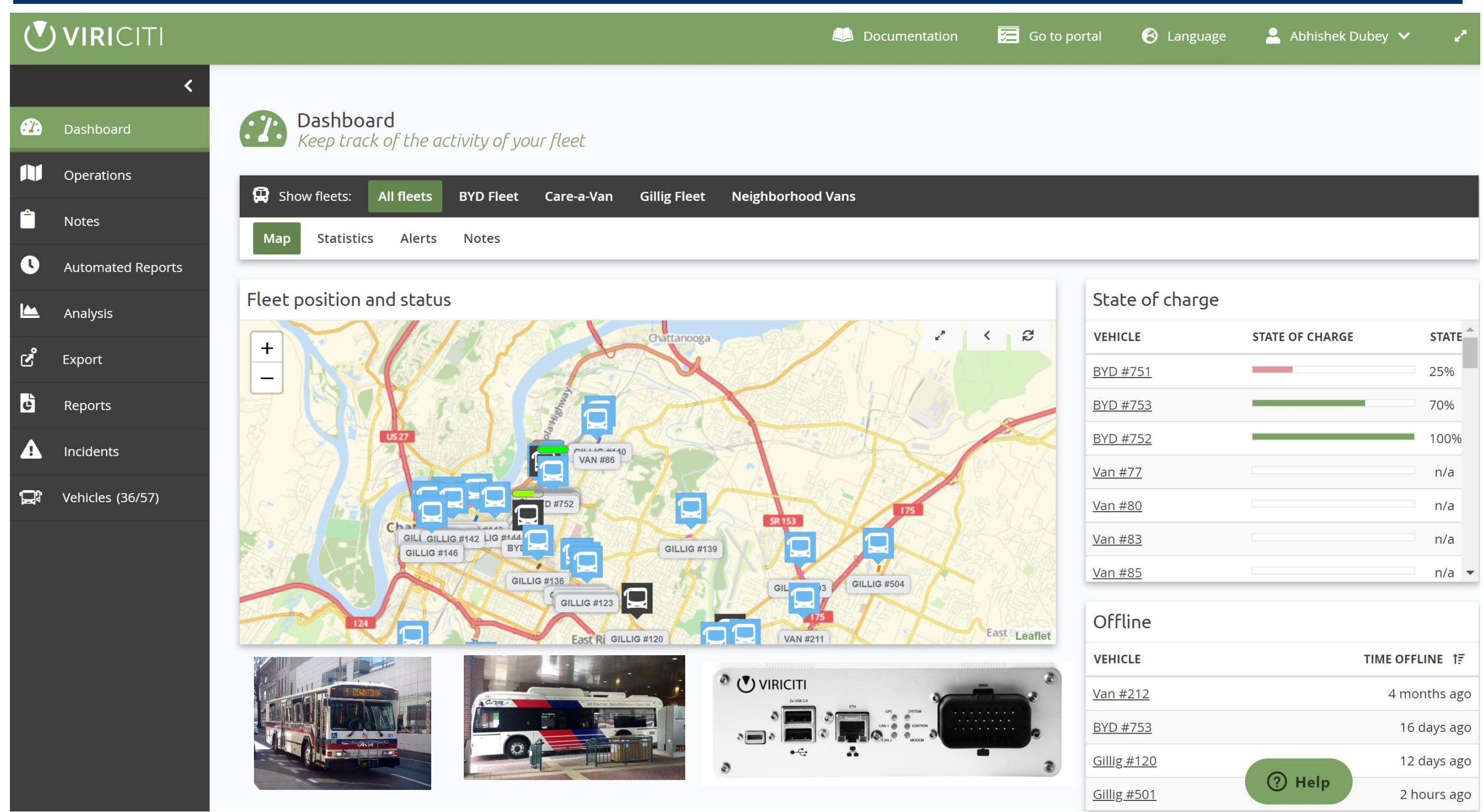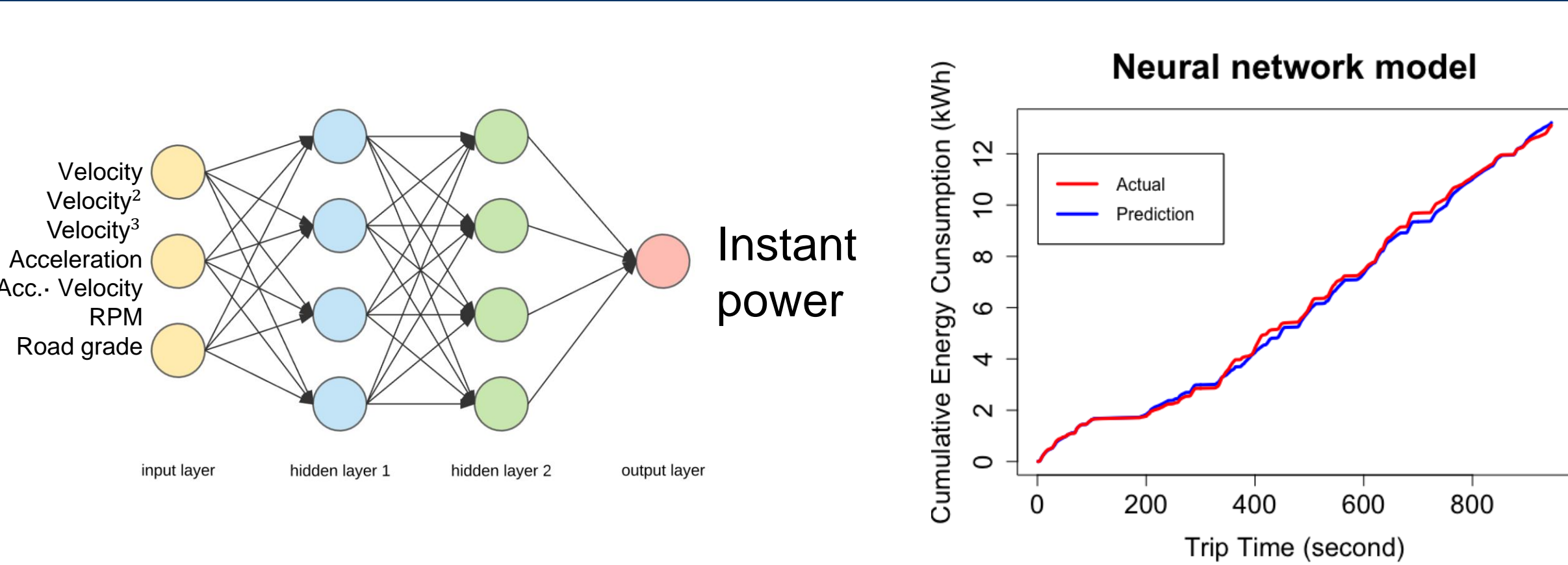

Key Deliverables

- Data Collection and AI Model Training Framework
- Data-Driven Predictors for Contextual Energy Consumption Prediction
- Generalized Predictors for Ensuring Transference for Application to Other Cities
- Operational Guidance system to improve transit agency performance and reduce energy consumption
- Network Guidance System to improve transportation mode selection process
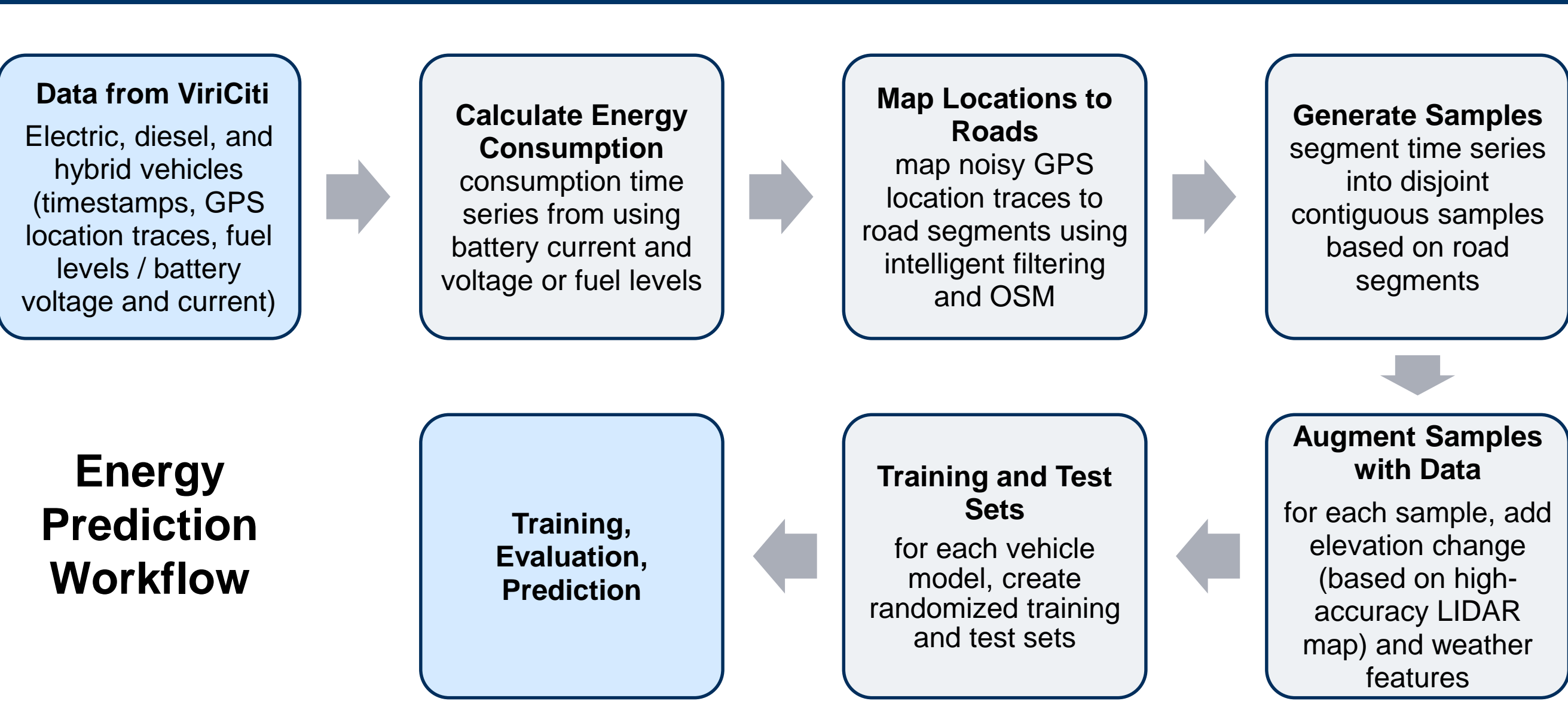
High Performance Data Store

- MongoDB (indexed)
- Neo4J (graph)
- etc.

Analysis

Material View [Not Only SQL]

Transaction Log [Apache Pulsar]

Spatial representation of data using a graph structure.

- Viriciti (bus)
- Traffic
- Real-time Bus Position
- Automated Passenger Count
- Weather

Source Connector

Data Sources [Various Forms]

1:1

The architecture can support different kinds of spatio-temporal queries

Goal	Data Sources Needed	Queries Needed	Optimal Data Structures
All buses in a region	- bus locations - bus route	- buses within polygon - aggregation across time	- spatial indexed data store
Identify bottlenecks on bus routes	- bus route - bus speed - surrounding traffic - road network infrastructure	- road network shape along route - traffic per road segment - bus route ID - bus speed - weather	- graph data store - spatial indexed data store

Micro Scale Energy Prediction (Initial Results)

Neural network model

Velocity
Velocity²
Velocity³
Acceleration
Acc.- Velocity
RPM
Road grade

input layer   hidden layer 1   hidden layer 2   output layer

Instant power

- Micro model is essential to evaluate energy benefits of traffic control strategies
- Existing models primarily adopted multiple linear regression models
- Limited (only 1-2) literature on micro model for electric bus

Macro Scale Energy Prediction (Initial Results)

Data from ViriCiti Electric, diesel, and hybrid vehicles (timestamps, GPS location traces, fuel levels / battery voltage and current)

→ **Calculate Energy Consumption** consumption time series from using battery current and voltage or fuel levels

→ **Map Locations to Roads** map noisy GPS location traces to road segments using intelligent filtering and OSM

→ **Generate Samples** segment time series into disjoint contiguous samples based on road segments

↓

Augment Samples with Data for each sample, add elevation change (based on high-accuracy LIDAR map) and weather features

← **Training and Test Sets** for each vehicle model, create randomized training and test sets

← **Training, Evaluation, Prediction**

Energy Prediction Workflow

Prediction example
- electric bus
- trained on 3 months of augmented data
- decision-tree based predictor

— Predicted Value
--- Actual Value

Electricity Consumption / Samples
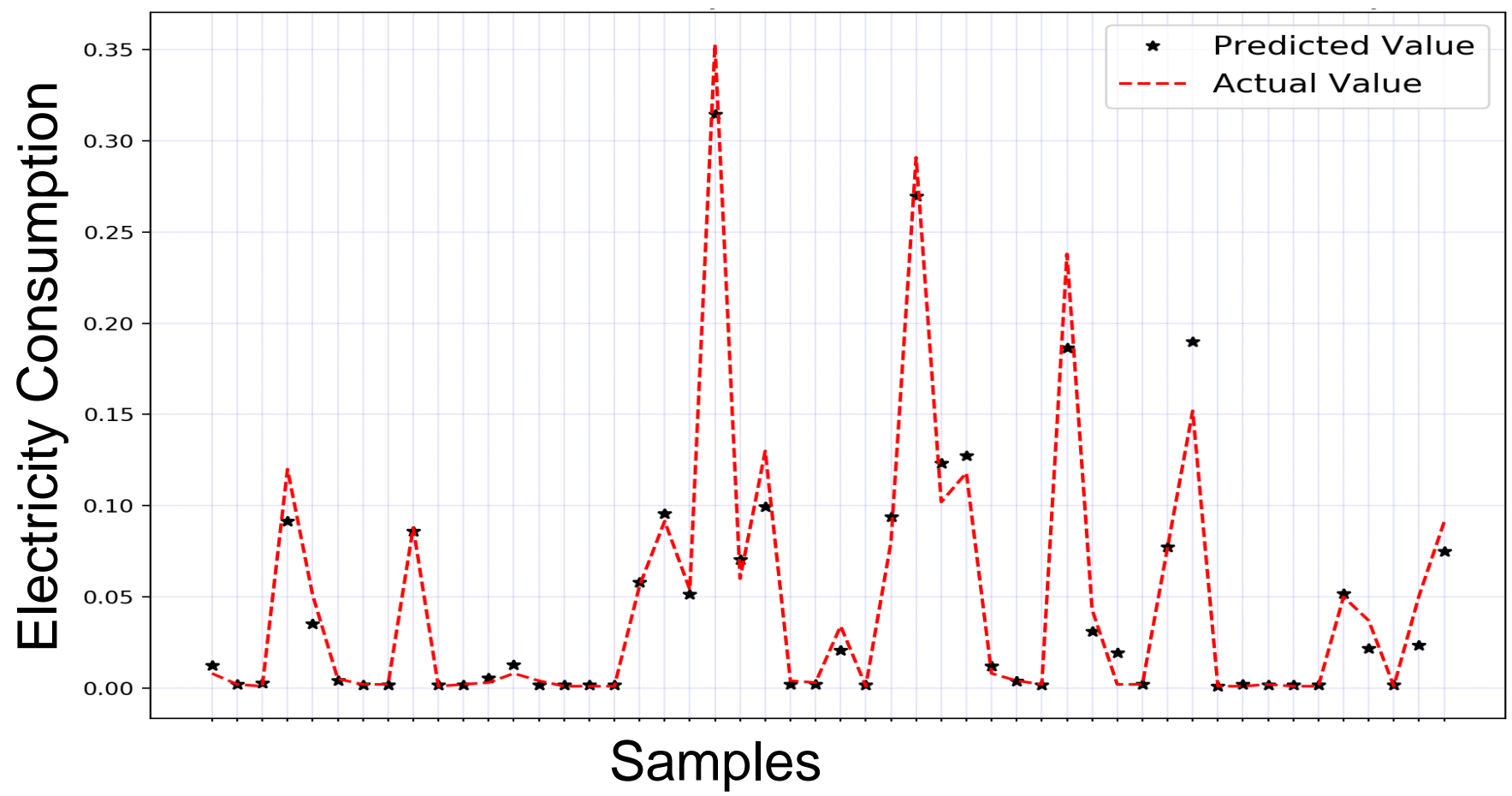
Accuracy gain from augmenting samples with weather data
- baseline without any weather data
- various parameters (e.g., temperature)
- all parameters combined

Prediction Error — Without weather / Wind speed / Precipitation / Temperature / Weather (combined)
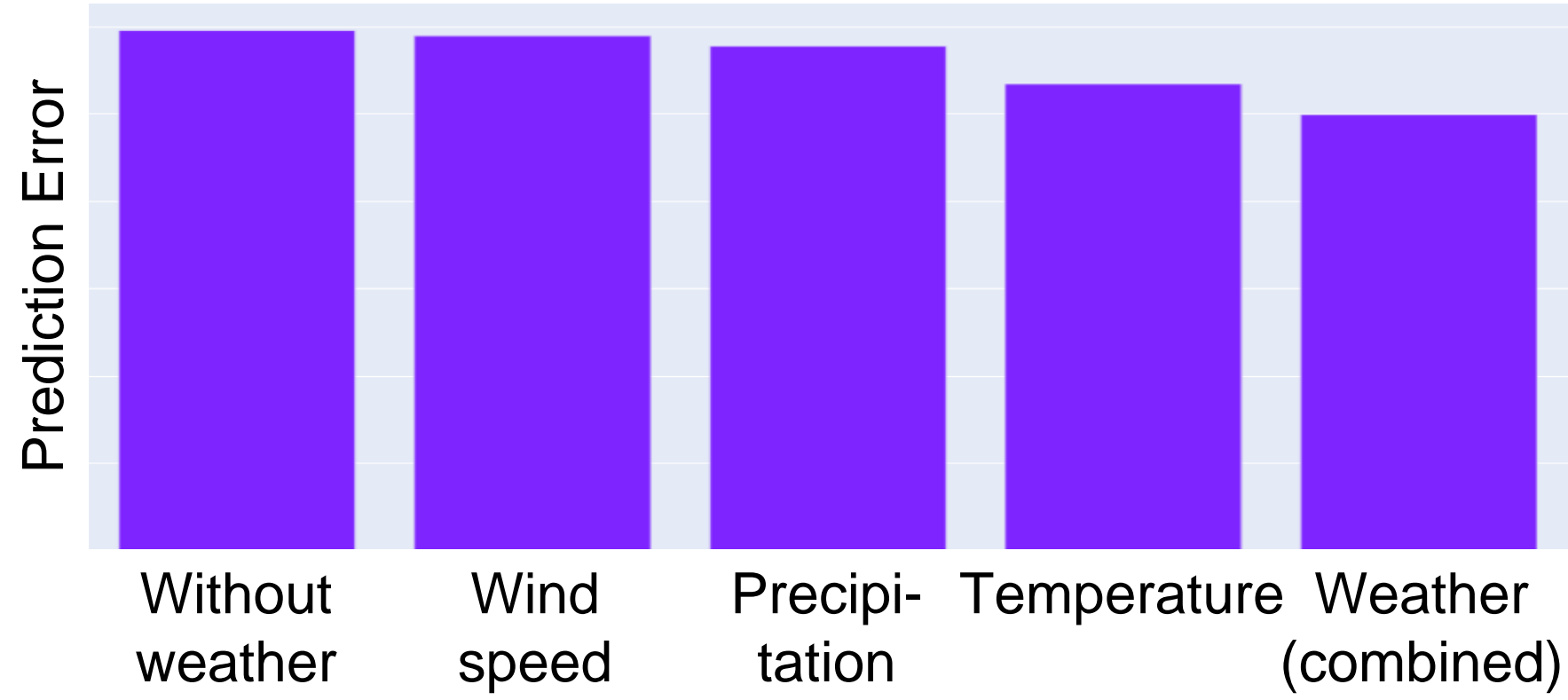
Project Impact and Takeaways

- The project will enable decreased operating costs through route and driver optimization
- It will also build first of its kind high-resolution dataset and machine learning models containing all information about engine idling status, engine temperature, engine speed, throttle, vehicle speed fuel level, engine temperature, road gradient, and driver behavior.

Source code: github.com/hdemma

Contact: Philip Pugliese: philippugliese@gocarta.org
Abhishek Dubey: abhishek.dubey@Vanderbilt.edu