## 2019 RESEARCH PROBLEM STATEMENT SOLICITATON

### I. PROBLEM TITLE

Transportation Data Analytics: Where Are We and Future Directions

### II. RESEARCH PROBLEM STATEMENT

Big data and data analytics have attracted substantial attention from both industry and academia across many disciplines and regions. State DOTs also own and generate a huge amount of data, including planning (e.g., population, employment, transit ridership), operations (e.g., loop detector, traffic camera, SPaT, parking), safety (e.g., crash records), and asset condition data (e.g., bus fleet maintenance, bridge and pavement conditions, traffic sign). These data sets are often collected and maintained by different divisions of DOTs, and such a practice may lead to duplication of efforts and underutilization of data. For example, sometimes one DOT division is not aware of the existence of a data set maintained by another division. This may result in the same data elements being collected and stored by multiple divisions. Also, data sets from different divisions could be but are usually not correlated for in-depth analysis. The power of the data is not fully understood and exploited as it should be.

With a complete understanding of what types of data DOTs have, the data collection, storage, and analysis efforts of different divisions may be coordinated and unified, and be conducted in a more cost-effective manner. This integrated approach will help to turn data into useful information and produce additional insightful results that support data-driven decision makings. It is also important for DOTs to learn from each other's successful practices of transportation data analytics. Given the advancements in connected vehicles, autonomous vehicles, smart cities, mobile sensors, and Internet of Things (IoT), it is anticipated that even more data will be generated. Thus, it is critical for DOTs to stay tuned of technology evolvements, understand their impacts on transportation data analytics, and avoid investing in soon-to-be-obsolete technologies.

While it is desirable to conduct a comprehensive review of all the data and data needs that state DOTs have, as the first step the focus of this problem statement is on data related to traffic operations, to ensure that the scope of work is feasible given the time and budget.

### III. RESEARCH OBJECTIVES

The objectives of this research include: (1) reviewing the practices of all DOTs in the New England area and developing a comprehensive inventory of data related to traffic operations, (2) providing recommendations on how to better collect, achieve, and analyze the data to support DOT decision making and how DOTs should prepare for future transportation data collection and analytics considering emerging technologies such as connected vehicles and IoT.

## IV. COST ESTIMATE

$180,000

## V. RESEARCH PERIOD

24 months

## VI. URGENCY AND PAYOFF POTENTIAL

Transportation agencies depend heavily on data to evaluate system performance and make decisions.  With technology advancements, data is being generated at increasing volume, velocity, and variety.  It is critical for transportation agencies to clearly understand the types of data they currently have and may need in the future; improve existing practices of data collection, storage, and analysis; and stay tuned of the impacts of technology advancements.  This is particularly important for Transportation Systems, Management & Operations (TSMO), which rely deeply on data and are undergoing rapid changes due to new data analysis models (e.g., deep learning) and technologies (e.g., connected and autonomous vehicles)

The proposed research will help DOTs reduce data collection and storage costs; generate more insightful findings for improved decision making; be prepared for the opportunities brought by emerging technologies; and avoid investing in soon-to-be-obsolete technologies.  The payoff of improved decision making can never be overstated.

## VII. PRELIMINARY LITERATURE SEARCH

*To avoid duplicating research already published or in progress, the submitter of the Problem Statement will perform a brief literature search prior to submitting the Problem Statement. This literature search can be conducted using the Transportation Research Board's TRID database, available at* https://trid.trb.org/. *The TRID database contains information on completed research as well as research in progress.*

***Attach a brief summary (1-3 paragraphs)*** *of the results of this literature search to the Problem Statement. The summary should describe how the subject of this Problem Statement would differ from or add to existing studies.*

A search using https://trid.trb.org with two keywords (big data, data analytics) resulted in 136 publications.  Many of them are journal papers and talk about specific statistical methods.  By changing the "Result Type" from "All publications and projects" to "Only projects", a total of 20 projects were identified.  These projects are listed below.  Among them, the most relevant ones are projects c, e, and g (see below).  Overall, transportation data analytics and big data are getting a lot of attention from both state DOTs and federal agencies.  None of the existing studies overlap with the proposed work.

Project c is to "design a prototype interactive, web-based platform that will assist decision makers at MoDOT to seamlessly integrate and analyze transportation datasets."

Project e is to "(1) To develop a framework for identifying, collecting, aggregating, analyzing, and disseminating data from emerging public and private transportation technologies. This framework will address, at a minimum, data from CV/AV deployments as well as other data linked to smart city and related transportation initiatives. (2) To outline a process for using this framework to help decision-makers incorporate data from emerging technologies into transportation planning and policy."

Project g is to "determine whether statistical modeling and/or machine learning can be applied to for estimating/predicting traffic conditions (e.g. vehicle miles of travel (VMT) and reliability) when it is conflated with other available data sets, such as demographics, income, highway capacity, and highway improvement projects. The secondary objective is to develop analytical methods that can be used by UDOT in the future to estimate traffic volumes and reliability for short-duration traffic count sites, based on other available data, and quantitative relationships."

The identified research projects are:

- *Improving Access and Management of Transit ITS Data. [Project].  Transit Cooperative Research Program, Federal Transit Administration. Start date: 8 Jan. 2019.  https://trid.trb.org/view/1577716*
- *DeepHyd: A Deep Learning-based Artificial Intelligence Approach for the Automated Classification of Hydraulic Structures from LiDAR and Sonar Data. [Project].  North Carolina Department of Transportation. Start date: 1 Aug. 2018.  https://trid.trb.org/view/1530096*
- *TITAN - An Interactive Web-based Platform for Transportation Data InTegration and Analytics. [Project].  Missouri Department of Transportation. Start date: 1 Jun. 2018.  https://trid.trb.org/view/1530055*
- *Impact of Transformational Technologies on Land Use and Transportation. [Project].  National Cooperative Highway Research Program, American*

*Association of State Highway and Transportation Officials (AASHTO), Federal Highway Administration. Start date: 1 May. 2018. https://trid.trb.org/view/1467317*

- *Framework for Managing Data from Emerging Transportation Technologies to Support Decision-Making. [Project]. National Cooperative Highway Research Program, American Association of State Highway and Transportation Officials (AASHTO), Federal Highway Administration. Start date: 24 Apr. 2018. https://trid.trb.org/view/1467318*

- *Big Data Visualization and Spatiotemporal Modeling of Aggressive Driving. [Project]. Office of the Assistant Secretary for Research and Technology. Start date: 1 Feb. 2018. https://trid.trb.org/view/1500801*

- *Big Transportation Data Analytics. [Project]. Office of the Assistant Secretary for Research and Technology. Start date: 15 Nov. 2017. https://trid.trb.org/view/1489715*

- *City-wide Transportation Information Estimation with Heterogeneous Urban Data. [Project]. Office of the Assistant Secretary for Research and Technology. Start date: 1 Sep. 2017. https://trid.trb.org/view/1489956*

- *Assessing the Impact of Game Day Schedule and Opponents on Travel Patterns and Route Choice using Big Data Analytics. [Project]. Nebraska Department of Transportation, Federal Highway Administration. Start date: 1 Jul. 2017. https://trid.trb.org/view/1474120*

- *Application of Data Science and Big Data Analytics in Underground Transportation Infrastructure. [Project]. University Transportation Center for Underground Transportation Infrastructure, Maptek, Office of the Assistant Secretary for Research and Technology. Start date: 1 Jun. 2017. https://trid.trb.org/view/1500814*

- *Vehicle Occupants and Driver Behavior: An Assessment of Vulnerable User Groups. [Project]. Office of the Assistant Secretary for Research and Technology. Start date: 1 Jun. 2017. https://trid.trb.org/view/1468650*

- *Using Advanced Analytics to Frame Vulnerable Road User Scenarios with Autonomous Vehicles. [Project]. Office of the Assistant Secretary for Research and Technology. Start date: 1 Mar. 2017. https://trid.trb.org/view/1457860*

- *Completing the Picture of Traffic Injuries: Understanding Data Needs and Opportunities for Road Safety. [Project]. Office of the Assistant Secretary for Research and Technology. Start date: 1 Mar. 2017. https://trid.trb.org/view/1457189*

- *Synthesis of Information Related to Airport Practices. Topic S03-14. Simulation of Airside and Landside Operations in Airport Planning. [Project]. Airport Cooperative Research Program, Federal Aviation Administration. Start date: 14 Dec. 2016. https://trid.trb.org/view/1439993*

- *Synthesis of Information Related to Airport Practices. Topic S03-14. Simulation of Airside and Landside Operations in Airport Planning. [Project]. Airport Cooperative Research Program, Federal Aviation Administration. Start date: 29 Oct. 2016. https://trid.trb.org/view/1428166*

- *Unifying Railcar Monitoring Sensor Data, Maintenance Records, and Railcar Usage Information through Big Data Processing for Optimizing Railcar Maintenance and Safety. [Project]. Nebraska Department of Roads, Advanced Transportation Technology Institute, Research and Innovative Technology Administration. Start date: 1 Oct. 2016. https://trid.trb.org/view/1429401*
- *2300(16-01) - National Performance Management Research Data Set (NPMRDS) – Speed Data Validation For Traffic Performance Measurement. [Project]. Oklahoma Department of Transportation. Start date: 1 Oct. 2016. https://trid.trb.org/view/1372618*
- *Laboratory for Energy Smart Systems (LESS). [Project]. Research and Innovative Technology Administration. Start date: 1 Apr. 2015. https://trid.trb.org/view/1354219*
- *Big Data Analytics to Aid Developing Livable Communities. [Project]. Department of Transportation. Start date: 1 Jul. 2014. https://trid.trb.org/view/1318413*
- *Development of a New Connected Eco-Driving Technology at Signalized Intersections with Adaptive Signal. [Project]. Research and Innovative Technology Administration, University Transportation Research Center. Start date: 1 Mar. 2014. https://trid.trb.org/view/1346331*

## VIII. RESEARCH KEY WORDS

Big Data, Data Analytics

## IX. ENDORSEMENT BY THE SPONSORING DOT (To be signed by the DOT representative to the NETC Advisory Committee through whom the Problem Statement is submitted.)

*By signing the endorsement, the DOT representative is certifying that:*
1. *The Problem Statement follows the required format.*
2. *The required literature search has been conducted.*
3. *The Problem Statement addresses a transportation issue of relevance to NETC and does not duplicate another Problem Statement being submitted at this time.*

_Ann Scholz_____     _New Hampshire_____
Name                                              DOT


_ann.scholz@dot.nh.gov_____     __(603) 271-1659____
Email                                             Tel.


_____     01-25-2019
Signature                                         Date

**X. ENDORSEMENT BY A SECOND EMPLOYEE OF THE SPONSORING DOT who agrees to chair the project's technical advisory committee (TAC) if the Problem Statement is selected for funding.** (To be signed by a DOT staff person who has technical knowledge of the project topic and is committed to the research outcome.)

DOT Technical Endorsement: *I agree to chair the project's Technical Advisory Committee if this Problem Statement is selected for funding by NETC.*

Susan M Klasen

Name

NH

DOT

Susan.Klasen@dot.nh.gov

Email

603 271-6862

Tel.

Susan M Klasen

Signature

1/25/19

Date