

Establishing a Benchmark Dataset for Autonomous Vehicles Using AI-Driven Smart Infrastructure

Overview/Project Description

The San Francisco County Transportation Authority (SFCTA) is seeking funding in the amount of \$1.7 million dollars through the Strengthening Mobility and Revolutionizing Transportation (SMART) grants program. Collaborating closely with the University of Washington (UW), AIWayision, and a consortium of public and private stakeholders, our primary objective of the Stage 1 Planning and Prototyping grant application is to deploy AI-driven smart infrastructure technologies to collect comprehensive data to evaluate Autonomous Vehicle (AV) fleet safety relative to the human driven fleet. One of the promises of AV technology is that it will improve roadway safety by reducing injuries and fatalities relative to human driven vehicles. The critical gap preventing a comparison and assessment of AV and human driving safety is the unavailability of a comprehensive dataset capturing AV and human driven vehicles operating in exactly the same geographies, time periods, and conditions.

The goals of our project are to:

- Demonstrate that it is possible to collect comprehensive and consistent operational, safety, and related data for both human driven and AV fleets; and to
- Demonstrate how to establish and compare human driver safety performance and AV safety performance using these data.

The project will identify comprehensive information describing both AV and human driven motor vehicle fleet operations, as well as other transportation system users such as transit, pedestrians, and bicyclists by leveraging AI-driven smart infrastructure technologies. Safety data will include a range of leading and lagging safety-related events such as crashes, near misses, unplanned stops, and first responder interactions. Critical context data will also be collected, including traffic counts and speeds, vehicle miles travelled (VMT), environmental conditions and roadway conditions.

The project will demonstrate how to collect comprehensive data for a subarea using camera-based instruments and applying AI detection and classification techniques which can be applied consistently across multiple Operational Design Domains (ODDs), multiple operators, and different geographic conditions. Methods for broad, comprehensive and exhaustive data collection and analysis are critical to understanding AV and human-driven

fleet safety. The project will establish a comprehensive benchmark dataset that includes all transportation system users and safety-related events for a subarea of San Francisco. The project will also demonstrate how to conduct a preliminary analysis of AV and human driven fleet safety performance, enhance AV fleet and human driven fleet safety, and guide and support future AV deployments and regulatory frameworks. San Francisco is uniquely positioned to demonstrate how to collect these data given the AV traffic on its streets.

Alignment with SMART Grants Program Goals and Objectives

The proposed project is primarily aligned with the SMART Grants Program's **Safety and Reliability** goal to improve the safety of systems for pedestrians, bicyclists, and the broader traveling public, and to improve emergency response.

The project's characteristics also align with other SMART Grant funding priorities, including

- **Fit, scale, and adoption:** The project is not simply “right-sized” but essential, given the changes that have affected San Francisco’s multimodal transportation system and its diverse users, including the introduction of AVs and Transportation Network Companies (eg. Uber, Lyft), increases in the use of delivery services, and changes in telework patterns. The project will expand AI capabilities to include a more inclusive set of system users, and is designed to be scaled in a Stage 2.
- **Data sharing, cybersecurity, and privacy:** A key product of the project is a multimodal safety benchmarking dataset. The project team will build upon decades of transportation safety experience to identify safety-relevant factors and new data items that can be collected independent of specific technologies. The team will apply best practices to protect individual privacy and meet cybersecurity standards.
- **Measurement and validation:** The project will demonstrate how to fill a critical gap: the data required to measure and compare AV and human driving safety. The project’s data will also be contributed to NHTSA’s AV TEST Initiative so that automated driving systems developers can continue to improve their systems, and to provide AV and human driving insights to the public.

Project Location

The project will be implemented in a subarea of San Francisco. It is anticipated that the subarea will be approximately twenty city blocks. The subarea will be selected based on a number of criteria, such as:

- **Density and diversity of transportation system users:** Areas with greater density and diversity of users will be prioritized.
- **Street network configuration:** Areas with a simpler street network will be prioritized because fewer instruments will be required at each intersection.

Freeway segments and freeway ramps will not be included in the Stage 1 demonstration project due to the increased complexity and associated costs. These facility types could be included in the expanded, scaled deployment proposed for a Stage 2 of this project.

The current proposed subarea is in the Mission District of San Francisco, roughly bounded by 16th Street, 24th Street, Valencia Street and South Van Ness Street. Exhaustive and complete coverage of the project area is a critical element of the study design, so the full extent of the project area cannot be established until the number of instruments required per intersection is determined.

Community Impact

The purpose of the proposed project is to demonstrate how to develop a benchmark data set by collecting comprehensive multimodal transportation system safety-related data, including human drivers and AVs, and to demonstrate how these data can be used to support “apples to apples” safety assessments of AVs and human drivers. Currently, it is not possible to do this because the information because the Operational ODDs for AVs are not comparable to the "ODDs" for human driving. While licensed human drivers are generally permitted to drive anywhere, anytime, on any facility, AV ODDs have, until recently, been limited in their operations by geographies, times, facilities and other conditions. Thus, the rates of different types of events for AVs and human drivers may not be comparable. However, collecting comprehensive data even for a small subarea may pose a risk to citizen privacy. To mitigate this, San Francisco adopted its Surveillance Ordinance, which offers a governance structure and additional safeguards for the deployment of any data collection or processing technologies. The project will be implemented in full compliance with San Francisco’s Surveillance Ordinance. Our partners at PATH will advise on best practices regarding cybersecurity and technology standards. Finally, the proposed prototype will not affect the quality or availability of jobs.

Technical Merit

Identification and Understanding of the Problem to Be Solved

Reliable, objective information about AV and human driving crash rates is critical in order to guide transportation safety and permitting policies. At present, no one has such information. AV companies have detailed information about their own operations but no comparable information about human driving, while policymakers and the general public do not have sufficiently robust information about either human driving or AV driving. This project will demonstrate how to establish a comprehensive dataset inclusive of all transportation system users, collected simultaneously, in the same geographies, and under the same conditions.

There is no comprehensive source for assessing the safety of human driving. NHTSA data is often compiled from exogenous sources that may not be collected or processed consistently. There may also be systematic biases in the human driving crash reporting data due to potential underreporting of minor crashes. VMT, a critical denominator in many safety analyses, is often estimated at an aggregate level and may not reflect the VMT that correspond to safety-related events reported.

There is also no comprehensive source for assessing the safety of AV driving. AV crash and VMT reporting is scattered across many regulatory entities, or not reported at all. For example, AV crashes are comprehensively reported to NHTSA, but VMT is not. California Department of Motor Vehicles (DMV) AV deployment permits require no reporting of either crashes or VMT. The California Public Utilities Commission (CPUC), which regulates AV commercial deployment, publishes VMT at the statewide level, despite the fact that AVs are operating in a number of different local jurisdictions, with radically different numbers of crashes. Local-level VMT information is redacted by the CPUC. Differences in data collection methods, data reporting specifications, and missing context data make it difficult to properly align data for analysis purposes.

The proposed project will demonstrate how to collect and analyze comprehensive data on human and AV driving. This is especially important in San Francisco, where AV crash reporting to NHTSA shows that San Francisco has more AV crashes than all other states, including the rest of California, combined.

Appropriateness of Proposed Solution

The effectiveness of the core technologies proposed as part of this project - AIWaysion's Mobile Unit for Sensing Traffic (MUST) device and their edge AI-based video analytics and AI/ML algorithms for traffic, road conditions, and safety data collection - have been demonstrated by numerous USDOT funded research and deployment projects in City of Bellevue, Washington, and for the Confederated Tribes and Bands of the Yakama Nation. The Yakama Nation project won the prestigious [2023 Innovative Project Award](#) from FHWA's Build a Better Mousetrap program. AIWaysion's technology was featured by the U.S. DOT Volpe Center's as a success story for "[leveraging AI-powered technology to gather traffic safety data](#)". The edge AI-based video analytics technologies proposed in this project have proved the ability to detect, classify, and track different road users such as cars, vans, buses, trucks, pedestrians, and cyclists with a high degree of fidelity, as well as roadway surface conditions, environmental conditions, and visibility conditions. These capabilities will be further extended to detect AVs, TNCs, transit vehicles, and first responders.

AIWaysion's edge AI-based video analytics technologies have also already demonstrated the ability to detect and classify intersection conflict events, and trajectory conflicts ("near

misses”). These capabilities will be further extended to include a broader range of “leading” and “lagging” indicators such as unplanned stops and first responder interactions. Critically, the proposed technology solution will also provide important context information, such as multimodal traffic volumes and turning movement counts by vehicle class (including AVs and trucks), as well as queues, speeding, and wrong way driving (Figure 1).



Figure 1. AI-driven smart infrastructure deployment for comprehensive data collection.

As shown in Figure 1, proposed technology solution can rapidly be scaled because the small instrument profile does not impose any infrastructure requirements other than location(s) to mount the instruments and a power source. The ease of installation has been demonstrated through numerous implementation projects. Stage 1 will prove that the technology can be successfully deployed, and a Stage 2 of the project could capture the broader range of ODDs of AV and human drivers.

This project will transform status quo data collection and reporting because no one – neither AV companies, regulators, the public, nor any stakeholder – has the reliable, consistent, objective information about AV and human driving crash rates that is critical in order to make sound transportation safety and permitting policies. The proposed solution is not only appropriate but is necessary for San Francisco, a dense urban environment with a diversity of users that is the center of AV deployment in the United States.

Expected Benefits

The proposed Stage 1 project is focused on the deployment of proven technologies to fill critical gaps in the information required to address some of the most pressing policy questions regarding the integration of AVs into complex transportation systems, and the relative safety of human and AV driven vehicles. The project can help build public trust in AV technologies by demonstrating transparent monitoring and reporting of the safety of all transportation system users. Ultimately, this will support decision makers as they develop

and evolve policies in the face of rapid technological advancements. The project can also help ensure that this technology benefits all community members, including underserved and vulnerable populations.

The project will develop multiple performance measures as part of the project, including:

1) Validation of AI-driven smart infrastructure data relative to other data available; 2) Validation of agent and event detection and classification by the AI-driven smart infrastructure data relative to empirical observations; and 3) Cost effectiveness of the AI-driven smart infrastructure deployment. Regarding cost-effectiveness, the technologies deployed and data resulting from the project can not only support assessments of safety of human and AV driving, but can also support for a broad range of other safety and non-safety related use cases and goals including the VisionZero and Justice40 initiatives.

The Stage 1 project provides an opportunity to comprehensively document and analyze human and AV driving safety in a way that is not currently possible. In San Francisco, this can help enhance safety, reduce congestion and emissions, improve travel time reliability, and help transit move, and ultimately help address the transportation insecurity associated with people being able to fulfill their daily needs more reliably and safely.

Project Readiness Overview

Feasibility of Workplan

The work program is structured into seven inter-related tasks. Overall project management will be led by SFCTA staff. The analysis needs, framework development, and data needs identification will be led by our academic partners (UW and UCB), with input from technical advisors, the SFCTA, and agency partners. Refinement of the AI detection and classification methods, as well as instrument deployment and data collection will be led by AIWayson and UW. Data analysis and evaluation will be led by our academic partners (UCB and UW) and technical advisors, while the final report will be produced by the SFCTA. The SFCTA will also lead all public outreach.

1. Program Management
2. Analysis Needs and Framework Development
3. Data Needs Identification
4. Detection and Classification Refinement
5. Deployment and Data Collection
6. Data Analysis and Evaluation
7. Final Report and Dissemination

The pilot will take place over 18 months, with approximately 3 months of design, 9 months for deployment, data collection and refinement; and 6 months for analysis and evaluation.

Community Engagement and Partnerships

SFCTA is committed to community-centered stakeholder engagement in all our work, and the proposed project location is entirely within or bordering an Equity Priority Community in San Francisco. Stakeholder outreach will occur at regular intervals during the project, but will be concentrated in the early project tasks to help support the development of the analysis needs and framework development as well as data needs identification, and in the later project tasks supporting data analysis and evaluation. The project team will also establish a Technical Advisory Committee, in order to collaborate with industry, public agencies, and the broader community to ensure that a comprehensive set of required data and analysis needs are identified.

Leadership and Qualifications

SFCTA has managed a number of multi-stakeholder federal US DOT grants in the past, including a federal US DOT ATCMTD grant (as a subrecipient to the SFMTA), a regional IDEA-SAV for Treasure Island AV Shuttle, (as subrecipient to MTC) (\$828,000 awarded in 2022); and the eFleet: Carsharing Electrified project (\$1.7M awarded in 2011).

The key partners for this pilot are:

AIWaysion: AIWaysion will serve as the technology partner for deploying its AI-driven smart infrastructure technologies and edge AI-based video analytics and AI/ML algorithms, and will collaborate with team members to build the benchmark dataset. Dr. Wei Sun has served as the PI for several US DOT SBIR projects, and is the technology partner for the Yakama Nation team's FY23 SMART grant.

University of Washington: Dr. Yinhai Wang, the project PI, will lead the development of video analytics algorithms, build benchmark datasets, and conduct safety performance evaluations. Dr. Wang directs UW-STARLab and the Pacific Northwest Transportation Consortium (PacTrans), Northwestern Tribal Technical Assistance Program Center (NW TTAP). Dr. Wang is a fellow of ASCE and ITE. He served as the 2019 president of ASCE T&DI, and chairs the TRB AED50 AI and Advanced Computing Applications Committee.

ITSA: role will be to help with outreach to industry and final report/findings dissemination task (organize/host webinars, conference panels).