Exhibit D – Research Project Requirement Template

Beyond Pedestrian Flow: Evaluating Urban Sidewalk Friction and Accessibility via Efficient State Space Models

Recipient/Grant (Contract) Number: The University of Texas at Austin/Grant # 69A3552344815 and 69A3552348320

Center Name: National Center for Understanding Future Travel Behavior and Demand (TBD)

Research Priority: Improving Mobility of People and Goods

Principal Investigator(s): Zhigang Zhu, Alison Conway

Project Partners: Professor Xuan Wang, Machine Learning, CUNY City Tech

Research Project Funding: $200,000 ($100,000 Federal and $100,000 CUNY Match)

Project Start and End Date: 06/01/2026 - 05/31/2027

Project Description: Traditional sidewalk performance evaluation has primarily relied on flow metrics, such as level of service, that treat pedestrians as uniform units, similar to vehicles on a highway. These methods typically rely on basic pedestrian intersection or screenline counts, which capture pedestrians at a single point location. More recently, advanced sensors such as video- and LIDAR have enabled automation of the counting process. Anonymized probe-based data sources have enabled observation of aggregated metrics like speed or travel time at a segment scale, but these data sources are biased and challenging to validate at a population scale. Walkability metrics typically assess the suitability of a street environment for pedestrian activity based on the presence or absence of specific neighborhood or street characteristics (e.g. business densities, greenery), usually using local land use or point-of-interest data, or more recently, Streetview imagery. While useful to capture high-level neighborhood pedestrian activity, these approaches fail to capture the complex, micro-level dynamics that occur in pedestrian spaces, where individuals walk for a wide range of purposes including but not limited to commuting, making deliveries, running errands, exercising, and socializing. A street that well serves the needs of one group of these users might present an undesirable environment for a disparate group, but current assessment strategies do not capture these differences. When user needs are not met, these individuals may experience congestion or delay, may be discouraged from choosing to walk or use specific routes, and may engage in dangerous behaviors such as encroaching on vehicle or bicycle lanes.

To address this gap, this project introduces a comprehensive framework for evaluating sidewalk performance by developing novel *friction metrics*. By moving beyond simple counting, we aim to quantify how different pedestrian groups and infrastructural elements interact in real-world scenarios. To process continuous, real-world spatial-temporal data, our architecture leverages the combined strengths of Visual Language Models (VLMs) and highly efficient State Space Models (SSMs). Traditional VLMs provide databanks of vast knowledge trained into their latent space and thus can perform a majority of novel tasks without requiring specific training data. However, traditional vision transformers struggle with extended sequence lengths. SSMs offer superior long-term memory retention. Working in tandem with VLMs, this integrative approach provides a scalable, nuanced method for understanding urban pedestrian dynamics and offering actionable insights for pedestrian infrastructure planning and accessibility improvement.

The project is structured around three primary objectives/tasks:

Objective 1: Semantic Categorization via Visual Language Models (VLMs): The first task involves processing CCTV camera feeds to achieve a deep contextual understanding of the sidewalk environment, relying on publicly available data feeds to minimize privacy concerns. Rather than merely detecting moving blobs, we will deploy SOTA VLMs to categorize pedestrians into distinct behavioral groups, such as commuters, workers, and social clusters. Furthermore, the VLMs will also identify and classify various temporary and permanent infrastructural hurdles (e.g., construction barriers, crowded vendors, garbage bags, and sidewalk conditions).

Objective 2: Development of SSM-Based "Friction Metrics": The second task focuses on developing the mathematical and computational framework for the "friction metrics". Using the rich semantic data generated by the VLMs, we will utilize SSMs operating on anomaly detection principles to analyze specific pedestrian interactions over time. The SSM will calculate relevant performance metrics for individual user groups (e.g. speed or path deviations for commuters and workers, continuous cohesion of social groups), as well as an overall "friction" caused when distinct behavioral groups navigate around each other and the identified infrastructural hurdles.

Objective 3: Walkability /Accessibility Evaluation and Sidewalk Planning Integration
The final task will synthesize the outputs of the pipeline to evaluate large-scale urban walkability and accessibility by creating and utilizing benchmark data. We will apply the friction metrics across various temporal conditions to pinpoint persistent time-specific accessibility bottlenecks and isolate the impacted user groups.

US DOT Priorities: This proposal addresses both of USDOT's core priorities of transformation and safety. The project proposes a novel technology-driven machine learning architecture that integrates VLMs and SSMs to efficiently derive unique performance metrics from publicly available CCTV data. Unlike commonly-used sidewalk performance evaluation approaches, this approach will account for the unique behaviors of distinct groups of sidewalk users, and will enable both user-specific and overall assessment of existing and proposed sidewalk conditions to inform design and management decisions. The outputs will serve as a diagnostic tool to identify locations and times in need of design (e.g. width, slope, pavement quality, curb cut) or management intervention (e.g. wayfinding signage, curb management, sidewalk use restrictions such as waste management or outdoor dining, prioritization of maintenance or enforcement actions), enabling agencies to better manage potential conflicts and discourage dangerous walking behaviors.

Outputs: In addition to a ***final project report***, this study is expected to produce a ***practical guide*** for implementing the proposed approach and for utilizing outputs for design and management decision-making. The project will produce a ***robust software framework*** utilizing VLMs to automatically extract and categorize pedestrian behaviors from CCTV feeds and will produce an ***open-source, computationally efficient algorithm*** for implementing the SSM-based friction metric calculation. Results will be further disseminated via academic articles in refereed journals and conference proceedings, and will be presented at relevant professional conferences. Codes will be shared through a GitHub repository, and deliverables will be disseminated via TBC, CCNY, and research team social media.

Outcomes/Impacts: The tools and practical guidance developed in this project will give transportation planners an immediately actionable approach for data-driven sidewalk design and management that recognizes the unique behaviors of different sidewalk users and the interactions between them. This project establishes foundational models and increases the body of knowledge for applied researchers in introducing a new approach to quantifying sidewalk performance. The collaboration of PI (AI/ML) and Co-PI (transportation), and the hands-on experience of the two PhD students in computer vision and machine learning, will ensure development of tools that are methodologically advanced and practically useful.

Final Research Report: A URL link to the final report will be provided upon completion of the project.