Exhibit D – Research Project Requirement Template

Potential Use of Large Language Models (LLMs) for Travel Behavior Survey Research

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): Chandra R. Bhat

Project Partners: N/A

Research Project Funding: $150,000 ($75,000 Federal + $75,000 matching funds)

Project Start and End Date: 6/1/2025 – 5/31/2026

Project Description: Traditional travel behavior surveys are resource-intensive and constrained by challenges such as small sample sizes, response bias, and high costs. With the rapid advancements in artificial intelligence, particularly LLMs, we now have the opportunity to explore novel, cost-efficient methods for creating synthetic data comparable to real-world survey results. However, the potential of LLMs to contribute to transportation research remains underexplored. This project seeks to evaluate the role of LLMs in replicating human decision-making in the context of travel behavior, investigate their strengths and limitations, and identify biases to develop methodologies for integrating them into survey-based design, data collection, and analysis workflows. This project is structured around three primary objectives:

1. **Comparative Benchmarking:**
   - Assess the quality of synthetic datasets generated by LLMs across three models (e.g., GPT-4, Claude, and Mistral) by comparing their statistical properties and behavioral narratives to the 2022 NHTS dataset.
   - Identify key areas of similitude and deviation to establish benchmarking protocols.
2. **Predictive Validation:**
   - Generate AI-based predictions for an upcoming stated-preference survey on Unmanned Aerial Vehicles before data collection occurs.
   - As human data collection proceeds, monitor similitude between AI-generated predictions for key behavioral questions and actual responses.
   - Compare all AI-generated predictions to human responses post-data collection to evaluate their validity.
3. **Sensitivity via Ablation Studies:**
   - Explore sensitivity to detailed prompts by varying input attributes (e.g., socio-demographics, built environment, personal attitudes).
   - Develop prompting best practices for improved reliability and consistency in survey integration.

By benchmarking LLM responses against real-world data, this research aims to enhance data collection mechanisms, investigate inherent biases in AI-generated surveys, and develop best practices for integrating AI into travel behavior research. This work has the potential to illuminate the potential benefits (low-cost, customizable, and constantly evolving tool) and pitfalls (biases, unreliability) of using today's LLMs in transportation data collection and modeling by investigating and evaluating how they complement traditional survey paradigms while ensuring transparency and accountability. The research outputs from this project will be added to the Travel Behavior Data (TBD) Hub funded through the TBD UTC and under

development till 2029. A separate tab for AI-generated silicone samples, benchmarks, and data comparisons for revealed and stated preference surveys will be added to the TBD hub providing transparent access to ongoing research.

US DOT Priorities: This research aligns with the Transformation Priority in USDOT's RD&T Strategic Plan (pages 50-62), specifically focusing on advancing innovative data collection and modeling mechanisms that leverage artificial intelligence for transportation research. By exploring the applicability of LLMs, the project supports research that amplifies the potential for transformative technologies in addressing current and future transportation challenges. Additionally, it addresses user needs considerations (pages 33-40) by identifying and mitigating biases in AI outputs, ensuring generalizability in synthetic data applications. These contributions directly enhance the TBD Center's focus on improving survey methodology and addressing emerging technological implications in transportation planning and policymaking.

This project contributes to the Technology Transfer Priority outlined in USDOT's RD&T Strategic Plan (pages 63-68). The developed benchmarking protocols, prompting workflows, and sensitivity evaluation tools are designed for public dissemination through open-source platforms, research publications, and conference presentations. Deliverables also include accessible guidelines and best practices for leveraging LLMs in transportation research. These outputs aim to foster greater adoption of AI-assisted survey methodologies across academic, governmental, and industry stakeholders, contributing to increased research efficiency, scalability, and decision-support capabilities in transportation systems.

Outputs: The proposed research study shall result in a paper that will be submitted to a relevant journal. Also, the paper will be submitted to and presented at relevant conferences and meetings. Additionally results from the effort will be included as part of the Travel Behavior Data (TBD) Hub, including: synthetic travel behavior datasets developed using LLMs, benchmarking protocols for comparing AI-generated and human survey data, open-source tools for integrating LLMs into survey methodologies, and best practices guide to mitigate biases in synthetic data generation.

Outcomes/Impacts: This research will enhance understanding of LLM applications in transportation surveys and data modeling. By validating AI-generated synthetic datasets and benchmarking their performance, the project will expand the body of knowledge in survey design and data collection mechanisms. The deliverables will evaluate whether broader adoption of AI tools in addressing real-world transportation challenges, such as mode choice modeling and behavior prediction, is warranted (as of today). Potential outcomes include improved survey cost efficiency, scalable data collection methods, and actionable protocols for mitigating AI biases, ensuring generalizability and reliability in synthetic data generation for transportation planning.

The project's findings are expected to advance transportation research by investigating the potential of LLMs in supplementing traditional survey techniques. Impacts include:

- Efficiency Gains: Reduced survey costs and faster data collection timelines.
- Knowledge Advancement: Greater awareness of AI tools in transportation systems.
- Practice Change: Improved evidence-based decision-making through AI-integrated survey workflows.

By exploring the transformative potential of AI, this research contributes to sustainable, cost-effective, and efficient transportation planning, addressing both current and emerging challenges in travel demand modeling.

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