limos lab for innovative mobility systems

Workshop on Big Data, AI and Transportation Planning Applications, University of Washington, Seattle, WA, May 30, 2025

LLM-Agent-Based Simulation for Travel Demand Forecasting*

Yafeng Yin, Ph.D.
Donald Cleveland Collegiate Professor of Engineering
Donald Malloure Department Chair of Civil and Environmental Engineering
University of Michigan, Ann Arbor

Joint work with **Tianming Liu, Jirong Yang, Yingnan Yan and Manzi Li @ University of Michigan*

Evolution of Travel Forecasting Models

Societal needs

- More cars need more highways
- More about humans, less about cars
- Technologies are transforming our travel

First generation (1950s-1970s)

SOCIAL PHYSICS APPROACH

Second generation (1970s-Present)

MICRO-BEHAVIORAL APPROACH

Next-generation forecasting models

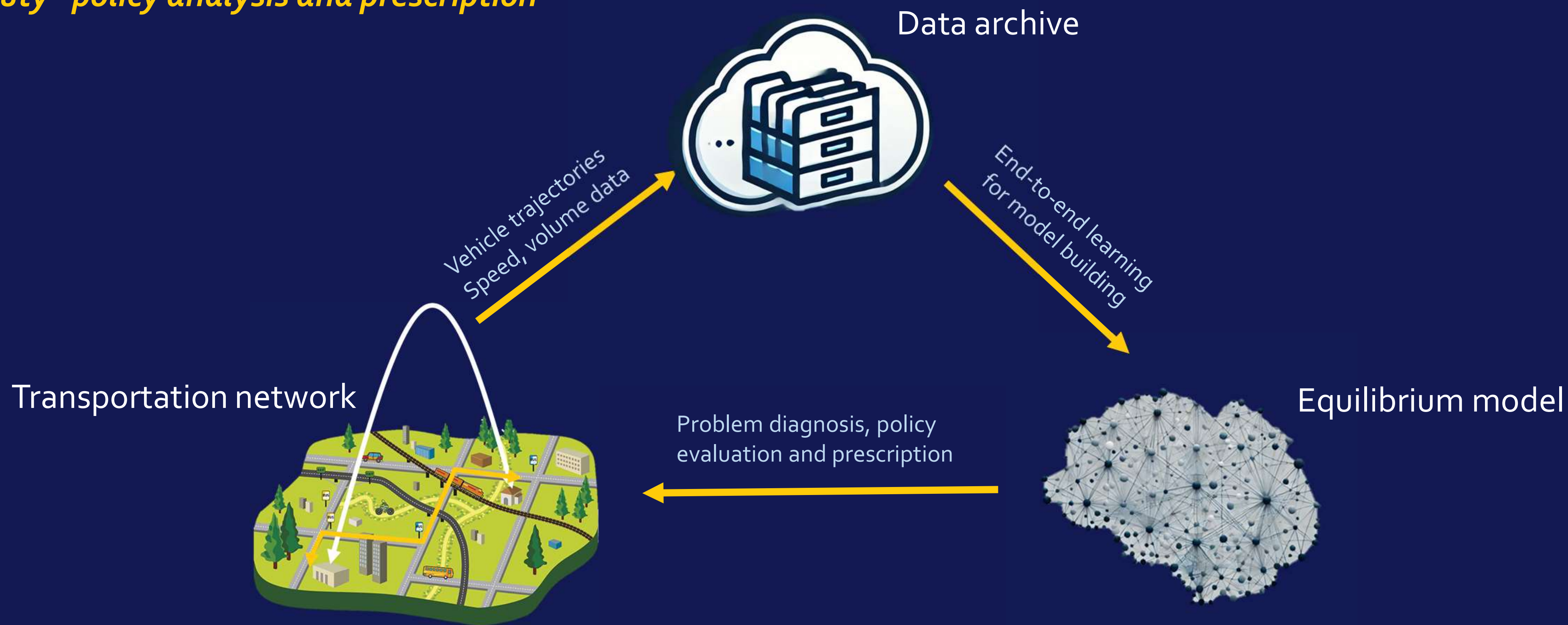
?

Technological advancements

- Operations research
  Game theory
  Computer
- Econometrics
  Behavioral science
  Microsimulation
- Artificial intelligence

Leveraging Big Traffic Data

End-to-end learning approach for constructing integrated network equilibrium models for "light-duty" policy analysis and prescription*



Liu, Z., **Yin, Y., Bai, F. and Grimm, D. (2023) End-to-end learning of user equilibrium with implicit neural networks. Transportation Research Part C, 150, 104085.*
*Liu, Z. and **Yin, Y**. (2025) End-to-end learning of user equilibrium: expressivity, generalization, and optimization, Transportation Science, https://doi.org/10.1287/trsc.2023.0489*

LLMs for Agent-Based Microsimulation

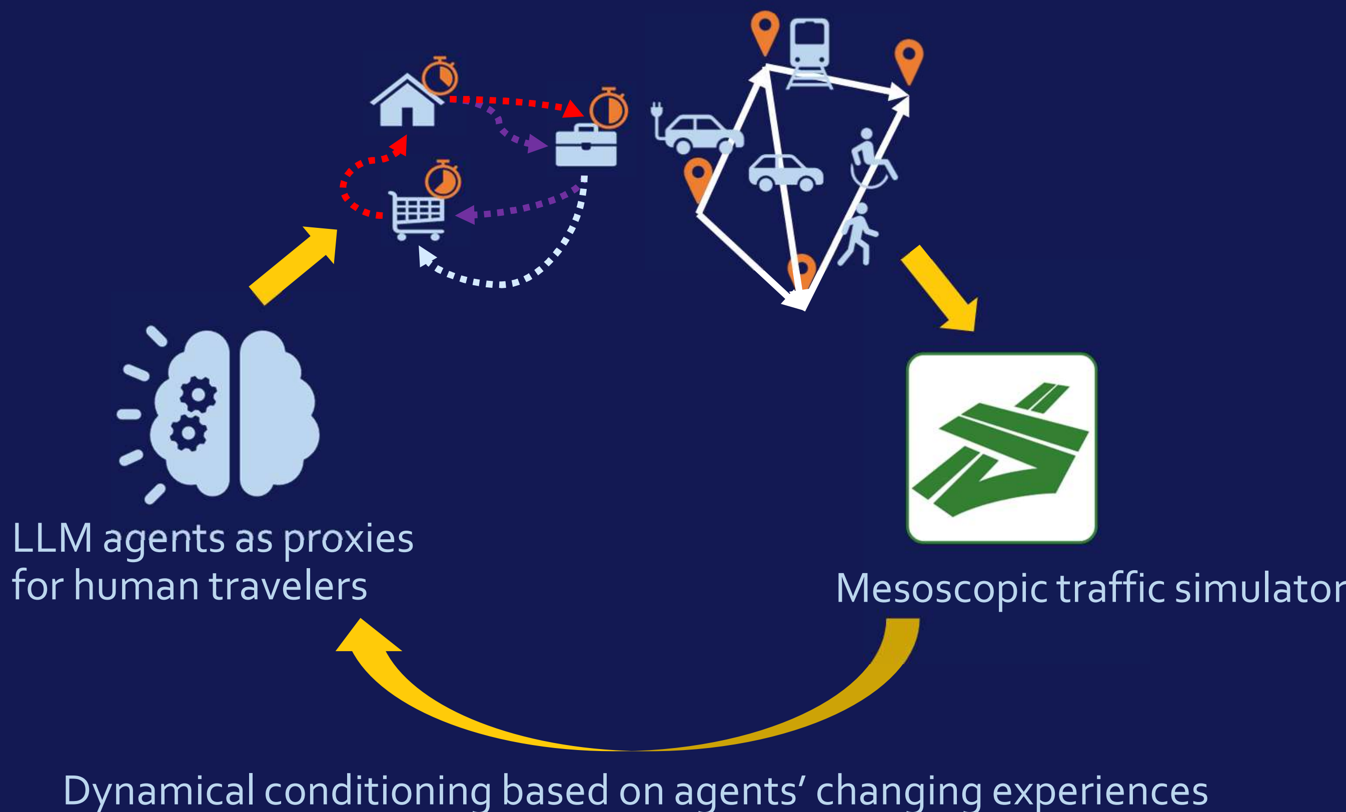
LLMs *encode a wide range of human behaviors from their training data. When prompted with a narrowly defined context, these models* **can generate believable, human-like responses and behaviors***

If LLMs can replicate travel behaviors, *then we can use LLM-powered agents to transform agent-based microsimulation for activity-based demand modeling. The models are designed for in-depth, comprehensive analysis of transportation plans and policies*

- **The decision rules for these agents emerge from the wealth of knowledge encapsulated in the LLM (aka parametric knowledge)**, *rather than being imposed by the modelers*

** Grossmann, et al., AI and the transformation of social science research, Science, 2023, 380 (6650).*
Hutson, M., Can AI chatbots replace human subjects in behavioral experiments. Science, 2023, 381(6654).

LLM-Agent-Based Simulation for Activity-Based Modeling



Can LLM Agents be Proxies for Human Travelers?

Evidence and insights from investigating the value-of-travel-time (VOTT) of LLM agents

Measuring VOTT

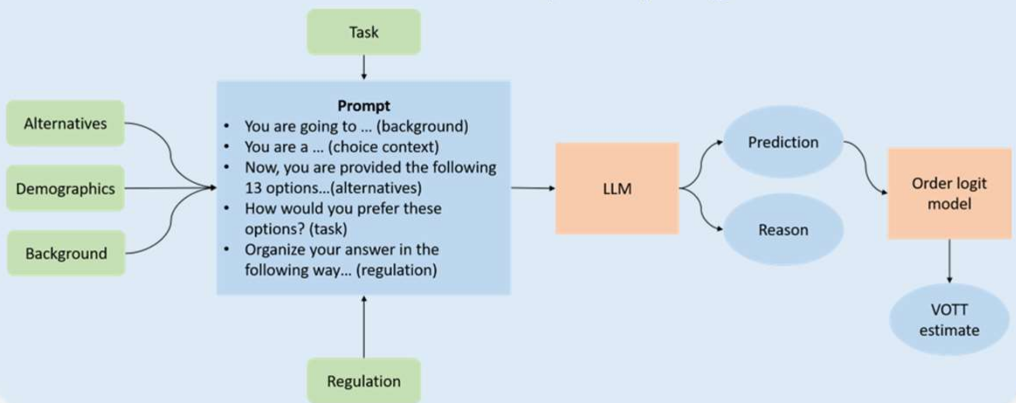
- LLM agents provide responses (ratings and a ranking of options) to route-choice surveys designed by Calfree et al (2001)*
- Use LLM's responses to calibrate a ranked-order logit model and measure VOTT:

$$U_{in} = \beta_c c_{in} + \beta_t t_{in} + \beta_g x_{in} + \epsilon_n$$
$$\text{VOTT} = \beta_t / \beta_c$$

Experiment Design

- Full factorial design experiment to control social-demographics and travel situation
- On each run we repeat survey 60 times on GPT 40 with temperature 1.

LLM zero-shot prompting



Factors	Levels			
Purpose	Leisure	Personal	Commute	Business
Age	25-29		55-59	
Sex	Male		Female	
Education	High-school		College	
Wage	$15/hour	$25/hour	$35/hour	$50/hour

*J. Calfee, C. Winston, R. Stempski (2001) Econometric issues in estimating consumer preferences from stated preference data: a case study of the value of automobile travel time, Review of Economics and Statistics, 83, pp. 699-707

LLM Agents' VOTT

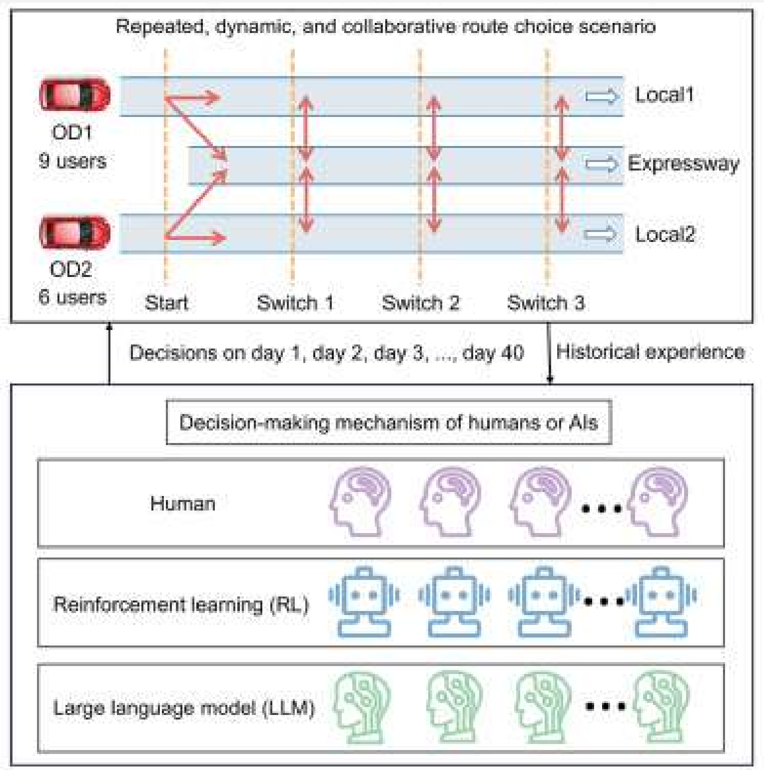
Factor	Level	GPT-4o	Factor	Level	GPT-4o
Purpose	Leisure	**$7.12/h**	Wage per hour	$15	**$6.47/h**
	Commute	**$8.54/h**		$25	**$7.80/h**
	Business	**$8.22/h**		$35	**$8.38/h**
				$50	**$8.77/h**

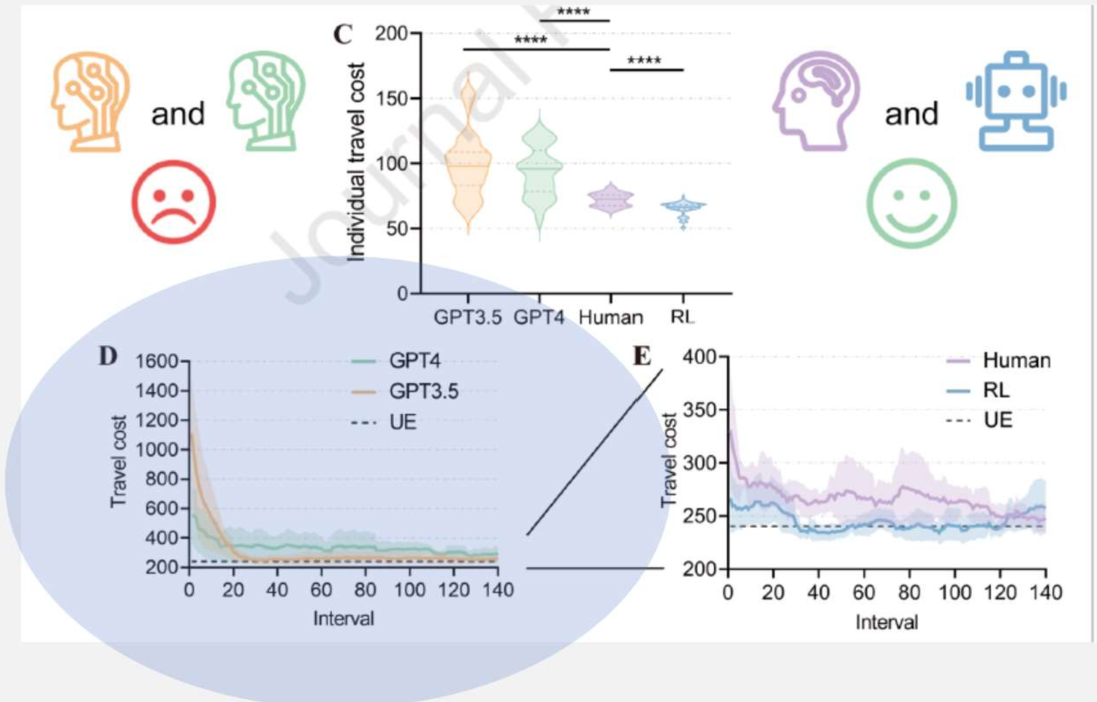
Purpose	Income Elasticity			
	USDOT	Binsuwadan et al. (2023)*	Shires and de Jong (2009)**	GPT-4o
Commute	**1**	**0.37**	**0.67**	**0.24**
Business	**1**	**0.53**	**0.47**	**0.21**
Leisure or other	**1**	**0.53**	**0.52**	**0.22**

**Binsuwadan, J., Wardman, M., de Jong, G., Batley, R., and Wheat, P. (2023). The income elasticity of the value of travel time savings: A meta-analysis, Transport Policy, Volume 136, 126-136.*
***Shires, J. D., & de Jong, G. C. (2009). An international meta-analysis of values of travel time savings. Evaluation and program planning, 32(4), 315-325..*

Learning and Adaptation?

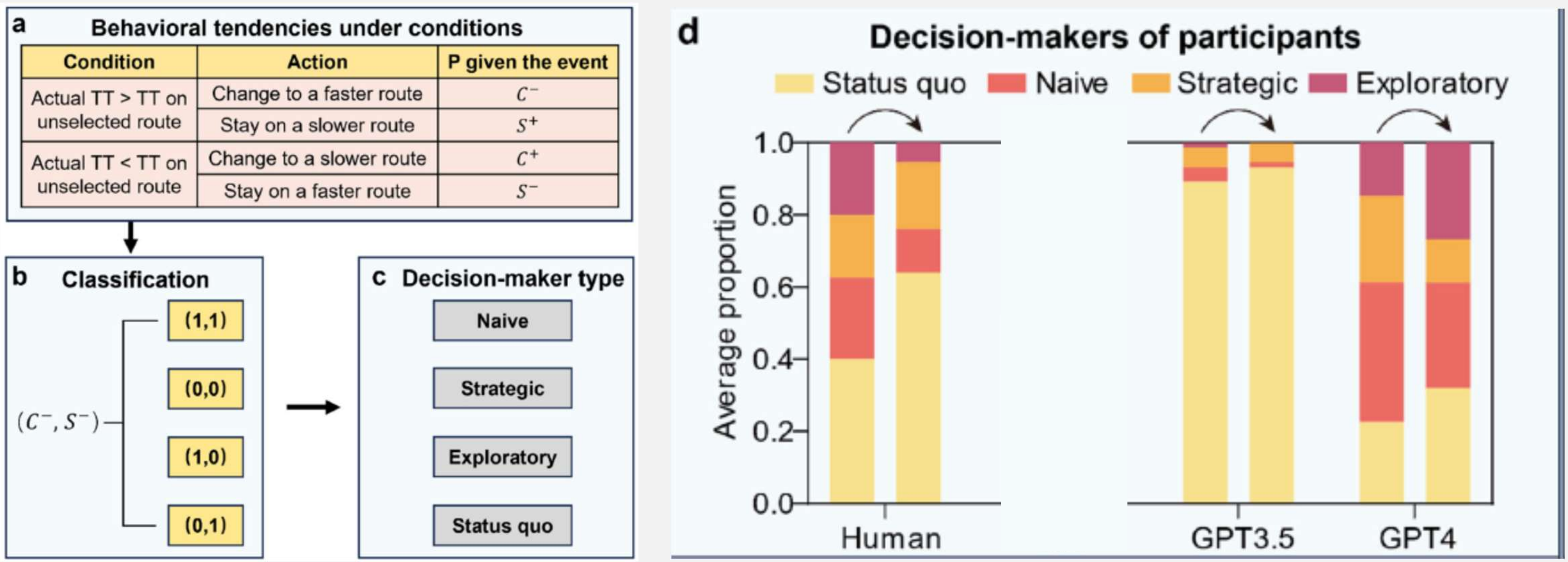
*Evidence and Insights from a Commuting Route Choice Lab Experiment**





*Wang, L., Jiang, Z., Hu, C., Zhao, J. , Zhu, Z. Chen, X., Liu, L., He, G., **Yin, Y.** and Lee D-H. (2025) Comparing AI and Human Decision-Making Mechanisms in Daily Collaborative Experiments, iScience. https://doi.org/10.1016/j.isci.2025.112711*

LLMs exhibit different adaptation behaviors.



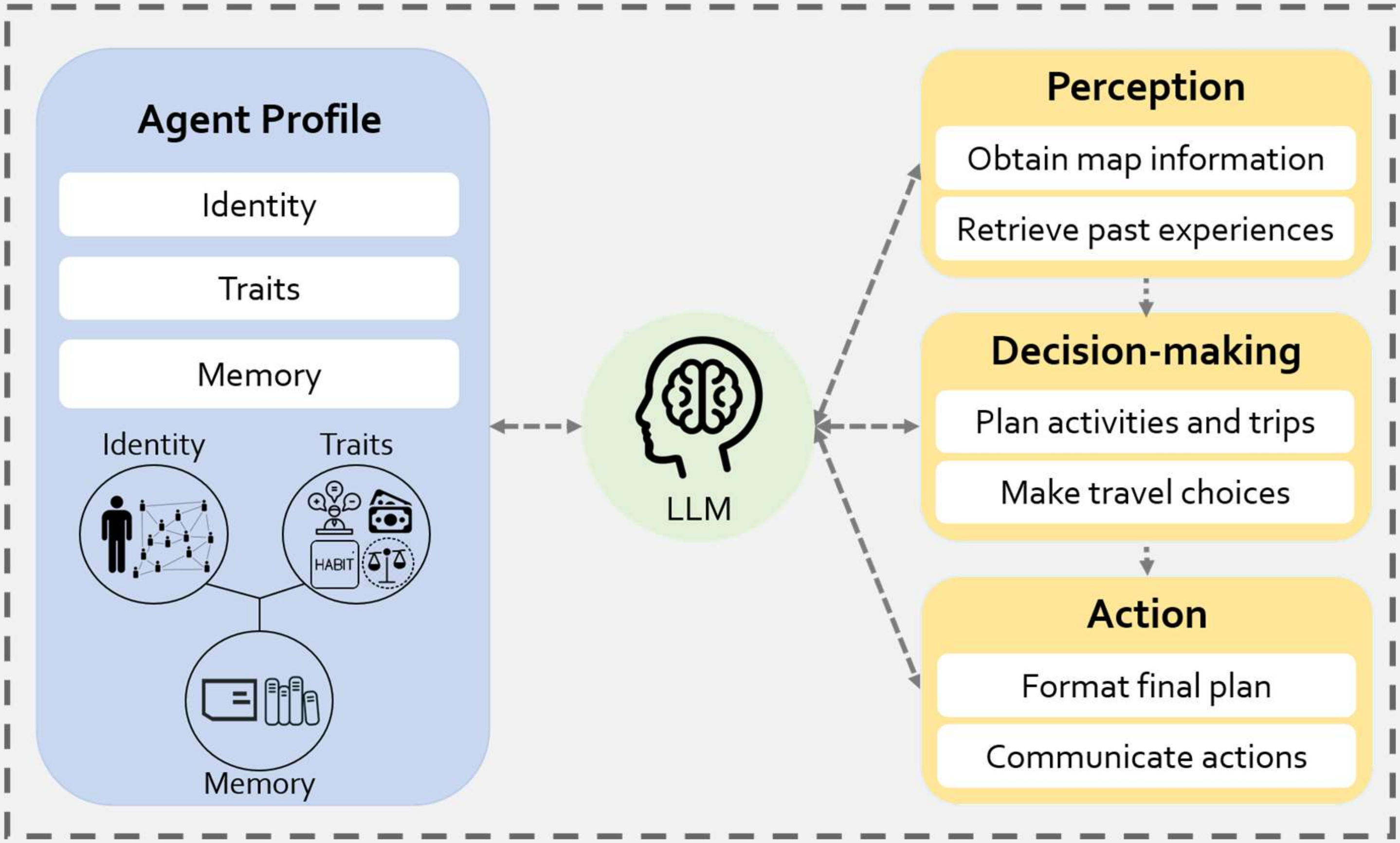
Observations

- LLM agents demonstrate human-like behavioral responses to variations in sociodemographic characteristics and travel choice contexts.
- However, incorporating these characteristics and contexts alone does not guarantee that LLM agents will replicate authentic human travel behavior without targeted training or behavioral alignment.*
- LLM agents can learn from historical experiences in a human-like manner, adjusting their travel decisions.
- A well-designed memory system is essential for enabling LLM agents to learn from past experiences and adapt their travel choices over time, as real travelers do.**

*Liu, T., Li. M. and **Yin, Y.** (2024) Can Large Language Models Capture Human Travel Behavior? Evidence and Insights on Mode Choice. SSRN: https://ssrn.com/abstract=4937575.

Liu, T., Yang, J. and **Yin, Y. (2025) LLM-ABM for Transportation: Assessing the Potential of LLM Agents in System Analysis. The 1st Workshop on AI for Urban Planning, 39th Annual AAAI Conference on Artificial Intelligence. March 3, 2025, Philadelphia, Pennsylvania, USA.

LLM Agents*



*Liu, T., Yang, J. and **Yin, Y.** (2025) Toward LLM-agent-based modeling of transportation systems: A conceptual framework. Artificial Intelligence for Transportation, https://doi.org/10.1016/j.ait.2025.100001. arXiv preprint arXiv:2412.06681.*

Prompting to Align LLM Agents' Behaviors with Human Behavior

Social demographics Choice context

Choices by LLM

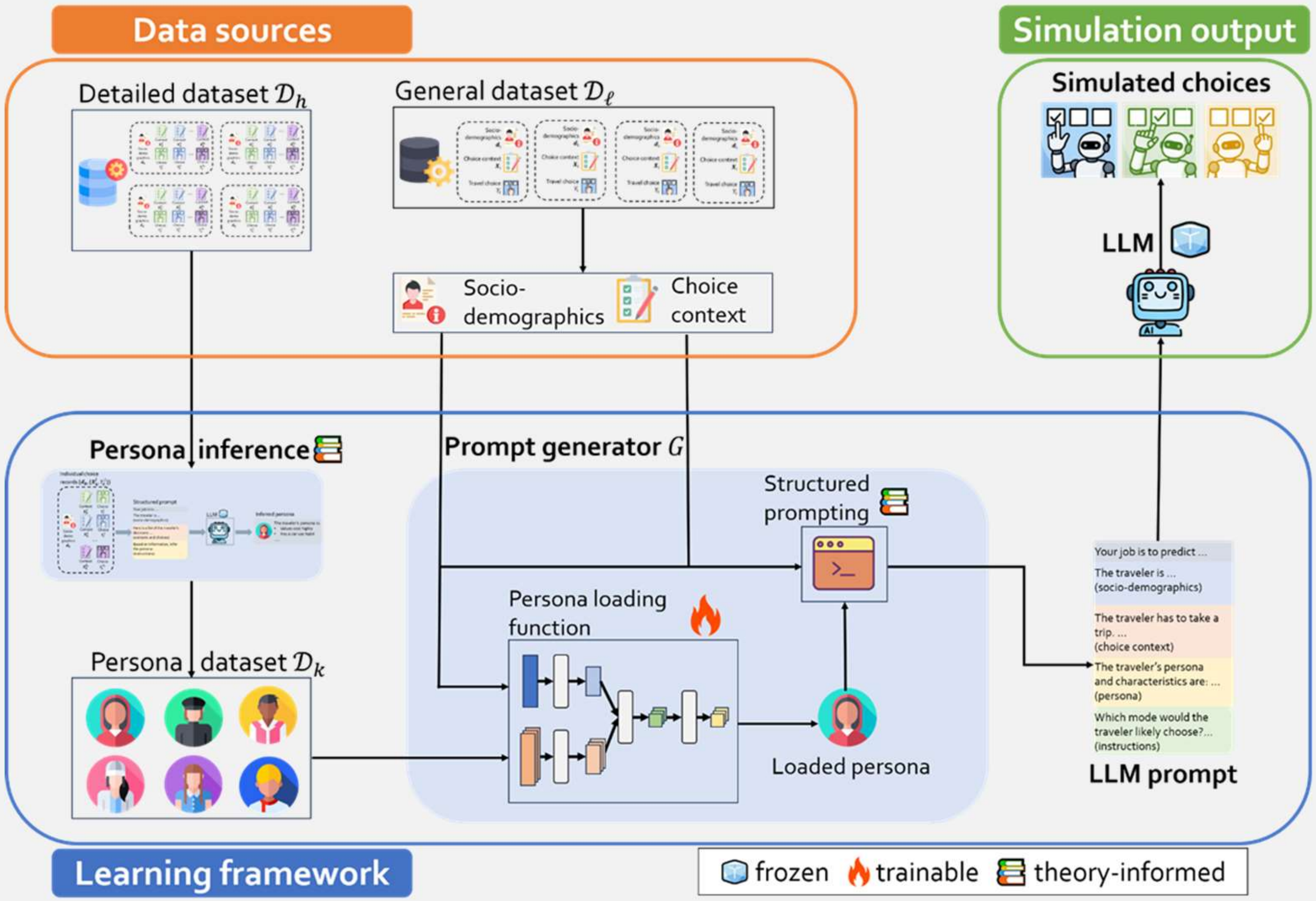
$$\hat{Y} = LLM(G_{\Phi}(d, X))$$

Prompt generation function Function parameters

LLM behavioral alignment problem:

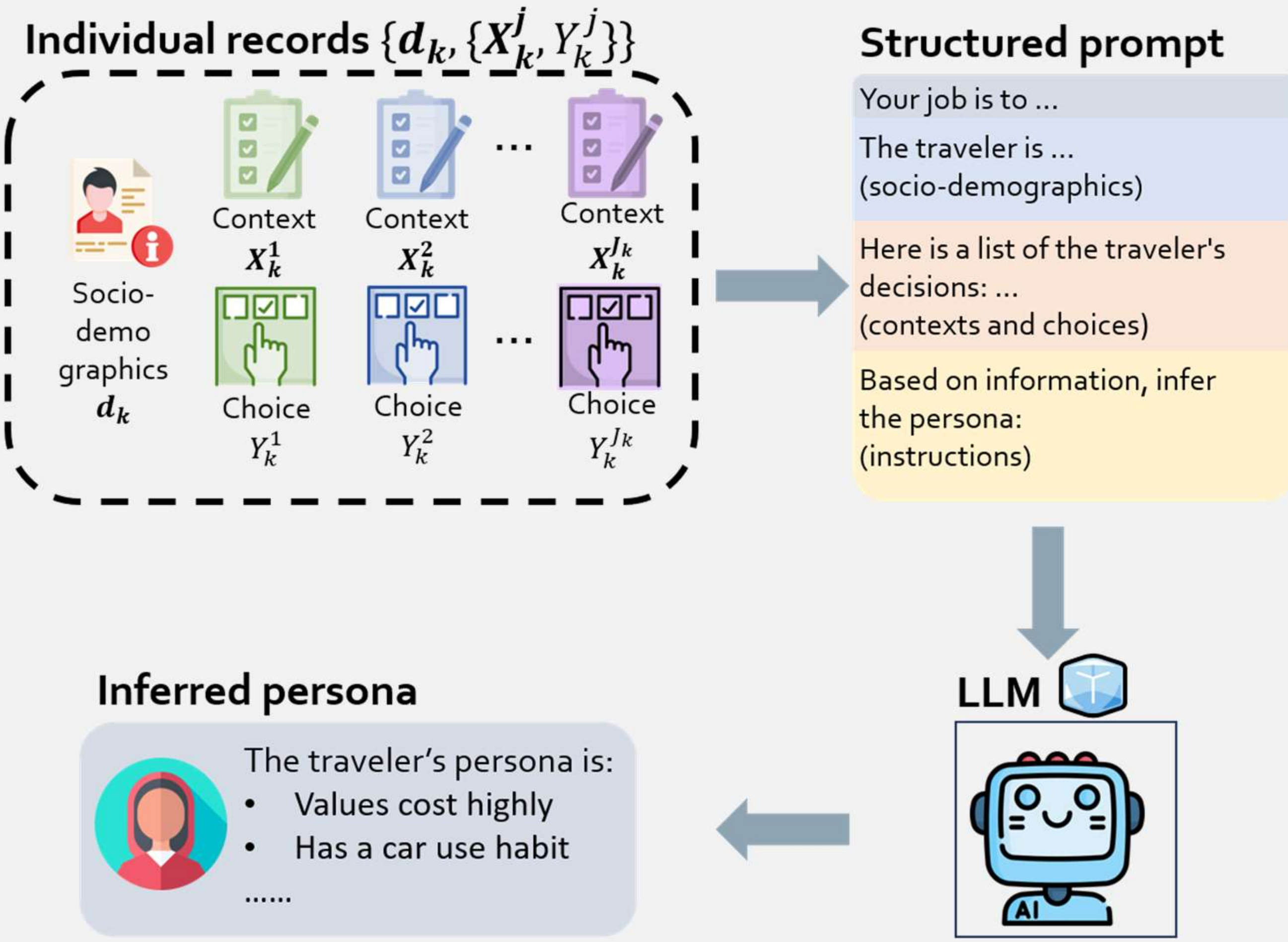
$$\max_{\Phi} \sum_{Y \in \mathcal{Y}} \iint_{\mathcal{X},\mathcal{D}} (\log \mathbb{P}[LLM(G_{\Phi}(d,X)) = Y])\, f_{\mathcal{P}}(d,X,Y) \mathrm{d}d \mathrm{d}X$$

Persona Inference and Loading Approach *

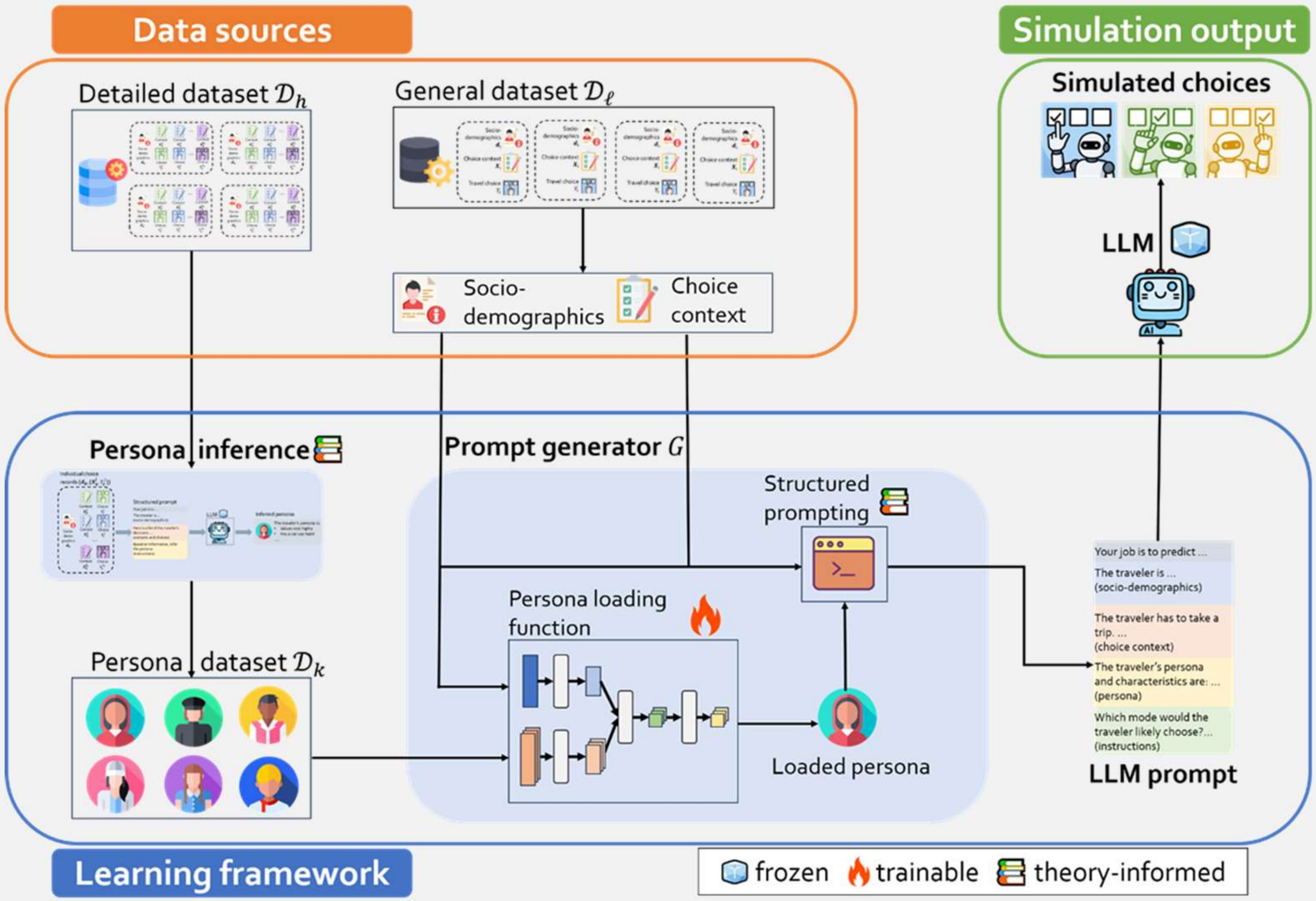


*Liu, T. and Li, M. and **Yin, Y.** (2025) Aligning LLM with human travel behavior: a persona-based embedding learning approach, https://arxiv.org/abs/2505.19003.*

Persona Inference

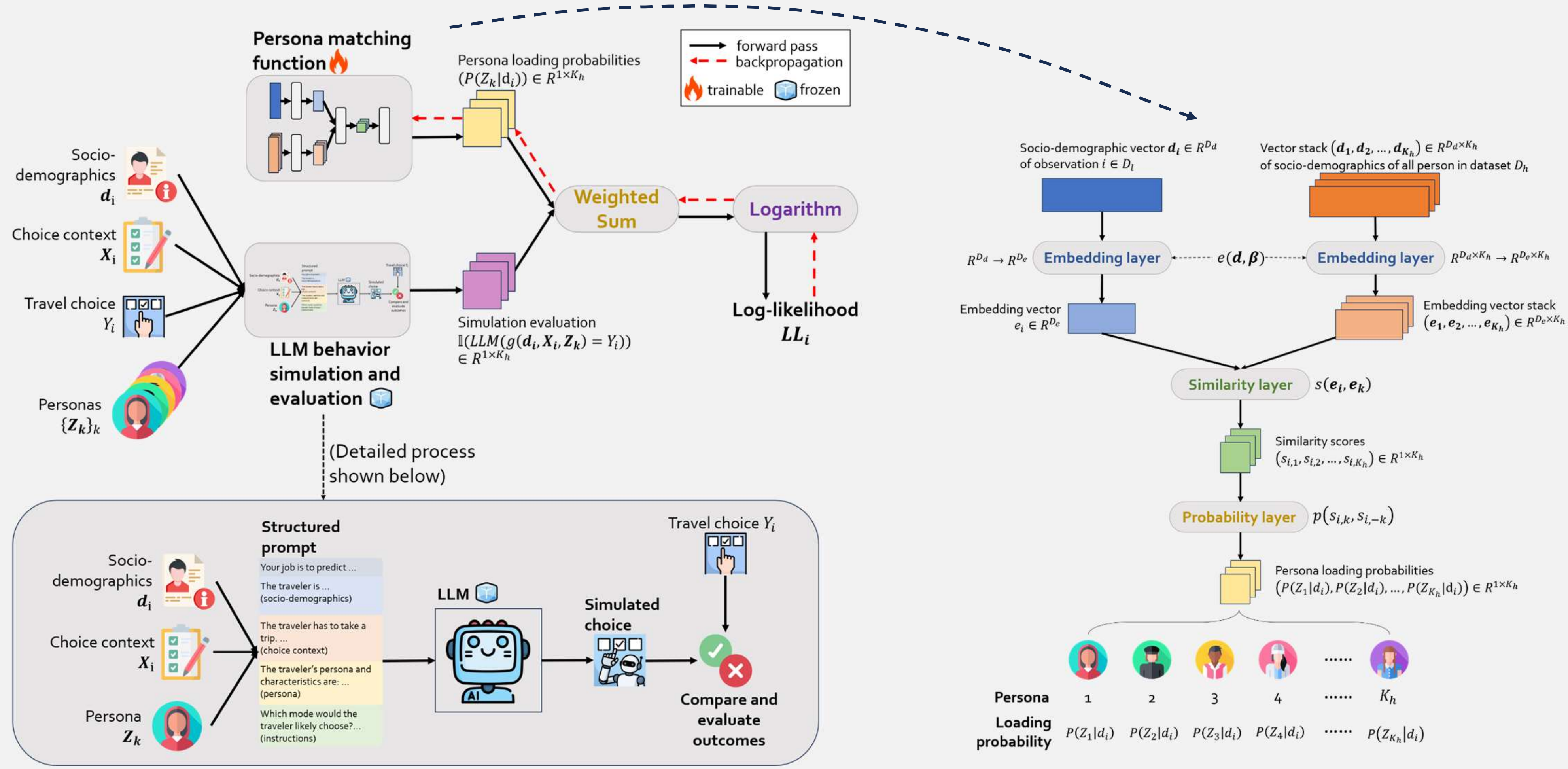


Persona Inference and Loading Approach *



*Liu, T. and Li, M. and **Yin, Y.** (2025) Aligning LLM with human travel behavior: a persona-based embedding learning approach, https://arxiv.org/abs/2505.19003.*

Training Persona Loading Function



Results

Example on Mode Choice using the Swissmetro Dataset*

Method	Mode split			Jensen-Shannon Divergence	Macro F1-score	Weighted F1-score
	Train	Swissmetro	Car			
Ground truth	6.0%	53.3%	40.7%	0.000	1.000	1.000
MNL	2%	77.5%	20.5%	0.483	0.474	0.606
Zero-shot LLM	3.3%	70.0%	26.7%	0.216	0.407	0.543
Few-shot LLM	3.5%	65%	31.5%	0.108	0.429	0.594
Same-group persona	3.5%	49.7%	46.8%	0.044	0.542	0.657
Our method	4.0%	51.7%	44.3%	**0.021**	**0.556**	**0.683**

**Bierlaire, M., K. Axhausen, and G. Abay, The acceptance of modal innovation: The case of Swissmetro. In Swiss Transport Research Conference, 2001*

Proof-of-Concept Demonstration

- 10 LLM agents, each representing a household

Agent

Identity

The household has dad, who works at the company from 8am to 5pm, …

Traits

The household is a typical suburban household and they do all their travels by driving.

Activity memory

- At day 1, mom go grocery shopping.
- At day 1, dad go to work.

…

Travel memory

The household is a typical suburban household and they do all their travels by driving.

Simulation Results

LLM agents can generate a household's daily activity and travel plan

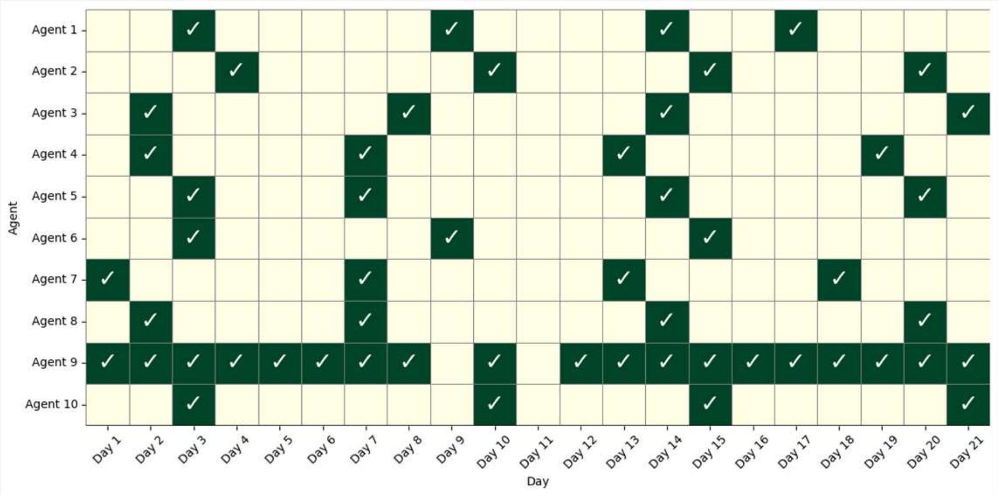
Plan Item	Description
1	**7:00 am: Dad departs for work.**
2	**7:15 am: Son departs for school.**
3	7:45 am: Dad arrives at work.
4	7:45 am: Son arrives at school.
5	8:00 am: Dad starts work, son starts school
6	10:30 am: Mom starts work from home.
7	3:30 pm: Mom finishes work from home.
8	4:00 pm: Son finishes school.
9	**4:15 pm: Son departs from school.**
10	4:45 pm: Son arrives home.
11	5:00 pm: Dad finishes work.
12	**5:15 pm: Dad departs from work.**

Trip	Origin	Destination	Departure Time	Person involved
1	Zone 1	Zone 2	7:00 am	Son
2	Zone 1	Zone 3	7:15 am	Dad
3	Zone 2	Zone 1	4:15 pm	Dad
4	Zone 3	Zone 1	5:15 pm	Son

Simulation Results

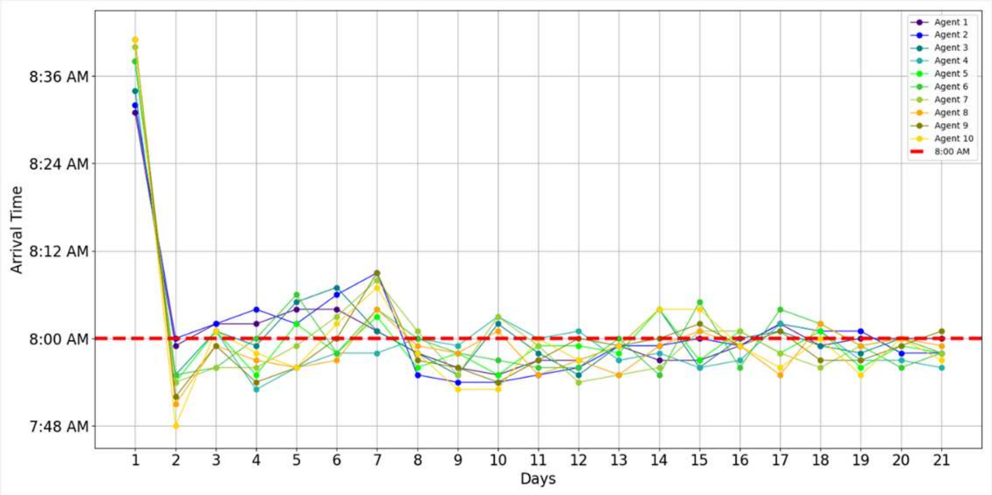
Agents organize activities based on their needs and past actions

When to do grocery shopping?

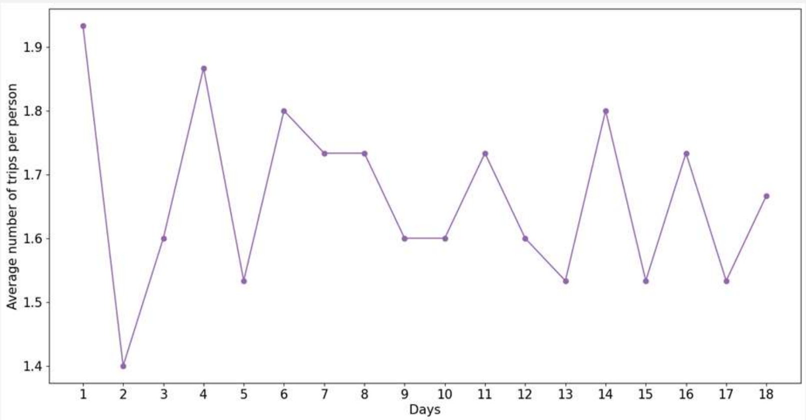


Agents learn from their travel experiences and adapt their travel choices accordingly

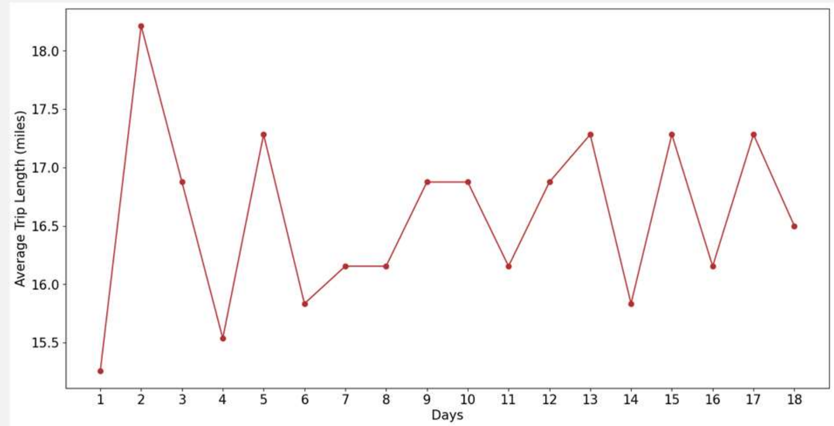
Father's departure time for work that starts at 8AM



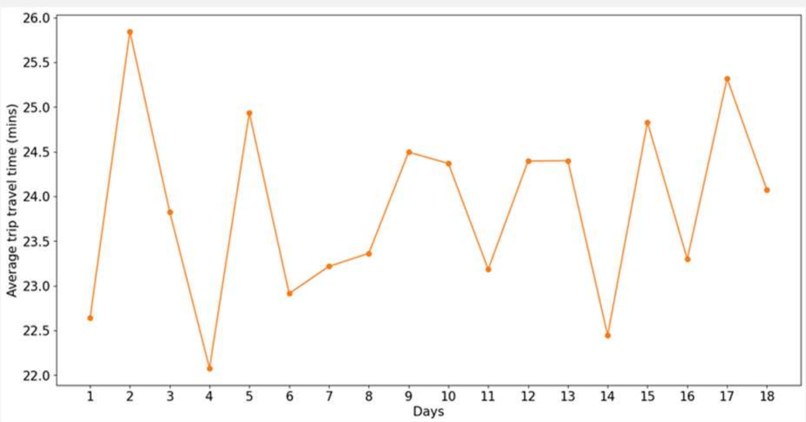
Simulation Results



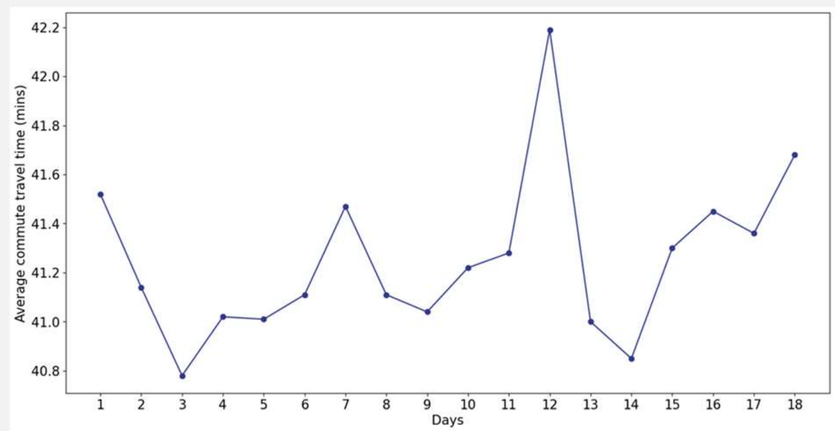
Average number of trips per person



Average trip length



Average trip time



Average morning commute time

Key Challenges

Behavioral Alignment

- ★ LLMs struggle to replicate natural randomness in human behavior
- ★ Risk of systematic biases due to skewed training data
- ★ Lack of integration of attitudinal variables (e.g., preferences, values)
- ★ Value alignment during training may distort behavioral realism

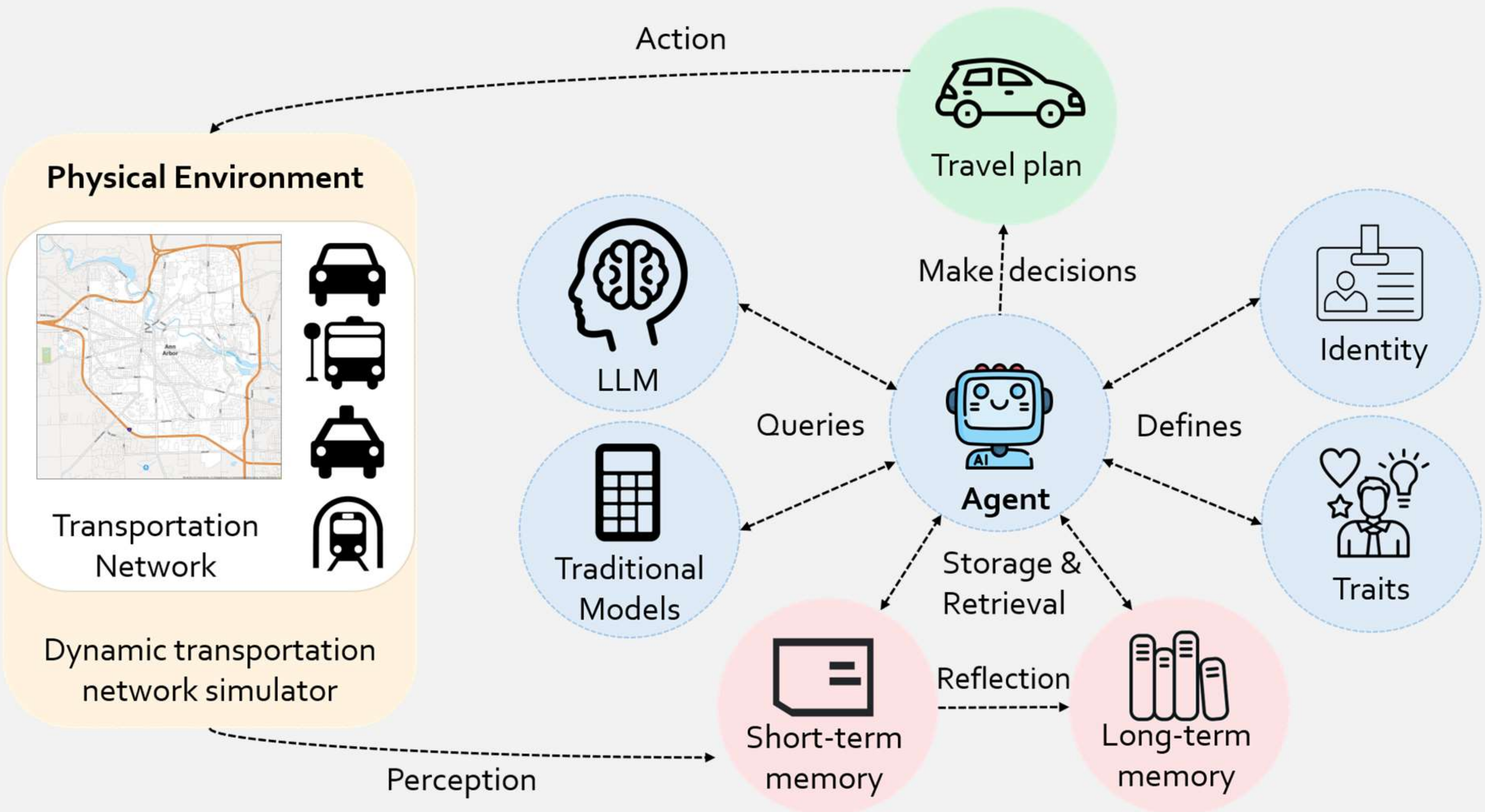
Validation

- ★ Need for rigorous micro-level and macro-level validation
- ★ Individual-level misalignment can amplify into system-level errors

Scalability

- ★ High computational cost for simulating large agent populations
- ★ Latency due to token-by-token LLM inference
- ★ Requires special techniques (e.g., batching, quantization) for large-scale deployment

LLM-augmented Agent-based Simulation



limos lab for innovative mobility systems

Thank You!

yafeng@umich.edu

Potential Advantages

Relaxed Behavioral Assumptions

★ Beyond rigid econometric/rule-based models
★ Captures nuanced, context-driven behaviors
★ Models bounded rationality in decision-making

Built-in Agent Learning and Adaptation

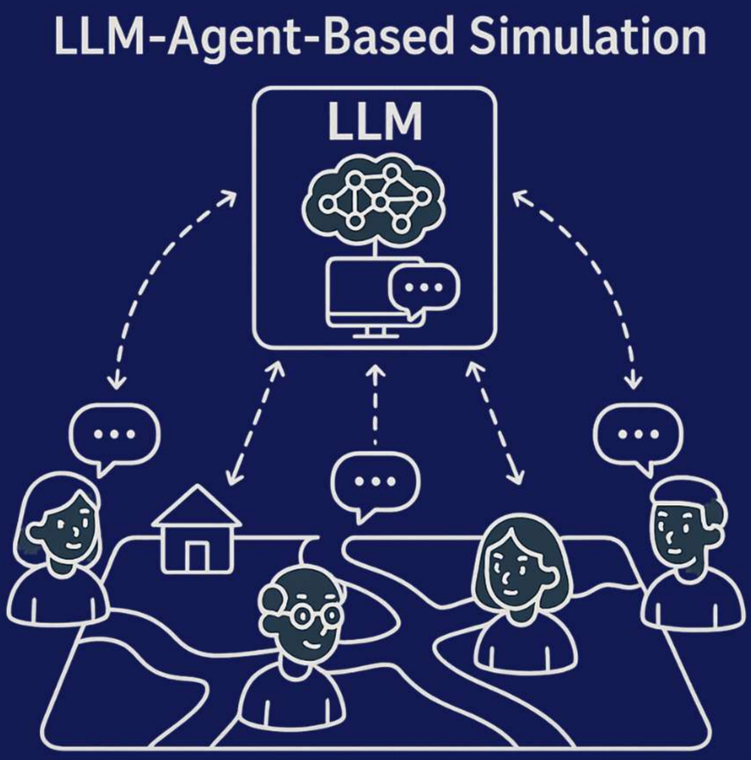
★ Agents adjust based on feedback and memory
★ Mimics human learning and adaptation
★ Supports behavioral convergence in simulation

Improved Use of Diverse Data Sources

★ Leverages LLM pretraining, reducing calibration needs
★ Uses unstructured/multimodal data (e.g., text, images)

Greater Flexibility in Scenario Evaluation

★ Behaviors adaptable via natural language prompts
★ No need for complex reprogramming
★ Enables rapid testing of alternative scenarios



Activity-based Modeling