limos lab for innovative mobility systems

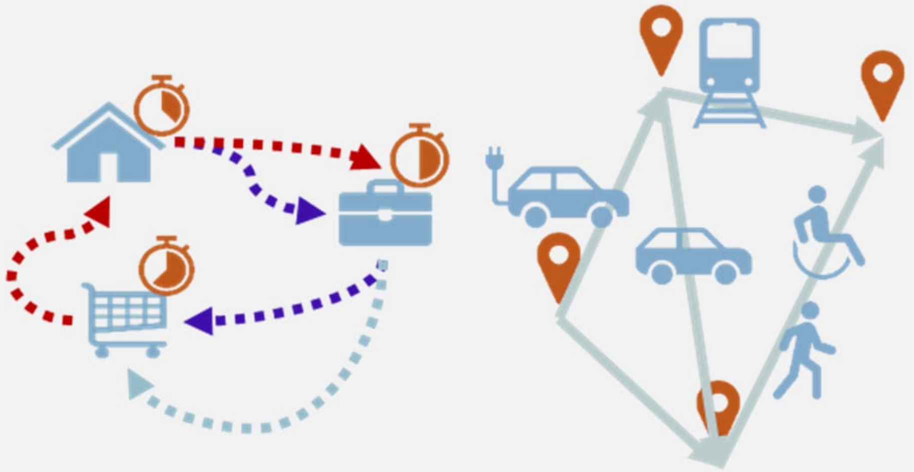
2025 Modeling Mobility Conference
Minneapolis, Minnesota, USA

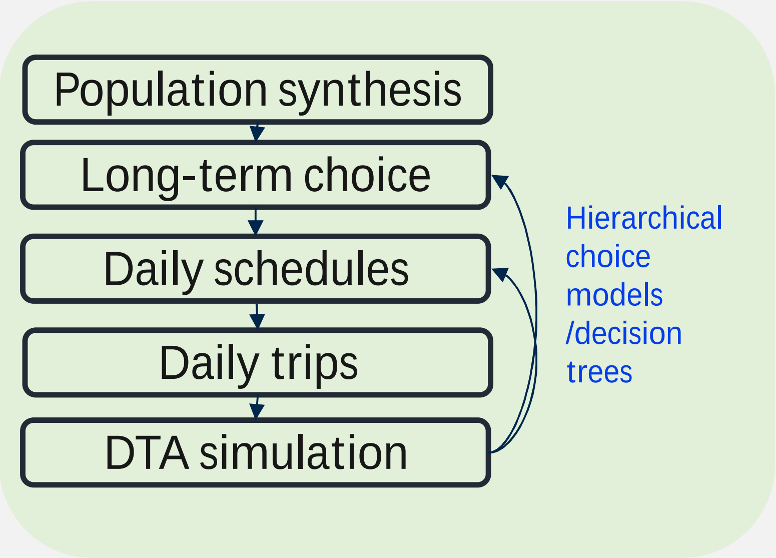
Integrating Activity-Based Transportation Models with Large Language Model Agents

Tianming Liu
Ph.D. candidate
Department of Civil and Environmental Engineering
University of Michigan

Activity-based Modeling

- The-state-of-the-art paradigm of travel demand modeling
- Recent efforts primarily leverage agent-based simulation



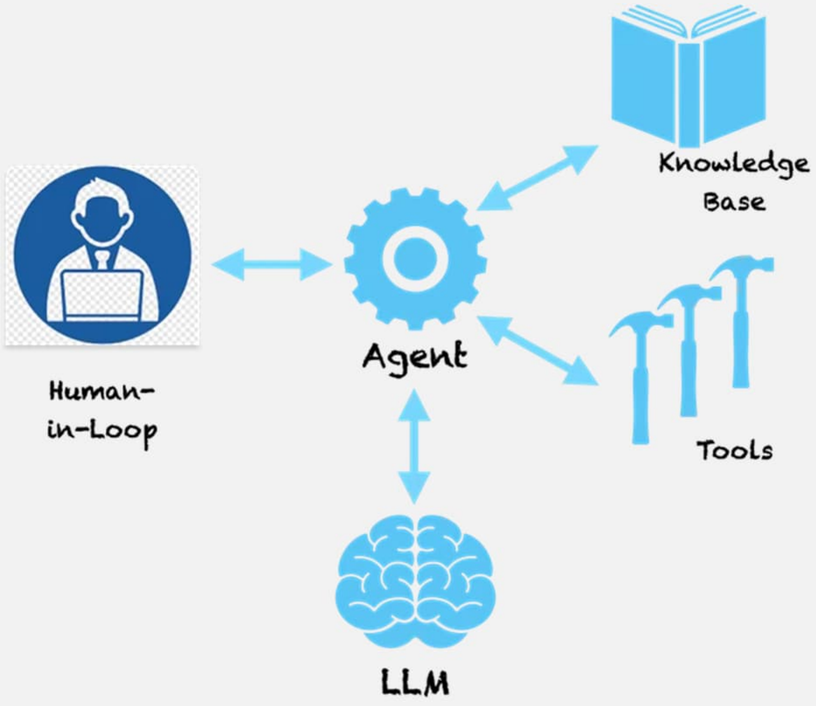


Agents follow hardcoded condition-action rules to schedule daily activities and make travel plans. They can learn, adapt, and improve their interactions with other agents as well as their dynamic environment.

tianmliu@umich.edu

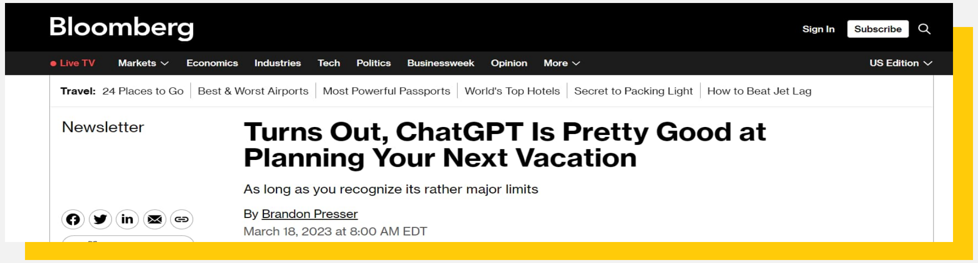
LLM Agents

LLM-based or powered agent is a system that can use an LLM as “brain” to reason through a problem, create a plan to solve the problem, and execute the plan with the help of a set of tools



ChatGPT

Powered by GPT-4, ChatGPT is a text-based, tool-augmented LLM agent that can assist with reasoning, planning, and problem-solving across various domains.



Bloomberg

Sign In | Subscribe

Live TV | Markets | Economics | Industries | Tech | Politics | Businessweek | Opinion | More | US Edition

Travel: 24 Places to Go | Best & Worst Airports | Most Powerful Passports | World's Top Hotels | Secret to Packing Light | How to Beat Jet Lag

Newsletter

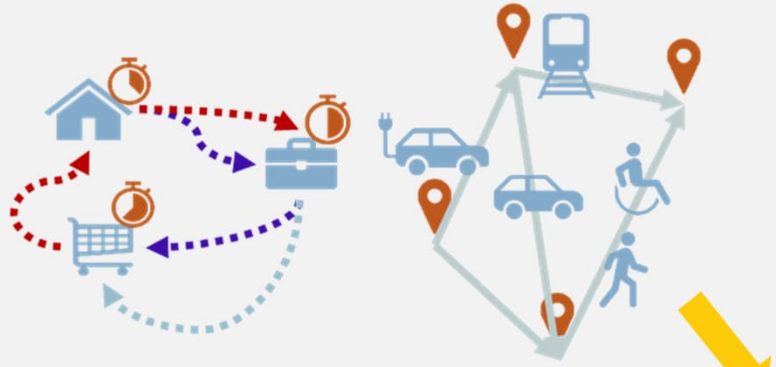
Turns Out, ChatGPT Is Pretty Good at Planning Your Next Vacation

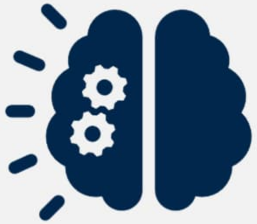
As long as you recognize its rather major limits

By Brandon Presser

March 18, 2023 at 8:00 AM EDT

LLM-Agent-Based Simulation for ABM





LLM agents as proxies
for human travelers

Mesoscopic traffic simulator

Dynamical conditioning based on agents’ changing
experiences

LLM as a Proxy of Human Behavior

- Key hypothesis: LLMs have the potential to act as “silicon samples”
- Main reason is that LLM has three key properties:

Imitation in learning

Human-like Interaction

Instruction-following and role-play

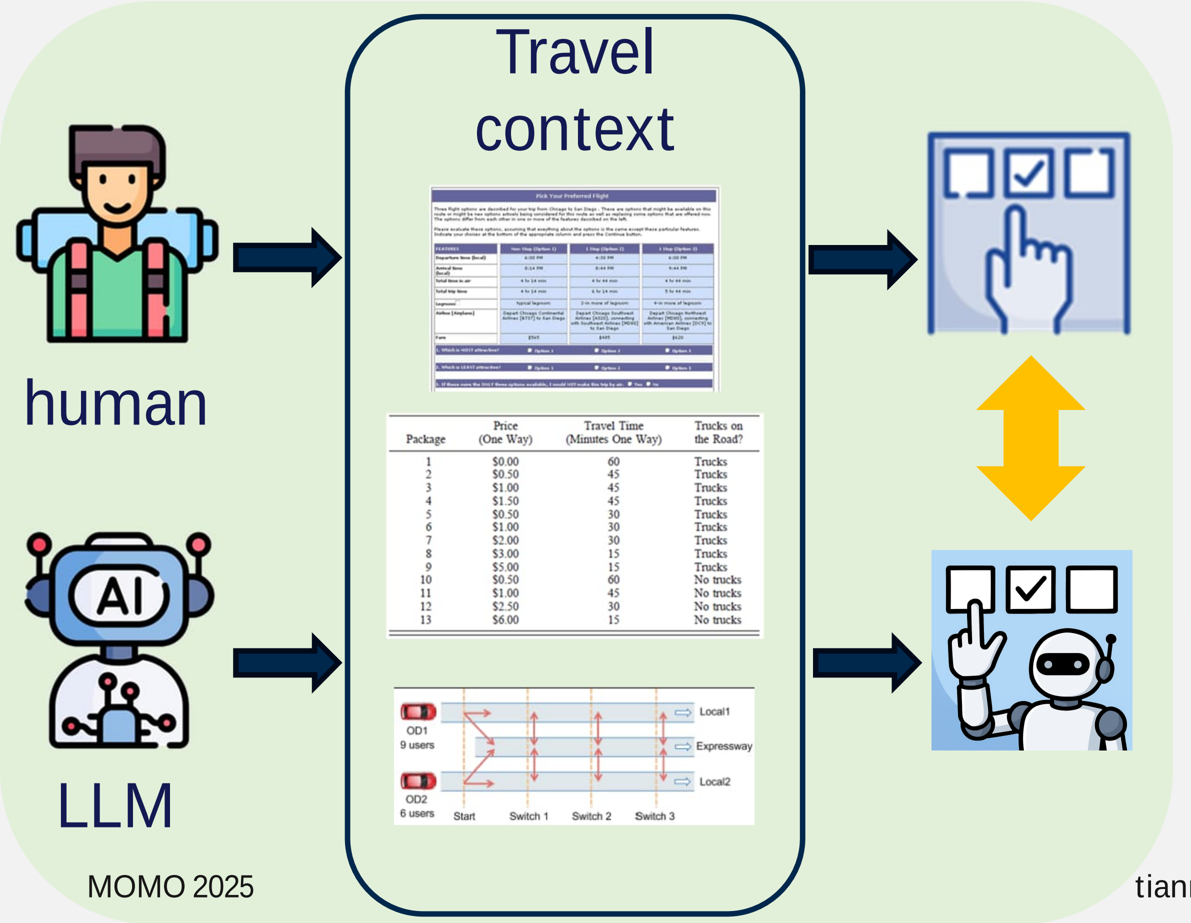
Critical questions: how much do they align with human travel behavior?

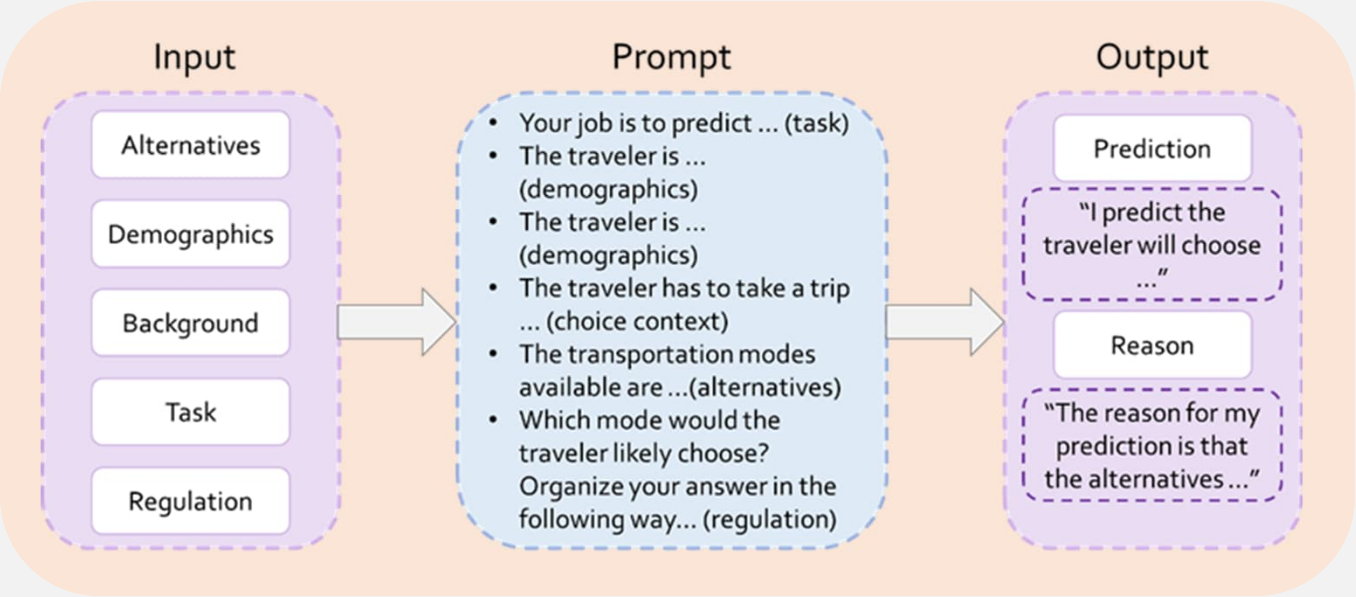
* Grossmann, et al., AI and the transformation of social science research, Science, 2023, 380 (6650).
** Hutson, M., Can AI chatbots replace human subjects in behavioral experiments. Science, 2023, 381(6654).

How do we evaluate alignment?

We conduct evaluation on choice and learning levels

- Basic methodology: treat LLM as an autonomous traveler, prompt them with the same information, compare LLM&human responses





LLM's VOTT

Following Calfee et al. (2001)*

- 13 options (packages) with varied travel time and cost
- Respondents are asked to provide ratings and a ranking of options
- Rankings are used to calibrate an ordered logit model

Package	Price (One Way)	Travel Time (Minutes One Way)	Trucks on the Road?
1	$0.00	60	Trucks
2	$0.50	45	Trucks
3	$1.00	45	Trucks
4	$1.50	45	Trucks
5	$0.50	30	Trucks
6	$1.00	30	Trucks
7	$2.00	30	Trucks
8	$3.00	15	Trucks
9	$5.00	15	Trucks
10	$0.50	60	No trucks
11	$1.00	45	No trucks
12	$2.50	30	No trucks
13	$6.00	15	No trucks

Experiment

- Full factorial design experiment
to control social-demographics and travel situation
- On each run we repeat survey 60 times
on GPT 4o with temperature 1.

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's VOTT

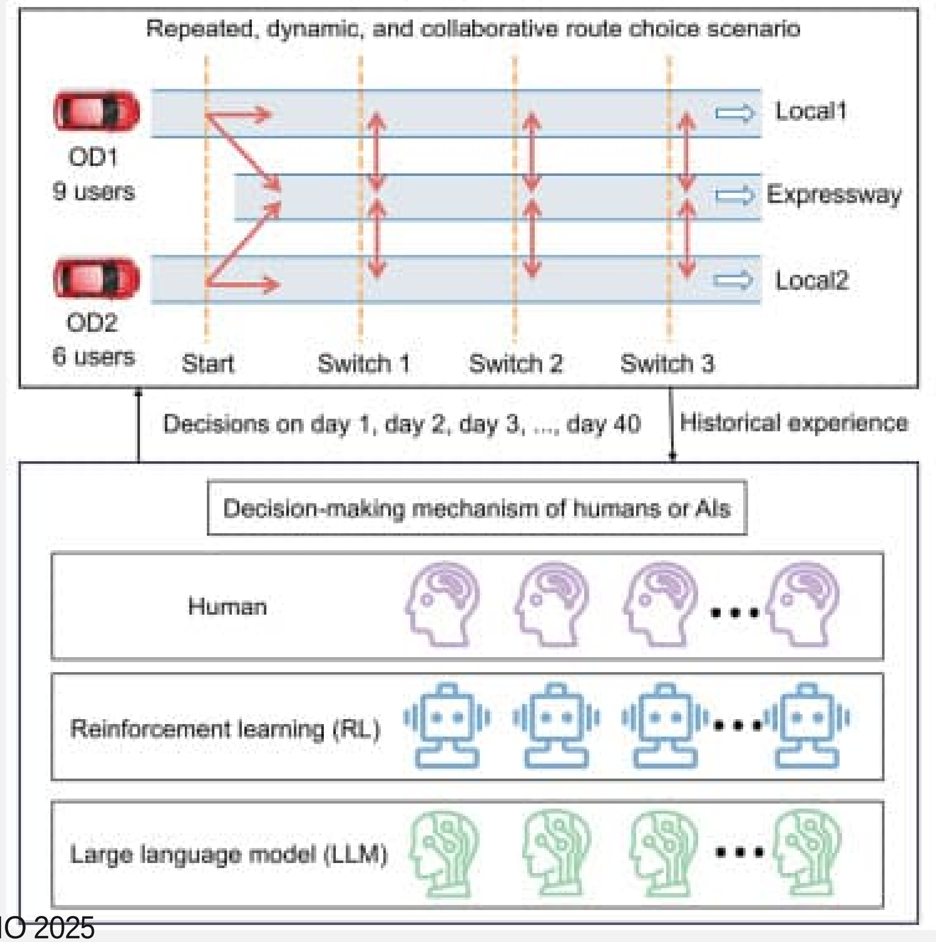
Factor	Level	GPT-4o	Factor	Level	GPT-4o
Purpose	Leisure	$7.12/h	Wage per hour	$15	$6.47/h
	Commute	$8.54/h		$25	$7.80/h
	Business	$8.22/h		$35	$8.38/h
	Personal	$7.88/h		$50	$8.77/h

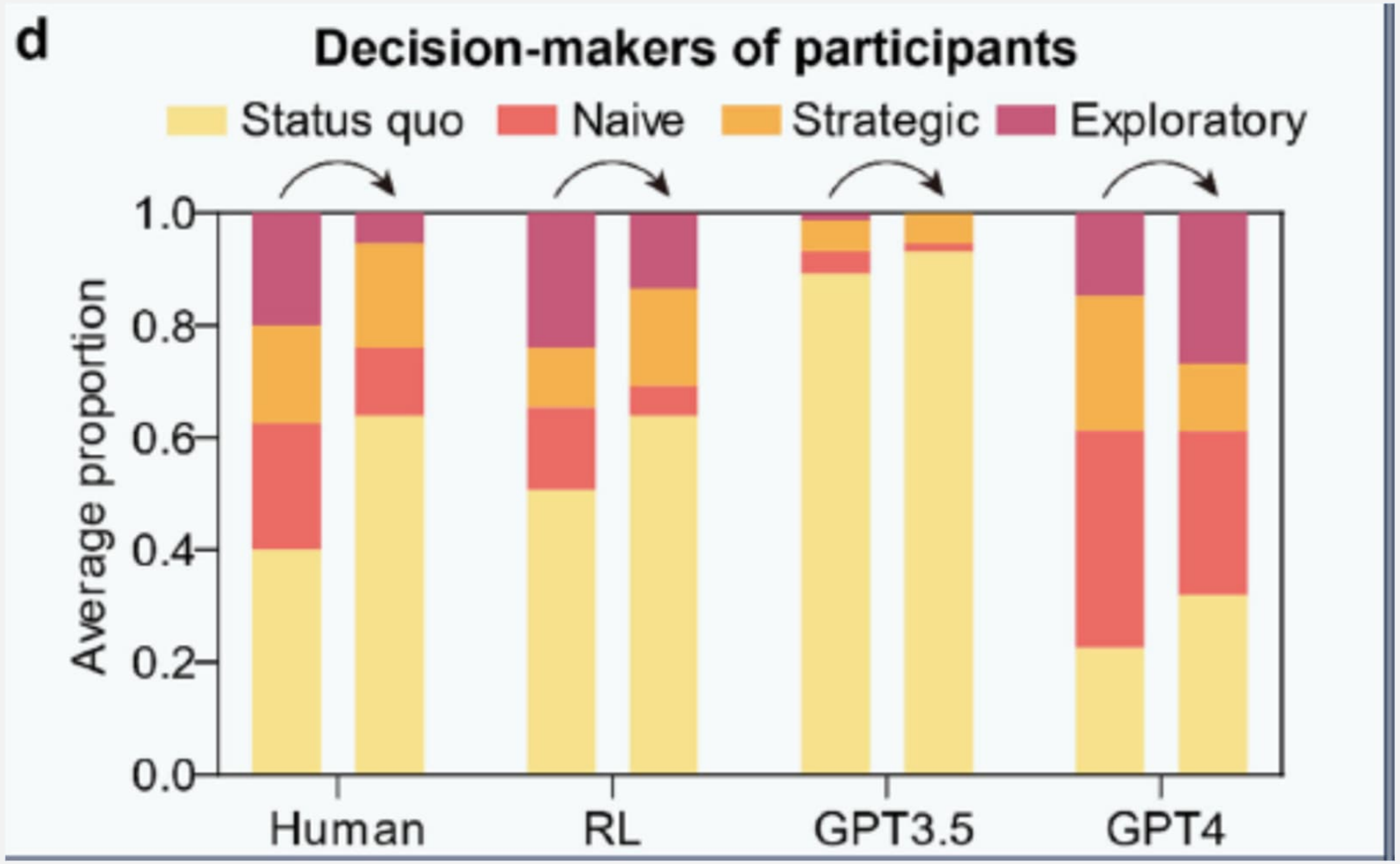
Purpose	Income Elasticity			
	USDOT	Binsuwadan et al. (2023)*	Shires and de Jong (2009)**	GPT-4o
Commute	1	0.37	0.67	0.24
Business	1	0.53	0.47	0.21
Leisure	1	0.53	0.52	0.22
Personal	1	0.53	0.52	0.22

**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.*

LLM’s learning and choice adaptation

- LLMs and humans are put into the same commuting route choice experiment
- LLMs exhibit significantly different adaptation behaviors than humans.

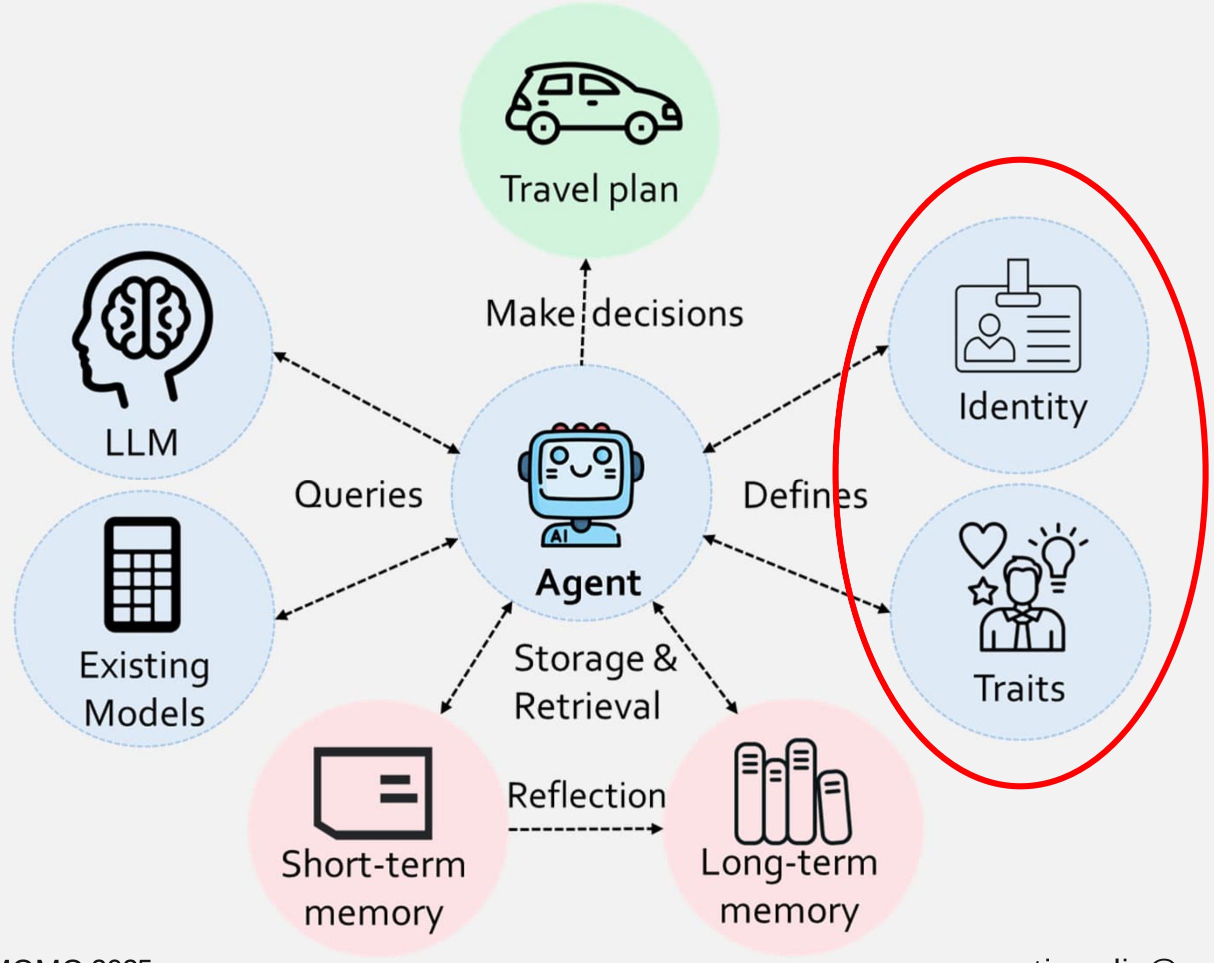


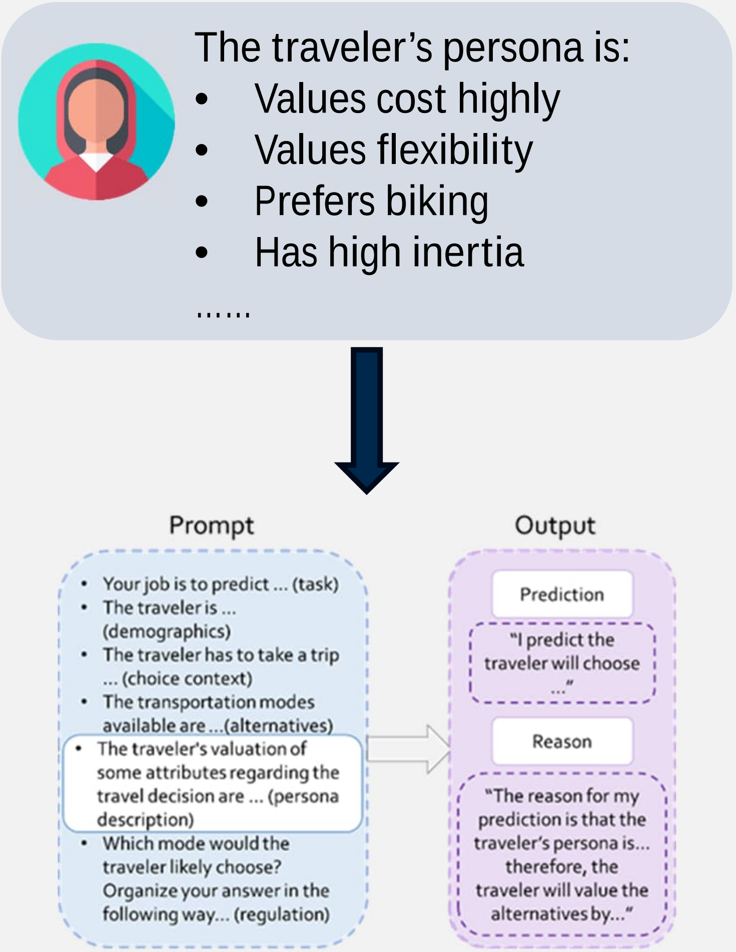


Humanize LLMs

Our key idea: persona

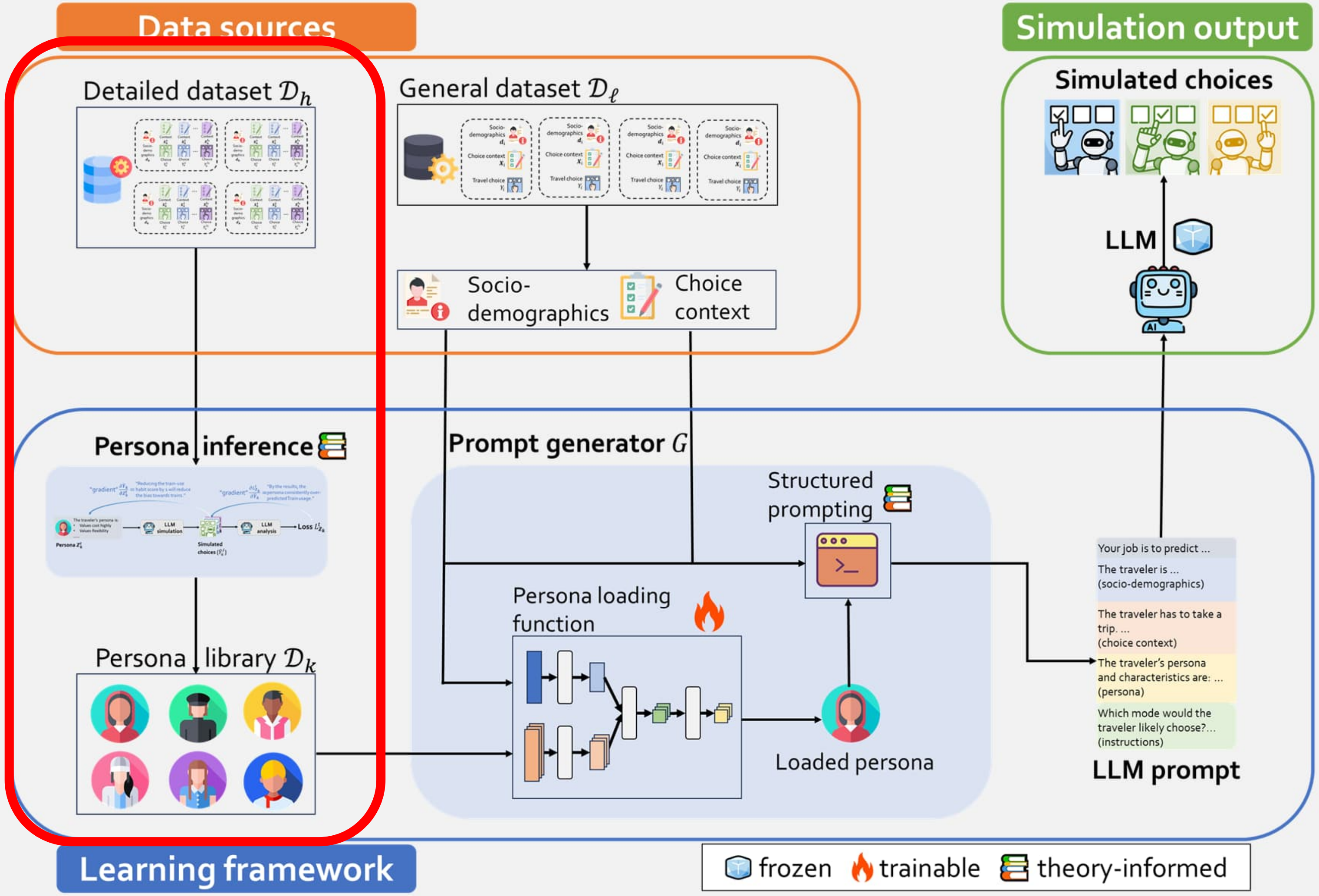
- Textual representation capturing the preferences and traits of the person





Alignment: choice modeling

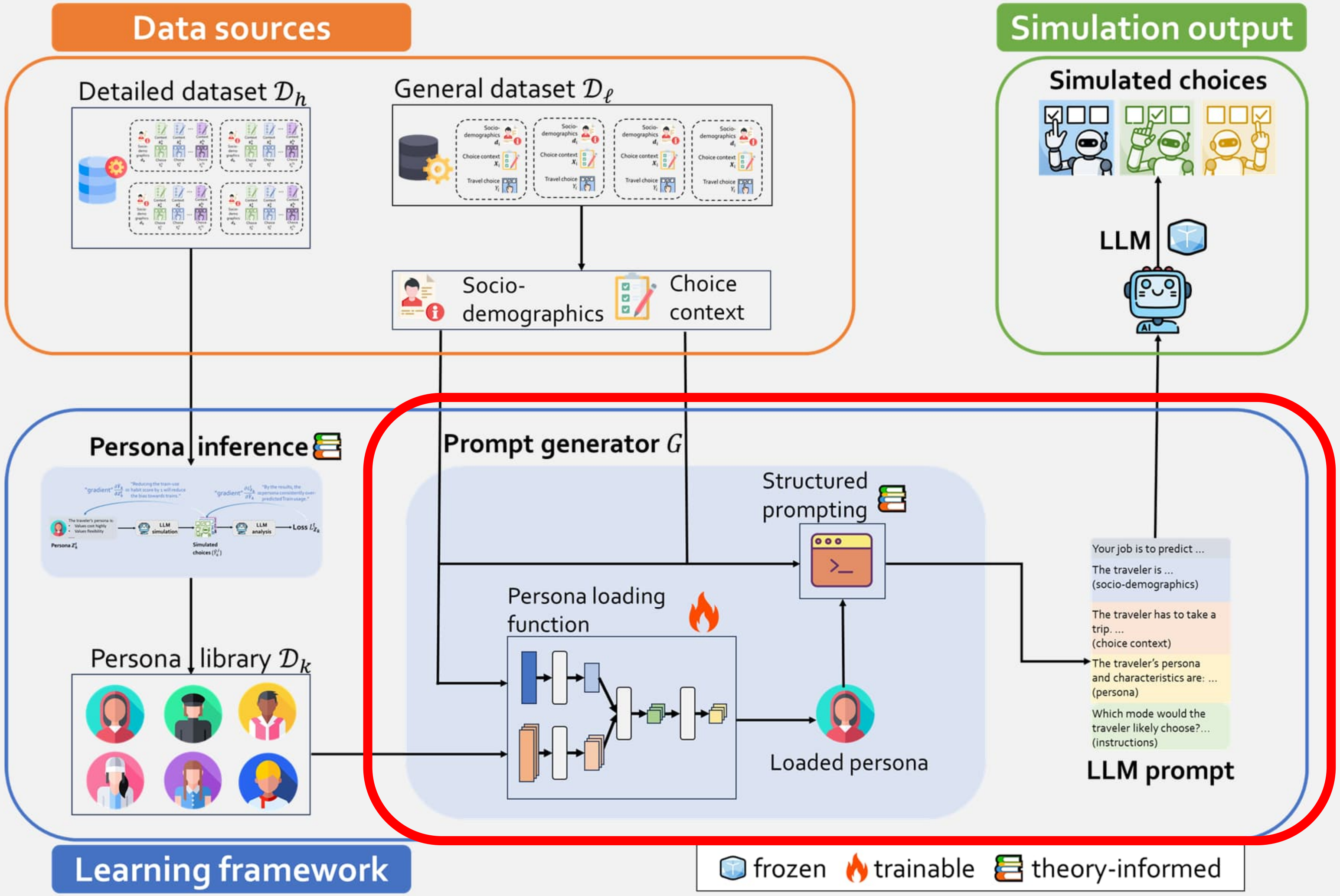
Our main approach involves two key steps:



Step 1:
Define and inversely learn a set of LLM persona from data

Alignment: choice modeling

Our main approach involves two key steps:



Step 1:
Define and inversely learn a set of LLM persona from data

Step 2:
Learn a persona loading function based using latent embeddings and underlying behavior similarities

Choice model: results

Method	Train	Mode Split Swissmetro	Car	JSD (in 0.1 bits)	Marco F1	Weighted F1
Ground truth	6.0%	54.0%	40.0%	0.000	1.000	1.000
MNL	1.5%	79.0%	19.5%	0.548	0.464	0.617
Zero-shot LLM	13.5%	73.5%	13.0%	0.735	0.438	0.542
Few-shot LLM	13.0%	59.5%	27.5%	0.188	0.446	0.579
Liu et al. 2024	5.5%	62.5%	32.0%	0.055	0.493	0.648
Our method	4.5%	60.0%	35.5%	0.029	0.541	0.691

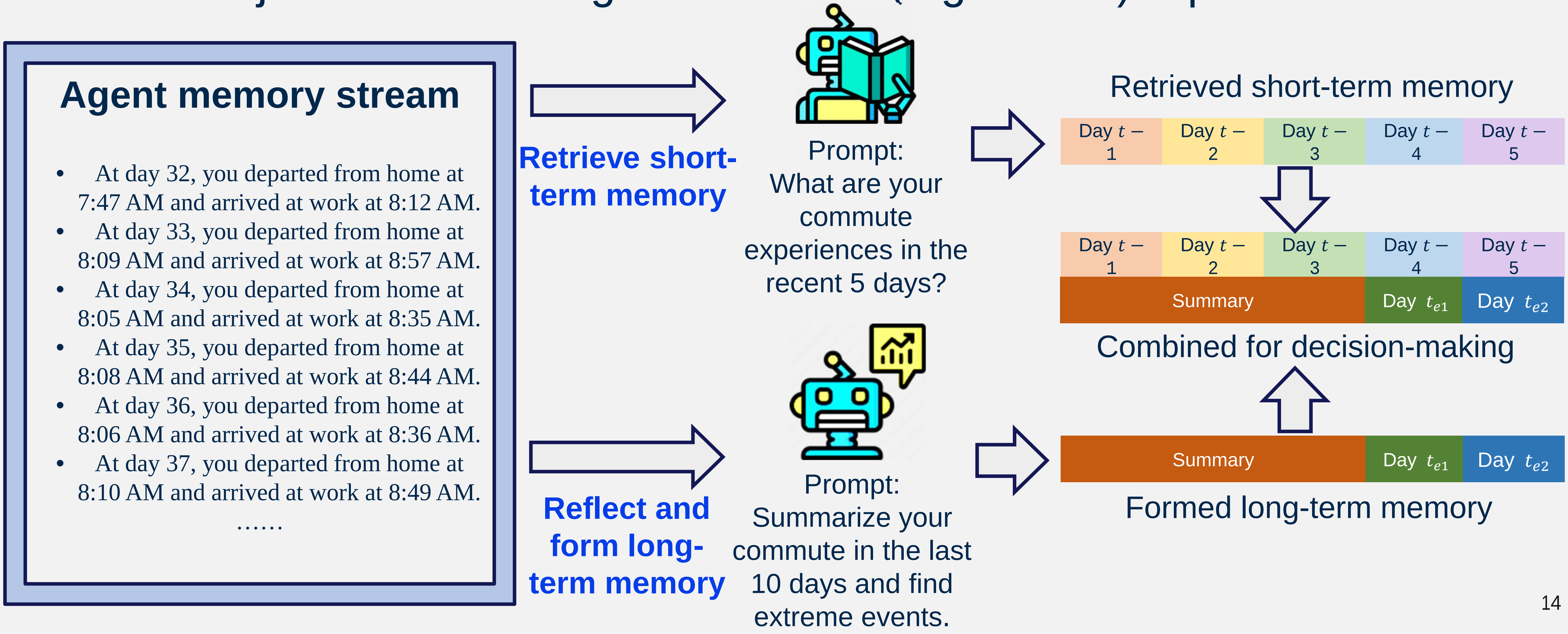
Experiment: Swissmetro mode choice dataset+GPT-4o

Takeaways: our method exceeds existing methods' performances in

- Generating a more realistic aggregate alternative share prediction
- Producing more accurate individual behavior prediction

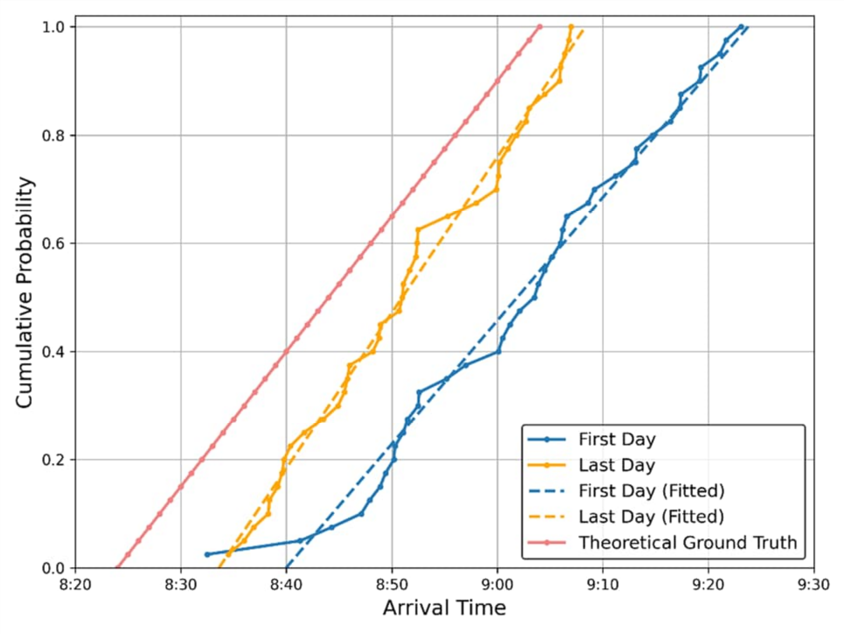
Alignment: day-to-day adaptation

- LLM agents use a combination of short-term memories and long-term memories when making their decisions.
- Other adjustments: adding human traits (e.g. inertia) in persona

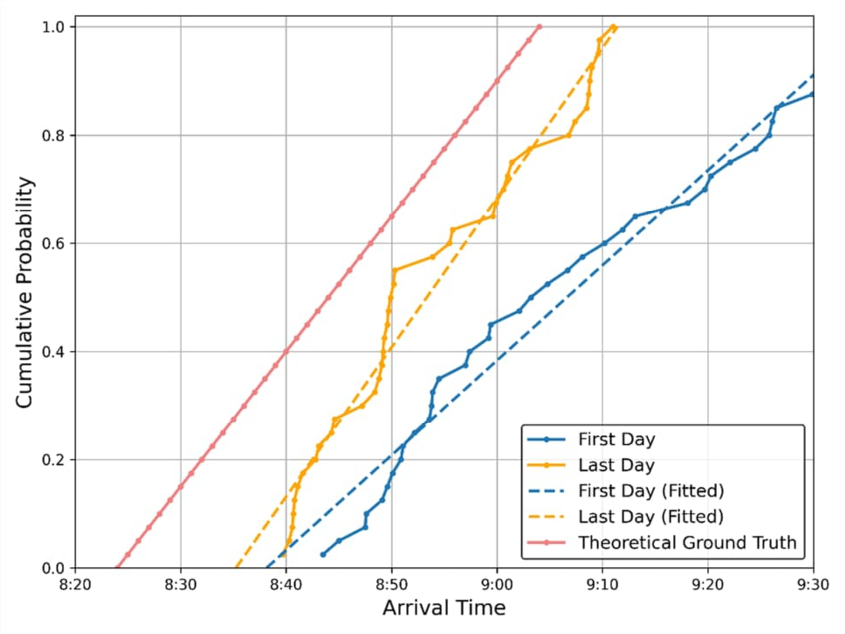


Day-to-day adaptation: results

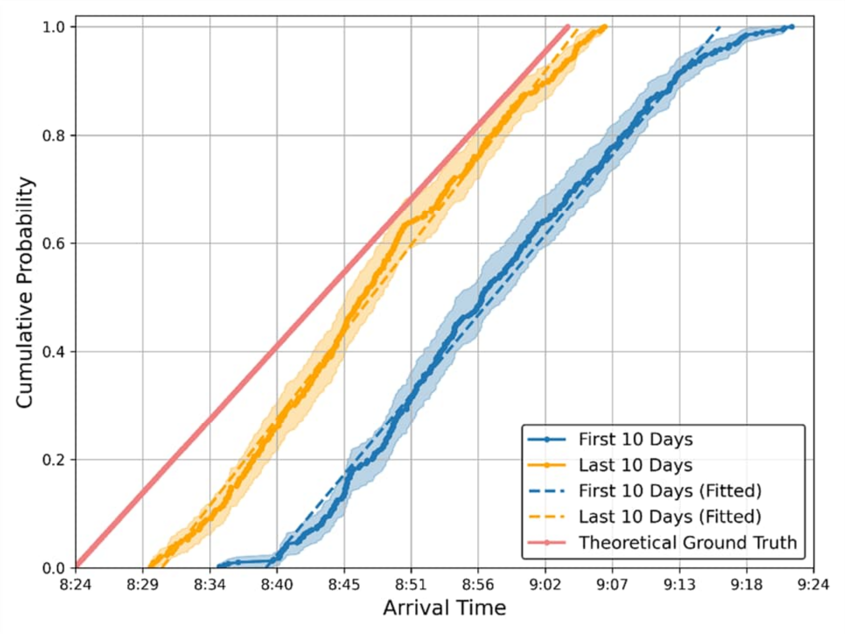
- We put our approach into a day-to-day departure choice context and compare it with other agents.
- Our proposal facilitates better learning behavior and converge most closely to ground truth (red line).



Vanilla agents



CoT-only agents



Proposed agents

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

Takeaways

Our idea: using LLM to augment agent-based modeling

Takeaways:

- LLMs alone can imitate some human trends in travel demand, but has significant limitations
- Persona and memory system can significantly enhance LLM agent's ability to simulate human travel
- LLM agents have potential but also requires further development

Check our papers online

vision

Mode choice eval

VOTT eval

choice alignment

learning eval

learning alignment

tianmliu@umich.edu

Thank You!

tianmliu@umich.edu

Future opportunities

Behavioral Alignment

- ★ More expansive and efficient distributional alignment method
- ★ Leverage multimodal and multi-source data
- ★ Identify the optimal mixture for hybrid modeling

Validation

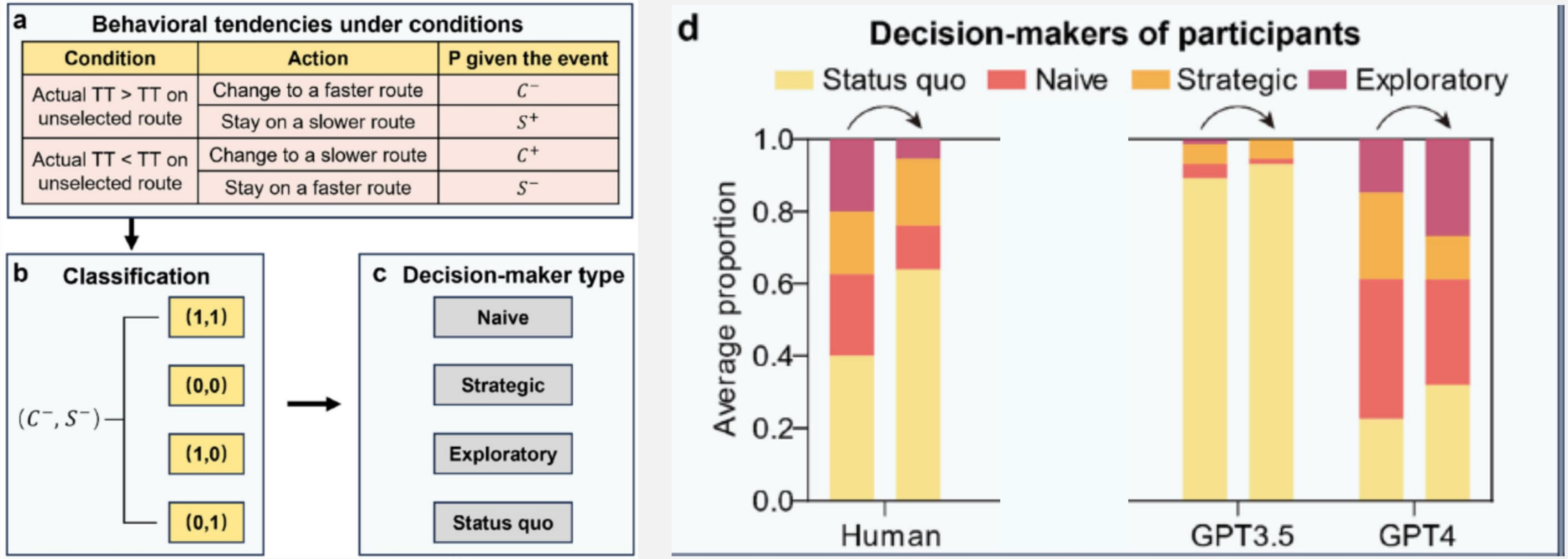
- ★ Expansive testing of LLM behavior
- ★ More evaluation and improvement of the value/need level

Scalability

- ★ Computational optimization (e.g. parallel computing)
- ★ Application of small language models

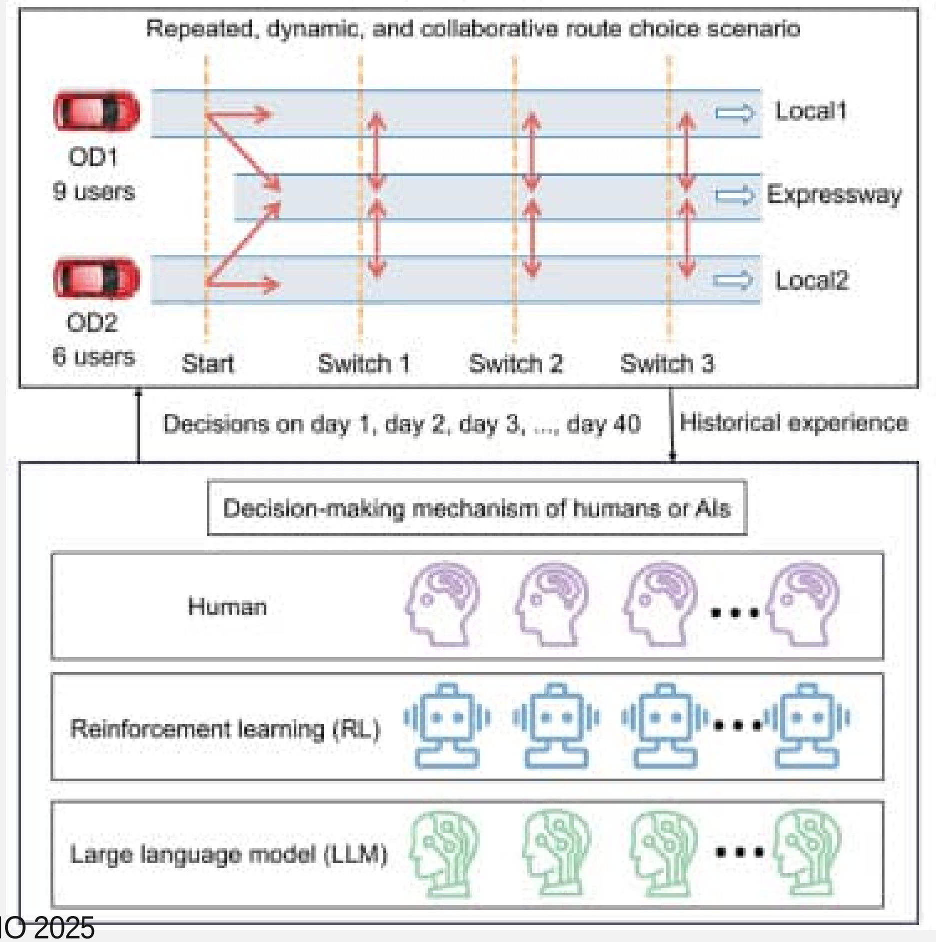
LLM's learning and choice adaptation

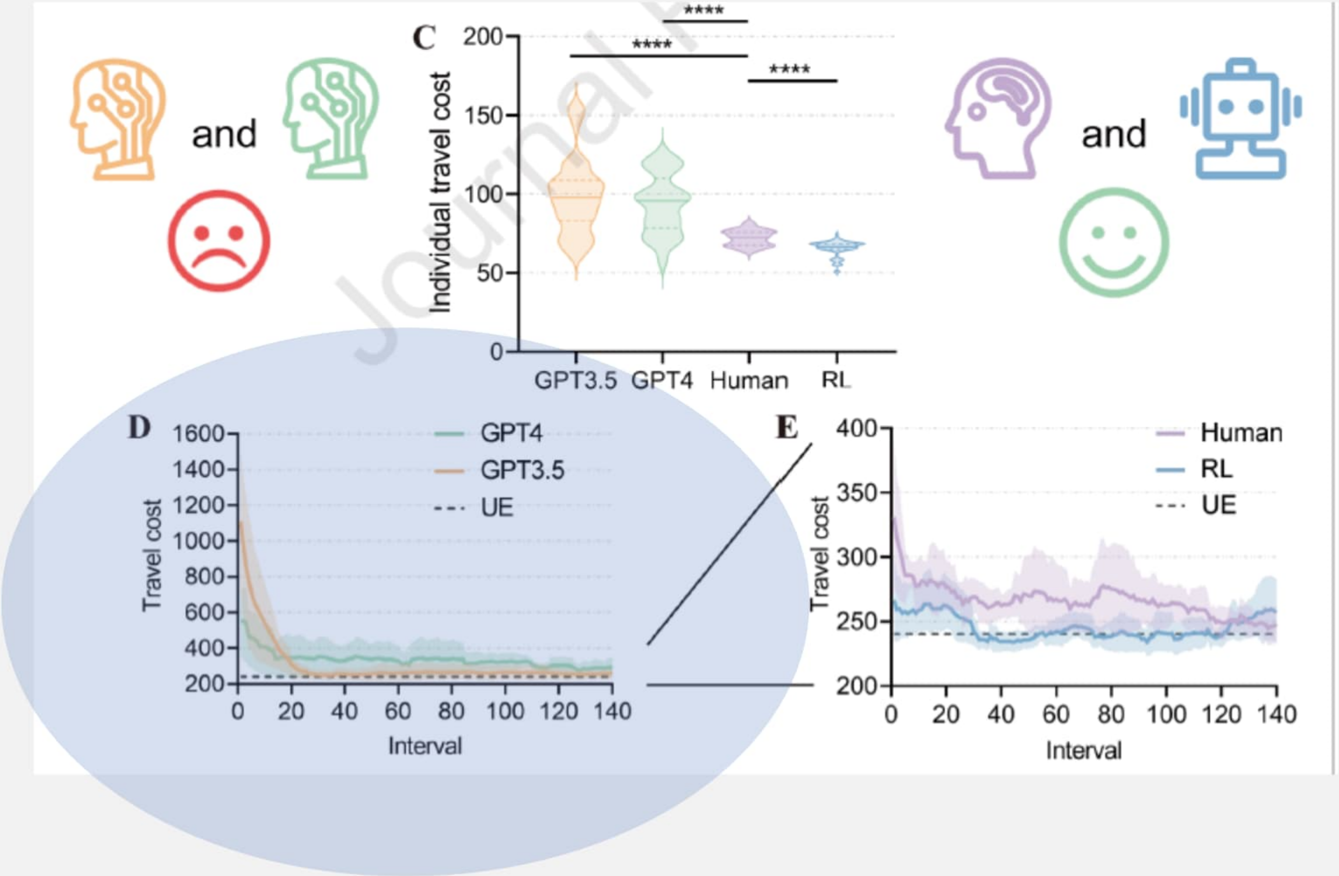
LLMs exhibit significantly different adaptation behaviors than humans.



LLM’s learning and choice adaptation

- LLMs and humans are put into the same commuting route choice experiment
- The resulting system dynamics is also vastly different

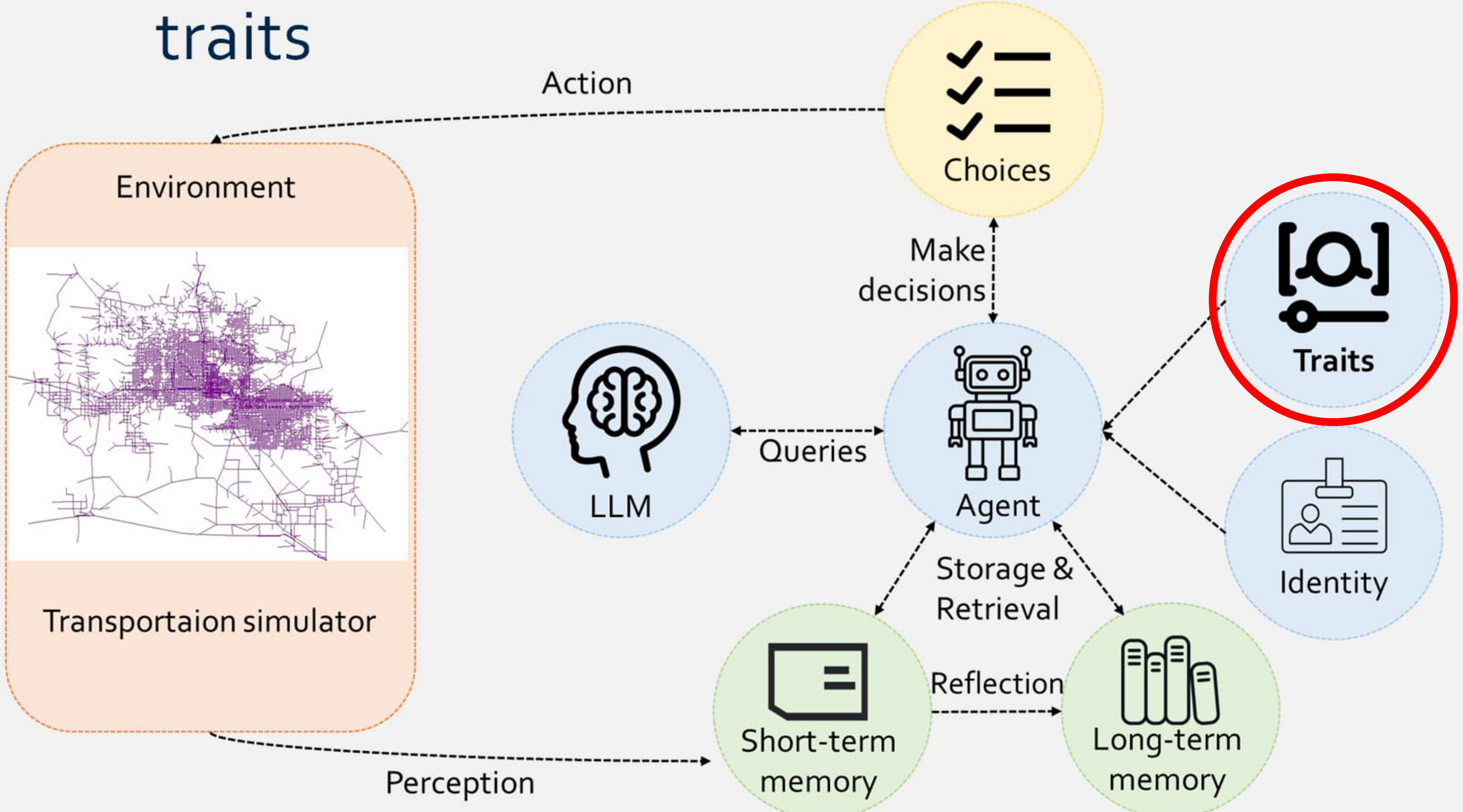




Alignment: day-to-day adaptation

We design the agents for them to have human-like decision making traits and memories:

- Key: enhancing persona with clear values and human decision-making traits



Key enhancements

Enhancements

- Valuation and trade-offs
- Behavioral inertia
- Theory of mind
- Chain of thought
- Self correction