# Mobility AI Agents and Networks

Haoxuan Ma, Yifan Liu, Qinhua Jiang, Brian Yueshuai He, Xishun Liao*, and Jiaqi Ma

***Abstract*—Intelligent vehicles and smart mobility systems are at the forefront of transportation evolution, yet effective management of these new mobility technologies and services are nontrivial. This letter proposes an Intelligent Mobility System Digital Twin (MSDT) framework as a solution. Our framework uniquely maps human beings and vehicles to AI agents and the mobility systems to AI networks, creating realistic digital simulacra of the physical mobility system. By integrating AI agents and networks, this framework offers unprecedented capabilities in prediction and automated simulation of the entire mobility systems, thereby improving planning, operations, and decision-making in smart cities.**

## I. INTRODUCTION

As TECHNOLOGY advances and economies grow, urbanization accelerates, leading to increased population densities [1]. This trend presents complex mobility management challenges for city decision-makers and governments. The concept of the digital twin emerges as a promising solution to these challenges [2]–[4]. Creating a digital twin of the physical mobility system facilitates more efficient planning, operations, and decision-making. It seamlessly integrates existing mobility models, enhancing the simulation and optimization of mobility networks. This integration is particularly valuable for intelligent vehicles and smart mobility systems [5], as it provides a comprehensive virtual environment to model and predict the behavior of both traditional vehicles and connected automated vehicles (CAVs). Digital twins enable the assessment of how these emerging technologies interact with existing infrastructure and influence traffic patterns, safety, and efficiency. This real-time analysis and predictive modeling capability allows for more responsive and adaptive management strategies, paving the way for the smooth integration of intelligent mobility solutions.

The key to establishing a mobility system digital twin (MSDT) lies in accurately modeling the intelligent behavior of all system components, such as humans, vehicles, and roadway networks. We propose Mobility AI agents and network modeling. In this MSDT, AI agents and AI networks form the core components. AI agents serve as simulacra of human beings, learning and replicating complex behavior patterns. These agents can perceive and understand their environment, making autonomous predictions and decision-making. AI networks, in turn, represent an intelligent and adaptive model representation of the transportation infrastructure system. These networks go beyond static mapping, incorporating real-time data and self-learning capabilities to reflect and predict the state of the mobility system dynamically. Together, AI agents and AI networks create a dynamic and realistic representation of mobility, enabling comprehensive modeling of complex mobility systems.

This letter presents a comprehensive conceptual framework for developing an MSDT that integrates diverse data sources, as illustrated in Fig. 1. Through the fusion and analytics of multifaceted data streams, we construct highly accurate digital representations that mirror the complexities of real-world mobility systems. This approach enables the creation of sophisticated AI agents and networks, forming a robust foundation for advanced simulation and analysis. The framework supports various applications, ranging from cutting-edge simulation tools to in-depth mobility system inference. These capabilities significantly enhance planning processes, inform policy formulation, and facilitate strategic decision-making. Moreover, the digital twin operates in a continuous feedback loop with the physical mobility system, allowing for real-time adjustments and long-term improvements of the digital twin for mobility management.

## II. SYSTEM FRAMEWORK OF MSDT

The MSDT is a comprehensive framework designed to enhance mobility analytics by integrating digital and physical systems. As illustrated in Fig. 1, this system relies on AI agents and AI networks to create a detailed digital replica of the physical mobility system, with the data warehouse serving as a critical foundation.

### A. Data Warehouse

Our framework begins by constructing a comprehensive data warehouse that feeds into the digital twin. This data warehouse encompasses three key categories: Foundation, Processed, and Synthetic Data. Foundation Data includes information collected from human subjects (e.g., surveys [6] and tracked GPS points [7]) and infrastructure (e.g., Open Street Map [8], traffic count data [9], crash data [10], and aerial videos and images). Processed Data, derived through data mining, can include aggregated traffic states of specific road segments in given time intervals and travel trajectories of individuals in areas of interest, including activity types, start/end times, and locations. Synthetic Data can be generated through simulation or pre-trained models, such as synthetic population information, synthetic travel trajectories, and multimodal networks. These datasets are instrumental in developing and validating AI agents and networks, and also perform use case analysis, such as post-processing to understand travel patterns or network congestion patterns.

Haoxuan Ma, Yifan Liu, Qinhua Jiang, Xishun Liao, and Jiaqi Ma are with the UCLA Mobility Lab under the Department of Civil and Environmental Engineering, University of California, Los Angeles, Los Angeles, USA.

Brian Yueshuai He is with the Department of Civil and Environmental Engineering, University of Louisville, Louisville, Kentucky, USA.

*Corresponding author: xishunliao@ucla.edu

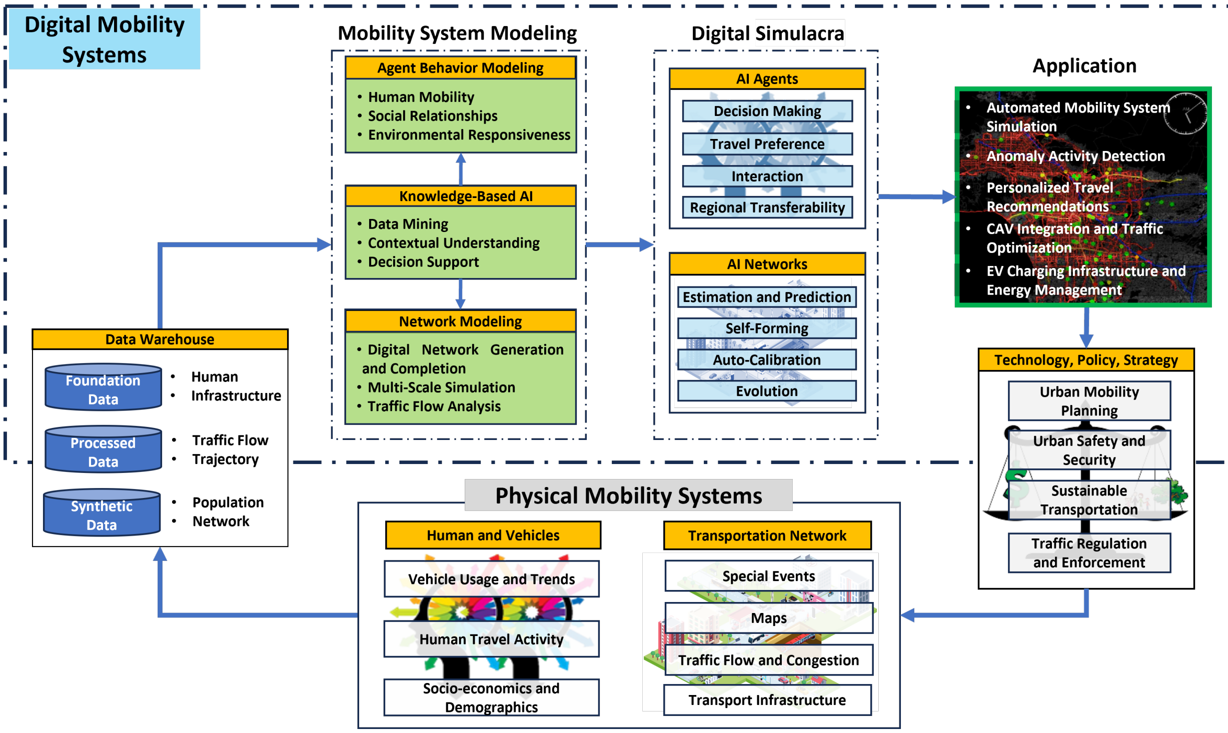


Fig. 1: Integrated intelligent mobility system: a digital twin framework with AI agent and network modeling

### *B. Mobility System Modeling*

At the core of our framework is the Mobility System Modeling process. It serves as a crucial bridge between the data warehouse and Digital Simulacra, automating the AI agents and networks to create an autonomous and intelligent digital twin of the mobility system. Agent Behavior Modeling is a critical component of our Mobility System Modeling, focusing on three key models: human mobility, social relationships, and environmental responsiveness. **Human Mobility** modeling examines how individuals with different socio-economic and demographic backgrounds exhibit various daily activities and travel patterns (considering that travel demand is derived demand from needed activities). This modeling helps predict and understand diverse mobility needs and preferences across the population [11]. **Social Relationships** modeling captures the interpersonal dynamics that influence travel decisions. It recognizes that an agent's activities are often affected by others, for example, friends coordinating get-togethers or family members arranging joint activities. This aspect is crucial for simulating the integration between AI agents. **Environmental Responsiveness** modeling examines how agents make decisions under diverse environmental conditions. These conditions range from traffic scenarios and weather conditions to special events. Ideally, the model should also be capable of reflecting any changes in the network, including the introduction of new mobility technologies and services such as EVs and AVs [12]–[15]. This modeling is essential for equipping AI agents with the ability to make realistic decisions. It forms the foundation of their decision-making capabilities and allows interaction with dynamic and constantly evolving AI networks.

Network modeling is another crucial component in the mobility system modeling process. It focuses on three main aspects: digital network generation and completion, multi-scale simulation, and traffic flow analysis. **Digital Network Generation and Completion** involves creating a digital representation of the physical mobility network using real-world data. However, recognizing that real-world data is often incomplete, an automatic completion functionality is in demand [16]. This AI-driven completion process utilizes data from nearby traffic networks, various infrastructure sources (e.g., loop detectors and image/video data), and human activity patterns to fill in the gaps. Next, **Multi-Scale Simulation** models can be built on the digital network, ranging from individual agent-based vehicle simulations to large-scale transportation system simulations. This multi-scale approach allows for a comprehensive understanding of the mobility system, capturing both micro-level interactions and macro-level patterns [17]–[20]. The third aspect, **Traffic Flow Analysis**, involves modeling traffic flow and congestion patterns, enabling the AI network to estimate and predict traffic conditions [21], [22].

Knowledge-based AI, such as large language models (LLMs) and vision-and-language models (VLMs), plays a role in enhancing mobility system modeling. They interpret and contextualize data from multiple sources, providing a comprehensive view of the mobility system. This equips AI agents and networks with the ability to extract crucial information, enabling them to understand environmental conditions and real-time mobility situations effectively.

### *C. Digital Simulacra: AI Agents and Networks*

*1) AI Agents:* AI agents in our digital twin represent a significant advancement in simulating human mobility behavior. Each agent is assigned a unique identity with specific socio-economic and demographic information, along with their residential, work, or school locations. Based on their identity, AI agents can make decisions according to complex situational factors such as traffic and weather conditions, as well as time constraints.

With these capabilities, AI agents can serve as the next generation of travel demand models. To generate travel demand, AI agents initially learn the activity patterns of each type of identity, such as work, school, or leisure. Then, they form travel preferences, including preferred routes and transportation modes. These choices adapt based on their past experiences and new situations.

Moreover, AI agents interact not only with other AI agents but also with the AI networks. For example, an agent might adjust its route based on real-time traffic conditions provided by the AI network, or the network might update its congestion predictions based on the collective behavior of multiple agents. This interaction simulates real-world social influences in transportation decisions, such as how individuals influence each other's travel decisions or how large events can cause shifts in traffic patterns.

The transferability of a pre-trained AI agent model allows for seamless behavior adaptation across different regions and cultures, making it versatile for various mobility systems. This transferability is achieved through fine-tuning techniques. We start with a base model trained on a large, diverse dataset, then adapt it to specific regions or cultures by training on smaller, localized datasets. This approach allows the model to retain general mobility patterns while learning region-specific behaviors, enabling efficient deployment across different urban environments with minimal additional training.

*2) AI Networks:* AI networks form the intelligent backbone of our digital mobility infrastructure. Their state estimation capabilities provide real-time insights into mobility system conditions, including areas with limited data due to signal loss, device malfunction, or low sensor coverage. AI networks can predict traffic considering multiple factors, such as weather and special events, to forecast future traffic states accurately. This prediction goes beyond traffic congestion, aiming to understand underlying causes and potential chain effects. For example, it can predict how a traffic incident on a major highway might lead to congestion on nearby arterial roads or how severe weather might affect public transit ridership and subsequently impact road traffic.

The generative capabilities (estimation and prediction functions) bestow the AI network with self-forming capability, allowing it to automatically generate pedestrian routes, bicycle paths, and other network elements based on trajectory data and observed movement patterns. The networks also feature auto-calibration, continuously fine-tuning parameters based on observed discrepancies without human intervention.

Additionally, their evolution capability enables adaptation to long-term changes in urban structure, policy, and human behavior, effectively "growing" with the region they represent. Hence, the AI network can adjust its model over time to accommodate significant changes such as new infrastructure developments, shifts in population density, or evolving transportation policies.

The synergy between AI agents and AI networks forms our Digital Simulacra. Once established, AI agents are loaded onto the AI network, creating a comprehensive and interactive digital ecosystem. This integration allows for sophisticated simulations where individual agent behaviors influence and are influenced by network conditions. For instance, AI agents' route choices and travel times are affected by the network's traffic predictions, while the collective behavior of agents shapes the network's state and evolution. This bidirectional interaction enables the digital twin to capture complex phenomena such as emergent traffic patterns, the impact of policy changes, and the ripple effects of local disturbances across the entire mobility system. Furthermore, this integration facilitates scenario testing, where changes to either agent behaviors or network conditions can be simulated to predict system-wide outcomes, providing valuable insights for urban planning and mobility management strategies.

### D. Applications

Digital Simulacra achieves digital autonomy and opens the door to numerous downstream applications.

- AI Agent Anomaly Detection: AI agents learn typical human movement patterns across various times, locations, and social contexts. These models help identify deviations from the norm, enabling the detection of anomalous activities within the extensive global human trajectory data [23], [24].
- Automated Mobility System Simulation: Integrating AI agents and AI networks enables comprehensive, automated simulations of entire mobility systems. AI networks generate and complete digital mobility networks. They then predict traffic patterns, performing dynamic traffic assignments based on real-time conditions. Simultaneously, AI agents function as next-generation, AI-based travel demand models [16], [17].
- New Mobility Intelligence: AI agents learn the travel preferences of individuals from various socio-demographic backgrounds, enabling them to offer personalized travel recommendations [25]. Additionally, AI networks can simulate the impact of CAVs on traffic flow, safety, and overall system efficiency. By modeling human interactions with CAVs, AI agents help optimize integration strategies and predict adoption rates [26].

## III. Progress and Results

Building on the comprehensive overview of the MSDT, this section provides a detailed examination of the AI agent and network modeling processes, as well as the outcomes and progress of ongoing research by the UCLA Mobility Lab. These processes form the core of our framework, enabling the creation of digital twins. The work started with LA as a testbed and is being extended to other places in the US and other parts of the world.

### A. Data Warehouse Construction

Developing a high-quality data warehouse adapts the data-centric AI and Machine Learning Operations (MLOps) concept [27], and it includes organizing, integrating, cleaning, and enhancing raw data, and annotating data with contextual information to make data more useful for analysis, discovery, and decision-making.

To build up the data warehouse, we collected and curated datasets from various sources:

- *Population Data*: American Community Survey [28], National Household Travel Survey [6], and Synthetic Population data from Southern California Association of Governments (SCAG) [29] for building AI agent profile.
- *Human Travel Behavior Data*: Household Travel Surveys [6], [30] and synthetic trajectory from LA-Sim [17] for AI agent travel activity modeling.
- *Location-Based Data*: GPS data provided by Veraset [7], and Point-of-interest (POI) data provided by Open Street Map [8] for learning the movement pattern of AI agent in the network.
- *Transportation Network*: Open Street Map for building road network [8], General Transit Feed Specification data (GTFS) for public transit behavior study [31].
- *Traffic Data*: PeMS [9] for traffic flow modeling, Regional Integrated Transportation Information System (RITIS) [32] and Work Zone Data Exchange (WZDx) [33] for work zone traffic impact modeling;
- *Stated/Revealed Preference Survey Data*: Southern California Autonomous Vehicle Preference Survey [34] for learning CAV impact on travel choices.

### B. Implementation for AI Agents

Gathering socio-demographic information, travel diaries, and trajectories is essential to construct AI agents. However, collecting such data using location-based services raises significant privacy concerns [35], particularly regarding the identification of specific POIs within trajectories. To address privacy concerns, AI agents are developed as anonymous digital simulacra of human beings, designed to emulate travel behaviors while safeguarding personal privacy. By masking the link to any specific person, these agents can process and simulate detailed information without leaking personal data.

*1) AI agent mobility pattern modeling:* We develop the Deep Activity Model [11], a generative deep learning approach to model human mobility patterns using socio-demographic information, travel diaries, and travel trajectories. The model is trained to learn the relationship between an AI agent's socio-demographic information and mobility (activity and travel) patterns. Importantly, the input data includes information about the AI Agent's household members and social network, allowing the model to learn interactive behaviors and their impact on activities across multiple AI agents. The Deep Activity Model learns the dependencies between activities, considering factors such as start and end times, duration, and location for each activity. It employs an auto-regressive generation method, sequentially creating a full day's travel activities starting from midnight. As shown in Fig. 2, this approach enables the model

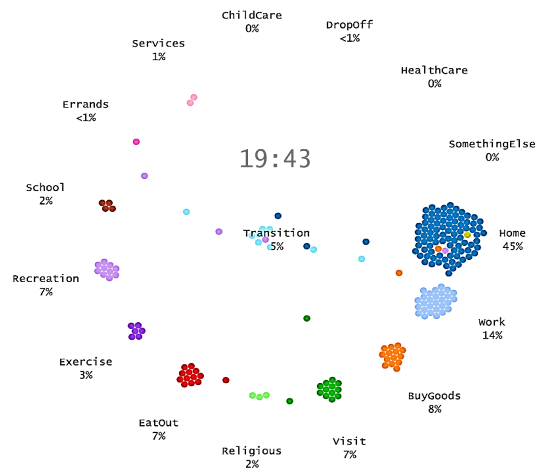


Fig. 2: Activity of 200 AI agents on a weekday at 19:43. Each bubble stands for an AI agent, and its travel behavior is displayed by the movement from one activity to another.

to generate realistic activities for AI agents by capturing the complex interdependency within households, social networks, and individual activities.

Our experiments validate the model's robustness and versatility. It demonstrates strong performance when fine-tuned with data from diverse regions, including California, the Puget Sound area, and Mexico City. It allows us to capture various travel behaviors and patterns across geographical and cultural contexts.

*2) AI Agent Modeling Empowered by Knowledge-Based AI:* To further enhance the intelligence of our AI agents, we implement principles of embodied AI [36], [37], making these agents highly sophisticated and capable of understanding complex contexts, reasoning about their environment, and making informed decisions. This advanced level of AI can be achieved through knowledge-based AI approaches.

LLMs, as one typical knowledge-based AI, are used to improve the reasoning capabilities of AI agents, as summarized in Fig. 3. We implement LLMs for data mining to interpret each stay point in a trajectory and annotate it with possible POIs and activities [38]. This process enriches time series data processing tasks with natural language processing, enabling AI agents to understand and contextualize trajectories. To address situations where social-demographic information is lacking, we harness the sophisticated reasoning capabilities of LLMs to reconstruct AI agent profiles. Building upon the annotated human trajectory data, LLMs infer uncovered AI agents' attributes, such as occupation, travel preferences, and income level. With these reconstructed profiles and annotated trajectories, LLMs can generate realistic activities for AI agents and guide the activity generation process. This approach allows our AI agents to exhibit more human-like decision-making in complex mobility scenarios, considering not only raw data but also contextualized and implied information.

### C. Implementation for AI Networks

*1) Agent-Based Simulation:* We have developed a large-scale simulation platform to obtain realistic travel trajectories

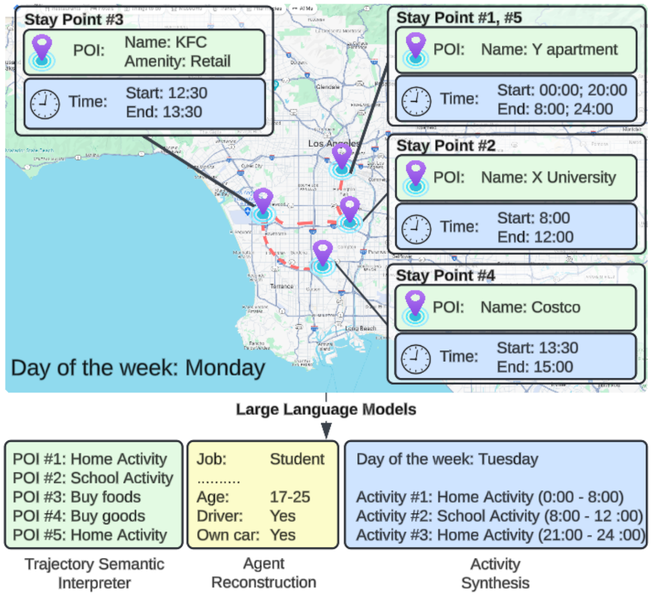


Fig. 3: Semantic trajectory analysis. (top) Raw trajectory data visualization and (bottom) analysis results using LLM. Agent reconstruction is achieved with multi-day trajectories.

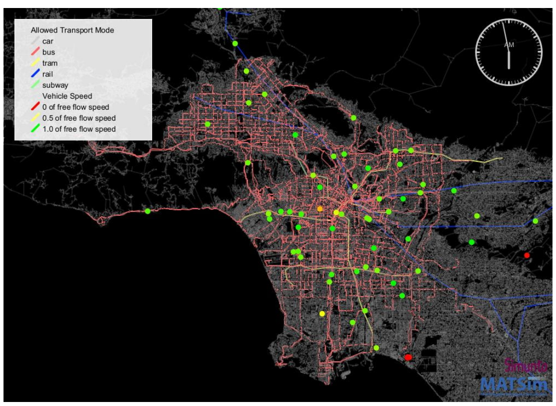


Fig. 4: Multi-modal mega-city macroscopic simulation in Los Angeles County. The movement of AI agents is presented in dots, while the traffic evolution is displayed in links.

and traffic conditions. Our Deep Activity Model generates realistic mobility patterns, and our agent synthesis includes diverse transportation modes for a comprehensive mobility representation [11]. As shown in Fig. 4, our network incorporates local road and public transit infrastructures, and AI agents are loaded into the network.

Our simulation utilizes a macroscopic agent-based simulation toolkit called MATSim [39], which provides a nuanced platform for integrating our detailed AI agent profiles and network models. A data-driven approach optimized the model by clustering highways based on traffic data. In future work, we will formulate this optimization problem with a meta-model for inference, enabling an auto-calibration process for the AI network.

To demonstrate the capabilities of our simulation platform, We apply our simulation and calibration pipeline for Los Angeles County [17]. The model incorporates detailed multi-modal networks, a million-level size of agents, and provisions for micro-mobility and electric vehicles. Through rigorous validation against real-world traffic counts and an extensive calibration process involving multiple iterations, the model demonstrated a robust capability to predict and analyze the traffic dynamics of a mega-city. This offers valuable insights for urban planners and policymakers to evaluate the impact of future infrastructural changes, policy implementations, and technological advancements on city-wide mobility and accessibility. The simulation outputs are intended to support decision-making processes to enhance transportation efficiency, reduce emissions, and improve overall urban mobility in densely populated urban areas.

*2) Network State Modeling:* Traffic state prediction is a crucial component of intelligent transportation systems (ITS) [40], enhancing the AI network's estimation and prediction capability. We have made significant breakthroughs in estimating and predicting traffic state in special events. For example, We developed a novel deep-learning model to predict traffic speed and incident likelihood during planned work zone events [41]. Also, we model the long-term congestion and short-term speed patterns during hurricane evacuations [42], [43]. These accurate and timely predictions, especially predictions under special weather and roadway conditions, enhance traffic management and support congestion prevention and mitigation efforts [44] [45].

*3) Interactions Between AI Agents and AI Networks:* The interaction between AI agents and AI networks, especially in estimating how AI agents' behavior affects networks, is central to our modeling approach. Our study examines how AI agents adopting CAVs and Electric Vehicle (EV) affect the network and alter mobility patterns [34], [46]–[48] in large urban transportation networks, using a traditional choice model-based techniques, i.e., activity-based models (ABMs). However, developing ABMs is often time-consuming and resource-intensive, requiring extensive parameter calibrations, and hence limits the adaptability and scalability of AI agents and networks. Therefore, a more efficient method is in demand. This MSDT framework provides new solutions to learning interaction patterns between AI agents and networks in a highly automated approach.

## IV. Future Works

Looking ahead, this MSDT framework consists of AI agents and networks that have the potential to revolutionize the way we model and analyze mobility systems. It can more accurately simulate the impacts of new technologies, policies, and infrastructure changes. Future work will focus on enhancing agent learning capabilities, improving the adaptability of AI agents and networks, and developing more sophisticated interaction models. This approach will provide planners with valuable insights for creating more efficient and sustainable mobility systems.

## References

[1] United Nations, "68% of the world population projected to live in urban areas by 2050, says un," 2018, accessed: 2024-06-17. [Online]. Available: https://www.un.org/development/desa/en/news/population/2018-revision-of-world-urbanization-prospects.html

[2] G. Salierno, L. Leonardi, and G. Cabri, "A big data architecture for digital twin creation of railway signals based on synthetic data," *IEEE Open Journal of Intelligent Transportation Systems*, 2024.

[3] Z. Wang, O. Zheng, L. Li, M. Abdel-Aty, C. Cruz-Neira, and Z. Islam, "Towards next generation of pedestrian and connected vehicle in-the-loop research: A digital twin co-simulation framework," *IEEE Transactions on Intelligent Vehicles*, 2023.

[4] Z. Hu, S. Lou, Y. Xing, X. Wang, D. Cao, and C. Lv, "Review and perspectives on driver digital twin and its enabling technologies for intelligent vehicles," *IEEE Transactions on Intelligent Vehicles*, vol. 7, no. 3, pp. 417–440, 2022.

[5] X. Han, Z. Meng, X. Xia, X. Liao, Y. He, Z. Zheng, Y. Wang, H. Xiang, Z. Zhou, L. Gao *et al.*, "Foundation intelligence for smart infrastructure services in transportation 5.0," *IEEE Transactions on Intelligent Vehicles*, 2024.

[6] U.S. Department of Transportation, "National Household Travel Survey (NHTS)," 2024. [Online]. Available: https://nhts.ornl.gov/

[7] Veraset, "Global mobility and location data provider," https://www.veraset.com/, 2024, accessed: 2024-06-17.

[8] OpenStreetMap contributors, "Planet dump retrieved from https://planet.osm.org ," https://www.openstreetmap.org, 2017.

[9] California Department of Transportation, "Pems," https://pems.dot.ca.gov, 2020.

[10] Federal Highway Administration (FHWA), "Highway Safety Information System (HSIS)," https://www.hsisinfo.org/, 2020, accessed: 2024-06-22.

[11] X. Liao, B. Y. He, Q. Jiang, C. Kuai, and J. Ma, "Deep activity model: A generative approach for human mobility pattern synthesis," *arXiv preprint arXiv:2405.17468*, 2024.

[12] S. Teng, L. Li, Y. Li, X. Hu, L. Li, Y. Ai, and L. Chen, "Fusionplanner: A multi-task motion planner for mining trucks via multi-sensor fusion," *Mechanical Systems and Signal Processing*, vol. 208, p. 111051, 2024.

[13] Z. Wang, X. Liao, C. Wang, D. Oswald, G. Wu, K. Boriboonsomsin, M. J. Barth, K. Han, B. Kim, and P. Tiwari, "Driver behavior modeling using game engine and real vehicle: A learning-based approach," *IEEE Transactions on Intelligent Vehicles*, vol. 5, no. 4, pp. 738–749, 2020.

[14] K. Messaoud, I. Yahiaoui, A. Verroust-Blondet, and F. Nashashibi, "Attention based vehicle trajectory prediction," *IEEE Transactions on Intelligent Vehicles*, vol. 6, no. 1, pp. 175–185, 2020.

[15] Z. Qin, A. Ji, Z. Sun, G. Wu, P. Hao, and X. Liao, "Game theoretic application to intersection management: A literature review," *IEEE Transactions on Intelligent Vehicles*, 2024.

[16] Intelligent Transportation Systems Joint Program Office. (2024) Complete Streets AI. U.S. Department of Transportation. [Online]. Available: https://its.dot.gov/csai/

[17] B. Y. He, Q. Jiang, H. Ma, and J. Ma, "Multi-agent multimodal transportation simulation for mega-cities: Application of los angeles," 2024.

[18] S. Hasan, J. Gorospe, S. Girs, A. A. Gómez, and E. Uhlemann, "Platoonsafe: An integrated simulation tool for evaluating platoon safety," *IEEE Open Journal of Intelligent Transportation Systems*, 2023.

[19] K. Yamamoto, R. Teng, and K. Sato, "Simulation evaluation of vehicle movement model using spatio-temporal grid reservation for automated valet parking," *IEEE Open Journal of Intelligent Transportation Systems*, vol. 4, pp. 261–266, 2023.

[20] G. O. Kagho and M. Balac, "Identifying and planning for group travellers in on-demand mobility models," *IEEE Open Journal of Intelligent Transportation Systems*, 2023.

[21] Z. Wang, P. Keo, and M. Saberi, "Real-time traffic state measurement using autonomous vehicles open data," *IEEE Open Journal of Intelligent Transportation Systems*, 2023.

[22] S. Zhang, C. Zhang, S. Zhang, and J. James, "Attention-driven recurrent imputation for traffic speed," *IEEE Open Journal of Intelligent Transportation Systems*, vol. 3, pp. 723–737, 2022.

[23] Intelligence Advanced Research Projects Activity. (2024) HAYSTAC: Hidden Activity Signal and Trajectory Anomaly Characterization. U.S. Office of the Director of National Intelligence. [Online]. Available: https://www.iarpa.gov/research-programs/haystac

[24] K. Sotiropoulos, L. Zhao, P. J. Liang, and L. Akoglu, "Adamm: Anomaly detection of attributed multi-graphs with metadata: A unified neural network approach," in *2023 IEEE International Conference on Big Data (BigData)*. IEEE, 2023, pp. 865–874.

[25] D. Yi, J. Su, L. Hu, C. Liu, M. Quddus, M. Dianati, and W.-H. Chen, "Implicit personalization in driving assistance: State-of-the-art and open issues," *IEEE Transactions on Intelligent Vehicles*, vol. 5, no. 3, pp. 397–413, 2019.

[26] S. Kitajima, H. Chouchane, J. Antona-Makoshi, N. Uchida, and J. Tajima, "A nationwide impact assessment of automated driving systems on traffic safety using multiagent traffic simulations," *IEEE Open Journal of Intelligent Transportation Systems*, vol. 3, pp. 302–312, 2022.

[27] A. Ng, "A chat with andrew on mlops: From model-centric to data-centric ai," 2021, youTube. [Online]. Available: https://www.youtube.com/watch?v=06-AZXmwHjo

[28] U.S. Census Bureau, "American Community Survey (ACS) Data," 2024. [Online]. Available: https://www.census.gov/programs-surveys/acs/data.html

[29] S. C. A. of Governments (SCAG), "2016 regional travel demand model and model validation report," Southern California Association of Governments, Tech. Rep., 2020, accessed: date-of-access.

[30] N. I. of Statistics and Geography, "Origin-destination survey in households of the metropolitan zone of the valley of mexico," https://en.www.inegi.org.mx/programas/eod/2017/, 2017.

[31] General Transit Feed Specification (GTFS), "General Transit Feed Specification (GTFS)," 2024. [Online]. Available: https://gtfs.org/

[32] Center for Advanced Transportation Technology Laboratory, "Regional Integrated Transportation Information System (RITIS)," 2024. [Online]. Available: https://www.cattlab.umd.edu/ritis/

[33] U.S. Department of Transportation, "Work Zone Data Exchange (WZDx)," 2024. [Online]. Available: https://www.transportation.gov/av/data/wzdx

[34] B. Y. He, Q. Jiang, and J. Ma, "Connected automated vehicle impacts in southern california part-i: Travel behavior and demand analysis," *Transportation research part D: transport and environment*, vol. 109, p. 103329, 2022.

[35] M. Gruteser and D. Grunwald, "Anonymous usage of location-based services through spatial and temporal cloaking," in *Proceedings of the 1st international conference on Mobile systems, applications and services*, 2003, pp. 31–42.

[36] J. Duan, S. Yu, H. L. Tan, H. Zhu, and C. Tan, "A survey of embodied ai: From simulators to research tasks," *IEEE Transactions on Emerging Topics in Computational Intelligence*, vol. 6, no. 2, pp. 230–244, 2022.

[37] B. Li, X. Li, Y. Cui, X. Bian, S. Teng, S. Ma, L. Fan, Y. Tian, F.-Y. Wang *et al.*, "Integrating large language models and metaverse in autonomous racing: An education-oriented perspective," *IEEE Transactions on Intelligent Vehicles*, 2024.

[38] Y. Liu, C. Kuai, H. Ma, X. Liao, B. Y. He, and J. Ma, "Semantic trajectory data mining with llm-informed poi classification," *arXiv preprint arXiv:2405.11715*, 2024.

[39] A. Horni, K. Nagel, and K. W. Axhausen, *The Multi-Agent Transport Simulation MATSim*, A. Horni, K. Nagel, and K. W. Axhausen, Eds. London: Ubiquity Press, 2016.

[40] Z. Ghandeharioun and A. Kouvelas, "Link travel time estimation for arterial networks based on sparse gps data and considering progressive correlations," *IEEE Open Journal of Intelligent Transportation Systems*, vol. 3, pp. 679–694, 2022.

[41] Q. Jiang, X. Liao, Y. Gong, and J. Ma, "An attention-based multi-context convolutional encoder-decoder neural network for work zone traffic impact prediction," *arXiv preprint arXiv:2405.21045*, 2024.

[42] Q. Jiang, B. Y. He, C. Lee, and J. Ma, "Deploying scalable traffic prediction models for efficient management in real-world large transportation networks during hurricane evacuations," *arXiv preprint arXiv:2406.12119*, 2024.

[43] U.S. DoT, "Integrated Modeling for Road Condition Prediction Phase 4: Final Report," 2024. [Online]. Available: https://rosap.ntl.bts.gov/view/dot/74645

[44] S. Zhang, C. Zhang, S. Zhang, and J. James, "Attention-driven recurrent imputation for traffic speed," *IEEE Open Journal of Intelligent Transportation Systems*, vol. 3, pp. 723–737, 2022.

[45] Z. Cheng, X. Wang, X. Chen, M. Trépanier, and L. Sun, "Bayesian calibration of traffic flow fundamental diagrams using gaussian processes," *IEEE Open Journal of Intelligent Transportation Systems*, vol. 3, pp. 763–771, 2022.

[46] Q. Jiang, B. Y. He, and J. Ma, "Connected automated vehicle impacts in southern california part-ii: Vmt, emissions, and equity," *Transportation research part D: transport and environment*, vol. 109, p. 103381, 2022.

[47] H. Ma, B. Y. He, T. Kaljevic, and J. Ma, "A two-sided model for ev market dynamics and policy implications," *arXiv preprint arXiv:2405.17702*, 2024.

[48] Q. Jiang, N. Zhang, B. Y. He, C. Lee, and J. Ma, "Large-scale public charging demand prediction with a scenario-and activity-based approach," *Transportation research part A: Policy and Practice*, vol. 179, p. 103935, 2024.