# UMFM 2025

Proceedings of the 1st ACM SIGSPATIAL International Workshop on
Urban Mobility Foundation Models
(UMFM 2025)
November 3, 2025, Minneapolis, Minnesota, USA


Editors:
Shushman Choudhury, Google Research, USA

Neha Arora, Google Research, USA

Cyrus Shahabi, University of Southern California, USA

**The Association for Computing Machinery, Inc.**
**1601 Broadway, 10th Floor**
**New York, NY 10019-7434**



**Notice to Past Authors of ACM-Published Article**

ACM intends to create a complete electronic archive of all articles and/or other material previously published by ACM. If you have written a work that was previously published by ACM in any journal or conference proceedings prior to 1978, or any SIG Newsletter at any time, and you do NOT want this work to appear in the ACM Digital Library, please inform permissions@acm.org, stating the title of the work, the author(s), and where and when published.

**ACM ISBN:** 979-8-4007-2188-5

Additional copies may be ordered prepaid from:
**ACM Order Department**
P.O. BOX 11405
New York, NY 10286-1405
Phone: 1-800-342-6626 (USA and Canada)
+1-212-626-0500 (Global)
Fax: +1-212-944-1318
E-mail: acmhelp@acm.org
Printed in the USA

# FOREWORD

The inaugural **ACM SIGSPATIAL International Workshop on Urban Mobility Foundation Models (UMFM'25)** marks a pivotal moment for the geospatial AI community. In recent years, *foundation models* have transformed fields such as natural language processing, computer vision, and biology by demonstrating how large-scale, self-supervised pre-training on heterogeneous data can yield models that generalize across tasks, domains, and even modalities. Yet, as global society becomes increasingly mobile, instrumented, and spatially interconnected, the **spatial and mobility domains have remained conspicuously underrepresented in this transformation**.

Spatial and urban mobility data possess unique characteristics that make them both challenging and essential for foundation-model research. Unlike text or images, spatio-temporal data encode **geometric relationships, human intent, and environmental context**, often across multiple scales, from individual movements to city-wide flow networks. Human mobility encapsulates how people interact with physical spaces, how infrastructure shapes accessibility, and how collective behavior responds to shocks such as disasters, pandemics, or policy changes. Developing *Spatial Foundation Models* (SFMs) and, more specifically, *Urban Mobility Foundation Models* (UMFMs) promises to provide **a unifying representation space** where geographic entities, trajectories, places, and infrastructure can be jointly modeled, queried, and reasoned about.

The motivation extends far beyond prediction. Foundation models for urban mobility can support **generalizable reasoning** about the built environment: simulating future movement scenarios, testing policy interventions, transferring insights between cities, and even embedding ethical principles such as equity, inclusion, and sustainability into data-driven decision support. UMFMs thus sit at the intersection of **AI, geography, urban science, and governance**, a domain where data scale meets societal relevance.
As foundation models reshape scientific disciplines, this workshop recognizes the urgency of establishing **geospatial analogues**, models capable of capturing the semantics of *where*, *when*, and *why* movement happens, while maintaining interpretability, fairness, and privacy.

The first edition of UMFM brings together eight short papers that collectively define this emerging area. They span theoretical frameworks, semantic representation learning, generative simulation, disaster behavior modeling, infrastructure analysis, and optimization theory. Together, they demonstrate that *the future of spatial computing will be foundation-driven*: models pre-trained on massive, multimodal mobility corpora that can reason about movement, transfer across geographies, and provide a common substrate for spatial AI research and applications.

## 1. Conceptual and Theoretical Foundations

Two papers establish the theoretical groundwork for defining what an urban mobility foundation model should be.

*Song Gao*'s **"Issues on Pre-Training and Applications of Urban Mobility Foundation Models"** offers a panoramic perspective on the methodological and ethical challenges of UMFMs. It examines the complexities of multimodal data integration, spatial zoning, and sampling, while highlighting the importance of interpretability, transparency, and geo-privacy. Gao emphasizes that UMFMs must be *human-centered*, sensitive to place, culture, and community, to ensure equitable decision support in planning and governance.

*Mohammad Hashemi and Andreas Züfle* advance this conceptual agenda in **"From Points to Places: Towards Human Mobility-Driven Spatiotemporal Foundation Models via Understanding Places."** They argue that the next generation of spatial foundation models must transcend discrete points of interest to embrace *places*, dynamic, socially meaningful regions shaped by human activity. Their formal framework defines places as hierarchically nested spatial entities, suggesting that foundation models should learn multi-scale, semantically rich representations of human–environment interactions.

## 2. Representation Learning for Mobility and Semantics

Moving from theory to practice, two contributions explore new representations of mobility that embed semantics, behavior, and multimodal context.

*Ryuichi Sudo, Keiichi Ochiai, and Hiroyuki Toda* present **"From Trajectories to Semantic Mobility: Context-Aware Embedding for Route Recommendation."** Their Transformer-based model transforms raw GPS traces into *semantic trajectories* enriched with behavioral features (speed, acceleration, direction) and environmental context (POI density, land use, and amenities). This representation captures both short- and long-term user preferences, enabling personalized route recommendations that reflect the *why* behind mobility choices rather than just the *where*.

Complementing this approach, *Min Namgung, JangHyeon Lee, Yijun Lin, and Yao-Yi Chiang* introduce **"Capturing Shared and Unique Information in Multimodal Region Representations for Urban Mobility Prediction."** Their framework disentangles shared and modality-specific information from diverse urban data, remote sensing imagery, textual descriptions, and mobility flows, to build regional embeddings that generalize across data types and geographies. By modeling both common and unique patterns, the work advances the goal of multimodal foundation models that understand cities as integrated systems.

## 3. Infrastructure and Multimodal Urban Contexts

A key insight emerging from UMFM research is that **the built environment is not static**: it evolves, interacts with human behavior, and shapes mobility patterns over time.

*Jon Atkins, Joao Rulff, Claudio T. Silva, and Maryam Hosseini* address this challenge in **"StreetTransformer: Precedent Retrieval and Multimodal Fusion with Vision-Language Models."** They develop a retrieval-augmented multimodal pipeline that fuses aerial imagery, infrastructure features, and planning documents to track street-level changes across 130,000 intersections in New York City. By linking detected changes to precedent projects, the system establishes a foundation for models that learn how design interventions alter urban mobility, a major step toward infrastructure-aware foundation modeling.

At a broader scale, *Mohamed Hemdan, Youssef Hussein, and Mohamed F. Mokbel*'s **"Large Language Models for Spatial Analysis Tasks"** surveys how large language models are being adapted for spatial reasoning across traffic, trajectory, and point-of-interest analysis. Their taxonomy categorizes approaches based on model adaptation strategies, prompting, fine-tuning, architecture-only training, and customized loss functions, providing a roadmap for integrating LLMs into geospatial modeling. The paper highlights emerging opportunities to combine linguistic understanding with spatial computation for tasks such as trajectory reasoning and region-level description.

## 4. Simulation, Prediction, and Optimization

The final cluster of papers demonstrates how foundation models can simulate, predict, and optimize mobility systems in dynamic or uncertain environments.

*DongHak Lee, Jiayi Weng, Vaidehi Raipat, and Takahiro Yabe*'s **"Evaluating the Predictability of Disaster Evacuation Behavior using LLM Agent Simulations"** uses large language models as *behavioral simulators*. Synthetic agents, parameterized by mobility histories and demographic data, emulate human decision-making during disasters such as the Marshall and Route Fires. Fine-tuned models exhibit improved alignment with real-world evacuation data, revealing that language-based reasoning can replicate elements of human collective behavior under stress.

Siyu Li et al. extends this predictive paradigm through **"Lagrangian Multiplier Approaches to Mobility Flow Optimization in Foundation Models."** By embedding physical conservation laws into pre-training objectives, they ensure that learned mobility representations obey the fundamental constraints of flow continuity. This approach bridges deep learning and physics-based optimization, demonstrating how foundation models can be both expressive and physically consistent, a necessary step toward trustworthy, simulation-ready mobility AI.

## 5. Outlook

Together, these eight papers illustrate the extraordinary breadth and promise of the emerging UMFM community. They span the full pipeline, from data foundations to ethical reflection, from semantic representation to multimodal integration, and from cognitive simulation to optimization theory. Across these dimensions, a common thread emerges: **the pursuit of unified, interpretable, and generalizable models that can learn the language of human movement and the structure of the city itself.**

Looking forward, several grand challenges remain: constructing benchmark datasets across cities and scales, developing efficient multimodal pre-training strategies, and safeguarding geo-privacy while ensuring transparency and fairness. Yet the works presented at UMFM'25 collectively demonstrate that spatial foundation modeling is not just an extension of AI, it is an evolution of how we understand urban life.

The **ACM SIGSPATIAL UMFM'25 Workshop** thus stands as both an origin and a manifesto: a commitment to building foundation models that see not only space, but meaning, models that capture the living pulse of cities and the mobility of the people who make them.

Shushman Choudhury, Google Research, USA

Neha Arora, Google Research, USA

Cyrus Shahabi, University of Southern California, USA

ACKNOWLEDGEMENTS

We extend our gratitude to the program committee for their vital efforts in reviewing submissions, ensuring the quality of the accepted papers.

## SPONSORS

CORPORATE SPONSORS:

Platinum

Silver

Google

Bronze

esri

THE SCIENCE OF WHERE

ORACLE

QUANTUM
GEOMETRY LLC

PUBLISHER SPONSORS:

Bronze

Springer

TRAVEL SPONSORS [USA STUDENTS]:

NSF