# Mobility-based infrastructure planning in the digital twin era

Ludovica Alessandria[1] Jinpeng Li[1] Markus Schläpfer[1]

[1]Columbia University, USE-Lab

**Abstract**

Human mobility shapes the demand for urban infrastructure services, including energy, transportation, and water. However, predicting how mobility patterns will evolve, especially in response to urban redevelopment or the introduction of new transit infrastructure, has remained challenging. Traditional models capture only coarse regularities, while recent deep-learning approaches often generalize poorly beyond the cities on which they were trained. This paper outlines a path toward a “Foundation Model for Urban Mobility,” capable of providing fine-grained, transferable predictions across diverse cities. The basis for this framework is the integration of principles from urban science into graph-based deep learning approaches, enabling the model to capture universal mobility regularities and allowing for zero-shot forecasting in data-scarce settings. This approach becomes fully operational when embedded within Urban Digital Twins (UDTs), where local UDTs refine it using high-resolution city data. Through federated learning, these locally adapted models are then aggregated into a shared cross-city representation, forming the Foundation Model for Urban Mobility while avoiding the exchange of sensitive data. We conclude by discussing the technical and governance challenges of integrating such models into UDTs, as well as their potential to support anticipatory, scalable, and privacy-preserving infrastructure planning.

**USE-Lab Working Paper #2025–12–001**
`https://urbansystems.civil.columbia.edu`

# Mobility-Based Infrastructure Planning in the Digital Twin Era

Ludovica Alessandria, Jinpeng Li and Markus Schläpfer

**Abstract** Human mobility shapes the demand for urban infrastructure services, including energy, transportation, and water. However, predicting how mobility patterns will evolve, especially in response to urban redevelopment or the introduction of new transit infrastructure, has remained challenging. Traditional models capture only coarse regularities, while recent deep-learning approaches often generalize poorly beyond the cities on which they were trained. This chapter outlines a path toward a 'Foundation Model for Urban Mobility', capable of providing fine-grained, transferable predictions across diverse cities. The basis for this framework is the integration of principles from urban science into graph-based deep learning approaches, enabling the model to capture universal mobility regularities and allowing for zero-shot forecasting in data-scarce settings. This approach becomes fully operational when embedded within Urban Digital Twins (UDTs), where local UDTs refine it using high-resolution city data. Through federated learning, these locally adapted models are then aggregated into a shared cross-city representation, forming the Foundation Model for Urban Mobility while avoiding the exchange of sensitive data. We conclude by discussing the technical and governance challenges of integrating such models into UDTs, as well as their potential to support anticipatory, scalable, and privacy-preserving infrastructure planning.

Markus Schläpfer
Columbia University, Department of Civil Engineering and Engineering Mechanics, New York, NY 10027, e-mail: m.schlaepfer@columbia.edu

Ludovica Alessandria
Sapienza University of Rome, Italy, and Columbia University, Department of Civil Engineering and Engineering Mechanics, New York, NY 10027

Jinpeng Li
Columbia University, Department of Civil Engineering and Engineering Mechanics, New York, NY 10027

# 1 Introduction

Human mobility lies at the core of urban life, driving not only the social and economic functioning of cities but also the demand for civil infrastructure services [1–3]. Indeed, daily energy, water, transportation and telecommunication needs are strongly correlated with the time-varying presence and movement of people (see Box 1). The planning of large-scale civil infrastructures in cities therefore requires adequate predictions of future mobility flows. An example is the planning of a new subway line such as the extension of the Second Avenue Subway in New York City. Predicting how such a new transit line will alter mobility patterns is crucial not only for determining the route alignment and capacity design, but also for assessing the impacts on other infrastructure systems, such as the energy supply, and whether they can accommodate the associated additional loads.

**Box 1: Mobility-informed energy infrastructure planning.**
High-resolution human mobility flow predictions can substantially improve infrastructure demand forecasts. This is particularly relevant for energy systems since much of urban energy use, particularly in residential, commercial, and transportation domains, is directly or indirectly shaped by where and when people are present [4, 5], see Fig. B1a. As an example, in a recent research project, mobility traces from anonymized large-scale mobile phone data were used to simulate future electric vehicle (EV) charging demand profiles in urban neighborhoods in Singapore, allowing to identify bottlenecks in the existing electricity supply network, see Fig. B1b.

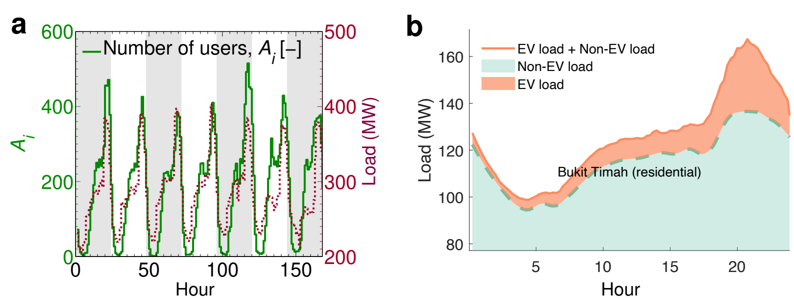


**Fig. B1. a** The presence of people is strongly correlated with the local electricity load. Shown are the number mobile phone users present in a neighborhood of Dakar, Senegal, versus the dynamic electricity demand (load) curve. Adapted from [4]. **b** Mobility-based electricity demand (load) predictions due to future EV charging in a residential neighborhood in Singapore. Adapted from [6].

However, despite this importance, – and despite the growing availability of large-scale data and rapid advances in computational methods – our ability to predict the mobility flows of people in cities remains surprisingly limited. The main reason is the complexity of the daily movement of people that results from a myriad of individual and ever-changing motivations: from daily commutes to work (which have recently declined due to work-from-home practices) and grocery shopping (some of which is being increasingly being replaced by home delivery services), to leisure and social activities like visiting a cinema or friends.

Traditional mobility modeling approaches like the well-known gravity model or agent-based models either struggle with fine-grained predictions or require extensive parameter calibrations. Recent advances in machine learning, especially in deep learning, offer significant gains in predictive accuracy, but these models typically generalize poorly beyond the specific cities and time periods on which they were trained. This limits their usefulness for forward-looking infrastructure planning in diverse urban contexts.

Among the emerging digital tools supporting urban planning and infrastructure management, Digital Twins (DTs) have gained increasing traction across multiple engineering and operational domains. When applied to urban systems, they evolve into Urban Digital Twins (UDTs), providing a platform capable of integrating dynamic data and simulation processes to enable more advanced and responsive decision-making. In this context, despite their potential operational centrality, urban mobility flows are often represented merely as an informational layer within UDTs used for decision-making. This gap significantly limits their effectiveness for anticipatory planning, as the redistribution of people in response to new infrastructures, policies, or disruptions constitutes one of the primary drivers of demand for urban infrastructure services. Developing a universal and transferable, and thus predictive, representation of mobility is therefore essential to strengthen the predictive and strategic value of next-generation UDT frameworks.

In this Chapter, we argue that we can now harness recent advances in deep learning to build a *lightweight foundation model* (LFM) for urban mobility, similar to what has been adopted in other fields like weather forecasting, and subsequently, in a forward-looking perspective, its potential integration into a broader UDT framework. This would transform UDTs from predominantly descriptive platforms into 'cognitive' infrastructures capable of supporting infrastructure planning through anticipatory governance. To that end, we first review the state-of-the art in human mobility modeling, and then develop a trajectory from deep learning models for individual cities to a lightweight foundation model that is applicable to diverse urban settings, and finally, its integration into a larger urban digital twin framework that can be used for infrastructure planning, in particular energy systems.

## 2 From Traditional Flow Models to Deep Learning Approaches

Mobility flows in cities are studied across a wide range of academic disciplines, including transportation engineering, computer science, complexity science, human geography, urban economics, and public health [7–11]. A basic representation is given by origin–destination (OD) flows, which quantify how many trips occur between different locations in a city over a given time period. Approaches for modeling

OD flows can be broadly categorized into traditional models and machine-learning based models.

## 2.1 Traditional Models

Traditional mobility flow models can be divided into 'parsimonious' models and Land Use and Transport Integrated (LUTI) models.

**Parsimonious models** generally aim at incorporating a minimal set of parameters and mechanisms, while being able to account for a broad range of empirical observations [12]. Among the earliest and most widely used are gravity models, which estimate OD flows based on the attractiveness of the origin and destination and the distance between them [13–15]. Extensions to the basic formulation often incorporate additional factors such as travel time [16]. Intervening opportunities models such as the radiation model [17] are a related class of models that predict mobility patterns based not on distance alone but on the number of opportunities between the origin and destination. Moreover, using a statistical model of individual mobility behavior [18, 19], the TimeGeo framework enrich sparse mobility data to extrapolate to entire populations [20]. While parsimonious models offer interpretability and analytical tractability [12], they often struggle to account for complex behavioral and contextual factors such as temporal variability in trip purposes or socio-demographic heterogeneity [21], which can lead to significant prediction errors in urban settings [22]. Recently, a universal mathematical law of human travel has been discovered [23] by leveraging an extensive dataset of mobile phone records. The underlying general principle is a nested hierarchy of central places [24], whereas higher-order centres with specialized functions (e.g., a jewelry store) – reflected in lower visitation frequencies and larger travel distances – also embrace non-specialized functions of lower-order centres (e.g., grocery stores) – reflected in higher visitation frequencies and shorter distances. The OD flow predictions based on this law (Fig. 1a) are still too coarse for detailed infrastructure applications but demonstrate that transferability with zero-shot predictions across different geographical regions are possible without parameter calibrations.

**Land Use Transport Interaction (LUTI) models** include agent-based and activity-based simulations and aim to represent the detailed interdependencies between land use and transportation systems. Tools such as MATSim and the BEAM–UrbanSim–ActivitySim framework generate synthetic populations and detailed daily activity schedules, typically based on household travel surveys [25, 26]. These schedules are then used as fixed inputs to simulate transportation responses under baseline and future scenarios. These models typically are highly granular but require a large number of assumptions and calibrated parameters, introducing the risk of overfitting and limiting transferability across cites.

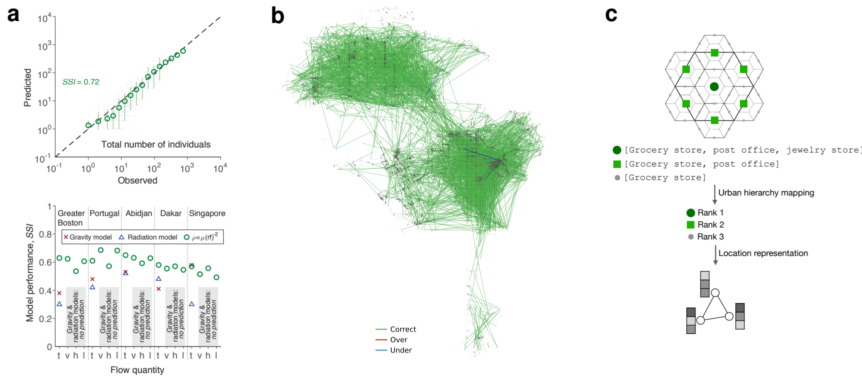


**Fig. 1 Hidden universalities in the movement of people allow for lightweight, zero-shot predictions**. **a**, The visitation law [23] enables globally transferable zero-shot mobility predictions using population density and geographic distance as the sole inputs. However, the spatial resolution is often too coarse for many infrastructure applications. Upper panel: predicted versus observed number of travelers from different origin locations to Back Bay, Boston (dashed line indicates perfect agreement). Lower panel: performance across world regions for number of trips (t), visitors (v), high-frequency visitors (h), low-frequency visitors (l). Model performance is measured by the Sørensen–Dice similarity (SSI). **b**, Graph-based deep learning models, such as `UrbanPulse` [27], achieve fine-grained predictions, but the geographic transferability is limited. **c** To enable predictions that are both fine-grained and globally transferable, it is promising to use a universal location representation that maps the region-specific semantics of attraction points to their rank within the hierarchy of central places.

## 2.2 Deep Learning Models

Deep learning offers a promising framework to address some of these challenges. By learning compact latent factors that encode spatial interactions and temporal evolution, deep models can capture nonlinear cross-effects among OD pairs and exogenous drivers [28]. Moreover, deep architectures provide several additional advantages such as a seamless integration of heterogeneous data sources, the ability to model long-range temporal dependencies without manual feature engineering, and the use of graph priors to represent network-constrained movement.

The spatio-temporal nature of OD flows has inspired architectures that explicitly separate and then recombine temporal and spatial components. Temporally, recurrent units, such as LSTM and GRU, handle burstiness and seasonality, while temporal convolutional networks (TCNs) with dilated kernels extend receptive fields efficiently. Sequence-to-sequence decoders with attention mechanisms further stabilize multi-step predictions. Spatially, graph convolutional networks (GCNs), graph attention networks (GATs), and diffusion convolutions propagate and transform features along adjacency structures defined by road connectivity, transit links, or functional proximity [27], see Fig. 1b. Problem-specific graph constructions, such as bipartite OD graphs or adaptive learned adjacency matrices, enhance the inductive bias for diverse data regimes. Integrative models like DCRNN [29], STGCN [30], Graph

WaveNet [31], and ST-Transformer [32] combine temporal and spatial modules, often incorporating exogenous signals such as weather, events, and land-use indicators through cross-attention or gating mechanisms. Attention mechanisms and transformer architectures extend these capabilities by allowing models to focus dynamically on the most relevant spatial–temporal context. This flexibility improves the modeling of long-range dependencies and supports pretraining across cities, followed by fine-tuning for local adaptation. Emerging large-scale spatio-temporal transformers integrate lightweight adapters or graph-aware positional encodings, balancing scalability with the need to respect network constraints.

Despite significant progress, several challenges must be addressed to make OD forecasting a routine and reliable tool:

- **Cross-city transferability.** Land use, infrastructural development and socio-economic characteristics vary widely across cities, altering both covariates and underlying mobility mechanisms. While domain adaptation and meta-learning techniques narrow these gaps, they often fail under unseen conditions.
- **Data scarcity and privacy.** High-fidelity mobility data, such as mobile device records, are privacy sensitive and tightly regulated.
- **Interpretability and uncertainty.** Decision-makers need models that provide explanations and uncertainty estimates. Many deep models remain opaque and poorly calibrated under distribution shifts. Techniques such as Bayesian deep learning [33] can enhance the interpretability of the results.
- **Scalability and efficiency.** Large spatio-temporal models demand significant computational resources, making them challenging to deploy.

## 3 A Lightweight Foundation Model for Cross-City Learning

As described in section 2, deep learning models offer powerful predictions for a specific city, on which they are trained, but struggle to make predictions in cities they have not seen before. At the same time, urban science, using concepts from complex systems theory, has uncovered a range of statistical regularities and general principles behind the apparent complexity of urban mobility patterns, which are remarkably robust across different geographies, cultures and levels of development [12, 23, 34]. It is therefore promising to integrate concepts from urban science into deep learning models (in particular, graph-based models) to enable flow predictions that are fine-grained, globally transferable and lightweight. For instance, a promising way to increase the generalizability is to replace the location-specific semantics of points-of-interests (e.g., ‘bakery’) with its rank in the hierarchy of central places [24] which, in principle, applies to any city, see Fig. 1c. This allows for the development of lightweight foundation models (LFM) pre-trained on diverse mobility patterns from multiple cities that are able to perform generalizable predictions even in data-scarce environments. The development of such a foundation model comprises the following main steps, illustrated in Fig. 2:

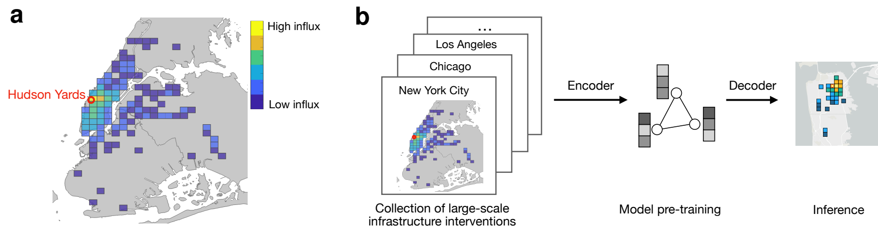


**Fig. 2 Towards a lightweight foundation model of urban mobility. a,** Mobility flows to Hudson Yards, a major urban redevelopment in New York City that opened in 2019. Mobility flows are derived from anonymized mobile phone data from 2023 (provided by citydata.ai). Brightness reflects flow intensity from a given cell to Hudson Yards. **b,** Constructing a benchmark dataset of infrastructure interventions across U.S. cities, enabling pretraining and transfer learning for a graph-based foundation model.

**Step 1 - Benchmark dataset development.** Assemble a multi-year, multi-city dataset capturing mobility responses to diverse infrastructure interventions. The resulting benchmark will enable robust foundation model pretraining and support out-of-sample generalization across cities.

**Step 2 - Foundation model architecture.** A graph-based deep learning architecture to capture functional relations between different POIs is particularly promising [27]. Shocks to urban mobility patterns like large-scale urban redevelopments can be encoded as node embeddings in terms of their ranks in the urban hierarchy, while changes in the transportation system can be encoded as edge embeddings in terms of travel times.

**Step 3 - Case study applications.** The developed framework is applied to real world cases, enabling direct insight into how planned or recent infrastructure changes may reshape urban mobility. A series of retrospective studies allows to validate the results with more traditional data (e.g., travel surveys, transit data, vehicle counts).

## 4 Vision for a Foundation Model as Part of a Larger Digital Twin

While the LFM captures universal patterns and enables zero-shot forecasting, its practical relevance for planners is enhanced when embedded within an operational urban framework. In this section, we propose, from a medium-term perspective, its potential integration into an UDT as the first layer of a multilayer cognitive architecture. The LFM functions as a universal prior, offering each city cross-city mobility representations that are then specialised through fine-tuning on small, but high-resolution local data. (e.g., from volunteers). Deep learning models inside each UDT transform these data into a local cognition layer. Federated learning (FL) mechanisms then combine local cognitions into a shared representation without centralising raw data. This creates a two-level architecture: a cross-city lightweight foundation model (LFM) that provides a general representation of movement dynamics and its operational extension through cognitive UDTs, which federate it into a unified Foundation Model for Urban Mobility (FM-UM). The following subsections describe this transition from representation to cognition and toward forms of collective urban intelligence, enabled by recent advances in Artificial Intelligence (AI) applied to the urban domain.

## 4.1 From Representation to Urban Cognition

Many urban models have historically been descriptive: maps, gravity formulations, and deterministic simulations portraying the city as a quasi-static, observable entity. An **Urban Digital Twin (UDT)**[1] instead represents a living, data-driven system able to sense, simulate, and interpret urban processes through heterogeneous data streams, multimodal simulations, and interactive visualization environments. Contemporary research recognizes that cities must be conceived as complex, self-organizing, and co-evolving systems whose dynamics emerge from innumerable interactions among agents, infrastructures, and environmental layers [36]. Within this framework, the UDT becomes an adaptive analytical environment that not only mirrors urban dynamics but learns from them. Most current DTs still reproduce infrastructure and flows without capturing the nonlinear feedbacks driving urban evolution [37]. Bridging this gap may be achieved by integrating deep-learning architectures capable of revealing hidden relationships and emergent patterns within the city. Through this 'cognitive layer', the digital twin evolves from a representational model into an intelligent system that interprets and adapts in real time, redefining the city as a system able to learn from its own dynamics [36].

## 4.2 From Data Fragmentation to Cognitive Integration

UDTs have historically faced fragmented and descriptive representations: data silos, semantic inconsistencies, and missing analytical pipelines limit their ability to generate actionable knowledge. DL provides a promising cognitive layer that synthesizes heterogeneous data streams from IoT networks, satellite imagery, environmental sensors into unified and adaptive latent representations. By integrating spatial cognition algorithms and cyber-physical systems, the DTs transitions from a passive simulation platform to an active learning architecture that co-evolves with the city it represents. In this setting, DP-models fuse multimodal signals into embeddings resilient to data sparsity and contextual variability, enabling real-time forecasting, continuous recalibration, and decision-support across multiple domains. A concrete illustration of this

[1] Definition of UDT: An Urban Digital Twin is a high-fidelity computational model of the city, a "replica" that represents its functional units explicitly across 3D space and time, leveraging high-resolution data, real-time streams, and bi-directional interaction between computational models and the physical city [35].

process can be observed in the *Blue City Project* of Lausanne, Switzerland, where DP architectures integrate mobility, environmental, and infrastructural data under uneven sensor coverage. Autoencoder-based techniques reconstruct incomplete traffic and energy-flow records, fill gaps in time-series data, and generate coherent latent representations despite fragmentation and semantic inconsistency, demonstrating how deep learning stabilizes disparate inputs and provides the cognitive layer required for **adaptive urban intelligence** [38]. This transformation allows the UDT to capture nonlinear feedbacks, learn from its own data, and convert raw information into synthetic knowledge, thus enabling the city to function as an adaptive, self-learning organism [39].

### 4.3 Mobility as the Cognitive Interface of the City

If deep learning constitutes the cognitive substrate of the UDT, urban mobility represents the domain in which such intelligence manifests most explicitly. Mobility flows embody the physical manifestation of the continuous interaction between human behaviour and the built environment: a high-frequency signal through which the city perceives shifts, tensions, patterns, and anomalies. Operational UDTs already integrate empirical routes and mobility simulations with environmental models, creating unified contexts in which learned mobility representations act as feedback signals for broader urban cognition. When predictive mobility models are embedded at the core of the twin, they serve as sensors of cognition, continuously testing hypotheses about how the city responds to interventions, disruptions, or routines. Indeed, the literature on UDTs and transport analytics shows a rapid convergence between AI, traffic forecasting, and infrastructure management, consolidating mobility's role as the city's cognitive interface [40] [41] [42]. By capturing the **feedback loop** between citizens and the built environment, mobility enables the digital twin to adapt and self-correct, embedding human decisions into computational reasoning. The result is a reflective rather than descriptive model-one in which human agency remains integral to sensing, learning, and evolution [42].

### 4.4 Towards a Foundation Model for Urban Mobility

A natural next step in this trajectory is a **Foundation Model for Urban Mobility (FM-UM)** capable of aggregating learning derived from multiple cities. The previously introduced LFM operates as a universal prior, offering an inter-city representation of mobility patterns that each UDT can specialize through fine-tuning. This distinction between the lightweight inter-city prior (LFM) and the federated model integrating local UDT learning (FM-UM) clarifies the respective roles of representation and collective adaptation. If mobility functions as a city's cognitive interface, linking DTs through FL enables a distributed intelligence that transcends local boundaries. Each UDT trains predictive models on local mobility-flow data while sharing only learned representations within a secure, decentralized network, exchanging knowledge rather than data and allowing cities to co-evolve through reciprocal learning. This approach aligns with research on FL-based DTs, where federated training mitigates privacy, latency, and bandwidth constraints in critical domains such as healthcare, transportation, IoT, and 5G, allowing DTs to operate as local intelligence substrates while contributing updates to a shared aggregation layer without exposing raw data [43]. Parallel work envisions foundation models suited to the nonlinear and generative nature of urban mobility[40], with large-scale architectures capable of unifying multi-source data, modes, and decision layers. Empirical advances on Large Flow Models already demonstrate the ability to represent flows, complete missing observations, estimate unmonitored areas, and forecast system evolution across mobility, energy, waste, materials, and biodiversity [38]. A complementary line of work explores how foundation models can be embedded within DT architectures to enhance predictive analytics, adaptive learning, and real-time system configuration across sectors [44]. Within this broader FL–DT landscape, the FM-UM emerges as a mobility-centered instance of a federated DT stack: local mobility twins learn from their own data while contributing to a shared latent space capturing transferable movement patterns [43].Integrated into UDTs, such models improve predictive accuracy, adaptive learning, and multi-scale planning. Federated exchange of flow embeddings enables cities to learn from one another without transferring sensitive information, reducing privacy and governance risks. In this configuration, the FM-UM therefore acts not as a global twin but as a decentralized *meta-cognitive layer* connecting many cities.The LFM provides the initial cross-city representational layer, while UDTs supply adaptive, context-specific cognition, generating local models that evolve from the shared prior. **Figure 3** illustrates this progression, from representational UDTs to local predictive systems via DP-based models, and from local cognitions to a cross-city FM-UM through FL. Despite its promise, the FM-UM raises operational challenges - data governance, privacy, computational demands, and data quality - echoing concerns already identified for FL–DT integration in *smart-city governance*, including the need to embed human oversight within automated decision loops [43]. Such considerations call for careful architectural design and coordinated frameworks across cities, which are discussed in the following sections.

### 4.5 From Predictive Systems to Cognitive Governance

As predictive capacity compounds, the locus of innovation shifts from computation to institutions. The promise lies in **anticipatory decision-making**: simulating policy effects on transits, low-emission zones, pricing, or street design before implementation, and recalibrating policies through observed outcomes. Yet evidence shows that sustainable impact depends less on technical sophistication than on transparent data governance, citizen consent, and traceable decision processes [45]. Leverag-

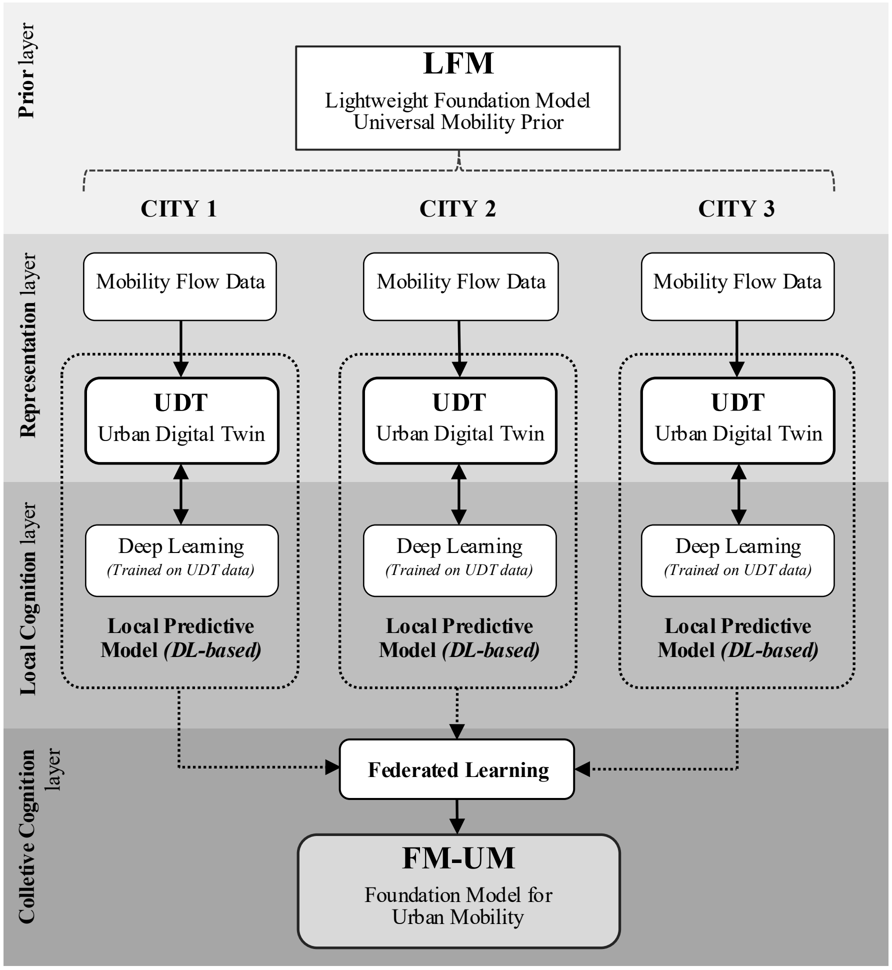


**Fig. 3 Multi-city cognitive pipeline for constructing a Foundation Model for Urban Mobility (FM-UM).** A lightweight foundation model provides an initial cross-city prior for urban mobility (Prior Layer), which is distributed to each city prior to local training. In each city, mobility flow data feed into the local Urban Digital Twin (UDT), where they preserve their representational function (Representation Layer). Within the UDT, Deep Learning (trained on UDT-specific data and refined by the universal prior) produces a local predictive model, which constitutes the Local Cognition Layer. These models are then incorporated into a Federated Learning process, in which only learned parameters and latent representations are shared. The aggregation of these elements gives rise to the Collective Cognition Layer, which defines the modelling structure underpinning the FM-UM.

ing the integrative intelligence of foundation models, cities can embed prediction directly into deliberative and policy workflows, transforming data into collective intelligence and anticipatory governance.Technically, UDTs integrate heterogeneous real-time data from IoT devices, sensors, satellite imagery, and municipal databases to simulate multi-domain interactions - mobility, energy, environment - thus supporting evidence-based planning [38]. Research on foundation-model integration within DT architectures confirms this potential: generalist models trained on multimodal datasets enhance decision-support, improve temporal responsiveness, and enable real-time system configuration [44]. These capacities extend prediction into actionable governance, allowing institutions to test and revise interventions through continuous human–AI feedback. While machine-learning architectures improve precision, genuine cognition requires two-way coupling between physical and virtual layers, with learning shaped by human feedback and environmental change. The potential of digital twins lies in supporting inclusive and accountable decisions rather than obscuring them behind algorithmic opacity [46]. Governance frameworks must balance innovation with protection, coupling foundation-model intelligence with explainability, auditability, and civic participation [45].

### 4.6 The Foundation of an Intelligent Urban Future: Systemic Vulnerabilities and Human–AI Coevolution

Interpreted as the city's analytical mind and its collective memory, UDTs augmented with foundation models promise a shift from reactive monitoring to continuous urban learning. Yet cities remain complex, multi-scale systems requiring rigorous abstraction, verification, and epistemic discipline [35]. Progress hinges on socio-technical integration that is interdisciplinary, contextually grounded, and procedurally embedded, while persistent bottlenecks in semantic interoperability, data quality, scalability, visualization, and governance remain unresolved [47] [48]. Moreover, these new AI approaches will inevitably participate in **human–AI coevolution**: where policy shapes data, data shape models, and models recursively influence policy, risking bias amplification without institutional safeguards [49].

However, the same mechanisms: progressive incorporation of foundation models (FMs) into UDTs, expands analytical, predictive, and configurational powerthat enable cognitive capacity, also introduce a multilayered spectrum of vulnerabilities. Alongside concerns regarding transparency, safety and the computational intensity of large-scale AI systems [44], research on FL–enabled DTs highlights persistent tensions in privacy, data quality, cross-domain integration, and information-flow regulation across distributed sensing architectures[43]. Within broader human–AI coevolution, these issues compound the epistemic and institutional risks inherent in transitioning toward cognitive urban infrastructures. These vulnerabilities, emerging from the integration of FMs, FL and UDTs ccan be clustered across three domains:

- **Design and Implementation Challenges of FM-DT Integration.** Embedding FMs as core representational components of DTs introduces substantial computational demands, architectural opacity, and privacy and safety risks associated with large multimodal datasets [44] [41]. Further limitations concern insufficient real-time processing, complex domain adaptation to urban infrastructures, and verification difficulties, including risks of off-label usage stemming from FM generalisation [43].

- **Operational Challenges in FL-DT Architectures.** Federated DT ecosystems face risks of privacy leakage through shared model updates, while heterogeneous and noisy data introduce systemic inaccuracies [43]. Additional obstacles include cross-domain integration across misaligned ontologies and spatial–temporal resolutions, information overload from expanding sensing layers, and limited explainability in federated deep-learning systems.
- **Epistemic and Socio-institutional Challenges in Cognitive Urban Systems.** The coevolution of human behaviour, institutional structures, and AI-mediated urban representations[49][42] introduces risks of recursive biases, reinforcement of socio-spatial inequalities, and normative tensions as optimisation shifts toward responsibility. Increasing epistemic opacity undermines public trust, while effective oversight requires reflexive governance capable of continuous auditing and recalibration.

Extending beyond mobility to the intertwined systems of energy, environment, infrastructure, and urban society, the future points toward a form of integrated urban intelligence in which computational models accumulate and transfer knowledge across contexts, while institutions safeguard transparency, equity, and accountability. Ultimately, the central challenge is not the perfection of prediction but the cultivation of a deeper awareness of human-AI coevolution enabling action on the city both as it exists today and as it is continuously becoming.

## References


[1] Michael Batty. *The New Science of Cities*. MIT Press, 2013.
[2] Luís MA Bettencourt. Introduction to urban science: evidence and theory of cities as complex systems. 2021.
[3] Alain Bertaud. *Order without Design How Markets Shape Cities*. MIT Press, 2024.
[4] Eduardo Alejandro Martinez-Cesena, Pierluigi Mancarella, Mamadou Ndiaye, and Markus Schläpfer. Using mobile phone data for electricity infrastructure planning, 2015. Preprint at https://arxiv.org/abs/1504.03899.
[5] Yuehan Yang, Zhexuan Tan, and Markus Schläpfer. Assessing the space-use efficiency of French cities by coupling city volumes with mobile data traffic. *Sustainable Cities and Society*, 124:106292, 2025.
[6] Jiazu Zhou, Seanglidet Yean, Tianyu Dong, Bu Sung Lee, and Markus Schläpfer. Estimating electric vehicle charging demand and its impact on the power grid using mobile phone data. In *IEEE 27th International Conference on Intelligent Transportation Systems (ITSC)*, pages 3076–3081, 2024.
[7] Juan de Dios Ortúzar and Luis G. Willumsen. *Modelling Transport*. Wiley, Hoboken, 4th edition, 2011.
[8] Hugo Barbosa, Marc Barthelemy, Gourab Ghoshal, Charlotte R. James, Maxime Lenormand, Thomas Louail, Ronaldo Menezes, José J. Ramasco,

Filippo Simini, and Marcello Tomasini. Human mobility: Models and applications. *Physics Reports*, 734:1–74, 2018.

[9] Massimiliano Luca, Gianni Barlacchi, Bruno Lepri, and Luca Pappalardo. A survey on deep learning for human mobility. *ACM Comput. Surv.*, 55(1), 2021.

[10] Luca Pappalardo, Ed Manley, Vedran Sekara, and Laura Alessandretti. Future directions in human mobility science. *Nature Computational Science*, 3(7):588–600, 2023.

[11] Can Rong, Jingtao Ding, and Yong Li. An interdisciplinary survey on origin-destination flows modeling: Theory and techniques. *ACM Comput. Surv.*, 57(1), 2024.

[12] Marc Barthelemy. The role of parsimonious models in addressing mobility challenges. *npj Sustainable Mobility and Transport*, 1(1):11, 2024.

[13] George Kingsley Zipf. The P1P2/D hypothesis: on the intercity movement of persons. *American Sociological Review*, 11:677–686, 1946.

[14] Alan Geoffrey Wilson. *Urban and regional models in geography and planning*. Wiley, New Jersey, 1974.

[15] Sven Erlander and Neil F. Stewart. *The Gravity Model in Transportation Analysis: Theory and Extensions*. CRC Press, Boca Raton, 1990.

[16] Alan Geoffrey Wilson. A family of spatial interaction models, and associated developments. *Environment and Planning A*, 3:1–32, 1971.

[17] Filippo Simini, Marta C. González, Amos Maritan, and Albert-László Barabási. A universal model for mobility and migration patterns. *Nature*, 484(7392):96–100, 2012.

[18] Marta C. González, César A. Hidalgo, and Albert-László Barabási. Understanding individual human mobility patterns. *Nature*, 453(7196):779–782, 2008.

[19] Chaoming Song, Tal Koren, Pu Wang, and Albert-László Barabási. Modelling the scaling properties of human mobility. *Nature Physics*, 6(10):818–823, 2010.

[20] Shan Jiang, Yingxiang Yang, Siddharth Gupta, Daniele Veneziano, Shounak Athavale, and Marta C. González. The timegeo modeling framework for urban mobility without travel surveys. *Proceedings of the National Academy of Sciences*, 113(37):E5370–E5378, 2016.

[21] Maxime Lenormand, Aleix Bassolas, and José J Ramasco. Systematic comparison of trip distribution laws and models. *Journal of Transport Geography*, 51:158–169, 2016.

[22] A. Paolo Masucci, Joan Serras, Anders Johansson, and Michael Batty. Gravity versus radiation models: On the importance of scale and heterogeneity in commuting flows. *Phys. Rev. E*, 88:022812, Aug 2013.

[23] Markus Schläpfer, Lei Dong, Kevin O'Keeffe, Paolo Santi, Michael Szell, Hadrien Salat, Samuel Anklesaria, Mohammad Vazifeh, Carlo Ratti, and Geoffrey B. West. The universal visitation law of human mobility. *Nature*, 593:522–527, 2021.

[24] Brian JL Berry. Cities as systems within systems of cities. *Papers in regional science*, 13(1):147–163, 1964.

[25] Paul Waddell. Urbansim: Modeling urban development for land use, transportation, and environmental planning. *Journal of the American Planning Association*, 68(3):297–314, 2002.

[26] Cristian Poliziani, A. Zachary Needell, Haitam Laarabi, Rashid Waraich, Annika Todd-Blick, K. Sydny Fujita, Nazanin Rezaei, D. Juan Caicedo, Carlos Guirado, C. Anna Spurlock, and Tom Wenzel. Simulating impacts from transit service enhancements in the san francisco bay area. *Transportation Research Record*, 2679(4):399–413, 2025.

[27] Hongrong Yang and Markus Schläpfer. UrbanPulse: A cross-city deep learning framework for ultra-fine-grained population transfer prediction, 2025. Preprint at https://arxiv.org/abs/2507.17924.

[28] Ons Aouedi, Van An Le, Kandaraj Piamrat, and Yusheng Ji. Deep learning on network traffic prediction: Recent advances, analysis, and future directions. *ACM Computing Surveys*, 57(6):1–37, 2025.

[29] Yaguang Li, Rose Yu, Cyrus Shahabi, and Yan Liu. Diffusion convolutional recurrent neural network: Data-driven traffic forecasting. *arXiv preprint arXiv:1707.01926*, 2017.

[30] Shengnan Guo, Youfang Lin, Ning Feng, Chao Song, and Huaiyu Wan. Attention based spatial-temporal graph convolutional networks for traffic flow forecasting. In *Proceedings of the AAAI conference on artificial intelligence*, volume 33, pages 922–929, 2019.

[31] Zonghan Wu, Shirui Pan, Guodong Long, Jing Jiang, and Chengqi Zhang. Graph wavenet for deep spatial-temporal graph modeling. In *International Joint Conference on Artificial Intelligence 2019*, pages 1907–1913. Association for the Advancement of Artificial Intelligence (AAAI), 2019.

[32] Hao Miao, Jiaxing Shen, Jiannong Cao, Jiangnan Xia, and Senzhang Wang. Mba-stnet: Bayes-enhanced discriminative multi-task learning for flow prediction. *IEEE Transactions on Knowledge and Data Engineering*, 35(7):7164–7177, 2022.

[33] Aike Steentoft, Bu-Sung Lee, and Markus Schläpfer. Quantifying the uncertainty of mobility flow predictions using Gaussian processes. *Transportation*, 51:2301–2322, 2024.

[34] Laura Alessandretti, Ulf Aslak, and Sune Lehmann. The scales of human mobility. *Nature*, 587(7834):402–407, 2020.

[35] Luís M. A. Bettencourt. Recent achievements and conceptual challenges for urban digital twins. *Nature Computational Science*, 4(3):150–153, 2024.

[36] G. Caldarelli, E. Arcaute, M. Barthelemy, M. Batty, C. Gershenson, D. Helbing, S. Mancuso, Y. Moreno, J. J. Ramasco, C. Rozenblat, A. Sánchez, and J. L. Fernández-Villacañas. The role of complexity for digital twins of cities. *Nature Computational Science*, 3(5):374–381, 2023.

[37] Gerhard Schrotter and Christian Hürzeler. The digital twin of the city of zurich for urban planning. *PFG –Journal of Photogrammetry, Remote Sensing and Geoinformation Science*, 88(1):99–112, 2020.

[38] J. Huang, S. E. Bibri, and P. Keel. Generative spatial artificial intelligence for sustainable smart cities: A pioneering large flow model for urban digital twin. *Environmental Science and Ecotechnology*, 24:100526, 2025.

[39] E. Nica, G. H. Popescu, M. Poliak, T. Kliestik, and O.-M. Sabie. Digital twin simulation tools, spatial cognition algorithms, and multi-sensor fusion technology in sustainable urban governance networks. *Mathematics*, 11:1981, 2023.

[40] Z. Zhou, Z. Gu, X. Qu, P. Liu, Z. Liu, and W. Yu. Urban mobility foundation model: A literature review and hierarchical perspective. *Transportation Research Part E*, 192:103795, 2024.

[41] Xinke Huang, Yuqin Jiang, and Ali Mostafavi. The emergence of urban heat traps and human mobility in 20 US cities. *npj Urban Sustainability*, 4(1):6, 2024.

[42] M. Batty. Digital twins in city planning. *Nature Computational Science*, 4:192–199, 2024.

[43] S. P. Ramu, P. Boopalan, Q.-V. Pham, P. K. R. Maddikunta, T. Huynh-The, M. Alazab, T. T. Nguyen, and T. R. Gadekallu. Federated learning enabled digital twins for smart cities: Concepts, recent advances, and future directions. *Sustainable Cities and Society*, 79:103663, 2022.

[44] A. Trantas and P. Pileggi. Digital twin and foundation models: A new frontier. In *Proceedings of the 16th International Conference on Agents and Artificial Intelligence (ICAART 2024)*, volume 3, pages 988–994. SCITEPRESS, 2024.

[45] M. Hooli. Digital twins and urban planning: Designing smarter, more inclusive cities. *European Journal of Computer Science and Information Technology*, 13(44):104–112, 2025.

[46] R. Roversi. Urban digital twins as socio-technical infrastructures for city regeneration and decarbonization. *IOP Conference Series: Earth and Environmental Science*, 1402:012065, 2024.

[47] S. Azadi, D. Kasraian, P. Nourian, and P. van Wesemael. What have urban digital twins contributed to urban planning and decision making? from a systematic literature review toward a socio-technical research and development agenda. *Smart Cities*, 8:32, 2025.

[48] C. Weil, S. E. Bibri, R. Longchamp, F. Golay, and A. Alahi. Urban digital twin challenges: A systematic review and perspectives for sustainable smart cities. *Sustainable Cities and Society*, 99:104862, 2023.

[49] D. Pedreschi, L. Pappalardo, E. Ferragina, R. Baeza-Yates, A.-L. Barabási, F. Dignum, V. Dignum, T. Eliassi-Rad, J. Kertész, A. Knott, Y. Ioannidis, P. Lukowicz, A. Passarella, A. S. Pentland, J. Shawe-Taylor, and A. Vespignani. Human–AI coevolution. *Artificial Intelligence*, 339:104244, 2025.