

AI in Smart Cities: A Systematic Literature Review and Open Traffic Datasets Investigation

Raafat Talib Hashim

Department of Computer Technical Engineering, Imam Al-Kadhum College, Baghdad, Dhi Qar, Iraq.

Abstract: This work represents a systematic review of the applications of Artificial Intelligence (AI) in environments defined as smart cities, with an emphasis on urban mobility and the open traffic datasets. The conducted review examines the existing studies published between 2019 and 2025, outlining the applications of ML, DL, RL and hybrid AI ranging from critical mobility-related tasks including traffic forecasting, route planning, demand estimation, anomaly detection and intelligent traffic signal control. The characteristics, availability, interoperability and features/limitations of open traffic datasets in smart cities that can be used to support AI-based transport systems are also examined by the study. Furthermore, we also performed an empirical exploration of the open traffic dataset in New York City (NYC), to showcase practical traffic patterns, time-based congestion behavior and segment-level traffic distribution. It concludes that AI-based mobility systems can increase the efficiency, sustainability and decision-making in transportation planning when they are backed by high quality open datasets. This is very important, but there are still serious challenges that remain in terms of standardization, privacy, interoperability and ethical governance by October 2023. The paper ends with unresolved research challenges and outlooks towards resilient, sustainable and data-driven urban mobility in smart cities.

Keywords: Smart Cities; Artificial Intelligence; Urban Mobility; Open Traffic Datasets; Intelligent Transportation Systems.

INTRODUCTION

Today, more than half the world's population is urban and more than three quarters of global energy consumption and GHG emissions originate from cities (UN, 2023). Cities have become a competitive marketplace for population, resources and investments, where city-level and metropolitan-level decisions significantly determine how cities look like and how they work (Ghareh Mohammadi *et al.*, 2021). The urgency of urban sustainability seems particularly clear in developing countries where rapid urban growth outstrips government capacity to plan; and also in developed countries faced with planning bottlenecks that make it impossible to build cities for people, polluting wastelands creating health hazards, low density sprawl need). These inequities may harm the effectiveness and cause safety-critical incidents for users. The concept of smart cities, which includes applications for service delivery, urban planning and development, efficiency in public administration, safety and quality of life has recently been pushed to the forefront. The rise of the Internet of Things (IoT) has facilitated both real-time collection and analysis of a city's data supporting urban informatics and participatory sensing. In Europe, the concept of Urban Mobility Data as formulated by the European Commission encompasses public-transport and traffic and travel-related datasets (comprising bus/train schedules, transport vehicle positions and cycling resources) and is already wide-spread. Urban mobility— transport for

people and goods in towns and cities—has therefore become an issue at the heart of.

End-to-end urban mobility describes the continuous and comprehensive movement of people and goods throughout city space where no distinct origin, destination, or space can be assigned. Urban Mobility Data refers to the corresponding information dimension and has been rapidly and widely adopted across multiple systems and services.

AI (artificial intelligence) denotes the capability of machines and computer systems to mimic human cognitive processes. Machine learning (ML) involves computer algorithms automatically learning from observations rather than being explicitly programmed. Within ML, deep learning (DL) describes algorithms à la Convolutional Neural Network (CNN), Recurrent Neural Network (RNN), and Trajectory-Aware Graph Convolutional Network (TAGCN), operating with large amounts of data across multiple layers (). Reinforcement learning (RL) concerns sequential decision-making agents interacting with an environment and making observations, aiming to maximize cumulative rewards (). RL has received increasing attention with the emergence of deep RL, combining classical RL ideas with DL techniques ().

A keyword-based strategy was devised to probe the scientific literature for trends and insights concerning the use of AI techniques, models, and

methods to address urban-mobility challenges in smart cities. Urban mobility forecasting, route planning, incident detection, demand estimation, travel-time prediction, and intelligent signal-control strategies were identified as key challenges in connected and automated traffic systems, and Smart Cities were recognized as vital socio-technical systems for tackling sustainable development (). The embedding of AI within Smart Cities holds significant promise for addressing large and complex problems linked to human daily-life issues (). A recent conceptual framework delineates Smart City Systems; references frames, definitions, and component-mapping clarifications for Smart City Systems, mobility, and traffic are available ().

Rationale for AI in Urban Mobility

Mobility is one of the largest contributors to greenhouse gas emissions, and by 2030, the number of vehicles on the world's roads is expected to increase by over 1 billion. In addition, nearly 70% of the world's population is projected to live in urban areas by 2050, with more than 40 megacities larger than 10 million inhabitants (Ghareh Mohammadi *et al.*, 2021). Therefore, urban mobility is a critical problem that must be addressed to optimize traffic flows, reduce operational costs, enable eco-friendly travel, and provide seamless transportation in smart cities (Yan & Li, 2023). AI is essential for analyzing large-scale traffic data to address complex urban mobility issues. This systematic review addresses common urban mobility problems and the available solutions using AI. Urban mobility refers to the movement of people in metropolitan areas through various means, such as public transport and private vehicles. Urban Mobility as a Service (MaaS) is a transport paradigm allowing users to plan, book, and pay for mobility services via a single application. Urban mobility vehicles include shared bicycles, cars, motorcycles, taxis and ridesharing, and public transportation, such as buses and subways.

Cities implement various systems to improve transportation efficiency. Traffic State Estimation (TSE) uses high-resolution, large-volume real-time data to estimate the entire traffic state, including travel time and traffic flow based on partial measurements. The TSE can be formulated as both a spatial and temporal problem. Traffic Forecasting (TF) involves prediction of future traffic given the current and past history time series of traffic measurements, MFD data and other context information. It could be performed

for short-term prediction (1 h ahead) and long-term prediction (> 1 h ahead). Traffic Control and Optimization Systems limit congestion by use of traffic signal control interventions with sequences of red and green light—containing multi-intersection scenarios. A node-based spatial-temporal graph signal prediction algorithm is proposed to model the temporal evolution of system states (i.e., vehicular mobility and traffic statuses) and infer the spatial dependency between nodes. Route Planning (RP) allows to plan trips from start point to target point on various vehicles by different criteria with use of geographical and past months traffic data. Rerouting systems provide alternative routes according to current traffic conditions. Incident Detection (ID) detects the presence of incidents and obstacles along the road, including traffic flow reduction due to accidents or road blockages. Anomaly Detection (AD) identifies particular traveling tracks or base stations to warn drivers to avoid congestion in advance. Urban mobility cannot be addressed using a single approach, and multiple aspects are highly intertwined; therefore, a multi-modality approach is proposed to combine hybrid models. (Liu *et al.*, 2023)

Data-driven Artificial Intelligence (AI) has the capacity to transform urban environments into smart cities that provide innovative and integrated sustainability solutions. Rapid urbanization has significantly impacted the development of large smart cities that accommodate a growing number of vehicles to ensure an adequate urban environment and urban mobility. Data-driven AI methods benefit urban mobility solutions such as traffic forecasting, traffic assignment, route planning, demand forecasting and distribution, vehicle-sharing system optimization, traffic state estimation, signal control, and incident detection. A Transform-based Traffic Speed Prediction (TTSP) framework predicts urban traffic speed and fine-tunes the model based on multi-scale spatial-temporal Graph Convolutional Networks (GCNs). Deep Learning Fusion Traffic State Estimation (DLFTSE) focuses on fusing calibrated GPS trajectory data and traffic flow data collected from a road network's flow-calibrated edge. A Density-aware Taxi Demand Prediction (DTP) model leverages density-variation characteristics of taxi demand through stage 1 clustering for different time periods. Urban Mobility-as-a-Service (UMaaS) platform planning is also addressed based on multimodal demand estimation. By analyzing multi-scale graph structure and multi-

source data, a Spatial-Temporal Multiscale Network (STMN) captures urban mobility evolution patterns considering local structure variations, temporal transformations and long-range dependencies. (Chintala, 2024)

Research Questions and Objectives

Development of cities Urbanization has led to the growth of populations and economic activities in cities over centuries. Such rapid urban development has contributed to high concentration of free passage, cars and cargoes in city areas with connected challenges on urban mobility. Smart cities are utopian visions to support the sustainable development of their growth by applying IoT, internet and telecommunication infrastructures, along with AI methods. Smart cities determine how resources are used in city and also quality of life with less pollution through monitoring of the city parameters which is happening at actual time. With such capabilities a range of mobility-related information can be gathered, like vehicular flow, demand or environmental data. The above complexity offers research avenues within Transportation Engineering (TE), applying a variety of operational research methods in combination with modelling methods (Ghany Ismaeel *et al.*, 2024).

Scope and Definitions

Urban mobility has recently become very significant in the design of smart cities as well as in their strategies. Urban mobility, smart cities, smart mobility and intelligent urban mobility are keywords frequently used interchangeably however there exist some differences. Urban mobility concerns the movement of people or goods in urban environment and is an important part for the metropolis functioning.(Ghareh Mohammadi *et al.*, 2021) Smart cities also combine technical, social and governance dimensions to confront urbanization problems and encourage sustainability. Smart mobility constitutes one of smart city components (Khurshudov, 2024). Intelligent urban mobility places an emphasis on technologies that exploit urban mobility data to enhance awareness of mobility/state and facilitate decision making (Elias Bibri *et al.*, 2024).

Rail transport, shipping, aviation, and unmanned aerial vehicles (drones) are excluded. Major algorithms (notably machine learning, deep learning, and reinforcement learning) are included; however, a high-level overview is provided. Hybrid and multimodal approaches that involve

multiple methods and/or datasets are considered as separate categories.

Systematic Review Rationale and Protocol Overview

The systematic literature review protocol is summarized as follows. A preliminary search using the keywords (artificial intelligence OR AI) AND (smart city OR smart mobility OR intelligent urban) AND (open dataset OR open data OR open traffic) has been executed to estimate the extent of relevant research and related trends. This survey covers scientific studies published between 2019 and 2025 addressing the application of AI techniques to urban mobility problems within a smart-city context based on the use of open traffic datasets that are freely available to any user. (Mach *et al.* 2025)

The systematic review has been structured around four broad steps:

Step (1) define clear, specific, and relevant review protocol;

Step (2) and select paper

(Step 3) extracting information related to Objectives,

Open Datasets used in AI image processing as well as the type of AI techniques used and architectures make use of; with final step being a synthesis step: compare obtained results from previously selected papers by problem category. The search is implemented with the following set of keywords in Scopus and Google Scholar databases between May and July 2024. Publications are identified, screened, and selected in two subsequent stages. Full article text is retrieved from the publications selected, the key information according to the protocol is extracted, and the information summarized in a structured format. (Mobayen & Sadeghi, 2022)

The protocol has undergone several iterations to arrive at the revised version presented above, motivating changes in the protocol or protocols. The modes of dissemination include the completion of a comprehensive systematic review of AI applications in the context of smart cities and a secondary survey of the alignment of the topic with the open traffic dataset theme (Elias Bibri *et al.*, 2024).

SMART CITIES: CONCEPTS AND ARCHITECTURES

To address a range of challenges posed by urban growth (e.g., rapid population growth, congested streets, and inefficient traffic management), cities

worldwide are moving towards becoming smart, defined as a city that “embraces technology and innovation to improve the lives of its citizens” (Ghareh Mohammadi *et al.*, 2021). The foundations of a smart city and a smart city in the field of mobility rely on the development of smart systems for a variety of applications (e.g., healthcare, energy, waste management, and the environment). In smart mobility, intelligent decision-making systems equipped with urban data offer real-time information and the capability to optimize the management of transportation systems (e.g., public transit, taxis, ridesharing, and traffic). Information relevant to making timely decisions and responding to certain destinations is voluminous, high-dimensional, and often incomplete, motivating the development of smart systems that capture, analyze, and leverage such data. Universal smart city decision-making frameworks accommodate these challenges by addressing data capture, analysis, and decision-making.

At a conceptual level, smart city reference architectures provide a thematic description of smart cities and highlight analysis subjects (Figure 1). The urban environment monitoring analysis subject indicates which elements can be monitored for large-scale urban analytics. The smart city analytical framework provides a finer decomposition in terms of urban data, analytics, and services rather than elements. For smart mobility, strategic, tactical, and operational decision supports are considered to enlarge the analysis scope to cover traffic forecasting. Extended urban dynamics candidate topics enlarge the monitored elements relevant to smart mobility. Smart city frameworks, compared to smart mobility frameworks, do not explicit data, sensing and interoperability layers, and governing aspects (Elias Bibri *et al.*, 2024).

Smart city architecture, per se, can be revisited from the viewpoint of mobility, where data-aware smart mobility reference architecture exhibits the difference-cycle and inter-city pair. Urban mobility data-aware differential characteristics help building proper models, capture raw data to preserve essential knowledge for propagation across urban environments and enhance urban fitness by economic-volume, spatial-distribution, and temporal-regularity descriptors. Inter-city traffic dynamic and structure analysis subject identifies the macrostructural aspects of mobility and provides insights into the design of multimodal deep neural networks fitted for Urban

Mobility task. Two smart city costs—urban mobility smart-city cost and dimension externality—reflect the global characteristics of a clean intelligence with an analytical model. (Mirzaie *et al.*, 2025)

Foundational Definitions and Evolution

Artificial intelligence (AI) has emerged at the forefront of transforming physical urban mobility, a primary concern in smart cities. Despite its incorporation into diverse and innovative applications, the full potential of AI remains unrealized due to significant technological, pedagogical, and institutional restrictions. Owing to the rapid development of urban passenger travel and vehicle ownership, the public concerns about inefficient transportation systems and increasingly urgent demand for sustainable development (Zhuge *et al.*, 2020), a considerable effort has been devoted to smart mobility, an essential aspect of smart cities and a most relevant research area (Ghareh Mohammadi *et al.*, 2021). In this respect, smart mobility minimizes journey times and the level of congestion as well as pollution effects on the environment through intelligent and efficient resource planning for movable objects based on knowledge while providing guidance to individuals in terms of specific routing choices, meanwhile enabling proactive traffic management. It encompasses the timely and precise communication of mobile users’ origin-destination (OD) demand and trip purpose, traffic-flow and vehicle-speed state forecasting, real-time incident detection and monitoring, route planning, and efficient signal-control optimization under a sustainability paradigm (Elias Bibri *et al.*, 2024).

Integrating open datasets, such as those from New York City (NYC), London, Uber Movement, and others, can further accelerate the advancement of these applications by facilitating the benchmarking of AI models and enabling scenario analysis under varying conditions. Due to the global rise of AI and growing interest in urban smart mobility, a systematic review covering the period 2019–2025 was conducted to provide an overview of this research landscape and make open-traffics-data datasets openly available to the community. (Anthony Jnr, 2025)

Reference Architectures for Smart Mobility

The development of smart cities requires reference architectures and interoperability methods for efficient communication of data across various mobility technologies. A smart city collects, compiles, analyses, and distributes data from

vehicles, buildings and pedestrians to improve mobility-related services and quality of life. Typical components involved in smart mobility include Intelligent Transport Systems, the Internet of Things, Connected Intelligent Transport Systems, Intelligent Urban Logistics and Mobility as a Service (Ghareh Mohammadi *et al.*, 2021).

A multilevel typology of smart mobility systems distinguishes personal transport, shared mobility and public transport. Mobility as a Service integrates public transport, shared mobility and personal transport, allowing travel between locations using diverse mobility services. A data-centric vision of smart mobility encompasses three levels and nine components. Smart mobility is positioned between smart road and smart government with these components: smart traffic management, traffic analytics, a traffic information publishing platform and intelligent road user. (Narayanan & Antoniou, 2023)

Data and information flow in smart mobility systems begins with large amounts of diverse traffic data from vehicles, road infrastructures and the environment. Data are processed to generate traffic information and knowledge, which are then published to road users. Considerations such as different standards and protocols needed to complete the interoperability function are necessary when multiple producers support a smart mobility system.

Data, Sensing, and Interoperability Layers

Smart cities undergo various transitions depending upon their historical background, the nature of the stakeholders involved, the nature of the enabling technologies, and societal pressures and needs. Rijkswaterstaat's perspective defines smart vehicles, smart transport systems, smart traffic, smart (mobility) hubs, and smart energy, extending these definitions into a general smart governance framework.

The various smart systems in cities, including smart mobility systems transfer data to each other and to users over multiple devices (e.g., PCs, tablets, smartphones). Smart systems and the information exchanged between them can be mapped to technical domains in the Internet of Things (IoT) systems. Smart vehicles can be equipped with many types of sensors (acoustic, electromagnetic, optical, pressure, temperature, etc.) to collect data at a low frequency (typically in minutes or seconds). Data from these sensors are preprocessed in heterogeneous environments in the

vehicular fog computing (VFC) layer or in the vehicular cloud processing (VCP) layer before being exfiltrated from the vehicles. Transport systems gather traffic data for route planning and guidance. Traffic, parking and reservation data is filtered based on the granularity and topology of the data exchanges and corresponding data processing options. Traffic flows, occupancy rates, sacrificial times (time-to-reach-destination), will then be analyzed in VFC or VCP before being integrated with other data sources if necessary. Road side pollution, noise and weather data are also collected by traffic management systems and transferred to provide environmental awareness to smart vehicle users. The (Ghareh Mohammadi *et al.*, 2021)

Governance, Ethics, and Stakeholder Alignment

Improving urban mobility through advanced technology raises governance, ethics, and stakeholder alignment concerns (Ziosi *et al.*, 2022). Citizens may seek to protect their privacy, autonomy, and access to resources when government authorities deploy new ICT systems, especially if they promise substantial operational efficiencies. Government digitally mediated direct interactions with citizens potentially heighten concerns where local authorities can 'track' in detail how citizens use public services and interact with the government itself.

Technologies that support the efficient operation of a smart city-wide service are now acknowledged as having the potential to accelerate existing inequality in access to and the opportunities offered by public services (Elias Bibri *et al.*, 2024). This potential for adverse societal effects has grown as the uptake of big data and AI systems in smart cities has driven an associated increase in algorithmically activated governance with limited human oversight. Concerns about algorithmic bias have thus arisen, stressing the necessity to ensure the responsible use of AI techniques that actively maintain stalking the support of sustainability and resource-efficient decision-making concerning service provision to wider smart city contexts.

AI TECHNIQUES IN URBAN MOBILITY

Artificial Intelligence (AI) is increasingly applied to various areas in smart cities, and the urban mobility sector is no exception. Artificial Intelligence can support effective policy related to COVID-19 through the analysis of mobility

behaviour (Ghareh Mohammadi *et al.*, 2021) by providing near real-time information to city decision-makers on traffic congestion and public transit mobility patterns (Albertengo & Hassan, 2018). AI based systems can enhance incident detection, support decision making in traffic management, assist with traffic flow prediction in an Urban Mobility as a Service (UMaaS) environment and improve routes for logistics transport. The role of AI for predicting and monitoring urban mobility, along with the analysis of open traffic datasets, reflects the acceleration in research and projects exploring AI based solutions to addressing growing problems in urban mobility as the number of vehicles and inhabitants grows and as the urban landscape continually changes.

Machine Learning forms a category of AI algorithms relying on inference and intelligence from data. Applications for Machine Learning in urban mobility encompass traffic forecasting (speed, volume, flow), ride-sharing demand forecasting, generating and analysing travel behaviour scenarios (home-work/school locations, mode of transport), predicting flows and times in public transport. Deep Learning constitutes a sub-category of Machine Learning relying on deep neural network architectures to predict outcomes from data. Applications for Deep Learning in urban mobility include sensing and perception and training neural networks to extract spatiotemporal features from traffic data. Reinforcement Learning defines interactions with a dynamic system through trial and error, in which an agent interacts within an environment and adapts to current states using an objective reward signal to learn. Applications for Reinforcement Learning in urban mobility concentration on optimisation and control, including adaptive signal control and urban mobility scheduling. Various studies demonstrate hybrid and multi-modal approaches that either integrate different AI techniques or combine diverse sources of data, such as satellite imagery, social networks, the physical road graph, and historical information.

Machine Learning for Traffic Forecasting and Demand Modeling

The transport sector plays an important role in a city's overall economy, carrying around 65% of the value of goods transported. Accurate forecasting of traffic state is thus crucial for smart cities and intelligent transport systems (ITS). Several machine learning models are adopted for short-term traffic forecasting and demand modelling in urban mobility applications, and they

can be generally classified into (i) supervised learning, (ii) unsupervised learning and (iii) reinforcement learning algorithms. Supervised learning techniques include regression-based approaches, Bayesian networks, support vector machines and decision trees. Unsupervised learning is associated with clustering algorithms like k-means and self-organising maps. Reinforcement learning enables dynamic control of traffic flow by adjusting traffic signal timing in response to real-time traffic data. (Albertengo & Hassan, 2018)

Deep Learning for Sensing, Perception, and Spatiotemporal Analysis

The prediction of multiple traffic flow times series is a significant concern for smart cities (Ghareh Mohammadi *et al.*, 2021). Highway traffic flows are often monitored based on ordinary traffic conditions, leaving the prediction of traffic volumes during nonstandard conditions a critical task (Ghany Ismaeel *et al.*, 2024). Moreover, spatiotemporal characteristics in the same dataset require an appropriate model for transportation systems that can capture spatial and temporal correlation features. Consequently, it is important to understand the characteristics of traffic-affected regions in smart cities by advances in precision forecasting. Consequently, as aforementioned the requirements above are to be justified with spatiotemporal information from coarse and fine granular levels in combination with a reasonable operation method suited to particular analytic requirements.

Recognizing the significance of urban mobility, there has been increased attention focused on the traversal characteristics of urban traffic and mobility flows on public bicycles. Riding duration analysis captures important factors influencing user demand preferences, such as the selection of starting and ending stations. The graphical representation of public bike flows and the characterization of their spatial distribution can help devise revenue-enhancement strategies for bike-sharing systems. The cycling-decision Chinese urban mobility dataset consists of approximately 77,000 bike-share records over a five-day period. It includes comprehensive bike-share records, individual trip ID, riding times, picking-up and dropping-off stations (longitude and latitude), and urban land-use information. Such an open cafe dataset comprises a large amount of accurately labelled user daily-cycling behaviour data of Chinese citizens and is suitable for exploring the temporal and spatial patterns,

factors impacting people's choice of riding, and the origin–destination demand prediction of national bike riding. The outdoor mobility dataset includes transcripts from 957 trips based on user records from a bike-sharing service. It contains real-time trajectory recordings of cyclists, which can be used to build bike-moving trajectory data models and extract diverse trajectory features.

Reinforcement Learning for Control and Optimization

Traffic signal control aims to minimize congestion and enhance road system efficiency through the effective management of traffic density and flow. Traffic density, which measures the number of vehicles within a certain space, reflects the degree of congestion. Traffic flow indicates the number of vehicles passing a specified section of road at a certain time. Thus, adjusting the timing of red, green, and orange lights is a vital method to regulate both the traffic flow and density (Guo *et al.*, 2019).

Modeling the traffic-control problem as a Markov decision process (MDP) enables the combination of dynamic programming with approximation. Such an MDP includes a six-dimensional observation vector containing the number of vehicles in each lane, the impending green-phase time remaining for every lane, and the queue length of each lane upon the green-phase turn-switching time. An improved DDPG (deep deterministic policy gradient) actor-critic algorithm with prioritized experience replay is additionally developed to enhance real-time performance. This algorithm is capable of selecting one phase from eight candidates. Urban traffic signal-control benchmarking further indicates satisfactory training efficiency and good adaptability to cross sections equipped with considerably different numbers of lanes (Korecki, 2023).

Hybrid and Multimodal AI Approaches

Urban mobility is a fundamental pillar of smart cities and relies heavily on advanced intelligence techniques to analyze diverse data sources. Recent years have witnessed growing research interest in hybrid and multimodal approaches that realize added value through model or data source combinations, but systematic overviews of these methods remain scarce. A curated collection of hybrid and multimodal models across diverse urban mobility subfields helps identify broad trends and generalize principles.

Hybrid models apply different artificial intelligence methods to tackle a problem, whereas multimodal models combine multiple sources of information to perform a task. Although data integration techniques may vary, scenarios requiring decision making based on deep learning–extracted features illustrate the value of both paradigms. Hybrid and multimodal studies occur across diverse topics: urban mobility forecasting, routing, demand estimation, spatiotemporal trajectory prediction, incident detection, and intelligent traffic-signal control.

Evaluation Metrics and Benchmarking Practices

Measuring the performance of urban mobility applications can rely on standard machine learning (ML) metrics to assess the accuracy of predicted values against ground truth data. However, these metrics alone can become limited for tasks that also require the determination of optimal routes or control signals, thus making evaluation difficult. Spotting these supplementary indicators would require explicit definition of proper targets in modeling scenario specifications, as well as breaking down the evaluation process into clearly-stated machine learning questions to foster more reproducible research and benchmark datasets targeting temporal and spatial mobility problems.

Another literature study on urban traffic-based AI city challenges provides benchmarking consideration rationale. Answering a variety of public, large-scale, complex urban problems ranging from wide-area vehicle detection, re-identification, multi-target tracking, to abnormal traffic detection, city-form cellphone-user-level spatiotemporal smooth prediction and multi-scale spatio-temporal city-wide crowd pre-looking and crowd motion controller design, these context-demand different, generalizable and quantitative metric-sets and comparative experiments are involved, which inspire possible benchmarking consideration for study scope (Naphade *et al.*, 2020).

In smart mobility, benchmarks are not solely based on data sets but also on public platforms that details algorithm resources and metric standards, to evoke competition and agreement. Various mobility Questions can be raised into interaction-Point or edge-level structure. Representative performance metrics are provided for large-scale metro passenger operation prediction and crowd flow generation, coming with openly-released codes and models (Jiang *et al.*, 2019).

OPEN TRAFFIC DATA IN SMART CITIES

In the context of smart cities, traffic data are often used to analyze road usage patterns, accessibility, congestion, bottlenecks, and route planning. Traffic data come in many forms (the type of data that refer to a certain phenomenon). Smart cities leverage traffic data from vehicles, public transportation, and road sensors to monitor the traffic conditions in real time. Open traffic datasets share traffic data from various cities for public reuse. Such datasets enable additional projects in urban computing.

Definitions of open datasets vary in the literature. The Open Knowledge Foundation defines an open dataset as (1) available in some form of digitally, (2) licensed in a manner that allows its reuse, (3) provided in appropriate format, and (4) published without any restriction to individual (Reza Honarvar & Sami, 2018). Ghareh Mohammadi points out that a dataset is released as an open dataset if the individual can access, use, modify, and share the dataset freely (Ghareh Mohammadi *et al.*, 2021). For open traffic datasets in specific, they should share traffic measurements, such as travel time, connected time, volume, speed, etc. Open datasets should also grant others free access under a certain license.

Characteristics of open traffic datasets differ from cities and transport systems. The difference also leads to variation in datasets for projects. Considerable work requires open traffic datasets to (1) measure the most common traffic conditions, (2) reveal diverse traffic conditions overcrowding or underused, and (3) capture traffic volume alterations even when all roads are busy (Albertengo & Hassan, 2018). Such alteration can indicate the functioning of the transport system, like traveling in an area with the highest density, intersections position, or choosing a shortcut route. Public datasets are important for urban and transportation development. Many urban and mobility studies demand data from extremely numerous cities to ensure smoothness and generalize the phenomenon.

Characteristics of Open Traffic Datasets

Open traffic datasets are essential for developing traffic flow prediction models and intelligent transportation systems (Albertengo & Hassan, 2018). with different sources such as sensor data, traffic counts, and travel times. They can help in understanding traffic patterns, validating models and optimizing systems. Waterborne traffic-flow

mobility/routing and kinematic characteristics knowledge is an important basis in the safety fast supervision, and its leading to more effective data-driven decision-making process.

Open traffic datasets may be treated with clustering techniques such as k-means which cluster road segments together according to properties associated with traffic behavior. Random forest algorithms can be used to predict unknown features which may be beneficial for traffic flow analysis. The clustering outcomes indicate a homogeneity among street segments, which is consistent not necessarily with usual road taxonomies but also with observed traffic practices. The future work could be focused on better traffic load profiles prediction such as when the profile can be predicted in a deep learning manner. Such profiles are invaluable to support navigation and traffic management in particular, if no floating car data is present. The calculated features from this analysis are also meaningful for applications in traffic prediction and urban planning.

Notable Open Datasets and Accessibility

Urban traffic dataset is widely defined as the measurement sets that consider various aspects such as traffic counts, travel times, trajectories and speeds (Ghareh Mohammadi *et al.*, 2021). Source datasets S6 are an invaluable resource and have been of the utmost importance when creating records in areas such as traffic prediction demand characterization, route planning incident detection [Li *et al.*, 2023b]. The open traffic data is published with an open license, which can be freely used and attributed by the owner of the opened data. We define a dataset as meeting this criteria if it was first published after January 1, 2019; has accrued at least 100 citations according to Google Scholar; and remains accessible as of May 31, 2023 at the same (unmodified) URL.

Open traffic datasets have unique features such as different temporal and spatial granularity, geographic coverage ranging from city blocks to the globe, data aggregation (e.g., individual sensors/zones/vehicle types) granularity, and licensing restrictions (Yazgan *et al.*, 2024). Many datasets qualify, such as those within the New York City, London, and Uber Movement initiatives.

The traffic datasets play critical roles in real-time traffic prediction, which can make different model architectures have accurate evaluations and

practical applications. The well-known benchmarks reveal the device architecture and data set to a greater extent, while being based upon synthetic or scale-down data. To the contrary, applied papers that used actual datasets often reported little information for any datasets they used/had access to, and thereby made it unnecessary or impossible for others to replicate their method on those datasets.

Data Quality, Preprocessing, and Privacy Considerations

Cities have always been organizers of everyday life, conditioned by the evolutive design of policies in science, technology, economy, and communication. At present, the increased availability of urban data, together with the latest advances in Artificial Intelligence (AI) and the pervasiveness of the Internet, raise high expectations on the ability to monitor, model, and design cities. Cities therefore do promise to improve those urban systems that allow cities to remain organized. Recent developments in AI have significantly increased its capabilities and widespread applications in various fields. Mobility systems have, in particular, benefitted from AI because of their impact on accomplishing the Sustainable development Goal 11 by reducing pollution, energy consumption, and a pressure on infrastructures. A review of the Artificial Intelligence trends in Smart Cities with respect to Urban Mobility and Data sets is provided by compiling and analyzing over 130 papers in the domain for a pre-defined, systematic literature review in smart cities and open data defined within the period of 2019–2025 (Reza Honarvar & Sami, 2018).

Data Fusion and Interoperability Challenges

Open data has been recognized as a foundation for smart cities. Nonetheless, the lack of standardization has created a disparity in data formats, data models, terminology, and data quality. Integrating heterogeneous data of distinct types, collected at various times and places by a multitude of monitoring devices, raises issues in estimation and prediction that must be addressed. Urban traffic monitoring systems generally run independently, frequently producing surface data that are incorporated into larger-scale models. Challenges such as traffic data interoperation between different sensing systems, data accessibility improvement and standardization of extraction and fusion methods are considered to be important (A. Kashinath *et al.*, 2021). Methods of real-time data fusion, that combine and connect

information from various (stationary or mobile) sensing environments broaden the scope of mathematical models to be applied, support real-time traffic-state estimation over a network of urban roads and allow prescriptive optimization in an effective manner (Ghareh Mohammadi *et al.*, 2021). Therefore, multi-source urban traffic data (including weather condition, road topology information, traffic reports and Incident Related-Impact) usually are represented by different styles and have different forms of relevance – It is inconvenient to integrate them together. For better traffic-sensing, the spatiotemporal data should not be simply merged due to different mechanism in monitoring flow. To solve these problems, we propose a two-stage multi-source fusion models based on spatial-temporal temporal graphs to align information data from multiple sensors by considering the geographical and temporal properties simultaneously.

SURVEY OF RECENT WORKS (2019–2025)

Smart cities seek to find solutions to the problems that they face—such as mobility, energy consumption and building use, to name just a few—by leveraging the most advanced information and communication technology within their physical structures. Given that the urban population is estimated to represent 68% of the world's population by 2050, predicting and estimating traffic in cities are still important topics to ensure efficient operation of smart-city transportation systems. (Elias Bibri *et al.*, 2024). Subsequently, real-time traffic incident detection, surveillance-camera monitoring, and video-anomaly monitoring improve mobility performance—essential for efficiency, sustainability, and safety.

Following the systematic literature review protocol, an analysis of 28 open-traffic dataset studies determines their contribution to modelling. Prominent studies catalog urban-mobility forecasting with AI traffic-prediction research concentrated on specific cities and datasets and limited-use open traffic data obtained from intelligent-transportation-system deployment.

Urban mobility becomes the subject of mathematical modelling—contributing to the modelling of innovative urban environments and information dissemination support undertaken in the study of three datasets enabling a variety of improvement avenues still unidentified in related literature. Observability and robustness of models

arise in intelligent connecting-vehicle systems nationwide while commonly available open-traffic-training, test, and continuous-stream datasets sought in large-scale freely accessible data prepared the working of multi-cities are essential.

Urban Mobility Forecasting and Traffic State Estimation

Urban mobility forecasting and traffic state estimation involve predicting traffic flow using various models and data analysis techniques. Approaches include Kalman filtering for dynamic traffic volume prediction, neural networks and non-parametric regression for short-term forecasts, and Bayesian network models. A number of researches have concluded that temporal data aggregation and adaptive control systems can better the results accuracy. More advanced techniques are composed by ITS models and data processing for real time traffic management and forecasting (Albertengo & Hassan, 2018).

Studies on urban traffic forecasting are mainly aimed at short-term control and prediction of future states for short time horizons, typically in the range from 5 min to a couple of hours. Statistical and deep learning models predict traffic at network junctions (road segments, sensors, intersections) in a geographical area of interest; reports can be spatially summarised for wider perspective views or urban planning. Performance measures such as root mean square error (RMSE) and mean absolute error (MAE) characterize accuracy.

Route Planning, Mobility as a Service, and Demand Responsive Systems

Route planning, Mobility-as-a-Service (MaaS), and demand-responsive systems concentrate on offering efficient solutions for shared and autonomous vehicles. Autonomous mobility-on-demand enables convenient transportation options, while MaaS integrates diverse travel methods into single routes. Cooperative vehicle-based routing supports collaborative dispatching and idle repositioning, minimizing overall costs. In demand-responsive systems, categorical passenger profiles enhance matching and accommodation, while predictive pickup and drop-off locations optimize the routing sequence. Models considering same-vehicle restrictions boost ride-sourcing efficiency by reducing distance and time deviations. Crowd-shared same-vehicle requests mitigate relocation and waiting time, benefitting new passengers or non-time-sensitive users, while prioritizing single-rider ride-hailing requests when

demand surges. Most investigations employ simulation-based platforms like SUMO and MATSim, with publicly accessible datasets facilitating contribution to the field. (S. El-Wakeel *et al.*, 2018) (Zardini *et al.*, 2021)

Incident Detection, Anomaly Monitoring, and Cloud-Edge Collaboration

Smart cities aim to improve daily life and public services while enhancing the security of citizens and drivers. One means to achieve these objectives is through an AI-based video surveillance system combined with a traffic-monitoring and anomaly-detection framework. The architecture encompasses three tiers: a ground level of smart lampposts equipped with cameras and sensors for real-time accident detection; a territorial level that integrates data from lampposts with external sources to coordinate alerts; and a training level oriented toward refining the system's accuracy. A preliminary experimental approach for anomaly detection within the framework employs the Faster R-CNN algorithm (Atzori *et al.*, 2021).

Warning of traffic anomalies is an important issue in intelligent urban transportation. Many challenges persist due to high complexity of the scene and lack of high quality data. One simple and effective detection pipeline consists of a pre-processing stage, a dynamic tracking module, and a post-processing stage. Video stabilisation, background modelling and vehicle detection are furthermore applied in the pre-processing stage to generate potential anomalies. The tracker locates the temporal start of anomalies by analyzing apparent motion patterns and spatiotemporal status. In post processing, we refine the temporal anomaly window. This model was top in traffic-anomaly detection on the NVIDIA AI CITY 2021 leaderboard (Zhao *et al.*, 2021).

Edge computing and cloud-edge collaboration enrich the urban-video-surveillance services by offloading simple analysis tasks to edge. This solution not only relieves the load of calculation on the cloud, but also reduces network transmission costs and enhance overall service response time. The integration of edge and cloud systems can result in timeiy, accurate, trustworthy and effcient traffic monitoring and incident detection periods.

Sustainability, Emissions, and Intelligent Signal Control

The mobility sector consumes the most important part of the energy use and greenhouse gas

emissions (GGEs) in comparison to other sectors, therefore ensuring sustainability is considered as a main challenge. Urbanisation has expanded in the last-50 years increasing congestion, vehicle hours of operation and non-renewable fuel consumption. Simulation results with the fine-grained analysis indicate that energy consumption and GGEs rise by approximately 8–10% even in the fictitious case where road length is enlarged and vehicle running amount grows by a factor of 10%. The traffic management is further supported by ITS and AI as well as open traffic data sets that mitigate energy consumption and GHG emission to meet the sustainable goal in Smart Cities. ITS deploys a wide range of sophisticated sensors and devices, such as Mobile, Highway, Vehicle, GPS, Controller, and others, for real-time monitoring and collecting comprehensive traffic information. Traffic flow predictor can be established by applying AI algorithms to sensed traffic data, which in turn can support the use of traffic congestion monitoring system to distribute green time effectively, thus contributing to energy and emission reduction (Culita *et al.*, 2020).

The rapid advance in the world is a transitional period. Artificial intelligence (AI) and open traffic datasets are increasingly popular and have been used frequently in intelligent transportation systems, green transportation, energy efficiency, and emission reduction (Rinchi *et al.*, 2024). AI de facto becomes an important enabling technology and an integral part in intelligent transportation systems. Sustainable urban mobility can be maximised through the appropriate use of artificial intelligence applications, algorithms, models, and open traffic datasets.

Case Studies and Open Dataset Usage Examples (with Real Datasets)

This section showcases examples of real-world applications of AI techniques and relevant open traffic datasets. Two case studies are presented in each example, demonstrating the use of open datasets for AI-based urban mobility solutions.

The first case study examines a study that forecasted the future state of the road network based on historical traffic speed data, focusing on a city in China. The model outperformed several comparative algorithms and out-of-the-box configurations when trained using an open dataset. The second case study presents a project that correlated urban music trends with different environments to recommend music based on users' locations in the metropolitan area of Madrid. The

approach employed data from the Open Data Portal of the City of Madrid, including traffic counts, air quality, and taxi data.

The first case study involved a study that monitored and modeled the spatiotemporal evolution of epidemic growth across different regions of COVID-19 for traffic flow prediction in a city in China. The model was trained with various datasets, including the Open Traffic Data Set from New York City. The second case study described a study on a sustainable intelligent transportation system in a Baltic city. The project investigated charts and outdoor temperature in relation to traffic data from an open data portal.

RESEARCH GAPS AND FUTURE DIRECTIONS

As discussed in the previous sections, the application of Artificial Intelligence (AI) in smart cities remains a fertile area of research. Within this field, the intelligent management of urban mobility is of primary importance. Although some insightful systematic reviews have already presented reports on the integration of AI techniques into smart urban mobility (Shaygan *et al.*, 2023) and on data-driven traffic modelling approaches, several literature gaps persist. Such analyses do not address how AI techniques make use of open traffic datasets, nor do they clarify the nature or characteristics of such datasets. These gaps hinder a comprehensive understanding of the adoption of AI-assisted traffic solutions leveraging publicly available open traffic data in smart cities.

Methodological Gaps in Systematic Synthesis

Chia *et al.* (2024) point out that past systematic reviews of smart cities with AI have not consistently reported specified search strategies, inclusion/exclusion criteria, and comparative dimensions. Although standard-setting reports highlight the importance of knowledge synthesis, rigorous synthesis of AI trends in smart cities remains scarce. The search uncovered a study that reviews AI techniques in smart cities from various angles, but does not systematically compare AI trends across subfields (Elias Bibri *et al.*, 2024) ; (Lim *et al.*, 2018).

Data Gaps, Standards, and Open Dataset Ecosystems

Smart cities draw attention due to their potential for economic, social, and environmental transformations. Smart mobility is a key challenge within this paradigm, with rapid changes in mobility services and urban transport systems that

have major implications for traffic congestion, air quality, safety, social equity, and energy consumption. Therefore, many transportation agencies and municipalities have adopted smart mobility strategies. The role of AI technologies in the urban mobility of smart cities is evident; AI techniques are applied to improve the performance of diverse mobility services. While AI is one of the most widely used technologies in smart cities, there are few studies on how AI has been adopted and traffic data have become essential datasets to enrich the domain of urban mobility. Recent work contributions related to understanding smart cities is analyzed in the present study with a focus on its mobility subdomains and the involvement of AI.

The systematic review complies with the preferred reporting items for systematic reviews and meta-analyses (PRISMA) guidelines, and is based on a strict protocol presented in section 4. The examined time period is 2019–2025 and the research is targeted on urban mobility and open traffic datasets. The chosen search strategy yielded a sample of 745 papers. Four criteria of inclusion are implemented to obtain 27 contributions of interest. The study is carried in the smart urban mobility subdomain, seeking to understand how open traffic data can be leveraged by AI techniques. Studying datasets is vital to understand the evolution of methods and section 3 includes a list of dataset properties, some relevant datasets, and open problems. A dataset is considered open if it is available free of charge and does not constrain its access or reuses, following the principles of the open data movement.

The effects of smart city, smart mobility, and AI technologies on urban mobility cannot be understood without examining traffic data sources. Smart traffic sensors are being deployed by cities and multinationals to monitor traffic flows in real time; nevertheless, the collection mechanisms, coverage, confidentiality, and usability of these traffic data sources vary widely according to sensor location, deployment period, and privacy concerns. Terrain with widespread persistent mobile data accounts for numerous demand-responsive mobility services, including app-based taxis and ride-sharing, as well as traffic-flow prediction, queuing modeling, and route planning. Epic global open datasets have emerged in relation to these services yet are scarcely exploited in the governance and management of urban mobility. Global open datasets, such as real-time aerial images and satellite maps, retain undeniable value when extensive region-specific datasets remain

unprovisioned. Multinational firms obtain access through extensive transports conducted at the global or regional levels, forwarding global coarse-grained delivery or ride-hailing datasets that inevitably shrink to micro urban datasets during the proposition of local governance concerning urban mobility. Nevertheless, the complete traffic characteristics required for either mobility service governance in smart cities or multimodal service association among different datasets remain obscure in the absence of an extensive overview.

Where the provision of large-scale datasets cannot keep pace with deployment time, design time, compilation time, or fragmentation, the utilization of worldwide datasets becomes an obligatory resort of smart city traffic modeling compilation. Mobility services in large-scale highly dynamic traffic demands, e.g. DataSET, continue to lack comprehensive global urban traffic datasets. Furthermore, datasets are typically proprietary and are either shared at subtasks or stripped of the spatial dimension to guarantee corporate privacy. Open datasets exploiting fresh readily accessible global wide tools for micro urban traffic or mobile service interactions can encourage traffic demand forecasting of smart-mobility services that exploit wide-range traces and enrich interim knowledge on the lining of independent datasets. Existing open datasets for traffic simulation, flow prediction, routing, taxi, ride-sharing, and so on, either remain proprietary, are scarce compared to accelerating developments, or cannot maintain a full spectrum across urban mobility. The connection and complementary supply across different mobility-service datasets form an important avenue towards urban-traffic understanding at spatial macro or time aggregation.

The datasets involved in the scope of understanding traffic patterns in smart-mobility urban-agenda circulation are classified to comply with the ensuing analysis on open urban-traffic datasets and the summarized stage of smart-city traffic processing tasks on urban mobility.

Ethical, Legal, and Social Implications

Smart city initiatives created in pursuit of an improved quality of life in urban areas raise numerous concerns regarding ethical, social, and legal implications that ought to be carefully addressed. While smart cities promise improvement of the living conditions in urban environments and enhancement of public services, several potential threats may arise in the

implementation of such initiatives. The digitization of daily processes in smart cities may risk excessive tracking and surveillance of citizens, thereby threatening their privacy and autonomy (Ziosi *et al.*, 2022). The development of smart streetlights or traffic cameras, for instance, may strengthen police capacity, leading to increased persecution of some urban populations and thus aggravating existing social inequalities. Technology could exacerbate social divides, for example with the launch of online government services that are easy to use for the digitally literate and yet shut out those who are not.

More generally, the adaptation to technology-embedded smart ecosystems has far-reaching social implications beyond specific projects; consider, for example the transformation of state-citizen relations with respect to the provision of community services and enjoyment of rights enshrined by laws. All these elements contribute to the formation of urban spaces and generate new demands for universities that are challenged to recognize more seriously the ethical, social and legal dimensions of their activities. Even the challenge of their role in smart cities initiatives itself stimulates the development of new concepts for working with such city-university relationships as a whole..

Sustainability and Resilience Considerations

Smart cities in sustainable development Smart cities play a crucial role in smart governance and sustainable urban development through the encouragement of the use of cutting-edge technologies and data-driven approaches to increase efficiency, improve standards of living, conserve resources and mitigate environmental challenges while securing durability over a long period (Elias Bibri *et al.*, 2024). And such cities apply AI and Internet of Things (AIoT) solutions in renewable energy, energy conservation, water resource management, waste management, sustainable transportation, pollution treatment and climate change adaptation as well as smart infrastructure. The transformation of human settlements into sustainable smart cities should be driven by strategic thinking and focusing on environmental and technological aspects. It is necessary to create assessment tools using performance indicators that provide the means to assess the effectiveness of sustainable smart city projects.

Resilience in urban systems entails the characterization of human mobility patterns

including clustering of daily activities, together with aids in mobile phone-derived data sources like Wi-Fi and social media to get richer insight on behavior (Bellini *et al.*, 2021). Methods such as k-means and hierarchical clustering can be useful for finding mobility patterns, and in turn to make decisions. The architecture of smart city is designed to provision service APIs and linked-data publication. Resilience is considered using approaches to manage risk and design, notably resilience led urban planning. Systemic and scenario perspectives inspire systems for sustainable and resilient cities.

Proposed Roadmap for Advancing AI in Urban Mobility

Modern times are a time of rapid development in urbanization, globalism and Internet. These driving factors create enormous volume of data, and when properly processed combined with efficient analytics, can render a significant improvement in urban management and quality of life. In this context, smart city frameworks has emerged to offer a new ways for urban management in compliance with sustainable development goals. Urban mobility is the cornerstone of all smart city domains and directly influences citizens' quality of life, traffic situation and environmental pollution. Artificial intelligence (AI) has been recognized as one of the key contributors to enhancing urban mobility. While there are a lot of studies which focus on AI dimensions in smart cities urban traffic mobility and open traffic datasets only get limited concerns.

In light of this, we present in this review a comprehensive exploration of AI trends for urban mobility to shed the lights on not only prominent AI application approaches but also relevant open traffic datasets. Critical components are discussed, such as AI modalities, open traffic datasets, urban mobility challenges and performance metrics for baseline generation and benchmarking. Moreover, possible direction of urban mobility AI are open traffic dataset release, multi-discipline joint-research pushing forward the integration of data sets, reinforcement learning for new urban mobility scene s and hybrid algorithm as well as soft sensor technique.

ANALYSIS OF THE NYC OPEN TRAFFIC DATASET

For illustration purposes, a public dataset on urban transportation from New York City was analyzed to further contextualize the theoretical standpoint. 8 The [original data] is over 10 years of data, but a

representative week (20/10/19- 20/10/25) was selected for the analysis conducted in this paper. It is an acceptable choice which could give a reasonable period of observation and reflecting the average traffic pattern in a large city. The subset consisted of 42,890 records after filtering and cleaning, and combining these with the full data set we were confident this would be adequate to detect common temporal and spatial patterns.

Dataset Characteristics

The attributes of the traffic data are in various dimensions features referring to time and space. The observed time is expressed as a combination of year, month, date and hour/minute/s to derive one concatenated timestamp for potential time series analysis. Other fields such as road segment ID, direction of travel, and intersecting street names contribute to a broader understanding of how congestion is distributed throughout the roadway system. The Vol column, giving the actual volume read at each reading is the "output factor" used for analysis.

Data Preparation

Before some statistical treatments (described later) were carried out, the dataset was slightly preprocessed, but this is enough to noise-reduction procedure when it comes to exploring for free. The individual time features were transformed into

a single, consolidated datetime, and then several secondary features derived: day of the week (DOW), calendar date (CD) and hour of the day. Rows where the timestamp was NaN, or not well formed were dropped. Because in practical systems, sensing results from real world are not perfect, such as the unexpected large drops (suching no actions) for sensing values due to resting sensors, slight anomalous were inevitable. However, the majority of data was well-structured and no aggressive outlier elimination was necessary.

Statistical Overview

The daily statistics for the week we chose are presented in Table 1. A clear trend occurs when looking at total daily volumes: the week progresses and traffic gradually increases, with Wednesday Oct 22 hitting a high of 1,126,364 vehicles in aggregate across all sensors. Certain days include extremely low totals (e.g., 7,572), which are undersired as they represent a day with not much sensor activity than a collapse in traffic. On average throughout the week, however, the mean daily volume stands at 139.36, telling us something about the rhythm of a standard working week in city. These scores are contained in the CSV files being output within preprocessing parties.

Table 1: Daily Traffic Statistics for NYC (19–25 October 2020)

date	Daily_Average_Volume	Daily_Total_Volume	Daily_Max_Volume	Daily_Min_Volume
10/19/2020	133.7428653	937270	865	0
10/20/2020	153.5452778	1105526	1243	0
10/21/2020	155.855	1122156	1296	0
10/22/2020	156.4394444	1126364	1340	0
10/23/2020	142.2087614	996599	800	0
10/24/2020	132.7417697	955608	1015	0
10/25/2020	100.96	7572	553	0

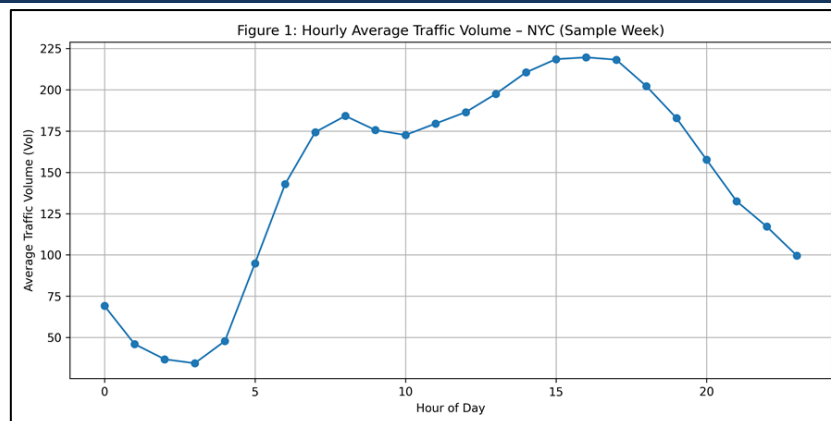
Temporal Traffic Patterns

Hourly Behavior (Figure 1)

When volumes are summed by hour, the conventional pattern of an urban day emerges. Traffic spikes quickly toward 7:00 AM, hits the local morning maximum of flow between 8:00 and 9:00, and eases off before jumping up again during the evening return period of roughly 4:00 to 6:00

PM. The lowest number of commits unsurprisingly occurred during the night, in the early morning hours between 1:00 and 5:00am. Such observations are not limited to New York and they capture general mobility behavior and can be used as a pivot while looking at any AI-based prediction or control model.

Average_Daily_Volume	Max_Daily_Total_Volume	Min_Daily_Total_Volume
139.3561598	1126364	7572



Day-of-Week Behavior (Figure 2)

Another consistent finding emerges from the day-level averages. The constraint from Monday to Friday is still distinctly larger than on the weekend, however Thursday is the busiest day. The drop on Saturday and Sunday reflects less commute, less business transports in general and a shift in daily rhythm within the city. Such a weekly profile is useful for any systems which require to predict demand and adjust control mechanisms accordingly

Segment-Level Observations

Asarathi, Veeranna, Nay Md congestion is rarely distributed on all the roads uniformly we analyzed

the dataset at road segment level. The results indicate that a small number of segments account for a significant amount of the total volume. The highest average vehicles/24 h for segments 140064, 140062 and 258712 was higher than all the other segment in the dataset. These are the most likely candidates for representing major arteries or intersections, and would be a natural choice in any urban planning/smart-city sensing strategy. Discovery of such hotspots is critical for developing effective signal control systems, evaluating the effects of re-routing policies and finding optimal locations for sensors.

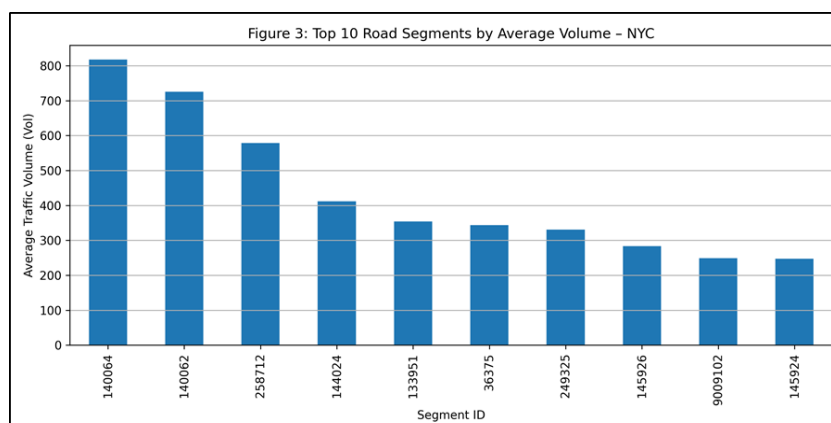
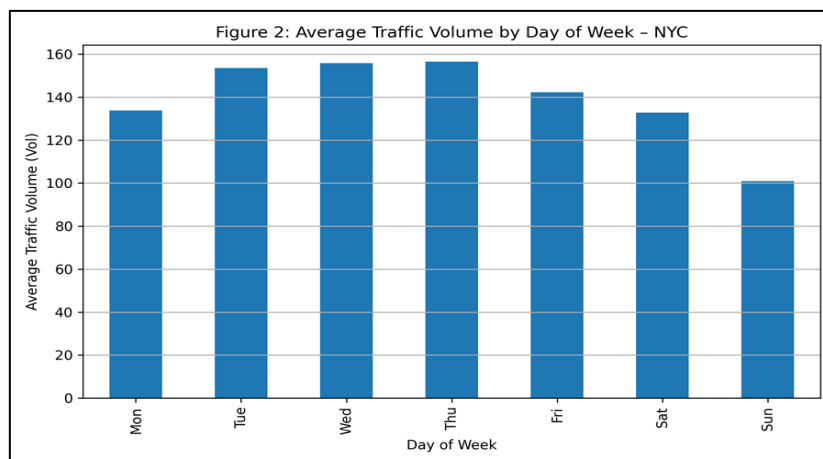


Table 2:Top 10 Road Segments by Average Volume

SegmentID	Vol
140064	817.5694444
140062	725.6979167
258712	578.8405546
144024	411.7487002
133951	354.2305026
36375	343.2478336
249325	330.3708839
145926	283.5771231
9009102	248.4939341
145924	247.4870017

Interpretation and Relevance to Smart-City Systems

These patterns, while confined to a single week, indicate the value that can be extracted from freely available traffic data. These strong temporal invariants—peaks within the day, workday vs. weekend differences and load concentration in specific segments—are a firm basis for AI models that could provide short-term forecasting or an adaptive control. Of greater concern, the results demonstrate that systems solutions for smart-city traffic require considerations of both temporal and spatial trade-offs. Such systems, aware of time-varying demand and network bottlenecks, should yield more precise interventions (by prediction modeling, dynamic signal tuning or the design of urban mobility plans).

As such, the analysis functions not just as a theoretical exercise but also as an indication of how open datasets can immediately facilitate ITS applications within contemporary smart-city environments.

CONCLUSION

Smart cities' prioritize the safety and well-being of citizens, drawing on ICT to better public health, pollution, energy supply, infrastructure and housing buildings and transport. As urbanization continues at an unprecedented rate, resulting in congested urban spaces with more than half of the planet's population living in cities and towns; high volume of traffic vehicles emissions and planning technology significantly contribute to challenges in urban mobility. Smart mobility encompasses a wide range of topics, but is generally focused on the pursuit of safer, more efficient, sustainable and accessible transportation systems.

The dynamics of mobility have been profoundly altered by the COVID-19 pandemic. Restrictions disrupted people's daily movements, schools and workplaces shifted to online modes, and the

number of vehicles on the road plummeted. In the post-epidemic recovery phase, either demand for personal transport increases or public transport remains suppressed. These shifts drive a surging interest in smart mobility research, such as demand forecasting, transportation mode detection, and vehicle routing. Machine Learning (ML) based on data analytics assumes the desired objective and is widely adopted. ML techniques, such as deep learning and reinforcement learning, are gaining significant opportunities in traffic safety, efficiency, and sustainability.

Open data promote transparency and accountability, encourage public participation, spur private investments, and stimulate innovation. Many cities are launching open-data initiatives and making unprecedented amounts of information publicly accessible. The data come from sensors, vehicle-monitoring systems, GPS devices, and semantic sensors. Many Smart-City projects are emerging, and existing projects are expanding significantly. The focus is on data management but is also exploring data-capturing and data-analytics issues. Urban mobility is essential for healthy functioning of cities. Mobility can range from individual vehicles to public transport, supply chain, and logistics. City planners often review traffic to identify problems and plan future transport accordingly. Open urban-mobility datasets can assist smart-city and mobility efforts. Open datasets originate from official-sensor data or thirdparty GPS data, while the degree of granularity varies. For services such as traffic forecasting, route planning, and incident detection, many datasets are available and are publicly accessible.

This study provides a systematic review of AI techniques in smart mobility and demonstrates sample usage scenarios based on open datasets. Examine AI trends in the smart-city domain, systematically surveying the literature from 2019

to 2025. The first part incorporates an overview of AI applications, while the latter part highlights open datasets in the context of smart cities. (Ghareh Mohammadi *et al.*, 2021)

REFERENCES:

1. Ghareh Mohammadi, F., Shenavarmasouleh, F., Hadi Amini, M., & Arabnia, H. R. "Data analytics for smart cities: Challenges and promises." *arXiv e-prints* (2021): arXiv-2109.
2. Yan, H., & Li, Y. "A survey of generative AI for intelligent transportation systems: Road transportation perspective." *arXiv preprint arXiv:2312.08248* (2023).
3. Liu, L., Porr, A., & Miller, H. J. "Realizable accessibility: evaluating the reliability of public transit accessibility using high-resolution real-time data: L. Liu *et al.*" *Journal of Geographical Systems* 25.3 (2023): 429-451.
4. Chintala, S. "Harnessing AI and BI for smart cities: Transforming urban life with data driven solutions." *International Journal of Science and Research (IJSR)* (2024).
5. Ismaeel, A. G., Mary, S. J., Anitha, C., Logeshwaran, J., Mahmood, S. N., Alani, S., & Shather, A. H. "An Enhanced Analysis of Traffic Intelligence in Smart Cities Using Sustainable Deep Radial Function." *arXiv preprint arXiv:2402.09432* (2024).
6. Khurshudov, A. "On defining smart cities using transformer neural networks." *arXiv preprint arXiv:2403.14639* (2024).
7. Bibri, S. E., Huang, J., Jagatheesaperumal, S. K., & Krogstie, J. "The synergistic interplay of artificial intelligence and digital twin in environmentally planning sustainable smart cities: A comprehensive systematic review." *Environmental science and ecotechnology* 20 (2024): 100433.
8. Mach, F., Koskinas, K. C., Roeters van Lennep, J. E., Tokgözoğlu, L., Badimon, L., Baigent, C., & Sabatine, M. S. "2025 Focused Update of the 2019 ESC/EAS Guidelines for the management of dyslipidaemias: Developed by the task force for the management of dyslipidaemias of the European Society of Cardiology (ESC) and the European Atherosclerosis Society (EAS)." *European heart journal* 46.42 (2025): 4359-4378.
9. Mobayen, M., & Sadeghi, M. "Prevalence and related factors of electrical burns in patients referred to Iranian medical centers: a systematic review and meta-analysis." *World journal of plastic surgery* 11.1 (2022): 3.
10. Mirzaie, M., Behkamal, B., Allahbakhsh, M., Paydar, S., & Bertino, E. "Ada-Context: adaptive context-aware grid-based approach for curation of data streams: M. Mirzaie." *Data Mining and Knowledge Discovery* 39.3 (2025): 22.
11. Anthony Jnr, B. "Sustainable mobility governance in smart cities for urban policy development—a scoping review and conceptual model." *Smart and sustainable built environment* 14.3 (2025): 649-671.
12. Narayanan, S., & Antoniou, C. "Shared mobility services towards Mobility as a Service (MaaS): What, who and when?." *Transportation research part A: policy and practice* 168 (2023): 103581.
13. Ziosi, M., Hewitt, B., Juneja, P., Taddeo, M., & Floridi, L. "Smart cities: reviewing the debate about their ethical implications." *AI & society* 39.3 (2024): 1185-1200.
14. Albertengo, G., & Hassan, W. "Short term urban traffic forecasting using deep learning." *ISPRS Annals of the Photogrammetry, Remote Sensing and Spatial Information Sciences* 4 (2018): 3-10.
15. Guo, M., Wang, P., Chan, C. Y., & Askary, S. "A reinforcement learning approach for intelligent traffic signal control at urban intersections." *2019 IEEE Intelligent Transportation Systems Conference (ITSC)*. IEEE, (2019).
16. Korecki, M. "Deep reinforcement meta-learning and self-organization in complex systems: Applications to traffic signal control." *Entropy* 25.7 (2023): 982.
17. Naphade, M., Wang, S., Anastasiu, D., Tang, Z., Chang, M. C., Yang, X., Zheng, L., Sharma, A., Chellappa, R., & Chakraborty, P. "The 4th AI City Challenge." (2020).
18. Jiang, R., Cai, Z., Wang, Z., Yang, C., Fan, Z., Song, X., & Shibasaki, R. "VLUC: An empirical benchmark for video-like urban computing on citywide crowd and traffic prediction." *arXiv preprint arXiv:1911.06982* (2019).
19. Honarvar, A. R., & Sami, A. "Multi-source dataset for urban computing in a Smart City." *Data in brief* 22 (2019): 222-226.
20. Albertengo, G., & Hassan, W. "Short term urban traffic forecasting using deep learning." *ISPRS Annals of the Photogrammetry, Remote Sensing and Spatial Information Sciences* 4 (2018): 3-10.
21. Li, H., Li, Y., Wang, H., Zeng, J., Xu, H., Cai, P., & Qiao, Y. "Open-sourced data ecosystem

- in autonomous driving: the present and future." *Scientia Sinica Informationis* 54.6 (2023): 1283.
22. Yazgan, M., Akkanapragada, M. V., & Zöllner, J. M. "Collaborative perception datasets in autonomous driving: A survey." *2024 IEEE Intelligent Vehicles Symposium (IV)*. IEEE, (2024).
 23. Kashinath, S. A., Mostafa, S. A., Mustapha, A., Mahdin, H., Lim, D., Mahmoud, M. A., & Yang, T. J. "Review of data fusion methods for real-time and multi-sensor traffic flow analysis." *IEEE Access* 9 (2021): 51258-51276.
 24. El-Wakeel, A. S., Noureldin, A., Hassanein, H. S., & Zorba, N. "IDriveSense: Dynamic route planning involving roads quality information." *2018 IEEE Global Communications Conference (GLOBECOM)*. IEEE, (2018).
 25. Zardini, G., Lanzetti, N., Pavone, M., & Frazzoli, E. "Analysis and control of autonomous mobility-on-demand systems." *Annual Review of Control, Robotics, and Autonomous Systems* 5.1 (2022): 633-658.
 26. Atzori, A., Barra, S., Carta, S., Fenu, G., & Podda, A. S. "Heimdall: an AI-based infrastructure for traffic monitoring and anomalies detection." *2021 IEEE International Conference on Pervasive Computing and Communications Workshops and other Affiliated Events (PerCom Workshops)*. IEEE, (2021).
 27. Zhao, Y., Wu, W., He, Y., Li, Y., Tan, X., & Chen, S. "Good practices and a strong baseline for traffic anomaly detection." *arXiv preprint arXiv:2105.03827* (2021).
 28. Culita, J., Caramihai, S. I., Dumitrache, I., Moisescu, M. A., & Sacala, I. S. "An hybrid approach for urban traffic prediction and control in smart cities." *Sensors* 20.24 (2020): 7209.
 29. Rinchi, O., Alsharoa, A., Shatnawi, I., & Arora, A. "The role of intelligent transportation systems and artificial intelligence in energy efficiency and emission reduction." *arXiv preprint arXiv:2401.14560* (2024).
 30. Shaygan, M., Meese, C., Li, W., Zhao, X. G., & Nejad, M. "Traffic prediction using artificial intelligence: Review of recent advances and emerging opportunities." *Transportation research part C: emerging technologies* 145 (2022): 103921.
 31. Lim, C., Kim, K. J., & Maglio, P. P. "Smart cities with big data: Reference models, challenges, and considerations." *Cities* 82 (2018): 86-99.
 32. Bellini, E., Bellini, P., Cenni, D., Nesi, P., Pantaleo, G., Paoli, I., & Paolucci, M. "An IOE and big multimedia data approach for urban transport system resilience management in smart cities." *Sensors* 21.2 (2021): 435.

Source of support: Nil; **Conflict of interest:** Nil.

Cite this article as:

Hashim, R. T. "AI in Smart Cities: A Systematic Literature Review and Open Traffic Datasets Investigation" *Sarcouncil Journal of Engineering and Computer Sciences* 4.11 (2025): pp 266-283.