Big Data Analytics in Smart Cities: Systems, Techniques, and Applications

Elias Dritsas, Maria Trigka, and Phivos Mylonas
Department of Informatics and Computer Engineering
University of West Attica, Egaleo 12243, Greece
{idritsas,mtrigka,mylonasf}@uniwa.gr

***Abstract*—This survey presents a system-level analysis of big data analytics in smart cities, bridging data sources, analytical techniques, and deployment infrastructures. It categorizes analytics into stream, batch, predictive, semantic, and explainable methods, and maps them to edge, fog, and cloud layers through distributed orchestration and standardized interfaces. Real-world applications across mobility, environment, safety, and citizen services are examined, emphasizing execution-aware design, interoperability, and compliance with ethical and regulatory frameworks. The survey highlights integration gaps in prior works and outlines future directions for transparent, scalable, and semantically enriched urban intelligence systems.**

***Index Terms*—Smart Cities, Big Data, Urban Intelligence, Architectures, Cross-Domain Analytics**

I. Introduction

The rapid deployment of urban sensors, mobile devices, and interconnected systems has generated massive volumes of heterogeneous, high-velocity data. Structured and unstructured streams originate from traffic sensors, environmental monitors, social media, smart meters, and user-facing applications. Extracting actionable insights from such data demands scalable infrastructures and advanced analytics. Big data analytics thus plays a central role in enabling urban intelligence across domains, such as mobility, pollution, energy, and emergency management [1].

Yet, the technical landscape remains fragmented. Scheduled, stream, predictive, semantic, and explainable analytics techniques are often applied in isolation, lacking architectural cohesion. Urban data systems operate across edge, fog, and cloud infrastructures, each posing distinct latency, privacy, and scalability challenges. A systematic analysis is needed to bridge analytics with execution architectures and align urban data practices with interoperability, accountability, and regulatory compliance requirements [2].

A. Motivation

Smart cities produce high-volume, heterogeneous data across transport, environment, safety, and energy domains. While prior studies address individual aspects of mobility analytics, Internet of Things (IoT) platforms, or policy models, a gap persists in connecting analytical methods with execution infrastructures. Urban intelligence requires more than algorithms: it depends on real-time data orchestration, semantic interoperability, and deployment-aware pipelines that respect societal and regulatory constraints.

Recent advances in stream processing, edge–fog–cloud computing, and explainable artificial intelligence (XAI) enable actionable urban analytics, yet existing literature remains fragmented. There is limited integration across analytics layers, architectural designs, and real-world deployments. This survey addresses this gap by mapping analytics techniques to system architectures, identifying deployment challenges, and outlining best practices for scalable, transparent, citizen-focused smart city systems.

B. Methodology

The study follows a structured and reproducible methodology to collect, select, and analyze recent literature on big data analytics for smart cities. Relevant works were systematically retrieved from five major repositories, IEEE Xplore, Elsevier ScienceDirect, SpringerLink, ACM Digital Library, and Google Scholar, focusing on peer-reviewed articles, conference papers, and surveys published from 2021 onward. Inclusion was limited to studies addressing big data analytics in urban contexts, across ingestion, processing, modeling, or deployment stages. Excluded were papers limited to low-level sensing, narrowly scoped optimizations, or theoretical models without system-level applicability. After de-duplication and abstract-level filtering, 49 publications were retained.

The research design comprised three phases. First, retrieval was based on keywords such as “big data analytics,” “smart cities,” and “urban intelligence.” Second, selected papers were thoroughly reviewed to extract metadata on analytics types (batch, stream, predictive, semantic, explainable), architectural layers (edge, fog, cloud), and domains (mobility, environment, safety, citizen services). Third, the works were mapped to a unified taxonomy and comparatively analyzed across execution scope, infrastructure, interoperability, and regulatory compliance. This process supports consistent and comprehensive insight into the current state of smart city analytics.

C. Contribution

This survey offers a structured classification of big data analytics techniques for smart cities, spanning stream, batch, predictive, semantic, and explainable analytics dimensions. These are mapped to deployment layers, edge, fog, and cloud, highlighting their system-level integration. A domain-specific application mapping is provided across mobility, environment, citizen services, and safety, supported by concrete use cases. The study also analyzes execution mechanisms and incorporates cross-cutting concerns such as interoperability, compliance, and explainability. A comparative analysis with recent surveys highlights integration gaps and supports the novelty of this end-to-end perspective.

Section II outlines key concepts, data types, system layers, and analytics dimensions. Section III presents the taxonomy of analytics techniques and their execution characteristics across the edge–fog–cloud stack. Section IV examines representative domains and real-world deployments. Section V compares prior surveys, discusses challenges, and outlines future trends. Section VI concludes the paper.

II. Data Sources and Infrastructures for Smart Cities

Smart city platforms are sustained by diverse urban data streams and a layered infrastructure that ensures reliable acquisition, processing, and integration. This section presents major categories of data sources and the architectural components that support scalable, real-time analytics and decision-making.

A. Urban Sensor Networks and Operational Systems

Physical sensors embedded in roadways, buildings, and environmental stations generate continuous structured data streams. Inductive loops and magnetometers measure traffic density and lane occupancy, while air quality stations monitor pollutants such as PM2.5 and NO_2. Smart meters transmit consumption profiles from water, gas, and electricity grids, often via supervisory control and data acquisition (SCADA) or advanced metering infrastructure (AMI) systems. These sources follow standardized timestamping and encoding, supporting periodic or event-triggered reporting [3], [4].

Public services generate semi-structured records via automatic vehicle location (AVL)-enabled vehicles, global positioning system (GPS) tracking, and route execution logs. Transportation networks produce vehicle locations and delay reports; waste management systems transmit bin status and geotagged collection events. Formats vary from comma-separated values (CSV) and JavaScript Object Notation (JSON) to proprietary telemetry, requiring schema harmonization during preprocessing [5].

Differences in sampling rates and delivery reliability necessitate preprocessing techniques such as time-window alignment, device synchronization, and buffering. Before storage, data undergoes normalization of units, reconciliation of timestamps, and metadata enrichment [6].

B. Citizen-Generated and Multimedia Data

Citizens contribute unstructured data through incident reports, mobile applications, and social media, complementing infrastructure-based sources. These inputs contain linguistic variation, incomplete metadata, and inconsistent geotagging. Natural language processing pipelines extract event types, locations, and temporal references, while geocoding and temporal inference enhance alignment with structured layers [7].

Video data from municipal closed-circuit television (CCTV) and mobile devices introduces additional challenges. Live feeds require on-device inference using convolutional models for object detection, behavioral analysis, and anomaly tracking. Due to bandwidth limits, only processed metadata (e.g., object counts, motion vectors) is transmitted upstream, while raw video is selectively stored for forensics and model retraining [8], [9].

C. Processing Architecture: Edge, Fog, and Cloud Layers

Urban data infrastructures employ hierarchical processing to balance latency and capacity. At the sensing site, edge nodes handle threshold-based triggers, low-latency filtering, or initial feature extraction. They communicate over lightweight IoT protocols such as message queuing telemetry transport (MQTT) or constrained application protocol (CoAP), operating under energy and bandwidth constraints [10], [11].

Fog nodes perform regional aggregation and mid-level analytics, executing spatiotemporal correlation and real-time coordination tasks, such as synchronizing adaptive traffic signals. These nodes act as intermediaries between local sensors and global systems, reducing the volume and redundancy of data transmitted to the cloud [12].

Cloud systems provide cross-domain integration and heavy analytics. Distributed file systems like Hadoop Distributed File System (HDFS) store raw and processed data. In contrast, not only Structured Query Language (NoSQL) engines such as MongoDB and Cassandra support scalable querying. Distributed stream processing frameworks perform both batch and real-time analytics. Directed acyclic graph (DAG)-based orchestrators coordinate tasks including ingestion, validation, model application, and export to external services [13], [14].

D. Interoperability, Privacy, and Governance

Smart city deployments face interoperability challenges across domains and devices. Next generation service interface linked data (NGSI-LD) and SensorThings application programming interface (API) define semantic models for context-aware data representation and querying. City geography markup language (CityGML) provides geometric and semantic metadata for urban objects, supporting fusion with simulation platforms [15].

Middleware platforms mediate between data producers and consumers. FIWARE Orion brokers expose unified APIs and manage dynamic discovery, registration, and routing. These ensure compatibility between heterogeneous systems without requiring complex coupling [16].

Data governance imposes strict access control, anonymization, and compliance mechanisms. Sensitive information, such as GPS traces, is anonymized using k-anonymity, spatial cloaking, or differential privacy. General data protection regulation (GDPR)-compliant systems incorporate consent-aware APIs, metadata registries, and auditable processing pipelines [17].

TABLE I
URBAN DATA SOURCES, PROCESSING LAYERS, AND ASSOCIATED TECHNOLOGIES.

Category	Data Types & Modalities	Processing Layer	Technologies & Standards
Sensor-Level Infrastructure	Traffic sensors, smart meters, air quality monitors	Edge, Fog, Cloud	MQTT, CoAP, Modbus, IEC 61850
Operational Systems	Transport logs, GPS traces, bin reports	Fog, Cloud	JSON, CSV, SCADA, AVL systems
Citizen-Generated Content	Reports, text, social media, CCTV streams	Edge (inference), Cloud (archival)	NLP, BERT, YOLO, social APIs
Analytics and Storage Infrastructure	Fused, cleaned multistream data	Fog (preprocessing), Cloud (analytics)	Kafka, Spark, Flink, MongoDB, Cassandra, HDFS
Interoperability and Governance	Context entities, metadata, access policies	Middleware, Brokers	NGSI-LD, SensorThings API, CityGML, FIWARE Orion, GDPR

Table I consolidates the main categories of data, processing layers, and associated technologies discussed in this section. It provides a reference model for how smart city infrastructures manage diverse, cross-domain information at scale.

III. BIG DATA ANALYTICS TECHNIQUES FOR SMART CITIES

This section analyzes the analytical pipelines that enable adaptive urban decision-making through high-throughput, context-sensitive computation. As smart city systems integrate domain-specific information layers, ranging from real-time sensor feeds to historical service records, they require scalable techniques tailored to distinct temporal granularities and operational latencies.

A. Stream Analytics for Real-Time Urban Intelligence

Stream analytics enables real-time responses to dynamic urban conditions by processing continuous data using frameworks like Apache Flink and Spark Structured Streaming. These systems handle event-time processing, in-flight aggregation, and temporal joins across potentially disordered or delayed data, essential for correlating events such as vehicle speed anomalies [18].

Their primary advantage lies in low-latency operation. In traffic systems, real-time GPS and sensor feeds inform congestion detection and adaptive signal control, while environmental systems trigger alerts or heating, ventilation, and air conditioning (HVAC) adjustments based on pollutant levels. Processing is often colocated with edge or fog nodes to reduce latency in location-sensitive scenarios [19].

Despite these benefits, challenges persist in achieving consistency and scalability under bursty loads. Mechanisms such as watermarking, checkpointing, and exactly-once semantics are essential to ensure correctness, especially in mission-critical urban deployments [20].

B. Batch Analytics and Temporal Pattern Discovery

Batch analytics retrospectively examines large urban datasets to uncover temporal patterns and latent correlations. Run on platforms like Spark, Hive, or Dask in data lakes, it supports clustering, association rules, and time-series decomposition [21].

Their strength lies in analytical depth. Utilities use multi-year smart meter data to detect anomalies or seasonal drifts, while transport agencies apply temporal clustering to assess past interventions or guide infrastructure planning. Integrating external variables, such as weather, holidays, or events, enriches context modeling [22].

Though not real-time, batch analytics are vital for strategic planning. Long-term forecasts on energy demand or policy effectiveness rely on models built from extended temporal windows, semantically labeled and spatially indexed [23].

C. Predictive Modeling and Forecast-Driven Optimization

Predictive analytics aims to forecast future urban states using statistical and machine learning models. Techniques range from univariate time-series methods (e.g., ARIMA, Prophet) to multivariate models like gradient boosting and Long Short-Term Memory (LSTM) networks that capture nonlinear dependencies [24].

Applications span multiple domains: traffic volume predictions enable demand-responsive transit, emergency visit forecasts inform health resource allocation, and short-term load forecasting supports grid management. Predictions also guide optimization pipelines, anticipated pedestrian flows adjust micro mobility pricing, while parking availability forecasts inform routing. In this context, predictions are not endpoints but inputs that steer real-time decisions and urban service adaptation [25].

D. Semantic Reasoning and Knowledge-Based Integration

Semantic reasoning applies symbolic logic and ontologies (e.g., sensor, observation, sample, and actuator ontology (SOSA) / semantic sensor network ontology (SSN), CityGML) to unify heterogeneous urban data. Middleware exposes reasoning over annotated entities and relations via RDF triples and rule engines [26].

This enables integration across sources such as elevation maps, rainfall sensors, and infrastructure data, supporting flood response through rule-based inference on propagation patterns and real-time thresholds. Semantic annotation facilitates automated fusion, provenance tracking, and explainability [27].

As cities adopt digital twins, semantic frameworks interlink simulation, telemetry, and registries, ensuring coherent, cross-domain interoperability and transparent decision support [28].

E. Explainable Models and Transparent Decision Support

As algorithmic decision-making permeates urban services, the demand for interpretability becomes fundamental. Black-box models, especially in deep learning, are often inappropriate in public-sector contexts unless augmented with explanation mechanisms. Methods such as SHapley Additive exPlanations (SHAP) and Local Interpretable Model-agnostic Explanations (LIME), counterfactual analysis, and gradient-based attribution offer insights into feature relevance and model behavior under perturbations [29].

TABLE II
BIG DATA ANALYTICS TECHNIQUES FOR SMART CITIES.

Analytical Technique	Target Data Types	Execution Tier	Related Technologies & Models
Stream Analytics	Real-time sensor data, GPS traces, environmental readings	Edge, Fog	Apache Flink, Spark Streaming, Kafka, MQTT
Batch Analytics	Historical logs, smart meter archives, urban service records	Fog, Cloud	Apache Spark, Hive, Dask, HDFS
Predictive Modeling	Multivariate time series, demand profiles, user mobility	Fog, Cloud	ARIMA, LSTM, XGBoost, Prophet
Semantic Reasoning	Domain-tagged records, spatial knowledge graphs	Cloud, Middleware	RDF, OWL, CityGML, SOSA/SSN, SPARQL
XAI	Model outputs, inferred decisions, sensitive predictions	Cloud, Decision Layer	SHAP, LIME, Counterfactuals, Attention Maps

Explainability supports two distinct goals in practical deployments: internal validation and external accountability. Engineers use XAI tools to identify misclassification patterns, investigate outliers, or monitor for concept drift. Simultaneously, decision-makers need interpretable justifications when models prioritize interventions, recommend zoning changes, or automate service eligibility [30].

Furthermore, public-facing applications must balance transparency with privacy. Visual summaries and natural-language explanations enhance user trust but must avoid leaking sensitive correlations. In this context, XAI is not merely a debugging tool but a bridge between statistical inference and policy legitimacy [31].

Table II summarizes the core categories of big data analytics employed in smart city contexts, mapping each technique to the corresponding data modalities, processing layers, and associated technologies.

IV. APPLICATIONS AND USE CASES

Smart city analytics materialize through domain-specific services, where contextual adaptation and operational relevance are essential. This section outlines representative applications in mobility, sustainability, public safety, and citizen services, illustrating how analytical methods align with temporal constraints, governance structures, and feedback loops unique to each domain.

A. Intelligent Mobility and Traffic Orchestration

Smart city mobility systems leverage real-time analytics to optimize flow, enhance safety, and cut emissions. Adaptive traffic signals use live GPS and loop detector data to adjust phases under low-latency constraints, often via fog-level processing. At the system level, multimodal transit platforms combine historical ridership, real-time vehicle locations, and disruption alerts to optimize routing. Predictive models guide dynamic fleet assignments and bus frequency adjustments, integrating micro-mobility data for responsive service delivery. Semantic reasoning further supports unified management across transport modes. Ontology-based models define route hierarchies and service priorities, enabling automated violation detection, emergency prioritization, and cross-agency coordination via interoperable data layers [32], [33].

B. Environmental Monitoring and Urban Sustainability

Sustainability applications focus on long-term environmental monitoring and mitigation. Air quality platforms use distributed sensors and predictive models to detect pollution hotspots, trigger alerts, and control public HVAC systems. Seasonal trends extracted via time-series analysis inform zoning and policy. Sensor-equipped waste bins report fill levels, enabling batch analytics to uncover inefficiencies and predictive routing to reduce overflows and fuel use. Incorporating weather and event data refines scheduling. Water systems integrate Water systems fuse SCADA streams with anomaly detection to spot contamination or leaks, combining edge detection, fog aggregation, and central validation for proactive maintenance and resource efficiency [34], [35].

C. Public Safety and Emergency Response

Urban safety systems employ real-time sensing and inference for early detection and response. CCTV and mobile signals support crowd monitoring, enabling anomaly detection for event safety and evacuation planning. Object recognition and trajectory analysis flag suspicious behavior or unauthorized access. Using semantic ontologies to infer risk zones and allocate resources, disaster platforms integrate weather, hydrological, and citizen data. Coupled with digital twins, these systems simulate scenarios and assess impact. Emergency services use historical data and environmental inputs to forecast incidents and pre-position assets. Given the stakes, such predictive decisions require transparency and auditability to ensure trust in life-critical deployments [36], [37].

D. Citizen Services and Participatory Platforms

Citizen-centric analytics in smart cities integrates participatory sensing, open data access, and personalized service delivery. Mobile applications enable users to report urban issues (e.g., noise, lighting, road defects), with natural language processing (NLP) pipelines extracting structured information and geo-referencing events for automated routing to relevant authorities. Open data portals provide real-time and historical datasets on mobility, air quality, and utilities, facilitating third-party innovation and transparent governance. Usage patterns and feedback analytics drive iterative service optimization and expose systemic inefficiencies. Personalized dashboards leverage urban data streams and XAI to generate context-aware alerts and behavioral recommendations. For instance, air quality warnings may prompt indoor activity suggestions, accompanied by interpretable justifications. Privacy-preserving methods such as differential privacy or federated aggregation ensure anonymity while maintaining data utility for fine-grained, citizen-level service adaptation [38], [39].

Table III consolidates the key urban domains addressed by big data analytics in smart city contexts. It highlights the main application areas, representative real-world use cases, analytical techniques employed, and the corresponding execution layers. The table illustrates how heterogeneous methods such as stream analytics, semantic reasoning, and XAI are deployed across the urban stack, from edge processing of sensor data to cloud-based decision support. Each use case anchors the analytics in a concrete service setting, showcasing data translation into actionable urban intelligence.

V. Discussion and Future Trends

Several recent surveys have explored big data analytics in smart cities, each emphasizing different layers of the system stack. [40] centers on policy adoption and socio-technical theory, neglecting analytics pipelines and infrastructure. [41] focuses on spatiotemporal clustering and mobility modeling, lacking architectural and semantic integration. [42] adopts a governance-oriented lens, highlighting temporal misalignments without engaging technical detail.

Similarly, [43] proposes a privacy- and consent-driven ethical framework with case studies but omits analytics architecture. [44] classifies applications like crime prediction and transport optimization without formal modeling or deployment depth. [45] surveys IoT–big data integration using layered architectures, yet offers limited discussion on analytic diversity and practical services. In contrast, the present survey provides an end-to-end perspective, spanning data ingestion, real-time and retrospective analytics, semantic modeling, and XAI, explicitly mapped to execution layers and urban domains.

By aligning big data techniques with infrastructure tiers (edge, fog, cloud) and key applications (mobility, sustainability, safety, citizen services), this work offers a structured blueprint for building interpretable, latency-aware, and fault-tolerant systems. It advances urban analytics by integrating statistical learning, semantic reasoning, and distributed orchestration, addressing prior literature gaps and enhancing theoretical clarity and practical relevance.

A key future trend in smart cities is the convergence of semantic interoperability and real-time analytics. Adopting knowledge graphs, ontologies (e.g., SOSA/SSN, CityGML), and standardized APIs (NGSI-LD, SensorThings) will enable seamless data exchange across domains like mobility, environment, and energy. Paired with reasoning engines, these tools will support cross-domain inference and automated decision-making. Their integration with DAG-based orchestration and distributed stream processing (e.g., Apache Flink, Spark Structured Streaming) will foster scalable, context-aware urban workflows [46], [47].

Equally crucial is deploying XAI in critical services. As predictive models inform traffic control, emergency response, and environmental alerts, interpretability becomes essential for engineers, officials, and citizens. Future platforms will embed XAI techniques, such as SHAP, LIME, counterfactuals, and attribution maps, into dashboards and user interfaces. With growing regulatory pressure (e.g., GDPR, AI Act), pipelines must ensure transparency and accountability by design, promoting auditable, human-centered intelligence over opaque model performance [48], [49].

VI. Conclusions

This survey examined big data analytics in smart cities through a system-oriented lens, integrating analytical techniques with infrastructure execution and real-world application domains. A five-dimensional taxonomy was introduced, covering stream, batch, predictive, semantic, and explainable analytics, and mapped across the edge–fog–cloud continuum. The analysis extended to practical deployments in mobility, environment, citizen services, and safety, highlighting the interplay between data processing models, architectural constraints, and regulatory considerations. By comparing existing surveys, this work identified fragmentation in prior approaches and addressed the lack of end-to-end perspectives. The findings offer a foundation for designing urban analytics systems that are execution-aware, interoperable, and aligned with accountability and transparency requirements. These insights can inform the design of future urban analytics platforms that are both scalable and context-aware.

References

[1] D. V. Ogunkan and S. K. Ogunkan, "Exploring big data applications in sustainable urban infrastructure: A review," *Urban Governance*, 2025.

[2] M. Trigka and E. Dritsas, "Edge and cloud computing in smart cities," *Future Internet*, vol. 17, no. 3, p. 118, 2025.

[3] E. Zadobrischi, "Intelligent traffic monitoring through heterogeneous and autonomous networks dedicated to traffic automation," *Sensors*, vol. 22, no. 20, p. 7861, 2022.

[4] S. Shahid, D. J. Brown, P. Wright, A. M. Khasawneh, B. Taylor, and O. Kaiwartya, "Innovations in air quality monitoring: Sensors, iot and future research," *Sensors*, vol. 25, no. 7, p. 2070, 2025.

[5] A. Abdelhalim and J. Zhao, "Computer vision for transit travel time prediction: an end-to-end framework using roadside urban imagery," *Public Transport*, pp. 1–26, 2024.

[6] A. Tawakuli, D. Kaiser, and T. Engel, "Transforming iot data preprocessing: A holistic, normalized and distributed approach," in *Proceedings of the 20th ACM Conference on Embedded Networked Sensor Systems*, 2022, pp. 1083–1088.

[7] D. Bastos, A. Fernández-Caballero, A. Pereira, and N. P. Rocha, "Smart city applications to promote citizen participation in city management and governance: A systematic review," in *Informatics*, vol. 9, no. 4. MDPI, 2022, p. 89.

[8] D. Ngo, H.-C. Park, and B. Kang, "Edge intelligence: A review of deep neural network inference in resource-limited environments," *Electronics*, vol. 14, no. 12, p. 2495, 2025.

[9] O. El Marai, S. Messinis, N. Doulamis, T. Taleb, and J. Manner, "Roads infrastructure digital twin: Advancing situational awareness through bandwidth-aware 360° video streaming and multi-view clustering," *IEEE Open Journal of Vehicular Technology*, 2025.

[10] A. Al-Dulaimy, M. Jansen, B. Johansson, A. Trivedi, A. Iosup, M. Ashjaei, A. Galletta, D. Kimovski, R. Prodan, K. Tserpes *et al.*, "The computing continuum: From iot to the cloud," *Internet of Things*, vol. 27, p. 101272, 2024.

[11] E. Dritsas and M. Trigka, "Machine learning for blockchain and iot systems in smart cities: A survey," *Future Internet*, vol. 16, no. 9, p. 324, 2024.

[12] K. Sambhav, "Fog computing in iot-driven smart city traffic control systems," *Smart Internet of Things*, vol. 2, no. 1, pp. 35–43, 2025.

[13] M. B. Ribeiro and K. R. Braghetto, "A scalable data integration architecture for smart cities: implementation and evaluation," *Journal of Information and Data Management*, vol. 13, no. 2, 2022.

[14] A. Marchese and O. Tomarchio, "Orchestrating dag workflows on the cloud-to-edge continuum." in *itaDATA*, 2023.

[15] V. González, L. Martín, J. R. Santana, P. Sotres, J. Lanza, and L. Sánchez, "Reshaping smart cities through ngsi-ld enrichment," *Sensors*, vol. 24, no. 6, p. 1858, 2024.

TABLE III
ANALYTICS TECHNIQUES AND REPRESENTATIVE USE CASES IN SMART CITY DOMAINS.

Urban Domain	Key Applications	Representative Use Cases	Analytical Techniques	Execution Layer(s)
Mobility and Traffic	Adaptive signaling, route optimization, congestion detection	Dynamic signal control in Barcelona, multimodal journey planner in Helsinki	Stream analytics, predictive modeling, semantic reasoning	Edge, Fog, Cloud
Environmental Sustainability	Air quality monitoring, smart waste collection, water leakage detection	AQI alert system in London, bin-level routing in Amsterdam, leak prediction in Singapore	Time-series analysis, anomaly detection, batch analytics	Fog, Cloud
Public Safety	Crowd analytics, emergency forecasting, incident triage	Real-time crowd control in Tokyo Olympics, flood response system in Jakarta, fire risk alerts in California	Video analytics, semantic integration, forecasting	Edge, Fog, Middleware
Citizen Services	Participatory reporting, open data publishing, personalized alerts	FixMyStreet in the UK, NYC Open Data portal, air quality notifications via Delhi AQ App	NLP, XAI, participatory analytics	Cloud, Decision Tier

[16] S. Loss, H. P. Singh, N. Cacho, and F. Lopes, "Using fiware and blockchain in smart cities solutions," *Cluster Computing*, vol. 26, no. 4, pp. 2115–2128, 2023.

[17] S. Sampaio, P. R. Sousa, C. Martins, A. Ferreira, L. Antunes, and R. Cruz-Correia, "Collecting, processing and secondary using personal and (pseudo) anonymized data in smart cities," *Applied Sciences*, vol. 13, no. 6, p. 3830, 2023.

[18] Q. Zou, W. Xiong, X. Wang, and F. Qin, "Research on real-time anomaly detection method of bus trajectory based on flink," *Electronics*, vol. 12, no. 18, p. 3897, 2023.

[19] S. A. Celtek and A. Durdu, "A novel adaptive traffic signal control based on cloud/fog/edge computing," *International journal of intelligent transportation systems research*, vol. 20, no. 3, pp. 639–650, 2022.

[20] M. K. Geldenhuys, B. J. Pfister, D. Scheinert, L. Thamsen, and O. Kao, "Khaos: Dynamically optimizing checkpointing for dependable distributed stream processing," in *2022 17th Conference on Computer Science and Intelligence Systems (FedCSIS)*. IEEE, 2022, pp. 553–561.

[21] M. Bilal, R. S. A. Usmani, M. Tayyab, A. A. Mahmoud, R. M. Abdalla, M. Marjani, T. R. Pillai, and I. A. Targio Hashem, "Smart cities data: Framework, applications, and challenges," *Handbook of smart cities*, pp. 113–141, 2021.

[22] P. Lin, Y. He, M. Pei, and R. Yang, "Data-driven spatial-temporal analysis of highway traffic volume considering weather and festival impacts," *Travel behaviour and society*, vol. 29, pp. 95–112, 2022.

[23] N. Maleki, O. Lundström, A. Musaddiq, J. Jeansson, T. Olsson, and F. Ahlgren, "Future energy insights: Time-series and dl models for city load forecasting," *Applied Energy*, vol. 374, p. 124067, 2024.

[24] P. Linardatos, V. Papastefanopoulos, T. Panagiotakopoulos, and S. Kotsiantis, "Co2 concentration forecasting in smart cities using a hybrid arima–tft model on multivariate time series iot data," *Scientific reports*, vol. 13, no. 1, p. 17266, 2023.

[25] S. Akhtar, S. Shahzad, A. Zaheer, H. S. Ullah, H. Kilic, R. Gono, M. Jasiński, and Z. Leonowicz, "Short-term load forecasting models: A review of challenges, progress, and the road ahead," *Energies*, vol. 16, no. 10, p. 4060, 2023.

[26] A. De Nicola and M. L. Villani, "Smart city ontologies and their applications: A systematic literature review," *Sustainability*, vol. 13, no. 10, p. 5578, 2021.

[27] M. Hofmeister, J. Bai, G. Brownbridge, S. Mosbach, K. F. Lee, F. Farazi, M. Hillman, M. Agarwal, S. Ganguly, J. Akroyd *et al.*, "Semantic agent framework for automated flood assessment using dynamic knowledge graphs," *Data-Centric Engineering*, vol. 5, p. e14, 2024.

[28] A. Shareef, B. Tomaš, and N. Vrček, "Semantic interoperability of digital twins in smart cities," in *Central European Conference on Information and Intelligent Systems*. Faculty of Organization and Informatics Varazdin, 2023, pp. 481–487.

[29] A. R. Javed, W. Ahmed, S. Pandya, P. K. R. Maddikunta, M. Alazab, and T. R. Gadekallu, "A survey of explainable ai for smart cities," *Electronics*, vol. 12, no. 4, p. 1020, 2023.

[30] F. Hinder, V. Vaquet, and B. Hammer, "One or two things we know about concept drift—a survey on monitoring in evolving environments. part b: locating and explaining concept drift," *Frontiers in AI*, vol. 7, p. 1330258, 2024.

[31] M. H. Kabir, K. F. Hasan, M. K. Hasan, and K. Ansari, "Explainable ai for smart city application: a secure and trusted platform," in *Explainable AI for Cyber Security: Next Generation AI*. Springer, 2022, pp. 241–263.

[32] M. Ashkanani, A. AlAjmi, A. Alhayyan, Z. Esmael, M. AlBedaiwi, and M. Nadeem, "A self-adaptive traffic signal system integrating real-time vehicle detection and license plate recognition for enhanced traffic management," *Inventions*, vol. 10, no. 1, p. 14, 2025.

[33] M. Scrocca, M. Grassi, M. Comerio, V. A. Carriero, T. Delgado Dias, A. Vieira Da Silva, and I. Celino, "Intelligent urban traffic management via semantic interoperability across multiple heterogeneous mobility data sources," in *International Semantic Web Conference*. Springer, 2024, pp. 218–235.

[34] K. Chatterjee, M. N. Thara, M. S. Reddy, N. Selvamuthukumaran, M. Priyadharshini, T. A. V. Reddy, S. Chakraborty, S. Mallik, M. A. Shah, and A. Li, "Ecosense: a revolution in urban air quality forecasting for smart cities," *BMC Research Notes*, vol. 18, p. 62, 2025.

[35] S. R. Laha, B. K. Pattanayak, and S. Pattnaik, "Advancement of environmental monitoring system using iot and sensor: A comprehensive analysis," *AIMS Environmental Science*, vol. 9, no. 6, pp. 771–800, 2022.

[36] Y. Myagmar-Ochir and W. Kim, "A survey of video surveillance systems in smart city," *Electronics*, vol. 12, no. 17, p. 3567, 2023.

[37] M. M. F. Ariyachandra and G. Wedawatta, "Digital twin smart cities for disaster risk management: a review of evolving concepts," *Sustainability*, vol. 15, no. 15, p. 11910, 2023.

[38] E. Przeybilovicz, M. A. Cunha, S. Geertman, C. Leleux, A. Michels, Z. Tomor, C. W. R. Webster, and A. Meijer, "Citizen participation in the smart city: findings from an international comparative study," *Local government studies*, vol. 48, no. 1, pp. 23–47, 2022.

[39] E. Dritsas and M. Trigka, "Federated learning for iot: A survey of techniques, challenges, and applications," *Journal of Sensor and Actuator Networks*, vol. 14, no. 1, p. 9, 2025.

[40] C. Davies, "Big data analytics for smart cities," *International Journal of Computing and Engineering*, vol. 6, no. 1, pp. 14–29, 2024.

[41] H. Huang, X. A. Yao, J. M. Krisp, and B. Jiang, "Analytics of location-based big data for smart cities: Opportunities, challenges, and future directions," *Computers, Environment and Urban Systems*, vol. 90, p. 101712, 2021.

[42] J. Kandt and M. Batty, "Smart cities, big data and urban policy: Towards urban analytics for the long run," *Cities*, vol. 109, p. 102992, 2021.

[43] V. Chang, "An ethical framework for big data and smart cities," *Technological Forecasting and Social Change*, vol. 165, p. 120559, 2021.

[44] E. Cesario, "Big data analytics and smart cities: applications, challenges, and opportunities," *Frontiers in big data*, vol. 6, p. 1149402, 2023.

[45] M. Talebkhah, A. Sali, M. Marjani, M. Gordan, S. J. Hashim, and F. Z. Rokhani, "Iot and big data applications in smart cities: recent advances, challenges, and critical issues," *IEEE Access*, vol. 9, pp. 55 465–55 484, 2021.

[46] A. Abid, J. Lee, F. Le Gall, and J. Song, "Toward mapping an ngsi-ld context model on rdf graph approaches: A comparison study," *Sensors*, vol. 22, no. 13, p. 4798, 2022.

[47] R. H. Pereira, "Lightweight data bridge for connecting self-service end-user analytic tools to ngsi-based iot systems," *Internet of Things*, vol. 25, p. 101125, 2024.

[48] X. Kong, S. Liu, and L. Zhu, "Toward human-centered xai in practice: A survey," *Machine Intelligence Research*, vol. 21, no. 4, pp. 740–770, 2024.

[49] K. Zhang, J. Zhang, P.-D. Xu, T. Gao, and D. W. Gao, "Explainable ai in deep reinforcement learning models for power system emergency control," *IEEE Transactions on Computational Social Systems*, vol. 9, no. 2, pp. 419–427, 2021.