# Bridging Knowledge Gaps in Smart City Intelligence Using Federated Explainable Agentic AI

Dr. Pandi Chiranjeevi[1], Prof. (Dr.) Sailesh Suryanarayan Iyer[1,2]

[1]Lincoln University College , Petaling Jaya, Selangor, Malaysia.
[1,2]Principal, Narnarayan Shastri Institute of Technology – Institute of Forensic Sciences and Cyber Security (Affiliated to NFSU), Ahmedabad – 382426, Gujarat, India

Email: pdf.chiranjeevi@lincoln.edu.my[1] , drsaileshiyer@gmail.com[2]

---

**Abstract**: Increasingly, urban smart city systems depend on data-driven intelligence for making informed decisions related to transportation, energy, and public administration. Nevertheless, the current strategies still lack cohesion. In this study, we introduce a Federated Explainable Agentic AI (FEA-AI)-based framework for smart cities that combines decentralized federated learning, explainable decision making, and agent-based autonomous reasoning in order to facilitate private local data analysis, interpretation of results, and flexible decision making in dispersed environments. The newly proposed framework has a lot of positive factors compared to centralized methods. The framework's value lies in the provision of certain foundations for smart urban management intelligence.
**Keywords:** *Federated intelligence, Explainable decision systems, Agentic AI, Privacy-aware learning, Urban decision support, Adaptive analytics*

## 1. Introduction

The complexity of modern cities and the increase in the amount of diverse information produced by gadgets, sensors, and Digital Infrastructure make necessary the introduction of intelligent and data-oriented decision support systems in smart cities. Such systems can be utilized in sectors such as transport, energy, environmental observation, medicine, and public sphere. At the same time, traditional approaches to centralized intelligence suffer from drawbacks connected with privacy, scalability, communication power requirements, and adaptability when functioning in complex urban conditions.

Although distributed and federated learning paradigms increase the level of privacy due to decentralization of data processing, they do not provide transparent and coordinated decision-making on their own [2], [3]. Explainable AI contributes to the interpretation and accountability of results of models but it is commonly treated separately from decentralized learning paradigms [4], [5]. Agentic AI uses machine learning to carry out independent decision making, but the combination of this approach with confidentiality guaranteeing learning and explainable intelligence has not been greatly explored yet [12], [13]. Thus, the existing paradigms offer solutions to separate problems, without being unified.

In order to overcome the mentioned constraints, this research introduces a Federated Explainable Agentic AI (FEA-AI) framework for smart city intelligence purposes. The framework incorporates federated learning in order to enable privacy-aware decentralized model building, explainable reasoning so that results can be interpreted in a transparent manner, and agentic intelligence in order to facilitate adaptive and free-willed decision making. This proposed framework is expected to ensure the implication of scaled, privacy-preserving, clear-cut, and adaptive decision support across a variety of smart cities.The main contributions of this study are:

1. Identification of crucial information gaps and weaknesses in current smart city intelligence methods.

2. Creation of a unified FEA-AI framework that integrates decentralized learning, understandable reasoning, and decision-making using self-sufficient agents.
3. Creating an expandable and privacy-sensitive framework for a transparent smart city decision-making unit.
4. Identifying a proper procedure which goes from distributed data gathering to federated learning and explainability and adaptive decision-making.
5. Creation of an experimental framework for testing prediction accuracy, communication expenses, data confidentiality, scalability, comprehensibility, and adaptability.

## 2. Knowledge Gaps

Today's smart city intelligence methods focus on specific issues like data centralization, privacy, clear explanation, and autonomous judgment, while these solutions are created separately, thus causing the separation and the lack of cooperation between various urban sectors.

Although distributed and federated learning ensures the protection of privacy through localized learning, their implementation in conjunction with coordinated decision support, explainability, and autonomous reasoning is still somewhat limited. The same is true for Explainable AI, which improves transparency and accountability but still remains an isolated or centralized solution. Although agent-based and adaptive systems can provide autonomous decision-making, these technologies have never been integrated with decentralized learning and explainable intelligence.

Taken together, all of these points indicate a critical gap in research – the absence of a singular intellectual framework in smart cities that will incorporate decentralized learning with privacy preservation, explainable reasoning, and autonomous actions [8], [12], [13], [14]. Another less significant research gap is that there are not enough experimental approaches for evaluating privacy, scalability, transparency, adaptability, and decision-support efficiency.

**Table 1:** Comparative Analysis of Existing Approaches and Identified Research Gaps in Smart City Intelligence

| Approach | Key Strength | Key Limitation | Missing Integration (Gap) |
|---|---|---|---|
| **Centralized Intelligence** | Centralized global data aggregation and coordinated control | Privacy risks, scalability limitations, single point of failure | Absence of integrated distributed and privacy-preserving intelligence within unified decision-support frameworks |
| **Federated Learning [2], [3], [10], [11]** | Privacy preservation through decentralized model training | High communication overhead during model synchronization and coordination complexity | Limited integration with explainability and autonomous decision-making mechanisms |
| **Explainable AI [4], [5]** | Improved transparency and interpretability of model decisions | Primarily applied in isolated or centralized environments | Lack of integration with distributed learning and agent-based adaptive systems |
| **Agentic AI [12], [13]** | Adaptive and goal-driven autonomous decision-making | Limited accountability and lack of real-world deployment maturity | Lack of integration with privacy-preserving and explainable intelligence frameworks |

The study shows that the current techniques resolve single issues but do not offer a consolidated answer including decentralized learning, transparency, and self-governing decision-making, which creates the background for the establishment of the proposed combined framework.

## 3. Proposed Methodology

The Federated Explainable Agentic AI framework integrates decentralized learning, explainable reasoning, and autonomous decision-making into an interconnected intelligent system for smart cities. The proposed methodology is comprised of five connected aspects, namely, edge data processing, federated learning, explainability, agentic intelligence, and decision support. At the edge level, several distributed sensors, smart devices, and urban infrastructures record specific types of data regarding transportation, energy, environment monitoring, and public services. The data remain available at the edge nodes only. Each node participates in the process by building a local model based on its own data, but all that gets transmitted to the aggregation server are local updates. The aggregation server integrates all the local updates to create a global model which is then sent to all participating nodes.

The results of local and global models are put through the explainability layer so that understandable information regarding the reasoning of the models can be obtained. The models' outcomes are then communicated to the autonomous agents that make decisions that lead to fulfilling the goals based on the current state of the system and the needs of the application. Then their results are delivered using the decision support block.

The stages of the whole process are: (1) collecting data from remote edge nodes, (2) training local models making use of domain-specific data, (3) sending model parameters to the aggregation server, (4) performing federated aggregation for an updated global model, (5) distributing the global model to participating nodes, (6) generating explanations of model output, (7) performing agent-based reasoning on the basis of the model results and explanations and (8) making adaptive decisions for smart city applications. The described cycle functions multiple times to adapt to the changing realities of the city and evolving operational requirements.

### 3.2 Algorithm Description

The method introduced is based on the principles of iterative learning and decision-making processes:

1. Create local models for each edge device.
2. Develop local models using domain-specific data.
3. Send model parameters to the aggregation server.
4. Conduct federated aggregation to modify the global model.
5. Deliver modified global model and distribute it to the nodes.
6. Apply explanation algorithms to obtain model solutions.
7. Perform agent-based decision-making through model analysis.
8. Produce adaptive decisions in the context of smart city application.

The iterative cycle continues until the target requirements are fulfilled.

### 3.3 Mathematical Model

Let there be N distributed nodes, each with local dataset $D_i$, where i=1,2,…,N.

**Local Model Training**

Each node optimizes a local objective function:

$$\min_{w_i} F_{i(w_i)} = \frac{1}{|D_i|} \sum_{(x,y)\in D_i} L(w_i;\ x,y)$$

**Global Model Aggregation**

The global model is obtained through weighted aggregation:

$$w^{t+1} = \sum_{i=1}^{N} \left( \frac{|D_i|}{\sum_{j=1}^{N} |D_j|} \right) w_i^t$$

**Explainability Function**

An interpretability function E(·) maps model outputs to explanations:

$$E(w,x) \rightarrow \text{interpretable output}$$

**Agentic Decision Function**

An agent selects action aaa based on state sss, model output, and explanation:

$$a = \pi(s,w,E(w,x))$$

where π represents the policy governing goal-driven decisions.

## 4. System Architecture

Figure 1 shows the recommended Federated Explainable Agentic AI system for smart city decision-making. The system comprises five layers that are interrelated and make possible the conversion of decentralized city data into transparent and flexible decisions.

**Smart City Edge Data Layer:** Various sensors, smart devices, and urban infrastructures gather diversity of information from different fields like traffic, energy, ecology, and public services. The information is processed locally making it possible to achieve decentralization and ensuring privacy.

**Federated Learning Layer:** Every edge point creates a local model using the available data. All that is shared with the common server are those updates of the model which are then processed by the server to vary the global model that will be shared with all edges.

**Explainability and Governance Layer:** Both local and global models will enhance the usability of the information.

**Agentic Intelligence Layer:** Independent agents will make use of the data obtained from the local and global models to make decisions in accordance with the current city situation.

**Intelligent Decision Support Interface:** The output of the decisions rendered is transformed into operations, allowing for monitoring, forecasting, planning, optimization, and decision-making activities.

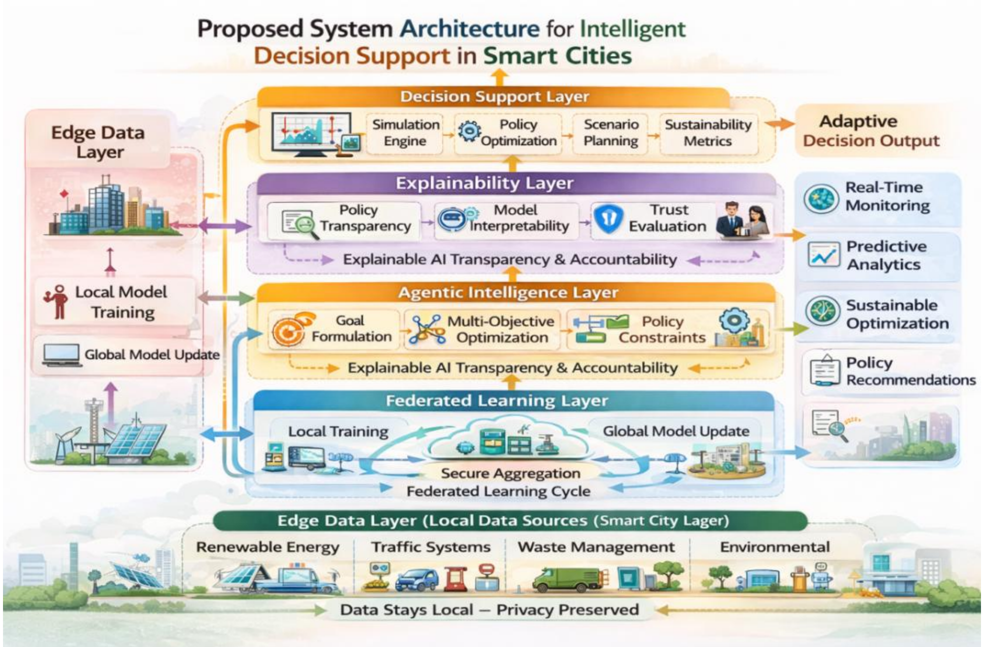


**Figure 1:** Proposed Federated Explainable Agentic AI Architecture for Smart City Decision Support

The five layers are designed to work together in a seamless manner, and through this approach local data is collected and processed, knowledge is learned through aggregation that uses data generated in several different locations, results of various models are interpreted, and intelligent agents arrive at optimal solutions for decision-making. Thanks to this architecture, the above-mentioned qualities of confidentiality, transparency, scalability, and adaptability are maintained within one smart city intelligence framework.

## 5. Workflow

The described workflow is an iterative process in which smart city data from various sources is transformed into adaptive and transparent decision support outputs thanks to federated learning, explainable reasoning, and agentic intelligence. The workflow is composed of nine stages.

Collecting information from edge nodes: Data is collected at local nodes, which consist of various decentralized sources such as transportation, energy, environment, public service, and other smart city systems.

**Local model training:** Individual edge nodes are able to use local facility-specific information for training its local model.

**Model parameter exchange:** The set of parameters produced by the local model is sent to the aggregation server.

**Federated aggregation:** The aggregation server combines the local model updates into one single model.
**Global model sharing:** The global model is sent back to edge nodes.
**Explaining prediction results:** Both local and global model outputs are processed so that they are able to provide some interpretation about the predictions or explanations being made.
**Agents performing decisions:** Agents utilize the results of the model along with the explaining model so that they can make the desired decisions in this environment.
**Generating outputs for supporting decision making:** The outputs are produced by the models to provide aid for monitoring, prediction, optimization, planning, and decision making purposes.
**Providing feedback for adapting to conditions:** New information helps adapt to the constantly changing conditions of the urban environment.

The workflow runs all the time, permitting the framework react to changes in urban conditions and also keeping decentralized learning, transparent reasoning, and adaptable decision-making.

## 6. Experimental Design

The framework has been designed to be used for its evaluation with representative smart city data in the areas of transportation, energy, environmental monitoring, and public services. These data were distributed among multiple edge nodes in order to reflect the heterogeneous and decentralized environment of smart cities. The categories of data and their purposes of evaluation are summarized in Table 2.

The evaluation of the framework comprises comparative evaluations of the new FEA-AI approach with centralized, distributed learning, and federated learning systems. It distributes the data across multiple nodes, implements baseline and newly developed methods, performs federated learning and aggregation, uses explainability and agent-based decision-making methodologies, and analyzes the resulting system behavior in comparison with the previous ones.

**Table 2:** Summary of Smart City Datasets Used for Evaluation

| Dataset Type | Data Source | Characteristics | Purpose |
|---|---|---|---|
| Transportation Data | Traffic sensors | Time-series, dynamic | Traffic prediction |
| Energy Data | Smart grids | Consumption patterns | Energy optimization |
| Environmental Data | IoT sensors | Air quality, temperature | Pollution monitoring |
| Public Service Data | Urban systems | Event-based | Service efficiency |

## 7. Expected Performance and Discussion

According to the proposed FEA-AI framework, fragmented smart city intelligence will be dealt with by its incorporation of decentralized learning, explainable reasoning, and autonomous decision-making into an integrated design. This will allow it to save data at distributed edge nodes rather than collecting raw data, thus helping to maintain enhanced privacy while allowing for collaborative learning across different types of urban environments [2], [10], [11].

It is presumed that decentralized architecture will ensure scalability by spreading computing tasks across participating nodes and enabling reduced reliance on centralized processing. Explainability paved the way for interpretation of outcomes, thus promoting transparency, responsibility, and credibility for the users.

The proposed framework is made to lessen the risks of privacy and reliance on one center for information processing, when contrasted with centralized methods. However, unlike conventional forms of decentralized technology, via federated aggregation, the collaborative use of information becomes possible due to the incorporation of explanation mechanisms and agents' capabilities into classical federated learning.

On the whole, the proposed framework offers a unified foundation for the emergence of smart city intelligence which is both adaptive and aware of privacy concerns. The thorough validation of the framework can be performed in the course of real-life experiments or simulations through the research design implemented in the present research.

**8. Conclusion and Future Scope**

To sum up, this paper presents a novel proposal for the unified Federated Explainable Agentic AI framework to close significant knowledge gaps in smart city intelligence. The proposed paradigm uses decentralized federated learning, explainable reasoning, and autonomous agent-based decision-making to facilitate the privacy-conscious, scalable, transparent, and efficient decision-making process in smart cities.

The framework carries out a predefined structure to utilize the edge data for transforming available observations into interpretable decisions by employing collaborative learning along with explanatory and agentic reasoning.

The next stage of the research will be focused on testing the framework through experimental studies in the real-world conditions and simulation-based environments. Additional studies will also cover the issues of communication efficiency, sophisticated explanation, agent cooperation, and applications in various technologies used in smart cities.

**References:**


[1] Zhang, L., et al. "Edge computing for internet of things: a survey." *IEEE Internet of Things Journal* 7.7 (2020): 6379-6394.

[2] Khan, Latif U., et al. "Federated learning for internet of things: Recent advances, taxonomy, and open challenges." *IEEE Communications Surveys & Tutorials* 23.3 (2021): 1759-1799.

[3] Nguyen, Van-Dinh, et al. "Efficient federated learning algorithm for resource allocation in wireless IoT networks." *IEEE Internet of Things Journal* 8.5 (2020): 3394-3409.

[4] Javed, Abdul Rehman, et al. "A survey of explainable artificial intelligence for smart cities." *Electronics* 12.4 (2023): 1020.

[5] Kostopoulos, Georgios, Gregory Davrazos, and Sotiris Kotsiantis. "Explainable artificial intelligence-based decision support systems: A recent review." *Electronics* 13.14 (2024): 2842.

[6] Ullah, Zaib, et al. "Applications of artificial intelligence and machine learning in smart cities." *Computer communications* 154 (2020): 313-323.

[7] Bosco, Gerardo, et al. "AI-driven innovation in smart city governance: achieving human-centric and sustainable outcomes." *Transforming Government: People, Process and Policy* 20.1 (2026): 98-123.

[8] Dubey, Praveer, and Mohit Kumar. "Integrating Explainable AI with Federated Learning for Next-Generation IoT: A comprehensive review and prospective insights." *Computer Science Review* 56 (2025): 100697.

[9] Ragab, Mahmoud, et al. "Advanced artificial intelligence with federated learning framework for privacy-preserving cyberthreat detection in IoT-assisted sustainable smart cities." *Scientific reports* 15.1 (2025): 4470.

[10] Fernández, J. Delgado, et al. "Scaling smart cities with federated learning." *Business & Information Systems Engineering* (2025).

[11] Alterkawi, Laila, and Fadi K. Dib. "Federated Learning for Smart Cities: A Thematic Review of Challenges and Approaches." *Future Internet* 17.12 (2025): 545.

[12] Pati, Ashis Kumar. "Agentic AI: a comprehensive survey of technologies, applications, and societal implications." *IEEE Access* (2025).

[13] Abou Ali, Mohamad, Fadi Dornaika, and Jinan Charafeddine. "Agentic AI: a comprehensive survey of architectures, applications, and future directions." *Artificial Intelligence Review* 59.1 (2025): 11.

[14] Tunduny, Titus, and Bernard Shibwabo. "Explainable AI Approaches in Federated Learning: Systematic Review." *JMIR AI* 5 (2026): e69985.