

An Introduction to Artificial Intelligence of Things for Smart Policing and Criminal Investigation

Chen-Yu Li 

Assistant Professor, Department of Information Management, Central Police University, Taoyuan, Taiwan

Correspondence should be addressed to Chen-Yu Li; cyli@mail.cpu.edu.tw

Received: 25 Sep 2026; **Revised:** 9 Oct 2026; **Accepted:** 15 Oct 2026; **Published Online:** 26 Oct 2026

Copyright © 2026 Chen-Yu Li. This is an open-access article distributed under the [Creative Commons Attribution 4.0 International License \(CC BY 4.0\)](https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

ABSTRACT- Artificial intelligence of things (AIoT) combines connected sensors with AI models to interpret observations from the physical world. We review its use in smart policing and criminal investigation across public safety, traffic management, search and rescue, officer safety, and environmental investigation. We then examine how large language models (LLMs) might extend these applications by interpreting requests, retrieving operational knowledge, and coordinating tools. An incident-response scenario follows a case from intake and dispatch through travel assistance, field sensing, and command-center coordination. The scenario separates language interpretation from sensor analysis, numerical optimization, and decisions reserved for police personnel. The design must account for ambiguous speech, device interoperability, mobile computation, unreliable information, security, privacy, and incident-state consistency during network outages. We identify research questions and outline controlled comparisons and ablation studies to assess the contribution of LLMs. This is an introductory review and a conceptual design; operational benefits remain to be established against current procedures and equivalent AIoT systems without LLM agents.

KEYWORDS- AIoT, Criminal Investigation, Internet of Things, Large Language Models, Smart Policing

I. INTRODUCTION

Police dispatch and criminal investigation depend on information gathered from several sources, often under time pressure. Connected cameras, wearable devices, and environmental sensors add field observations to this process. The growth in IoT connections [1] and forecasts of their economic impact [2] indicate the scale of connected sensing. For police use, the central issue is how an observation reaches the right person, with enough context to guide a response.

The Internet of Things (IoT) provides sensing and connectivity. Artificial intelligence (AI) interprets the resulting data through recognition, prediction, and other analytical methods. Their combination is known as artificial intelligence of things (AIoT) [3]. Developments in edge computing, networking, and AI [4]-[10] have widened the available processing options. Police applications need an arrangement that delivers results

within the required time and makes their meaning clear to officers and command personnel.

An AIoT system must connect a measurement to a specific task. A camera image may be processed into a plate identifier for a vehicle check; road measurements may be used to forecast traffic and guide a control decision. In both cases, sensor quality and model error affect the information available to the officer. The design also needs to specify how that information is presented and what action it can reasonably inform.

This article presents a narrative review of AIoT for policing and criminal investigation, with an emphasis on LLM integration. The review focuses on literature relevant to sensing and communication architectures, task-specific analysis, and language-based interaction and tool coordination. Foundational IoT and AIoT research [3], [11], [21] provides the technical context, while application and deployment literature [23]-[29] and research on LLM-enabled systems [30]-[36] inform the discussion of operational requirements and integration challenges. The review is selective in scope, and the proposed incident-response workflow remains a conceptual design requiring empirical evaluation.

The remainder of this article is organized as follows. Section II outlines the sensing, communication, and computational foundations of AIoT. Section III reviews five policing application areas, describing their data, analytical methods, and deployment requirements. Section IV examines LLM integration through an incident-response scenario and discusses the associated research challenges. Section V concludes with priorities for evaluation.

II. TECHNICAL FOUNDATIONS OF AIOT

Designing an AIoT application involves choices about sensing, data transmission, computation, and use of the results. This section introduces the components needed to make those choices in a police setting.

A. Internet of Things

IoT connects physical devices with sensing, computing, communication, or actuation functions [3], [11]. Police deployments may use fixed cameras, vehicle equipment, wearables, or environmental monitoring stations. Their outputs include images, positions, physiological signals, and pollutant measurements. Gateways and application

services receive these observations for storage and use by authorized personnel.

A common IoT architecture has perception, network, and application layers [11], as shown in Figure 1 [12]. Sensors in the perception layer acquire the observations. Platforms such as Raspberry Pi, ESP32, and Arduino are useful for prototyping, with the choice determined by the sensor interfaces and processing load. Image acquisition needs adequate resolution and processing capacity; environmental measurement needs sensors suited to the quantity being measured.

The network layer transports data between devices, gateways, and services. Bluetooth, Wi-Fi, LTE, and low-power wide-area networks offer different coverage, delay,

and power characteristics [11]. A deployment must match these characteristics to its data volume and reporting frequency. A short sensor message and a continuous video stream impose substantially different demands. Gateways connect devices with different interfaces, while backhaul links carry their data to a remote service or command center.

The application layer stores and displays data, manages alerts, and connects to operational systems [11]. A police application might place an observation on a map, associate it with an incident, or record the response. Data acquisition alone does not complete the task. The application must present the observation in a form that an officer can interpret and use.

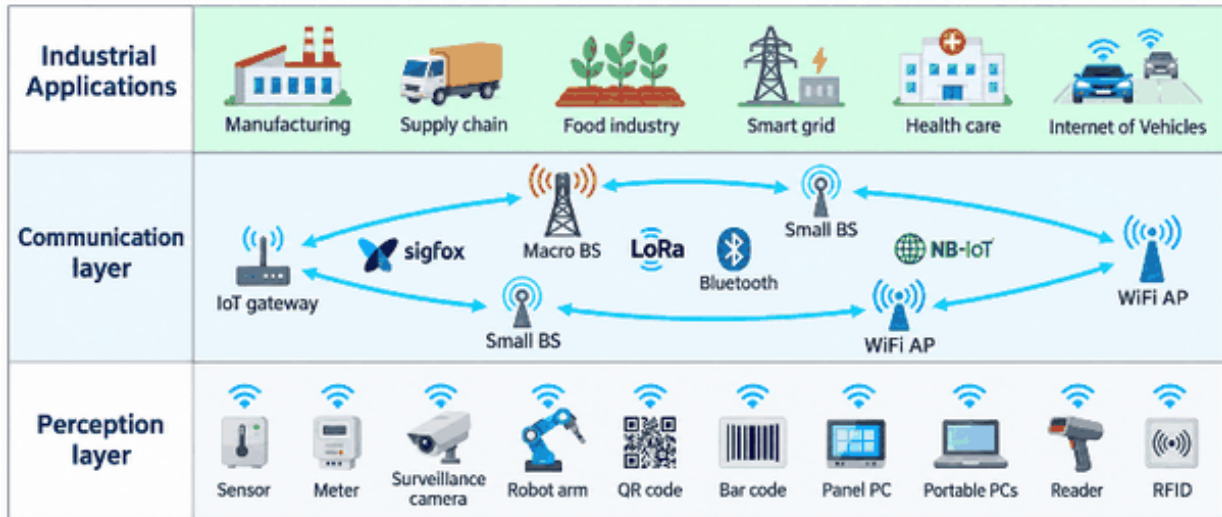


Figure 1: Three-layer IoT architecture [12]

B. Artificial Intelligence

AI covers computational methods for recognition, prediction, reasoning, and decision support. Machine learning (ML) develops models from data [13], [14]. The intended output should be defined before selecting a model. Vehicle recognition, traffic forecasting, and detection of unusual sensor readings require different training data and evaluation criteria.

ML includes supervised, unsupervised, and reinforcement learning [15]. Supervised learning uses labeled examples for classification or prediction. Unsupervised learning finds patterns in unlabeled data, including clusters and unusual observations. Reinforcement learning develops a policy from interaction with an environment and feedback on actions, often using simulation to explore control strategies. The available data and the operational objective determine which approach is appropriate.

Deep learning uses multilayer neural networks to learn representations from data [16]-[18]. Convolutional neural networks (CNNs) extract image features, and long short-term memory (LSTM) networks model sequences. Graph neural networks (GNNs), including graph convolutional networks (GCNs), represent relationships among connected entities. These methods offer possible approaches to camera analytics, traffic or physiological monitoring, and analysis of road networks, respectively.

Transformer-based generative adversarial networks (GANs) offer further methods for representation learning and data generation [19]. Their use in policing depends on the input data and the required output. Predictive performance must be considered alongside computational cost, robustness, and whether officers can interpret the results.

C. Artificial Intelligence of Things

AIoT applies AI models to data acquired and communicated by IoT devices [3], [20], [21]. Processing may run on a device, at an edge gateway, or in the cloud. Local processing can produce results near the sensor, while back-end services can combine observations from several locations. Placement depends on delay, connectivity, computing resources, and data-handling requirements. Relaying a measurement or applying a fixed threshold does not itself require AI; the AI component must have a defined analytical role.

Healthcare, intelligent vehicles, robotics, UAVs, and industrial systems provide examples of this combination [6], [7], [20]-[22]. They connect measurements to an analytical task and then to a decision. The police scenarios below use the same structure, separating functions documented in the cited work from extensions suggested here.

III. POTENTIAL AIOT APPLICATIONS IN POLICING

The five application areas below describe what is sensed, how the data are analyzed, and how the results might be used. Figure 2 Figure 3 Figure 4 Figure 5 Figure 6

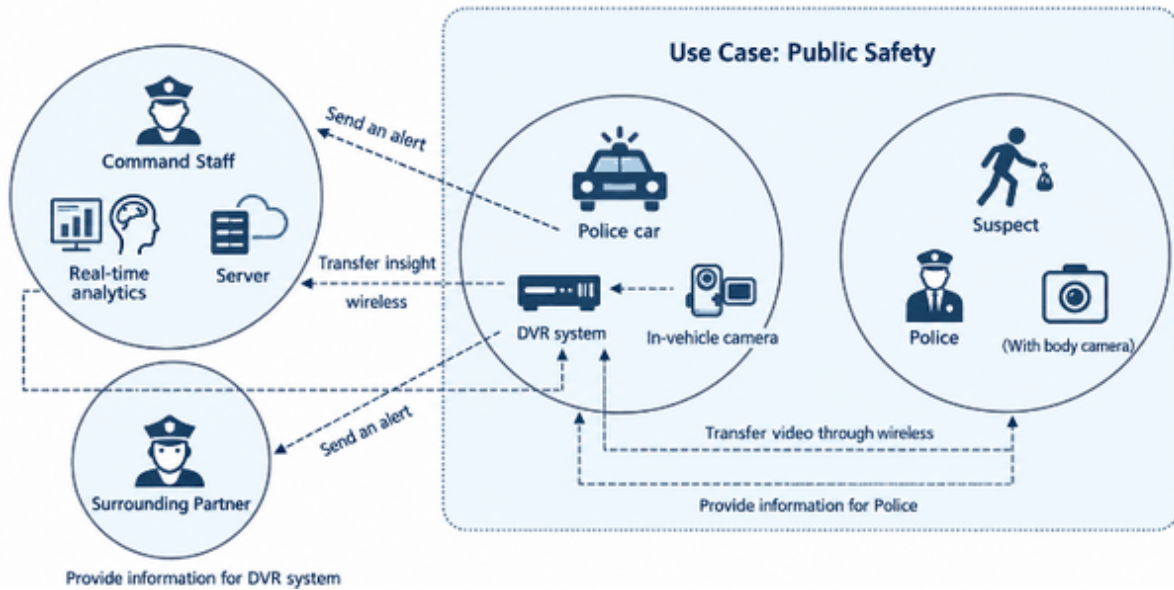


Figure 2: Architecture for public safety monitoring [23].

Connected cameras and image models can assist vehicle identification and public safety monitoring. Figure 2 shows an architecture that links field information and analytical services to police personnel [23]. Automatic license plate recognition converts vehicle images into identifiers for comparison with records. Video models can flag activity for review, but a match or activity label still needs assessment in context. Cameras may be fixed or vehicle-mounted, with processing on the device, at the edge, or in a

back-end service. Mobile use introduces motion, occlusion, changing illumination, and unreliable connectivity. Evaluation should therefore measure recognition errors, false alerts, reporting delay, and the workload of reviewing outputs.

B. Traffic Management

Traffic models use road observations to estimate congestion or forecast demand.

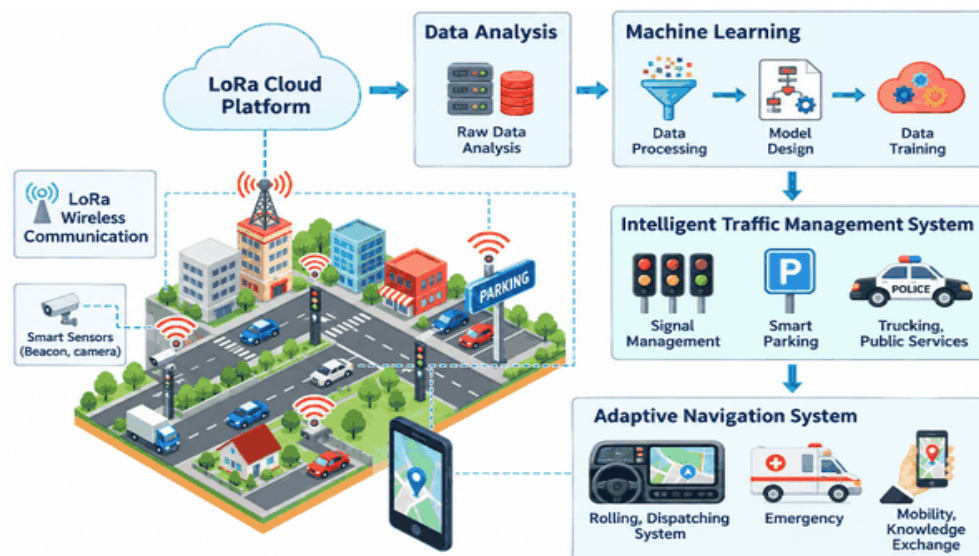


Figure 3: Traffic management using LoRa and machine learning [24]

Figure 3 illustrates traffic management using LoRa communication and ML [24]. Roadside equipment can supply vehicle counts or speed measurements, depending on its sensors. Models use patterns over time and across locations to inform signal planning and route guidance. LoRa may carry short sensor reports, whereas camera

streams need a different assessment of network capacity. Evaluation should compare travel delay and prediction error with existing procedures, including missing data and reporting delays. A route recommendation also needs network-level assessment because reducing delay for one vehicle can shift congestion elsewhere.

C. Emergency Search and Rescue

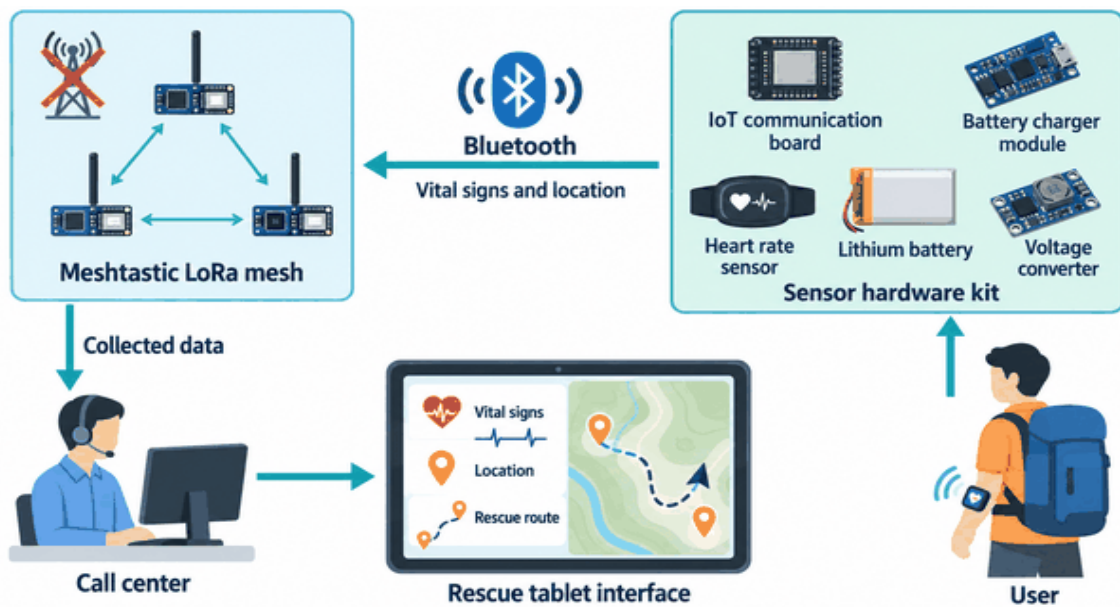
Mountain search and rescue requires communication, location reporting, and information about a casualty's condition when cellular coverage is unavailable. Chung et al. [25] propose a wearable IoT-based system that combines Meshtastic LoRa mesh communication with physiological monitoring, as illustrated in Figure 4. The proposed nodes use GPS to obtain coordinates and exchange location reports and text messages directly or through neighboring nodes. Android Team Awareness Kit (ATAK) provides a map-based interface for displaying personnel positions, logged messages, and disaster

conditions. The local Meshtastic mesh does not depend on the cellular coverage required by NB-IoT [26] or the conventional LoRaWAN gateway-to-server arrangement [27].

This design provides a sensing and communication basis for AIoT, allowing information to be relayed to rescue teams or a call center through available mesh links. An AI extension could combine terrain, last-known positions, physiological trends, and incoming observations to prioritize search areas or routes and identify events requiring review.



(a) Mountain rescue application scenario with wearable monitoring and LoRa mesh communication.



(b) Wearable sensing, LoRa mesh communication, and rescue information display.

Figure 4: Wearable IoT-based mountain rescue system [25]

D. Officer Safety

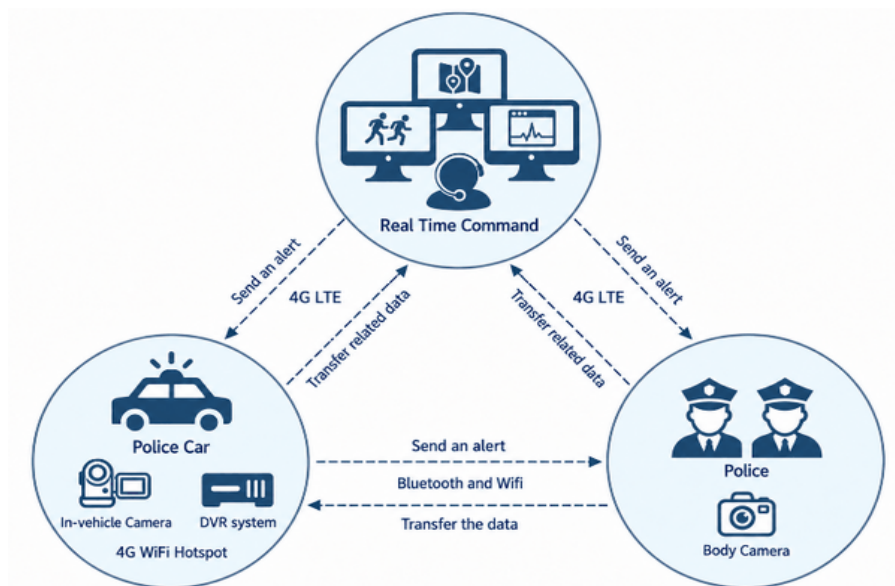


Figure 5: Field-to-command communication for officer safety [23].

Wearable measurements, motion data, and speech can help command personnel assess an officer's situation. Figure 5 shows the exchange of field information with a command function [23]. An analytical service could identify combinations of observations that warrant an alert. Heart rate, for example, must be interpreted against activity and the officer's baseline because exertion and stress may produce similar changes. A body-worn camera can add visual context, although timely transmission depends on the connection. The value of combining these signals needs to be established under field conditions.

Defensive-equipment notifications and spoken requests for assistance provide further context. Recording equipment use is an IoT function; an AI model could interpret that event with other observations or recognize a request in noisy audio. Command personnel would assess the alert and decide on assistance. Relevant measures include missed events, false alarms, speech-recognition errors under operational noise, reporting delay, and review workload.

E. Technology Assisted Criminal Investigation

Environmental measurements and camera observations can help organize investigative leads. One possible use is identifying unusual patterns near a suspected illicit drug manufacturing site. The PM-monitoring architecture in Figure 6 [28] illustrates environmental data collection and analysis. Detecting drug-related substances requires separately selected and validated sensors, with attention to calibration, interference, background sources, and uncertainty [29]. General water-quality anomalies likewise need appropriate analytical methods and, where necessary, confirmatory sampling to identify a substance. Temporal or multimodal analysis could rank events for review by combining environmental patterns with contextual observations. Fixed-threshold alarms can also notify personnel, without constituting AI analysis. Since legitimate activities can produce similar anomalies, the result should be a lead accompanied by its observations and uncertainty. Corroboration remains necessary. Evaluation should test whether ranking improves lead assessment while controlling false positives under realistic background conditions.

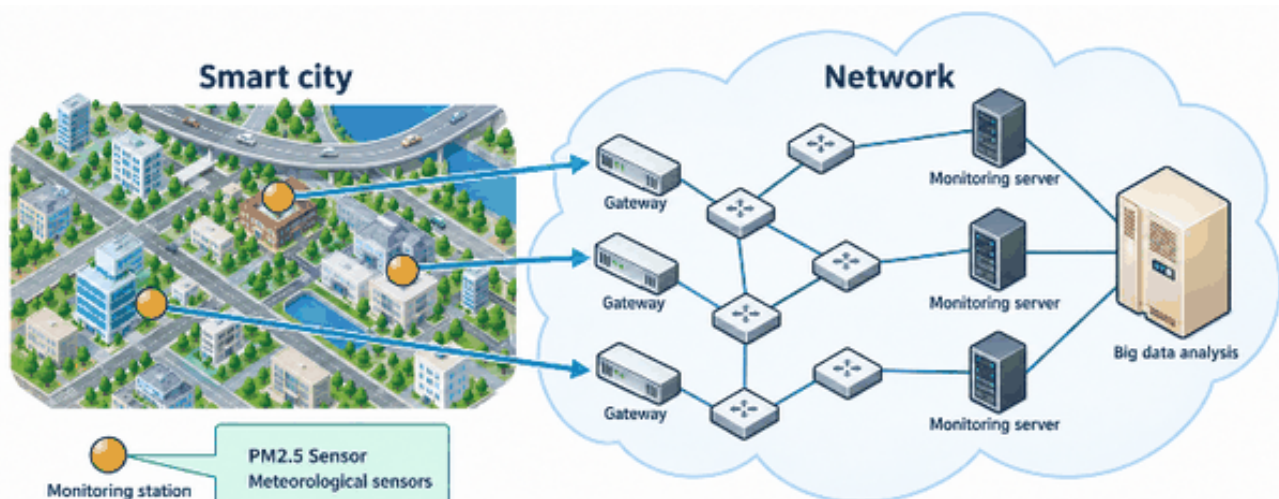


Figure 6: Environmental sensing and PM analysis [28].

IV. INTEGRATING LLMs WITH AIOT FOR POLICING

The applications in Section III analyze sensor data for defined tasks. LLM integration [30] offers a way to connect these outputs to officers' requests, case knowledge, and available services as an incident develops. It also places language interpretation in the path between a report and a tool call, creating opportunities for errors to affect later actions. We examine this integration through an incident-response scenario. The workflow is a conceptual design informed by research in other application domains; it has not been evaluated in police operations.

A. LLM Roles and Operational Control

An officer's request may refer to the current incident, omit a device name, or require several services in sequence. An LLM agent could interpret the request in context, retrieve relevant case categories and guidance, and choose among services such as unit-status queries and route planning. SAGE [31] provides a precedent in smart homes, where an autonomous LLM agent interprets flexible requests and resolves references to devices. The police scenario applies this idea to a changing incident record and a restricted set of operational services.

LLM agents can combine planning, tool use, memory, and task execution [30]. Here, the agent would organize information, retrieve authorized records, query read-only services, and prepare recommendations. Assigning a unit, sending backup, or changing a deployment would remain subject to approval by authorized personnel. Permissions must be enforced outside the model. Section IV-C discusses the security basis for these controls; their effectiveness would still need to be tested.

Information captured at intake should remain available through dispatch and field response. New field observations can revise the briefing and inform command decisions, with links to the original report.

Cameras, microphones, location sensors, and motion sensors acquire observations, and specialized models extract speech, visual, and activity features. The LLM organizes these results with case records, approved guidance, and current unit status. Dispatch and routing tools calculate options under defined constraints. Police personnel decide on the response.

The practical hypothesis is that this arrangement would reduce repeated searches across case history, traffic data, and unit-status systems. Speech recognition, speech synthesis, and data integration are available without LLMs. The empirical question is whether context-sensitive interpretation and selection of service sequences add enough benefit to justify the LLM's cost, delay, and verification burden. Comparisons must use the same sensors and underlying services.

B. Incident Response Scenario

Below Table 1 sets out the agent and personnel roles at four stages of this conceptual scenario.

1. **Incident intake and dispatch-** At intake, an agent would extract the reported location, event, people involved, urgency indicators, and unresolved questions from a call or an in-person report. Retrieval-augmented generation (RAG) provides access to past case categories and current response guidance [30]. Past cases would be reference material, with

recommendations checked against current procedures. A dispatch tool would combine confirmed incident details with live traffic, vehicle availability, officer qualifications and training, and current assignments. Unconfirmed locations and resource states would remain marked as uncertain. The dispatcher would review the options and missing information, then approve or revise the assignment.

2. **Travel assistance-** Once a unit is assigned, a mobile interface would deliver a short incident briefing and spoken route updates from the command-center agent. A routing service would calculate the route from current road conditions and applicable constraints. Minimum distance and minimum estimated arrival time are different objectives, so the interface must identify which is used and label travel times as estimates. The agent would report the service's results. In electric-vehicle charging, Zhang et al. [32] use RAG to translate natural-language objectives and constraints into a mixed integer linear programming formulation and executable solver code. We adopt the same division of responsibility between language interpretation and numerical optimization as a design principle for the police scenario.
3. **On-scene sensing and interpretation-** At the scene, a smartphone, portable or body-worn video recorder, and connected sensors would supply authorized audio, video, location, and motion data. Speech-recognition and sensor models would produce timestamped transcripts and event descriptions. The agent would relate an officer's request to the active incident, identify missing information, and prepare a situation update, extending Section III-D. Interpretation needs the speaker, preceding dialogue, mission status, and sensor context. Bystander speech, quotations, and background voices must be distinguished from authenticated officer commands. An alert about danger would remain an assessment to be confirmed where feasible.
4. **Command-center coordination-** Command personnel would receive summaries of new observations, changes since the previous update, and possible requests for assistance. Validated urgent alerts must have a path independent of lengthy LLM generation. Authorized personnel would decide whether to send backup, request medical assistance, or change the assignment. Requests and acknowledgments would be logged as the incident record develops. Original recordings and sensor data would remain separate from the summaries, with links that allow each update to be checked against its sources.

Table 1: Agent and Personnel Roles in the Incident Response Scenario

Stage	Main inputs	Proposed agent role	Personnel role
1. Intake and dispatch	Live report; past incident types and handling methods; current traffic; vehicle status; officer capability; task load	Structure the report, retrieve similar cases, and explain dispatch options computed by a dispatch tool	Dispatcher approves or revises the assignment
2. En route	Vehicle position; road conditions; updated report	Present a concise briefing and spoken route updates computed by a routing service	Officer drives and acknowledges; command center monitors
3. On scene	Audio, spoken commands, and motion data from a smartphone or portable recorder	Associate requests with the active incident and assemble a situation update	Officer retains control of the scene; confirms or cancels alerts
4. Command coordination	Relayed updates; unit availability; records	Summarize changes and present a possible need for assistance	Commander decides; the record supports review

C. Integration Challenges

A misinterpreted report can lead to the wrong record being retrieved and then to an unsuitable service call. Errors thus cross the boundaries between sensing, interpretation, and action. Some constraints come from IoT infrastructure; others arise from the LLM's reasoning and tool use. Table 2 lists design options and measures for examining these problems.

- 1. Context-dependent language:** Field speech varies with speaker, accent, noise, stress, and rapidly changing events. The difference between “send another unit” and “we do not need another unit” depends on reliable recognition of negation and context. SAGE [31] handles flexible requests and device disambiguation in smart homes. With GPT-4, it achieved 76% success on a 50-task smart-home benchmark, compared with 30% for the best tested LLM-enabled baseline. These results apply to that task set and configuration. Police encounters add uncertain speakers and consequences for field response. The interface needs clarification, explicit uncertainty labels, and a manual communication option. Testing should record missed assistance requests, false alerts, and incorrect associations between speech and incidents.
- 2. Device interoperability and operational state:** Recorders, phones, vehicle systems, and sensors differ in protocols, functions, and data formats. Costa et al. [33] demonstrate local discovery and peer-to-peer interoperability in a consumer-IoT prototype. A police implementation would additionally need consistent device identities, timestamps, measurement units, capability descriptions, and authenticated interfaces, together with command coordination. The LLM must map a request to a function the device actually supports in its current state. External checks must verify permissions and arguments, then confirm the result. A requested operation and a completed operation must be recorded as different states.
- 3. Mobile edge deployment and delay.** Officer devices and patrol vehicles move between connections and computing resources. Appropriately equipped vehicles or nearby infrastructure can host edge processing. Liu et al. [34] jointly study LLM deployment and task offloading under stochastic requests, evaluating the approach through simulations of dynamic networks. Police workloads would add variation in retrieval and tool-call sequences as incidents develop. Local feature

extraction and compact models are candidates for time-sensitive tasks; heavier analysis could use authorized edge or back-end resources when available. Delay must be measured over the full path from sensing and transmission through queueing, retrieval, tool execution, and generation. Tail latency, missed deadlines, energy use, and operation during outages all matter.

- 4. Resources and information reliability.** Section IV of [30] discusses resource constraints, model optimization, prompting, fine-tuning, retrieval, evaluation, truthfulness, ethics, and security. In a police application, compression or selective transmission may discard information that changes the meaning of a report. Retrieved records need sources and version or validity information, and briefings must separate observations from estimates and missing data. Fine-tuning and retrieval alone cannot ensure factual accuracy. The survey [30] reports SensorQA mobile-sensor question-answering accuracy as low as 31.22% for certain configurations and describes difficulties in reasoning over long periods. That figure is a result from selected SensorQA configurations, reported through the survey, rather than an accuracy estimate for police tasks or LLMs generally.
- 5. Agent and network security.** Chhabra et al. [35] review attacks on agent planning, tool use, memory, and interaction with external content, as well as defenses. In this scenario, transcripts of scene audio, retrieved records, and device messages are untrusted inputs. Their contents must not change tool permissions or workflow rules. Controls outside the model should authenticate commands, restrict actions, validate execution, and keep an audit record. Huang et al. [36] evaluate cloud-edge collaborative DDoS detection in LLM-enabled AIoT using specified traffic datasets and network scenarios. Those results require testing on police networks before operational use. DDoS detection addresses attacks on communication availability, whereas command authentication and tool authorization address misuse of the agent. Both network availability and the integrity of agent-selected actions need evaluation.
- 6. Privacy, governance, and accountability.** Continuous recording can capture bystanders and victims as well as officers. The survey [30] describes exposure through cloud or third-party services and sensitive information

inferred from combined sensor data. Costa et al. [33] retain context data within the local network in their consumer-IoT architecture. Local processing can limit transmission, but access and retention controls remain necessary. A deployment must define the data collected, what is transmitted, who can access it, and how long it is retained. Generated summaries need links to the original recordings. Investigative assessments should be checked against those observations.

7. Incident context and intermittent connectivity.

Reports may reach the command center after the event

they describe. Each observation, recommendation, and acknowledgment therefore needs an incident identifier, source, event time, and version or status. The time of recording or generation must remain distinct from the event time. After reconnection, pending requests must be reconciled with acknowledgments before an earlier recommendation is acted on. The unresolved problem is how to preserve useful incident context without allowing stale information or duplicate requests to produce inappropriate actions.

Table 2: Integration Challenges and Evaluation Criteria

Challenge	Operational concern	Design options	Example measures
Context-dependent language	Negation; noisy or stressed speech	Clarification; uncertain-event labels; manual fallback	Missed assistance requests; false alerts; speech-to-incident association errors
Interoperability and state	Mixed vendors; requested versus completed operations	Authenticated interfaces; external validation; acknowledged state	Integration effort; execution-check failures
Mobile edge deployment and delay	Variable connectivity; variable tool-call sequences	Place functions on device, edge, or back end by urgency	Tail latency; missed deadlines; energy; operation when disconnected
Resources and reliability	Hallucination; lost critical information	Sourced and dated retrieval; separate observations from estimates	Briefing accuracy; factual error rate; override rate
Agent and network security	Injected instructions; tool misuse; attacks on availability	Permissions outside the model; authenticated commands; audit records	Injection success rate; detection performance on police networks
Privacy and accountability	Bystander recording; opaque reasoning	Local processing; retention limits; decision logs	Leakage incidents; audit completeness
Incident context and connectivity	Stale observations; delayed acknowledgments; duplicate requests	Incident IDs; event times; versioned state; reconciliation	Stale-action rate; duplicate actions; reconciliation delay

D. Research Questions and Evaluation

The scenario raises six research questions.

1. **Incident context.** How can an agent maintain a consistent, correctable incident record when observations and assignments change at different rates, edge and cloud outputs disagree, or a connection fails during a response?
2. **Computation placement.** Where should inference, retrieval, and tool execution run across officer devices, vehicle computers, edge servers, and back-end services to meet response deadlines?
3. **Tool authorization.** How can the system respond to varied requests while ensuring that each consequential action is authorized, verified, and traceable to its inputs?
4. **Human decision authority.** Which decisions can the agent recommend or execute, and which require a person? The boundary needs to account for urgency, reversibility, and the consequences of error, with evaluation involving dispatchers and field personnel.
5. **Need for an LLM.** Which functions gain from language reasoning, and which can be handled by a smaller, validated model at lower energy, delay, and verification cost [30], [34]?
6. **Police-task benchmarks.** The evaluations discussed in [30] and [31] concern mobile-sensor questions and

smart-home control. A benchmark for incident intake, dispatch, and field interpretation needs defined tasks, independent reference labels, ambiguous cases, and explicit completion criteria. Data collection and annotation also need privacy controls and independent review.

A controlled study would compare current procedures, AIoT without an LLM agent, and the integrated system using the same sensor inputs, case records, dispatch and routing tools, and resource conditions. Ablation studies could remove retrieval or agent-based tool coordination to measure their separate contributions. Reference labels must be assigned independently of model outputs. In addition to task success, the study should record unsupported statements, unnecessary actions, clarification requests, and operator corrections.

Evaluation should measure dispatch appropriateness, briefing accuracy, arrival-time estimation, assistance-request detection, operator workload, and task completion. Test cases need incomplete reports and ambiguous speech, congestion and changes in unit availability, device faults and disconnections, and malicious inputs. Human intervention, driver attention, access control, and traceability of summaries must be assessed with model performance. Any gain in coordination should be weighed against delay, cost, and uncertainty.

V. CONCLUSION

The applications reviewed here connect field observations to police decisions through sensing, communication, and analysis. A plate identifier, traffic forecast, safety alert, or environmental anomaly requires contextual assessment before it informs a police decision. Sensor quality, model error, and the way results are presented remain central to that assessment.

LLM integration could make the interaction less dependent on fixed commands and link incident information to relevant knowledge and services. The scenario in Section IV assigns the LLM interpretation and coordination tasks, leaves numerical calculations to dedicated tools, and reserves consequential decisions for authorized personnel. Its operational benefits remain untested. Ambiguous speech, uncertain device state, disconnected operation, and untrusted inputs are specific failure modes that a police implementation must address.

Further work should test incident interpretation, continuity of the incident record during outages, and verification of tool calls in a controlled prototype. Comparisons with existing procedures and equivalent AIoT systems would establish whether LLM assistance improves response quality or reduces workload. Latency, energy use, operator corrections, and traceability must be measured alongside those outcomes.

REFERENCES

- [1] Lionel Sujay Vailshery, "Number of Internet of Things (IoT) connections worldwide from 2022 to 2023, with forecasts from 2024 to 2033," Statista.
- [2] M. Chui, M. Collins, and M. Patel, "The Internet of Things: Catching up to an accelerating opportunity," Nov. 2021. Available from: <https://tinyurl.com/24nkv6vn>
- [3] J. Zhang and D. Tao, "Empowering things with intelligence: A survey of the progress, challenges, and opportunities in artificial intelligence of things," *IEEE Internet Things J.*, vol. 8, no. 10, pp. 7789–7817, May 2021. Available from: <https://doi.org/10.1109/JIOT.2020.3039359>
- [4] Y. Shi, K. Yang, T. Jiang, J. Zhang, and K. B. Letaief, "Communication-efficient edge AI: Algorithms and systems," *IEEE Commun. Surveys Tuts.*, vol. 22, no. 4, pp. 2167–2191, 2020. Available from: <https://doi.org/10.1109/COMST.2020.3007787>
- [5] Ahmed, G. Jeon, and F. Piccialli, "From artificial intelligence to explainable artificial intelligence in Industry 4.0: A survey on what, how, and where," *IEEE Trans. Ind. Informat.*, vol. 18, no. 8, pp. 5031–5042, Aug. 2022. Available from: <https://doi.org/10.1109/TII.2022.3146552>
- [6] Alshehri and G. Muhammad, "A comprehensive survey of the Internet of Things (IoT) and AI-based smart healthcare," *IEEE Access*, vol. 9, pp. 3660–3678, 2021. Available from: <https://doi.org/10.1109/ACCESS.2020.3047960>
- [7] W. Tong, A. Hussain, W. X. Bo, and S. Maharjan, "Artificial intelligence for vehicle-to-everything: A survey," *IEEE Access*, vol. 7, pp. 10823–10843, 2019. Available from: <https://doi.org/10.1109/ACCESS.2019.2891073>
- [8] X. Zhou, C. Chai, G. Li, and J. Sun, "Database meets artificial intelligence: A survey," *IEEE Trans. Knowl. Data Eng.*, vol. 34, no. 3, pp. 1096–1116, Mar. 2022. Available from: <https://doi.org/10.1109/TKDE.2020.2994641>
- [9] M. E. Morochó Cayamcela and W. Lim, "Artificial intelligence in 5G technology: A survey," in *Proc. 9th Int. Conf. Inf. Commun. Technol. Convergence (ICTC)*, 2018, pp. 860–865. Available from: <https://doi.org/10.1109/ICTC.2018.8539642>
- [10] X. H. Li et al., "A survey of data-driven and knowledge-aware explainable AI," *IEEE Trans. Knowl. Data Eng.*, vol. 34, no. 1, pp. 29–49, Jan. 2022. Available from: <https://doi.org/10.1109/TKDE.2020.2983930>
- [11] J. Lin, W. Yu, N. Zhang, X. Yang, H. Zhang, and W. Zhao, "A survey on Internet of Things: Architecture, enabling technologies, security and privacy, and applications," *IEEE Internet Things J.*, vol. 4, no. 5, pp. 1125–1142, Oct. 2017. Available from: <https://doi.org/10.1109/JIOT.2017.2683200>
- [12] Lerario and A. Varasano, "An IoT smart infrastructure for S. Domenico Church in Matera's 'Sassi': A multiscale perspective to built heritage conservation," *Sustainability*, vol. 12, no. 16, Art. no. 6553, Aug. 2020. Available from: <https://doi.org/10.3390/SU12166553>
- [13] M. Soori, B. Arezoo, and R. Dastres, "Artificial intelligence, machine learning and deep learning in advanced robotics, a review," *Cognitive Robotics*, vol. 3, pp. 54–70, Jan. 2023. Available from: <https://doi.org/10.1016/J.COGR.2023.04.001>
- [14] Collins, D. Dennehy, K. Conboy, and P. Mikalef, "Artificial intelligence in information systems research: A systematic literature review and research agenda," *Int. J. Inf. Manage.*, vol. 60, Art. no. 102383, Oct. 2021. Available from: <https://doi.org/10.1016/J.IJINFOMGT.2021.102383>
- [15] F. Morales and H. J. Escalante, "A brief introduction to supervised, unsupervised, and reinforcement learning," in *Biosignal Processing and Classification Using Computational Learning and Intelligence: Principles, Algorithms, and Applications*. Academic Press, 2022, pp. 111–129. Available from: <https://doi.org/10.1016/B978-0-12-820125-1.00017-8>
- [16] P. Pham, L. T. T. Nguyen, W. Pedrycz, and B. Vo, "Deep learning, graph-based text representation and classification: A survey, perspectives and challenges," *Artif. Intell. Rev.*, vol. 56, no. 6, pp. 4893–4927, Jun. 2023. Available from: <https://doi.org/10.1007/S10462-022-10265-7/METRICS>
- [17] Venugopal, J. Tollich, M. Fairbank, and A. Scherp, "A comparison of deep-learning methods for analysing and predicting business processes," in *Proc. Int. Joint Conf. Neural Netw. (IJCNN)*, 2021. Available from: <https://doi.org/10.1109/IJCNN52387.2021.9533742>
- [18] U. A. Bhatti, H. Tang, G. Wu, S. Marjan, and A. Hussain, "Deep learning with graph convolutional networks: An overview and latest applications in computational intelligence," *Int. J. Intell. Syst.*, vol. 2023, no. 1, Art. no. 8342104, Jan. 2023. Available from: <https://doi.org/10.1155/2023/8342104>
- [19] S. R. Dubey and S. K. Singh, "Transformer-based generative adversarial networks in computer vision: A comprehensive survey," *IEEE Trans. Artif. Intell.*, 2024. Available from: <https://doi.org/10.1109/TAI.2024.3404910>
- [20] S. Mahendra Kothari et al., *Research Trends in Artificial Intelligence: Internet of Things*. Bentham Science Publishers, 2023.
- [21] K. S. Awaisi, Q. Ye, and S. Sampalli, "A survey of industrial AIoT: Opportunities, challenges, and directions," *IEEE Access*, 2024. Available from: <https://doi.org/10.1109/ACCESS.2024.3426279>
- [22] P. McEnroe, S. Wang, and M. Liyanage, "A survey on the convergence of edge computing and AI for UAVs: Opportunities and challenges," *IEEE Internet Things J.*, vol. 9, no. 17, pp. 15435–15459, Sep. 2022. Available from: <https://doi.org/10.1109/JIOT.2022.3176400>
- [23] C. H. Huang, T. C. Chou, and S. H. Wu, "Towards convergence of AI and IoT for smart policing: A case of a mobile edge computing-based context-aware system," *J. Global Inf. Manage.*, vol. 29, no. 6, 2021. Available from: <https://doi.org/10.4018/JGIM.296260>
- [24] S. B. Seo and D. Singh, "Smart town traffic management system using LoRa and machine learning mechanism," *IEEE Technol. Policy Ethics*, vol. 3, no. 6, pp. 1–4, Jun. 2022. Available from: <https://doi.org/10.1109/NTPE.2018.9778109>

- [25] M.-Y. Chung, Y.-D. Lu, C.-H. Lee, S.-Y. Kuo, and L.-B. Chen, "A wearable IoT-based rescue system using LoRa mesh network and physiological monitoring for mountain emergency response," in Proc. 2025 IEEE 14th Global Conf. Consumer Electron. (GCCE), 2025, pp. 1102–1103. Available from: <https://doi.org/10.1109/GCCE65946.2025.11274898>
- [26] GSMA, "NB-IoT deployment guide to basic feature set requirements," ver. 3, Jun. 2019. Available from: <https://www.gsma.com/solutions-and-impact/technologies/internet-of-things/wp-content/uploads/2019/07/201906-GSMA-NB-IoT-Deployment-Guide-v3.pdf>
- [27] LoRa Alliance, "LoRaWAN back-end interfaces," ver. 1.0, 2017. Available from: <https://resources.lora-alliance.org/document/lorawan-back-end-interfaces-v1-0>
- [28] C. T. Yang, H. W. Chen, E. J. Chang, E. Kristiani, K. L. P. Nguyen, and J. S. Chang, "Current advances and future challenges of AIoT applications in particulate matters (PM) monitoring and control," J. Hazard. Mater., vol. 419, Art. no. 126442, Oct. 2021. Available from: <https://doi.org/10.1016/J.JHAZMAT.2021.126442>
- [29] U.S. Environmental Protection Agency, "Frequent questions about air sensors." Available from: <https://www.epa.gov/air-sensor-toolbox/frequent-questions-about-air-sensors>
- [30] Sarhaddi et al., "LLMs and IoT: A comprehensive survey on large language models and the Internet of Things," IEEE Open J. Commun. Soc., vol. 7, pp. 3585–3616, 2026. Available from: <https://doi.org/10.1109/OJCOMS.2026.3680064>
- [31] D. Rivkin et al., "AIoT smart home via autonomous LLM agents," IEEE Internet Things J., vol. 12, no. 3, pp. 2458–2472, Feb. 2025. Available from: <https://doi.org/10.1109/JIOT.2024.3471904>
- [32] Zhang, R. Zhang, W. Zhang, D. Niyato, Y. Wen, and C. Miao, "Advancing generative artificial intelligence and large language models for demand side management with Internet of Electric Vehicles," IEEE Internet Things J., vol. 13, no. 13, pp. 28825–28840, Jul. 2026. Available from: <https://doi.org/10.1109/JIOT.2026.3685302>
- [33] N. Costa, A. Kraft, and A. Pereira, "Feasibility of an autonomous, cloud-less and controller-less architecture for consumer IoT," IEEE Sens. Rev., vol. 3, pp. 620–634, 2026. Available from: <https://doi.org/10.1109/SR.2026.3720252>
- [34] J. Liu, J. Wang, Y. Huang, and C. Wang, "LLM deployment strategies on mobile edge servers for dynamic uncertain user requests," IEEE Trans. Netw. Serv. Manag., vol. 23, pp. 5832–5850, 2026. Available from: <https://doi.org/10.1109/TNSM.2026.3708677>
- [35] Chhabra, S. Datta, S. K. Nahin, and P. Mohapatra, "Agentic AI security: Threats, defenses, evaluation, and open challenges," IEEE Access, vol. 14, pp. 49455–49482, 2026. Available from: <https://doi.org/10.1109/ACCESS.2026.3675554>
- [36] Huang, J. Yan, X. Fan, and H. Zhou, "Multi-scenario cloud-edge collaborative DDoS detection in LLM-enabled AIoT," IEEE Trans. Netw. Sci. Eng., vol. 13, pp. 3790–3809, 2025. Available from: <https://doi.org/10.1109/TNSE.2025.3637740>