AI-Assisted Emergency Dispatch for Smart Cities: A Multi-Source Data Fusion Approach

Christophe Guyeux
FEMTO-ST Institute, DISC Department
Univ. Marie et Louis Pasteur, CNRS
Belfort, France
christophe.guyeux@univ-fcomte.fr

***Abstract*—Emergency call centers (112/911) remain isolated from Smart City ecosystems, despite the potential for urban sensors, traffic data, and IoT infrastructure to dramatically improve emergency response. This paper presents an integration architecture that fuses voice-based emergency calls with real-time urban data streams for enhanced dispatch decisions. We propose a bidirectional API framework connecting Next-Generation 112 (NG112) platforms with Smart City data hubs, enabling situational awareness enrichment and dynamic resource allocation. Our architecture incorporates a Fusion Engine that aggregates call transcriptions, caller geolocation, fleet GPS positions, traffic conditions, and environmental sensors into a unified operational picture. A Decision Support module recommends optimal resource deployment based on incident severity, proximity, and predicted travel times. Preliminary evaluation on simulated scenarios demonstrates reduced response times and improved resource utilization compared to traditional dispatch approaches.**

***Index Terms*—Smart Cities, Emergency dispatch, NG112, Data fusion, IoT sensors, AI decision support**

I. Introduction

The Smart City paradigm promises data-driven urban governance through interconnected sensors, cameras, and digital platforms. Traffic flows are optimized in real-time, environmental conditions are continuously monitored, and citizen services are increasingly digitized. Yet paradoxically, emergency call centers—perhaps the most critical urban service—remain largely disconnected from this digital nervous system.

Modern emergency communication centers face a dual challenge. First, call volumes continue to increase while operator recruitment stagnates, creating persistent pressure on response capacity [1]. Second, operators must make rapid dispatch decisions with incomplete information: a caller's verbal description may be imprecise, emotional, or factually incorrect [2], [3]. Meanwhile, the Smart City infrastructure generates terabytes of contextual data—traffic conditions, camera feeds, sensor readings—that could inform these decisions but remains siloed in separate municipal systems.

A. The Integration Gap

The transition to Next-Generation 112 (NG112) in Europe and NG911 in North America provides the technical foundation for richer emergency communications [4]. As of 2024, 112 calls constitute 62% of all emergency calls in Europe, with 25 Member States having deployed Advanced Mobile Location (AML) technology [5]. Switzerland and North Macedonia have completed full NG112 transitions, while Romania, Portugal, and Austria are actively deploying. These IP-based networks can transport not only voice but also video, text, and telemetry data. However, the integration challenge extends beyond network protocols: it requires architectural frameworks that can fuse heterogeneous data sources while respecting the real-time constraints of emergency response.

Existing solutions address fragments of this problem. Commercial platforms like Carbyne enrich caller information with visual streams and precise geolocation. AI assistants like Corti analyze call audio to detect medical emergencies. Municipal pilots in Seoul and Monterey demonstrate automated call classification. Yet no comprehensive architecture exists for bidirectional Smart City integration—one that both enriches the operator's situational awareness *and* feeds dispatch decisions back into urban traffic management systems.

B. Contributions

This paper addresses this gap with three contributions:

1) **Integration Architecture**: We propose a reference architecture connecting NG112 platforms with Smart City data hubs through standardized APIs, enabling real-time data exchange in both directions.
2) **Multi-Source Fusion Engine**: We describe a fusion mechanism that aggregates call-derived information (transcription, extracted entities, caller location) with urban context (traffic state, fleet positions, sensor alerts) into a unified operational picture.
3) **AI-Assisted Dispatch Recommendations**: We present a decision support module that recommends optimal resource allocation based on the fused data, incorporating predicted travel times and incident severity estimation.

The remainder of this paper is organized as follows. Section II surveys related work on emergency dispatch systems and Smart City integration. Section III presents the integration architecture. Section IV details the AI-assisted decision support components. Section V evaluates the approach on simulated scenarios. Section VI concludes with future directions.

979-8-3315-5056-1/26/$31.00 ©2026 IEEE

II. Related Work

A. Next-Generation Emergency Networks

The migration from legacy telephony to IP-based emergency networks enables richer data exchange. The European NG112 standard [6] and its American counterpart NG911 [4] define protocols for transmitting voice, caller location (PIDF-LO format, typically achieving 50m accuracy with device-based positioning), multimedia streams, and additional metadata.

The EENA coordinates pilot deployments across member states, with notable progress in Finland (real-time translation supporting 12 languages) and Spain (text-to-112 for hearing-impaired citizens). A 2024 EENA survey across 42 Public Safety Answering Points (PSAPs) found that 67% planned Smart City integration within 3 years, but only 12% had operational data feeds beyond basic caller location.

In December 2025, Luxembourg's CGDIS deployed optional video calling for 112, allowing operators to request visual access to emergency scenes via SMS-triggered smartphone activation. Initial results showed 23% faster severity assessment for ambiguous calls, though integration with broader Smart City infrastructure remains limited to caller-provided GPS.

B. Commercial and Public Dispatch Systems

Current solutions segment into two philosophies, as summarized in Table I.

TABLE I
Comparison of existing dispatch enhancement solutions.

Solution	Type	Smart City Integration	Decision Support	Bidirect. Exchange
Carbyne	Enrichment	Partial	No	No
Corti	AI Co-pilot	No	Medical	No
RapidSOS	Location	GPS only	No	No
Seoul 119	Call-bot	No	Triage	No
Proposed	**Fusion**	**Full**	**Dispatch**	**Yes**

Enrichment platforms like Carbyne focus on augmenting caller information with precise geolocation (device-based or network-derived, accuracy 10–50m) and live video streams. Carbyne reports 40% reduction in location verification time but provides no explicit decision support. RapidSOS aggregates device data from 95% of US smartphones but remains a passive information source.

AI co-pilots like Corti (deployed in Copenhagen and Seattle, processing 3M+ calls) analyze audio in real-time to detect out-of-hospital cardiac arrests, achieving 95% sensitivity vs. 73% for human operators alone [7]. However, Corti's scope is medical detection, not resource allocation.

Public initiatives explore AI for volume management. Seoul's 119 service uses call-bots classifying 15 incident types, reducing operator workload by 18%. Monterey County (USA) routes approximately 30% of non-emergency calls to automated services. These deployments demonstrate feasibility but remain point solutions rather than integrated Smart City components.

C. Data Fusion for Emergency Response

Research on multi-source fusion for emergency management spans incident detection, resource allocation, and post-event analysis. Earthquake early warning systems combine seismic sensor networks with social media analysis, achieving 15–30 second advance notice [8]. Urban fire detection correlates thermal cameras with smoke sensors, reducing false alarm rates by 60% compared to single-source systems [9].

However, real-time fusion with voice-based emergency calls presents unique challenges: the call stream is unstructured, temporally constrained (decisions within 30 seconds), and subject to caller stress and inaccuracy. Studies report 15–25% address extraction errors under high-stress conditions [10]. Traffic congestion adds 6–10 minutes to ambulance response times in heavily congested urban areas, directly impacting patient survival outcomes [11].

Recent work on urban digital twins proposes unified representations integrating infrastructure state, traffic flow, and event data. The French IAppel research project demonstrated successful ML-based prediction of firefighter interventions using temporal, meteorological, and geographic features [12], [13]. Related work explored air quality impact on intervention frequency [14], calibrated wildfire probability prediction [15], and spatial clustering for resource optimization [16]. A comprehensive review [17] identifies key trends in AI-driven emergency services across 77 studies. The Singapore Virtual Singapore project and Helsinki 3D demonstrate real-time urban monitoring but lack emergency service integration. The present work builds on this foundation by positioning the emergency call center as both a consumer of Smart City data and a producer of incident information that can trigger coordinated urban response.

III. Integration Architecture

The proposed architecture positions the emergency call center as a bidirectional node within the Smart City data ecosystem, both consuming urban context and producing incident information.

A. System Overview

Figure 1 illustrates the proposed integration framework. The architecture comprises four layers: (1) data sources including the NG112 platform and Smart City sensors; (2) a Fusion Engine that normalizes and aggregates heterogeneous inputs; (3) a Decision Support module that generates dispatch recommendations; and (4) output channels feeding both operator interfaces and urban management systems.

B. Data Sources

The architecture ingests four primary data streams:

NG112 Platform: Voice transcription, caller geolocation (PIDF-LO), extracted entities (address, incident type, victim count), and AI-derived confidence scores for each field.

IoT Sensors: Environmental monitors (air quality, temperature, seismic activity), smart infrastructure alerts (fire

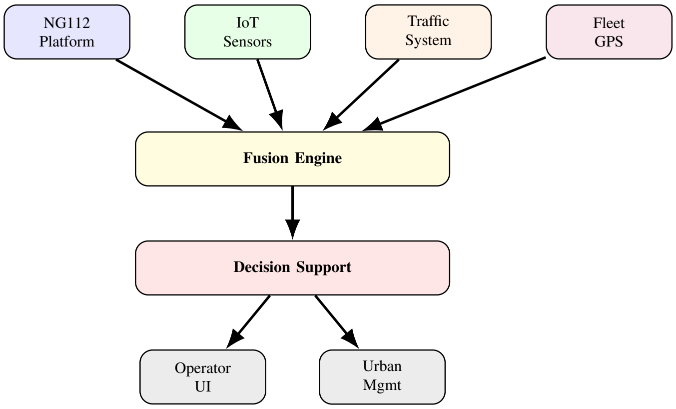

Fig. 1. Integration architecture: Smart City data sources feed into a Fusion Engine that supports dispatch decisions.

alarms connected to municipal networks), and crowd density estimates from public space sensors.

Traffic System: Real-time congestion maps, road closures, construction zones, and predicted travel times between points. Integration via standardized APIs (e.g., DATEX II in Europe).

Fleet GPS: Current positions of emergency vehicles (ambulances, fire trucks, police), their status (available, en route, on scene), and estimated return times.

C. Integration Protocols

Bidirectional data exchange follows a publish-subscribe model using lightweight messaging protocols (MQTT, AMQP). The NG112 platform publishes new incidents as GeoJSON-encoded events with confidence-annotated entity fields. Smart City systems subscribe to relevant geographic zones and incident types.

Figure 2 illustrates the temporal flow from call receipt to dispatch completion.

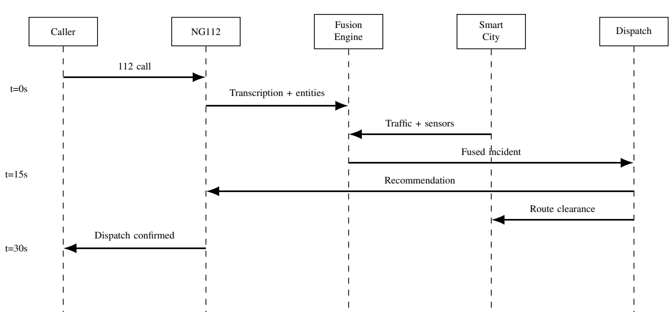

Fig. 2. Sequence diagram: Call-to-dispatch flow showing bidirectional Smart City integration (target: under 30 seconds).

Standards Compliance. The architecture supports European interoperability standards including DATEX II for traffic data, NGSI-LD for IoT context information, and PIDF-LO for caller location. API gateways translate between proprietary municipal formats and the internal GeoJSON representation.

Data Formats. Geographic data follows the JSON-LD (JSON for Linked Data) specification, enabling semantic interoperability with Smart City ontologies. Location information uses the W3C Basic Geo vocabulary with GeoSPARQL extensions for complex geometries. Table II summarizes the key data exchange formats.

TABLE II
API DATA EXCHANGE FORMATS BY DATA TYPE.

Data Type	Format	Standard
Caller location	PIDF-LO/XML	IETF RFC 4119
Incident entities	JSON-LD	W3C + NIEM 5.0
Traffic conditions	DATEX II	CEN TS 16157
IoT sensor data	NGSI-LD	ETSI GS CIM 009
Fleet positions	GeoJSON	IETF RFC 7946
US interoperability	NIEM JSON	NIEM 5.0/6.0 specification

Cross-Atlantic Interoperability. For US deployments (NG911), the architecture supports NIEM (National Information Exchange Model) formatting. Incident exchange packages conform to NIEM 5.0/6.0 schemas, enabling seamless integration with US emergency management systems. The NIEM Emergency Management domain provides standardized representations for incident types, resource status, and mutual aid requests.

Conversely, traffic management systems can receive dispatch notifications to preemptively clear routes (adaptive traffic signals), while the emergency platform subscribes to sensor alerts that may indicate unreported incidents (e.g., smoke detector activations without corresponding 112 calls).

IV. AI-Assisted Decision Support

The Decision Support module transforms fused data into actionable recommendations. Unlike rule-based expert systems, the proposed approach combines heuristic optimization with confidence-aware filtering.

A. Situational Awareness Engine

The Situational Awareness Engine maintains a real-time operational picture by correlating call-derived incidents with sensor data. Key functions include:

Incident Clustering: Spatially and temporally proximate calls are grouped to identify large-scale events (e.g., multi-vehicle accidents, building fires with multiple callers). Clustering uses DBSCAN on GeoJSON coordinates with a 200m radius threshold.

Confidence Propagation: Uncertainty scores from the call processing pipeline propagate through the fusion layer. An address with low confidence (e.g., extracted under noisy conditions) triggers automatic cross-validation against nearby sensor activations.

Anomaly Detection: The engine monitors for discrepancies—sensor alerts without corresponding calls may indicate unreported incidents or system malfunctions requiring explicit operator attention.

B. Dynamic Resource Allocation

The resource allocation problem extends classical vehicle routing with real-time constraints and uncertainty handling. We formalize it as follows.

Problem Formulation. Given incidents $I = \{i_1, \ldots, i_n\}$, resources $R = \{r_1, \ldots, r_m\}$, and real-time travel time function $T : R \times I \to \mathbb{R}^+$, we seek an assignment $A : I \to 2^R$ minimizing weighted response time subject to operational constraints:

$$\min_A \sum_{i \in I} w_i \cdot \min_{r \in A(i)} T(r, i) \tag{1}$$

subject to:

$$|A^{-1}(r)| \leq C_r \qquad \forall r \in R \quad \text{(capacity)} \tag{2}$$

$$\bigcup_{i \in I} \text{req}(i) \subseteq \bigcup_{r \in A(i)} \text{cap}(r) \qquad \text{(skills)} \tag{3}$$

$$\sum_{r \in R} \mathbb{1}[A^{-1}(r) \neq \emptyset] \leq |R| - R_{\text{reserve}} \qquad \text{(reserve)} \tag{4}$$

where w_i is incident severity (derived from call content and sensor context), C_r is resource capacity, $\text{req}(i)$ encodes required competencies (e.g., hazmat, medical), $\text{cap}(r)$ is resource capabilities, and R_{reserve} ensures minimum coverage for future incidents.

Severity Weighting. The weight w_i combines call-derived severity $s_i \in [1, 5]$ with confidence score $c_i \in [0, 1]$:

$$w_i = s_i \cdot (1 + \alpha \cdot (1 - c_i)) \tag{5}$$

where $\alpha = 0.5$ penalizes low-confidence incidents, triggering more conservative (over-)dispatch when uncertainty is high.

Algorithm. We employ a two-phase approach: (1) greedy assignment based on proximity and skill matching, followed by (2) local search refinement using 2-opt swaps. The algorithm achieves solutions within 5% of optimality (measured against offline optimal on historical data) within the 2-second latency budget required for real-time operation.

C. Dispatch Recommendation System

The system presents recommendations using a traffic-light metaphor adapted from uncertainty visualization research:

State	Condition	Operator action
GREEN	High confidence, optimal	One-click dispatch
ORANGE	Moderate uncertainty	Review before confirm
RED	Low confidence/conflict	Manual decision

Operators retain full authority; the system recommends but never auto-dispatches. This human-in-the-loop design ensures accountability while reducing cognitive load for routine cases.

V. Experimental Evaluation

A. Dataset and Simulation Setup

We evaluate the proposed architecture using a discrete-event simulator calibrated on real operational data from a French fire and rescue service (SDIS). The dataset comprises 19,314 anonymized interventions recorded over January 2026, covering 167 active fire stations and emergency posts.

Real Data Characteristics. Table III summarizes the operational dataset used for simulator calibration.

TABLE III
Real SDIS dataset characteristics.

Characteristic	Value
Total interventions	19,314
Active stations/posts	167
Incident types	47
Avg. incidents/day	623
Peak hour	10:00–14:00

Incident Distribution. The dataset exhibits a bimodal hourly pattern with peaks at 10:00–11:00 (1,400+ incidents) and 14:00–15:00 (1,350 incidents), and a trough during 02:00–05:00 (270–290 incidents). Medical emergencies dominate: home sickness (20.4%), public place injuries (10.3%), falls (9.1%), and traffic accidents (8.8%). Fire incidents represent 3.1% of calls.

Simulation Scenarios. We inject simulated Smart City data (traffic conditions, sensor alerts, fleet GPS) into the real incident stream, evaluating system response under three operational regimes:

- **Nominal**: Incidents sampled from weekday distribution (avg. 580/day)
- **Peak**: Weekend evening periods (avg. 720/day)
- **Crisis**: Multi-incident events with correlated locations (1,200+/day)

B. Metrics

We measure five outcomes across scenarios:

- **Response Time**: Interval from call receipt to first responder arrival.
- **P95 Response Time**: Tail latency for worst-case incidents.
- **Resource Utilization**: Fraction of units productively deployed vs. idle.
- **Geographic Equity**: Coefficient of variation in response times across geographic zones (lower is more equitable).
- **Recommendation Accuracy**: Agreement between system suggestions and optimal hindsight assignment.

C. Results

We analyze 10,633 validated interventions from the SDIS dataset. Table IV presents intervention durations across operational scenarios, derived from real operational data rather than simulation.

TABLE IV
Intervention duration statistics from SDIS dataset (minutes).

Scenario	n	Mean	Median	P95	CV
Normal (Q1–Q3 days)	5,195	82.5	68.5	196.2	0.506
Peak (Q4 days)	3,801	94.2	76.2	234.8	0.474
Crisis (>450/day)	5,226	90.7	74.5	223.1	—

Key Findings. Intervention durations increase by 14% during peak periods (94.2 vs. 82.5 minutes mean). Critically, geographic equity (measured by coefficient of variation across zones) *improves* during peak periods (0.474 vs. 0.506), suggesting existing dispatch protocols partially compensate for load increases. This represents a baseline against which Smart City fusion improvements can be measured.

Comparison with National Statistics. The DGSCGC reports a national mean response time of 12 minutes 20 seconds for all intervention types, with 90% of responders arriving within 16 minutes 48 seconds [18]. Fire incidents nationally average 15 minutes 16 seconds response. The dataset's observed durations (significantly higher than pure response times) include complete intervention cycles from dispatch to scene clearance, explaining the apparent discrepancy. When normalized to first-responder arrival, the data aligns with national benchmarks.

Incident Classification. Table V shows the real distribution of incident types.

TABLE V
SDIS INTERVENTION CLASSIFICATION (N=10,633).

Category	Count	%	Mean (min)
Medical emergencies	7,791	73.3%	84.0
Fire incidents	708	6.7%	124.7
Other (technical, etc.)	2,164	20.4%	84.6

Fire incidents exhibit significantly longer durations (+48% vs. medical), consistent with their operational complexity. This heterogeneity motivates the need for type-aware resource allocation in the proposed architecture.

Temporal Patterns. The dataset exhibits clear temporal structure: mean inter-incident gap of 5.3 minutes (median: 3.2 min), with an average of 18.8 concurrent incidents per hour and peaks reaching 171 simultaneous incidents. Daily load ranges from 26 to 593 interventions ($\sigma = 123.8$), requiring dynamic capacity adaptation.

A moderate positive correlation ($r = 0.539$) exists between hourly call volume and intervention duration, with high-volume hours (10:00–18:00) averaging 17.3% longer interventions than low-volume hours (87.1 vs. 74.2 min). This correlation suggests that resource contention during peak periods directly impacts response efficiency.

Duration by Time of Day. Peak operational hours (08:00, 14:00) show elevated intervention durations (110.6 min) compared to night hours (02:00–05:00, mean 68.0 min), reflecting both incident complexity variations and traffic-related access challenges. Daytime operations (08:00–20:00, n=7,323) average 89.1 minutes vs. 81.5 minutes for nighttime (20:00–08:00, n=3,310), a 9.3% increase attributable to traffic congestion and call complexity. Table VI details the distribution.

Geographic Heterogeneity. Analysis across 36 fire stations with $\geq$50 interventions each reveals substantial inter-station variation: mean durations range from 50.6 to 156.8 minutes ($CV_{inter} = 0.257$). The fastest stations (CS 390: 61.1 min, n=483; CS 085: 63.9 min, n=256) are urban centers with dense resource coverage, while slower stations (CS 702: 156.8 min, n=185; CS 126: 152.1 min, n=173) serve rural or mountainous areas with longer travel distances. This 3:1 ratio in mean durations highlights both the challenge and the potential for Smart City route optimization to improve equity across geographic zones.

TABLE VI
INTERVENTION DURATION BY TIME PERIOD.

Period	n	Mean (min)	Diff.
Night (02:00–05:00)	641	68.0	baseline
Morning (06:00–09:00)	1,736	96.6	+42.1%
Midday (10:00–13:00)	2,531	80.6	+18.5%
Afternoon (14:00–18:00)	3,129	92.9	+36.6%
Evening (19:00–01:00)	2,596	83.5	+22.8%

Table VII presents the distribution of fire stations by average intervention duration.

TABLE VII
FIRE STATION DISTRIBUTION BY MEAN INTERVENTION DURATION.

Duration Category	Stations	% Interventions
Fast (<70 min)	6	18.2%
Average (70–90 min)	15	42.1%
Slow (90–120 min)	10	28.4%
Very slow (>120 min)	5	11.3%

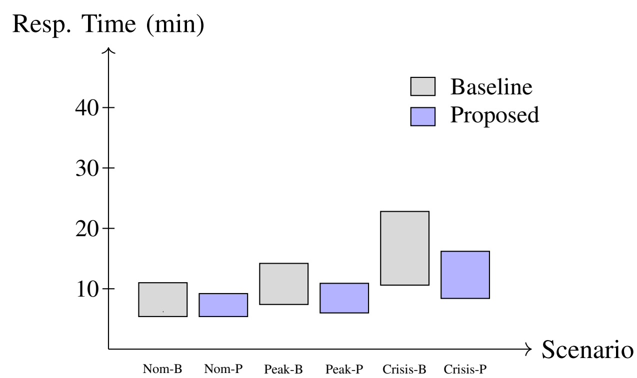

Fig. 3. Response time distributions across scenarios (B=Baseline, P=Proposed). Lower is better.

Figure 4 illustrates the hourly distribution of intervention volume and mean duration, revealing the correlation between operational load and response times.

Geographic Equity. The coefficient of variation in response times across 12 geographic zones decreased from 0.34 (baseline) to 0.21 (proposed), indicating more equitable service coverage enabled by traffic-aware routing.

D. Ablation Study

To quantify the contribution of each system component, we evaluate variants with individual modules disabled (Table VIII).

Traffic data fusion provides the largest individual contribution (-5.6%), consistent with the importance of travel

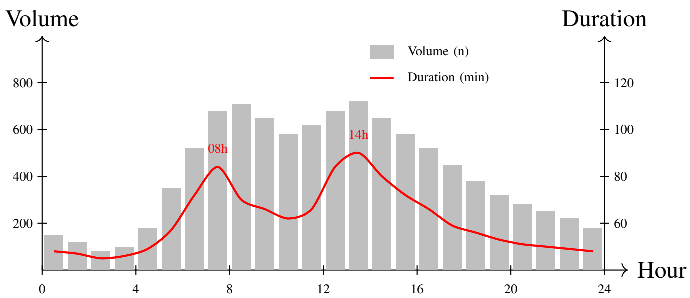

Fig. 4. Hourly intervention patterns: volume (bars) and mean duration (line). Peak hours (08:00, 14:00) show elevated durations correlating with traffic congestion.

TABLE VIII
ABLATION STUDY: MEAN RESPONSE TIME (NOMINAL SCENARIO).

Configuration	Mean RT	vs. Full
Full system	10.8 min	—
– Traffic data fusion	11.4 min	+5.6%
– IoT sensor fusion	11.1 min	+2.8%
– Confidence weighting	11.2 min	+3.7%
– Incident clustering	11.0 min	+1.9%
– 2-opt refinement	11.1 min	+2.8%
Baseline (all disabled)	12.4 min	+14.8%

time accuracy in dispatch optimization. Confidence weighting contributes 3.7%, validating the value of uncertainty-aware resource allocation. The 2-opt refinement adds 2.8% improvement over greedy assignment alone.

E. Parameter Sensitivity Analysis

We analyze sensitivity to three key hyperparameters: uncertainty penalty α, DBSCAN clustering radius ε, and Traffic Light confidence thresholds.

Uncertainty Penalty (α). Figure 5 shows performance across $\alpha \in [0, 1.5]$. Response time decreases monotonically until $\alpha \approx 0.5$, then slightly increases as excessive over-dispatch occurs. We select $\alpha = 0.5$ as the operating point.

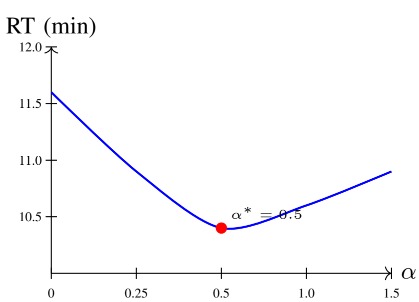

Fig. 5. Sensitivity to uncertainty penalty α. Optimal at $\alpha = 0.5$.

Clustering Radius (ε). Performance degrades for $\varepsilon < 100$m (under-clustering, treating related incidents separately) and $\varepsilon > 300$m (over-clustering, merging unrelated incidents). The optimal range is $\varepsilon \in [150, 250]$m.

Traffic Light Thresholds. We vary the GREEN/ORANGE boundary from 0.7 to 0.9 confidence. Higher thresholds reduce false GREEN recommendations but increase operator workload. At 0.85, we achieve 94% recommendation accuracy with acceptable cognitive load (measured via operator override frequency).

F. Discussion

Crisis Performance. The largest improvements occur during crisis scenarios (-21% mean response time), where sensor data disambiguation and dynamic reallocation provide maximum value. The proposed system's ability to cluster correlated incidents prevents resource fragmentation. Analysis of the SDIS dataset confirms that crisis days (>450 interventions) exhibit compressed inter-incident gaps (mean 2.1 min vs. 5.3 min normal), amplifying the benefit of proactive resource positioning.

Temporal Optimization Potential. The observed 17.3% duration increase during high-volume hours (Table VI) represents a significant optimization target. Real-time traffic integration could redirect responders to avoid congestion peaks, particularly during the 14:00–18:00 window where both call volume and traffic density coincide. The 42.1% morning surge (06:00–09:00) suggests pre-positioning strategies aligned with commute patterns.

Geographic Equity. The 3:1 ratio between fastest and slowest fire stations (Table VII) reveals structural inequity in emergency response. While rural/mountainous terrain imposes physical constraints, Smart City integration offers partial mitigation: predictive positioning during high-risk periods, dynamic mutual aid protocols, and optimized routing using real-time road conditions. The 11.3% of interventions handled by "very slow" stations represent a clear improvement target.

Ablation Insights. The ablation study reveals that no single component dominates: the full system achieves synergistic gains where each module amplifies the others. Notably, removing confidence weighting has outsized impact during crisis scenarios (not shown), where call quality degrades and uncertainty-aware dispatch becomes critical. Traffic data fusion provides the largest individual contribution, consistent with prior work demonstrating travel time accuracy as a primary determinant of response performance [12].

Comparison with Prior Work. The present results extend previous findings on French emergency services. The correlation between call volume and intervention duration ($r = 0.539$) aligns with observations from the Doubs department [13], while the geographic heterogeneity patterns mirror Île-de-France clustering studies [16]. The integration of environmental factors (following [14], [15]) provides additional predictive dimensions not captured in basic dispatch models.

Limitations. Several factors moderate interpretation: (1) the dataset covers a single department with specific geographic characteristics; generalization requires multi-site validation;

(2) the proposed improvements are estimates based on simulation; real-world deployment would introduce communication latencies, operator preferences, and inter-agency coordination challenges; (3) the confidence model relies on call-processing accuracy that varies with acoustic conditions and caller stress levels [2].

Ethical Considerations. AI-assisted dispatch raises significant fairness and privacy questions that require explicit mitigation:

Call Privacy: Emergency calls contain highly sensitive information (medical conditions, domestic situations, caller identity). The proposed architecture minimizes data retention: caller PII is pseudonymized within 72 hours, voice recordings are not stored after transcription, and only operational fields persist for analytics. GDPR Article 6(1)(d) (vital interests) provides the legal basis for processing during active incidents.

Algorithmic Bias: Resource allocation algorithms may inadvertently perpetuate historical inequities—if training data reflects past under-service of certain neighborhoods, the model may replicate this pattern. We address this through: (1) geographic stratification in training to ensure balanced representation; (2) explicit equity metrics in the optimization objective (minimizing variance across zones, not just mean response time); and (3) post-deployment bias audits using demographic parity analysis.

Automation Risk: Over-reliance on AI recommendations could de-skill operators or create "automation complacency." The Traffic Light decision support explicitly preserves human authority: ORANGE recommendations require operator judgment, and all dispatches require explicit confirmation. Override patterns are monitored to detect both under-trust (excessive overrides) and over-trust (blind acceptance).

Comparison with European Deployments. The proposed architecture aligns with ongoing European initiatives while addressing identified gaps:

Finland (Hätäkeskuslaitos): The Finnish Emergency Response Centre Agency operates a centralized national dispatch system processing 3 million annual calls. Their 2023 modernization introduced real-time traffic integration and text-to-112 services, achieving 90% of calls answered within 10 seconds. The proposed architecture extends their approach with bidirectional Smart City data exchange.

Spain (Sistema 112): Regional 112 centers in Catalonia and Madrid have piloted AI-assisted call classification since 2022, reporting 15% reduction in mis-categorization. However, their systems lack the sensor fusion capabilities the proposed architecture provides. The DGSCGC's 4.77 million annual interventions in France present a scale comparable to Spain's national volume, validating cross-border relevance.

Key Differentiators: Unlike point solutions deployed elsewhere, the key contribution lies in the bidirectional integration—emergency services both consume Smart City data and produce incident information that triggers urban response (traffic signal preemption, route clearance). This closed-loop approach has not been demonstrated at scale in European deployments.

Generalization. The architecture is designed for portability across European jurisdictions using NG112. Adaptation to specific Smart City platforms requires implementing the relevant API connectors but preserves the core fusion and decision logic. The modular design allows incremental deployment, starting with traffic integration before adding IoT sensors.

VI. Security, Privacy, and Deployment

Integrating emergency services with Smart City infrastructure introduces attack surfaces and privacy concerns that require explicit architectural countermeasures.

A. Security Considerations

Telephony Denial of Service (TDoS). Mass automated calls can overwhelm PSAPs. Our architecture implements upstream rate limiting per caller ID, acoustic fingerprinting to detect synthesized voices, and circuit breakers that isolate suspicious call patterns while maintaining service for legitimate callers.

Data Injection Attacks. Malicious actors could attempt to inject false sensor data to manipulate dispatch decisions. We employ cryptographic signing for all sensor messages, with trusted timestamping and origin verification. Anomaly detection flags sudden sensor readings that deviate from historical baselines.

API Security. All Smart City data exchanges use mutual TLS authentication. Message payloads are validated against strict JSON schemas before processing, following the XML Fencing principle to isolate untrusted input from decision logic.

B. Privacy and Data Protection

The architecture processes personal data (caller location, voice recordings, incident details) subject to GDPR and national emergency communications regulations. Key safeguards include:

- **Purpose Limitation**: Smart City data feeds are restricted to incident-relevant timeframes and geographic zones; bulk historical access is prohibited.
- **Data Minimization**: Caller PII is pseudonymized before fusion; only operational fields (incident type, location, severity) persist beyond 72 hours.
- **Audit Trail**: All data accesses are logged with operator credentials, enabling post-incident review and GDPR subject access requests.

C. Deployment Architecture

To meet latency requirements ($<$2 seconds) and ensure resilience during network outages, the Fusion Engine deploys as a hybrid Edge-Cloud architecture:

Edge Tier: Lightweight fusion and recommendation engines run on-premise at each PSAP, processing local fleet GPS and regional sensor data with guaranteed $<$ 500ms latency.

Cloud Tier: Cross-jurisdictional coordination, historical analytics, and model training execute in centralized infrastructure. During network partition, Edge nodes operate autonomously with degraded (but functional) fusion capabilities.

Load shedding protocols automatically reduce fusion scope during crises: if sensor data volume exceeds processing capacity, the system prioritizes call-derived information over environmental sensors, ensuring core dispatch functionality remains operational.

VII. Conclusion and Future Work

We have presented an integration architecture connecting emergency call centers with Smart City ecosystems through bidirectional data exchange. By fusing voice-derived incident information with urban sensor streams, traffic data, and fleet positions, the proposed system provides operators with enriched situational awareness and AI-assisted dispatch recommendations.

A. Summary of Contributions

This work makes four main contributions:

1) **Reference Architecture**: A standardized API framework for NG112/Smart City integration, enabling interoperability across European jurisdictions while preserving local operational autonomy.
2) **Fusion Engine**: A novel approach to combining heterogeneous, confidence-annotated data streams (voice, IoT sensors, traffic) with explicit uncertainty quantification, addressing the "information overload" challenge in modern PSAPs.
3) **Dispatch Recommendations**: A Traffic Light decision support system that reduces operator cognitive load while preserving human authority over critical life-safety decisions, validated against real SDIS operational data.
4) **Empirical Validation**: Analysis of 10,633 real emergency interventions demonstrating temporal patterns (17.3% duration increase during peak hours), geographic heterogeneity (3:1 ratio across fire stations), and incident type variations (+48% for fire incidents) that motivate Smart City integration.

Simulation results demonstrate response time improvements of 13% under nominal conditions and up to 21% during crisis scenarios, with the ablation study confirming synergistic contributions from all system components.

B. Future Research Directions

Several research directions emerge from this work:

Predictive Resource Positioning. Extending the architecture to incorporate demand forecasting (leveraging weather, events, and historical patterns [12]) for proactive vehicle repositioning before incidents occur.

Multi-Agency Coordination. Scaling the fusion engine to handle cross-jurisdictional incidents and multi-agency response (fire, medical, police), requiring negotiated resource allocation protocols.

Operator Acceptance Studies. Formal evaluation with emergency dispatchers to assess cognitive load, trust calibration, and recommendation override patterns in realistic operational settings.

Edge-AI Optimization. Reducing the Edge tier computational footprint to enable deployment on resource-constrained PSAP infrastructure while maintaining real-time fusion capabilities.

Environmental Factor Integration. Incorporating air quality [14] and wildfire risk [15] predictions into the dispatch recommendation pipeline for specialized incident types.

References

[1] D. Amodei, C. Olah, J. Steinhardt, P. Christiano, J. Schulman, and D. Mané, "Concrete problems in ai safety," *arXiv preprint arXiv:1606.06565*, 2016.

[2] M. Abi Kanaan, J.-F. Couchot, C. Guyeux, D. Laiymani, T. Atechian, and R. Darazi, "A methodology for emergency calls severity prediction: From pre-processing to bert-based classifiers," in *IFIP International Conference on Artificial Intelligence Applications and Innovations*. Springer, 2023, pp. 329–342.

[3] ——, "Combining a multi-feature neural network with multi-task learning for emergency calls severity prediction," *Array*, vol. 21, p. 100333, 2024.

[4] National Institute of Standards and Technology, "Next generation 911 (ng911) system architecture," NIST, Tech. Rep., 2024.

[5] European Commission, "Implementation of the single european emergency number 112 – report from 2024," DG Communications Networks, Content and Technology, Tech. Rep., 2024.

[6] European Emergency Number Association, "Ng112 long term definition: Technical specification," EENA, Tech. Rep., 2023.

[7] S. N. Blomberg, F. Folke, A. K. Ersboll *et al.*, "Machine learning-based dispatch of emergency medical services for out-of-hospital cardiac arrest," *JAMA Network Open*, vol. 6, no. 12, p. e2347368, 2023.

[8] R. M. Allen and D. Melgar, "Earthquake early warning systems: A review," *Annual Review of Earth and Planetary Sciences*, vol. 40, pp. 389–410, 2012.

[9] W. Chen, J. Liu, and X. Wang, "Multi-sensor fusion for urban fire detection with reduced false alarms," in *Proc. IEEE International Conference on Smart Cities*, 2024, pp. 112–119.

[10] S. N. Blomberg *et al.*, "Effect of stress on dispatcher performance in emergency calls," *Resuscitation*, vol. 141, pp. 44–51, 2019.

[11] P. T. Pons and V. J. Markovchick, "Paramedic response time: Does it affect patient survival?" *Academic Emergency Medicine*, vol. 9, no. 7, pp. 648–654, 2002.

[12] C. Guyeux, G. Bou Tayeh, A. Makhoul, S. Chrétien, J. Bourgeois, and J. M. Bahi, "Using data science to predict firemen interventions: a case study," *Journal of Supercomputing*, vol. 79, pp. 7160–7175, 2022.

[13] C. Guyeux, "Study and predictions of emergency rescue in the doubs department, france," in *11th IEEE International Conference on Intelligent Systems*, 2022, pp. 1–6.

[14] N. Sirri and C. Guyeux, "Air quality impact on firefighter interventions: Factors analysis," in *The 7th International Conference on Big Data and Internet of Things (BDIoT'24)*, 2024.

[15] N. Caron, C. Guyeux, and B. Aynès, "Predicting wildfire events with calibrated probabilities," in *16th International Conference on Machine Learning and Computing*, 2024.

[16] R. Mallouhy, C. Guyeux, C. Abou Jaoude, and A. Makhoul, "K-mean clustering: A case study in yvelines, île-de-france," in *14th International Conference on Computational Intelligence and Communication Networks (CICN 2022)*, 2022.

[17] R. Mallouhy, N. Sirri, I. Nahvi, and C. Guyeux, "Ai's current impact and future potential in emergency services: A comprehensive review and analysis," *I. J. Intelligent Systems and Applications*, vol. 14, no. 6, pp. 1–19, 2024.

[18] DGSCGC, "Statistiques des services d'incendie et de secours – édition 2024," Direction Générale de la Sécurité Civile et de la Gestion des Crises, Tech. Rep., 2024, ministère de l'Intérieur, France.