A Geospatial Data-Driven Platform for Real-Time Urban Monitoring and Crowd Management

Valentina Franzoni
Department of Mathematics and Computer Science
University of Perugia, Italy
valentina.franzoni@unipg.it
ORCID: 0000-0002-2972-7188

Andrea Chiancone
WiseTown Department
TeamDev Ltd, Perugia, Italy
a.chiancone@teamdev.it
ORCID: 0000-0002-9818-1537

Mariana Minascurta
WiseTown Department
TeamDev Ltd, Perugia, Italy
m.minascurta@teamdev.it

Cristiana Mapelli
WiseTown Department
TeamDev Ltd, Perugia, Italy
c.mapelli@teamdev.it

***Abstract*—**Urban environments face growing pressure from high-density scenarios generated by public events, tourism, mobility dynamics and emergencies, where reliable, timely and integrated information is essential for operational decision-making, crowd management and public safety. Yet urban data remain fragmented across heterogeneous systems, hindering a coherent real-time understanding of city dynamics. This work presents WiseTown® Situation Room, a cloud-native geospatial platform that integrates heterogeneous urban data streams for real-time monitoring and crowd management from Perugia Digital Twin, and introduces its most novel component: an affective crowd-sound analytics module that classifies the collective emotional state of a crowd, directly from ambient sound, for an early, privacy-preserving indicator of rising tension. We describe the platform architecture and report a real-world test during Eurochocolate, a large annual festival in the historic center of Perugia. The results indicate that multi-source geospatial intelligence, augmented with affective sensing, improves situational awareness, operational coordination and decision support in both planned events and emergencies.

***Keywords*—**Smart Cities; GeoAI; Crowd Monitoring; Affective Computing; Crowd Sound; Urban Intelligence; GIS

I. INTRODUCTION

Large-scale urban events and emergencies force cities to adapt operational monitoring to rapidly evolving conditions. Public gatherings in historical centers create localized overcrowding, traffic instability and safety-critical situations that fragmented infrastructures struggle to manage. Urban computing increasingly combines heterogeneous sensing infrastructures and spatial analytics to support data-driven city management [1] through distributed sensors, traffic systems, mobile devices, institutional services and citizen-participation platforms, while advances in GeoAI have expanded the capacity to process and contextualize large-scale urban data streams [2]. Operational use, however, remains constrained by fragmentation and limited interoperability.

To address these challenges, WiseTown® Situation Room, developed by TeamDev, integrates multi-source urban data from the Digital Twin of Perugia, Italy, within a unified geospatial framework for scenario-oriented monitoring and actionable intelligence in event management and civil protection. Beyond integration, the main novel and challenging contribution of this work is the inclusion of affective crowd-sound analysis (Section III) in the operational loop. Section IV reports as a case study a concrete deployment during the Eurochocolate festival in Perugia.

II. THE WISETOWN SITUATION ROOM PLATFORM

WiseTown® Situation Room is a cloud-native geospatial platform that supports real-time urban monitoring through the integration, correlation and visualization of heterogeneous data streams within a modular, multi-layer architecture (Fig. 1). Data are acquired from four source families: distributed IoT sensors capturing noise and *crowd sound*, traffic, air quality, occupancy and environmental conditions; public social-media and web channels reporting citizen-perceived issues; citizen reports submitted through mobile and web apps [3]; and institutional external APIs and open-data feeds such as weather, mobility and public events. An API gateway and ETL ingestion pipelines standardize these streams through interoperable connectors.

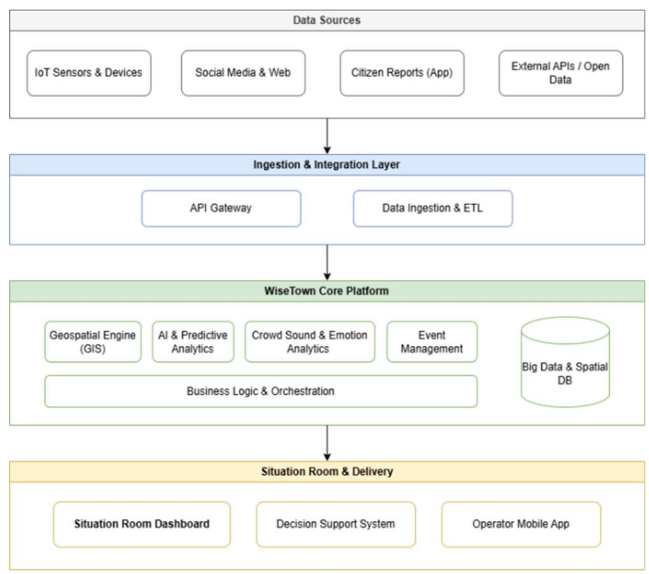

Fig 1: Workflow of the system architecture

A core intelligence layer transforms raw data into actionable knowledge. It combines a geospatial (GIS) engine for geolocation and spatial correlation, AI and predictive analytics for pattern and trend detection, the crowd-sound and emotion analytics module of Section III, event and alarm management, and a big-data spatial repository, all coordinated by a business-logic and orchestration component. The output layer delivers a real-time Situation Room dashboard, a decision-support system that recommends prioritized actions, and an operator mobile app for field coordination. A space-time geo-intelligence layer contextualizes information by both location and temporal evolution [4], so that operators can define dynamic areas of interest—squares, access points, transport hubs, evacuation routes—for scenario-oriented monitoring of crowd density, traffic conditions and anomalous aggregations.

III. AFFECTIVE CROWD-SOUND ANALYTICS

Conventional crowd monitoring relies on counting people and measuring flows and density: it reveals where and how many, but not the collective emotional state—whether a dense gathering is festive, indifferent or turning tense. We integrate into the Situation Room an affective classifier that treats the crowd as a collective individual and infers its emotion from the sound it produces: cheering and clapping signal approval, booing and shouting signal disapproval, and background chatter is neutral [5].

Ambient audio is loudness-normalized (EBU R128, −23 LUFS) and segmented by a sliding window into overlapping one-second blocks; each block is converted into a frequency–amplitude spectrogram image. Spectrograms feed an AlexNet convolutional neural network pre-trained on ImageNet and fine-tuned by transfer learning on the crowd-sound domain through spectrograms, so that only a comparatively small labeled dataset is needed. A per-clip majority vote over block predictions yields the crowd-emotion label. The approach reaches over 90% accuracy with the Log and Mel auditory scales—about 10% above the corresponding state-of-the-art individual speech baseline.

Emotion carried by crowd sound is largely pre-cognitive—hisses, boos, modulated screams—so it is less mediated than text-based social sentiment and can flag rising tension before it becomes visible in flow data. Within the Situation Room, the classifier will run on microphone streams georeferenced to areas of interest: a shift toward disapproval in a saturated square can raise an early alert, trigger the decision-support system, and be cross-correlated with density, traffic and citizen reports. Because it uses aggregate ambient sound rather than images or identifiable speech, it complements visual monitoring while preserving privacy.

IV. CASE STUDY: EUROCHOCOLATE, PERUGIA

Eurochocolate is the largest chocolate festival in Europe, held every autumn since 1994 in the historic center of Perugia; for about ten days its narrow medieval streets and squares fill with visitors, producing exactly the high-density, complex-mobility conditions that challenge real-time monitoring. The Situation Room was deployed in an experimental version to monitor and manage visitor flows across the event. Heterogeneous, mostly georeferenced data were fused: e.g., area video surveillance, geo-referenced social-media posts, IoT streams, and anonymized and aggregated mobile-network data, in full compliance with privacy regulations. From these sources the platform reconstructed inbound and outbound flows on the main road and rail access routes and across the internal mobility system (shuttles, minimetro, buses, cars, parking), quantified congestion peaks and the most-used routes, distinguished weekday from weekend attendance, profiled visitors by gender, age band and regional origin, and compared figures with previous editions. It also captured the intersection dynamics arising when Eurochocolate overlaps with other city activities. The resulting hour-by-hour picture was shared with public-safety institutions to plan corrective actions and improve accessibility and safety in later editions. This deployment provides the flow-monitoring backbone into which the affective crowd-sound layer of Section III is being integrated, so that quantitative flow intelligence is complemented by emotion-aware situational awareness.

Fig. 2 : WiseTown® Situation Room monitoring dashboard.

Beyond event management through the Digital Twin of Perugia, the platform is being extended as a civil-protection tool across all the Umbria Region, supporting evacuation coordination, hydrogeological-hazard monitoring and large-scale urban risk assessment across both planned and unplanned high-density contexts.

V. CONCLUSIONS

We presented WiseTown® Situation Room, a geospatial data-driven platform for real-time urban monitoring and crowd management that integrates heterogeneous urban data within a unified geo-intelligence framework, and we introduced affective crowd-sound analytics as its main novel contribution: a spectrogram-based deep-transfer-learning classifier that adds the collective emotional dimension to conventional flow monitoring. The Eurochocolate deployment demonstrated enhanced situational awareness and operational coordination in a challenging historic-city setting. Future work will consolidate the on-line integration of the crowd-sound module with density and flow analytics, extend the emotional model, and incorporate AI-based anomaly detection to strengthen adaptive urban monitoring and decision support.

REFERENCES

[1] Y. Zheng et al., “Urban computing: concepts, methodologies, and applications,” *ACM TIST*, vol. 5, no. 3, pp. 1–55, 2014.

[2] D. G. Costa et al., “Achieving sustainable smart cities through geospatial data-driven approaches,” *Sustainability*, vol. 16, no. 2, 640, 2024.

[3] P. Cardullo and R. Kitchin, “Being a ‘citizen’ in the smart city,” *GeoJournal*, vol. 84, no. 1, pp. 1–13, 2019.

[4] W. Li, “GeoAI: where machine learning and big data converge in GIScience,” *J. Spatial Information Science*, vol. 20, pp. 71–77, 2020.

[5] V. Franzoni, G. Biondi and A. Milani, “Emotional sounds of crowds: spectrogram-based analysis using deep learning,” *Multimedia Tools and Applications*, vol. 79, pp. 36063–36075, 2020.