



---

## MARVEL: Multimodal Extreme Scale Data Analytics for Smart Cities Environments

Eva Papadogiannaki<sup>1#</sup>, Despina Kopanaki<sup>1</sup> and Sotiris Ioannidis<sup>1,2\*</sup>

<sup>1</sup> Foundation for Research and Technology – Hellas (FORTH)

<sup>2</sup> Technical University of Crete (TUC)

# Presenting author: Eva Papadogiannaki, email: epapado@ics.forth.gr

\* Corresponding author: Sotiris Ioannidis, email: sotiris@ics.forth.gr

---

### ABSTRACT

The Smart City paradigm considers the city as a complex and dynamic system involving different interconnected spatial, social, economic, and physical processes subject to temporal changes and continually modified by human actions. One critical aspect is to encapsulate the complexity of a city and support accurate, cross-scale, and in-time predictions based on the ubiquitous spatiotemporal data of high-volume, high-velocity, and of high-variety. To address this challenge, MARVEL delivers a disruptive Edge-to-Fog-to-Cloud (E2F2C) ubiquitous computing framework that enables multi-modal perception and intelligence for audio-visual scene recognition and event detection in a smart city environment. MARVEL aims to collect, analyse and data mine multi-modal audio-visual data streams of a Smart City and help decision-makers to improve the quality of life and services to the citizens. This is achieved via: (i) fusing large scale distributed multi-modal audio-visual data in real-time; (ii) achieving fast time-to-insights; (iii) supporting automated decision making at all levels of the E2F2C stack; and (iv) delivering a personalised federated learning approach, where joint multi-modal representations and models are co-designed and improved continuously through privacy-aware sharing of personalised fog and edge models of all interested parties. Of course, privacy preservation is a key objective, so MARVEL aims to achieve these goals without violating ethical and privacy limits in an AI-responsible manner.

Since the whole MARVEL platform consists of 29 technological components with a wide range of functionalities and roles, the components have been grouped into seven subsystems to achieve coherence in the framework presentation: (i) Sensing and perception, (ii) Security, privacy, and data protection, (iii) Data management and distribution, (iv) Audio, visual, and multimodal AI, (v) Optimized E2F2C processing and deployment, (vi) E2F2C infrastructure and (vii) system outputs. Greenroads Malta is the first pilot operating in the transport domain. The goal is to utilize tools capable of detecting possible anomalous or dangerous situations, activate the proper countermeasures or inform other drivers and the control room and to provide traffic analytics. For the Municipality of Trento, the use cases focus on monitoring some public spaces in order to identify anomalous or potentially dangerous situations. For the third pilot, the University of Novi Sad experiment will support the other two pilots (Malta and Trento) by providing a specific small scale use case and data for further processing. The goal is to evaluate the potential of drones in the monitoring of large public events.

Some of the components that FORTH brings into the MARVEL project participate in the *security, privacy, and data protection* and the *optimized E2F2C processing and deployment* subsystems. The first component, *EdgeSec VPN*, aims to secure the communications of the MARVEL components that exchange data across the network. The second component, *EdgeSec TEE*, aims to shield the execution of applications that process sensitive user data by using Trusted Execution Environments. The third component, *GPURegex*, accelerates the audio captioning processing in a streaming fashion.