

**Ph.D. in Information Technology  
Thesis Defense**

**February 3<sup>rd</sup>, 2025  
at 14.30 pm**

**Sala Conferenze Emilio Gatti – building 20**

**Carlo Alberto BONO – XXXVII Cycle**

**Enhancing Situational Awareness through Real-Time Social Media Data Analysis**

Supervisor: Prof. Barbara Pernici

**Abstract:**

Social media, with its pervasive and immediate nature and by generating an unprecedented flow of information, has revolutionized the way people experience and share reality, creating opportunities and challenges that span across multiple domains on a global scale. Given their real-time nature, social media platforms are ideal for accessing up-to-date, actionable information during time-critical events, such as emergencies. However, the complex and unstructured nature of social media-sourced data makes it challenging to extract valuable insights. This thesis addresses the need for effective data preparation methodologies to fully leverage social media, focusing on the unique challenges posed by its content during emergency events, with the ultimate goal of extracting critical, relevant information from vast and undifferentiated streams.

Research on the relevance of data is conducted by exploring several key areas: the distinctive characteristics of social media content, the unpredictability and variability of events, and the need for tailorable data filtering; it does so by investigating methods for isolating relevant content, leveraging multimedia attachments to enhance situational awareness, detecting events of interest, adapting data filtering to evolving scenarios.

Four guiding questions drive this research:

(i) How can systematic approaches support the extraction of relevant data from social media streams, accounting for context-dependent requirements, and in particular during emergency events?

(ii) How can a language-independent, social media-based, context-aware detection system for emergency scenarios be designed?

(iii) How can continuous model adaptation during emergency events be achieved?

(iv) What information can be inferred from media attachments, and how can they enhance available data in emergency scenarios?

This thesis contributes to the field according to the investigated areas, by (i) proposing a methodology for designing data preparation pipelines, analyzing the dimensions for requirements and constraints and supporting the design with a human-in-the-loop enhancement procedure, (ii) investigating language-agnostic, customizable methods for early detection and characterization of

large-scale events, (iii) studying adaptive filtering techniques through the lens of bounded labelling resources, and (iv) proposing innovative uses of multimedia attachments in social media posts, also leveraging recent advances in LLM technologies, to enrich the available information.

These contributions aim to enhance situational awareness and the actionable value of gathered data, ultimately improving concrete decision-making processes such as disaster response and mitigation efforts. Through real-world case studies and empirical validation, this research aims to bridge the gap between current methodologies and practical needs for time- and resource-constrained data discovery and filtering, proposing validation scenarios in the field of emergency management and offering practical solutions for real-time social media data analysis.

## **PhD Committee**

Prof. Maristella Matera, **Politecnico di Milano**

Prof. Terje Gjosater, **University of Adger**

Prof. Matteo Palmonari, **Università di Milano-Bicocca**