UNIVERSIDAD
DE GRANADA

Facultad de Ciencias Económicas y Empresariales
Programa de Doctorado en Ciencias Económicas y Empresariales
Departamento de Organización de Empresas

Towards an Open Source Intelligence and Machine Learning-Based Platform for Automation in Disaster Management Situational Awareness Reporting

TESIS DOCTORAL

PRESENTADA POR:

Klaus Dieter Schwarz

Directores: Prof. Dr. Daniel Arias Aranda
Prof. Dr. Michael Hartmann

Granada, 02.2025

Editor: Universidad de Granada. Tesis Doctorales
Autor: Klaus Dieter Schwarz
ISBN: 978-84-1195-840-0
URI: https://hdl.handle.net/10481/105527

Acknowledgements

I would like to thank my esteemed supervisors, Prof. Dr. Daniel Arias-Aranda and Prof. Dr. Michael Hartmann, for their unparalleled expertise and guidance in shaping the research topic and their continuous support along the way. I would also like to express my sincere gratitude to my research colleagues from various disciplines and friends who have been an incredible source of inspiration and fruitful exchange. I am deeply grateful to my family for their unwavering emotional support. I would also like to thank the reviewers, the idea generators and all those who played a crucial role in the success of this project. I also thank all those who said I couldn’t do it. Their collective wisdom and encouragement were indispensable to my academic growth and the completion of this thesis.

Dear Dante, may the countless nights spent working be an inspiration to you. Dear Franziska, thank you for your patience, endurance and support.

Resumen

Esta tesis doctoral presenta el desarrollo de una plataforma modular basada en aprendizaje automático para la Inteligencia de Fuentes Abiertas (OSINT) destinada a la generación automatizada de informes de conciencia situacional en la gestión de desastres. A través de tres estudios progresivos, la investigación demuestra cómo implementar una plataforma escalable para analizar datos de microblogs mediante la combinación de detección de eventos, clusterización, clasificación zero-shot, generación aumentada por recuperación (RAG) y grafos de conocimiento. El primer estudio establece la base metodológica para la detección automatizada de eventos utilizando SimHash para la agrupación de mensajes similares y extracción de características mediante TF-IDF para identificar contenido relevante en datos heterogéneos. El segundo estudio mejora este enfoque implementando incrustación de frases con detección de comunidades para lograr una agrupación semántica más precisa e integrando grafos de conocimiento para una mejor estructuración e interpretación de la información extraída. El tercer estudio culmina en un sistema OSINT completamente modular (ODET) que incorpora clasificación zero-shot y generación aumentada por recuperación dentro de una arquitectura basada en agentes, permitiendo una adaptación flexible a diversos escenarios de desastres. La evaluación mediante casos de estudio reales, incluyendo el huracán Harvey y el terremoto de Turquía de 2023, demuestra una alta alineación factual con fuentes de referencia (AlignScore hasta 89%) y confirma la viabilidad de implementación en diferentes entornos de hardware, desde servidores en la nube hasta equipos de consumo. Esta tesis contribuye significativamente al avance de la recopilación automatizada de inteligencia en situaciones de crisis, superando los desafíos tradicionales de los procesos OSINT manuales mediante una plataforma modular, eficiente y adaptable que mejora la velocidad y precisión de la conciencia situacional en la gestión de desastres.

Abstract

This dissertation presents the development of a modular machine learning-based Open Source Intelligence (OSINT) platform for automated situational awareness reporting in disaster management. Through three progressive studies, the research demonstrates how a scalable system for analyzing microblogging data can be implemented by combining event detection, clustering, zero-shot classification, retrieval-augmented generation (RAG), and knowledge graphs. The first study establishes the methodological foundation for automated event detection using SimHash clustering and TF-IDF feature extraction. The second study enhances this approach by implementing sentence embedding with community detection for improved semantic grouping and integrating knowledge graphs for better structuring and interpretation of extracted information. The third study culminates in a fully modular OSINT system (ODET) that incorporates zero-shot classification and retrieval-augmented generation within an agent-based architecture, enabling flexible adaptation to various disaster scenarios. Evaluation through real-world case studies, including Hurricane Harvey and the 2023 Turkey earthquake, demonstrates high factual alignment with reference sources (AlignScore up to 89%) and confirms deployment feasibility across different hardware environments. The dissertation contributes to advancing automated intelligence gathering in crisis situations by overcoming traditional challenges of manual OSINT processes through an efficient, adaptable platform that enhances the speed and accuracy of situational awareness in disaster management.

Table of Contents

List of Figures

List of Tables

Chapter 1

Introduction

Disasters pose a significant challenge to societies worldwide [1, 2]. Natural disasters, technical incidents, and humanitarian crises require prompt and effective responses to mitigate damage and preserve human life. A key factor in the success of disaster response efforts is the ability to obtain an up-to-date and accurate assessment of the situation in real-time [3, 4, 5, 6]. Rapid situational awareness is paramount for the coordination of rescue operations, efficient distribution of resources, and prioritization of affected areas.

Conventional disaster management strategies depend on official reports and the direct observations of first responders [1, 2]. Although these methods provide valuable information, they are often slow, time-consuming, and geographically constrained. In practice, it can take hours or even days before official reports on the extent of a disaster are available, limiting the ability to respond in a timely manner [7, 3]. This can result in inefficient relief and potentially poor decision-making. In light of the escalating challenges posed by climate change, expanding urbanization, and the rise of geopolitical uncertainties, the need for an automated data-driven solution has become increasingly pressing [1, 2].

The increasing volume and fragmentation of disaster-related information have a significant impact on response capabilities. Millions of digital interactions occur during disaster events and responders struggle to extract relevant information from a flood of irrelevant or conflicting data. Traditional methods are unable to efficiently process this massive flow of information or do not consider the possibility, making it difficult to gain a holistic view of the situation and leading to delays and potential misinformation [4, 5, 8]. At the same time, the widespread availabil-

ity of publicly shared data offers an opportunity to improve disaster assessments, provided that they can be systematically analyzed and structured in a way that supports rapid and informed decision-making [3, 9, 10].

1.1 Background

The conventional approach to information collection in disaster management is labor-intensive and heavily relies on manual processes [1, 2]. Authorities and aid organizations typically employ standardized report formats to collect and process data by drawing on expert opinions, direct observations, or structured reports. These processes require a substantial degree of human analysis, which is conditionally scalable in disaster situations.

Another notable challenge pertains to the unstructured nature of information flows [8, 11]. While official reports adhere to a specific structure, valuable information, especially from the civilian population, is fragmented, heterogeneous, and difficult to interpret. In disaster situations, those affected often exchange initial impressions and observations on digital platforms; however, these data are rarely collected or systematically analyzed [3, 4, 5, 6]. Because the information-gathering process is time-consuming and labor-intensive, collecting initial impressions from civilians is often deemed too expensive and time-consuming. The absence of automated analysis processes necessitates manual filtration of the relevant content, a process that is both time-consuming and susceptible to errors.

Additionally, there are substantial data gaps in regions with inadequate infrastructure where sensors or official reporting chains are absent [9]. Under such circumstances, alternative sources of information can serve as valuable supplements. This prompts consideration of whether Open Source Intelligence (OSINT) – the systematic utilization of publicly available data – can contribute to enhancing disaster response.

OSINT for Disaster Management

Open Source Intelligence (OSINT) describes a systematic analysis of publicly accessible data sources to generate actionable information. Traditionally, OSINT has been used in intelligence analyses to gather information from media reports, social networks, satellite images, and other freely accessible sources [12, 13]. While OSINT has been used primarily in cybersecurity, military strategy, and crime fighting, it could also be a valuable addition to disaster management.

The primary advantage of OSINT is the abundance of real-time data, which, in theory, enables a more rapid assessment of the situation than traditional reporting methods. Social media has firmly established itself as an immediate platform for documenting events, with eyewitness reports, images, and videos, often providing valuable insights into the current situation in a disaster area, even before official agencies can respond [3, 4, 5, 6].

However, OSINT remains underdeveloped and unstandardized in disaster contexts [10]. The utilization of open data sources introduces substantial challenges, particularly concerning data quality, information overload, and disinformation [4, 5, 8, 11]. In its conventional form, OSINT is predominantly human-centered, and analysts manually evaluate the information and determine its relevance. Consequently, the current state of OSINT restricts its effectiveness as a real-time solution for disaster management by adding an extra layer of complexity and cost. The integration of automation and machine learning (ML)-based methods is imperative for enhancing the systematic application of OSINT in disaster management.

Social Media for Disaster Management

In recent years, social media has emerged as a significant source of information during disasters, although it is often underutilized [3, 9, 10]. Platforms such as Twitter/X facilitate the dissemination of information regarding disasters in real-time, frequently preceding the publication of official reports. Eyewitnesses, affected individuals, aid organizations, and news agencies use social media to disseminate initial assessments,

images, videos, and warnings. This direct and unfiltered form of communication can provide valuable information about the extent of an event, affected areas, and urgent needs of the population [10, 5].

Social media is not only a useful communication tool for authorities and aid organizations but also a potential data source for automated analysis [3, 10]. In the event of a disaster, activities on platforms such as Twitter/X skyrocket as people search for information, share their own observations, or post pleas for help [14, 3, 4, 5, 6]. In regions where official reporting is delayed or limited, social media can serve as a valuable supplementary source of information, offering rapid situational overviews. The interaction between users on social media platforms can also play a crucial role in verifying information, particularly rumors, or in providing contextual details that official agencies can use to enhance the credibility of information. Mechanisms, such as hashtags and geotags, facilitate targeted data searches, thereby streamlining the analysis process.

However, the use of social media as a data source is challenging [4, 5, 8, 11]. The sheer volume of unstructured data, often characterized by a high volume of misinformation, sensational reports, and manipulative content, complicates the identification of the relevant information. Furthermore, the absence of uniform data structures and standardized formats hinders data processing automation [8, 5]. Consequently, the development of intelligent algorithms to filter and structure social media data is imperative for its effective utilization in disaster management.

The utilization of social media for situation assessment in the event of a disaster demonstrates the potential of OSINT; however, it also introduces novel challenges. Suitable methods for automated event detection, analysis and verification of information must be developed to employ social media as a reliable data source in disaster management. This necessitates the implementation of machine learning (ML)-based methods that are capable of distinguishing relevant content from irrelevant or misleading posts.

Limitations of OSINT and Social Media for Disaster Management

Although OSINT is a valuable addition to disaster management when combined with social media, there are significant challenges to this approach. Open-data sources are not subject to quality control and often contain inaccurate, contradictory, duplicate, or misleading information. Although official reports are provided by verified institutions, OSINT data are compiled from a wide range of sources that cannot be easily verified. One of the greatest challenges is the flooding of information [4, 5, 8, 11]. In disaster situations, millions of users simultaneously generate content, making it difficult to extract relevant information. Many of these posts are redundant, misleading, or of no direct use in disaster management; therefore, efficient filtering is required. Without appropriate mechanisms for structuring and prioritizing the data, the content relevant to decision-making remains unused.

Additionally, the spread of misinformation is a key concern [4, 5]. Especially in disaster situations, false or manipulated content is circulated intentionally or unintentionally. Rumors spread quickly and create false impressions of the situation, leading to poor decisions by emergency services and aid organizations. Automated bots and targeted disinformation campaigns have exacerbated this issue.

In addition to data quality, another challenge is the lack of structure in OSINT information [8, 15, 10]. While traditional intelligence analysis uses well-defined metadata and standardized formats, OSINT data are fragmented, unsystematic, and heterogeneous. Social media posts have different spellings, incomplete contextual information, missing time, and location information. This variability makes the automatic processing difficult.

Machine Learning-Based OSINT for Disaster Management

Traditionally, OSINT analysis is performed manually, making the process time consuming and difficult to scale [12, 13]. Intelligence agencies use specialized analysts to review, verify, and crosscheck open source

information with other data sources. This approach requires significant human resources and is not suitable for disaster management because of the need for timely analysis, which cannot be achieved using manual methods.

To effectively use OSINT in disaster management, automated processes are needed to detect, extract, structure, and verify relevant data. These processes are based on ML-based algorithms that are capable of systematically analyzing large amounts of data.

Analyzing open data sources is challenging because it involves the processing of large amounts of unstructured information [8, 15, 10]. Machine learning (ML) offers new ways to analyze unstructured content, recognize patterns, and extract relevant information efficiently. Advances in AI technology, particularly transformer models, have the potential to significantly improve the quality of analysis by enabling more precise and scalable processing of large amounts of text. These developments have the potential to significantly increase the efficiency of the automated processing and structuring of OSINT data.

Despite the growing recognition of Open-Source Intelligence (OSINT) as a valuable tool for rapidly acquiring and accurately assessing situational awareness in disaster management, significant gaps persist in the systematic operationalization and automation of analysis processes. Conventional OSINT methods are frequently executed manually or constrained by rule-based procedures, rendering them unsuitable for the analysis of voluminous, unstructured data from social media and other open sources. This finding underscores the need for innovative approaches to address these limitations. This dissertation proposes a comprehensive solution to address these challenges by integrating advanced methods based on machine learning and modular architectures and aims to enhance the real-time analysis and processing of disaster data to facilitate more efficient and effective decision-making processes in disaster management.

1.2 Research Objectives and Research Questions

The objective of this dissertation is to develop an ML-based Open Source Intelligence (OSINT) platform for automated situational awareness in disaster management. The combination of machine learning (ML) and OSINT techniques is expected to automate the processing of large amounts of unstructured data, thereby enabling efficient, scalable, and precise detection of disaster events. The platform should be capable of extracting and structuring relevant information from microblogging data and providing automated reports.

In the context of this dissertation, three scientific studies were conducted and published that examine various aspects of the development of this platform.

#	Title	Journal	Quality Indicators
1	Towards Automated Situational Awareness Reporting for Disaster Management—A Case Study [16]	Sustainability	Impact Factor: 3.3 (JCR 2024); Q2
2	Harnessing OSINT in Disaster Management: Transforming Microblogging Posts Into Insightful Data Through Text Embedding [17]	Proceedings of International Conference on Theoretical and Applied Computing (Springer)	Peer-Reviewed Book Chapter
3	AI-Enhanced Disaster Management: A Modular OSINT System for Rapid Automated Reporting [18]	Applied Sciences	Impact Factor: 2.5 (JCR 2024); Q1

Table 1.1: Overview of the publications included in this dissertation

Each study addressed a specific research question that arises from the requirements for automation, scalability, and accuracy of OSINT-based systems in disaster management. The following section presents these research questions and the scientific contributions of the three studies.

Q1: **Can OSINT and ML be used for automated situational awareness reporting in disaster management?**
The first research question deals with the basic integration of OS-

INT into disaster management. The aim is to investigate how publicly available data sources, particularly microblogging platforms such as Twitter/X, can be used for the automated detection and analysis of disaster events.

- This question was investigated in Study 1, in which an event detection system was developed to enable the automated analysis of microblogging data. The methodology is based on TF-IDF for feature extraction, a SimHash-based clustering approach for grouping similar content, and a comparative evaluation of multiple ML-based summarization techniques for condensing the detected events.
- Study 2 builds on these results and improves event detection using text embedding and clustering algorithms. This allows for the semantic grouping of events, making the analysis more flexible and accurate. Furthermore, knowledge graphs were integrated to structure detected events and enhance analysis.
- Study 3 further integrates these approaches into a modular AI-based OSINT system (ODET). In addition to zero-shot classification, knowledge graphs, and retrieval-augmented generation (RAG), a modular agent-based architecture was implemented to enable scalable and automated disaster reporting.

Q2: **How can the accuracy of OSINT and ML-based automated situational awareness reporting for disaster management be improved?**

The second research question addresses the challenge of designing scalable OSINT-based systems that can efficiently process large amounts of data.

- In Study 1, an initial pipeline for processing microblogging data was developed. This lays the foundation for automation but also highlights the limitations in flexibility and scalability, as it relies on predefined term-based event detection methods.
- Study 2 focuses on improving event detection using embedding and clustering methods. This enables the system to analyze large amounts of unstructured data more efficiently and

structure the extracted information using knowledge graphs for improved interpretability.

- Study 3 presents the system as a complete solution for automated generation of situational awareness reports. Performance and deployment feasibility were examined by testing different computing environments, from powerful cloud servers to consumer hardware (MacBook Air M3). The modular architecture allows rapid adaptation to different disaster scenarios and ensures scalability across various hardware configurations.

Q3: **How can the scalability of OSINT and ML-based solutions for situational awareness reporting in disaster management be improved?**
The third research question addresses the accuracy of OSINT-based disaster analysis and examines how modern AI methods can help reliably identify and structure relevant information.

- Study 1 evaluated different methods for automated event detection as a proof of concept and investigated the feasibility of summarizing detected event clusters into situational awareness reports.

- Study 2 increased the accuracy by using text embedding and community detection-based clustering. By introducing semantic text embedding and community detection for event clustering, events can be interpreted and structured more accurately, allowing irrelevant or redundant data to be filtered out more efficiently.

- Finally, Study 3 integrated several AI techniques to further improve the accuracy and contextualization of disaster data. The combination of retrieval-augmented generation (RAG), knowledge graphs, and modular agent-based processing optimizes the analyzability of the data. The agent-based architecture of ODET increased its scalability, ensured long-term adaptability, and allowed flexible integration of new models. The quality of the generated reports was compared with Wikipedia event

summaries, achieving a high level of agreement (AlignScore up to 89%), confirming the effectiveness of the developed platform.

This dissertation contributes to the improvement of situational awareness reporting in disaster management through incremental development of a salable and modular ML-based OSINT platform. Study 1 laid the foundation for event detection and reporting. Study 2 improved the scalability and accuracy of the analysis by using advanced machine learning techniques. Finally, Study 3 integrated these advances into a modular and scalable system (ODET).

1.3 Delimitation

This dissertation deals with the development and evaluation of an Open Source Intelligence (OSINT) and machine learning-based platform for automated situational awareness reporting in disaster management. The focus is on the automated processing, analysis, and structuring of unstructured data from open sources, especially social media, to support decision making in crisis situations. AI-based methods, such as text embedding, clustering, and zero-shot classification, are used to extract relevant information and transform it into actionable situational awareness. This dissertation examines the entire process, from data aggregation to analysis and reporting, and evaluates the performance of the developed approaches based on empirical case studies.

This dissertation exclusively uses microblogging data (Twitter/X), specifically prebuilt freely available datasets, as a source for disaster analysis. Other digital or traditional news sources, such as newspapers, blogs, or TV reports, were not considered. This allowed for a focused investigation of the suitability of machine learning-based OSINT analysis for processing unfiltered, high-frequency social media data in real-time. Furthermore, no direct API access via Twitter/X API was used.

This dissertation focuses explicitly on the technological implementation and evaluation of OSINT-based analysis methods in disaster management. Sociological, political science, economic studies, business studies, and social analyses of OSINT or a disaster management platform

in the broader sense are not the subject of this dissertation. Similarly, there will be no discussion on legal, ethical, or any of the above aspects, and this dissertation focuses solely on the possibility of technical implementation.

This dissertation focused exclusively on disaster-related information. The analysis of misinformation, disinformation, and deliberate manipulation was not part of the research. While the system is technically capable of filtering irrelevant or potentially misleading content, the focus is on automated detection, structuring, and reporting of relevant events.

This dissertation develops and evaluates a modular AI architecture for OSINT analysis. It does not investigate a fully automated end-to-end solution but rather a modular and explicable AI system that enables transparent and traceable analysis.

This dissertation exclusively deals with text-based data processing using machine learning and natural language processing (NLP). Methods for image or video analysis have not yet been developed or evaluated. Furthermore, the system is based on existing open language models (e.g., zero-shot classification, RAG, knowledge graphs) without training its own language models from scratch.

The methods were explicitly designed for disaster management scenarios. Other application areas such as political event analysis, market surveillance, and cyber security have not been investigated. However, the developed platform is modular and flexible so that future adaptations for other purposes are possible.

Another focus was on the practical applicability and scalability of the developed methods. The platform was designed to operate efficiently on consumer hardware (e.g., MacBook Air M3) and cloud environments. The evaluation is based on case studies of real disaster events, with comparability to established information sources, such as Wikipedia, ensured by quantitative metrics.

1.4 Progress and Integration of Research

The research in this dissertation has evolved incrementally over several years, with each publication integrating new technological advances and advancing existing methodologies. The three publications that form the core of this dissertation logically build upon each other and incrementally contribute to the development towards an automated modular OSINT platform for situational awareness in disaster management.

The following publications are included in this dissertation:

Study 1: Towards Automated Situational Awareness Reporting for Disaster Management—A Case Study [16]

The first study laid the technical foundation for the use of Open Source Intelligence (OSINT) in disaster management. It focuses on the feasibility of automated event detection and text summarization of microblogging data. To this end, a system was developed that uses term frequency-inverse document frequency (TF-IDF) for feature extraction, employs SimHash-based clustering to group similar messages, and evaluates multiple extractive text summarization methods.

This initial study demonstrates the feasibility of automated situational awareness report generation using microblogging data and provides a processing pipeline for disaster information extraction. However, scalability and accuracy were still limited due to the reliance on term-based feature extraction and predefined lexicons for event detection.

Study 2: Harnessing OSINT in Disaster Management: Transforming Microblogging Posts Into Insightful Data Through Text Embedding [17]

The second study builds on the results of the first study and extends the methodology by using advanced machine learning techniques. While the first study relied on TF-IDF for feature extraction and lexicon-based event detection, this study introduced the utilization of text embedding and community detection-based clustering algorithms to capture and group events more accurately.

A major advancement is the use of semantic similarity models, which allow events to be grouped based on content rather than just term frequency-based methods. This significantly improves the scalability

and accuracy of the system by making event detection more robust to linguistic variation. In addition, knowledge graphs were introduced to structure extracted information and enhance interpretability.

Study 3: AI-Enhanced Disaster Management: A Modular OSINT System for Rapid Automated Reporting [18]

The third study summarizes the progress of previous work and develops a modular OSINT platform for disaster management. This platform, called the Open Source Intelligence Disaster Event Tracker (ODET), integrates the methods developed in the first and second studies into a scalable, agent-based architecture. It enhances event detection and situational awareness reporting through zero-shot classification, retrieval-augmented generation (RAG), knowledge graphs, and modular AI agents.

The developed platform was evaluated in two real-world case studies (Hurricane Harvey, Turkey Earthquake 2023) and achieved an AlignScore of up to 89% compared to Wikipedia event summaries. Furthermore, the scalability and adaptability of the system were demonstrated by deploying ODET on both consumer hardware (MacBook Air M3) and cloud infrastructure (A100 GPU in Google Cloud).

1.5 Theoretical and Technical Contributions

This dissertation provides both theoretical and technical contributions to the automation of Open Source Intelligence (OSINT) in disaster management. Although OSINT is traditionally used in intelligence contexts, its potential for automated disaster analysis has not been systematically explored [12]. This dissertation demonstrates that OSINT-based methods are not only suitable for analyzing microblogging data but also provide an efficient, scalable, and accurate solution for timely disaster reporting.

Technologically, this dissertation combines several machine learning-based methods into a modular framework, including zero-shot classification, retrieval-augmented generation (RAG), knowledge graphs, text

embedding and community detection for event grouping. These methods not only improve event detection but also enable structured and contextual reporting, overcoming the limitations of rule-based OSINT approaches.

Theoretical Contributions

A key theoretical contribution of this dissertation is the systematic application of OSINT in disaster management. While OSINT has traditionally been used in cybersecurity, intelligence, and investigative journalism, this thesis demonstrates that it can also be used to automatically collect and analyze disaster information. This differs from previous works, which have mostly been limited to general social media analysis without explicitly applying OSINT methods.

Furthermore, this dissertation demonstrates the effectiveness of text embedding and community detection for event detection. Although these methods have been used in other contexts, their application to event detection, and thus in the broader context of disaster management, has been little explored. This dissertation shows that they not only provide an efficient and accurate method for disaster event detection and grouping, but also enable the direct reusability of the generated vectors in RAG. This creates a synergistic advantage because the same vectors can be used for both event detection and generative reporting.

Additionally, this dissertation demonstrates that the use of quantized Instruct LLMs for zero-shot classification is an efficient method for the automated analysis of large amounts of unstructured data. Instead of relying on specifically trained classification models, the system uses generic language models that can be deployed ad hoc for specific classification tasks by targeting them using structured prompts and grammar rules. Model quantization also optimizes the processing power, increases flexibility, and enables resource-efficient processing on a variety of hardware, from high-performance cloud instances to consumer devices. This significantly expands the potential applications of OSINT analysis and enables rapid adaptation to new information needs in disaster management.

Finally, this dissertation demonstrates that a modular agent-based system is a useful approach for automated OSINT analysis. Although classical OSINT analysis relies on rule-based systems or manual processing, this dissertation provides a fully automated solution that adaptively combines different AI models and a scalable and flexible platform for disaster analysis.

Technical Contributions

The main technical contribution of this thesis is the development of the Open Source Intelligence Disaster Event Tracker (ODET). This system combined event detection, semantic clustering, zero-shot classification, and RAG-based reporting in a modular architecture. This enables fully automated and scalable OSINT analysis of disaster events.

To ensure an accurate and scalable OSINT analysis, ODET integrates several advanced techniques. The following sections describe the key technical components of the system and how they contribute to the automated processing of the disaster data.

Text Embedding and Clustering-Based Event Detection

In this approach, event detection is performed using text embedding that map message content into high-dimensional semantic vector space. These vectors were then clustered into thematically related groups using community detection-based clustering methods. In contrast to traditional rule-based methods, this approach enables the flexible and dynamic identification of relevant events, even in highly variable data sources, without the need for manual pretraining on similar events. Clustering ensures that only data with similar content are merged, increasing the processing efficiency and relevance of the extracted information while evening out story deviations and misinformation.

Zero-Shot Classification for Dynamic Categorization

The zero-shot classification used in this dissertation enables the automatic categorization of textual content without the need for prior training in specific disaster scenarios. This is achieved using a pretrained

quantized model in conjunction with a JSON-driven grammar structure that allows the system to adapt flexibly and quickly to new and unknown events. In contrast to traditional approaches that require extensive collection and labeling of training data, this method significantly improves the efficiency and scalability of automatic categorization.

Knowledge Graphs for Structured Semantic Analysis

Knowledge graphs were used for structured analysis of the extracted data. They link extracted entities, such as locations, timestamps, and event types, enabling a deeper semantic analysis of the data. The use of graphs helps to identify patterns and correlations that are not apparent in traditional text-based analysis. This method supports precise modeling of events and their causal relationships, enabling a more detailed and contextual assessment of the situation.

Retrieval-Augmented Generation (RAG) for Enhanced Reporting

Another innovative approach in this dissertation is the use of Retrieval-Augmented Generation (RAG) to improve the quality and consistency of generated reports. This involves integrating relevant external knowledge sources into the generation process to produce a more accurate and contextualized text output. RAG combines the benefits of knowledge retrieval with generative language models to improve the accuracy and consistency of the automatically generated reports. In combination with knowledge graphs, this approach enables the continuous improvement of analysis results by integrating aggregated cluster-overlapping knowledge, thereby increasing the reliability and depth of the generated situational awareness reports.

Modular System Design for Scalability and Flexibility

The developed system is characterized by a modular architecture that allows flexible exchange and adaptation of the analysis modules. This flexibility ensures that the system is suitable for different disaster scenarios, as well as for future extensions and technological developments. Scalability is ensured by the use of cloud-based technologies and optimizations for processing large amounts of data, which means that

the system can operate in real-time, even if the data volume suddenly increases during a crisis.

1.6 Thesis Structure

This dissertation consists of an introduction, three scientific publications, an English and a German conclusion chapter. This documents the publication-by-publication development process of an OSINT-based machine learning platform for automated situational awareness reporting in disaster management. Each chapter represents a central step in this development, from initial event detection to the integration of advanced classification methods towards a modular and scalable platform for disaster situational awareness reporting.

Chapter 1 – Introduction

This first chapter introduces the topic of the thesis, describes the relevance of Open Source Intelligence (OSINT) in disaster management, and presents central research questions. It explains the scientific and technical contributions of this dissertation and shows how individual scientific works are embedded in the overall context of the research.

Chapter 2 – Towards Automated Situational Awareness Reporting for Disaster Management—A Case Study

This chapter presents the first study, which lays the foundation for this dissertation. It explores the feasibility of Open Source Intelligence (OSINT) for automated disaster event detection and situational awareness reporting. The study focuses on processing microblogging data, particularly from Twitter/X, and introduces a system that applies TF-IDF for feature extraction, SimHash-based clustering to group similar messages, and a comparative evaluation of multiple extractive text summarization methods. The chapter first outlines the challenges of disaster intelligence, followed by a literature review of event detection and text summarization. The proposed approach is then introduced, detailing the preprocessing steps, event detection mechanism, and summarization methods. An experiment was conducted to evaluate the system using real-world disaster data, and the results were analyzed in

terms of event detection accuracy and reporting efficiency. The chapter concludes with a discussion of the study's contributions and limitations, emphasizing its role as a proof-of-concept for automated OSINT-driven disaster monitoring.

Chapter 3 – Harnessing OSINT in Disaster Management: Transforming Microblogging Posts Into Insightful Data Through Text Embedding

Building on the findings of the first study, this chapter enhances the scalability and accuracy of OSINT-based disaster analysis by incorporating advanced machine learning techniques. The study transitions from TF-IDF-based feature extraction and SimHash-based clustering to text embedding and community detection-based clustering, significantly improving event grouping and structured information representation through knowledge graphs. This chapter follows a structured approach, beginning with a discussion of OSINT's role in disaster management and the potential of text embedding for event analysis. It then presents a refined methodology, introducing semantic similarity models, zero-shot classification for event categorization without prior training, and Retrieval-Augmented Generation (RAG) techniques for improved contextual understanding. A case study demonstrated the effectiveness of these methods for structuring large-scale disaster data. The chapter concludes with an evaluation of the enhancements compared with the first study, highlighting improvements in scalability, adaptability, and classification accuracy.

Chapter 4 – AI-Enhanced Disaster Management: A Modular OSINT System for Rapid Automated Reporting

This chapter integrates the advances from previous studies into a fully modular OSINT system, the Open Source Intelligence Disaster Event Tracker (ODET). It begins with an overview of the technical challenges in scaling automated disaster intelligence and the need for a modular, agent-based architecture. It introduces ODET, detailing its flexible structure, AI-driven processing pipeline, and key components, including knowledge graphs, retrieval-augmented generation (RAG) for enhanced reporting, and zero-shot classification for dynamic event categorization.

The methodology section explains the system deployment, computational requirements, and adaptability across disaster scenarios. The chapter also presents two case studies—Hurricane Harvey and the 2023 Turkey Earthquake—to evaluate ODET's performance in real-world disaster contexts. The evaluation highlights the system's ability to generate highly accurate situational awareness reports, achieving an AlignScore of up to 89% when benchmarked against Wikipedia data. The chapter concludes by emphasizing ODET's modularity, scalability, and real-world applicability of ODET, positioning it as a robust AI-driven OSINT platform for disaster response.

Chapter 5 – Conclusions and Future Works

Chapter 5 synthesizes the findings of this dissertation, evaluating their collective contributions to OSINT-based disaster management. It begins with "Research Questions Revisited," assessing how the developed methodologies have addressed the core challenges of automation, accuracy, and scalability in situational awareness reporting. The chapter then presents a synthesis of key insights, highlighting the strengths and limitations of the proposed approaches. This includes discussions on system scalability, data availability, and methodological constraints. The final section outlines future research directions, suggesting enhancements such as improved misinformation detection, integration of multimodal data, and further automation of the reporting process.

Chapter 6 – Zusammenfassung und Ausblick

Chapter 6 provides a summary of the dissertation's findings and an outlook on future developments, written in German, to ensure broader accessibility. The chapter follows a structure similar to Chapter 5. It begins with a reflection on the research questions, assessing how the developed methodologies have addressed the core challenges of automation, accuracy, and scalability in situational awareness reporting. The chapter then presents a synthesis of key insights, highlighting the strengths and limitations of the proposed approaches. This includes discussions on system scalability, data availability, and methodological constraints. The final section outlines future research directions, suggesting enhancements such as improved misinformation detection,

integration of multimodal data, and further automation of the reporting process.

Chapter 7 – Conclusión y Futuros Trabajos

Chapter 7 provides a summary of the dissertation's findings and an outlook on future developments, written in Spanish, to ensure broader accessibility. This chapter is equivalent in content to Chapter 6, but presented in Spanish. The chapter follows a structure similar to Chapter 5. It begins with a reflection on the research questions, assessing how the developed methodologies have addressed the core challenges of automation, accuracy, and scalability in situational awareness reporting. The chapter then presents a synthesis of key insights, highlighting the strengths and limitations of the proposed approaches. This includes discussions on system scalability, data availability, and methodological constraints. The final section outlines future research directions, suggesting enhancements such as improved misinformation detection, integration of multimodal data, and further automation of the reporting process.

Chapter 2

Towards Automated Situational Awareness Reporting for Disaster Management—A Case Study[1]

2.1 Synopsis

Disasters do not follow a predictable timetable. Rapid situational awareness is essential for disaster management. People witnessing a disaster in the same area and beyond often use social media to report, inform, summarize, update, or warn each other. These warnings and recommendations are faster than traditional news and mainstream media. However, due to the massive amount of raw and unfiltered information, the data cannot be managed by humans in time. Automated situational awareness reporting could significantly and sustainably improve disaster management and save lives by quickly filtering, detecting, and summarizing important information. In this work, we aim to provide a novel approach towards automated situational awareness reporting using microblogging data through event detection and summarization. Therefore, we combine an event detection algorithm with different summarization libraries. We test the proposed approach against data from the Russo-Ukrainian conflict to evaluate its real-time capabilities and determine how many of the events that occurred could be highlighted. The results reveal that the proposed approach can outline significant events. Further research can be carried out to improve short-text summarization and filtering.

[1]This chapter is based on the open-access publication: K. Schwarz, D. Arias Aranda, M. Hartmann, "Towards Automated Situational Awareness Reporting for Disaster Management—A Case Study," Sustainability, MDPI, 2023. DOI: https://doi.org/10.3390/su15107968. Reused under the CC BY 4.0 license. Impact Factor: 3.3 (JCR 2024); Q2

2.2 Problems and Challenges

A disaster is a sudden accident or natural catastrophe that causes significant damage or loss of life [19]. Disasters can occur suddenly anywhere on Earth. In addition to natural disasters, there are also human-made disasters such as war [20]. Some might happen once, while others could be recurrent disasters such as earthquakes or tsunamis [21]. Even though it is impossible to avoid them, we must contain their consequences and act quickly once they occur. To achieve this, societies need to be as well-prepared as possible. Hence, disaster management refers to the totality of all coordinated measures in the areas of disaster prevention, disaster preparedness, disaster management, and post-disaster recovery, including the ongoing evaluation of actions taken in these areas. The phases of disaster management are structured in a continuous cycle — the Disaster Management Cycle — as depicted in Figure 2.1. The Disaster Management Cycle illustrates the ongoing process of reducing the impact of disasters. The overarching goal of disaster management is to reduce or avoid potential losses, provide rapid and appropriate assistance to victims, and ensure the fast rebuilding of infrastructure. The entire cycle includes the design of public policies and plans to mitigate the impacts on people, property, and infrastructure [22]. According to the definition of Al-Madhari et al. [19], disaster management actors are involved in the immediate response and recovery phases when a disaster occurs. For the most part, these are humanitarian organizations. The four phases of disaster management outlined here(in addition to the disaster itself) do not regularly occur in isolation. Often, the phases of the cycle overlap, and their lengths depend heavily on the severity of the disaster.

The Disaster Management Cycle includes four phases. These are prevention, preparedness, response, and recovery. The prevention phase contains measures to reduce vulnerability to the consequences of a disaster. These include, for example, efforts to protect against injury, loss of life, and property. Strengthening public infrastructure, medical care, and other actions that increase a community's resilience to disasters are also included.

The preparedness phase is about understanding how a disaster might

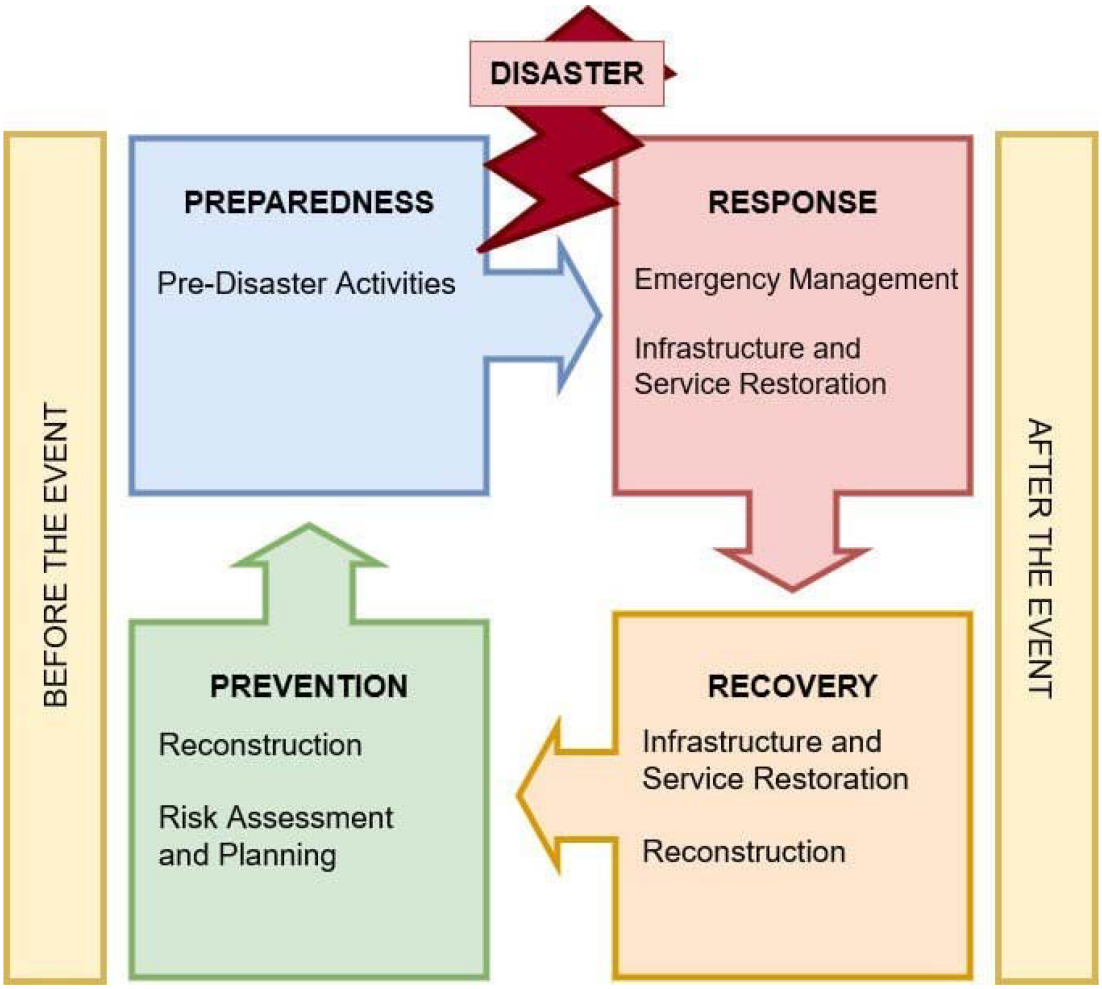

Figure 2.1: The Disaster Management Cycle

affect the community and how to build the capacity to respond to and recover from a disaster through education, outreach, and training. This may include business community involvement, pre-disaster strategic planning, and other logistical preparedness activities.

The response phase is the direct response to a disaster. It addresses the immediate threats the disaster poses through actions such as saving lives, meeting humanitarian needs such as food, shelter, clothing, public health, and safety, as well as cleaning up, assessing damage, and resource distribution. As the mission progresses, the focus shifts from addressing the immediate emergency to conducting repairs, restoring utilities, establishing public services, and completing cleanup. The most pressing emergency issues are assessed and addressed. There

is often a degree of chaos during this phase, which can last for a while depending on the disaster's nature and the amage's extent.

The recovery phase is the fourth and final phase in the Disaster Management Cycle. It is typically the most long-term phase in the cycle and usually lasts for decades. It requires thoughtful strategic planning and actions to address a disaster's more severe or permanent impacts. In this work, we focus on the preparedness phase. The preparedness phase has been chosen, as we take the first step towards automated situational awareness reporting using microblogging data through event detection and summarization. The approach we propose was developed using real data from past war events. The system proposed with this work is settled within the training phase of the preparedness cycle. Meanwhile, the long-term goal of this research is to develop a system that delivers sustainable improvement to the recovery phase of the Disaster Management Cycle. The National Incident Management System (NIMS) [23] defines preparedness as "a continuous cycle of planning, organizing, training, equipping, exercising, evaluating, and taking corrective action in an effort to ensure effective coordination during incident response".

This cycle depicts a broader system to prevent, respond to, recover from, and mitigate natural disasters, acts of terrorism, and other human-made disasters. There are five phases of the preparedness cycle [24]. These are:

- Plan
  - The planning phase basically covers the entire cycle. The strategic and operational planning carried out in this section defines requirements and provides a benchmark for evaluating implementations. The planning elements show which capacities must be available during a major disaster [25].
- Organize/Equip
  - Organization and equipment provide the human and technical resources needed to build and meet modernization and sustainability requirements. Organization and equipment include identifying the competencies and skills that staff should have

and ensuring that an organization has the proper personnel. It also includes identifying and procuring standard equipment that an organization might need in emergencies [25].

- Train
  - Training provides first responders, homeland security officials, emergency management officials, private and nongovernmental partners, and other personnel with the knowledge, skills, and abilities needed to perform key tasks required during a specific emergency situation [24].
- Exercise
  - Exercises enable entities to identify strengths and incorporate them within best practices to sustain and enhance existing capabilities. They also provide an objective assessment of gaps and shortfalls within plans, policies, and procedures to address areas for improvement prior to a real-world incident [25].
- Evaluate
  - The final phase of the preparedness cycle is evaluation. In this phase, organizations gather lessons learned, develop improvement plans, and pursue corrective actions to address gaps and deficiencies identified during exercises or real-world events [24].

Situational awareness is an essential point in most phases of the cycle. Accordingly, it becomes essential in the response phase. Targeted actions can only be taken with a precise overview of the situation. As pointed out by Arias-Aranda et al., flaws in the constant monitoring of all actions cause critical failures in achieving crucial milestones of humanitarian actions [26]. Therefore, fast and undistorted situational awareness reports can contribute to the improvement of humanitarian responses in a sustainable way.

Evaluating microblogging sites such as Twitter/X has become essential to obtaining situational awareness in disaster situations [27, 28, 29, 30, 31, 32, 33]. As Phengsuwan et al. pointed out in their Survey

about the use of social media data in disaster management,“Several publications have proposed the exploitation of social media data for disaster management, with Twitter/X being one of the most significant social media data sources used for disaster management. The temporal and spatial information extracted from Twitter/X is critical information for supporting decision-making in disaster management” [34].

The nature of these microblogging services enables their users to publish messages with a limited number of characters. However, this limited number of characters allows people to distribute status updates quickly and easily. These are not exclusively important or relevant messages. Most status updates are highly irrelevant and concern social or private matters. However, recent research shows that this type of social media yields critical information in disaster situations because people publish, share, and report critical information [35, 36, 37].

Nonetheless, the many news stories during a disaster event, the accompanying social statements worldwide, and the many social messages of those who have something to say that is unrelated to the topic are a problem in the evaluation of this goldmine that is yet to be exploited [12]. However, the number of status updates that occur and are relevant during disasters holds great potential for helping in disaster management in the long term. Occurring events can be found and tracked during a disaster.

According to the definition of Hasan et al. [38], an event in the context of social media is “[...] an occurrence of interest in the real world which instigates a discussion on the event-associated topic by various users of social media, either soon after the occurrence or, sometimes, in anticipation of it”.

By bundling and following these events, a news stream in the form of a story evolves, which is essential for decision-makers and first responders. This particularly timely stream of information can also help related work in a variety of other disaster-related fields and phases, such as the use of unmanned aerial vehicles (UAVs) for reconnaissance [39], vehicle routing and relief supply distribution [40], as well as volunteer assignment [41], to determine mission locations, mission requirements, and success rates. Furthermore, recordings of on-site

information can help to understand complex situations in retrospect. Finally, the resulting records can help ensure adequate preparation and training in the future after a disaster. However, this valuable opportunity comes with its problems. Even after filtering irrelevant messages, personal communications, spam, and background noise, much relevant information remains. Although unmanageable to organize manually, these pieces of information need to be logically summarized. Of course, evaluating and summarizing this information by hand would require a large workforce and a considerable amount of time and resources. This would be unsustainable since we are dealing with vast amounts of data in this context.

To assess the following research questions, this work aims to provide a novel approach towards automated situational awareness reporting using microblogging data through event detection and summarization, by combining an event detection algorithm with different summarization libraries.

- Is the system able to process the data of the previous time window in the current time window?

- What percentage of the events in a dataset that is as realistic as possible is the system able to detect?

- How could the results be improved?

This chapter addresses the first step towards automated situational awareness reporting for disaster management. As previously described, Twitter/X provides one of the most significant sources for this purpose. Although we think our approach can be applied to other text-focused social media, they are outside the scope of this chapter. The rest of this chapter is ordered as follows. In Section 2.4, we introduce the architecture of our proposed system. Section 2.5 and Section 2.6 explain the design of our experiments and discuss the results. Finally, we summarize the work in Section 2.7 and give an outlook.

2.3 Related Works

To discuss related works, the following subdivision is made. The first part discusses works with related goals in the same disaster management/preparedness cycle phases. The second part discusses works that describe related components. Therefore, Section 2.3.1 presents different event detection approaches, while Section 2.3.2 presents different approaches for automated text summarization. To conclude, in Section 2.3.3 a summary is given.

Works with related goals can be divided into two major categories. The first category represents ontology-based approaches [42, 43, 44], while the second category concerns AI-based approaches [45, 46, 47, 48].

Nguyen et al. [49] propose an interpretable classification-summarization framework that first classifies tweets into different disaster-related categories, and then, summarizes those tweets. Instead of focusing on performance measures, their proposed approach focuses on the decision-making process to give decision-makers an interpretable context. To achieve this, they employ a BERT-based multi-task learning approach and labeling. Accordingly, there are several limitations. In contrast to our approach, real-time capabilities are sacrificed to achieve higher accuracy. While we focus on delivering undistorted information fast to provide decision-makers with rapid situational awareness reports, their work mainly focuses on providing explanations or rationales. However, fast, undistorted reports can be used to better simulate the conditions of a disaster and to better train individuals in the appropriate behavior. It allows emergency personnel to practice reacting quickly until more detailed reports are available.

Mukherjee et al. [50] introduce an approach called MTLTS, a multi-task framework to obtain trustworthy summaries from crisis-related microblogs. Unlike other researchers in the field, they concentrate on a supervised approach. While they claim that their approach generalizes well across domains, outperforms the strongest baselines for the auxiliary verification/rumor detection task, and achieves high scores overall in the verified ratio of summary tweets, this approach requires time-consuming training and the labeling of datasets. The latter, itself,

requires a time-consuming process, and thus, a high level of human resources.

Rudra et al. [51] propose a framework that classifies tweets to extract situational information, and then, summarizes this information. Their proposed approach relies on their observation that tweets in disaster situations often contain disaster-related and non-disaster-related information alongside numerical information such as casualties. While this approach is developed to meet rapid performance, it does not perform event detection, and thus, relies on a keyword search provided by the Twitter/X API.

2.3.1 Event Detection

To ensure targeted and sustainable automated awareness reporting in disaster management, the initial detection and tracking of events is an important starting point. It is necessary to find out what happened, and when and where it happened. The microblogging service Twitter/X has emerged as the most popular and widely used service to obtain information about real-world events [52].

The works of Hasan et al. [38] and Li et al. [53] provide a comprehensive overview of the field of event detection methods applied to streaming data from Twitter/X. For this purpose, representative social media event detection techniques are extracted and presented in a categorized overview according to common features. Subsequently, various aspects of the subtasks and challenges related to event detection are discussed. Our proposed approach uses these results as a basis and overview of the commonly used event detection algorithms. For this purpose, we use a complex, large, real-world dataset for evaluation that is publicly available. In contrast to the presented work, we focus on fast, automated situational awareness reporting. The approaches presented in the included works often have entirely different goals and are only tested on small artificial or non-public datasets.

Unankard et al. [52] propose the approach of Location-Sensitive Emerging Event Detection (LSED). This approach is used to detect emerging hotspot events from microblog messages. The goal of the message detection in this work is specifically to help governments or organizations

prepare for and respond to unexpected events and disasters. The approach is also based on correlations between the user's location and the event's location obtained from microblog messages. The researchers report that this approach provides better real-world event leverage results than traditional TF-IDF and hashtag approaches. Still, they also disclose that it is "[...] difficult to effectively and efficiently process a large number of noisy messages." [52]. This leads to significant variations and errors in the results, which needs further optimization in future work. Knowing the location and time associated with the user and event can be vital to increase the event detection accuracy and tracking performance. The distinction from this work is that it refers to the early detection of possible disaster scenarios. It is not intended to be used for news reporting or even report writing, but solely for government agencies to issue early warnings.

Osborne et al. [54] propose a system based on local sensitive hashing (LSH). As in our proposed approach, the main focus is on disaster detection. The LSH approach is not new and is derived from previous work [55]. Since the accuracy of LSH is not high enough by itself, a content classifier based on a passive-aggressive algorithm is developed and trained. Similar to our proposed system, a dictionary is also used, which is used to help detect relevant tweets in the disaster domain. However, in contrast to our work, a weakly supervised approach is used here to learn new words successively.

2.3.2 Text Summarization

Summarizing Twitter/X messages is complicated. While the brevity of microblogging-based messages can have advantages in event detection, it poses many problems in summarization [56]. Short messages with a high repetition factor pose significant problems for standard algorithms. Rudrapal et al. [56] state that standard summarization algorithms can be roughly divided into two camps: those based on summary content or those based on events. Content-based approaches can be divided into two categories. So-called abstract summarization tries to create entirely new text from given text. This new text may contain components that were not present in the original text. Furthermore, there is the extractive summary. Here, the features are weighted again,

and essential parts of the text are extracted. Text parts with lower weights are omitted in this type of algorithm. Again, event category-based approaches can also be divided into two categories: generic and domain-specific approaches.

Li et al. [27] present two summarization approaches for events previously detected by the system. The first approach uses semantic types of event-related terms and ranks the messages based on the terms computed from the analysis of these semantic terms. The second approach uses graph convolutional networks and tweet relation graphs to detect hidden tweet features. Furthermore, Li et al. state that "There are many studies on text summarization, but only a few of them focus on social media" [27].

In this context, we propose, for our approach, the benchmarking of three standard libraries, namely SpayCi [57], Gensim [58], and NLTK [59], to evaluate the absolute baseline.

2.3.3 Summary

This section first discussed work with related objectives in the same disaster management/preparedness cycle phases. This was followed by a discussion of work describing related components. Each of the previous works shows a related component or approach. However, the works dealing with related components mainly show solutions at different disaster management/preparedness cycle phases or do not deal with disaster-based objectives. The approaches are either unsuitable in real-time, represent only an experiment, or are tested on small or non-public datasets.

On the other hand, works with related goals always have a deep neural network-based component in at least one part. As Li et al. state in their work, the disadvantage of deep neural networks is that they have velocity issues when used to discover new events from data streams in real-time. This work focuses on delivering undistorted information fast to provide decision-makers with rapid situational awareness reports. Fast, undistorted reports can be used to better simulate the conditions of a disaster and to better train individuals in the appropriate behavior.

Our approach allows emergency personnel to practice reacting quickly until more detailed reports are available.

2.4 Proposed Approach

The system we propose in this work consists of three main components that work together as follows. The first component involves loading and sorting data from the Twitter/X API. The data loaded in this way are first cleaned of usernames, hashtags, and other so-called stop words via a cleaning process. This is followed by the operation of the event detection and clustering module. At this stage, the previously cleaned data are assessed and rated using a term frequency-inverse document frequency algorithm and checked against a Dictionary containing unique disaster-related words. Simultaneously every tweet is hashed using a sophisticated hash function. Afterward, highly ranked data are clustered. Finally, the data are summarized. The system uses the Twitter/X API to preprocess simple filtering tasks. To achieve this, only relevant keywords, hashtags, and locations are tracked and retrieved from the Twitter/X API. The data are always fetched within a time window for processing. In the course of this work, the data from such a time window are called a chunk. A chunk contains all tweets from a given time window. One of the goals of the modules of our system is to process the messages of the last time window in the given time window. The following section describes the components of the system, as depicted in Figure 2.2.

2.4.1 Preprocessing

In this first processing step, data previously retrieved from the Twitter/X API are normalized, presorted, and cleaned. Since the text data to be processed are unfiltered raw data, it is to expected that they will contain a high proportion of noise and symbols that cannot be processed. Weblinks, hashtags, and so-called retweets,as well as usernames, which usually start with an @ symbol in Twitter/X data, are removed as a first step. Additionally, so-called stop words are filtered. Stop words are words whose processing is problematic because they occur particularly

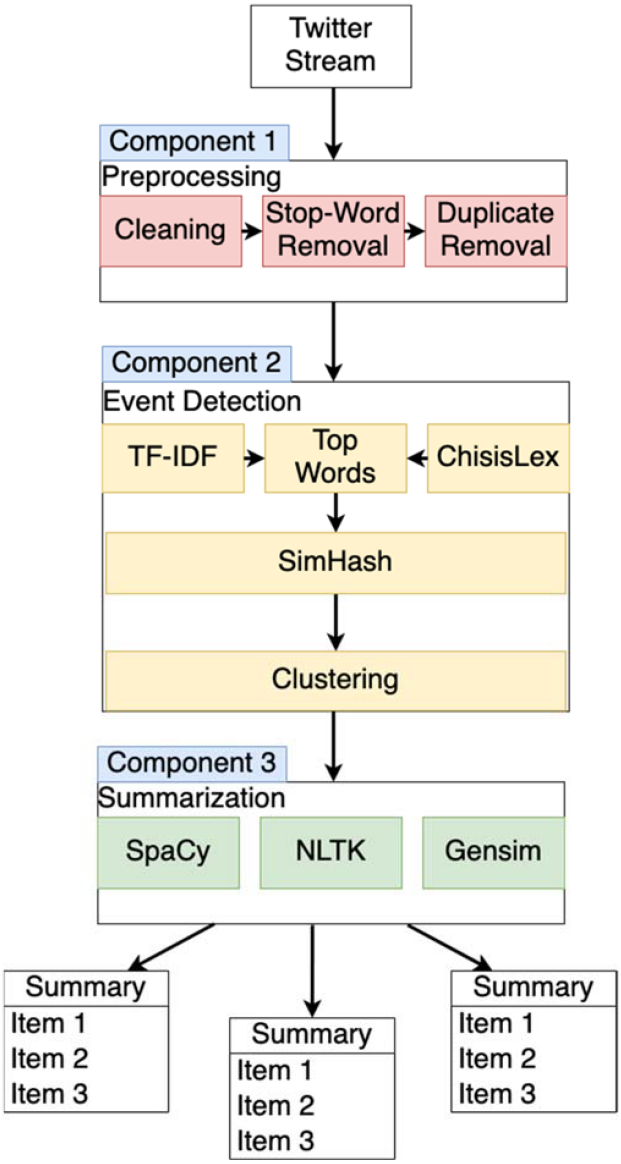

Figure 2.2: The proposed architecture

frequently but have no deeper meaning. Especially in emotional crisis situations, profanity can occur, so these words are also filtered.

Furthermore, only English-language messages are processed at this research stage, which is why all messages in other languages are filtered out. The filtering of the data is performed in three steps. In the first step, all messages are cleaned using the regular expressions described above, and capitalization is removed. After this, a specially created stop word dictionary is used to pre-filter the messages. The second step of filtering is performed using a stop word list from the SpaCy library [57].

Finally, duplicates are filtered by employing the Levenshtein distance [60]. The filtered data are then passed to the event detection stage.

2.4.2 Event Detection

The data are systematically evaluated for latent events in this second step. In the first step, the term frequency–inverse document frequency (tf-idf) [61] is calculated over the entire chunk. This is a statistical measure for assessing the relevance of terms in documents. To achieve this, first, the relative term frequency is calculated to prevent particularly frequently appearing words from being rated exceptionally highly. Secondly, the inverse document frequency is calculated. Here, the specificity of a term is measured. The matching occurrence of a rare term leads to a higher evaluation.

In our system, this first step creates a list of top words. This list of top words is now matched with CrisisLex [62]. CrisisLex contains a list of words that occur particularly frequently in disaster situations. Simultaneously, a hash is formed over each tweet in the current chunk using the SimHash algorithm [63]. Simhash is a so-called locality-sensitive hashing method (LSH). Unlike cryptographic hashes, where the slightest difference leads to significantly different hash values, with LSH, similar inputs lead to close hash values. This closeness can later be calculated easily and quickly as it only represents a one-dimensional distance. The advantages of the SimHash algorithm are a very efficient computation time, and that it takes so-called features in hashed elements into account so that unique semantic properties are not lost during the hashing process.

Relevant tweets are now filtered from the entire chunk using matching and scoring based on the previously selected top words. The previously created hash value is used to refine this set of relevant messages further and evaluate their significance.

The system now possesses an overview of the properties of the current chunk. At this stage, relevant tweets can be clustered according to their hash values. In our proposed system, we assume that important events generate many similar messages. Close hash values indicate similar

messages. These similar messages are bundled into clusters. A larger cluster potentially represents more critical events.

These clusters are now passed on to the summary module. At this point, the previously determined term frequency–inverse document frequency (tf-idf) is reset. Only the list of hashes of tweets that made it into the final clusters is kept, in order to merge similar clusters into a time-divided story later.

2.4.3 Summarization

According to the problems described in Section 2. We use three well-known standard libraries at this point to assess a baseline. For this reason, our system relies on so-called Extractive Summarization. Here, features are weighted again, and essential text parts are extracted. An advantage of this approach is that the previously created clusters do not have to be filtered. Due to the fuzziness of the extraction, duplicates are not significantly noticeable. Furthermore, former top words can be reused at this stage. For the summarization processes, the previously mentioned libraries SpayCi [57], Gensim [58], and NLTK [59] are used again. The verbal set of a cluster is now processed further. First, essential words are again determined and contextualized with the whole text. Afterward, in this way, weighted sentences are summarized. Since our approach aims to assess a baseline of fast summarization, no further filtering is conducted. Finally, clusters with similar hash values are merged in future time windows to obtain a chronological story.

2.5 Experiment

Testing automation for situational awareness reporting under realistic conditions is difficult. First, a reliable source of disaster events is needed. As McCreadie et al. [64] have pointed out, Wikipedia is not only well-known to contain newsworthy events, but is also a good source for disaster event information and a commonly used benchmark in related research [65, 66, 67].

The timeline of current events provided by Wikipedia was used as an evaluation basis to establish a measurement scale [68]. Wikipedia

maintains a broad list of events, including those found in the dataset we used. As a disclaimer, we want to mention that this list of events and their real-world backgrounds are not verifiable by us. It remains unclear whether the list is complete. Due to the nature of Wikipedia, which allows anyone to edit the records stored there, Wikipedia could only be used as a rough basis.

As a next step, a baseline of actual events that occurred in September 2022 was needed to evaluate the system. In order to achieve this, the timeline mentioned above was assessed by hand to extract every event that occurred in that timeframe.

Many available datasets contain either irrelevant data or are limited to data that include disaster events exclusively and,therefore, are already filtered to a high percentage. To test the proposed system adequately, we used the Ukraine Conflict Twitter/X Dataset [69]. This dataset has been updated daily since 23 February 2022, and is expected to continue receiving daily updates until 30 January 2023 [69]. It was published under the CC0 Public Domain license and currently includes about 16 gigabytes of records on various hashtags. The publishers of the dataset specify the “Twitter-Verse” as the geospatial coverage. The dataset is highly relevant to the nature of the proposed system.

Nevertheless, it contains misinformation, spam, duplicates, advertisements, noise, and private conversations. To test the proposed system, we used all the data of September 2022 from the above mentioned dataset. As a final step, the proposed system was executed three times consecutively, and each time, another summarization library was tested.

The partial dataset we used contains about two million tweets. The proposed system approached the whole month of September in timeframes of 24 h. For each timeframe, the proposed system started detecting events at exactly 00:00 and stopped at 24:00. To detect multiple co-occurring events. The dataset was continuously evaluated in chunks. The chunk being evaluated always represented a period of precisely 15 min.

2.6 Discussion

In this work, we aimed to provide a novel approach towards automated situational awareness reporting using microblogging data through event detection and summarization. Therefore, we combined an event detection algorithm with different summarization libraries. Furthermore, we tried to answer the following questions:

- Is the system able to process the data of the previous time window in the current time window?
- What percentage of the events in a dataset that is as realistic as possible is the system able to detect?
- Which of the three libraries for automated text summarization provides the best results?
- How could the results be improved?

Regarding the first question, “Is the system able to process the data of the previous time window in the current time window?”, this research shows that the proposed system can perform the task of processing the given data in the given time window very well. It even works faster than in the assigned time window. This result proves the practical feasibility of the proposed system and that it allows for the reliable processing of data in a given time frame.

Furthermore, we asked questions regarding what percentage of actual occurring events the proposed system can find in the given dataset, and which of the three libraries provided the best results. In order to be able to assess these questions, it was first necessary to determine the percentage of matches between events that actually happened and events detected by the summarization system on a given day. For this purpose, it was tested using a close-to-reality dataset. Table 2.1 shows the result of this comparison. The left column of the table shows the respective date of each day in the dataset analyzed by the system. Each day, we started detecting events at exactly 00:00 and stopped at 24:00 that day. In the second column is displayed the total number of tweets for that day. To the right of this column is the number of events that actually occurred on that day.

Table 2.1: Evaluation Results Text Summarization for September 2022

Date	**Num. of Tweets**	**Num. of Events**	**SpaCy**	**Gensim**	**NLTK**
Sep 01	48262	2	2	0	0
Sep 02	53668	4	4	2	3
Sep 03	40732	2	1	0	2
Sep 04	40151	2	3	1	2
Sep 05	44695	0	2	0	1
Sep 06	61619	2	1	0	1
Sep 07	46496	1	3	1	1
Sep 08	47830	0	0	0	0
Sep 09	70401	1	2	0	1
Sep 10	64953	6	9	3	4
Sep 11	67526	8	7	6	8
Sep 12	62944	2	5	3	2
Sep 13	63866	2	3	2	2
Sep 14	75392	2	2	1	2
Sep 15	67230	1	1	0	2
Sep 16	63125	1	2	0	1
Sep 17	49592	1	2	0	2
Sep 18	42426	1	1	0	1
Sep 19	49332	2	3	0	2
Sep 20	57409	3	5	1	3
Sep 21	104313	4	6	3	4
Sep 22	78507	0	2	0	0
Sep 23	77368	3	3	2	3
Sep 24	63866	0	0	0	0
Sep 25	52998	0	1	0	0
Sep 26	60625	1	1	1	1
Sep 27	68866	4	5	3	3
Sep 28	80527	3	3	2	2
Sep 29	66984	3	4	3	2
Sep 30	69930	3	3	1	1
	184,1633	64	134.375%	54.6875%	87.5%

To evaluate the events that actually occurred, we used the data from the “Timeline of the Russian Invasion of Ukraine 2022: Phase 3” of the Wikipedia current events website [68]. This timeline of the third

phase of the Russian invasion of Ukraine in 2022 covers the period from 29 August 2022, when Ukrainian forces retook significant territory in counterattacks in southern and eastern Ukraine, to the present day.

On the positive side, Wikipedia briefly summarizes each event, which allows for benchmarking of the events covered by the proposed system. On the negative side, Wikipedia does not list temporal information about the events. Thus, we were not able to determine the delay with which events could be detected for the first time or the delay with which they appeared in the report.

To evaluate the events that actually occurred, we manually extracted the data from the above timeline for the month of September. Then, we ran the system proposed in this work with one of each of the three libraries to detect and summarize events from the dataset. We were then able to compare how many hits each of the libraries achieved to determine what percentage of the actual number of events matched the number of events detected in the summaries.

The number of actual events varied from 0 to 8. A total of 1,841,633 tweets could be counted on the 30 evaluated days of September. The number of actual events that were manually extracted is 64 in the whole month of September. The SpaCy library detected 134.375 percent events, the Gensim library detected 54.6875 percent, and the NLTK library detected 87.5 percent. This means the SpaCy library achieved performance far above the expected average compared the actual number of events that were manually extracted.

However, the results obtained do not always correspond with those that can be read in the Wikipedia timeline, and include lots of sidenotes and unimportant information. The Gensim library performed the worst. With only 54.6875 percent, it detected just over half of the minimum detectable war events. The NLTK library achieved a good average with a performance of 87.5 percent. The most interesting result was achieved with the SpaCy library. Here, we were able to outperform Wikipedia by 34.375 percent. We see that the total number of tweets in a day ranged from 40,151 (on 4 September 2022) to 104,313 (on 21 September 2022). This variance of 64,162 tweets comes from the fact that war events do not happen every day.

To answer the last question regarding how the results could be improved, we assessed every report that the system produced. Every report consisted of multiple textual summarizations consisting of several sentences, where each summarization visualized a time window of 15 min. These time windows included multiple events that occurred simultaneously. These co-occurring events then formed stories over multiple time windows.

We manually read through every report to assess ways to improve the results. We can easily see the progression and importance of stories on days such as 16 September, with 63,125 tweets. Stories, as we find them in the media, can be well tracked in the summarizations and number of messages on Twitter/X. Take the example of September 16 mentioned above, which, according to Wikipedia, holds one significant war event. The SpaCy and NLTK summaries report a powerful explosion in Luhansk in the late morning of September 16 in the time window between 11:45 and 12:00 UTC. The events are clearly traceable; the location, the number of victims, and their names and affiliations are traceable in the text. In addition, SpaCy's early afternoon summary in the time window between 14:15 and 16:30 UTC reports that there were deaths and injuries from shelling in Velyka Kostromka Oblast. The text gives the exact number of dead and injured victims, as well as the exact cause and origin of the event.

The detection rates perform weakly when there is a lot of propaganda, spam, and hate speech among the messages. This significantly distorts the results. The dataset used has a special significance regarding its connection to the conflict in Ukraine. Unlike natural disasters, here, there are several parties with different views. The view of the attacker and that of the defense. At this point, the best way to improve the summarization results is to improve the filtering for hate speech, spam, and sentiment.

2.7 Summary

In this work, we provided a novel approach towards automated situational awareness reporting using microblogging data through event detection and summarization. We combined an event detection algorithm

with different summarization libraries. We then tested the proposed approach against data from the Russo-Ukrainian conflict to evaluate its real-time capabilities, and tried to answer multiple questions regarding its performance. In the process of this work, we evaluated our system using a dataset containing 16 GB of data and several million tweets. The system was shown to be able to detect significant events in almost every case in the process. However, minor events and side notes caused problems and could not always be appropriately detected, or were largely over-detected at times. A large amount of spam, noise, and hate speech also proved problematic. The automated awareness reports allowed conclusions to be drawn about the real-world events the system detected. However, propaganda and hate speech could also be clearly recognized for several days or hours with low message rates. Future work on this topic could focus intensively on spam filtering, sentiment analysis, rumors, and fake news detection. A possible future proposal should also address social data such as likes, retweets, sentiment, and social scoring to prevent hate speech from getting out of hand. We believe our solution is quite suitable for the creation of quick initial reports and can contribute to sustainable and improved social awareness in disaster management.

Chapter 3

Harnessing OSINT in Disaster Management: Transforming Microblogging Posts Into Insightful Data Through Text Embedding[1]

3.1 Synopsis

This chapter introduces a novel method for leveraging Open Source Intelligence (OSINT) in disaster management by analyzing social media data, notably from microblogging sites. Utilizing text embedding, event detection, and knowledge graphs, our technique converts Twitter/X messages into actionable insights, filtering out irrelevant content. Addressing the challenge of large-scale social media data's volume and hate speech, we employ an approach for sub-event detection by utilizing techniques like text embedding combined with clustering to categorize messages, simplifying data analysis, and enhancing pattern detection. Our approach surpasses traditional methods by automating data filtering and cleansing, reducing manual effort, ensuring data quality, and facilitating efficient resource utilization. It aids in identifying event-related clusters, tracking their development, and generating knowledge graphs for swift, accurate disaster management. This chapter details our clustering, cleansing, and processing strategies, setting the stage

[1]This chapter is reproduced with permission from Springer Nature: K. Schwarz, R. Creutzburg, M. Hartmann, D. Arias-Aranda, "Harnessing OSINT in Disaster Management: Transforming Microblogging Posts Into Insightful Data Through Text Embeddings," Springer eBook, 2025. DOI: https://doi.org/10.1007/978-981-97-6957-5_17. Licensed under Springer Nature License No. 5977080499639.

for future discussions on their application in comprehensive situation report generation.

3.2 Problems and Challenges

With the global increase in natural and man-made disasters, there is an urgent need for innovative disaster management approaches. Traditional methods often struggle to process large amounts of data during disaster events. Therefore, the integration of Open Source Intelligence (OSINT), microblogging platforms, and Large Language Models (LLMs) offers an opportunity to revolutionize disaster response mechanisms [70]. OSINT provides valuable information for tactical and strategic decision-making by accessing various public data sources, while microblogging platforms, especially Twitter/X, have become indispensable for rapid information dissemination and communication during crisis [70]. Furthermore, the advent of LLMs and artificial intelligence in the analysis of social media data has opened new avenues for gaining critical insights that enrich disaster management strategies [70].

The field of disaster management has recently benefited from significant advances in Open Source Intelligence (OSINT) and artificial intelligence (AI). One example is the integration of remote sensing data with two-dimensional directional wavelet analysis and open geospatial techniques [71]. In addition, the use of human sensor networks through Web 2.0, which has been highlighted as a means for people from around the world to participate in disaster relief, emphasizes the role of crowdsourcing and citizen participation in disaster management [72].

The role of big data analytics and the Internet of Things (IoT) in disaster management has been emphasized with the emergence of modern data analytics, services, and communication technologies that provide disaster management systems with new supportive data sources and fast, cost-effective data processing tools to support decision-making at all stages of a disaster [73]. In addition, a review of big data and processing frameworks for disaster response applications has highlighted the need for an efficient and real-time disaster response system to minimize the impact of disasters on people [74].

In the dynamic field of disaster management, the integration of Open Source Intelligence (OSINT), microblogging platforms, and large language models (LLMs) has transformed disaster management, revolutionizing situational awareness and response strategies [75]. OSINT, which uses public data sources, provides critical information for tactical and strategic decisions in crisis scenarios [75]. Microblogging platforms, such as Twitter/X, are central to the rapid dissemination of information and improve communication in crisis situations [76]. In addition, the integration of LLMs and AI in the analysis of social media data has provided new opportunities to gain valuable insights that enrich disaster management [77].

The synergy of these technologies has not only improved access to real-time data and communication but also fostered a new paradigm in environmental policy and disaster management [78]. Enhanced analytical capabilities through artificial intelligence, including LLMs, have significantly improved situational awareness and led to more effective disaster management.

The ongoing integration of OSINT, microblogging, and advanced AI technologies will further transform disaster management [79]. This integration promises to improve decision-support systems, especially in the context of smart cities, and expand preparedness and response capabilities [13]. The application of open geospatial models and AI in disaster warning and hazard analysis provides a new era for the management of complex emergency situations [80].

The approach outlined in this chapter offers significant advantages over traditional social media data analysis in disaster management. It automates the filtering and cleansing of datasets, minimizes the need for manual intervention, and ensures data quality. This method improves the detection of events and patterns, reduces the complexity of datasets, and enables more efficient use of resources. Autonomous Agents are used for the creation of knowledge graphs and situational awareness reports to improve the speed and accuracy of disaster management.

The remainder of this chapter is organized as follows. In Section 3.2, *Background and Importance of OSINT in Disaster Management*, we

explain the integration of text embedding with clustering algorithms, a strategy that we use to extract relevant information from large social media datasets. In this subsection, we explain the selection of specific embedding and clustering methods and highlight their effectiveness in identifying relevant events. In section 3.2, *Overview of Social Media as a Data Source*, we examine the increasing relevance of social media platforms in gathering actionable intelligence during crises. The focus then shifts to a real-world application in section 3.2 *The Russo-Ukrainian Conflict: A Case Study in OSINT Utilization*, where we analyze a specific instance to illustrate the practical implications and effectiveness of OSINT. Finally, section 3.2, *Purpose and Scope of this Chapter* outlines the objectives and boundaries of our research, providing a clear roadmap for what this chapter seeks to achieve. This introductory section is designed to provide a comprehensive overview of the critical aspects of OSINT in disaster management, framing the relevance and trajectory of our research.

Background and Importance of OSINT in Disaster Management

Open Source Intelligence (OSINT) refers to the collection and analysis of information from publicly available sources to provide actionable intelligence. OSINT has become increasingly important in the context of disaster management because it can provide important information for decision-making in crisis scenarios [12]. The open nature of OSINT enables the collection, processing, and correlation of data from multiple sources, which is particularly valuable in disaster management where real-time information is essential for effective response strategies [12]. In addition, the use of OSINT in disaster management is critical for planning and responding to natural disasters in advance to protect human life and infrastructure and minimize the risk of damage [13, 81, 82].

When combined with AI, OSINT becomes even more powerful for disaster management. The integration of AI technologies, such as large language models (LLMs), into OSINT has significantly improved situational awareness and led to more effective disaster management. AI

can maximize the use of social media data for disaster management, enabling authorities to make quick and effective decisions in the event of a disaster [83]. In addition, the technical and methodological improvement of hazard and disaster research, including the integration of AI, is considered a crucial aspect of disaster management [79].

The combination of OSINT and AI also offers opportunities to improve decision-support systems, particularly in the context of smart cities, and to enhance preparedness and response capabilities. The application of open geospatial models and AI in disaster warning and hazard analysis heralds a new era in the management of complex emergencies. In addition, the use of AI, the Internet of Things, and digital twin technologies in disaster risk management is an emerging field of research that holds great promise for improving disaster resilience and response capabilities [84].

This research was motivated by the fact that OSINT, with its open and collaborative nature, is becoming increasingly important in disaster management, especially when combined with AI technologies. The combination of OSINT and AI has the potential to improve decision-support systems, enhance preparedness and response capabilities, and ultimately contribute to more effective disaster management.

Social media platforms are rich repositories of real-time user-generated data that are invaluable in crisis situations. The following subsection focuses on these platforms and examines their role as important sources of data for disaster management.

Overview of Social Media as a Data Source

Social media has become an important source of data for disaster management scenarios, especially in recent years, for example, situational awareness reports on disasters such as wars, earthquakes, or tsunamis [85]. The use of social media in disaster relief has evolved significantly and is playing an increasingly important role in identifying security issues, locating displaced victims, and finding or providing assistance to those in need [86]. Social media platforms have become powerful tools for publishing disaster-related information and providing govern-

ments and relief organizations with real-time data for better disaster management [86].

In recent years, the presence of social media in disaster management has changed significantly, as shown by the evolution from a low presence in the 2008 Wenchuan earthquake to a strong force in the 2013 Ya'an earthquake [85]. Social media have played an important role in disaster management by enabling the public to contribute to disaster monitoring by reporting incidents related to disaster events [34]. In addition, social media has been used to inform emergency services of individual requests and reports, and there is great potential to aggregate this data and combine it with other relevant data sources to characterize affected areas [87].

However, the use of social media for disaster management is problematic. Concerns have been raised about the information quality of social media data, including variable data quality, the need for intelligent management of a large volume of social media feeds, and the need for manual review and verification of data [88]. Despite these challenges, social media, particularly platforms such as Twitter/X, have become an important component of situational awareness during disasters [89].

The use of social media data in disaster management is also a subject of research focused on understanding the quality of social media data in big data architectures and assessing the factors that influence the use of social media in disaster management by underserved communities [90][91]. In addition, the potential of social media data to provide customer insights for business decision-making has been recognized, highlighting the diverse applications of social media data beyond disaster management [90].

Social media have become an indispensable source of data for disaster management. It provides real-time information, enables public participation in disaster monitoring, and facilitates communication between the affected people and emergency services. Despite these challenges, the transformative role of social media in disaster management remains clear. After exploring the central role of social media as an appropriate data source for disaster management, the following section focuses on the practical application of these concepts in real-world scenarios.

The Russo-Ukrainian Conflict: A Case Study in OSINT Utilization

In the Russo-Ukrainian conflict, Open Source Intelligence (OSINT) and artificial intelligence (AI), particularly social media intelligence and natural language processing (NLP) using large language models (LLMs), are being used effectively for crisis management. OSINT, the collection and analysis of publicly available information, has been instrumental in providing real-time data and information for crisis management in the region. The integration of AI, particularly LLMs, has significantly improved disaster management capabilities in the context of wars, leading to more effective responses and decisions [79].

In the specific context of the Russo-Ukrainian conflict, social media intelligence has played a crucial role in providing valuable information for disaster management. Social media platforms are used to collect information and reports from affected areas, enabling authorities to monitor the situation, identify crisis areas, and coordinate relief efforts [92]. The use of AI systems to analyze data from social media platforms, such as Facebook, Twitter/X, and Instagram, has proven particularly effective in providing real-time information and situational awareness during conflicts [92].

In addition, LLMs such as BERT and the GPT series have demonstrated their ability to process and analyze large amounts of textual data, including war-related social media content [93]. These tools are used to gain valuable insights, identify trends, and understand public sentiment, contributing to a more comprehensive understanding of the evolving situation in the region.

Overall, recent research highlights the effective use of AI in disaster management and, in particular, the use of these technologies during the Russian-Ukrainian conflict. It also highlights the important role of AI in all phases of disaster management, leading to faster, more accurate, and better-equipped responses [79]. In addition, the application of AI technologies helps improve decision support systems, increase preparedness, and improve responsiveness in the context of conflict [79].

In conclusion, OSINT and AI, especially social media intelligence and LLMs, are effectively used for disaster management in the Russian-Ukrainian conflict. These technologies play a critical role in providing real-time data, improving situational awareness and decision-making, and ultimately contributing to more effective disaster management in the region.

Purpose and Scope of this Chapter

This project aims to explore and establish a novel approach using Open Source Intelligence (OSINT) in the field of disaster management, with a particular focus on the analysis of large datasets from social media platforms such as Twitter/X. We demonstrate how text embedding, in conjunction with advanced data processing techniques, can transform unstructured microblogging data into insightful situational awareness reports that are critical for effective disaster management and response.

To accomplish this, we focus on several key areas. Our first focus was on the use of embedding and clustering techniques for event detection. This involves identifying and grouping similar messages to reveal patterns and relationships in the data that may not be immediately apparent in manual analysis. The representation of these messages in a high-dimensional vector space facilitates the identification and tracking of events and topics that emerge during disasters.

In addition, the implementation of retrieval-augmented generation techniques was discussed. This approach is central to overcoming several problems associated with large language models, which are often a limitation when processing large datasets. By augmenting data retrieval, specific problems around the context windows of LLMs are treated.

Finally, several steps can be automated by developing and deploying LLM-based agents. These agents, based on a single large language model, play important roles in the processing of events, interpretation of knowledge graphs, and generation of situational awareness reports.

Another essential aspect of our methodology is the creation of knowl-

edge graphs. These graphs are not only a means of visualizing data but also an important tool for summarizing and interpreting the connections between events, findings, and relationships in the data. They enable the visualization of complex patterns and relationships that would otherwise remain hidden in the mass of data.

It is important to note that while this chapter lays the groundwork for the use of OSINT- and LLM-based techniques in disaster management, it is part of a larger research project. The real discovery, although not explicitly mentioned, is the potential of this method to revolutionize the way disaster management professionals use social media data. By transforming large unstructured datasets into structured, meaningful knowledge graphs, our approach aims to significantly improve the efficiency and effectiveness of disaster response and management.

In summary, this chapter not only presents a new approach for automated disaster data analysis but also aims to pave the way for future research and development in this area. The findings and methods presented here represent another step toward a more autonomous, data-driven approach to disaster management that takes full advantage of OSINT and AI technologies.

The remainder of this chapter is organized as follows. Section 3.3, *Literature Review*, discusses previous work in OSINT for disaster management, the role and challenges of social media in disaster response, advancements in text embedding, and the application of knowledge graphs. Section 3.4, *Proposed Approach* details our techniques for event detection, retrieval augmented generation, and use of LLM agents in constructing knowledge graphs. Finally, 3.5, *Experiment, Discussion, Summary, and Future Research*, presents our evaluation approach, discusses results, acknowledges limitations and potential improvements, summarizes our findings, and suggests directions for future research.

3.3 Literature Review

This section presents a comprehensive review of the existing literature in the realm of Open Source Intelligence (OSINT) in disaster management. It begins by examining 3.3.1 *Previous Work in OSINT for Disaster*

Management, focusing on pivotal studies and their contributions to the field. This is followed by an analysis of 3.3.2 *Social Media's Impact in Disaster Response and Challenges* on the role of social media and its specific challenges. Then 3.3.3 *Advances in Data Analysis: Text Embeddings and Clustering*, to show how recent technological advances have improved data processing capabilities. Finally, this subsection explores the 3.3.4 *Application of Knowledge Graphs in Disaster Management*, highlighting their importance in synthesizing and visualizing complex data sets. This literature review forms the basis for understanding the current OSINT landscape in disaster management and the potential of new technologies in this area.

3.3.1 Previous Work in OSINT for Disaster Management

The field of Open Source Intelligence (OSINT) has made significant contributions to disaster management by providing valuable information and tools for decision-making and response strategies. Several important studies have advanced the understanding and application of OSINT in disaster management, and are briefly presented and discussed below.

Jung et al. [13] proposed a conceptual framework for an intelligent decision-support system for disaster management in smart cities and highlighted the role of OSINT in disaster management. This chapter highlights the importance of using intelligent systems in disaster management, which can be improved by integrating the OSINT for data collection and analysis.

Similarly, Pastor-Galindo et al. [12] emphasized the benefits of OSINT in collecting, processing, and correlating information from cyberspace to generate knowledge, which is particularly important in disaster management scenarios. This chapter highlights the potential of the OSINT to provide valuable insights from open sources that can be critical to disaster preparedness and response.

Yu et al. [94] discussed the establishment of coordinating systems and platforms, such as the Sahana Free and Open Source Disaster Management System, for managing massive amounts of information during

disasters and highlighted the practical application of Open Source Tools in disaster management. This demonstrates the tangible impact of Open Source Platforms in facilitating effective disaster response.

Shah et al. [73] highlighted the utilization of various open source big data analytics tools within the scope of disaster management, indicating the growing relevance of Open Source Technologies in this field. This underscores the increasing integration of Open Source Analytics into disaster management practices.

Moreover, [78] discussed the emergence of a new philosophy of 'open source environmental governance and disaster management, emphasizing the linkage between disaster science, green computing, and open source movements, which has implications for the application of OSINT in disaster management. This highlights the broader implications of open source principles in shaping disaster management strategies.

In addition, Levy et al. [95] highlighted the promise of artificial intelligence applications for disaster risk management, indicating the potential synergy between AI and OSINT in enhancing disaster preparedness and response. This suggests the potential of integrating AI-driven insights with OSINT for more effective disaster management.

Overall, these studies collectively emphasize the significant contributions of OSINT to the current understanding and application of disaster management. They highlight the practical implications of leveraging Open Source Intelligence, big data analytics, and artificial intelligence to enhance disaster preparedness, response, and risk management.

3.3.2 Social Media's Impact in Disaster Response and Challenges

The role of social media in disaster management has been extensively researched, with studies highlighting its benefits and challenges.

Vieweg, Sarah, et al. [96] emphasized the integration of social media communication into the rapid assessment of sudden-onset disasters, highlighting its potential for real-time information gathering and situational awareness. Similarly, Nazir et al. [97] discussed the rise of digital humanitarian networks (DHN) in Southeast Asia, emphasizing

the compatibility of emerging social media and networks with intelligent data-centric systems for effective disaster management planning.

Kusumo et al. [98] highlighted the role of social media in providing a platform for sending alerts, identifying critical needs, and focusing responses, particularly in the spatial planning of flood evacuation shelters in Jakarta, Indonesia. Additionally, a study on the role of social media in disaster management in Bangladesh [99] discussed the positive roles of social media during the COVID-19 pandemic, such as practicing citizen journalism, creating social awareness, and maintaining communication during the emergency period.

However, research has also focused on the challenges associated with social media in disaster management. Cvetković et al. [100] examined the influence of demographic factors on the effectiveness of social media in forming the public regarding disaster risks, highlighting the need to address potential disparities in access and engagement. Ahmed et al. [101] enhanced the understanding of the complexity of community engagement through social media by conducting an empirical study on the existence of multiple communities of practice (MCoPs) during the Queensland floods.

In summary, studies on the role of social media in disaster management have highlighted the potential of social media for rapid assessment, information dissemination, and community engagement. They also highlight the unique benefits of social media, such as real-time communication and citizen engagement, while acknowledging challenges, such as inequalities in access and the complexity of citizen engagement.

3.3.3 Advancements in Data Analysis Techniques: Text Embedding

The field of artificial intelligence (AI) involving large language models (LLMs) and text embedding has made significant progress in recent years, especially in the context of disaster management. Several studies have focused on the use of large language models, such as BERT, FlauBERT, and ELMo, for various disaster management applications, such as information extraction, named person recognition, and social media analysis. Kozlowski et al. [102] experimented with classical

pre-trained embedding and contextual multilingual and French text embedding based on the BERT and FlauBERT language models to classify French tweets in environmental crises. Similarly, Zhang et al. [103] used text embedding of language models, including ELMo representations, to recognize named individuals in the context of cultural relics and Chinese text.

Moreover, the use of LLMs has been extended to social media analyses for disaster management. Silaa et al. [104] developed models to identify actionable information from disaster-related tweets using BERT, a graph attention network, and a relation network. Kumar et al. [105] focused on leveraging NLP approaches to validate the veracity of catastrophe-related information published on social media and enhance disaster response and management initiatives. This highlights the potential of language models in improving situational awareness and communication during disasters.

Prudhomme et al. [106] showed that the application of language models can be extended to disaster recovery and resilience. In addition, they developed a meta-model that provides domain-specific language for disaster management expertise, emphasizing the importance of modeling and simulation in disaster scenarios. Additionally, Cerna et al. [107] explored the use of LLMs for predicting peaks in firefighter interventions due to rare events to demonstrate the potential of NLP techniques in disaster prediction and response.

Overall, studies and findings in the areas of AI, LLMs, and text embedding, especially in the context of disaster management, show a growing interest in using AI for various tasks such as information extraction, social media analysis, disaster response, and resilience. These advances promise to improve disaster response, communication, and situational awareness through effective use of language models.

3.3.4 Application of Knowledge Graphs in Disaster Management

Ji et al. [108] provided a comprehensive review of knowledge graphs, covering topics such as knowledge graph representation learning, knowledge acquisition, and knowledge-aware applications. This survey

highlighted recent breakthroughs and future research directions, emphasizing the potential of knowledge graphs to facilitate data synthesis and visualization for disaster management.

Boné et al. [109] introduced "DisKnow," a system that utilizes natural language processing and knowledge graphs to extract disaster-related knowledge from socially driven sources. By employing knowledge graphs to present connected insights, the system offers real-time visual information about disasters and affected stakeholders, thereby enhancing the crisis management processes through effective data synthesis and visualization.

Son et al. [110] developed a knowledge graph for data management related to flooding disasters using open data. The paper emphasizes the importance of knowledge graphs in resolving heterogeneity among various disaster data and providing interoperability among domains, thus contributing to improved data synthesis and visualization for disaster management.

Ge et al. [111] focused on disaster prediction knowledge graphs based on multisource spatiotemporal information. This paper highlighted the need for knowledge graph models that focus on strategies and rules based on common semantics to provide specific disaster-handling strategies, thereby emphasizing the advantages of knowledge graphs in supporting decision-making through effective data synthesis and visualization in disaster emergency scenarios.

Wang et al. [112] introduced "GeoGraphViz," which is a geographically constrained 3D force-directed graph for knowledge graph visualization. The study emphasizes the creation of a knowledge graph connecting environmental datasets related to natural disasters, agriculture, and soil properties, highlighting the advantages of knowledge graphs in understanding the environmental impacts of disasters and visualizing complex relationships within disaster-related data.

In conclusion, the reviewed studies demonstrated the significant potential of knowledge graphs in data analysis related to disaster management. By leveraging knowledge graphs, researchers and practitioners

can effectively synthesize and visualize disaster-related data, leading to improved decision-making and disaster management processes.

3.4 Proposed Approach

In this section, we present the details of implementing our proposed approach, explaining the rationale and execution of each component. In 3.4.1, *Text Embedding and Clustering Techniques for Event Detection*, we explain the integration of text embedding with clustering algorithms, a strategy we use to extract relevant information from large social media datasets. In this subsection, we explain the selection of specific embedding and clustering methods and highlight their effectiveness in identifying relevant events. In 3.4.1, *Knowledge Graph Construction*, we explain the methods used to build comprehensive knowledge graphs with large language models. This section explains how these graphs are structured and used to visually represent complex data sets to enable more intuitive analysis and decision-making in disaster management. Finally, in 3.4.2, *Retrieval Augmented Generation (RAG)*, we discuss the integration of retrieval techniques to augment the generation capabilities of our models and highlight the synergistic benefits of this approach to improve data relevance and accuracy. We also concisely outline our use of autonomous agents.

3.4.1 Text Embedding and Clustering Techniques for Event Detection

The main goal of integrating text embedding and clustering techniques for event detection is to efficiently extract relevant information from large social media datasets. Embedding, which represent linguistic entities, such as phrases or words, as vectors in a high-dimensional space, make it easier for the model to understand and compare linguistic tokens. This process is important for combining discrete and continuous linguistic aspects, particularly in the dynamic and often ambiguous context of social media data.

In this chapter, embeddings serve dual purposes. First, it enables the storage and searchable indexing of data in vector databases enriched

with metadata from microblogging platforms. These metadata include details, such as reposts, comments, likes, and user information, which are invaluable for pattern recognition and location detection. Second, the clustering of these normalized vector embedding allows the data to be segmented into individual events. Each cluster represents a unique current event that can be filtered for relevance and tested for its applicability in disaster management. This process is critical to shifting through the vast, diverse, and often noisy social media landscape in order to gain actionable insights.

Selection of Text Embedding

For our embeddings, we selected models using the Massive Text Embedding Benchmark (MTEB) [113], specifically the all- MiniLM-L6-v2 and bge-base-en-v1.5 models. The all-MiniLM-L6-v2 model offers a good balance between speed and accuracy and is therefore suitable for initial clustering processes. Although the bge-base-en-v1.5 model is less efficient, it offers higher accuracy, which is critical for subsequent steps, such as call extension generation and situation report generation. Furthermore, we discuss future enhancements.

Clustering Techniques and Event Detection

Our clustering approach is based on the dot product and cosine similarity coupled with a threshold and simple sorting mechanism. The clustering process depends on the embedding generated by LLMs, with different models requiring different thresholds. This relationship underscores the need for a sophisticated understanding of the properties of embeddings to cluster and distinguish event-related messages effectively.

The core of event detection is the sorting of normalized vectors of length one derived from text embedding. This method facilitates the identification of clusters that are then analyzed for hate speech and disaster relevance, which are important indicators of the importance of an event in the context of disaster management. The fact that our approach does not detect fake news is a deliberate limitation in focusing our efforts on the most important aspects of event detection.

We used a text classifier to detect hate speech. This is a specially trained model that efficiently analyzes and classifies short texts in hate speech. At the time of writing, we are using the "Text Moderation" model from KoalaAI.

Constraints and Challenges

A major limitation of our approach is the trade-off between the model size and processing efficiency. Larger embedding models provide more accurate and detailed embeddings but at the cost of higher computational complexity. Our biggest challenge is to balance the need for a detailed analysis with a manageable computational capacity. However, for dedicated operations teams, this limitation may be less important.

Future enhancements

The landscape of large-scale language models is rapidly evolving, and newer models offer improvements in speed, efficiency, and accuracy. We expect these advances to significantly improve our approach and contribute to more effective and efficient disaster management strategies.

Knowledge Graph Construction

The construction of knowledge graphs plays a central role in disaster management projects by improving the quality of situational awareness reports. These graphs are used extensively by autonomous agents to generate queries for residual augmented generation (RAG) and produce comprehensive summaries. The repeated use of knowledge graphs throughout the reporting process underscores their importance in the overall methodology.

Data Integration and Preparation

The foundation of our knowledge graphs is the careful processing of event clusters and creation of metadata. Using a special grammar that we applied to our language model, we structured the metadata in a JSON format, as shown in Fig. 3.1. The key elements of this structure are indicators such as "disaster_related" to determine the relevance of a

message to disaster, "hatespeech_detected" to identify and ignore hate speech, "location_mentioned" to identify the actual location described, and "summarized_headline" to summarize the content of the message.

```
schema = {
    'type': 'object',
    'properties': {
        'disaster_related': {'type': 'boolean'},
        'hatespeech_detected': {'type': 'boolean'},
        'location_mentioned': {'type': 'string'},
        'summarized_headline': {'type': 'string'}
    },
    'required': ['summarized_headline', 'location_mentioned', '
       hatespeech_detected', 'disaster_related']
}
```

Figure 3.1: JSON Schema for Metadata Generation.

Methodology for Creating the Graphs

To create these knowledge graphs, we followed a methodical approach in which we converted each message of a cluster into topic-relationship-target triplets. This transformation created a branched network of interconnected information nodes. Furthermore, triplets can serve as a means of weighting and linking. The structure of these triplets allows for a nuanced representation of complex disaster scenarios and makes them indispensable tools for insightful data analysis.

Impact on situational awareness

Knowledge graphs play a critical role in the creation of situational awareness reports. By providing a structured and interconnected dataset, they enable autonomous agents to efficiently navigate a vast information landscape. This structured approach ensures that situational awareness reports are not only comprehensive but also contextual and accurate, improving the decision-making process in disaster management. In addition, the triplets we created allow weighting and, thus, finding, linking, and weighting individual subtopics within a topic cluster.

3.4.2 Retrieval Augmented Generation (RAG)

The field of Retrieval Augmented Generation (RAG) was excluded from this work. This work was primarily intended to serve as a basis for

further research. The current focus of this work is to create the necessary conditions for the use of RAG with maximum accuracy in future applications.

RAG plays a crucial role in the post-event detection phase of our system. After event detection, sorting, filtering, and knowledge graph generation, RAG is used to generate situation reports. These reports are essential for comprehensive disaster management and response.

A key feature of the proposed approach is the use of autonomous agents. These agents use the metadata associated with each message within an event cluster, as well as insights from knowledge graphs, to formulate specific contextual queries. The nature of these queries depends on the type of disaster that is being investigated. An agent who understands the structure of situational awareness reports coordinates the other agents and ensures that the reports are as complete as possible. The complicated mechanism underlying this process will be discussed in detail in future studies.

The key point of the RAG in our framework is its application in the preparation of refined management reports. Their accuracy and level of detail are determined by the interaction of several factors: the knowledge graph, characteristics of the detected event, and automatically generated questions. This complex interplay is crucial for the adaptation of the RAG to specific requirements of disaster management and situation analysis.

We now discuss our experiments, methods, and results. We then critically analyze the results in the discussion and consider their implications in a broader scientific context. This leads us to the summary, in which we briefly summarize the most important discoveries and findings. Finally, the chapter ends with the outlook, in which we look to the future and speculate on future research directions and potential applications arising from our findings, laying the groundwork for further investigation and exploration in this area.

3.5 Experiment, Discussion, Summary and Future Research

In this section, we first present our Evaluation Approach, where we discuss the use of the extensive Twitter/X dataset on the Ukrainian conflict as a robust platform to test our system's effectiveness, focusing on metadata completeness, knowledge graph plausibility, and data processing efficiency. Second, in the Discussion of the Results section, we highlight the capabilities of our system and its challenges. In the Limitations and Opportunities section, we acknowledge the limitations of relying on publicly available datasets, the difficulties in verifying event authenticity, and proposing future improvements. Finally, the recommendations for the Future Research section suggest possibilities for future researchers and express our interest in expanding the scope and capabilities of our disaster management approach.

3.5.1 Evaluation Approach

To evaluate the effectiveness of our proposed system, we used an extensive Twitter/X dataset on the Ukrainian conflict [69]. This dataset, updated daily from February 23, 2022, provided a robust platform to test our approach, particularly in contrast to the limited data access via the Twitter/X API. We deliberately avoided using the Twitter/X API because it is highly restrictive [114, 115, 116].

Our evaluation standards focused on the completeness of the metadata, plausibility of the constructed knowledge graphs, and efficiency of data processing. Following the methodology in [16], we divided each day into 15-minute intervals and analyzed each segment in a similar time frame. This approach ensures that the delay in reporting is no more than 15 min, which is in line with our goal of processing and analyzing the data in real-time.

3.5.2 Discussion of the Results

The results showed that our system could process data from an entire day in less than 15 min, far exceeding our original goal. The

analyzed metadata and knowledge graphs were largely plausible (over 90% accuracy) and complete. Problems occur when processing tweets with the excessive use of hashtags, but these can be mitigated by a simple filtering mechanism. The problems with hate speech documented in [16] can also be effectively solved using our approach with two filters. However, assessing the authenticity of individual events proved challenging owing to widespread propaganda during the conflict, emphasizing the need for neutral and comprehensive data sources for an accurate assessment.

3.5.3 Limitations and Opportunities

One of the main limitations of this chapter is that we relied on publicly available datasets, which may not contain all relevant data owing to API limitations. Additionally, verifying the authenticity of events in a politically charged and propaganda-driven environment remains a challenge. Future improvements could include the development of more sophisticated filters for hashtag sanitization and working with independent agencies for more neutral and comprehensive data collection. Verification of authenticity is critical to the robustness and reliability of our disaster management system.

3.5.4 Recommendations for Future Research

Future research should focus on expanding the scope and capabilities of our disaster management approach. A key recommendation is to improve hashtag-filtering mechanisms to address the challenges posed by the excessive use of hashtags in social media datasets. This refinement significantly improves the accuracy and relevance of metadata extraction and knowledge graph creation. In future work, we will focus more intensively on Retrieval Augmented Generation (RAG) and optimize its integration into our existing system to improve the quality and efficiency of situational awareness reporting. In addition, we want to work intensively on the automated presentation of situations by autonomous agents. These improvements will not only increase the robustness of our system but also extend its applicability to a wider range of disaster management scenarios.

Chapter 4

AI-Enhanced Disaster Management: A Modular OSINT System for Rapid Automated Reporting[1]

4.1 Synopsis

This chapter presents the Open Source Intelligence Disaster Event Tracker (ODET), a modular platform that provides customizable endpoints and agents for each processing step. ODET enables the implementation of AI-enhanced algorithms to respond to various complex disaster scenarios. To evaluate ODET, we conducted two case studies using unmodified AI models to demonstrate its base performance and potential applications. Through our case studies on Hurricane Harvey and the 2023 Turkey earthquake, we show how complex tasks can be quickly broken down with ODET while achieving a score of up to 89% using the AlignScore metric. ODET enables compliance with Berkeley protocol requirements by ensuring data privacy and using privacy-preserving processing methods. Our results demonstrate that ODET is a robust platform for the long-term monitoring and analysis of dynamic environments and can improve the efficiency and accuracy of situational awareness reports in disaster management.

[1]This chapter is based on the open-access publication: K. Schwarz, K. Bollens, D. Arias-Aranda, M. Hartmann, "AI-Enhanced Disaster Management: A Modular OSINT System for Rapid Automated Reporting," Applied Sciences, MDPI, 2024. DOI: https://doi.org/10.3390/app142311165. Reused under the CC BY 4.0 license. Impact Factor: 2.5 (JCR 2024); Q1

4.2 Problems and Challenges

Natural and man-made disasters pose increasingly complex challenges to disaster management. With the increasing frequency and intensity of such events, rapid and accurate information gathering is critical for timely decision-making and action [1, 2]. In this context, microblogging platforms such as Twitter/X have proven to be valuable sources of real-time information, often reacting to unfolding situations faster than traditional news sources by broadcasting eyewitness observations [3, 4, 5, 6]. However, these platforms present challenges [4, 5]. Large amounts of unstructured data must be processed, and relevant information must be filtered from a multitude of irrelevant or erroneous posts [8, 11].

Automated systems using artificial intelligence (AI) offer a promising approach to accelerate and optimize the analysis of such data volumes. In particular, techniques such as text embedding, community-based clustering, and zero-shot classification make it possible to identify semantically relevant content and group events and present them in a structured way without the need for models that have been specifically trained for the respective event [8, 11]. These techniques have the potential to form the basis for efficient and rapid reporting of disasters.

This chapter presents the Open Source Intelligence Disaster Event Tracker (ODET), a modular platform that provides customizable endpoints and agents for each processing step. ODET enables the implementation of AI-enhanced algorithms to respond to various complex disaster scenarios. To evaluate ODET, we conducted two case studies using unmodified AI models to demonstrate its base performance and potential applications. Through our case studies on Hurricane Harvey and the 2023 Turkey earthquake, we show how complex tasks can be quickly broken down with ODET while achieving a score of up to 89% using the AlignScore metric [117]. ODET enables compliance with Berkeley protocol requirements [118] by ensuring data privacy and using privacy-preserving processing methods.

Disaster management includes prevention, preparedness, response, and recovery, which together form the disaster management cycle [19]. ODET can be used in the response and recovery phases to provide

quick and accurate situational reports. In addition, ODET offers potential for the training phase by providing realistic scenarios and data-driven analysis. This not only improves the preparation and training processes but can also be used to simulate real disasters and fine-tune models.

This chapter contributes to the advancement of OSINT methods in the field of disaster management by introducing ODET, a modular platform that can be used to enable faster situational awareness in disaster scenarios.

The remainder of this chapter is organized as follows: In Section 4.3, the current state of research on Open Source Intelligence (OSINT) and artificial intelligence in disaster management is discussed in detail. Section 4.4 details the materials and methods used. Section 4.5 discusses the execution of case studies of Hurricane Harvey and the 2023 Turkey earthquake using ODET. Finally, in Section 4.6, we present and discuss the main results.

4.3 Literature Review

Our research builds on earlier work on Open Source Intelligence (OSINT) for disaster management, sentence embedding, community detection-based clustering, and zero-shot classification using large language models (LLMs).

4.3.1 Open Source Intelligence in Disaster Management

Social media, particularly Twitter/X, is a valuable tool for gaining situational awareness during disasters by aggregating information from numerous sources in real-time [3]. This information can be crucial for crisis responders but may not be presented in a useful format, necessitating platforms such as SMDRM (social media disaster risk management) for streamlining processing [3]. In disaster-prone countries, such as Indonesia, India, or the Philippines, Twitter/X reactions can provide insights into community responses, with sentiment analysis and classification methods aiding in understanding these reactions [9]. During specific disasters, such as earthquakes, Twitter/X discussions

can offer vital knowledge for rescue planning, with multimodal analysis helping identify important disaster topics [14]. Moreover, Twitter/X data can be used to monitor disasters and predict hazards using machine learning techniques that aid in filtering and categorizing related tweets [10].

Arapostathis et al. [7] developed a method to rapidly process and visualize Twitter/X data linked to fire events, therefore reducing the processing time to less than 2 h. Lorini et al. [3] created a platform called SMDRM (social media disaster risk management) to streamline the near real-time processing of text and images from Twitter/X during specific events. Yahya et al. [9] analyzed Twitter/X sentiments related to disaster management in Indonesia and found approximately 2081 positive and 605 negative sentiments with an accuracy of 84%. Can et al. [14] investigated Twitter/X discussions related to the 2020 Izmir earthquake by examining the spatiotemporal distribution of earthquake rescue and non-rescue terms. Baig et al. [10] explored the use of Twitter/X in disaster research by focusing on recent machine learning, deep-learning, and disaster-prediction techniques.

In summary, this research has demonstrated the utility of Twitter/X as a real-time source of information during disasters, with applications ranging from situational awareness and community response analysis to hazard prediction and rescue planning. Studies have employed methods such as sentiment analysis, multimodal analysis, and machine learning techniques to process and extract valuable insights from Twitter/X data, with notable success in reducing the processing time and improving the accuracy. However, there appears to be a gap in the research that focuses on the dynamic evolution of Twitter/X discussions throughout the entire disaster life cycle.

4.3.2 Sentence Embedding, Community Detection-Based Clustering, and Zero-Shot Classification Using Large Language Models

Sentence embedding captures the semantic meaning of sentences and is crucial in entity matching by transforming semantic meaning into a high-dimensional vector space [119]. Community detection-based

clustering is a fast method for grouping related entities represented by their vector embeddings [119, 120]. In the context of open source investigations, sentence embedding, and community detection-based clustering are used to extract and analyze information, with the goal of automating the process and maintaining a high alignment with reality [15]. Zero-shot classification using large language models (LLMs) has been applied across various domains to address the challenges associated with limited annotated datasets. Mugeni et al. [119] addressed entity matching in record linkage using context-aware sentence embeddings and graph clustering. Liu et al. [120] employed node embeddings and infection subgraphs for real-time rumor containment. Moura et al. [121] found that transformer-based models outperformed classical approaches for dialog data annotation, whereas Liang et al. [122] proposed an unsupervised Graph Clustering Network (GCN)-based community detection method. Kalra et al. [123] improved user clustering by integrating social context features into transformer models. In the zero-shot classification domain, Galeano et al. [124] explored LLM-based tasks, such as summarization and classification, whereas Arco et al. [125] combined LLM predictions with supervised models for better classification outcomes. Dr'apal et al. [126] developed an LLM framework for legal experts to assist with data coding and classification. Wang et al. [127] validated the performance of GPT models in text classification, and Raja et al. [128] automated article classification using LLMs.

In summary, recent research has demonstrated the utility of sentence embedding, community detection-based clustering, and LLMs in various applications, including entity matching for data quality improvement, real-time rumor containment in social networks, and open source investigations. In particular, zero-shot text classification allows rapid adaptation to new tasks without the need for task-specific training data. Studies have explored LLM-based subtasks, aggregated predictions, and collaboration with other models while also validating and comparing their capabilities with state-of-the-art methods across a range of tasks.

4.3.3 Related Works

This section summarizes recent related works and evaluates them based on their methodologies, performance metrics, and datasets used. We evaluated each study based on several criteria: the effectiveness of the technology in processing large-scale social media data, the sophistication of the algorithms used for classification or clustering, and the clarity of the reported performance metrics. Additionally, the size of the dataset is considered, as larger datasets generally offer more reliable assessments of a system's performance.

Afyouni *et al.* [129] developed an approach for event discovery by utilizing unsupervised machine learning and Natural Language Processing (NLP) algorithms to compute events' lifetime and spatial spanning. Their incremental clustering technique employs temporal sliding windows to update the discovered topic clusters with incoming social streams (e.g., tweets). They evaluated their system on a subset of 130,096 tweets and achieved an F1-score of 0.86, with a default number of five clusters.

Karimiziarani and Moradkhani [45] utilized Natural Language Processing (NLP) techniques, including sentiment analysis, topic modeling, and text classification, to process and analyze social media data during Hurricane Ian. Although specific numerical performance scores, such as accuracy, F1-score, and precision, were not mentioned, the system was applied to 21 million tweets collected between 24 September 2022 , and 6 October 2022, during the hurricane.

Ni *et al.* [130] developed a system using Natural Language Processing (NLP) techniques such as Bi-LSTM and Conditional Random Fields (CRF) for extracting structural information from textual emergency plans. The system incorporates scenario matching and semantic relevance matching to generate appropriate emergency plans (EPs) from historical data, specifically for unconventional emergencies. It achieved an F1-score of 0.941 and was tested on 9891 emergency plans.

Afyouni *et al.* [131] proposed a system for event detection from social media. Their system integrated a semantic keyword generation tool using KeyBERT for dataset preparation. Event detection is performed by

CNN and bidirectional LSTM, whereas hierarchical density-based spatial clustering is used for location-inference of events. They conducted experiments over a streamed batch of 4703 tweets, achieving an F1-score of 0.785 in their supervised approach and 0.7 in their unsupervised approach.

Huang *et al.* [132] developed a system integrating Natural Language Processing (NLP), the Analytic Hierarchy Process (AHP), the Entropy Weight Method (EWM), and the Grey TOPSIS method to quickly assess the relief needs of disaster areas and accurately distribute relief materials. The evaluation focused on relief demand urgency rather than specific numerical performance metrics, such as F1 or accuracy. The system was tested on 251,547 posts collected from disaster-affected areas.

Contreras *et al.* [133] evaluated a pre-trained sentiment analysis (SA) model developed by the no-code machine learning platform MonkeyLearn. The model was tested on 695 tweets using hashtags related to the Albanian Earthquake posted between 26 November 2019 and 3 February 2020. The model achieved an accuracy of 63%, indicating that it is acceptable for quick estimation of text polarity during the emergency and early recovery phases after an earthquake.

Zhou *et al.* [134] compared several NLP models, including BERT, RoBERTa, DistilBERT, XLNet, and ALBERT, to identify which model performs best in classifying rescue request tweets. The best-performing model was BERT with a CNN classifier, achieving an F1-score of 0.919. The evaluation was conducted on 3191 manually labeled tweets from Hurricane Harvey.

Wahid *et al.* [135] employed Latent Dirichlet Allocation (LDA) for topic modeling and BERT embedding for feature extraction, followed by deep-learning models (ANN, CNN, LSTM) to classify social media data for crisis response. The system achieved F1-scores between 77% and 83.1% across the different deep-learning models. The datasets included 70,000 disaster-related tweets across seven crises and 10 million COVID-19 pandemic tweets collected between 1 March and 30 April 2020.

Contreras *et al.* [136] performed sentiment analysis (SA) to assess post-disaster recovery on the 10th anniversary of L'Aquila's earthquake using Twitter/X. Polarity was first defined using a supervised classification based on experts' rules and Grammarly tones and then compared to the outcome of an unsupervised classification using the pre-trained SA machine learning algorithm developed by MonkeyLearn. The analysis, conducted on 4349 tweets collected from 4 April to 10 April 2019, achieved an overall accuracy of 57%, with a misclassification rate of 43%. The authors suggest that these results can serve as a benchmark for comparing other post-disaster recovery processes using the same Twitter/X-based sentiment analysis on future anniversaries.

Bashir *et al.* [137] conducted a sentiment analysis of tweets manually, categorizing them into eight broader sentiment categories. Orange Data Mining Software was used in an undisclosed version to visualize positive, negative, and neutral sentiments, whereas VOSviewer in an undisclosed version was employed to visualize word frequency in the tweets. This chapter analyzed 13,156 English tweets posted during the first 27 days of the attack. Performance metrics, such as F1-scores, were not measured because the analysis was performed manually.

Behl *et al.* [138] compared supervised learning approaches for multi-class classification of Twitter/X data, categorizing tweets into three classes: resource need, resource availability, or others. The system achieved a macro average F1-score of 0.61 and a weighted average F1-score of 0.85. The dataset comprised 2274 tweets, with 194 classified as resource needs, 125 as resource availability, and 1955 as others. This study noted a limitation in that the classification accuracy depended heavily on the type of resource being asked for, as demonstrated by the model's poor performance when trained on earthquake data and tested on COVID-19 data for the availability class.

Lian *et al.* [139] proposed a system using a density-based K-means algorithm for information extraction, ROST Emotion Analysis for sentiment recognition, MS-DICA for opinion classification, and a multi-agent system (MAS) to control false online information during natural disasters. The system achieved a reduction in false information by 2.05%

and 3.02%, respectively, using different strategies. This evaluation was conducted using 42,557 samples.

Karami *et al.* [140] developed a framework called Twitter Situational Awareness (TwiSA) that leverages text mining methods, including sentiment analysis and topic modeling, to enhance situational awareness for disaster preparedness, response, and recovery. While the study does not provide specific F1-scores or accuracy metrics, it demonstrates effectiveness in identifying relevant public concerns during disasters. The system was tested using 217,074 negative tweets.

Farnaghi *et al.* [141] introduced dynamic spatiotemporal tweet mining as a method for dynamic event extraction from geotagged tweets in large study areas. This approach employs a modified version of the ordering point clustering algorithm to address the heterogeneity of Twitter/X data and uses Word2Vec, GloVe, and FastText embedding to capture both syntactic and semantic similarities. This method was used to monitor the current state of disasters through spatiotemporal and textual analyses of geotagged tweets. The system was evaluated using tweets collected during Hurricane Florence from 12 to 19 September 2018, covering North and South Carolina. The performance, measured using the silhouette coefficient, yielded 0.561 for GloVe, 0.531 for FastText, and 0.516 for Word2Vec, indicating moderate clustering quality.

Gulnerman and Karaman [142] explored the best combination of filtering, mapping, and spatial accuracy methods for social media data, particularly in emergency situations. The goal was to develop a system capable of creating an incidence map shortly after a disaster. The system was evaluated on 4395 manually labeled tweets and achieved accuracies ranging from 36% to 49% for mapping disaster-related incidents.

Fan *et al.* [143] proposed a hybrid machine learning pipeline that uses Named Entity Recognition (NER) for detecting locations mentioned in posts, a location fusion approach to extract coordinates and remove noise, and a fine-tuned BERT model for classifying posts into humanitarian categories. In addition, graph-based clustering is used to identify credible situational information. The system, which aims to uncover disaster events across various locations from social media posts, was

tested on 95,606 relevant tweets from Hurricane Harvey and achieved an accuracy of 75.37%.

Singh *et al.* [144] utilized Natural Language Processing (NLP) techniques, including topic modeling with Latent Dirichlet Allocation (LDA) and Bayesian change point detection, to analyze emotions over time in disaster-related tweets. Although the study did not provide specific numerical metrics such as F1-scores or accuracy, the system was applied to 32,909 tweets.

4.4 Materials and Methods

This chapter describes the materials and methods used in this study. The first section provides a detailed overview of the ODET platform, the AI models used, and the hardware and computational resources required. The following section describes the processes and procedures necessary for analyzing and processing data. The aim is to present the approach transparently and understandably to provide a basis for the implementation and results presented later. Accordingly, data preparation and anonymization are first described, followed by text moderation and filtering procedures. In the next section, the techniques used for event detection are explained, followed by a description of the zero-shot classification process. Finally, the concepts and construction of knowledge graphs and the process of retrieval-augmented generation (RAG) are explained. This chapter concludes by explaining how compliance with the Berkeley protocol was achieved.

4.4.1 ODET

The Open Source Intelligence Disaster Event Tracker (ODET) is based on a modular architecture that allows for flexible adaptation to different scenarios. The system is divided into endpoints and agents. ODET supports specific endpoints, such as API endpoints for connecting external data sources, classification endpoints for filtering irrelevant or inappropriate content, text inference endpoints for semantic analysis, and embedding endpoints for creating textual representations. ODET also provides defined agent types: the API agent for data collection,

the classification agent for categorization and filtering, the embedding agent for semantically grouping content, the knowledge graph agent for extracting relationships, and the summary agent for generating detailed reports. These agent types can be flexibly combined and linked in the processing pipeline to satisfy specific requirements. The modular design makes it easy to replace or extend individual components, keeping ODET future-proof and compatible with new technologies or models.

In the case studies, unmodified AI models were used to show that ODET is powerful, even without additional training steps. The GIST-all-MiniLM-L6-v2 model was used for sentence embedding to generate high-dimensional text representations and capture semantic relationships between the analyzed data. The quantized model Mistral-Nemo-Instruct-2407-Q6_K_L was used for zero-shot classification and other agents, such as the knowledge graph agent and the summary agent. This model is not a pre-trained classifier but a basic Instruct model that is dynamically adapted by ODET using GGBNF grammar and specially defined prompts to generate zero- or multi-shot classifiers and other specialized outputs. The context window size was 65,536 tokens ($n\_ctx = 8192 \times 8$) with a batch size of 512 tokens (n_batch = 512), and 10 GPU layers (n_gpu_layers = 10) were used. For hate-speech detection, the KoalaAI TextModeration model was used, which analyzes content with a batch size of 64 tokens (batch_size = 64) while considering truncations (truncation = True).

Computing requirements vary depending on the ODET deployment scenario. In the case studies, the system successfully ran on a MacBook Air with an Apple Silicon M3 processor and 24 GB of shared RAM, although this configuration was not optimal. Using the Mistral-Nemo-Instruct-2407-Q6_K_L quantized model with a context window of up to 65,536 tokens and a batch size of 512 tokens typically requires at least 10 GB of RAM for model inference alone. Additional memory is required to process the input data and to perform parallel tasks, such as creating embeddings or performing classifications, bringing the total memory requirement to approximately 16–20 GB of RAM. On the CPU side, at least 8 cores are recommended to efficiently support multi-threaded processing. Running on an A100 GPU in the Google cloud proved to be

much more powerful, processing the entire pipeline at a much higher speed and with lower latency.

The ability to run ODET on both consumer hardware and cloud-based endpoints renders the system highly scalable and flexible. This scalability not only allows the system to adapt to varying resource availability but also provides the basis for rapid deployment in real-world disaster situations where time and availability are critical.

4.4.2 Data Preparation and Anonymization

The dataset used for the case study on Hurricane Harvey was obtained from Kaggle and includes tweets posted during the disaster between 20 and 30 August 2017. The dataset includes information such as tweet ID, number of likes, replies, retweets, timestamps, and tweet content, and was published by the original creator on Kaggle under the CC0 license [145].

For the 2023 Turkey earthquake case study, a dataset shared on Kaggle was used. According to the original author, the dataset was collected using the Tweepy software in an undisclosed version and the Twitter/X API on a daily basis during the disaster event. Tweepy is an open source Python package that facilitates access to the Twitter/X API. The dataset consists of public tweets containing information about the earthquake, including fields such as ID, text, hashtags, source, retweets, and retweet status [146].

To protect user privacy, all identifiable information, such as author names and mentions, was removed from the tweets. This anonymization ensures that the analysis is based solely on the content of the tweets and that no personal data are revealed. The datasets contained no specific user data or account information other than mentions or usernames. Accordingly, the anonymization procedure can be limited to the removal of mentions and author names. Subsequently, the data preparation process first converts the timestamps into a uniform date format to ensure that all tweets have a standardized time reference. Care was taken to remove invalid entries, such as tweets without a timestamp or tweets without text content. The next step is to clean the data. This involves using regular expressions to remove various

unwanted strings, including HTML entities, URLs, and the mentions of other users. The goal of this step is to reduce the text to core content and avoid possible disturbances during the analysis. Furthermore, the use of the upper and lower cases is standardized to ensure that the same words are not recognized as different terms if they are written differently.

4.4.3 Text Moderation and Filtering

In disaster management, it is crucial that the analyzed data are free of offensive or irrelevant content, as such content can negatively affect the accuracy and reliability of situation reports. For this purpose, the KoalaAI text moderation model was integrated into the preprocessing pipeline [147]. The KoalaAI classifier was chosen primarily for its speed and efficiency, as it allows us to quickly filter out unwanted content, such as hate speech, which could otherwise corrupt final reports. Although a more comprehensive text classification model can be used for filtering, the performance of the KoalaAI classifier is sufficient for the task at hand, particularly given the time-critical nature of disaster management.

Filtering out hate speech is not only a technical necessity but also an ethical imperative. The KoalaAI text classifier was configured to filter out tweets identified as hate speech. This filtering step is critical because hate speech, if included in the final reports, could undermine the credibility of situational awareness reports, potentially leading to distrust and misinformation. In the broader ODET platform, the filtering model can be swapped or even omitted, depending on the specific requirements of the case. In other studies, such as those on the conflicts in Ukraine, it was necessary to filter additional content types. This makes it possible to adapt it to complex scenarios. Even in the Hurricane Harvey case study, additional hate-speech filtering was necessary, as the moderation measures implemented by Twitter in 2017 were not sufficiently effective. In ODET, text moderation and filtering were performed in batches to process large numbers of tweets efficiently. The Huggingface transformer framework was used to create a pipeline for text classification [148]. Each tweet was submitted to the model for classification, and a label was assigned as a result, indicating

whether the tweet contained hate speech. Consequently, the hate-speech-labeled tweets were removed from the dataset. This ensured that the dataset was cleaned from irrelevant and disrespectful content for analysis.

4.4.4 Semantic Analysis and Clustering

The GIST-all-MiniLM-L6-v2 model [149] was integrated into the data pipeline by directly applying it to the pre-filtered and moderated data. After preprocessing, which involves filtering out hate speech, the tweets are converted into high-dimensional vector representations and stored in a vector database. These vectors store the semantic relationships between the tweets. In addition, a client for a persistent vector database was created to efficiently store the generated vectors and enable access for further steps.

ODET begins to process data in chunks to divide large amounts of data into manageable sections. These chunks represent time slots that can be preselected. This makes it possible to view data from historical records as live data by examining only ever looking at the currently relevant time slot. The preprocessed, filtered, and vectorized tweets were stored in a vector database. Each chunk is added to a separate collection, with the name of the collection being generated dynamically (e.g., “collection0”, ”collection1”). If a collection does not yet exist, it is created by the client.

This process involves adding vectors and IDs to the collections in the database to enable efficient and rapid searches for semantically similar content. This mechanism ensures that the data are structured and prepared for the subsequent semantic analysis. Finally, the file management process (handle_list_file) ensures that the information in the collections remains consistent and available for further analysis.

After transforming the tweets into embeddings, a community detection algorithm was used to identify semantically similar groups of tweets. This is performed by calculating the cosine similarity between the vectors, which measures the degree of content similarity between tweets. Tweets with high content similarity generate a similarity value close to one, whereas those with less similarity generate lower values. To ensure

that only relevant semantic relationships are captured, similarity values below a threshold of 0.1 are set to zero.

This process is performed in small steps to keep the calculations memory-efficient. The calculated similarity matrix was converted into a sparse matrix and stored in a compressed format. This data structure allows efficient processing of large amounts of data because only the relevant connections between tweets are stored.

A similarity matrix serves as the basis for the community detection algorithm, which groups semantically similar tweets and identifies and eliminates overlapping communities. Thus, significant topics in the data can be identified and used to generate situation reports. The minimum cluster size depends on the number of tweets analyzed. For smaller datasets, the minimum size can be reduced accordingly, whereas larger datasets require larger clusters.

4.4.5 Zero-Shot Classification and Information Extraction

For the case studies presented, zero-shot classification was realized using a 6-bit quantized version of the Mistral-Nemo-Instruct-2407 model that was specifically optimized for instructional tasks by its creators [150]. This model is a fine-tuned version of Mistral-Nemo-Base-2407, which was jointly developed by Mistral AI and NVIDIA. It is designed to ensure both high accuracy and efficiency and can be easily replaced by more advanced models in future studies. We used this model in a quantized but unmodified version. The quantized version of this model was specifically chosen for efficient use of memory and computing power. The model required approximately 10 GB of main memory with a near-perfect quality, making it particularly suitable for use in a wide range of hardware configurations.

A key aspect of ODET is its ability to freely exchange models and adapt them to specific requirements. The goal was to develop a system that does not rely on specially trained models, and can, therefore, be flexibly applied to different future disasters. This architecture allows for the exchange of AI models, embedding models, and classifiers to ensure

that researchers can benefit from new developments or integrate their own models without being tied to specific technologies.

Another central component is the ability to customize the prompts and grammar used to control the model output. The flexibility of the system allows researchers to define specific output requirements realized through JSON schema-based grammar. The schema used for the case studies consisted of disaster-event-related, location-related, and headline-related components. These components ensure that the model's output always includes information on whether a cluster is related to an event of interest, where the event is occurring, and which headline is the most appropriate. Using this structured grammar, we can ensure that the results are consistent and analyzable. The flexibility of this methodology allows us to define different questions and output formats for different use cases.

A specially developed prompt template that builds upon the grammar utilized earlier was used for the classification. As an example, the template for the Hurricane Harvey case study contains specific instructions that prompt the model to provide concise answers to questions such as, "Does the tweet refer to a hurricane?", "Where is the event located?" "What would be an appropriate headline?". This structured approach can be changed to specifics regarding the event of interest and ensures that the classified data are consistent and easy to analyze. The model was dynamically generated and adapted to a specific use case to provide accurate and meaningful results.

During the analysis, all clusters were classified thematically. Clusters that were classified as irrelevant were discarded to ensure the quality and meaningfulness of the final situation report. The cluster formation and processing method helps to increase the efficiency of data processing by processing large numbers of tweets in a shorter time and with fewer computational resources. The results of this process were stored in Python dictionaries and were subsequently used for further analysis. Data were filtered based on the output JSON values, excluding records classified as irrelevant.

Occasional misclassifications do not have a significant impact on the overall context of the data analysis. The clustering process in ODET

consists of two distinct phases. First, the vector embeddings of the tweets are generated using a sentence embedding model that transforms the data into a high-dimensional vector space in which semantic relationships are preserved. These embeddings are then sorted and clustered using a cosine similarity threshold such that only semantically relevant tweets are part of the same cluster.

Choosing an appropriate threshold for cosine similarity is a trivial procedure performed empirically by the user. This is performed by taking a set of samples of the computed similarities between vector representations generated with the desired sentence embedding model. Based on these similarity values, a threshold is chosen to ensure that only semantically related content is grouped into a cluster.

ODET supports all sentence embedding models supported by the SentenceTransformers library. This wide selection provides flexibility for different use cases and allows the user to choose models that are best suited to a particular problem.

Even a threshold that is not optimally chosen has little impact on the final result in the ODET pipeline. This is because the threshold is only used to cluster similar tweets.

After clustering, a zero-shot classification process is performed to assign thematic labels to each cluster. This comprehensive approach ensures that every sample within a cluster is processed, therefore minimizing the effects of misclassification. Given the large amount of data and robust clustering methodology, uncertainties, including hallucinations and misclassifications, must exceed a very high threshold to have a negative impact on the final results. In addition, a knowledge graph is generated in the further course of data processing, and retrieval-augmented generation is used, which provides an additional structure and context and further increases the robustness of the results.

4.4.6 Knowledge Graph Construction

To extract the relationships between entities, a similar methodology was used, in which the AI model was again supported by a specially defined

grammar. This grammar specifies that for each relationship between entities, three key elements are extracted:

- the ‘subject — which represents the main entity,
- the ‘relation — which describes the relationship between the entities,
- and the ‘object — which represents the target of the relationship.

Each element is defined in the schema as a string-based property. Grammar ensures that the model structures the relevant information and outputs it as triples, which is the basis for creating a knowledge graph.

These triples are the building blocks of a knowledge graph based on extracted entities and their relationships. In a knowledge graph, an entity such as “Hurricane Harvey” is linked as a subject to other entities such as “flooding” (object) via a relationship such as “caused”. This triple forms the basis of the semantic representation of the data. These links make it easy to represent the complex relationships within tweets. The knowledge graph connects these relationships to create a holistic understanding of events that go beyond mere data points.

Next, the previously classified clusters were divided into smaller groups to allow more efficient processing. The tweets contained in each group were combined into a single string by continuously adding the content of the tweets. The batch size can be defined freely. As soon as a group reached the defined batch size, the block was passed to the model for analysis. This structure facilitates the processing of large amounts of data into manageable batches.

The extracted data were returned as structured JSON responses, which were then organized in a Python dictionary. Each cluster was analyzed using the model, and the resulting triples were stored in the dictionary under the appropriate index. This process makes it possible to systematically document the relationships between the extracted entities and map them in a knowledge graph.

4.4.7 RAG and Summarization

The retrieval-augmented generation (RAG) process is critical for generating accurate and comprehensive summaries from large amounts of text [151, 152, 153]. This process combined the extraction of relevant information with the generation of new content to create meaningful and contextually relevant summaries. ChromaDB plays a central role in storing and querying the generated vector embeddings [154]. Previously generated vector embeddings were stored in the ChromaDB database, and relevant documents were efficiently retrieved through targeted queries. These queries allowed the collection of contextually relevant information for the subsequent summary. The headings from the zero-shot classification were used as inputs for the queries. At any given time, ChromaDB contains data only from the current hour. Agents are unable to look into the past or future.

The summarization process was performed using specialized agents that summarized the retrieved information. These agents used previously created knowledge graphs and contextual data to generate concise and consolidated summaries. The summarization agents interpret the knowledge graphs for each cluster (topic) to provide detailed situational analysis. These summaries have been consolidated into comprehensive reports.

4.4.8 Berkeley Protocol Conformance

This section explains how our technical process conforms to the standards and principles of the Berkeley protocol for open source digital investigations [118].

Anonymization and Privacy

We anonymized the tweets in our data-processing process by removing all user-related information. This includes the deletion of usernames, locations, and other identifying features. This ensured that the results could not be traced back to the authors of the tweets. Our process processes tweets mechanically, without storing or mining user information.

Methodological Principles of the Berkeley Protocol

The Berkeley protocol emphasizes the importance of accuracy, data minimization, and data security in open source digital investigations. Our approach meets these criteria in several ways.

- Accuracy:
  - Using advanced NLP techniques such as sentence embedding and community detection, we ensured that the data were accurately analyzed, and relevant events were correctly identified.
  - Using ChromaDB to store and query vector data ensures that only the most relevant information is used for analysis.
- Data minimization:
  - We processed only the necessary information by filtering out irrelevant and potentially harmful content, such as hate speech, using a classification model from KoalaAI [147].
  - Our approach limits the analysis to hourly intervals, with only the currently relevant hourly data stored in ChromaDB. This minimizes the amount of data and increases the focus on the current events.
- Security:
  - By isolating data on an hourly basis, we prevent our agents from looking into the past or future, ensuring the integrity of the data and the accuracy of the analysis.
  - Anonymizing the data and using only machine processing protects user privacy and ensures that no personal data are stored or analyzed.

4.5 Case Studies

This section explains the implementation of the Open Source Intelligence Disaster Event Tracker (ODET) for the case studies of Hurricane Harvey and the 2023 Turkey earthquake. This describes how ODET

was applied step-by-step to the respective datasets to extract relevant information and create automated reports.

Hurricane Harvey, which hit the coastal regions of Texas and Louisiana hard in August 2017, caused significant damage and economic losses. The storm caused widespread flooding and took a heavy toll on human lives, making it one of the most devastating natural disasters in US history [155]. By comparison, the earthquake in Turkey in February 2023 presents a very different scenario: measuring 7.8 on the Richter scale, it devastated large parts of the country, causing severe building damage and thousands of deaths [156]. The two events present different but equally challenging scenarios for applying ODET, demonstrating how flexibly the system can be adapted to analyze different disaster events and efficiently provide relevant information.

This section explains the basic structure of ODET in the first step, starting with setting up the endpoints for data collection, including configuring an API endpoint for the Hurricane Harvey dataset. In the second step, various agents are configured in detail. The third step describes the structure of the data-processing pipeline and explains how to execute it to enable continuous data analysis. Finally, benchmarking of the reports is described, including establishing a basis for comparison and selection of suitable metrics to evaluate the generated content. Figure 4.1 illustrates each step of ODET's structure.

4.5.1 Step 1: Endpoint Configuration

The endpoints form the basis of ODET's modular architecture and enable researchers to create their own modules and flexibly adapt ODET to new technologies or data sources. The ability to create custom endpoints makes ODET future-proof. For example, if a new powerful large language model (LLM) is released, a custom endpoint for text inference can be implemented. Similarly, the structure allows for easy integration of new datasets into the ODET. To integrate a new dataset, a REST API endpoint that provides specific data in the desired format can be made available. ODET supports four main types of endpoints:

1. API endpoints—Enables the connection of external data sources (e.g., web APIs such as Twitter/X or self-hosted datasets).

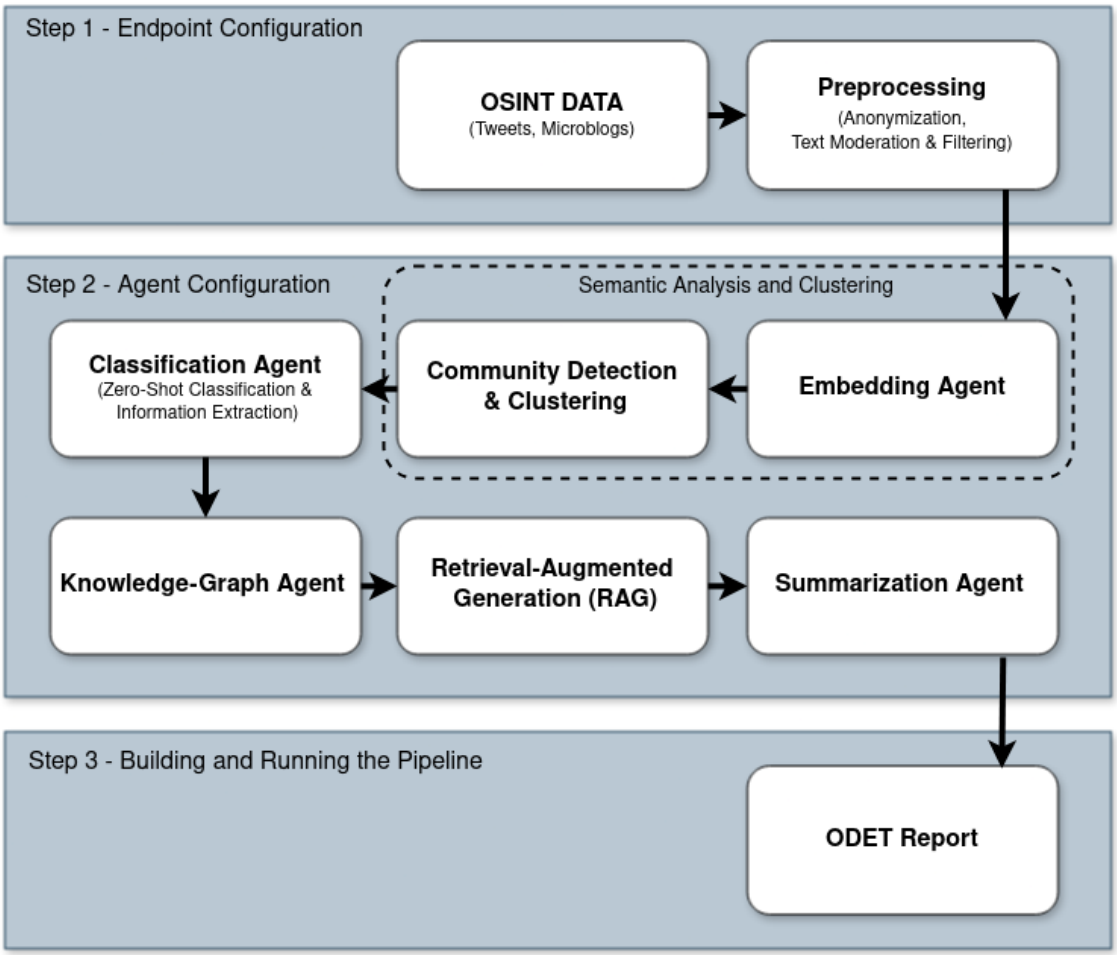

Figure 4.1: Structure of ODET for automated disaster event tracking and reporting.

2. Classify endpoints—Responsible for classification models, for example, filtering irrelevant or inappropriate content.

3. Text Inference Endpoints—Use inference models to derive specific information from text.

4. Embedding endpoints—Creation of text embedding to enable semantic similarities and clustering.

As shown in Figure A.1 we created an API endpoint for the Hurricane Harvey dataset that makes tweets accessible during hurricanes. These endpoints allow flexible querying of data at defined time intervals and are further processed by downstream agents. Aligned with this example, the API endpoint was implemented for the 2023 Turkey Earthquake study.

4.5.2 Step 2: Agent Configuration

In ODET, agents are used as central building blocks to perform various data-processing tasks. Each agent performs a specific task, such as classifying data, creating text embedding, or creating knowledge graphs. These agents operate sequentially, with each agent processing the results of the previous agent. ODET's modular design allows agents to be flexibly adapted, replaced, or extended to integrate future requirements and developments in AI technology.

API Agent Configuration

The API agent is the starting point of the data pipeline in ODET and is used to connect and retrieve data. As shown in Figure A.2, several parameters that influence the query and timing of datasets can be specified. While parameters can be set individually, for the case studies presented, the parameters start, offset, and interval were used. The ability to define parameters freely allows researchers to work with a wide range of time-series data and datasets from microblogging platforms.

- Start Time—This determines when the analysis of the dataset begins. For the Hurricane Harvey case study, 3:00 p.m. on 25 August 2017 was chosen to synchronize the start of reporting with the initial impact of the storm.
- Interval—This interval divides the data into regular hourly slices (60 min), resulting in hourly reports.
- Offset—This parameter determines the time of day the analysis begins. In this case, the offset was set to 3:00 p.m. The first data considered by the agent was the data created after 3:00 p.m.
- Input and Output—The input and output options of the API agent indicate that it is at the beginning of the pipeline and has no previous input data.

Hate-Speech Filter

The configuration of the classification agent is the next step in the pipeline, as shown in Figure A.3. This filter is a key part of data

processing to remove inappropriate content. Enabling the hate-speech filter ensures that any inappropriate content that is not relevant to the analysis of the disaster event is removed. This helps to ensure the accuracy and reliability of the reports. This step is not mandatory, and researchers may choose not to include it. However, for the Hurricane Harvey case study, as well as the 2023 Turkey earthquake case study, a hate-speech filter was enabled.

- Filter label—The filter is set to remove all tweets classified as "hate speech" by labeling them with an "H". ODET supports several different filters and labels through settings such as these.
- Input options—The hate-speech filter uses the API agent's output. ODET allows the data to be seamlessly transferred from the API agent using the "Map from previous" configuration option.
- Output options—The output data are passed to the next agent in the pipeline after filtering. It is decidable to either silently pass on the data or output debug information.

Embedding Agent

The embedding agent is a central element of the pipeline and is used to semantically group tweets and organize them for further processing. As shown in Figure A.8, the agent uses vector embeddings to group tweets based on their similarity. The input data for this agent comes from a previously set up hate-speech filtering agent.

- Similarity Threshold—The cosine similarity threshold between the tweets was set to 0.84. This means that only tweets with a high degree of content similarity were grouped into clusters.
- Minimum Cluster Size—To ensure that only relevant groups were formed, a minimum of 100 tweets per cluster was set. The cluster size allows for the identification of trends. A relevant event is expected to generate numerous similar messages.
- Grouping of Cluster Elements—Large clusters can contain thousands of similar tweets. To further compress them, tweets from a cluster can be grouped. This has the advantage of minimizing the number of queries to computationally intensive AI models and

increasing efficiency. This is particularly useful for processing large amounts of similar content in a cluster as a single unit because fewer queries need to be made.

Classification Agent

The classification agent is responsible for filtering irrelevant clusters and assigning an appropriate label to each relevant cluster. As shown in Figure A.5, this agent allows the configuration of a zero-shot classifier based on an LLM. The classification agent is critical for further processing of the clusters. It uses a pre-trained language model driven by JSON-based grammar. This grammar defines the output parameters and allows the model to determine the relevance of the clusters based on prompt instructions.

- Grammar and JSON schema—Grammars define the outputs of the language model. They rely on a prompt template adopting its fields. Grammars were created as JSON schema and translated by ODET into a format that can be processed by an AI model. Both case studies used a schema that included a Boolean value for detecting disaster relevance. An example is shown in Figure A.4.
- Prompt template and objective—This template provides specific instructions for creating the headline, rating the relevance, and identifying the location. The goal of the agent is to determine the relevance of the clusters by answering the question of whether the content is disaster-related. It is possible to answer multiple questions simultaneously to save steps. In the context of the case studies presented in this work, the agent was also asked for a possible location for the described context as well as for a headline summarizing the events. The results are output in a predefined format and structured based on the defined JSON grammar. An example is shown in Figure A.9.
- Mapper and object filter—The classification agent has a mapper that maps objects and a filter that removes irrelevant clusters based on the classification results. If a cluster was marked as irrelevant (“false”), it was removed from the pipeline. An example

is shown in Figure A.7. The mapper and object filters can be used to create complex filter rules that remove unwanted objects or process them in a targeted manner. On the other hand, objects can be grouped or processed in a targeted manner. This was made possible by the generated JSON output.

- Input and Output—The agent receives its input data from the embedding agent and outputs filtered and classified clusters with associated headings. The data are then passed to the next agent.

Knowledge Graph Agent

The knowledge graph agent is a central component of the ODET pipeline that not only identifies relationships between entities but also refines them through retrieval-augmented generation (RAG). RAG makes it possible to extract relevant information directly from the original data, which improves the accuracy of the reports because the generated summaries are supported by real data. Furthermore, the knowledge graph agent makes it possible to capture the dynamics of events by representing not only static but also changing relationships. This feature helps to track the progression of a disaster event in real time and to analyze the impact of decisions or external factors on the event. Combining the extracted knowledge graphs with RAG allows ODET to create more accurate and up-to-date reports that are invaluable to decision-makers in crisis situations. The input for this agent comes from the results of the classification agent, and the output is processed in the last step of the pipeline. This agent uses the same text inference endpoint as the classification endpoint and differs only in the grammar, prompt, and RAG settings. An example of the agent configuration is shown in Figure A.6.

- Grammar and JSON schema—The agent analyzes the semantic relationships between entities based on the context of tweets. ODET allows the definition of a JSON grammar that specifies the structure of the output data as triples (subject, relation, and object). These triples are supported by a prompt similar to that used by the classification agent. An example is shown in Figure A.10.
- Retrieval-augmented generation (RAG)—The knowledge graph

agent combines the generated graphs with RAG using the generated headlines to access the original data set and collect further information. This ensures higher accuracy and increases the validity of the reports.

- Grouping of entities—Since it can be assumed that messages with similar content and a similar creation time refer to a single event, the agent groups such messages. This reduces the number of requests for the model and optimizes processing speed. Grouping similar data points increased the efficiency of knowledge graph generation.
- Visualization of results—The knowledge graph agent creates the results as a structured network of entities in the form of triples, which supports the creation of reports and simplifies data analysis. The generated knowledge graphs can be displayed graphically in the report if desired, making it easier to identify correlations and validate relationships.

Summarization Agent

The summarization agent is the final step in the pipeline for Hurricane Harvey and the 2023 Turkey earthquake case studies, where the summary reports are generated based on the processed data. As shown in Figure A.11, the agent can efficiently summarize relevant information using previously generated knowledge graphs and retrieval-augmented generation (RAG) data in the system. Although the summarization agent represents the final step in the case studies presented in this work, it can be complemented by additional post-processing agents in other applications. The modularity and flexibility of ODET allow researchers to adapt the structure of their pipelines to the specific needs of their projects.

- Objective and reasoning—Researchers can provide detailed instructions to the agent and specify what information is included in the report and how it should be structured. As shown in Figure A.12, the agent was instructed to summarize the content from the knowledge graph and create an appropriate headline for each report.

- Markdown output—The agent outputs the reports in the Markdown format, which allows for flexible integration with external systems and reporting formats.

4.5.3 Step 3: Building and Running the Pipeline

The third and final step in ODET involves building and running the report generator, which allows researchers to configure agents in a specific order and generate consecutive reports. These reports are defined in what is called the "report area," where researchers can name the report, add a detailed description, and specify the exact order of agents to be processed. The description clarifies the focus of the report to ensure that it meets the objectives of the research project.

A key aspect of report creation in ODET is the ability to select and sequence agents flexibly. This is presented in a user-friendly interface, where researchers can drag agents from the list of available agents into the selected agent box. Multiple agents can be used in a pipeline or multiple times in a report to account for different aspects of data evaluation. This provides a high degree of flexibility as researchers can use the order and selection of agents to perform various analyses tailored to the specific needs of the investigation.

Another important feature of ODET is its ability to customize the frequency of report execution. The system allows researchers to repeat reports at regular intervals, which is particularly useful when processing time-series data. The execution frequency can be set to a minute, hour, day, or month. This makes ODET useful for both analyses that run through historical data and real-time analyses, in which it is important to constantly integrate and analyze new data. This ensures that researchers have up-to-date information and can make quick decisions based on the latest data.

In addition to flexibility in execution frequency, the ODET also provides complete interactive control over the agents during reporting. Any agent that is part of the pipeline can be paused, restarted, or re-executed to ensure an optimal operation. This allows researchers to adjust and try agents until they produce the desired results. This interactive capability makes the ODET a valuable tool for iterative report

development and improvement. Another useful feature is the ability to regenerate the entire report, which is especially helpful when changes have been made or when the report needs to be re-run after agent customization.

The generated reports provide detailed outputs, organized in a way similar to Jupyter notebooks. This means that each step of the pipeline is represented in a report, and researchers can review the output of each agent before proceeding to the next step. These reports can include both text-based and graphical outputs, and researchers have the option to view results in a Markdown format or other suitable formats to prepare and visually display the data for analysis.

Overall, this step in ODET enables the creation of flexible and scalable reports designed for both one-time analysis and continuous real-time monitoring of disaster events. The ability to interactively customize agents and continuously refresh reports ensures that ODET remains a powerful tool that meets the changing needs of disaster researchers. An example of the reported setup is presented in Figure A.13.

4.6 Results and Discussion

This section presents and discusses the results of the case studies performed to evaluate ODET's performance in terms of consistency with established reference sources. In addition, we discuss implications for disaster management. The central evaluation approach is based on measuring the content consistency between reports generated by ODET and English-language Wikipedia articles on Hurricane Harvey and the 2023 Turkey earthquake using the AlignScore metric. Finally, the strengths and weaknesses of the system were analyzed to identify opportunities for improvement and future applications.

4.6.1 AlignScore

To evaluate the quality of the reports generated by ODET in our case studies, the metric AlignScore was used to show the validity of the results. This metric makes it possible to compare machine generated reports with an established reference source and to measure the

alignment of the content elements. AlignScore provides a systematic and objective metric for evaluating the accuracy and completeness of automatically generated content compared to a reliable source [117] by measuring its structural and semantic alignment. The advantage of this metric is that it is possible to measure both content scope and semantic accuracy, which provides a more comprehensive insight into the system's performance. AlignScore considers not only the surface area but also the depth and context of the information. AlignScore is mathematically described as:

$$AlignScore(o, l) = \text{mean}_{j} \max_{i} \text{alignment}(o'_i, l'_j), \tag{4.1}$$

where o refers to the context , l represents the claim, o'_i is the collection of context chunks, l'_j is the collection of claim sentences, and alignment$(\cdot)$ denotes the likelihood that the model assigns to the Aligned label in a 3-class classification setup [117].

The goal of ODET is to generate situational awareness reports from unstructured disaster data. This task requires a robust evaluation platform that can assess how well generated outputs reflect real-world events. Compared to traditional metrics such as F1-score, AlignScore is well suited for ODET evaluation because it focuses on factual consistency rather than discrete classification accuracy. AlignScore addresses this need by holistically evaluating the semantic fidelity of the generated text to the reference material to identify nuanced discrepancies such as omissions, factual distortions, or misrepresentations [117].

Throughout the case studies of the 2023 Turkey Earthquake and Hurricane Harvey, Wikipedia was a valuable source for evaluating reports. A direct timeline of events was not available; therefore, Wikipedia pages had to be relied upon as a comprehensive source of information. These pages contain well-researched and detailed neutral information that has been edited and added by numerous authors over the years. Although Wikipedia contains retrospective insights and information that only became available after the event, it provides a reliable basis for comparison.

To evaluate the case studies, each of the 58 ODET-generated reports for the Hurricane Harvey case study and 131 reports for the 2023 Turkey

Earthquake case study were analyzed in a final summary report and then compared to the corresponding Wikipedia article. The high level of agreement indicates that the ODET-generated reports were largely accurate despite the limitations of live data and the limited database.

A final result of 0.89 (Hurricane Harvey) and 0.84 (2023 Turkey earthquake) was achieved, matching the generated ODET reports and the corresponding Wikipedia articles. These values confirm a remarkable agreement between the two sources, especially considering the different natures of the data sources. While Wikipedia is based on retrospective research and extensive source studies, ODET reports reflect information gathered in real time during the event.

4.6.2 Summary of Key Insights and Implications

In this chapter, we present ODET, a modular platform that allows both research and disaster management communities to experiment with different algorithms and develop innovative solutions. Our case studies show that ODET is not only a flexible platform but also capable of generating highly accurate and structured reports that enable rapid situational assessment.

ODET serves as a platform for researchers to deploy and customize AI-based algorithms to generate reports tailored to specific scenarios. We demonstrated that an appropriate algorithm based on text embedding, community clustering, and zero-shot classification can be efficiently integrated into the ODET. Our case studies on Hurricane Harvey and the 2023 Turkey earthquake show that ODET also works reliably under the challenges of real-world data sources, such as tweets. With AlignScore values of 0.89 and 0.84, respectively, the system demonstrated its ability to extract relevant and consistent information in a dynamic environment.

ODET's modular design allows for easy replacement of individual components or models, making the platform not only future-proof but also highly adaptable. This allows researchers to integrate new technologies or models and to test their own approaches without being tied to specific technologies. This flexibility means that ODET can run on standard

hardware, as well as in high-performance cloud environments, making it suitable for both small and large datasets.

The results presented here show that ODET is a promising platform for quickly generating situation reports that provide decision-makers with the information they need. At the same time, ODET underscores its importance as a research platform that enables the development and testing of innovative AI-based algorithms in a controlled environment.

Future work should focus on further improving ODET, particularly on how to deal with misinformation and integrate new approaches to increase the accuracy of reports. Another well-known unsolved problem is hallucinations, which we are currently trying to minimize in the reports generated using a multi-step clustering and labeling approach as well as RAG and knowledge graphs, as these technologies provide a stronger link between content and underlying data. However, hallucinations cannot be eliminated at this time because they are a known unsolved problem in AI-based text generation. Further research is needed to develop more effective mechanisms for detecting and avoiding hallucinations. This is crucial to meet the growing demands of disaster management systems and to realize the full potential of ODET as a flexible and scalable platform.

Chapter 5

Conclusions and Future Works

This dissertation describes the development of a modular machine learning-based OSINT platform for automated situational awareness reporting in disaster management. It shows how a scalable platform for analyzing microblogging data for automated situational awareness reporting in disaster management can be realized by combining event detection, clustering, zero-shot classification, retrieval-augmented generation (RAG), and the use of knowledge graphs.

The publication-by-publication development of this platform has been documented in three key studies.

Study 1: Towards Automated Situational Awareness Reporting for Disaster Management—A Case Study [16]

The first study laid the methodological foundation for automated event detection in disaster situations. A system that uses SimHash for clustering similar messages was developed to efficiently process large amounts of text. In combination with TF-IDF-based feature extraction, relevant content was identified from the heterogeneous microblogging data. Subsequent analyses using extractive text summarization methods (including SpaCy, Gensim, and NLTK) were used to condense event-related information. The evaluation based on the Russo-Ukrainian conflict focused on identifying event occurrences and demonstrated the feasibility of the approach for event detection, while also highlighting the challenges of filtering irrelevant content.

Study 2: Harnessing OSINT in Disaster Management: Transforming Microblogging Posts Into Insightful Data Through Text Embedding [17]

The second study extended the methodological approach by combining

sentence embedding with community detection to enable more precise semantic grouping of events. In addition, knowledge graphs were integrated to structure the extracted information and enhance interpretability. This combination of methods improved the quality of event detection and enabled a more structured representation of disaster-related insights.

Study 3: AI-Enhanced Disaster Management: A Modular OSINT System for Rapid Automated Reporting [18]

In the third study, the developed methods were combined into a fully modular OSINT system. In addition to the combination of zero-shot classification, community detection-based clustering, and knowledge graph-based analysis, retrieval-augmented generation (RAG) for structured reporting was integrated to improve the quality of reporting. Furthermore, an agent-based architecture allows for flexible system expansion and modular integration of individual components. The evaluation based on real case studies (Hurricane Harvey, Turkey Earthquake 2023) showed that the system not only achieves a high alignment with Wikipedia event summaries (AlignScore up to 89%), but also demonstrates deployment feasibility across different hardware environments.

This dissertation demonstrates that the application of OSINT and machine learning in disaster management not only enables efficient automation of data analysis but also contributes to accelerated and high-quality situational awareness. Simultaneously, advances in automated event detection using sentence embedding, community detection, and zero-shot classification have been presented, as well as advances in retrieval-augmented generation (RAG) and the integration of knowledge graphs to improve the quality and relevance of situational awareness reporting in disaster management.

In the following, a synthesis of the key scientific findings is presented, which combines the methodological contributions of the three studies presented throughout this dissertation and places them in the overall context. Subsequently, the limitations of the developed approaches are discussed, addressing both the methodological and technical challenges. Finally, promising research directions for further development of OSINT-based disaster analysis are outlined.

5.1 Research Questions Revisited

In this section, the central research questions of this dissertation are answered by bringing together the key findings of the entire investigation.

Q1: **Can OSINT and ML be used for automated situational awareness reporting in disaster management?**
Based on the research conducted, it can be concluded that OSINT and ML-based can be used for automated situational awareness reporting. Relevant events can be extracted from unstructured microblogging data, semantically grouped, and processed in a structured manner. This enables continuous recording of disaster events and their development based on publicly available information. A key finding is that the use of knowledge graphs and retrieval-augmented generation (RAG) to contextualize the extracted data significantly contributes to improving quality. Nevertheless, the challenge remains in reliably filtering out irrelevant or misleading information.

Q2: **How can the accuracy of OSINT and ML-based automated situational awareness reporting for disaster management be improved?**
A significant improvement in accuracy was achieved by combining the different techniques. The transition from classical statistical methods, such as TF-IDF, to modern embedding methods has significantly refined the semantic analysis of data. The use of community detection algorithms has improved the grouping of thematically related events, minimized redundancies, and increased the consistency of content. Additionally, zero-shot classification has proven particularly valuable because it allows for flexible, data-independent classification of events without the need for extensive manual training datasets. This helps to capture new and unforeseen event types more quickly.

Q3: **How can the scalability of OSINT and ML-based solutions for situational awareness reporting in disaster management be improved?**

The results obtained from this research show that scalability can be achieved through an agent-based architecture that allows different AI models to be integrated and exchanged in a modular and flexible manner. This allows dynamic adaptation to various requirements, such as the processing of large amounts of data or the use of different hardware environments. It has been successfully demonstrated that the system can be used on both high-performance cloud servers and consumer hardware (e.g., MacBook Air M3), allowing a high degree of deployment flexibility. This shows that the developed methods work reliably not only in research environments, but also in real-world application scenarios.

This dissertation has shown that Open Source Intelligence (OSINT) in conjunction with modern machine learning methods provides a powerful basis for automated situational awareness reporting in disaster management. By combining several innovative methods, from community detection to zero-shot classification and retrieval-augmented generation, a system that is both accurate and flexible has been developed. It is particularly clear that the quality of situational awareness reporting depends not only on technical implementation, but also on the ability to efficiently filter and meaningfully structure large amounts of unstructured data. While the developed methods already show a high level of performance, challenges remain, particularly with regard to the detection of misinformation and further automation of the prioritization of information. The system developed in this process created a scalable, modularly expandable OSINT framework that forms the basis for future research and practical applications. The results underscore the need to further develop OSINT-based situational awareness reporting, particularly through improved mechanisms for quality assurance of the analyzed data and efficient integration of new data sources. To classify the insights gained holistically, the following section brings together the central methodological and conceptual aspects. This synthesis considers the progress made in the context of existing research approaches and highlights the overarching implications of this dissertation.

5.2 Synthesis

This dissertation demonstrates how Open Source Intelligence (OSINT) can be used with machine learning methods to automatically generate situational awareness reports in disaster management. While OSINT has mainly been used in intelligence contexts, this dissertation shows that these principles can also be effectively applied in humanitarian crisis management. The three scientific publications of this dissertation deal with the step-by-step development, evaluation, and optimization of a modular ML-based OSINT system. It was shown that the combination of event detection, zero-shot classification, retrieval-augmented generation (RAG), and modular system architecture enables a flexible, scalable, and highly automated solution for disaster analysis. The synthesis of these results demonstrates not only the scientific advances in OSINT-based analysis but also how machine learning can help overcome the challenges of traditional, manually performed OSINT processes.

The first study laid the methodological foundation for automated situation reporting in disaster management by using microblogging data for event detection and synthesis. This study presents a system that uses SimHash for clustering similar messages to efficiently identify disaster-related events in social media data. In addition, Term Frequency-Inverse Document Frequency (TF-IDF) is used for feature extraction to highlight relevant content from a heterogeneous and noisy dataset. Events are grouped by clustering messages with similar SimHash representations and aggregating thematically related posts into coherent event clusters. These clusters were then analyzed using extractive text summarization methods, specifically SpaCy, Gensim, and NLTK, to evaluate their suitability for condensing event-related information. The evaluation was carried out using data from the Russo-Ukrainian conflict, where the detected events were compared with a Wikipedia timeline to measure the effectiveness of automated event detection. The results show that the proposed methodology detected a higher number of events than those listed in Wikipedia, with SpaCy providing the best summarization results. At the same time, the study highlights the challenges of filtering out irrelevant or misleading content and underscores the need for further refinement of disaster-related OSINT processing.

Building on the findings of the first study, the second study extended the methodological approach by combining sentence embedding with community detection for more precise semantic clustering. While the first study used SimHash for clustering based on term similarity, the method presented here allows for a more refined semantic grouping of events. In addition, knowledge graphs were integrated into the system to structure extracted information and enhance interpretability. This approach differs from classical machine learning methods that rely on large labeled datasets. The study shows that the combination of semantic clustering and structured event representation can improve recognition accuracy. It also introduces the next step toward automated information prioritization by extracting relevant content from large volumes of unstructured data and structuring it for situational awareness reporting. The results demonstrate that OSINT-based disaster analysis can be automated while improving event organization and interpretability at scale.

The third study integrated the developed methods into a fully modular and scalable OSINT system. Building on the methods introduced in previous work, it extends the system to include retrieval-augmented generation (RAG) for structured reporting and knowledge graphs for contextual analysis of the extracted information. While previous work has focused on the detection and semantic clustering of events, the focus here is on the end-to-end situational awareness reporting platform. The combination of zero-shot classification, community detection-based clustering, and knowledge graph-based analysis enables the detailed and contextual processing of relevant information. In addition, an agent-based architecture is implemented, allowing for flexible adaptation and modular extensibility of the system. The evaluation of the system using real case studies, including Hurricane Harvey and the 2023 Turkey earthquake, shows that the developed OSINT framework not only achieves high factual alignment in event detection (AlignScore up to 89%) but also enables consistent and traceable situational awareness reporting. Thus, this study represents a decisive step in the development of a fully automated ML-based OSINT system for disaster management.

The presented research shows that the integration of Open Source

Intelligence (OSINT) into disaster management enables a significant improvement in situational awareness. Whereas traditional methods rely on manual data evaluation, modern approaches use machine learning to automatically extract and process information from microblogging data. This reduces the time required to extract relevant content and enables faster decision-making. The use of text embedding, community detection, clustering, and zero-shot classification forms the technological basis for an efficient and scalable analysis. In particular, the automation of data processing is a major step forward, not only increasing speed, but also ensuring consistent and traceable analysis of large volumes of data.

The Open Source Intelligence Disaster Event Tracker (ODET) system is characterized by its modularity and scalability, allowing it to be flexibly adapted to different scenarios. The separation of the processing steps into dedicated agents and endpoints makes it easy to exchange models and technologies. This ensures continuous development and adaptability to the new requirements. The scalability of the system allows it to be used in low-performance devices and high-performance cloud instances. An evaluation based on real case studies shows that the ODET remains effective even with untrained models and achieves a high degree of consistency with reference reports. This architecture ensures that ODET remains applicable over the long term and can be used in various application scenarios.

A key aspect of this research is the objective evaluation of the generated reports to ensure the quality and reliability of automated situation assessment. To this end, the AlignScore metric was used to quantify the consistency of the content with established reference sources, such as Wikipedia. Case studies on the coverage of Hurricane Harvey and the 2023 earthquake in Turkey show that the system achieves a high level of content consistency and provides a sound basis for decision-making. This underscores the importance of machine-based situational awareness in disaster management and demonstrates that OSINT-based methods can provide a reliable alternative to traditional approaches. The combination of modular architecture, scalable methods, and scientifically based quality assessment makes ODET a powerful tool for automated intelligence gathering in crisis situations.

Despite these successes, there remain open research questions and challenges. A central problem remains in the detection and handling of misinformation and in targeted disinformation campaigns. While the developed methods can group and classify content according to semantic criteria, the question remains as to how to reliably detect false or misleading information automatically. The current generation of Generative large language models also carry the risk of hallucination, that is, the generation of plausible-sounding but factually incorrect content. Further development of methods for verifying and ensuring the quality of generated reports is, therefore, an important future research focus.

In conclusion, this dissertation shows that ML-based OSINT methods are viable and efficient solutions for automated situation assessment in disaster management. By combining different ML techniques with modular system architectures, a solution that is both flexible and powerful can be developed. The synthesis of the results shows that OSINT analysis can be accelerated and qualitatively improved by machine learning. The methods presented in this dissertation have the potential to fundamentally change the way crisis information is collected, analyzed, and reported.

5.3 Limitations

Despite ODET's high performance in automated analysis of microblogging data, there are some methodological and technical limitations that require further research and development.

A key challenge is that the detection of fake news and hallucinations remains an open challenge. ODET does not use explicit mechanisms to verify the extracted information, and does not check whether certain statements come from reliable or manipulated sources. Although the aggregated analysis of clusters partially compensates for bias, there is still the risk that false or misleading information will be included in automated reporting. Language models that can generate syntactically correct but factually incorrect content are particularly problematic in this context. Although the system attempts to minimize such effects through structured prompts and specialized processing steps, this is

a general problem with the use of generative language models that cannot be completely eliminated.

Another important issue is the reliance on text-based data. ODET only processes text-based information and does not consider additional data formats, such as images, videos, or audio recordings. In crisis situations, visual or audio content can be a valuable addition to text messages, especially when there are language barriers or misunderstandings. Therefore, the addition of multimodal analysis could improve the reliability of situation assessment.

Furthermore, the granularity of the event representation can be improved. While the clustering process allows for reliable detection and structuring of larger events, very small or locally limited incidents may be lost in large clusters or may not be captured as independent events. This is especially true for subtle developments that may only be recognized as coherent over longer periods of time. The current methodology ensures that global or widespread events are detected efficiently, but fine-tuning for microevent detection remains an open challenge.

5.4 Future Works

While ODET already provides a powerful and flexible platform for the automated analysis of microblogging data in disaster management, there are several areas where future research could further develop and optimize the system. Owing to its modular architecture, ODET can be extended in a targeted manner without compromising core functionality or scalability. In particular, improving event granularity, integrating multimodal data, and optimizing validation mechanisms for fake news and hallucinations are important starting points for future development.

One of the most important open research questions is the validation of collected information. The detection of fake news and reduction of hallucinations in the generated reports are essential for the reliability of OSINT analysis. Although the current system already provides some robustness against individual anomalies through clustering and majority voting, an additional layer of validation could further improve the value of the generated reports. A possible solution might be an in-depth

network review of social media accounts to circumvent the distribution of fake news and the implementation of more sophisticated validation layers. Regarding the remaining problems of hallucinations, however, it might take a whole new generation of AI-Models to eliminate this risk.

Although this methodology includes text filtering and classification, the detection of misinformation remains an under-researched challenge. The proliferation of fake news, manipulated content, and adversarial misinformation campaigns requires a more robust strategy. Future research should address adversarial testing frameworks, bias mitigation strategies, and trustworthiness assessment mechanisms to improve the data reliability in crisis scenarios.

Another important area for future research is the extension of ODET to multimodal data processing. Thus far, the system has focused exclusively on text-based information, but in crisis situations, images, videos, and audio data also play a crucial role. Because ODET is modular, a separate data handler can be developed to integrate multimodal models. For example, existing multimodal language models can be used to extend the zero-shot classification to visual content. This would make it possible to automatically analyze not only text messages but also images from disaster areas or relevant videos from social networks. Future research should explore how multimodal fusion techniques can improve event detection and situational awareness by combining different types of data for more holistic analysis.

In addition to the extension to multimodal data, another possible challenge is the granularity of event representation. While the current clustering method provides reliable results for larger events, refinement of the method can help capture smaller or more localized events more accurately. Adaptive clustering, which dynamically distinguishes large and small events, is a promising approach. The use of hybrid clustering methods that consider semantic similarities and temporal sequences can enable more precise distinctions.

Technological developments in the area of language models and AI-based clustering methods have opened up further possibilities for optimizing ODET. Future multimodal models could allow for more accurate detection and processing of different types of data, whereas more

efficient clustering algorithms would allow for even more accurate structuring of events.

ODET is designed to run in the cloud or on consumer hardware available in disaster scenarios, making it flexible for deployment based on the available infrastructure. It is not designed for mobile or edge computing environments, but rather works on existing computing resources at disaster sites. Unlike other OSINT-based disaster intelligence tools, ODET implements a modular endpoint architecture that allows efficient load balancing. This means that load-intensive modules, such as large language models (LLMs), can run in the cloud, while lightweight agents run on consumer hardware such as a MacBook. This distributed architecture ensures scalability, adaptability, and efficient resource utilization in different environments. Future work should focus on optimizing deployment workflows, improving orchestration strategies, and evaluating performance in live disaster simulations to validate the effectiveness of real-world crisis responses further.

ODET already has a scalable, modular architecture in which computation-intensive tasks can be offloaded to cloud-based endpoints, while lightweight agents run on local, consumer-grade hardware. Rather than requiring additional optimization for scalability, future research could focus on developing domain-specific endpoints that improve ODET's adaptability to different disaster scenarios. The flexibility of endpoint integration allows researchers to tailor the system to their specific needs, without modifying the core platform. Therefore, future studies should investigate the best practices for endpoint implementation, security considerations, and strategies for seamless integration with external data sources.

It is important to emphasize that ODET is not an incremental improvement in an existing system but a novel end-to-end platform for OSINT-based disaster situational awareness. While previous work has addressed individual components such as event detection, summarization, and knowledge graphs, no other system integrates these elements into a fully automated, modular, and scalable solution such as ODET. Current disaster intelligence tools either require manual intervention, lack real-time processing, or lack the flexible and adaptable architecture

of the ODET. Future research projects may explore further integrations, but the fundamental innovation of ODET as an extensible platform continues to set it apart from the existing approaches.

In summary, several promising research directions can be identified that can further exploit the potential of ODET. Refinement of the clustering methods could lead to more precise event identification, whereas integration of a multimodal data wrapper would significantly expand the analysis capabilities. Simultaneously, an improved validation layer can improve the detection of fake news and reduce hallucinations in the generated reports. Future research should focus on efficiently integrating these extensions into existing systems without compromising ODET's scalability or modular flexibility.

Además de la extensión a datos multimodales, otro posible desafío es la granularidad de la representación de eventos. Si bien el método de clustering actual proporciona resultados fiables para eventos más grandes, el refinamiento del método puede ayudar a capturar eventos más pequeños o más localizados con mayor precisión. El clustering adaptativo, que distingue dinámicamente entre eventos grandes y pequeños, es un enfoque prometedor. El uso de métodos de clustering híbridos que consideren similitudes semánticas y secuencias temporales puede permitir distinciones más precisas.

Los desarrollos tecnológicos en el área de modelos de lenguaje y métodos de clustering basados en IA han abierto más posibilidades para optimizar ODET. Los futuros modelos multimodales podrían permitir una detección y procesamiento más preciso de diferentes tipos de datos, mientras que algoritmos de clustering más eficientes permitirían una estructuración aún más precisa de los eventos.

ODET está diseñado para ejecutarse en la nube o en hardware de consumo disponible en escenarios de desastre, haciéndolo flexible para su implementación según la infraestructura disponible. No está diseñado para entornos de computación móvil o edge, sino que funciona en recursos computacionales existentes en los sitios de desastre. A diferencia de otras herramientas de inteligencia de desastres basadas en OSINT, ODET implementa una arquitectura modular de endpoints que permite un equilibrio de carga eficiente. Esto significa que los módulos que

consumen muchos recursos, como los large language models (LLMs), pueden ejecutarse en la nube, mientras que los agentes ligeros se ejecutan en hardware de consumo como un MacBook. Esta arquitectura distribuida garantiza escalabilidad, adaptabilidad y utilización eficiente de recursos en diferentes entornos. El trabajo futuro debería centrarse en optimizar los flujos de trabajo de implementación, mejorar las estrategias de orquestación y evaluar el rendimiento en simulaciones de desastres en vivo para validar aún más la eficacia de las respuestas a crisis del mundo real.

ODET ya cuenta con una arquitectura escalable y modular en la que las tareas que requieren un uso intensivo de cómputo pueden descargarse a endpoints basados en la nube, mientras que los agentes ligeros se ejecutan en hardware local de grado consumidor. En lugar de requerir una optimización adicional para la escalabilidad, la investigación futura podría centrarse en desarrollar endpoints específicos de dominio que mejoren la adaptabilidad de ODET a diferentes escenarios de desastre. La flexibilidad de la integración de endpoints permite a los investigadores adaptar el sistema a sus necesidades específicas, sin modificar la plataforma central. Por lo tanto, los estudios futuros deberían investigar las mejores prácticas para la implementación de endpoints, consideraciones de seguridad y estrategias para una integración perfecta con fuentes de datos externas.

Es importante enfatizar que ODET no es una mejora incremental en un sistema existente, sino una plataforma novedosa de extremo a extremo para la conciencia situacional de desastres basada en OSINT. Si bien trabajos anteriores han abordado componentes individuales como la detección de eventos, la sumarización y los knowledge graphs, ningún otro sistema integra estos elementos en una solución completamente automatizada, modular y escalable como ODET. Las herramientas actuales de inteligencia de desastres requieren intervención manual, carecen de procesamiento en tiempo real o carecen de la arquitectura flexible y adaptable de ODET. Los proyectos de investigación futuros pueden explorar más integraciones, pero la innovación fundamental de ODET como plataforma extensible continúa diferenciándolo de los enfoques existentes.

En resumen, se pueden identificar varias direcciones de investigación prometedoras que pueden explotar aún más el potencial de ODET. El refinamiento de los métodos de clustering podría conducir a una identificación más precisa de eventos, mientras que la integración de un wrapper de datos multimodales ampliaría significativamente las capacidades de análisis. Simultáneamente, una capa de validación mejorada puede mejorar la detección de noticias falsas y reducir las alucinaciones en los informes generados. La investigación futura debería centrarse en integrar eficientemente estas extensiones en los sistemas existentes sin comprometer la escalabilidad o la flexibilidad modular de ODET.

Chapter 6

Zusammenfassung und Ausblick

Diese Dissertation beschreibt die Entwicklung einer modularen, ML-basierten OSINT-Plattform zur automatisierten Lageberichterstattung im Katastrophenmanagement. Sie zeigt, wie durch die Kombination von Ereigniserkennung, Clustering, Zero-Shot-Klassifikation, Retrieval-Augmented Generation (RAG) und der Verwendung von Wissensgraphen ein skalierbares System zur Analyse von Microblogging-Daten für die automatisierte Lagedarstellung im Katastrophenmanagement realisiert werden kann.

Die schrittweise Entwicklung dieser Plattform wurde in drei Schlüsselstudien dokumentiert.

Studie 1: Towards Automated Situational Awareness Reporting for Disaster Management—A Case Study [16]

Die erste Studie fungierte als methodische Grundlage für die automatisierte Ereigniserkennung in Katastrophensituationen. Es wurde ein System entwickelt, das SimHash zur Gruppierung ähnlicher Nachrichten verwendet und in Kombination mit TF-IDF-basierter Merkmalsextraktion relevante Inhalte aus den heterogenen Microblogging-Daten identifiziert. Die anschließende Analyse der erkannten Ereignisse erfolgte mithilfe extraktiver Textzusammenfassungsmethoden (u. a. SpaCy, Gensim und NLTK), um die Eignung dieser Verfahren für die Verdichtung von Ereignisinformationen zu bewerten. Die Auswertung anhand von Daten zum russisch-ukrainischen Konflikt diente dazu, erkannte Ereignisse mit einer Wikipedia-Zeitleiste zu vergleichen und die Effektivität der Methode zur Ereigniserkennung zu bewerten. Dabei zeigte sich, dass die Methode eine größere Anzahl von Ereignissen als in Wikipedia

gelistet detektierte, gleichzeitig jedoch die Herausforderung bestand, irrelevante Inhalte zuverlässig zu filtern.

Studie 2: Harnessing OSINT in Disaster Management: Transforming Microblogging Posts Into Insightful Data Through Text Embedding [17]

In der zweiten Studie wurde der methodische Ansatz erweitert, indem Sentence Embedding mit Community Detection kombiniert wurden, um eine präzisere semantische Gruppierung von Ereignissen zu ermöglichen. Darüber hinaus wurden Wissensgraphen integriert, um die extrahierten Informationen zu strukturieren und ihre Interpretierbarkeit zu verbessern. Die Kombination dieser Methoden führte zu einer qualitativen Verbesserung der Ereignisgruppierung und ermöglichte eine strukturierte Darstellung der identifizierten Informationen.

Studie 3: AI-Enhanced Disaster Management: A Modular OSINT System for Rapid Automated Reporting [18]

In der dritten Studie erfolgte die Kombination der entwickelten Methoden zu einem vollständig modularen OSINT-System. Neben der Kombination von Zero-Shot-Klassifizierung, Community Detection-basiertem Clustering und wissensgraphbasierter Analyse wurde Retrieval-Augmented Generation (RAG) für die strukturierte Lageberichterstattung integriert, um die Qualität der Lageberichterstattung zu verbessern. Darüber hinaus ermöglicht eine agentenbasierte Architektur eine flexible Erweiterbarkeit des Systems und den modularen Austausch einzelner Komponenten. Die Evaluierung anhand realer Fallstudien (Hurrikan Harvey, Erdbeben 2023 in der Türkei) zeigte, dass das System nicht nur eine hohe faktische Übereinstimmung mit Wikipedia-Zusammenfassungen (AlignScore bis zu 89%) erreicht, sondern auch in unterschiedlichen Hardwareumgebungen effizient eingesetzt werden kann.

Die vorliegende Arbeit demonstriert, dass die Implementierung von OSINT und maschinellem Lernen im Katastrophenmanagement nicht nur eine effiziente Automatisierung der Datenanalyse ermöglicht, sondern auch zu einer beschleunigten und qualitativ hochwertigen Lageberichterstattung beiträgt. Darüber hinaus werden Fortschritte bei der automatisierten Ereigniserkennung durch Sentence Embedding, Com-

munity Detection und Zero-Shot Klassifizierung sowie Fortschritte bei der Retrieval-Augmented Generation (RAG) und der Integration von Wissensgraphen zur Verbesserung der Qualität und Relevanz der Lageberichterstattung im Katastrophenmanagement vorgestellt.

Im Folgenden wird eine Synthese der wichtigsten wissenschaftlichen Erkenntnisse präsentiert, die die methodischen Beiträge der drei Studien zusammenführt und in den Gesamtkontext der Forschung einordnet. Anschließend werden die Grenzen der entwickelten Ansätze diskutiert, wobei sowohl die methodischen als auch die technischen Herausforderungen behandelt werden. Abschließend werden vielversprechende Forschungsrichtungen für die Weiterentwicklung der OSINT-basierten Katastrophenanalyse skizziert.

6.1 Reflexion der Forschungsfragen

In diesem Abschnitt werden die zentralen Forschungsfragen dieser Dissertation beantwortet, indem die wichtigsten Ergebnisse der gesamten Untersuchung zusammengefasst werden.

Q1: **Können OSINT und ML zur automatisierten Lageberichterstattung im Katastrophenmanagement eingesetzt werden?** Basierend auf den durchgeführten Untersuchungen kann geschlussfolgert werden, dass OSINT und ML-basierte Verfahren zur automatisierten Lageberichterstattung eingesetzt werden können. Relevante Ereignisse können aus unstrukturierten Microblogging-Daten extrahiert, semantisch gruppiert und strukturiert verarbeitet werden. Dies ermöglicht eine kontinuierliche Erfassung von Katastrophenereignissen und deren Entwicklung auf Basis öffentlich zugänglicher Informationen. Eine wichtige Erkenntnis ist, dass der Einsatz von Wissensgraphen und Retrieval-Augmented Generation (RAG) zur Kontextualisierung der extrahierten Daten wesentlich zur Qualitätsverbesserung beiträgt. Dennoch bleibt die Herausforderung, irrelevante oder irreführende Informationen zuverlässig herauszufiltern.

Q2: **Wie kann die Genauigkeit der OSINT- und ML-basierten automatisierten Lageberichterstattung für das Katastrophen-**

management verbessert werden?
Eine deutliche Verbesserung der Genauigkeit wurde durch die Kombination verschiedener Techniken erreicht. Der Übergang von klassischen statistischen Methoden wie TF-IDF zu modernem Text Embedding hat die semantische Analyse der Daten erheblich verfeinert. Der Einsatz von Algorithmen zur Community Detection hat die Gruppierung thematisch verwandter Ereignisse verbessert, Redundanzen minimiert und die inhaltliche Konsistenz erhöht. Darüber hinaus hat sich die Zero-Shot-Klassifikation als besonders wertvoll erwiesen, da sie eine flexible, datenunabhängige Klassifizierung von Ereignissen ermöglicht, ohne dass umfangreiche manuelle Trainingsdatensätze erforderlich sind. Dadurch können neue und unvorhergesehene Ereignistypen schneller erfasst werden.

Q3: **Wie kann die Skalierbarkeit von OSINT- und ML-basierten Lösungen für die Lageberichterstattung im Katastrophenmanagement verbessert werden?**
Die Forschungsergebnisse zeigen, dass Skalierbarkeit durch eine agentenbasierte Architektur erreicht werden kann, die es erlaubt, verschiedene KI-Modelle modular und flexibel zu integrieren und auszutauschen. Dies ermöglicht eine dynamische Anpassung an unterschiedliche Anforderungen, wie z.B. die Verarbeitung großer Datenmengen oder die Nutzung unterschiedlicher Hardware-Umgebungen. Es konnte erfolgreich gezeigt werden, dass das System sowohl auf Hochleistungs-Cloud-Servern als auch auf Consumer-Hardware (z.B. MacBook Air M3) eingesetzt werden kann, was ein hohes Maß an Flexibilität bei der Bereitstellung ermöglicht. Dies zeigt, dass die entwickelten Methoden nicht nur in Forschungsumgebungen, sondern auch in realen Anwendungsszenarien zuverlässig funktionieren.

Diese Dissertation hat gezeigt, dass Open-Source-Intelligence (OSINT) in Verbindung mit modernen Methoden des maschinellen Lernens eine leistungsfähige Grundlage für die automatisierte Lageberichterstattung im Katastrophenmanagement bietet. Durch die Kombination mehrerer innovativer Methoden, von Community Detection über Zero-Shot-Klassifikation bis hin zu Retrieval-Augmented Generation, wurde ein System entwickelt, das sowohl präzise als auch flexibel ist.

Besonders deutlich wurde, dass die Qualität der Lageberichterstattung nicht nur von der technischen Umsetzung abhängt, sondern auch von der Fähigkeit, große Mengen unstrukturierter Daten effizient zu filtern und sinnvoll zu strukturieren. Während die entwickelten Methoden bereits ein hohes Leistungsniveau aufweisen, bestehen insbesondere bei der Erkennung von Fehlinformationen und der weiteren Automatisierung der Priorisierung von Informationen noch Herausforderungen.

Das in diesem Prozess entwickelte System schuf eine skalierbare, modular erweiterbare OSINT-Plattform, welche die Grundlage für zukünftige Forschung und praktische Anwendungen bildet. Die Ergebnisse unterstreichen die Notwendigkeit der Weiterentwicklung der OSINT und ML-basierten Lageberichterstattung, insbesondere durch verbesserte Mechanismen zur Qualitätssicherung der analysierten Daten und zur effizienten Integration neuer Datenquellen.

Um die gewonnenen Erkenntnisse ganzheitlich einordnen zu können, werden im folgenden Abschnitt die zentralen methodischen und konzeptionellen Aspekte zusammengeführt. Diese Synthese betrachtet die erzielten Fortschritte im Kontext bestehender Forschungsansätze und hebt die übergreifenden Implikationen dieser Dissertation hervor.

6.2 Synthese

Die vorliegende Dissertation demonstriert die Anwendung von Open Source Intelligence (OSINT) in Kombination mit maschinellen Lernmethoden zur automatisierten Erstellung von Lageberichten im Kontext des Katastrophenmanagements. Während OSINT in der Vergangenheit vornehmlich im nachrichtendienstlichen Bereich eingesetzt wurde, zeigt diese Arbeit, dass die zugrunde liegenden Prinzipien auch im humanitären Krisenmanagement effektiv anwendbar sind. Die drei wissenschaftlichen Publikationen dieser Arbeit befassen sich mit der schrittweisen Entwicklung, Evaluierung und Optimierung eines modularen, auf maschinellem Lernen basierenden OSINT-Systems. Die Ergebnisse zeigen, dass die Kombination aus Ereigniserkennung, Zero-Shot-Klassifizierung, Retrieval-Augmented-Generation (RAG) und modularer Systemarchitektur eine flexible, skalierbare und hochautomatisierte Lösung für die Katastrophenanalyse ermöglicht. Die Syn-

these dieser Ergebnisse demonstriert nicht nur den wissenschaftlichen Fortschritt in der OSINT-basierten Analyse, sondern auch den Beitrag des maschinellen Lernens zur Bewältigung der Herausforderungen traditioneller, manuell durchgeführter OSINT-Prozesse.

Die erste Studie legte die methodische Grundlage für die automatisierte Lageberichterstattung im Katastrophenmanagement, indem sie Microblogging-Daten zur Ereigniserkennung und -synthese verwendete. Die entwickelte Methode nutzt SimHash zur Gruppierung ähnlicher Nachrichten, um latente Ereignisse in Social-Media-Daten effizient zu identifizieren. Zusätzlich kommt Term Frequency-Inverse Document Frequency (TF-IDF) zur Merkmalsextraktion zum Einsatz, um relevante Inhalte aus einem heterogenen und verrauschten Datensatz hervorzuheben. Die Ereignisse werden gruppiert, indem Nachrichten mit ähnlichen SimHash-Darstellungen geclustert und thematisch verwandte Beiträge zu kohärenten Ereignisclustern zusammengefasst werden. Die Cluster wurden dann mit Methoden der extraktiven Textzusammenfassung, insbesondere SpaCy, Gensim und NLTK, analysiert, um ihre Eignung für die automatisierte Verdichtung von Ereignisinformationen zu bewerten. Die Auswertung erfolgte anhand von Daten aus dem Russland-Ukraine-Konflikt, bei dem die erkannten Ereignisse mit einer Wikipedia-Zeitleiste verglichen wurden, um die Effektivität der automatisierten Ereigniserkennung zu messen. Die Ergebnisse zeigen, dass die vorgeschlagene Methodik eine größere Anzahl von Ereignissen als in Wikipedia gelistet detektierte, wobei SpaCy die besten Ergebnisse bei der Textzusammenfassung lieferte. Gleichzeitig hebt die Studie die Herausforderungen beim Herausfiltern irrelevanter oder irreführender Inhalte hervor und unterstreicht die Notwendigkeit einer weiteren Verfeinerung OSINT-gestützter Lageberichterstattung.

Aufbauend auf den Erkenntnissen der ersten Studie erweiterte die zweite Studie den methodischen Ansatz, indem sie Sentence Embedding mit der Erkennung von Communities kombinierte, um ein präziseres semantisches Clustering zu ermöglichen. Während in der ersten Studie SimHash für die Gruppierung ähnlicher Nachrichten verwendet wurde, ermöglicht die hier vorgestellte Methode eine verfeinerte semantische Ereignisgruppierung. Darüber hinaus wurden Wissensgraphen in das System integriert, um die extrahierten Informationen

zu strukturieren und ihre Interpretierbarkeit zu verbessern. Dieser Ansatz unterscheidet sich von klassischen maschinellen Lernmethoden, die auf große, gelabelte Datensätze angewiesen sind. Die Ergebnisse der Studie zeigen, dass die Kombination aus semantischem Clustering und strukturierten Wissensgraphen die Organisation von Ereignisdaten verbessert. Darüber hinaus markiert sie einen witeren Schritt in Richtung strukturierter Informationsaufbereitung, indem relevante Inhalte aus großen Mengen unstrukturierter Daten extrahiert und in einer geordneten Form für die weitere Verarbeitung bereitgestellt werden. Die Ergebnisse zeigen weiterhin, dass OSINT-gestützte Lageberichterstattung nicht nur automatisiert, sondern auch effizient und in großem Maßstab durchgeführt werden kann.

Die dritte Studie integrierte die entwickelten Methoden in ein vollständig modulares und skalierbares OSINT-System. Aufbauend auf den in früheren Arbeiten vorgestellten Methoden erweitert es das System um Retrieval-Augmented Generation (RAG) für strukturierte Lageberichterstattung und Wissensgraphen für die kontextbezogene Analyse der extrahierten Informationen. Während frühere Arbeiten sich auf die Erkennung und Community Detection-basierte Gruppierung von Ereignissen konzentrierten, liegt der Schwerpunkt hier auf der End-to-End Lageberichterstattungsplattform. Die Kombination aus Zero-Shot-Klassifizierung, Community Detection-basiertem Clustering und wissensgraphbasierten Analysetechniken ermöglicht die detaillierte und kontextbezogene Verarbeitung relevanter Informationen. Darüber hinaus wird eine agentenbasierte Architektur implementiert, die eine flexible Anpassung und Erweiterbarkeit des Systems ermöglicht. Die Evaluierung des Systems anhand realer Fallstudien, darunter der Hurrikan Harvey und das Erdbeben in der Türkei im Jahr 2023, zeigt, dass das entwickelte OSINT-Framework nicht nur eine hohe faktische Übereinstimmung mit Wikipedia-Zusammenfassungen (AlignScore bis zu 89%) erreicht, sondern auch eine konsistente und nachvollziehbare Erstellung von Lageberichten ermöglicht. Die vorliegende Studie markiert somit einen entscheidenden Schritt in der Entwicklung eines vollautomatisierten, auf maschinellem Lernen basierenden OSINT-Systems für das Katastrophenmanagement.

Die Forschung zeigt, dass die Integration von Open Source Intelli-

gence (OSINT) in das Katastrophenmanagement zu einer signifikanten Verbesserung des Lagebewusstseins führt. Während traditionelle Methoden auf manueller Datenauswertung beruhen, nutzen moderne Ansätze maschinelles Lernen, um Informationen aus Microblogging-Daten automatisch zu extrahieren und zu verarbeiten. Dies reduziert die Zeit, die für die Extraktion relevanter Inhalte benötigt wird, und ermöglicht eine schnellere Entscheidungsfindung. Die technologische Grundlage für eine effiziente und skalierbare Analyse bildet die Verwendung von Texteinbettung, Community-Erkennung, Clustering und Zero-Shot-Klassifizierung. Ein signifikanter Fortschritt ist insbesondere die Automatisierung der Datenverarbeitung, die nicht nur die Geschwindigkeit erhöht, sondern auch eine konsistente und nachvollziehbare Analyse großer Datenmengen gewährleistet. Der Open Source Intelligence Disaster Event Tracker (ODET) zeichnet sich durch seine Modularität und Skalierbarkeit aus, wodurch eine flexible Anpassung an verschiedene Szenarien ermöglicht wird. Die Trennung der Verarbeitungsschritte in dedizierte Agenten und Endpunkte erleichtert den Austausch von Modellen und Technologien und gewährleistet eine kontinuierliche Weiterentwicklung und Anpassungsfähigkeit an neue Anforderungen. Die Skalierbarkeit des Systems gewährleistet den Einsatz in Geräten mit geringer Leistung und in Cloud-Instanzen mit hoher Leistung. Eine Evaluierung anhand realer Fallstudien zeigt, dass ODET auch mit ungeschulten Modellen effektiv bleibt und eine hohe Übereinstimmung mit Referenzberichten erzielt. Diese Architektur stellt sicher, dass ODET langfristig anwendbar bleibt und in verschiedenen Anwendungsszenarien eingesetzt werden kann.

Ein wichtiger Aspekt dieser Forschung ist die objektive Bewertung der generierten Berichte, um die Qualität und Zuverlässigkeit der automatisierten Lagebeurteilung sicherzustellen. Zu diesem Zweck wurde die AlignScore-Metrik verwendet, um die inhaltliche Konsistenz mit etablierten Referenzquellen wie Wikipedia zu quantifizieren. Fallstudien, die die Berichterstattung über den Hurrikan Harvey und das Erdbeben in der Türkei im Jahr 2023 untersuchen, belegen, dass das System eine bemerkenswerte inhaltliche Konsistenz aufweist und eine solide Grundlage für die Entscheidungsfindung bietet. Diese Erkenntnisse unterstreichen die Bedeutung maschineller Situationserkennung

im Katastrophenmanagement und demonstrieren, dass OSINT-basierte Methoden eine verlässliche Alternative zu herkömmlichen Ansätzen darstellen können. Die Kombination aus modularer Architektur, skalierbaren Methoden und wissenschaftlich fundierter Qualitätsbewertung macht ODET zu einem leistungsstarken Werkzeug für die automatisierte Informationsbeschaffung in Krisensituationen.

Trotz dieser Erfolge bleiben offene Forschungsfragen und Herausforderungen bestehen. Ein zentrales Problem besteht nach wie vor in der Erkennung und Handhabung von Fehlinformationen und gezielten Desinformationskampagnen. Obwohl die entwickelten Methoden dazu in der Lage sind, Inhalte anhand semantischer Kriterien zu gruppieren und zu klassifizieren, ist es nach wie vor schwierig, fehlerhafte oder irreführende Informationen automatisch zuverlässig zu erkennen. Die derzeitigen generativen großen Sprachmodelle bergen zudem das Risiko von Halluzinationen, d. h. der Erzeugung von plausibel klingenden, aber sachlich falschen Inhalten. Die Weiterentwicklung von Methoden zur Überprüfung und Sicherung der Qualität generierter Berichte ist daher ein wichtiger zukünftiger Forschungsschwerpunkt.

Zusammenfassend zeigt diese Dissertation, dass ML-basierte OSINT-Methoden praktikable und effiziente Lösungen für die automatisierte Lagebeurteilung im Katastrophenmanagement sind. Die Kombination verschiedener ML-Techniken mit modularen Systemarchitekturen ermöglicht die Entwicklung einer Lösung, die sowohl flexibel als auch leistungsstark ist. Die Synthese der Ergebnisse zeigt, dass die OSINT-Analyse durch maschinelles Lernen beschleunigt und qualitativ verbessert werden kann. Die in dieser Dissertation präsentierten Methoden haben das Potenzial, die Art und Weise der Sammlung, Analyse und Meldung von Kriseninformationen grundlegend zu verändern.

6.3 Einschränkungen

Trotz der hohen Leistung von ODET bei der automatisierten Analyse von Microblogging-Daten gibt es einige methodische und technische Einschränkungen, die weitere Forschung und Entwicklung erfordern. Eine zentrale Herausforderung besteht darin, dass die Erkennung von Fake News und Halluzinationen nach wie vor als ungelöst gilt. ODET ver-

wendet keine expliziten Mechanismen zur Überprüfung der extrahierten Informationen und überprüft nicht, ob bestimmte Aussagen aus zuverlässigen oder manipulierten Quellen stammen. Obwohl die aggregierte Analyse von Clustern Verzerrungen teilweise ausgleicht, besteht weiterhin das Risiko, dass falsche oder irreführende Informationen in die generierten Lageberichte aufgenommen werden. Sprachmodelle, die syntaktisch korrekte, aber sachlich falsche Inhalte erzeugen können, sind in diesem Zusammenhang besonders problematisch. Obwohl das System versucht, solche Effekte durch strukturierte Eingabeaufforderungen und spezialisierte Verarbeitungsschritte zu minimieren, handelt es sich hierbei um ein allgemeines Problem bei der Verwendung generativer Sprachmodelle, das nicht vollständig beseitigt werden kann.

Ein weiterer wichtiger Punkt ist die Einschränkung auf textbasierten Daten. ODET verarbeitet nur textbasierte Informationen und berücksichtigt keine zusätzlichen Datenformate wie Bilder, Videos oder Audioaufnahmen. In Krisensituationen können visuelle oder Audioinhalte eine wertvolle Ergänzung zu Textnachrichten sein, insbesondere wenn es Sprachbarrieren oder Missverständnisse gibt. Daher könnte die Hinzufügung einer multimodalen Analyse die Zuverlässigkeit der Lagebeurteilung verbessern.

Darüber hinaus kann die Granularität der Ereignisdarstellung verbessert werden. Während der Clustering-Prozess eine zuverlässige Erkennung und Strukturierung größerer Ereignisse ermöglicht, können sehr kleine oder lokal begrenzte Vorfälle in großen Clustern verloren gehen oder nicht als unabhängige Ereignisse erfasst werden. Dies gilt insbesondere für subtile Entwicklungen, die möglicherweise erst über längere Zeiträume als kohärent erkannt werden. Die derzeitige Methodik stellt sicher, dass globale oder weit verbreitete Ereignisse effizient erkannt werden, aber die Feinabstimmung für die Erkennung von Mikroereignissen bleibt eine offene Herausforderung.

6.4 Ausblick

Obwohl ODET bereits eine leistungsstarke und flexible Plattform für die automatisierte Analyse von Microblogging-Daten im Katastrophen-

management bietet, gibt es mehrere Bereiche, in denen zukünftige Forschung das System weiterentwickeln und optimieren könnte. Aufgrund seiner modularen Architektur kann ODET gezielt erweitert werden, ohne die Kernfunktionalität oder Skalierbarkeit zu beeinträchtigen. Insbesondere die Verbesserung der Ereignis-Granularität, die Integration multimodaler Daten und die Optimierung von Validierungsmechanismen für Fake News und Halluzinationen sind wichtige Ansatzpunkte für die zukünftige Entwicklung.

Eine der komplexesten offenen Forschungsfragen ist die Validierung der gesammelten Informationen. Die Erkennung von Fake News und die Reduzierung von Halluzinationen in den generierten Berichten sind für die Zuverlässigkeit der OSINT-Analyse von entscheidender Bedeutung. Obwohl das aktuelle System durch Clustering und Mehrheitsentscheidungen bereits eine gewisse Robustheit gegenüber einzelnen Anomalien bietet, könnte eine zusätzliche Validierungsebene den Wert der generierten Berichte weiter verbessern. Eine mögliche Lösung könnte eine eingehende Netzwerküberprüfung von Social-Media-Konten sein, um die Verbreitung von Fake News zu verhindern, sowie die Implementierung ausgefeilterer Validierungsebenen. Was die verbleibenden Probleme von Halluzinationen betrifft, so könnte es jedoch eine ganz neue Generation von KI-Modellen erfordern, um dieses Risiko zu beseitigen.

Obwohl diese Methodik das Filtern und Klassifizieren von Texten umfasst, ist die Erkennung von Falschinformationen nach wie vor eine unzureichend erforschte Herausforderung. Die Verbreitung von Fake News, manipulierten Inhalten und gegnerischen Desinformationskampagnen erfordert eine robustere Strategie. Zukünftige Forschung sollte sich mit Rahmenbedingungen für Gegentests, Strategien zur Eindämmung von Voreingenommenheit und Mechanismen zur Bewertung der Vertrauenswürdigkeit befassen, um die Verlässlichkeit von Daten in Krisenszenarien zu verbessern.

Ein weiterer wichtiger Bereich für zukünftige Forschung ist die Erweiterung von ODET auf multimodale Datenverarbeitung. Bisher konzentriert sich das System ausschließlich auf textbasierte Informationen, in Krisensituationen spielen aber auch Bilder, Videos und Audiodaten eine entscheidende Rolle. Da ODET modular aufgebaut ist,

kann ein separater Datenhandler entwickelt werden, um multimodale Modelle zu integrieren. Beispielsweise können bestehende multimodale Sprachmodelle verwendet werden, um die Zero-Shot-Klassifikation auf visuelle Inhalte auszuweiten. Damit wäre es möglich, nicht nur Textnachrichten, sondern auch Bilder aus Katastrophengebieten oder relevante Videos aus sozialen Netzwerken automatisch zu analysieren. Zukünftige Forschung sollte untersuchen, wie multimodale Fusionstechniken die Ereigniserkennung und das Situationsbewusstsein verbessern können, indem verschiedene Datentypen für eine ganzheitlichere Analyse kombiniert werden.

Neben der Erweiterung auf multimodale Daten liegt eine weitere mögliche Herausforderung in der Granularität der Ereignisdarstellung. Während die derzeitige Clustering-Methode zuverlässige Ergebnisse für größere Ereignisse liefert, kann eine Verfeinerung der Methode dazu beitragen, kleinere oder stärker lokalisierte Ereignisse genauer zu erfassen. Ein vielversprechender Ansatz ist das adaptive Clustering, das dynamisch zwischen großen und kleinen Ereignissen unterscheidet. Der Einsatz hybrider Clusterverfahren, die semantische Ähnlichkeiten und zeitliche Abfolgen berücksichtigen, kann genauere Unterscheidungen ermöglichen.

Technologische Entwicklungen im Bereich von Sprachmodellen und KI-basierten Clustering-Methoden könnten weitere Möglichkeiten zur Optimierung von ODET eröffnen. Zukünftige multimodale Modelle könnten eine genauere Erkennung und Verarbeitung verschiedener Datentypen ermöglichen, während effizientere Clustering-Algorithmen eine noch genauere Strukturierung von Ereignissen erlauben würden.

ODET ist für die Ausführung in der Cloud oder auf Consumer-Hardware konzipiert, die in Katastrophenszenarien zur Verfügung gestellt werden kann, und ist daher flexibel für den Einsatz auf der Grundlage der verfügbaren Infrastruktur. Es ist nicht für mobile oder Edge-Computing-Umgebungen konzipiert, sondern arbeitet mit vorhandenen Computerressourcen an Katastrophenstandorten. Im Gegensatz zu anderen OSINT-basierten Katastropheninformations-Tools implementiert ODET eine modulare Endpunkt-Architektur, die eine effiziente Lastverteilung ermöglicht. Das bedeutet, dass lastintensive Module wie Large Lan-

guage Models (LLMs) in der Cloud ausgeführt werden können, während leichtgewichtige Agenten auf Consumer-Hardware wie einem MacBook laufen. Diese verteilte Architektur gewährleistet Skalierbarkeit, Anpassungsfähigkeit und effiziente Ressourcennutzung in unterschiedlichen Umgebungen. Zukünftige Arbeiten sollten sich auf die Optimierung von Bereitstellungsprozessen, die Verbesserung von Orchestrierungsstrategien und die Bewertung der Leistung in Live-Katastrophensimulationen konzentrieren, um die Wirksamkeit der Krisenreaktion in der Praxis weiter zu validieren.

ODET verfügt bereits über eine skalierbare, modulare Architektur, in der rechenintensive Aufgaben auf Cloud-basierte Endpunkte ausgelagert werden können, während leichtgewichtige Agenten auf lokaler, benutzerfreundlicher Hardware ausgeführt werden. Anstelle einer weiteren Optimierung der Skalierbarkeit könnte sich die zukünftige Forschung auf die Entwicklung domänenspezifischer Endpunkte konzentrieren, die die Anpassungsfähigkeit von ODET an verschiedene Katastrophenszenarien verbessern. Die Flexibilität der Endpunktintegration ermöglicht es den Forschern, das System an ihre spezifischen Bedürfnisse anzupassen, ohne die Kernplattform ändern zu müssen. Zukünftige Studien sollten daher Best Practices für die Endpunktimplementierung, Sicherheitsaspekte und Strategien für die nahtlose Integration mit externen Datenquellen untersuchen.

Es ist wichtig zu betonen, dass es sich bei ODET nicht um eine inkrementelle Verbesserung eines bestehenden Systems handelt, sondern um eine neuartige End-to-End-Plattform für OSINT-basierte Katastrophenlageerkennung. Während sich frühere Arbeiten mit einzelnen Komponenten wie Ereigniserkennung, Synthese und Wissensgraphen befassten, integriert kein anderes System diese Elemente in eine vollautomatische, modulare und skalierbare Lösung wie ODET. Aktuelle Katastropheninformations-Tools erfordern entweder manuelle Eingriffe, bieten keine Echtzeitverarbeitung oder verfügen nicht über die flexible und anpassungsfähige Architektur von ODET. Zukünftige Forschungsprojekte können weitere Integrationen untersuchen, aber die grundlegende Innovation von ODET als erweiterbare Plattform unterscheidet es weiterhin von bestehenden Ansätzen.

Zusammenfassend lassen sich mehrere vielversprechende Forschungsrichtungen identifizieren, die das Potenzial von ODET weiter ausschöpfen können. Eine Verfeinerung der Clustering-Methoden könnte zu einer präziseren Ereignisidentifizierung führen, während die Integration einer multimodalen Datenverpackung die Analysefähigkeiten erheblich erweitern würde. Gleichzeitig kann eine verbesserte Validierungsschicht die Erkennung von Fake News verbessern und Halluzinationen in den generierten Berichten reduzieren. Die zukünftige Forschung sollte sich darauf konzentrieren, diese Erweiterungen effizient in bestehende Systeme zu integrieren, ohne die Skalierbarkeit oder modulare Flexibilität von ODET zu beeinträchtigen.

Chapter 7

Conclusión y Futuros Trabajos

Esta tesis describe el desarrollo de una plataforma modular basada en machine learning para OSINT (Inteligencia de Fuentes Abiertas) destinada a la generación automatizada de informes de conciencia situacional en la gestión de desastres. Demuestra cómo se puede implementar una plataforma escalable para analizar datos de microblogs para la generación automatizada de informes de conciencia situacional en la gestión de desastres, mediante la combinación de detección de eventos, clusterización, clasificación zero-shot, generación aumentada por recuperación (RAG), y el uso de grafos de conocimiento.

El desarrollo de esta plataforma, publicación por publicación, ha sido documentado en tres estudios clave.

Estudio 1: Towards Automated Situational Awareness Reporting for Disaster Management—A Case Study [16]

El primer estudio estableció la base metodológica para la detección automatizada de eventos en situaciones de desastre. Se desarrolló un sistema que utiliza SimHash para agrupar mensajes similares con el fin de procesar eficientemente grandes volúmenes de texto. En combinación con la extracción de características basada en TF-IDF, se identificó contenido relevante a partir de datos heterogéneos de microblogs. Análisis posteriores utilizando métodos de resumen extractivo de texto (incluyendo SpaCy, Gensim y NLTK) se emplearon para condensar información relacionada con eventos. La evaluación basada en el conflicto ruso-ucraniano se centró en identificar ocurrencias de eventos y demostró la viabilidad del enfoque para la detección de eventos, al tiempo que destacó los desafíos de filtrar contenido irrelevante o engañoso.

Estudio 2: Harnessing OSINT in Disaster Management: Transforming Microblogging Posts Into Insightful Data Through Text Embedding [17]

El segundo estudio amplió el enfoque metodológico combinando incrustación de frases con detección de comunidades para permitir una agrupación semántica más precisa de eventos. Mientras que el primer estudio utilizó SimHash para la agrupación basada en similitud de términos, el método presentado aquí permite una clasificación semántica más refinada a nivel de eventos. Además, se integraron grafos de conocimiento en el sistema para estructurar la información extraída y mejorar su interpretabilidad. Este enfoque difiere de los métodos clásicos de aprendizaje automático que dependen de grandes conjuntos de datos etiquetados manualmente. El estudio muestra que la combinación de agrupamiento semántico y representación estructurada de eventos puede mejorar significativamente la precisión en el reconocimiento de eventos. También introduce el siguiente paso hacia la priorización automatizada de la información, al extraer contenido relevante de grandes volúmenes de datos no estructurados y organizarlos de manera coherente para la generación de informes de conciencia situacional.

Estudio 3: AI-Enhanced Disaster Management: A Modular OSINT System for Rapid Automated Reporting [18]

En el tercer estudio, los métodos desarrollados se integraron en un sistema OSINT completamente modular. Además de la combinación de clasificación zero-shot, agrupamiento basado en detección de comunidades y análisis basado en grafos de conocimiento, se incorporó la generación aumentada por recuperación (RAG) para mejorar la calidad de los informes estructurados. Asimismo, una arquitectura basada en agentes permite la expansión flexible del sistema y la integración modular de componentes individuales. La evaluación basada en casos de estudio reales (Huracán Harvey, Terremoto de Turquía 2023) demostró que el sistema no solo logra una alta alineación con los resúmenes de eventos de Wikipedia (AlignScore hasta 89%), sino que también demuestra viabilidad de implementación en diferentes entornos de hardware.

Esta tesis demuestra que la aplicación de OSINT y aprendizaje au-

tomático en la gestión de desastres no solo permite una automatización eficiente del análisis de datos, sino que también contribuye a una conciencia situacional acelerada y de alta calidad. Simultáneamente, se han presentado avances en la detección automatizada de eventos mediante incrustación de frases, detección de comunidades y clasificación zero-shot, así como avances en generación aumentada por recuperación (RAG) y la integración de grafos de conocimiento para mejorar la calidad y relevancia de los informes de conciencia situacional en la gestión de desastres.

A continuación, se presenta una síntesis de los hallazgos científicos clave, que combina las contribuciones metodológicas de los tres estudios presentados a lo largo de esta tesis y los sitúa en el contexto general. Posteriormente, se discuten las limitaciones de los enfoques desarrollados, abordando tanto los desafíos metodológicos como técnicos. Finalmente, se esbozan direcciones de investigación prometedoras para el desarrollo ulterior del análisis de desastres basado en OSINT.

7.1 Revisión de las Preguntas de Investigación

En esta sección, se responde a las preguntas centrales de investigación de esta tesis mediante la integración de los hallazgos clave de toda la investigación.

Q1: **¿Se pueden utilizar OSINT y aprendizaje automático para la generación automatizada de informes de conciencia situacional en la gestión de desastres?** Basándose en la investigación realizada, se puede concluir que OSINT y los sistemas basados en aprendizaje automático pueden utilizarse eficazmente para la generación automatizada de informes de conciencia situacional. Los eventos relevantes pueden extraerse de datos no estructurados de microblogs, agruparse semánticamente y procesarse de manera estructurada. Esto permite el registro continuo de eventos de desastre y su desarrollo basado en información públicamente disponible. Un hallazgo clave es que el uso de grafos de conocimiento y generación aumentada por recuperación (RAG)

para contextualizar los datos extraídos contribuye significativamente a mejorar la calidad de los informes. No obstante, persiste el desafío de filtrar de manera fiable la información irrelevante o engañosa, un tema que sigue siendo un área activa de investigación.

Q2: **¿Cómo se puede mejorar la precisión de los informes automatizados de conciencia situacional basados en OSINT y aprendizaje automático para la gestión de desastres?** La precisión mejoró significativamente gracias a la combinación de diversas técnicas avanzadas. La transición de métodos estadísticos clásicos, como TF-IDF, a métodos modernos de incrustación ha refinado significativamente el análisis semántico de los datos. Esta evolución metodológica permite capturar relaciones semánticas más sutiles entre diferentes fragmentos de información, mejorando así la capacidad del sistema para identificar contenido relacionado incluso cuando no comparte términos específicos. El uso de algoritmos de detección de comunidades ha mejorado la agrupación de eventos temáticamente relacionados, minimizado redundancias y aumentado la consistencia del contenido. Además, la clasificación zero-shot ha demostrado ser particularmente valiosa porque permite una clasificación flexible e independiente de datos de los eventos sin necesidad de extensos conjuntos de datos de entrenamiento manual. Esto ayuda a capturar nuevos tipos de eventos imprevistos con mayor rapidez y precisión.

Q3: **¿Cómo se puede mejorar la escalabilidad de las soluciones basadas en OSINT y aprendizaje automático para la generación de informes de conciencia situacional en la gestión de desastres?** Los resultados obtenidos de esta investigación muestran que la escalabilidad puede lograrse mediante una arquitectura basada en agentes que permite integrar e intercambiar diferentes modelos de IA de manera modular y flexible. Esto permite una adaptación dinámica a diversos requisitos, como el procesamiento de grandes cantidades de datos o el uso de diferentes entornos de hardware. Se ha demostrado con éxito que el sistema puede utilizarse tanto en servidores cloud de alto rendimiento como en hardware de consumo (por ejemplo, MacBook Air M3), lo que permite un alto grado de flexibilidad en la implementación.

La arquitectura distribuida facilita que los módulos con alta demanda computacional, como los grandes modelos de lenguaje (LLMs), puedan ejecutarse en la nube, mientras que los agentes más ligeros pueden operar en hardware de consumo local. Esto confirma que los métodos desarrollados son robustos tanto en entornos de investigación como en aplicaciones reales.

Esta tesis ha demostrado que la Inteligencia de Fuentes Abiertas (OSINT) en conjunto con métodos modernos de aprendizaje automático proporciona una base poderosa para la generación automatizada de informes de conciencia situacional en la gestión de desastres. Mediante la combinación de varios métodos innovadores, desde detección de comunidades hasta clasificación zero-shot y generación aumentada por recuperación, se ha desarrollado un sistema que es tanto preciso como flexible. Resulta particularmente evidente que la calidad de los informes de conciencia situacional depende no solo de la implementación técnica, sino también de la capacidad para filtrar eficientemente y estructurar de manera significativa grandes cantidades de datos no estructurados. Si bien los métodos desarrollados ya muestran un alto nivel de rendimiento, persisten desafíos, particularmente en lo que respecta a la detección de información errónea y la mayor automatización de la priorización de la información. El sistema desarrollado en este proceso creó un marco OSINT escalable y modularmente expandible que forma la base para futuras investigaciones y aplicaciones prácticas. Los resultados subrayan la necesidad de seguir desarrollando la generación de informes de conciencia situacional basados en OSINT, particularmente mediante mecanismos mejorados para la garantía de calidad de los datos analizados y la integración eficiente de nuevas fuentes de datos. Para clasificar de manera holística los conocimientos adquiridos, la siguiente sección reúne los aspectos metodológicos y conceptuales centrales. Esta síntesis considera los avances realizados en el contexto de los enfoques de investigación existentes y destaca las implicaciones generales de esta tesis.

7.2 Síntesis

Esta tesis demuestra cómo la Inteligencia de Fuentes Abiertas (OSINT) puede integrarse con métodos de aprendizaje automático para generar automáticamente informes de conciencia situacional en la gestión de desastres. Si bien OSINT se ha utilizado principalmente en contextos de inteligencia, esta tesis demuestra que estos principios también pueden aplicarse de manera efectiva en la gestión de crisis humanitarias. Las tres publicaciones científicas de esta tesis tratan sobre el desarrollo, evaluación y optimización progresiva de un sistema modular de OSINT basado en aprendizaje automático. Se demostró que la combinación de detección de eventos, clasificación zero-shot, generación aumentada por recuperación (RAG) y una arquitectura de sistema modular permite una solución flexible, escalable y altamente automatizada para el análisis de desastres. La síntesis de estos resultados demuestra no solo los avances científicos en el análisis basado en OSINT, sino también cómo el aprendizaje automático puede ayudar a superar los desafíos de los procesos OSINT tradicionales realizados manualmente.

El primer estudio estableció la base metodológica para la generación automatizada de informes de situación en la gestión de desastres mediante el uso de datos de microblogs para la detección y síntesis de eventos. Este estudio presenta un sistema que utiliza SimHash para agrupar mensajes similares con el fin de identificar eficientemente eventos relacionados con desastres en datos de redes sociales. Además, se utiliza Term Frequency-Inverse Document Frequency (TF-IDF) para la extracción de características con el fin de destacar contenido relevante de un conjunto de datos heterogéneo y ruidoso. Los eventos se agrupan mediante el clustering de mensajes con representaciones SimHash similares y la agregación de publicaciones temáticamente relacionadas en clusters coherentes de eventos. Estos clusters fueron luego analizados utilizando métodos de resumen extractivo de texto, específicamente SpaCy, Gensim y NLTK, para evaluar su idoneidad para condensar información relacionada con eventos. La evaluación se llevó a cabo utilizando datos del conflicto ruso-ucraniano, donde los eventos detectados se compararon con una línea de tiempo de Wikipedia para medir la efectividad de la detección automatizada de eventos. Los

resultados muestran que la metodología propuesta detectó un mayor número de eventos que los enumerados en Wikipedia, con SpaCy proporcionando los mejores resultados de resumen. Al mismo tiempo, el estudio destaca los desafíos de filtrar contenido irrelevante o engañoso y subraya la necesidad de un mayor refinamiento del procesamiento OSINT relacionado con desastres.

Partiendo de los hallazgos del primer estudio, el segundo estudio amplió el enfoque metodológico combinando incrustación de frases con detección de comunidades para un clustering semántico más preciso. Mientras que el primer estudio utilizó SimHash para el clustering basado en la similitud de términos, el método presentado aquí permite una agrupación semántica más refinada de los eventos. Además, se integraron grafos de conocimiento en el sistema para estructurar la información extraída y mejorar la interpretabilidad. Este enfoque difiere de los métodos clásicos de aprendizaje automático que dependen de grandes conjuntos de datos etiquetados. El estudio muestra que la combinación de clustering semántico y representación estructurada de eventos puede mejorar la precisión del reconocimiento. También introduce el siguiente paso hacia la priorización automatizada de la información mediante la extracción de contenido relevante de grandes volúmenes de datos no estructurados y su estructuración para informes de conciencia situacional. Los resultados demuestran que el análisis de desastres basado en OSINT puede automatizarse mientras se mejora la organización de eventos y la interpretabilidad a escala.

El tercer estudio integró los métodos desarrollados en un sistema OSINT completamente modular y escalable. Basándose en los métodos introducidos en trabajos anteriores, amplía el sistema para incluir generación aumentada por recuperación (RAG) para informes estructurados y grafos de conocimiento para el análisis contextual de la información extraída. Mientras que los trabajos anteriores se han centrado en la detección y el clustering semántico de eventos, el enfoque aquí está en la plataforma de informes de conciencia situacional de extremo a extremo. La combinación de clasificación zero-shot, clustering basado en detección de comunidades y análisis basado en grafos de conocimiento permite el procesamiento detallado y contextual de información relevante. Además, se implementa una arquitectura basada en

agentes, permitiendo una adaptación flexible y extensibilidad modular del sistema. La evaluación del sistema utilizando casos de estudio reales, incluidos el huracán Harvey y el terremoto de Turquía de 2023, muestra que el marco OSINT desarrollado no solo logra una alta alineación factual en la detección de eventos (AlignScore hasta 89%) sino que también permite informes de conciencia situacional consistentes y trazables. Por lo tanto, este estudio representa un paso decisivo en el desarrollo de un sistema OSINT basado en aprendizaje automático completamente automatizado para la gestión de desastres.

La investigación presentada muestra que la integración de Open Source Intelligence (OSINT) en la gestión de desastres permite una mejora significativa en la conciencia situacional. Mientras que los métodos tradicionales dependen de la evaluación manual de datos, los enfoques modernos utilizan aprendizaje automático para extraer y procesar automáticamente información de datos de microblogs. Esto reduce el tiempo necesario para extraer contenido relevante y permite una toma de decisiones más rápida. El uso de incrustación de texto, detección de comunidades, clustering y clasificación zero-shot forma la base tecnológica para un análisis eficiente y escalable. En particular, la automatización del procesamiento de datos es un gran paso adelante, no solo aumentando la velocidad, sino también asegurando un análisis consistente y trazable de grandes volúmenes de datos.

El sistema Open Source Intelligence Disaster Event Tracker (ODET) se caracteriza por su modularidad y escalabilidad, permitiendo que se adapte de manera flexible a diferentes escenarios. La separación de los pasos de procesamiento en agentes y endpoints dedicados facilita el intercambio de modelos y tecnologías. Esto asegura un desarrollo continuo y adaptabilidad a los nuevos requisitos. La escalabilidad del sistema permite su uso en dispositivos de bajo rendimiento e instancias cloud de alto rendimiento. Una evaluación basada en casos de estudio reales muestra que el ODET permanece efectivo incluso con modelos no entrenados y logra un alto grado de consistencia con informes de referencia.

Es importante destacar que ODET no representa una mejora incremental sobre un sistema existente, sino que constituye una plataforma

completamente nueva y de extremo a extremo para la conciencia situacional en desastres basada en OSINT. Mientras que investigaciones previas han abordado componentes individuales, como la detección de eventos, la generación de resúmenes y el uso de grafos de conocimiento, ningún otro sistema ha integrado estos elementos en una solución totalmente automatizada, modular y escalable como ODET.

Un aspecto clave de esta investigación es la evaluación objetiva de los informes generados para garantizar la calidad y fiabilidad de la evaluación automatizada de situaciones. Con este fin, se utilizó la métrica AlignScore para cuantificar la consistencia del contenido con fuentes de referencia establecidas, como Wikipedia. Los estudios de caso sobre la cobertura del huracán Harvey y el terremoto de 2023 en Turquía muestran que el sistema logra un alto nivel de consistencia de contenido y proporciona una base sólida para la toma de decisiones. Esto subraya la importancia de la conciencia situacional basada en IA en la gestión de desastres y demuestra que los métodos basados en OSINT pueden proporcionar una alternativa fiable a los enfoques tradicionales. La combinación de arquitectura modular, métodos escalables y evaluación de calidad científicamente fundamentada hace de ODET una herramienta poderosa para la recopilación automatizada de inteligencia en situaciones de crisis.

A pesar de estos éxitos, siguen existiendo preguntas de investigación abiertas y desafíos. Un problema central sigue siendo la detección y el manejo de información errónea y en campañas de desinformación dirigidas. Si bien los métodos desarrollados pueden agrupar y clasificar contenido según criterios semánticos, queda la cuestión de cómo detectar de manera fiable y automática información falsa o engañosa. La generación actual de modelos de lenguaje grandes generativos también conlleva el riesgo de alucinación, es decir, la generación de contenido que suena plausible pero que es factualmente incorrecto. El desarrollo adicional de métodos para verificar y garantizar la calidad de los informes generados es, por lo tanto, un importante foco de investigación futura.

7.3 Limitaciones

A pesar del alto rendimiento de ODET en el análisis automatizado de datos de microblogs, existen algunas limitaciones metodológicas y técnicas que requieren mayor investigación y desarrollo.

Un desafío clave es que la detección de noticias falsas y alucinaciones sigue siendo un problema abierto. ODET no implementa mecanismos explícitos para verificar la información extraída ni para evaluar si ciertas declaraciones provienen de fuentes confiables o manipuladas. Si bien el análisis agregado de clústeres ayuda parcialmente a compensar sesgos, aún existe el riesgo de que información falsa o engañosa sea incluida en los informes automatizados. Este problema es particularmente crítico en el uso de modelos de lenguaje generativo, que pueden producir contenido sintácticamente correcto pero factualmente incorrecto. Aunque el sistema minimiza estos efectos mediante el uso de indicaciones estructuradas y pasos de procesamiento especializados, esta sigue siendo una limitación inherente al uso de modelos generativos que no puede eliminarse por completo.

Otro tema importante es la dependencia de datos basados en texto. ODET solo procesa información basada en texto y no considera formatos de datos adicionales, como imágenes, videos o grabaciones de audio. En situaciones de crisis, el contenido visual o auditivo puede ser un valioso complemento a los mensajes de texto, especialmente cuando existen barreras lingüísticas o malentendidos. Por lo tanto, la adición de análisis multimodal (es decir, texto junto con imágenes, video y audio) podría mejorar la fiabilidad de la evaluación de situaciones, proporcionando un panorama más completo de los eventos en tiempo real.

Además, la granularidad de la representación de eventos puede mejorarse. Si bien el proceso de clustering permite la detección y estructuración fiable de eventos más grandes, los incidentes muy pequeños o localmente limitados pueden perderse en clusters grandes o no captarse como eventos independientes. Esto es especialmente cierto para desarrollos sutiles que solo pueden reconocerse como coherentes durante períodos más largos. La metodología actual asegura que los

eventos globales o generalizados se detecten de manera eficiente, pero el ajuste fino para la detección de microeventos sigue siendo un desafío abierto. Un mayor perfeccionamiento de los algoritmos de agrupación y análisis temporal podría mejorar la capacidad del sistema para capturar y diferenciar eventos a menor escala.

7.4 Trabajos Futuros

Si bien ODET ya proporciona una plataforma potente y flexible para el análisis automatizado de datos de microblogs en la gestión de desastres, existen varias áreas donde la investigación futura podría desarrollar y optimizar aún más el sistema. Debido a su arquitectura modular, ODET puede extenderse de manera específica sin comprometer la funcionalidad principal o la escalabilidad. En particular, mejorar la granularidad de eventos, integrar datos multimodales y optimizar los mecanismos de validación para noticias falsas y alucinaciones son puntos de partida importantes para el desarrollo futuro.

Una de las preguntas de investigación abiertas más importantes es la validación de la información recopilada. La detección de noticias falsas y la reducción de alucinaciones en los informes generados son esenciales para la fiabilidad del análisis OSINT. Aunque el sistema actual ya proporciona cierta robustez contra anomalías individuales mediante clustering y votación por mayoría, una capa adicional de validación podría mejorar aún más el valor de los informes generados. Una posible solución sería un análisis exhaustivo de la red de cuentas en redes sociales para detectar la propagación de noticias falsas y mejorar la validación con métodos más sofisticados. Sin embargo, con respecto a los problemas restantes de alucinaciones, podría ser necesaria una generación completamente nueva de modelos de IA para eliminar este riesgo.

Aunque esta metodología incluye filtrado y clasificación de texto, la detección de información errónea sigue siendo un desafío poco investigado. La proliferación de noticias falsas, contenido manipulado y campañas adversarias de desinformación requiere una estrategia más robusta. La investigación futura debería abordar marcos de pruebas adversarias, estrategias de mitigación de sesgos y mecanismos de

evaluación de confiabilidad para mejorar la fiabilidad de los datos en escenarios de crisis.

Otra área importante para la investigación futura es la extensión de ODET al procesamiento de datos multimodales. Hasta ahora, el sistema se ha centrado exclusivamente en información basada en texto, pero en situaciones de crisis, las imágenes, los videos y los datos de audio también juegan un papel crucial. Debido a que ODET es modular, se puede desarrollar un módulo específico para integrar modelos multimodales. Por ejemplo, los modelos de lenguaje multimodales existentes pueden utilizarse para extender la clasificación zero-shot al contenido visual. Esto haría posible analizar automáticamente no solo mensajes de texto sino también imágenes de áreas de desastre o videos relevantes de redes sociales. La investigación futura debería explorar cómo las técnicas de fusión multimodal pueden mejorar la detección de eventos y la conciencia situacional combinando diferentes tipos de datos para un análisis más holístico.

ODET está diseñado para ejecutarse tanto en la nube como en hardware de consumo disponible en escenarios de desastre, lo que lo hace flexible para su implementación según la infraestructura disponible. No está diseñado para entornos de computación móvil o edge computing, sino que opera sobre los recursos informáticos existentes en los sitios afectados. A diferencia de otras herramientas de inteligencia para desastres basadas en OSINT, ODET implementa una arquitectura modular de puntos de acceso que permite un equilibrio eficiente de carga. Esto significa que los módulos con alta demanda computacional, como los grandes modelos de lenguaje (LLMs), pueden ejecutarse en la nube, mientras que los agentes ligeros pueden operar en hardware de consumo, como un MacBook. Esta arquitectura distribuida garantiza escalabilidad, adaptabilidad y uso eficiente de los recursos en distintos entornos.

En conclusión, esta tesis demuestra que los métodos OSINT basados en aprendizaje automático son soluciones viables y eficientes para la evaluación automatizada de situaciones en la gestión de desastres. Mediante la combinación de diferentes técnicas de ML con arquitecturas de sistemas modulares, se puede desarrollar una solución que sea

tanto flexible como potente. La síntesis de los resultados muestra que el análisis OSINT puede acelerarse y mejorarse cualitativamente mediante aprendizaje automático. Los métodos presentados en esta tesis tienen el potencial de cambiar fundamentalmente la forma en que se recopila, analiza e informa la información de crisis, estableciendo un nuevo paradigma para la inteligencia automatizada en la gestión de desastres.

References

[1] M. S. Hackathon, *Leveraging AI for Natural Disaster Management: Takeaways From The Moroccan Earthquake*, 2023. DOI: 10.48550/ARXIV.2311.08999. [Online]. Available: https://arxiv.org/abs/2311.08999 (cit. on pp. 7, 8, 70).

[2] M. Kejriwal, G. Fang, and Y. Zhou, "A Feasibility Study of Open-Source Sentiment Analysis and Text Classification Systems on Disaster-Specific Social Media Data," in *2021 IEEE Symposium Series on Computational Intelligence (SSCI)*, IEEE, Dec. 2021, pp. 1–8. DOI: 10.1109/ssci50451.2021.9660089. [Online]. Available: http://dx.doi.org/10.1109/SSCI50451.2021.9660089 (cit. on pp. 7, 8, 70).

[3] V. Lorini, P. Salamon, and C. Castillo, "SMDRM - Social Media for Disaster Risk Management," Mar. 2021. DOI: 10.5194/egusphere-egu21-15012. [Online]. Available: http://dx.doi.org/10.5194/egusphere-egu21-15012 (cit. on pp. 7–10, 70–72).

[4] T. J. Papadimos *et al.*, "Real-time Alert Framework for Fire Incidents Using Multimodal Event Detection on Social Media Streams," in *International Conference on Information Systems for Crisis Response and Management*, 2022. [Online]. Available: https://ingenious-first-responders.eu/wp-content/uploads/2022/09/ISCRAM_2022.pdf (cit. on pp. 7–11, 70).

[5] M. Abduraimov, "CRISIS MANAGEMENT IN THE DIGITAL AGE," *QO'QON UNIVERSITETI XABARNOMASI*, vol. 11, pp. 16–18, Jun. 2024, ISSN: 2181-1695. DOI: 10.54613/ku.v11i11.943. [Online]. Available: http://dx.doi.org/10.54613/ku.v11i11.943 (cit. on pp. 7–11, 70).

[6] X. Zhou and L. Chen, “Migrating social event recommendation over microblogs,” *Proceedings of the VLDB Endowment*, vol. 15, no. 11, pp. 3213–3225, Jul. 2022, ISSN: 2150-8097. DOI: 10.14778/3551793.3551864. [Online]. Available: http://dx.doi.org/10.14778/3551793.3551864 (cit. on pp. 7–10, 70).

[7] S. G. Arapostathis, “Utilising Twitter for disaster management of fire events: steps towards efficient automation,” *Arabian Journal of Geosciences*, vol. 14, no. 8, Apr. 2021, ISSN: 1866-7538. DOI: 10.1007/s12517-021-06768-2. [Online]. Available: http://dx.doi.org/10.1007/s12517-021-06768-2 (cit. on pp. 7, 72).

[8] B. O. Adelakun, B. O. Antwi, A. Ntiakoh, and A. O. Eziefule, “Leveraging AI for sustainable accounting: Developing models for environmental impact assessment and reporting,” *Finance & Accounting Research Journal*, vol. 6, no. 6, pp. 1017–1048, Jun. 2024, ISSN: 2708-633X. DOI: 10.51594/farj.v6i6.1234. [Online]. Available: http://dx.doi.org/10.51594/farj.v6i6.1234 (cit. on pp. 7–12, 70).

[9] R. H. Yahya, W. Maharani, and R. Wijaya, “Disaster Management Sentiment Analysis Using the BiLSTM Method,” *JURNAL MEDIA INFORMATIKA BUDIDARMA*, vol. 7, no. 1, p. 501, Jan. 2023, ISSN: 2614-5278. DOI: 10.30865/mib.v7i1.5573. [Online]. Available: http://dx.doi.org/10.30865/mib.v7i1.5573 (cit. on pp. 8, 9, 71, 72).

[10] H. Farman Ali Baig, N. Islam, and A. Alim, “Exploring Exploring Machine Learning and Deep Learning Approaches for Disaster Prediction and Management: A Survey of Different Approaches: A Survey of Different Approaches for Disaster prediction and management through tweets,” *Pakistan Journal of Engineering, Technology & Science*, vol. 11, no. 1, pp. 45–73, Jan. 2024, ISSN: 2222-9930. DOI: 10.22555/pjets.v11i1.1004. [Online]. Available: http://dx.doi.org/10.22555/pjets.v11i1.1004 (cit. on pp. 8–12, 72).

[11] M. Aljebreen, B. Alabduallah, M. M. Asiri, A. S. Salama, M. Assiri, and S. S. Ibrahim, “Moth Flame Optimization With Hybrid

Deep Learning Based Sentiment Classification Toward ChatGPT on Twitter," *IEEE Access*, vol. 11, pp. 104 984–104 991, 2023, ISSN: 2169-3536. DOI: 10.1109/access.2023.3315609. [Online]. Available: http://dx.doi.org/10.1109/access.2023.3315609 (cit. on pp. 8–11, 70).

[12] J. Pastor-Galindo, P. Nespoli, F. Gomez Marmol, and G. Martinez Perez, "The Not Yet Exploited Goldmine of OSINT: Opportunities, Open Challenges and Future Trends," *IEEE Access*, vol. 8, pp. 10 282–10 304, 2020, ISSN: 2169-3536. DOI: 10.1109/access.2020.2965257. [Online]. Available: http://dx.doi.org/10.1109/ACCESS.2020.2965257 (cit. on pp. 9, 11, 19, 32, 51, 57).

[13] D. Jung, V. Tran Tuan, D. Quoc Tran, M. Park, and S. Park, "Conceptual Framework of an Intelligent Decision Support System for Smart City Disaster Management," *Applied Sciences*, vol. 10, no. 2, p. 666, Jan. 2020, ISSN: 2076-3417. DOI: 10.3390/app10020666. [Online]. Available: http://dx.doi.org/10.3390/app10020666 (cit. on pp. 9, 11, 50, 51, 57).

[14] A. Burak Can, I. Burak Parlak, and T. Acarman, "A New Method of Automatic Content Analysis in Disaster Management," in *2022 10th International Symposium on Digital Forensics and Security (ISDFS)*, IEEE, Jun. 2022, pp. 1–5. DOI: 10.1109/isdfs55398.2022.9800778. [Online]. Available: http://dx.doi.org/10.1109/isdfs55398.2022.9800778 (cit. on pp. 10, 72).

[15] M. J. P. Canon, L. L. Maceda, and N. M. Flores, "Contextual Understanding of Academic-Related Responses Based on Enhanced Word Embeddings, Clustering, and Community Detection," *Journal of Advances in Information Technology*, vol. 15, no. 6, pp. 693–703, 2024, ISSN: 1798-2340. DOI: 10.12720/jait.15.6.693-703. [Online]. Available: http://dx.doi.org/10.12720/jait.15.6.693-703 (cit. on pp. 11, 12, 73).

[16] K. Schwarz, D. A. Aranda, and M. Hartmann, "Towards Automated Situational Awareness Reporting for Disaster Management—A Case Study," *Sustainability*, vol. 15, no. 10, p. 7968, May 2023, ISSN: 2071-1050. DOI: 10.3390/su15107968. [Online]. Available:

http://dx.doi.org/10.3390/su15107968 (cit. on pp. 13, 18, 67, 68, 101, 115, 129).

[17] K. Schwarz, R. Creutzburg, M. Hartmann, and D. Arias-Aranda, “Harnessing OSINT in Disaster Management: Transforming Microblogging Posts Into Insightful Data Through Text Embeddings,” in *Proceedings of International Conference on Theoretical and Applied Computing*. Springer Nature Singapore, 2025, pp. 195–213, ISBN: 9789819769575. DOI: 10.1007/978-981-97-6957-5_17. [Online]. Available: http://dx.doi.org/10.1007/978-981-97-6957-5_17 (cit. on pp. 13, 18, 101, 116, 130).

[18] K. Schwarz, K. Bollens, D. Arias Aranda, and M. Hartmann, “AI-Enhanced Disaster Management: A Modular OSINT System for Rapid Automated Reporting,” *Applied Sciences*, vol. 14, no. 23, p. 11 165, Nov. 2024, ISSN: 2076-3417. DOI: 10.3390/app142311165. [Online]. Available: http://dx.doi.org/10.3390/app142311165 (cit. on pp. 13, 19, 102, 116, 130).

[19] A. F. Al-Madhari and A. Z. Keller, “Review Of Disaster Definitions,” *Prehospital and Disaster Medicine*, vol. 12, no. 1, pp. 17–21, Mar. 1997, ISSN: 1945-1938. DOI: 10.1017/s1049023x0003716x. [Online]. Available: http://dx.doi.org/10.1017/s1049023x0003716x (cit. on pp. 28, 70).

[20] J. Zibulewsky, “Defining Disaster: The Emergency Department Perspective,” *Baylor University Medical Center Proceedings*, vol. 14, no. 2, pp. 144–149, Apr. 2001, ISSN: 1525-3252. DOI: 10.1080/08998280.2001.11927751. [Online]. Available: http://dx.doi.org/10.1080/08998280.2001.11927751 (cit. on p. 28).

[21] I. Hanna Sawalha, “Behavioural response patterns: an investigation of the early stages of major incidents,” *foresight*, vol. 20, no. 4, pp. 337–352, Aug. 2018, ISSN: 1463-6689. DOI: 10.1108/fs-12-2017-0073. [Online]. Available: http://dx.doi.org/10.1108/fs-12-2017-0073 (cit. on p. 28).

[22] J. Malilay *et al.*, “The Role of Applied Epidemiology Methods in the Disaster Management Cycle,” *American Journal of Public Health*,

vol. 104, no. 11, pp. 2092–2102, Nov. 2014, ISSN: 1541-0048. DOI: 10.2105/ajph.2014.302010. [Online]. Available: http://dx.doi.org/10.2105/AJPH.2014.302010 (cit. on p. 28).

[23] A. I. Anderson, D. Compton, and T. Mason, "Managing in a Dangerous World—The National Incident Management System," *Engineering Management Journal*, vol. 16, no. 4, pp. 3–9, Dec. 2004, ISSN: 2377-0643. DOI: 10.1080/10429247.2004.11415260. [Online]. Available: http://dx.doi.org/10.1080/10429247.2004.11415260 (cit. on p. 30).

[24] U. P. Gujral *et al.*, "Preparedness cycle to address transitions in diabetes care during the COVID-19 pandemic and future outbreaks," *BMJ Open Diabetes Research & Care*, vol. 8, no. 1, e001520, Jul. 2020, ISSN: 2052-4897. DOI: 10.1136/bmjdrc-2020-001520. [Online]. Available: http://dx.doi.org/10.1136/bmjdrc-2020-001520 (cit. on pp. 30, 31).

[25] A. Ramsbottom, E. O'Brien, L. Ciotti, and J. Takacs, "Enablers and Barriers to Community Engagement in Public Health Emergency Preparedness: A Literature Review," *Journal of Community Health*, vol. 43, no. 2, pp. 412–420, Aug. 2017, ISSN: 1573-3610. DOI: 10.1007/s10900-017-0415-7. [Online]. Available: http://dx.doi.org/10.1007/s10900-017-0415-7 (cit. on pp. 30, 31).

[26] D. Arias Aranda, L. M. Molina Fernandez, and V. Stantchev, "INTEGRATION OF INTERNET OF THINGS AND BLOCKCHAIN TO INCREASE HUMANITARIAN AID SUPPLY CHAINS PERFORMANCE," *DYNA*, vol. 96, no. 6, pp. 653–658, Nov. 2021, ISSN: 1989-1490. DOI: 10.6036/10067. [Online]. Available: http://dx.doi.org/10.6036/10067 (cit. on p. 31).

[27] K. Rudra, N. Ganguly, P. Goyal, and S. Ghosh, "Extracting and Summarizing Situational Information from the Twitter Social Media during Disasters," *ACM Transactions on the Web*, vol. 12, no. 3, pp. 1–35, Jul. 2018, ISSN: 1559-114X. DOI: 10.1145/3178541. [Online]. Available: http://dx.doi.org/10.1145/3178541 (cit. on pp. 31, 37).

[28] R. Lamsal and T. V. V. Kumar, “ Twitter-Based Disaster Response Using Recurrent Nets,” in *Research Anthology on Managing Crisis and Risk Communications*. IGI Global, Jul. 2022, pp. 613–632, ISBN: 9781668471463. DOI: 10.4018/978-1-6684-7145-6.ch031. [Online]. Available: http://dx.doi.org/10.4018/978-1-6684-7145-6.ch031 (cit. on p. 31).

[29] W. Mukhtiar, W. Rizwan, A. Habib, Y. S. Afridi, L. Hasan, and K. Ahmad, *Relevance Classification of Flood-related Twitter Posts via Multiple Transformers*, 2023. DOI: 10.48550/ARXIV.2301.00320. [Online]. Available: https://arxiv.org/abs/2301.00320 (cit. on p. 31).

[30] M. Karimiziarani, K. Jafarzadegan, P. Abbaszadeh, W. Shao, and H. Moradkhani, “Hazard risk awareness and disaster management: Extracting the information content of Twitter data,” *Sustainable Cities and Society*, vol. 77, p. 103 577, Feb. 2022, ISSN: 2210-6707. DOI: 10.1016/j.scs.2021.103577. [Online]. Available: http://dx.doi.org/10.1016/j.scs.2021.103577 (cit. on p. 31).

[31] M. Karimiziarani and H. Moradkhani, “Social Response and Disaster Management: Insights from Twitter Data Assimilation on Hurricane Ian,” *SSRN Electronic Journal*, 2022, ISSN: 1556-5068. DOI: 10.2139/ssrn.4292734. [Online]. Available: http://dx.doi.org/10.2139/ssrn.4292734 (cit. on p. 31).

[32] R. Lamsal, A. Harwood, and M. R. Read, “Socially Enhanced Situation Awareness from Microblogs Using Artificial Intelligence: A Survey,” *ACM Computing Surveys*, vol. 55, no. 4, pp. 1–38, Nov. 2022, ISSN: 1557-7341. DOI: 10.1145/3524498. [Online]. Available: http://dx.doi.org/10.1145/3524498 (cit. on p. 31).

[33] B. Takahashi, E. C. Tandoc, and C. Carmichael, “Communicating on Twitter during a disaster: An analysis of tweets during Typhoon Haiyan in the Philippines,” *Computers in Human Behavior*, vol. 50, pp. 392–398, Sep. 2015, ISSN: 0747-5632. DOI: 10.1016/j.chb.2015.04.020. [Online]. Available: http://dx.doi.org/10.1016/j.chb.2015.04.020 (cit. on p. 31).

[34] J. Phengsuwan *et al.*, “Use of Social Media Data in Disaster Management: A Survey,” *Future Internet*, vol. 13, no. 2, p. 46, Feb. 2021, ISSN: 1999-5903. DOI: 10.3390/fi13020046. [Online]. Available: http://dx.doi.org/10.3390/fi13020046 (cit. on pp. 32, 53).

[35] T. Sakaki, M. Okazaki, and Y. Matsuo, “Earthquake shakes Twitter users: real-time event detection by social sensors,” in *Proceedings of the 19th International Conference on World Wide Web*, ser. WWW ’10, ACM, Apr. 2010. DOI: 10.1145/1772690.1772777. [Online]. Available: http://dx.doi.org/10.1145/1772690.1772777 (cit. on p. 32).

[36] S. Shan, F. Zhao, Y. Wei, and M. Liu, “Disaster management 2.0: A real-time disaster damage assessment model based on mobile social media data—A case study of Weibo (Chinese Twitter),” *Safety Science*, vol. 115, pp. 393–413, Jun. 2019, ISSN: 0925-7535. DOI: 10.1016/j.ssci.2019.02.029. [Online]. Available: http://dx.doi.org/10.1016/j.ssci.2019.02.029 (cit. on p. 32).

[37] S. A. H. B. S. Muzamil *et al.*, “Proposed Framework for the Flood Disaster Management Cycle in Malaysia,” *Sustainability*, vol. 14, no. 7, p. 4088, Mar. 2022, ISSN: 2071-1050. DOI: 10.3390/su14074088. [Online]. Available: http://dx.doi.org/10.3390/su14074088 (cit. on p. 32).

[38] M. Hasan, M. A. Orgun, and R. Schwitter, “A survey on real-time event detection from the Twitter data stream,” *Journal of Information Science*, vol. 44, no. 4, pp. 443–463, Mar. 2017, ISSN: 1741-6485. DOI: 10.1177/0165551517698564. [Online]. Available: http://dx.doi.org/10.1177/0165551517698564 (cit. on pp. 32, 35).

[39] X. Yu, C. Li, and G. G. Yen, “A knee-guided differential evolution algorithm for unmanned aerial vehicle path planning in disaster management,” *Applied Soft Computing*, vol. 98, p. 106 857, Jan. 2021, ISSN: 1568-4946. DOI: 10.1016/j.asoc.2020.106857. [Online]. Available: http://dx.doi.org/10.1016/j.asoc.2020.106857 (cit. on p. 32).

[40] P. Rabiei and D. Arias-Aranda, “Introducing a novel multi-objective optimization model for vehicle routing and relief supply distribution in post-disaster phase: combining fuzzy inference systems with NSGA-II and NRGA,” in *2021 6th International Conference on Transportation Information and Safety (ICTIS)*, IEEE, Oct. 2021, pp. 1226–1243. DOI: 10.1109/ictis54573.2021.9798556. [Online]. Available: http://dx.doi.org/10.1109/ICTIS54573.2021.9798556 (cit. on p. 32).

[41] P. Rabiei, D. Arias-Aranda, and V. Stantchev, “Introducing a novel multi-objective optimization model for volunteer assignment in the post-disaster phase: Combining fuzzy inference systems with NSGA-II and NRGA,” *Expert Systems with Applications*, vol. 226, p. 120 142, Sep. 2023, ISSN: 0957-4174. DOI: 10.1016/j.eswa.2023.120142. [Online]. Available: http://dx.doi.org/10.1016/j.eswa.2023.120142 (cit. on p. 32).

[42] P. K. Garg, R. Chakraborty, and S. K. Dandapat, “OntoDSumm: Ontology-Based Tweet Summarization for Disaster Events,” *IEEE Transactions on Computational Social Systems*, vol. 11, no. 2, pp. 2724–2739, Apr. 2024, ISSN: 2373-7476. DOI: 10.1109/tcss.2023.3266025. [Online]. Available: http://dx.doi.org/10.1109/tcss.2023.3266025 (cit. on p. 34).

[43] P. K. Garg, R. Chakraborty, and S. K. Dandapat, *EnDSUM: Entropy and Diversity based Disaster Tweet Summarization*, 2022. DOI: 10.48550/ARXIV.2203.01188. [Online]. Available: https://arxiv.org/abs/2203.01188 (cit. on p. 34).

[44] K. Wu, L. Li, J. Li, and T. Li, “Ontology-enriched multi-document summarization in disaster management using submodular function,” *Information Sciences*, vol. 224, pp. 118–129, Mar. 2013, ISSN: 0020-0255. DOI: 10.1016/j.ins.2012.10.019. [Online]. Available: http://dx.doi.org/10.1016/j.ins.2012.10.019 (cit. on p. 34).

[45] M. Karimiziarani and H. Moradkhani, “Social response and Disaster management: Insights from Twitter data Assimilation on Hurricane Ian,” *International Journal of Disaster Risk Reduction*,

vol. 95, p. 103 865, Sep. 2023, ISSN: 2212-4209. DOI: 10.1016/j.ijdrr.2023.103865. [Online]. Available: http://dx.doi.org/10.1016/j.ijdrr.2023.103865 (cit. on pp. 34, 74).

[46] S. Banerjee, S. Mukherjee, S. Bandyopadhyay, and P. Pakray, "An extract-then-abstract based method to generate disaster-news headlines using a DNN extractor followed by a transformer abstractor," *Information Processing & Management*, vol. 60, no. 3, p. 103 291, May 2023, ISSN: 0306-4573. DOI: 10.1016/j.ipm.2023.103291. [Online]. Available: http://dx.doi.org/10.1016/j.ipm.2023.103291 (cit. on p. 34).

[47] N. Saini, S. Saha, and P. Bhattacharyya, "Microblog summarization using self-adaptive multi-objective binary differential evolution," *Applied Intelligence*, vol. 52, no. 2, pp. 1686–1702, May 2021, ISSN: 1573-7497. DOI: 10.1007/s10489-020-02178-1. [Online]. Available: http://dx.doi.org/10.1007/s10489-020-02178-1 (cit. on p. 34).

[48] F. Vitiugin and C. Castillo, "Cross-Lingual Query-Based Summarization of Crisis-Related Social Media: An Abstractive Approach Using Transformers," in *Proceedings of the 33rd ACM Conference on Hypertext and Social Media*, ser. HT '22, ACM, Jun. 2022, pp. 21–31. DOI: 10.1145/3511095.3531279. [Online]. Available: http://dx.doi.org/10.1145/3511095.3531279 (cit. on p. 34).

[49] T. H. Nguyen and K. Rudra, "Towards an Interpretable Approach to Classify and Summarize Crisis Events from Microblogs," in *Proceedings of the ACM Web Conference 2022*, ser. WWW '22, ACM, Apr. 2022. DOI: 10.1145/3485447.3512259. [Online]. Available: http://dx.doi.org/10.1145/3485447.3512259 (cit. on p. 34).

[50] R. Mukherjee *et al.*, "MTLTS: A Multi-Task Framework To Obtain Trustworthy Summaries From Crisis-Related Microblogs," in *Proceedings of the Fifteenth ACM International Conference on Web Search and Data Mining*, ser. WSDM '22, ACM, Feb. 2022, pp. 755–763. DOI: 10.1145/3488560.3498536. [Online]. Available: http://dx.doi.org/10.1145/3488560.3498536 (cit. on p. 34).

[51] K. Rudra, S. Ghosh, N. Ganguly, P. Goyal, and S. Ghosh, "Extracting Situational Information from Microblogs during Disaster Events: a Classification-Summarization Approach," in *Proceedings of the 24th ACM International on Conference on Information and Knowledge Management*, ser. CIKM'15, ACM, Oct. 2015. DOI: 10.1145/2806416.2806485. [Online]. Available: http://dx.doi.org/10.1145/2806416.2806485 (cit. on p. 35).

[52] S. Unankard and W. Nadee, "Sub-Events Tracking from Social Network based on the Relationships between Topics," in *2020 Joint International Conference on Digital Arts, Media and Technology with ECTI Northern Section Conference on Electrical, Electronics, Computer and Telecommunications Engineering (ECTI DAMT & NCON)*, IEEE, Mar. 2020, pp. 1–6. DOI: 10.1109/ectidamtncon48261.2020.9090732. [Online]. Available: http://dx.doi.org/10.1109/ectidamtncon48261.2020.9090732 (cit. on pp. 35, 36).

[53] Q. Li, Y. Chao, D. Li, Y. Lu, and C. Zhang, "Event Detection from Social Media Stream: Methods, Datasets and Opportunities," in *2022 IEEE International Conference on Big Data (Big Data)*, IEEE, Dec. 2022, pp. 3509–3516. DOI: 10.1109/bigdata55660.2022.10020411. [Online]. Available: http://dx.doi.org/10.1109/bigdata55660.2022.10020411 (cit. on p. 35).

[54] M. Osborne *et al.*, "Real-Time Detection, Tracking, and Monitoring of Automatically Discovered Events in Social Media," in *Proceedings of 52nd Annual Meeting of the Association for Computational Linguistics: System Demonstrations*, Association for Computational Linguistics, 2014, pp. 37–42. DOI: 10.3115/v1/p14-5007. [Online]. Available: http://dx.doi.org/10.3115/v1/p14-5007 (cit. on p. 36).

[55] S. Petrović, M. Osborne, and V. Lavrenko, "Streaming First Story Detection with application to Twitter," in *Human Language Technologies: The 2010 Annual Conference of the North American Chapter of the Association for Computational Linguistics*, R. Kaplan, J. Burstein, M. Harper, and G. Penn, Eds., Los Angeles,

California: Association for Computational Linguistics, Jun. 2010, pp. 181–189. [Online]. Available: https://aclanthology.org/N10-1021/ (cit. on p. 36).

[56] D. Rudrapal, A. Das, and B. Bhattacharya, "A Survey on Automatic Twitter Event Summarization," *J. Inf. Process. Syst.*, vol. 14, pp. 79–100, 2018. DOI: 10.3745/JIPS.02.0079. [Online]. Available: https://api.semanticscholar.org/CorpusID:4404250 (cit. on p. 36).

[57] M. Honnibal and I. Montani, "spaCy 2: Natural language understanding with Bloom embeddings, convolutional neural networks and incremental parsing," 2017 (cit. on pp. 37, 39, 41).

[58] R. Rehurek and P. Sojka, "Gensim–Python Framework for Vector Space Modelling," *NLP Centre, Faculty of Informatics, Masaryk University, Brno, Czech Republic*, vol. 3, no. 2, 2011 (cit. on pp. 37, 41).

[59] W. Wagner, "Steven Bird, Ewan Klein and Edward Loper: Natural Language Processing with Python, Analyzing Text with the Natural Language Toolkit," *Language Resources and Evaluation*, vol. 44, pp. 421–424, 2010. [Online]. Available: https://api.semanticscholar.org/CorpusID:207104425 (cit. on pp. 37, 41).

[60] G. Navarro, "A guided tour to approximate string matching," *ACM Computing Surveys*, vol. 33, no. 1, pp. 31–88, Mar. 2001, ISSN: 1557-7341. DOI: 10.1145/375360.375365. [Online]. Available: http://dx.doi.org/10.1145/375360.375365 (cit. on p. 40).

[61] A. Rajaraman and J. D. Ullman, "Data Mining," in *Mining of Massive Datasets*. Cambridge University Press, Oct. 2011, pp. 1–17. DOI: 10.1017/cbo9781139058452.002. [Online]. Available: http://dx.doi.org/10.1017/cbo9781139058452.002 (cit. on p. 40).

[62] A. Olteanu, C. Castillo, F. Diaz, and S. Vieweg, "CrisisLex: A Lexicon for Collecting and Filtering Microblogged Communications in Crises," *Proceedings of the International AAAI Conference on Web and Social Media*, vol. 8, no. 1, pp. 376–385, May 2014, ISSN: 2162-3449. DOI: 10.1609/icwsm.v8i1.14538. [Online].

Available: http://dx.doi.org/10.1609/icwsm.v8i1.14538 (cit. on p. 40).

[63] G. S. Manku, A. Jain, and A. Das Sarma, "Detecting near-duplicates for web crawling," in *Proceedings of the 16th international conference on World Wide Web*, ser. WWW'07, ACM, May 2007, pp. 141–150. DOI: 10.1145/1242572.1242592. [Online]. Available: http://dx.doi.org/10.1145/1242572.1242592 (cit. on p. 40).

[64] R. McCreadie, C. Macdonald, and I. Ounis, "Insights on the Horizon of News Search," 2010. [Online]. Available: http://terrierteam.dcs.gla.ac.uk/publications/richardmSSM2010.pdf (cit. on p. 41).

[65] C. Li, A. Sun, and A. Datta, "Twevent: segment-based event detection from tweets," in *Proceedings of the 21st ACM international conference on Information and knowledge management*, ser. CIKM'12, ACM, Oct. 2012. DOI: 10.1145/2396761.2396785. [Online]. Available: http://dx.doi.org/10.1145/2396761.2396785 (cit. on p. 41).

[66] G. Stilo and P. Velardi, "Efficient temporal mining of micro-blog texts and its application to event discovery," *Data Mining and Knowledge Discovery*, vol. 30, no. 2, pp. 372–402, May 2015, ISSN: 1573-756X. DOI: 10.1007/s10618-015-0412-3. [Online]. Available: http://dx.doi.org/10.1007/s10618-015-0412-3 (cit. on p. 41).

[67] M. Osborne, S. Petrovic, R. McCreadie, C. Macdonald, I. Ounis, and S. Petrovic, "Bieber no more: First Story Detection using Twitter and Wikipedia," 2012. [Online]. Available: https://www.dcs.gla.ac.uk/~craigm/publications/osborneTAIA2012.pdf (cit. on p. 41).

[68] *Timeline of the 2022 Russian Invasion of Ukraine Phase 3—Wikipedia*, Apr. 2023. [Online]. Available: https://en.wikipedia.org/wiki/Timeline_of_the_2022_Russian_invasion_of_Ukraine:_phase_3 (cit. on pp. 41, 44).

[69] BwandoWando, *Ukraine Conflict Twitter Dataset*, 2023. DOI: 10.34740/KAGGLE/DSV/4898954. [Online]. Available: https://www.kaggle.com/dsv/4898954 (cit. on pp. 42, 67).

[70] M. Bouafia and K. Zahari R, "Naturalistic Decision-Making in Natural Disasters: An Overview," *Journal of Geography & Natural Disasters*, vol. 07, no. 01, 2017, ISSN: 2167-0587. DOI: 10.4172/2167-0587.1000186. [Online]. Available: http://dx.doi.org/10.4172/2167-0587.1000186 (cit. on p. 49).

[71] Y.-B. Lin, Y.-P. Lin, D.-P. Deng, and K.-W. Chen, "Integrating Remote Sensing Data with Directional Two-Dimensional Wavelet Analysis and Open Geospatial Techniques for Efficient Disaster Monitoring and Management," *Sensors*, vol. 8, no. 2, pp. 1070–1089, Feb. 2008, ISSN: 1424-8220. DOI: 10.3390/s8021070. [Online]. Available: http://dx.doi.org/10.3390/s8021070 (cit. on p. 49).

[72] W. Yuan, D. Guan, E.-N. Huh, and S. Lee, "Harness Human Sensor Networks for Situational Awareness in Disaster Reliefs: A Survey," *IETE Technical Review*, vol. 30, no. 3, p. 240, 2013, ISSN: 0256-4602. DOI: 10.4103/0256-4602.113522. [Online]. Available: http://dx.doi.org/10.4103/0256-4602.113522 (cit. on p. 49).

[73] S. A. Shah, D. Z. Seker, S. Hameed, and D. Draheim, "The Rising Role of Big Data Analytics and IoT in Disaster Management: Recent Advances, Taxonomy and Prospects," *IEEE Access*, vol. 7, pp. 54 595–54 614, 2019, ISSN: 2169-3536. DOI: 10.1109/access.2019.2913340. [Online]. Available: http://dx.doi.org/10.1109/access.2019.2913340 (cit. on pp. 49, 58).

[74] S. P. Cumbane and G. Gidófalvi, "Review of Big Data and Processing Frameworks for Disaster Response Applications," *ISPRS International Journal of Geo-Information*, vol. 8, no. 9, p. 387, Sep. 2019, ISSN: 2220-9964. DOI: 10.3390/ijgi8090387. [Online]. Available: http://dx.doi.org/10.3390/ijgi8090387 (cit. on p. 49).

[75] M. Imran, C. Castillo, F. Diaz, and S. Vieweg, “Processing Social Media Messages in Mass Emergency: A Survey,” *ACM Computing Surveys*, vol. 47, no. 4, pp. 1–38, Jun. 2015, ISSN: 1557-7341. DOI: 10.1145/2771588. [Online]. Available: http://dx.doi.org/10.1145/2771588 (cit. on p. 50).

[76] Y. Tim, S. L. Pan, P. Ractham, and L. Kaewkitipong, “Digitally enabled disaster response: the emergence of social media as boundary objects in a flooding disaster,” *Information Systems Journal*, vol. 27, no. 2, pp. 197–232, Oct. 2016, ISSN: 1365-2575. DOI: 10.1111/isj.12114. [Online]. Available: http://dx.doi.org/10.1111/isj.12114 (cit. on p. 50).

[77] T. Li *et al.*, “Data-Driven Techniques in Disaster Information Management,” *ACM Computing Surveys*, vol. 50, no. 1, pp. 1–45, Mar. 2017, ISSN: 1557-7341. DOI: 10.1145/3017678. [Online]. Available: http://dx.doi.org/10.1145/3017678 (cit. on p. 50).

[78] J. Levy and B. Pandey, “INTELLIGENT DISASTER MANAGEMENT AND GREEN COMPUTING IN THE BIG DATA ERA: POLICIES, PERSPECTIVES AND PROSPECTS,” *3C Tecnología_Glosas de innovación aplicadas a la pyme*, pp. 15–27, Nov. 2019, ISSN: 2254-4143. DOI: 10.17993/3ctecno.2019.specialissue3.15-27. [Online]. Available: http://dx.doi.org/10.17993/3ctecno.2019.specialissue3.15-27 (cit. on pp. 50, 58).

[79] S. K. Abid *et al.*, “Toward an Integrated Disaster Management Approach: How Artificial Intelligence Can Boost Disaster Management,” *Sustainability*, vol. 13, no. 22, p. 12 560, Nov. 2021, ISSN: 2071-1050. DOI: 10.3390/su132212560. [Online]. Available: http://dx.doi.org/10.3390/su132212560 (cit. on pp. 50, 52, 54).

[80] D. Salluri, K. Bade, and G. Madala, “Object Detection Using Convolutional Neural Networks for Natural Disaster Recovery,” *International Journal of Safety and Security Engineering*, vol. 10, no. 2, pp. 285–291, Apr. 2020, ISSN: 2041-9031. DOI: 10.18280/ijsse.100217. [Online]. Available: http://dx.doi.org/10.18280/ijsse.100217 (cit. on p. 50).

[81] Á. Lorca, M. Çelik, Ö. Ergun, and P. Keskinocak, "An Optimization-Based Decision-Support Tool for Post-Disaster Debris Operations," *Production and Operations Management*, vol. 26, no. 6, pp. 1076–1091, Jun. 2017, ISSN: 1937-5956. DOI: 10.1111/poms.12643. [Online]. Available: http://dx.doi.org/10.1111/poms.12643 (cit. on p. 51).

[82] H. Mahfooz Ul Haque, K. Saleem, and A. Salman Khan, "Modeling belief-desire-intention reasoning agents for situation-aware formalisms," *Concurrency and Computation: Practice and Experience*, vol. 35, no. 15, May 2021, ISSN: 1532-0634. DOI: 10.1002/cpe.6417. [Online]. Available: http://dx.doi.org/10.1002/cpe.6417 (cit. on p. 51).

[83] M. R. Faisal, I. Budiman, F. Abadi, M. Haekal, M. K. Delimayanti, and D. T. Nugrahadi, "Using Social Media Data to Monitor Natural Disaster: A Multi Dimension Convolutional Neural Network Approach with Word Embedding," *Jurnal RESTI (Rekayasa Sistem dan Teknologi Informasi)*, vol. 6, no. 6, pp. 1037–1046, Dec. 2022, ISSN: 2580-0760. DOI: 10.29207/resti.v6i6.4525. [Online]. Available: http://dx.doi.org/10.29207/resti.v6i6.4525 (cit. on p. 52).

[84] E. A. Obonyo and L. A. Ouma, "Enhancing the resilience of low-income housing using emerging digital technologies," *IOP Conference Series: Earth and Environmental Science*, vol. 1101, no. 9, p. 092 013, Nov. 2022, ISSN: 1755-1315. DOI: 10.1088/1755-1315/1101/9/092013. [Online]. Available: http://dx.doi.org/10.1088/1755-1315/1101/9/092013 (cit. on p. 52).

[85] L. X. Li, "Involvement of Social Media in Disaster Management During the Wenchuan and Ya'an Earthquakes," *Asian Journal for Public Opinion Research*, vol. 1, no. 4, pp. 249–267, Aug. 2014, ISSN: 2288-6168. DOI: 10.15206/ajpor.2014.1.4.249. [Online]. Available: http://dx.doi.org/10.15206/ajpor.2014.1.4.249 (cit. on pp. 52, 53).

[86] Z. S. Dong, L. Meng, L. Christenson, and L. Fulton, "Social media information sharing for natural disaster response," *Natural*

Hazards, vol. 107, no. 3, pp. 2077–2104, Feb. 2021, ISSN: 1573-0840. DOI: 10.1007/s11069-021-04528-9. [Online]. Available: http://dx.doi.org/10.1007/s11069-021-04528-9 (cit. on pp. 52, 53).

[87] T. Schempp, H. Zhang, A. Schmidt, M. Hong, and R. Akerkar, "A framework to integrate social media and authoritative data for disaster relief detection and distribution optimization," *International Journal of Disaster Risk Reduction*, vol. 39, p. 101 143, Oct. 2019, ISSN: 2212-4209. DOI: 10.1016/j.ijdrr.2019.101143. [Online]. Available: http://dx.doi.org/10.1016/j.ijdrr.2019.101143 (cit. on p. 53).

[88] S. A. Shah, D. Z. Şeker, and H. Demirel, "A Framework for Enhancing Real-time Social Media Data to Improve Disaster Management Process," *Proceedings of the ICA*, vol. 1, pp. 1–5, May 2018, ISSN: 2570-2092. DOI: 10.5194/ica-proc-1-101-2018. [Online]. Available: http://dx.doi.org/10.5194/ica-proc-1-101-2018 (cit. on p. 53).

[89] H. To, S. Agrawal, S. H. Kim, and C. Shahabi, "On Identifying Disaster-Related Tweets: Matching-Based or Learning-Based?" In *2017 IEEE Third International Conference on Multimedia Big Data (BigMM)*, IEEE, Apr. 2017, pp. 330–337. DOI: 10.1109/bigmm.2017.82. [Online]. Available: http://dx.doi.org/10.1109/bigmm.2017.82 (cit. on p. 53).

[90] A. Immonen, P. Paakkonen, and E. Ovaska, "Evaluating the Quality of Social Media Data in Big Data Architecture," *IEEE Access*, vol. 3, pp. 2028–2043, 2015, ISSN: 2169-3536. DOI: 10.1109/access.2015.2490723. [Online]. Available: http://dx.doi.org/10.1109/access.2015.2490723 (cit. on p. 53).

[91] T. Ramakrishnan, L. Ngamassi, and S. Rahman, "Examining the Factors that Influence the Use of Social Media for Disaster Management by Underserved Communities," *International Journal of Disaster Risk Science*, vol. 13, no. 1, pp. 52–65, Feb. 2022, ISSN: 2192-6395. DOI: 10.1007/s13753-022-00399-1. [Online].

Available: http://dx.doi.org/10.1007/s13753-022-00399-1 (cit. on p. 53).

[92] H. S. Munawar, F. Ullah, S. Qayyum, S. I. Khan, and M. Mojtahedi, "UAVs in Disaster Management: Application of Integrated Aerial Imagery and Convolutional Neural Network for Flood Detection," *Sustainability*, vol. 13, no. 14, p. 7547, Jul. 2021, ISSN: 2071-1050. DOI: 10.3390/su13147547. [Online]. Available: http://dx.doi.org/10.3390/su13147547 (cit. on p. 54).

[93] M. Hirano, M. Suzuki, and H. Sakaji, *llm-japanese-dataset v0: Construction of Japanese Chat Dataset for Large Language Models and its Methodology*, 2023. DOI: 10.48550/ARXIV.2305.12720. [Online]. Available: https://arxiv.org/abs/2305.12720 (cit. on p. 54).

[94] M. Yu, C. Yang, and Y. Li, "Big Data in Natural Disaster Management: A Review," *Geosciences*, vol. 8, no. 5, p. 165, May 2018, ISSN: 2076-3263. DOI: 10.3390/geosciences8050165. [Online]. Available: http://dx.doi.org/10.3390/geosciences8050165 (cit. on p. 57).

[95] D. Velev and P. Zlateva, "CHALLENGES OF ARTIFICIAL INTELLIGENCE APPLICATION FOR DISASTER RISK MANAGEMENT," *The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, vol. XLVIII-M-1–2023, pp. 387–394, Apr. 2023, ISSN: 2194-9034. DOI: 10.5194/isprs-archives-xlviii-m-1-2023-387-2023. [Online]. Available: http://dx.doi.org/10.5194/isprs-archives-xlviii-m-1-2023-387-2023 (cit. on p. 58).

[96] S. Vieweg, C. Castillo, and M. Imran, "Integrating Social Media Communications into the Rapid Assessment of Sudden Onset Disasters," in *Social Informatics*. Springer International Publishing, 2014, pp. 444–461, ISBN: 9783319137346. DOI: 10.1007/978-3-319-13734-6_32. [Online]. Available: http://dx.doi.org/10.1007/978-3-319-13734-6_32 (cit. on p. 58).

[97] U. Nazir, N. Sulaiman, and S. K. Abid, "Rise of Digital Humanitarian Network (DHN) in Southeast Asia: Social Media Insights

for Crisis Mapping in Disaster Risk Reduction (DRR),” *International Journal of Safety and Security Engineering*, vol. 11, no. 5, pp. 573–583, Oct. 2021, ISSN: 2041-904X. DOI: 10.18280/ijsse.110509. [Online]. Available: http://dx.doi.org/10.18280/ijsse.110509 (cit. on p. 58).

[98] A. N. L. Kusumo, D. Reckien, and J. Verplanke, “Volunteered Geographic Information (VGI) for the Spatial Planning of Flood Evacuation Shelters in Jakarta, Indonesia,” in *GIS in Sustainable Urban Planning and Management*. CRC Press, Dec. 2018, pp. 307–324, ISBN: 9781315146638. DOI: 10.1201/9781315146638-18. [Online]. Available: http://dx.doi.org/10.1201/9781315146638-18 (cit. on p. 59).

[99] M. M. Kobiruzzaman, “Role of Social Media in Disaster Management in Bangladesh Towards the COVID-19 Pandemic: A Critical Review and Directions,” *International Journal of Education and Knowledge Management*, Mar. 2020, ISSN: 2616-5198. DOI: 10.37227/ijekm-2021-03-39. [Online]. Available: http://dx.doi.org/10.37227/ijekm-2021-03-39 (cit. on p. 59).

[100] V. M. Cvetković and A. Nikolić, “The Role of Social Media in the Process of Informing the Public About Disaster Risks,” Jan. 2023. DOI: 10.20944/preprints202301.0430.v1. [Online]. Available: http://dx.doi.org/10.20944/preprints202301.0430.v1 (cit. on p. 59).

[101] A. Ahmed and S. Sinnappan, “The role of Social media during Queensland floods: An Empirical Investigation on the Existence of Multiple Communities of Practice (MCoPs),” *Pacific Asia Journal of the Association for Information Systems*, pp. 1–22, 2013, ISSN: 1943-7544. DOI: 10.17705/1pais.05201. [Online]. Available: http://dx.doi.org/10.17705/1pais.05201 (cit. on p. 59).

[102] D. Kozlowski *et al.*, “A three-level classification of French tweets in ecological crises,” *Information Processing & Management*, vol. 57, no. 5, p. 102 284, Sep. 2020, ISSN: 0306-4573. DOI: 10.1016/j.ipm.2020.102284. [Online]. Available: http://dx.doi.org/10.1016/j.ipm.2020.102284 (cit. on p. 59).

[103] M. Zhang, G. Geng, and J. Chen, “Semi-Supervised Bidirectional Long Short-Term Memory and Conditional Random Fields Model for Named-Entity Recognition Using Embeddings from Language Models Representations,” *Entropy*, vol. 22, no. 2, p. 252, Feb. 2020, ISSN: 1099-4300. DOI: 10.3390/e22020252. [Online]. Available: http://dx.doi.org/10.3390/e22020252 (cit. on p. 60).

[104] V. Silaa, F. Masui, and M. Ptaszynski, “A Method of Supplementing Reviews to Less-Known Tourist Spots Using Geotagged Tweets,” *Applied Sciences*, vol. 12, no. 5, p. 2321, Feb. 2022, ISSN: 2076-3417. DOI: 10.3390/app12052321. [Online]. Available: http://dx.doi.org/10.3390/app12052321 (cit. on p. 60).

[105] M. Sadanandam and B. V. P. Kumar, “Hate Speech and Reality Check Analysis of Disaster Tweets using BERT Deep Learning Model,” *Mathematical Statistician and Engineering Applications*, vol. 71, no. 2, pp. 673–683, Mar. 2022, ISSN: 2094-0343. DOI: 10.17762/msea.v71i2.2251. [Online]. Available: http://dx.doi.org/10.17762/msea.v71i2.2251 (cit. on p. 60).

[106] C. Prudhomme, C. Cruz, and F. Boochs, “Semantic and Logic Modeling of Disaster Simulation for Multi-agent Systems,” *International Journal of Modeling and Optimization*, pp. 198–204, Aug. 2019, ISSN: 2010-3697. DOI: 10.7763/ijmo.2019.v9.709. [Online]. Available: http://dx.doi.org/10.7763/ijmo.2019.v9.709 (cit. on p. 60).

[107] S. Cerna, C. Guyeux, and D. Laiymani, “The usefulness of NLP techniques for predicting peaks in firefighter interventions due to rare events,” *Neural Computing and Applications*, vol. 34, no. 12, pp. 10 117–10 132, Feb. 2022, ISSN: 1433-3058. DOI: 10.1007/s00521-022-06996-x. [Online]. Available: http://dx.doi.org/10.1007/s00521-022-06996-x (cit. on p. 60).

[108] S. Ji, S. Pan, E. Cambria, P. Marttinen, and P. S. Yu, “A Survey on Knowledge Graphs: Representation, Acquisition, and Applications,” *IEEE Transactions on Neural Networks and Learning Systems*, vol. 33, no. 2, pp. 494–514, Feb. 2022, ISSN: 2162-

2388. DOI: 10.1109/tnnls.2021.3070843. [Online]. Available: http://dx.doi.org/10.1109/tnnls.2021.3070843 (cit. on p. 60).

[109] J. Boné, M. Dias, J. C. Ferreira, and R. Ribeiro, "DisKnow: A Social-Driven Disaster Support Knowledge Extraction System," *Applied Sciences*, vol. 10, no. 17, p. 6083, Sep. 2020, ISSN: 2076-3417. DOI: 10.3390/app10176083. [Online]. Available: http://dx.doi.org/10.3390/app10176083 (cit. on p. 61).

[110] J. Son, C.-S. Lim, H.-S. Shim, and J.-S. Kang, "Development of Knowledge Graph for Data Management Related to Flooding Disasters Using Open Data," *Future Internet*, vol. 13, no. 5, p. 124, May 2021, ISSN: 1999-5903. DOI: 10.3390/fi13050124. [Online]. Available: http://dx.doi.org/10.3390/fi13050124 (cit. on p. 61).

[111] X. Ge *et al.*, "Disaster Prediction Knowledge Graph Based on Multi-Source Spatio-Temporal Information," *Remote Sensing*, vol. 14, no. 5, p. 1214, Mar. 2022, ISSN: 2072-4292. DOI: 10.3390/rs14051214. [Online]. Available: http://dx.doi.org/10.3390/rs14051214 (cit. on p. 61).

[112] S. Wang, W. Li, and Z. Gu, "GeoGraphViz: Geographically constrained 3D force-directed graph for knowledge graph visualization," *Transactions in GIS*, vol. 27, no. 4, pp. 931–948, Apr. 2023, ISSN: 1467-9671. DOI: 10.1111/tgis.13053 (cit. on p. 61).

[113] N. Muennighoff, N. Tazi, L. Magne, and N. Reimers, *MTEB: Massive Text Embedding Benchmark*, 2022. DOI: 10.48550/ARXIV.2210.07316. [Online]. Available: https://arxiv.org/abs/2210.07316 (cit. on p. 63).

[114] S. Aghababaei and M. Makrehchi, "Activity-based Twitter sampling for content-based and user-centric prediction models," *Human-Centric Computing and Information Sciences*, vol. 7, no. 1, Jan. 2017, ISSN: 2192-1962. DOI: 10.1186/s13673-016-0084-z. [Online]. Available: http://dx.doi.org/10.1186/s13673-016-0084-z (cit. on p. 67).

[115] C. Byun, H. Lee, Y. Kim, and K. Ko Kim, " Twitter data collecting tool with rule-based filtering and analysis module," *International*

Journal of Web Information Systems, vol. 9, no. 3, pp. 184–203, Aug. 2013, ISSN: 1744-0084. DOI: 10.1108/ijwis-04-2013-0011. [Online]. Available: http://dx.doi.org/10.1108/ijwis-04-2013-0011 (cit. on p. 67).

[116] V. Suarez-Lledo and J. Alvarez-Galvez, "Assessing the Role of Social Bots During the COVID-19 Pandemic: Infodemic, Disagreement, and Criticism," *Journal of Medical Internet Research*, vol. 24, no. 8, e36085, Aug. 2022, ISSN: 1438-8871. DOI: 10.2196/36085. [Online]. Available: http://dx.doi.org/10.2196/36085 (cit. on p. 67).

[117] Y. Zha, Y. Yang, R. Li, and Z. Hu, "AlignScore: Evaluating Factual Consistency with A Unified Alignment Function," in *Proceedings of the 61st Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers)*, Association for Computational Linguistics, 2023. DOI: 10.18653/v1/2023.acl-long.634. [Online]. Available: http://dx.doi.org/10.18653/v1/2023.acl-long.634 (cit. on pp. 70, 98).

[118] *Berkeley Protocol on Digital Open Source Investigations: A Practical Guide on the Effective Use of Digital Open Source Information in Investigating Violations of International Criminal, Human Rights and Humanitarian Law*. United Nations, Jun. 2022, ISBN: 9789210053433. DOI: 10.18356/9789210053433. [Online]. Available: http://dx.doi.org/10.18356/9789210053433 (cit. on pp. 70, 87).

[119] J. B. Mugeni and T. Amagasa, "A graph-based blocking approach for entity matching using pre-trained contextual embedding models," in *Proceedings of the 37th ACM/SIGAPP Symposium on Applied Computing*, ser. SAC '22, ACM, Apr. 2022, pp. 357–364. DOI: 10.1145/3477314.3507689. [Online]. Available: http://dx.doi.org/10.1145/3477314.3507689 (cit. on pp. 72, 73).

[120] W. Liu, C. Xie, and S. Zong, "A rumor source identification method based on node embeddings and community detection in social networks," in *2023 Eleventh International Conference on Advanced Cloud and Big Data (CBD)*, IEEE, Dec. 2023, pp. 104–

109. DOI: 10.1109/cbd63341.2023.00027. [Online]. Available: http://dx.doi.org/10.1109/cbd63341.2023.00027 (cit. on p. 73).

[121] A. Moura, P. Lima, F. Mendonça, S. S. Mostafa, and F. Morgado-Dias, "On the Use of Transformer-Based Models for Intent Detection Using Clustering Algorithms," *Applied Sciences*, vol. 13, no. 8, p. 5178, Apr. 2023, ISSN: 2076-3417. DOI: 10.3390/app13085178. [Online]. Available: http://dx.doi.org/10.3390/app13085178 (cit. on p. 73).

[122] Y. Liang, J. Zhu, W. Ye, and S. Gao, "Region2Vec: community detection on spatial networks using graph embedding with node attributes and spatial interactions," in *Proceedings of the 30th International Conference on Advances in Geographic Information Systems*, ser. SIGSPATIAL '22, ACM, Nov. 2022, pp. 1–4. DOI: 10.1145/3557915.3560974. [Online]. Available: http://dx.doi.org/10.1145/3557915.3560974 (cit. on p. 73).

[123] S. Kalra, Y. Sharma, M. Agrawal, S. Mantri, and G. Chauhan, "Impact of Transformer-Based Models and User Clustering in Early Fake News Detection in Social Media," in *Proceedings of the 12th International Conference on Pattern Recognition Applications and Methods*, SCITEPRESS - Science and Technology Publications, 2023. DOI: 10.5220/0011684000003411. [Online]. Available: http://dx.doi.org/10.5220/0011684000003411 (cit. on p. 73).

[124] S. Rojas-Galeano, "Zero-Shot Spam Email Classification Using Pre-trained Large Language Models," 2024. DOI: 10.48550/ARXIV.2405.15936. [Online]. Available: https://arxiv.org/abs/2405.15936 (cit. on p. 73).

[125] F. M. Plaza-del-Arco, D. Nozza, and D. Hovy, *Wisdom of Instruction-Tuned Language Model Crowds. Exploring Model Label Variation*, 2023. DOI: 10.48550/ARXIV.2307.12973. [Online]. Available: https://arxiv.org/abs/2307.12973 (cit. on p. 73).

[126] J. Drápal, H. Westermann, and J. Savelka, "Using Large Language Models to Support Thematic Analysis in Empirical Legal Studies," 2023. DOI: 10.48550/ARXIV.2310.18729. [Online]. Available: https://arxiv.org/abs/2310.18729 (cit. on p. 73).

[127] Z. Wang, Y. Pang, and Y. Lin, *Large Language Models Are Zero-Shot Text Classifiers*, 2023. DOI: 10.48550/ARXIV.2312.01044. [Online]. Available: https://arxiv.org/abs/2312.01044 (cit. on p. 73).

[128] H. Raja *et al.*, *Using Large Language Models to Automate Category and Trend Analysis of Scientific Articles: An Application in Ophthalmology*, 2023. DOI: 10.48550/ARXIV.2308.16688. [Online]. Available: https://arxiv.org/abs/2308.16688 (cit. on p. 73).

[129] I. Afyouni, A. Khan, and Z. Al Aghbari, "E-ware: a big data system for the incremental discovery of spatio-temporal events from microblogs," *Journal of Ambient Intelligence and Humanized Computing*, vol. 14, no. 10, pp. 13 949–13 968, Jun. 2022, ISSN: 1868-5145. DOI: 10.1007/s12652-022-04104-4. [Online]. Available: http://dx.doi.org/10.1007/s12652-022-04104-4 (cit. on p. 74).

[130] W. Ni, Q. Shen, T. Liu, Q. Zeng, and L. Xu, "Generating textual emergency plans for unconventional emergencies — A natural language processing approach," *Safety Science*, vol. 160, p. 106 047, Apr. 2023, ISSN: 0925-7535. DOI: 10.1016/j.ssci.2022.106047. [Online]. Available: http://dx.doi.org/10.1016/j.ssci.2022.106047 (cit. on p. 74).

[131] I. Afyouni, A. Khan, and Z. A. Aghbari, "Deep-Eware: spatio-temporal social event detection using a hybrid learning model," *Journal of Big Data*, vol. 9, no. 1, Jun. 2022, ISSN: 2196-1115. DOI: 10.1186/s40537-022-00636-w. [Online]. Available: http://dx.doi.org/10.1186/s40537-022-00636-w (cit. on p. 74).

[132] L. Huang, P. Shi, and H. Zhu, "An integrated urgency evaluation approach of relief demands for disasters based on social media

data," *International Journal of Disaster Risk Reduction*, vol. 80, p. 103 208, Oct. 2022, ISSN: 2212-4209. DOI: 10.1016/j.ijdrr.2022.103208. [Online]. Available: http://dx.doi.org/10.1016/j.ijdrr.2022.103208 (cit. on p. 75).

[133] D. Contreras, S. Wilkinson, E. Alterman, and J. Hervás, "Accuracy of a pre-trained sentiment analysis (SA) classification model on tweets related to emergency response and early recovery assessment: the case of 2019 Albanian earthquake," *Natural Hazards*, vol. 113, no. 1, pp. 403–421, Mar. 2022, ISSN: 1573-0840. DOI: 10.1007/s11069-022-05307-w. [Online]. Available: http://dx.doi.org/10.1007/s11069-022-05307-w (cit. on p. 75).

[134] B. Zhou *et al.*, "VictimFinder: Harvesting rescue requests in disaster response from social media with BERT," *Computers, Environment and Urban Systems*, vol. 95, p. 101 824, Jul. 2022, ISSN: 0198-9715. DOI: 10.1016/j.compenvurbsys.2022.101824. [Online]. Available: http://dx.doi.org/10.1016/j.compenvurbsys.2022.101824 (cit. on p. 75).

[135] J. A. Wahid *et al.*, "Topic2Labels: A framework to annotate and classify the social media data through LDA topics and deep learning models for crisis response," *Expert Systems with Applications*, vol. 195, p. 116 562, Jun. 2022, ISSN: 0957-4174. DOI: 10.1016/j.eswa.2022.116562. [Online]. Available: http://dx.doi.org/10.1016/j.eswa.2022.116562 (cit. on p. 75).

[136] D. Contreras, S. Wilkinson, N. Balan, and P. James, "Assessing post-disaster recovery using sentiment analysis: The case of L'Aquila, Italy," *Earthquake Spectra*, vol. 38, no. 1, pp. 81–108, Aug. 2021, ISSN: 1944-8201. DOI: 10.1177/8755293021103648 6. [Online]. Available: http://dx.doi.org/10.1177/875529302110 36486 (cit. on p. 76).

[137] S. Bashir *et al.*, "Twitter chirps for Syrian people: Sentiment analysis of tweets related to Syria Chemical Attack," *International Journal of Disaster Risk Reduction*, vol. 62, p. 102 397, Aug. 2021, ISSN: 2212-4209. DOI: 10.1016/j.ijdrr.2021.102397. [Online].

Available: http://dx.doi.org/10.1016/j.ijdrr.2021.102397 (cit. on p. 76).

[138] S. Behl, A. Rao, S. Aggarwal, S. Chadha, and H. Pannu, " Twitter for disaster relief through sentiment analysis for COVID-19 and natural hazard crises," *International Journal of Disaster Risk Reduction*, vol. 55, p. 102 101, Mar. 2021, ISSN: 2212-4209. DOI: 10.1016/j.ijdrr.2021.102101. [Online]. Available: http://dx.doi.org/10.1016/j.ijdrr.2021.102101 (cit. on p. 76).

[139] Y. Lian, Y. Liu, and X. Dong, "Strategies for controlling false online information during natural disasters: The case of Typhoon Mangkhut in China," *Technology in Society*, vol. 62, p. 101 265, Aug. 2020, ISSN: 0160-791X. DOI: 10.1016/j.techsoc.2020.101265. [Online]. Available: http://dx.doi.org/10.1016/j.techsoc.2020.101265 (cit. on p. 76).

[140] A. Karami, V. Shah, R. Vaezi, and A. Bansal, " Twitter speaks: A case of national disaster situational awareness," *Journal of Information Science*, vol. 46, no. 3, pp. 313–324, Mar. 2019, ISSN: 1741-6485. DOI: 10.1177/0165551519828620. [Online]. Available: http://dx.doi.org/10.1177/0165551519828620 (cit. on p. 77).

[141] M. Farnaghi, Z. Ghaemi, and A. Mansourian, "Dynamic Spatio-Temporal Tweet Mining for Event Detection: A Case Study of Hurricane Florence," *International Journal of Disaster Risk Science*, vol. 11, no. 3, pp. 378–393, May 2020, ISSN: 2192-6395. DOI: 10.1007/s13753-020-00280-z. [Online]. Available: http://dx.doi.org/10.1007/s13753-020-00280-z (cit. on p. 77).

[142] A. G. Gulnerman and H. Karaman, "Spatial Reliability Assessment of Social Media Mining Techniques with Regard to Disaster Domain-Based Filtering," *ISPRS International Journal of Geo-Information*, vol. 9, no. 4, p. 245, Apr. 2020, ISSN: 2220-9964. DOI: 10.3390/ijgi9040245. [Online]. Available: http://dx.doi.org/10.3390/ijgi9040245 (cit. on p. 77).

[143] C. Fan, F. Wu, and A. Mostafavi, “A Hybrid Machine Learning Pipeline for Automated Mapping of Events and Locations From Social Media in Disasters,” *IEEE Access*, vol. 8, pp. 10 478–10 490, 2020, ISSN: 2169-3536. DOI: 10.1109/access.2020.2965550. [Online]. Available: http://dx.doi.org/10.1109/access.2020.2965550 (cit. on p. 77).

[144] N. Singh, N. Roy, and A. Gangopadhyay, “Analyzing The Emotions of Crowd For Improving The Emergency Response Services,” *Pervasive and Mobile Computing*, vol. 58, p. 101 018, Aug. 2019, ISSN: 1574-1192. DOI: 10.1016/j.pmcj.2019.04.009. [Online]. Available: http://dx.doi.org/10.1016/j.pmcj.2019.04.009 (cit. on p. 78).

[145] D. Gordon, *Hurricane Harvey Tweets*, Sep. 2017. [Online]. Available: https://www.kaggle.com/datasets/dan195/hurricaneharvey (cit. on p. 80).

[146] G. Preda, *Turkey Earthquake Tweets*, Feb. 2023. [Online]. Available: https://www.kaggle.com/datasets/gpreda/turkey-earthquake-tweets/data (cit. on p. 80).

[147] KoalaAI, *KoalaAI/Text-Moderation*, Published: Hugging Face repository, 2024 (cit. on pp. 81, 88).

[148] T. Wolf *et al.*, *HuggingFace’s Transformers: State-of-the-art Natural Language Processing*, 2019. DOI: 10.48550/ARXIV.1910.03771. [Online]. Available: https://arxiv.org/abs/1910.03771 (cit. on p. 81).

[149] A. V. Solatorio, *GISTEmbed: Guided In-sample Selection of Training Negatives for Text Embedding Fine-tuning*, 2024. DOI: 10.48550/ARXIV.2402.16829. [Online]. Available: https://arxiv.org/abs/2402.16829 (cit. on p. 82).

[150] T. Rebedea, R. Dinu, M. N. Sreedhar, C. Parisien, and J. Cohen, “NeMo Guardrails: A Toolkit for Controllable and Safe LLM Applications with Programmable Rails,” in *Proceedings of the 2023 Conference on Empirical Methods in Natural Language Processing: System Demonstrations*, Association for Computational

Linguistics, 2023. DOI: 10.18653/v1/2023.emnlp-demo.40. [Online]. Available: http://dx.doi.org/10.18653/v1/2023.emnlp-demo.40 (cit. on p. 83).

[151] P. Zhao *et al.*, *Retrieval-Augmented Generation for AI-Generated Content: A Survey*, 2024. DOI: 10.48550/ARXIV.2402.19473. [Online]. Available: https://arxiv.org/abs/2402.19473 (cit. on p. 87).

[152] H. Yu, A. Gan, K. Zhang, S. Tong, Q. Liu, and Z. Liu, "Evaluation of Retrieval-Augmented Generation: A Survey," in *Big Data*. Springer Nature Singapore, 2025, pp. 102–120. DOI: 10.1007/978-981-96-1024-2_8. [Online]. Available: http://dx.doi.org/10.1007/978-981-96-1024-2_8 (cit. on p. 87).

[153] W. Su, Y. Tang, Q. Ai, Z. Wu, and Y. Liu, *DRAGIN: Dynamic Retrieval Augmented Generation based on the Information Needs of Large Language Models*, 2024. DOI: 10.48550/ARXIV.2403.10081. [Online]. Available: https://arxiv.org/abs/2403.10081 (cit. on p. 87).

[154] V. Gupta and S. R. P, *Leveraging open-source models for legal language modeling and analysis: a case study on the Indian constitution*, 2024. DOI: 10.48550/ARXIV.2404.06751. [Online]. Available: https://arxiv.org/abs/2404.06751 (cit. on p. 87).

[155] E. S. Blake, C. Landsea, and E. J. Gibney, "The Deadliest, Costliest, and Most Intense United States Tropical Cyclones from 1851 to 2010 and Other Frequently Requested Hurricane Facts)," *NOAA Technical Memorandum NWS NHC-6.National Oceanic and Atmospheric Administration, National Weather Service, National Hurricane Center*, 2011. [Online]. Available: https://www.nhc.noaa.gov/pdf/nws-nhc-6.pdf (cit. on p. 89).

[156] B. Maletckii, E. Astafyeva, S. A. Sanchez, E. A. Kherani, and E. R. de Paula, "The 6 February 2023 Türkiye Earthquake Sequence as Detected in the Ionosphere," *Journal of Geophysical Research: Space Physics*, vol. 128, no. 9, Sep. 2023, ISSN: 2169-9402. DOI:

10.1029/2023ja031663. [Online]. Available: http://dx.doi.org/10.1029/2023ja031663 (cit. on p. 89).

Appendix

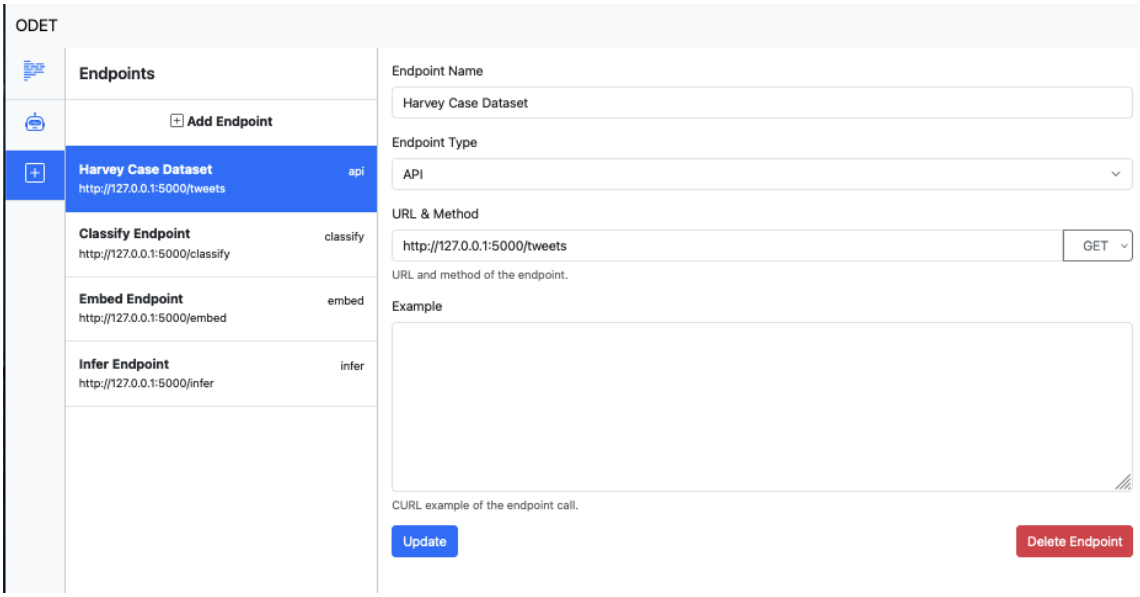

Figure A.1: Overview of the endpoints user interface for the Hurricane Harvey case study in ODET.

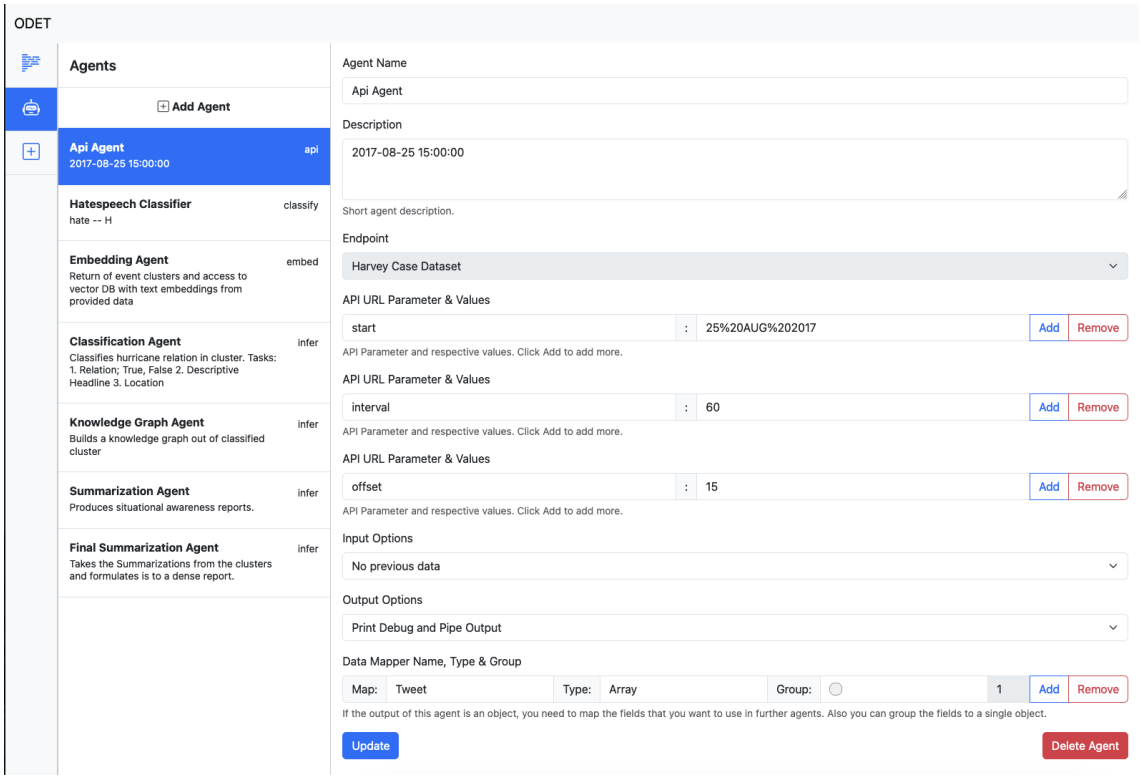

Figure A.2: Overview of the API agent user interface for the hurricane Harvey case study in ODET.

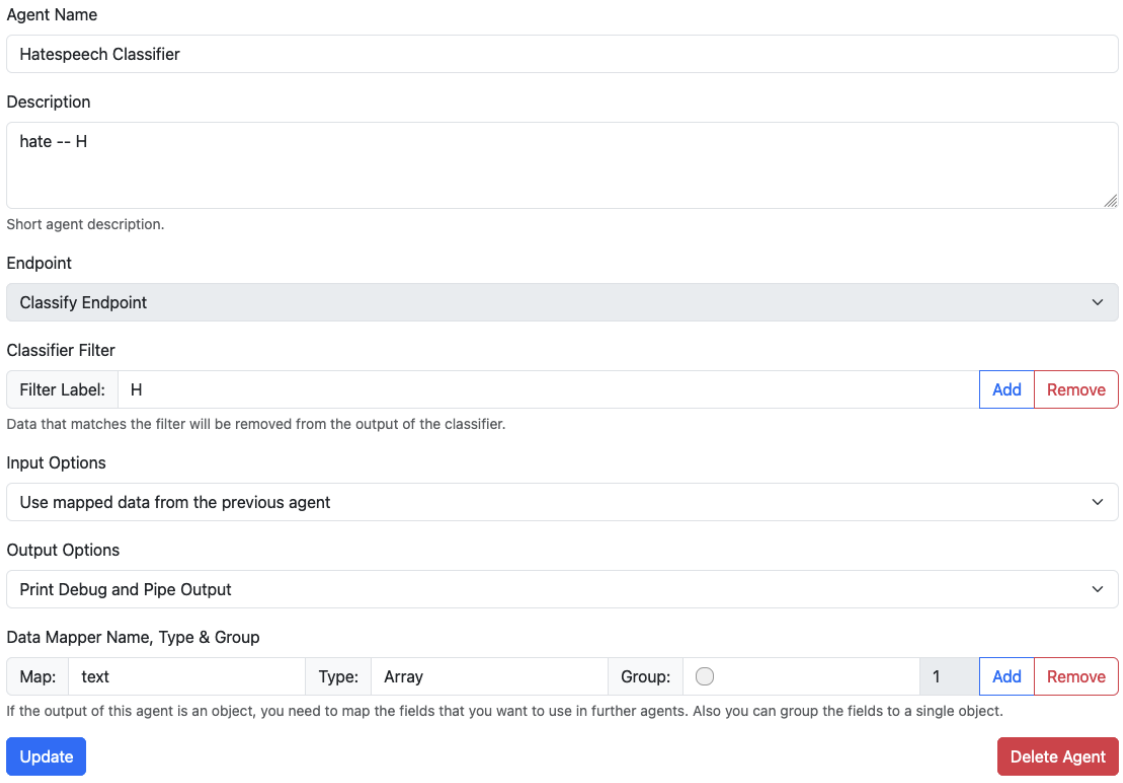

Figure A.3: Overview of the hate-speech filter agent user interface for the hurricane Harvey case study in ODET.

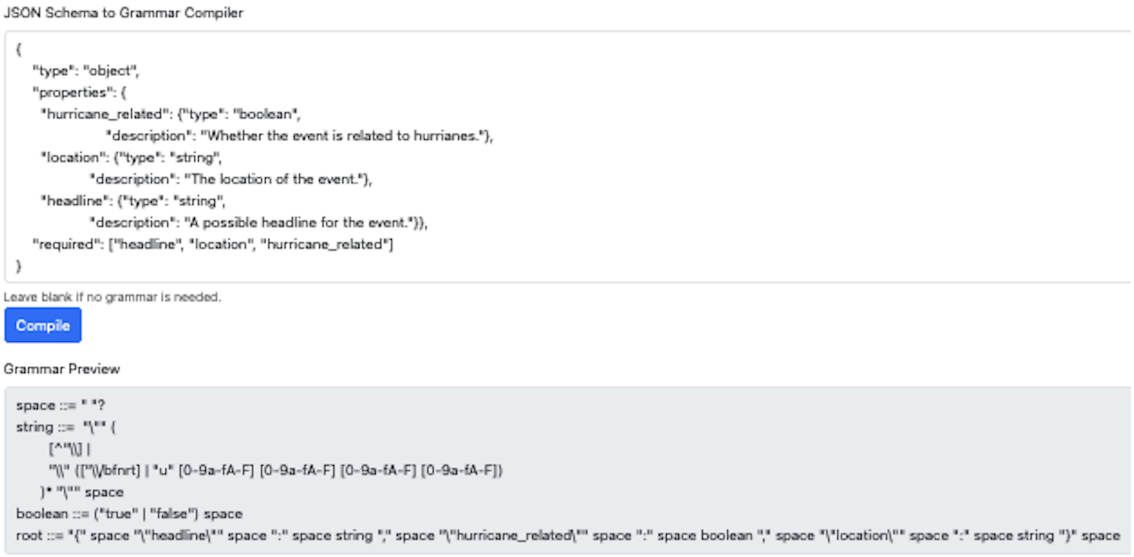

Figure A.4: Classification Agent—Grammar and JSON schema example.

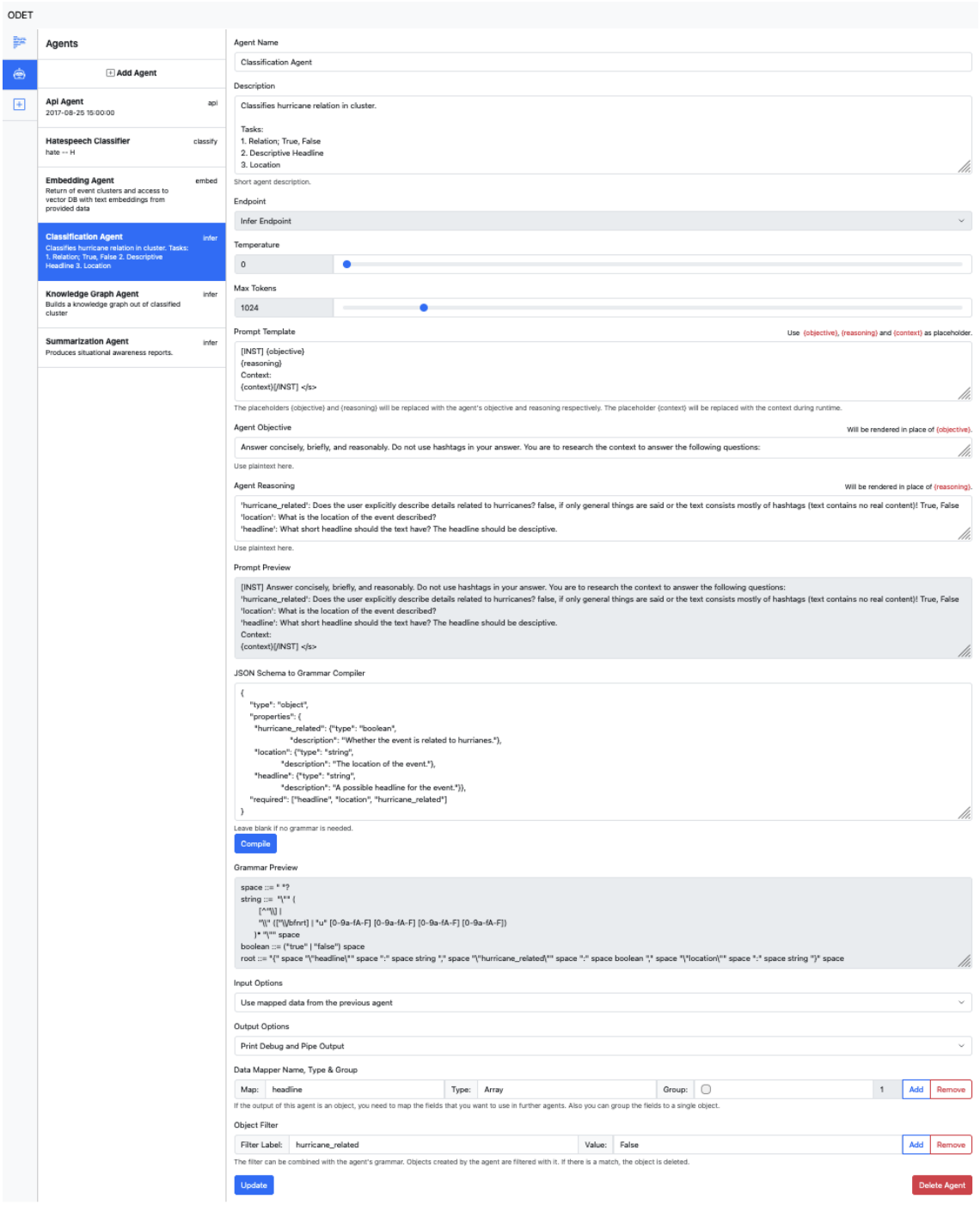

Figure A.5: Overview of the classification agent user interface for the Hurricane Harvey case study in ODET.

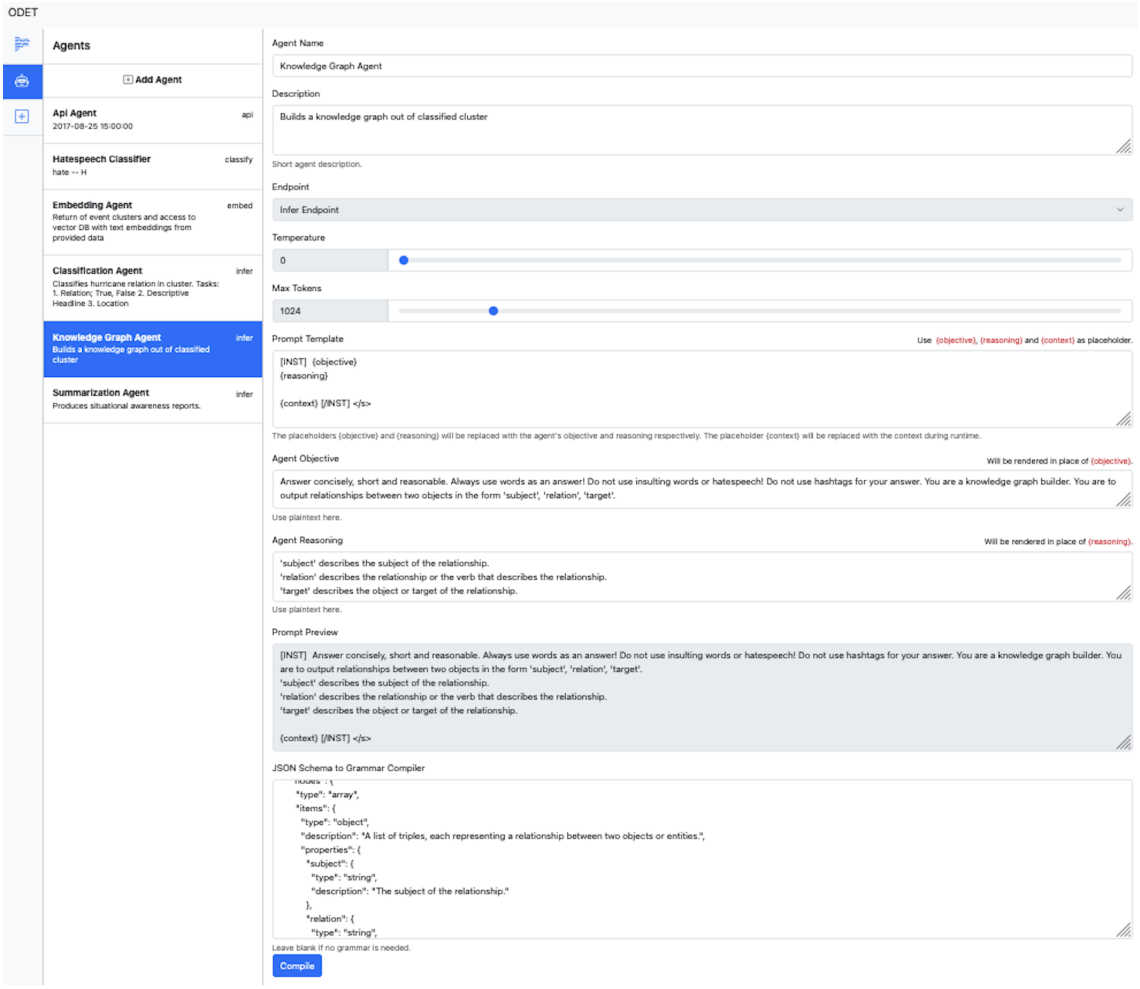

Figure A.6: Overview of the knowledge graph agent user interface for the Hurricane Harvey case study in ODET.

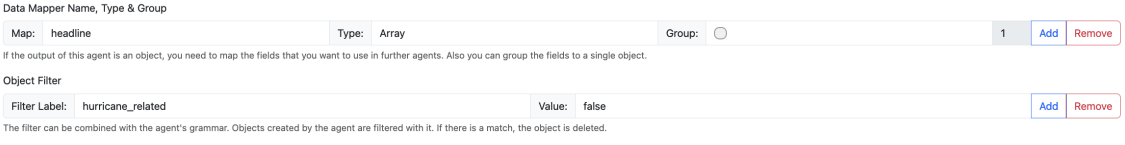

Figure A.7: Classification agent—mapper and object filter example.

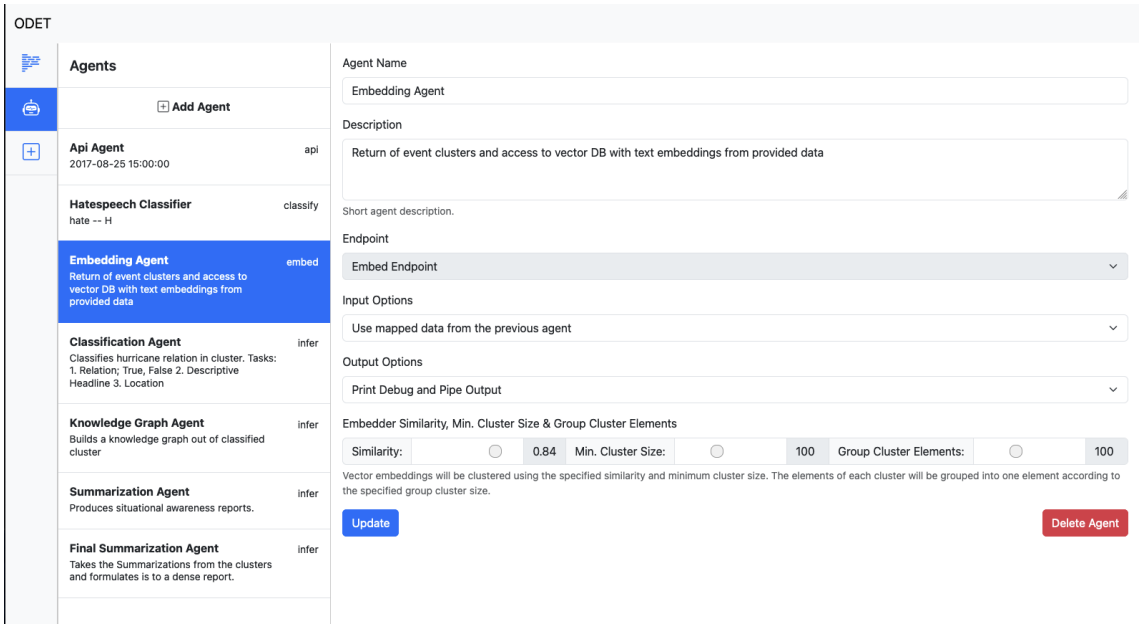

Figure A.8: Overview of the embedding agent user interface for the hurricane Harvey case study in ODET.

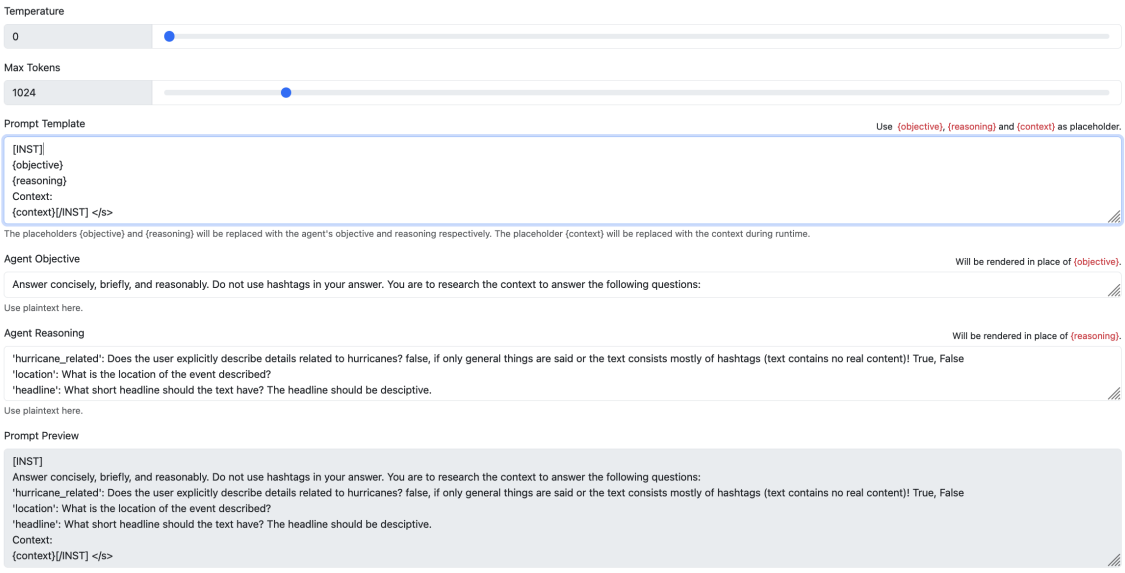

Figure A.9: Classification Agent—Prompt template and objective example.

Figure A.10: Knowledge Graph Agent—Grammar and JSON schema example.

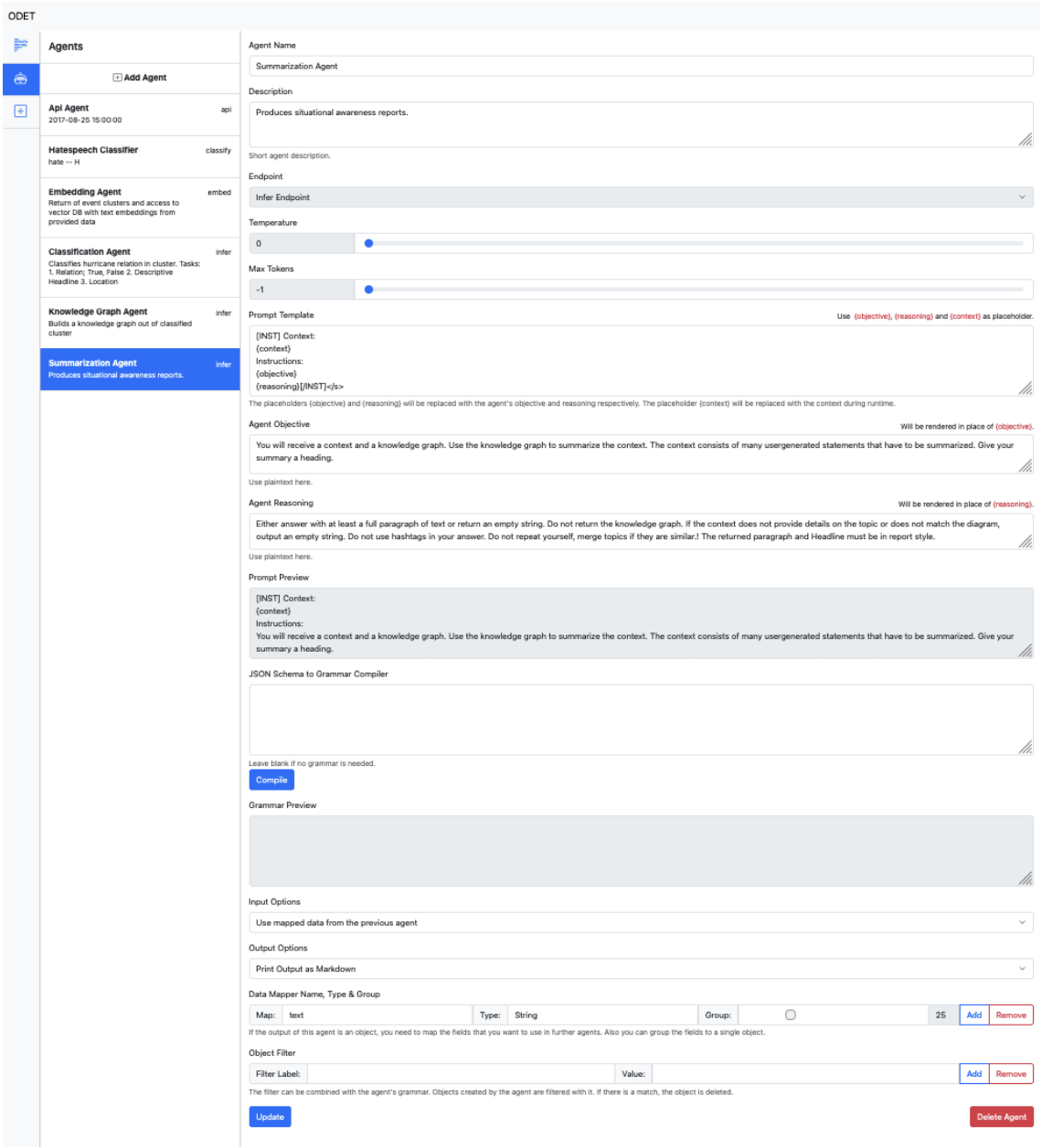

Figure A.11: Overview of the summarization agent user interface for the Hurricane Harvey case study in ODET.

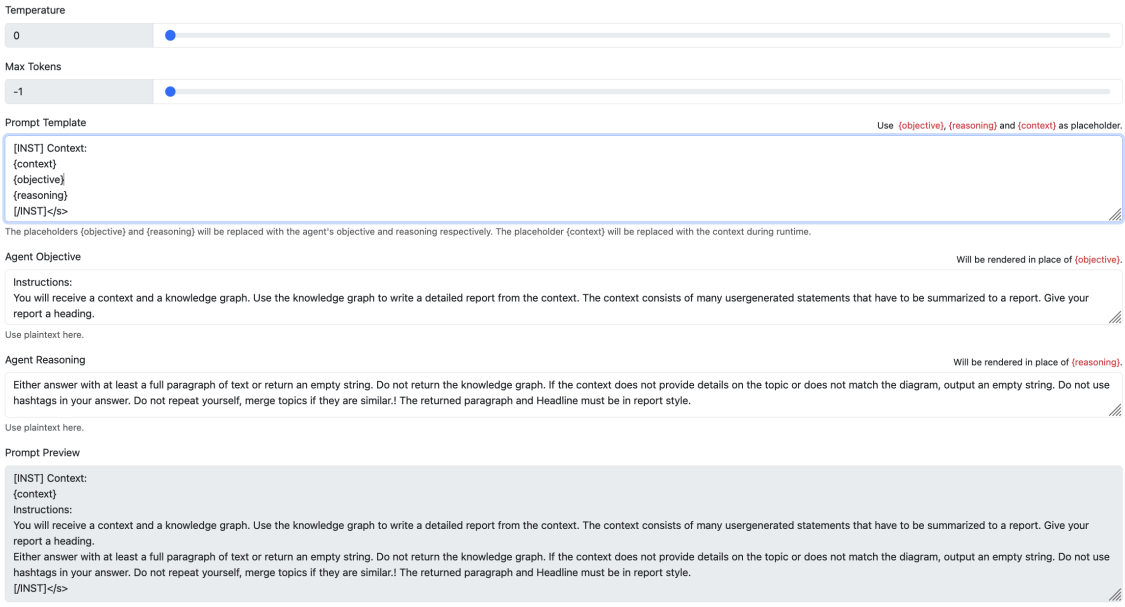

Figure A.12: Summarization Agent—Prompt example.

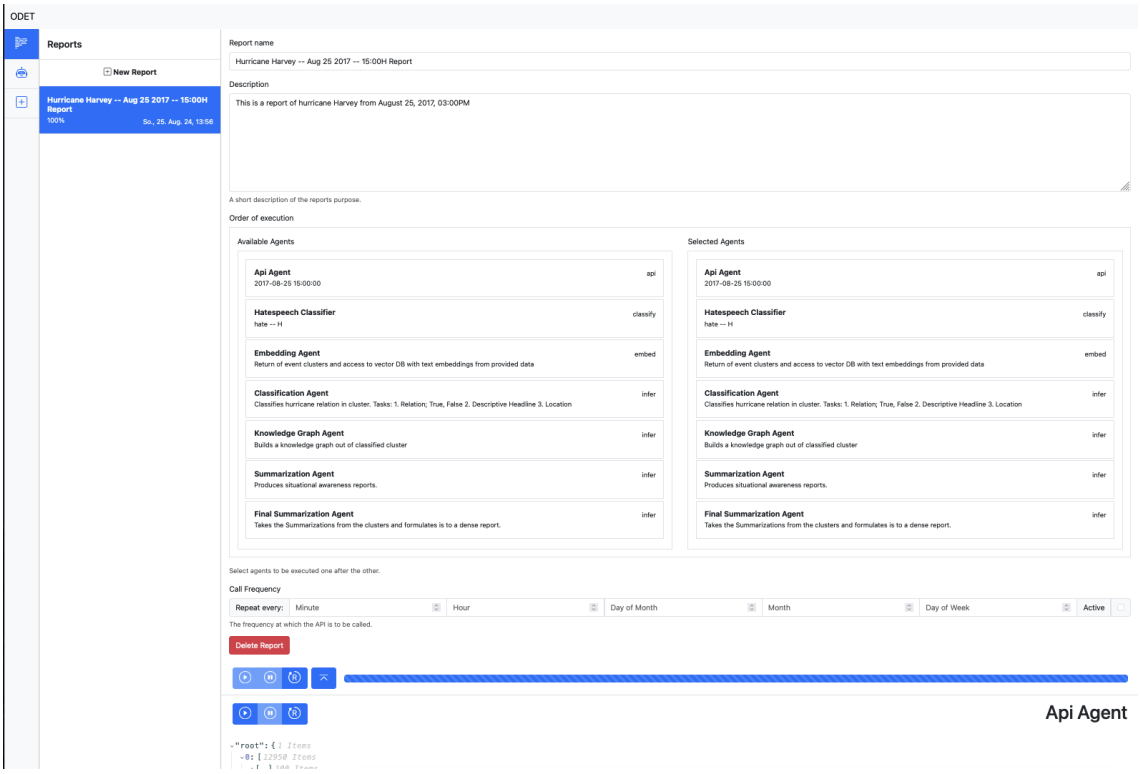

Figure A.13: Overview of the report configuration user interface for the Hurricane Harvey case study in ODET.