Social Network Analysis and Mining (2024) 14:154
https://doi.org/10.1007/s13278-024-01308-8

ORIGINAL ARTICLE

# Twitter analysis in emergency management: recent research and trends

**Alireza Arvandi[1] · Jon Rokne[2] · Reda Alhajj[3]**

Received: 1 August 2023 / Revised: 31 May 2024 / Accepted: 11 July 2024


**Abstract**
A disaster is an unexpected event with negative consequences for individuals and societies. Typically it is interfering with a community's or society's ability to function. Disasters can be man-made events, accidental events, or a natural catastrophes. These events create emergencies that require rapid responses. Timely information about the emergencies has to be obtained so that the responses can aid in dealing with the emergencies. One source of timely information about disasters is provided by social media postings which often provide on-site information about a disaster. An ideally suited social media tool for disaster information is Twitter due to the sort message format. This format enables the rapid composition of short messages by entities close to a disaster describing the nature of the disaster. The contents of the messages can then be used to guide emergency responses. The aim of this paper is to review the research on the usage of Twitter for emergency management that has been published so far. There are three steps required when using messages for information about disasters. The first step is to collect the data contained in the messages. Then the data has to be preprocessed for unification of format, duplication etc. Finally relevant information has to be extracted. These steps are considered in this paper reviewing the use of Twitter for emergency management. Papers using Twitter published within the past 5 years have been included.

**Keywords** Disaster · Social media · Twitter · Emergency management

## 1 Introduction

Social media platforms play an important role in modern life. Among these media, one of the currently most important ones is Twitter, which is used for quick information sharing and communication. Many people, especially younger adults, use Twitter daily (Pewresearch 2021). Twitter has also experienced tremendous growth in users over the past few years, and its influence in real life is becoming significant, affecting and shaping public opinion (Gaisbauer et al. 2009). This snapshot of how Twitter is used might change with the take-over by Elon Musk (Southern 2022). This paper reviews the usage of Twitter in disaster management up to the take-over.

Twitter tends to be used more frequently during disasters than during normal day-to-day situations (Xu 2020). For example, people may use Twitter to communicate with their loved ones in the disaster area to ensure they are safe and secure (Long et al. 2016). They share information about food, water, and other needs (Murakami et al. 2019). They also share information about infrastructure damages and problems (Thekdi and Chatterjee 2019).

Communication during emergencies using television and radio has the disadvantage of one-way communication from the source of information to the recipient. Twitter and other social media platforms do not have this limitation and can therefore be used to inform and share information about the emergency. By using Twitter during emergencies, anyone can be a potential source of information or someone that collects and distributes information.

✉ Alireza Arvandi
alireza.arvandi@ucalgary.ca

Jon Rokne
rokne@ucalgary.ca

Reda Alhajj
alhajj@ucalgary.ca

[1] Department of Computer Science, University of Calgary, Calgary, Alberta, Canada

[2] Department of Heath Informatics, University of Southern Denmark, Odense, Denmark

[3] Department of Computer Engineering, Istanbul Medipol University, Istanbul, Turkey

Published online: 07 August 2024

A few surveys on the topic of emergency management exist. However, they primarily focus on specific domains of emergency management, like prediction (Huang et al. 2021), improving community resilience (Rachunok et al. 2021), or methods used in a specific phase when handling a disaster (Ogie et al. 2022). In Saroj and Pal (2020), the authors cover the technicalities involved when using social media data, which is more related to what is discussed in this paper. Their approach was to cluster previous researchers' emergency management methods.

In this paper, the focus is on utilizing tweets as a valuable source of information for disaster management. Unlike previous studies that primarily analyze the methods employed, this study examines the various sources of information utilized in light of disaster management goals. The papers reviewed are categorized into three main groups based on the sources of information: the content of tweets, the Twitter network, and the metadata associated with tweets.

The aim is to analyze previous research concerning the utilization of social media, specifically Twitter, as an information conduit for disaster management. While prior studies have predominantly focused on specific domains within emergency management, this paper encompasses information relevant to all domains covered in previous research. The main goal is to provide valuable insights into using Twitter as a communication and information-sharing platform during times of crisis. Researchers can benefit from this paper by understanding the various aspects of the information provided through Twitter data and how this data has been previously used, specifically in the context of disaster management. The paper offers a comprehensive perspective on the information sources and the necessary details to consider before selecting an approach. This facilitates informed decision-making and encourages the exploration of alternative approaches for emergency management.

The major contributions of this paper are as follows:

- It offers valuable insights to researchers on addressing emergency management challenges using Twitter.
- It provides guidance and ideas to researchers unfamiliar with this domain, outlining different steps to be taken.
- It provides detailed explanations of the data collection step.
- It provides in-depth explanations of data pre-processing and the reasons behind its necessity.
- It explores various sources of information and demonstrates how researchers can effectively utilize them.

The rest of the paper is structured as follows: Section 2 describes different approaches to data collection from Twitter for disaster management. In Sect. 3, the approaches to pre-processing of information from Twitter are described. In Sect. 4, papers are clustered based on the information type used for disaster management by providing their approaches for each analysis. A conclusion is provided in Sect. 5.

## 2 Data collection

Since over 500 million tweets are posted on Twitter each day (Internetlivestats 2021), collecting relevant information from these tweets is a challenge for researchers. In what follows, different approaches for gathering information, the sizes of the datasets explored, and the collection durations are discussed. The number of tweets collected about an event depends on several factors, including the number of hashtags considered, the observation period, and the size of the event. The approaches to gathering data from Twitter and analyzing it are: (1) using available public datasets of tweets, (2) developing web scrapers or using existing web scrapers to collect tweets and associated information, and (3) utilizing third-party tools to access and filter data.

### 2.1 Using Twitter APIs

Twitter APIs allow researchers to access the data contained in their applications. These APIs enable gathering real-time and historical data based on different rules. Most of the papers surveyed used these APIs with keyword-based filtration (Bec and Becken 2021; Yue et al. 2021; Jin and Spence 2021; Thekdi and Chatterjee 2019; Schempp et al. 2019; Pourebrahim et al. 2019; Kumar and Singh 2019; Kankanamge et al. 2020; Gulesan et al. 2021; Li et al. 2021) to gather tweets related to an event or topic. Some papers filter and query tweets based on their authors (Wong and Jensen 2020; Lucas and Landman 2021; Eachus and Keim 2020). Another approach used by researchers for collecting data with APIs is based on language, location, and time (Bec and Becken 2021; Yue et al. 2021; Wong and Jensen 2020; Pourebrahim et al. 2019; Bhavaraju et al. 2019; Forati and Ghose 2022). Twitter APIs are public and free with some limitations on tweets captured, number of rules applied, and full-archive searches (Twitter 2022).

#### 2.1.1 The advantages of using Twitter APIs

- *Real-time and historical data access* Twitter APIs allow researchers to gather both real-time and past tweets, thus providing valuable insights into ongoing events and trends.
- *Filtered data collection* Researchers can use keyword-based filtration and other parameters to collect tweets specifically related to their event or topic of interest.

#### 2.1.2 Limitations

- *Limit on captured tweets* Twitter APIs limit the number of tweets that can be captured, restricting the scope of analysis for larger events or high-volume discussions.
- *Restricted rules and search capabilities* Researchers are constrained by the rules and search capabilities provided by Twitter APIs, which may not always cater to their specific research needs.

#### 2.1.3 Potential challenges

- *Learning curve and technical implementation* Utilizing Twitter APIs effectively requires knowledge of API documentation, authentication procedures, and programming skills to retrieve the data and process it.
- *Ethical considerations and compliance* Researchers need to ensure that they comply with Twitter's terms of service, privacy policies, and ethical guidelines when collecting and using data obtained through the APIs.

### 2.2 Scraping Twitter

Some studies such as Dong et al. (2021), Dereli et al. (2021) have scraped Twitter for data collection. This approach involves collecting information from the Twitter website by parsing HTML (HyperText Markup Language) pages to extract the information. Developers can also develop a scraper or use tools such as Selenium (Dereli et al. 2021) to scrape Twitter, which could be used as an alternative to the APIs without their limitations. It is challenging and time-consuming to develop such a web scraper, especially since Twitter blocks multiple HTTP responses, which might result in long run times.

#### 2.2.1 Advantages

- *Access to comprehensive data* Scraping Twitter allows researchers to gather data directly from its website, providing access to a wide range of information, including text, images, and user profiles.
- *Flexibility and customization* Researchers can develop custom web scrapers or use tools like Selenium to tailor the data collection process according to their specific requirements.
- *Independence from Twitter API limitations* Scraping Twitter bypasses the limitations imposed by the Twitter APIs, enabling researchers to capture larger volumes of data without restrictions on tweet counts.

#### 2.2.2 Limitations

- *Technical complexity and resource requirements* Developing and maintaining a web scraper can be technically challenging and resource-intensive, requiring expertise in web scraping techniques, programming, and handling potential obstacles such as blocking measures implemented by Twitter.
- *Legal and ethical considerations* Scraping Twitter may raise legal concerns and ethical considerations, as it involves accessing and extracting data from a platform governed by terms of service and privacy policies.

#### 2.2.3 Potential challenges

- *Continuous adaptation to website changes* Twitter frequently updates its website structure and layout, which may require regular modifications to the web scraper to ensure compatibility and data extraction reliability.
- *Data quality and consistency* Scraped data may be prone to inconsistencies or inaccuracies due to changes in website formatting, variations in HTML structure, or intermittent connectivity issues.

### 2.3 Using datasets

Many researchers have investigated the outcomes of high-impact emergencies and disasters, resulting in abundant data available for these events. They have generally used public datasets of tweets and archives of general data from the adverse events (Jamali et al. 2020; Kitazawa and Hale 2021; Behl et al. 2021; Forati and Ghose 2022; Bhullar et al. 2022; Ngamassi et al. 2022; Cheong and Babcock 2021). In Table 1, papers using such datasets are included. These papers use existing datasets that are based on various events or on events with a commonality such as a specific topic. The datasets have been gathered in previous studies by researchers (Cheong and Babcock 2021), provided by third-party vendors and companies (Pont-Sorribes et al. 2020; Contreras et al. 2022), or they can be public datasets for a specific topic. The datasets have varying sizes depending on the topic and event. It can be a very large dataset containing billions of tweets (Jamali et al. 2020) or it can be on the order of millions (Pont-Sorribes et al. 2020; Kitazawa and Hale 2021; Forati and Ghose 2022). Also, datasets covering specialized narrow topics may contain less than 5 thousand tweets (Contreras et al. 2022).

#### 2.3.1 Advantages

- *Availability of pre-collected data* Researchers can leverage existing datasets that cover specific events or topics, saving time and effort in data collection.

**Table 1** Data collection

| References | Data collection strategy | Dataset description | Dataset size | Temporal scope |
|---|---|---|---|---|
| Jamali et al. (2020) | Using datasets | A large dataset containing billions of geotagged tweets | Not specified exactly (billions of tweets) | October 25, 2012 to October 25, 2014 - 2 years |
| Wang and Guo (2021) | Using third-party tools and applications; Crimson hexagon software | Based on keywords. A dataset was gathered using Crimson hexagon software | 28,236 related tweets | October 21, 2015 to December 31, 2016 |
| Whittingham et al. (2020) | Using third-party tools and applications; Netlytic | Gathered tweets using Netlytic for three different topics | 103,001 tweets | 3 different timelines |
| Bec and Becken (2021) | Twitter API | Gathered tweets based on location | 1134 eligible posts | Not specified |
| Lauran et al. (2020) | Using third-party tools and applications; Coosto | All tweets sent between July 17 and October 5, 2017, related to the fipronil case in the Netherlands are scraped from the Coosto | 87,454 tweets | July 17, 2017 to October 5, 2017 |
| Yue et al. (2021) | Twitter API | Used keywords and location filters to find related tweets | 64,990 tweets | July 29, 2015 to August 29, 2015 |
| Wong and Jensen (2020) | Not specified | Tracked activity in Singapore relating to COVID19 from a dataset of 3.47 million global tweets using hashtags related to the coronavirus | Analysis was conducted for a final dataset of 200,002 | January 29, 2020 to February 28, 2020 |
| Jin and Spence (2021) | Using third-party tools and applications; NodeXL Pro | The tweets related to Hurricane Maria were collected by searching keywords | 46,335 tweets | September 16, 2017 to October 3, 2017 |
| Lucas and Landman (2021) | Twitter API | Anti-slavery organizations' tweets showing the activities related to COVID-19 filtered by hashtags | 7601 tweets | January 1, 2020 to June 30, 2020 |
| Pont-Sorribes et al. (2020) | Using datasets | A Spanish company Pirendo specializing in the monitoring and collection of Twitter data collected tweets related to the Ebola crisis in Spain and the Germanwings crisis | 1,334,608 tweets | Not specified |
| Yang et al. (2019) | Twitter API | Tweets from all users whose location in their user profile was designated as Houston through Twitter's PowerTrack API were collected | 5,662,137 tweets | August 25, 2017 to August 31, 2017 |
| Pourebrahim et al. (2019) | Twitter API | 13.7 Million tweets were collected to create two datasets based on keywords and location | 13.7 Million tweets | October 22, 2012 to November 7, 2012 |
| Bhavaraju et al. (2019) | Using datasets | A stream of data was collected over multiple years by a group dedicated to preserving online digital content | Not specified | 2013 to 2017 |

**Table 1** (continued)

| References | Data collection strategy | Dataset description | Dataset size | Temporal scope |
|---|---|---|---|---|
| Kumar and Singh (2019) | Twitter API | Tweets related to earthquakes were collected using keywords from Twitter Streaming API | 103,384 tweets | October 20, 2017 to March 15, 2018 |
| Kankanamge et al. (2020) | Twitter API | Tweets with keywords were collected | 12,910 tweets | December 25, 2010 to January 25, 2011 |
| Zhai et al. (2020) | Twitter API | A bounding box-based query was used to gather all tweets in a specific area | 2.2 Million tweets | September 10, 2018 to Sep 20, 2018 |
| Beedasy et al. (2020) | Twitter API | A set of historical tweets was identified based on a search strategy built on a set of keywords using Twitter PowerTrack API | 876,298 tweets | April 20, 1010 to August 20, 2010 |
| Jamali et al. (2020) | Not specified | Geotagged tweets were obtained for the first and second year | 109.1 Million tweets | October 25, 2012 to October 25, 2014 |
| Devaraj et al. (2020) | Twitter API | Tweets were provided by Twitter GNIP based on a query specified by researchers | 2,072,715 tweets | August 22, 2017 to August 29, 20217 |
| Gulesan et al. (2021) | Twitter API | Retrieved tweets from Twitter based on keywords | Not specified | Not specified |
| Li et al. (2021) | Twitter API | Used keywords to retrieve tweets | 30,851,895 tweets related to the pandemic lockdown and 10,082,646 tweets related to reopening | April 21, 2020 to July 21, 2020 |
| Kitazawa and Hale (2021) | Using datasets | They used a dataset created using snowball sampling | More than 5 million tweets | A period of more than six years |
| Chen and Lim (2021) | Twitter API | Typhoon-relevant information was collected using Twitter Search API | 10,996 tweets | 16:00 August 21, 2017 to 15:59 August, 30, 2017 |
| Behl et al. (2021) | Using datasets | A public dataset of Nepal 2015 and Italy 2016 earthquakes | 51,846 tweets in the Nepal earthquake dataset and 70,897 in the Italy earthquake dataset | Not specified |
| Mohanty et al. (2021) | Twitter API | Using Twitter API tweets located in the geospatial bounding box were retrieved | 784K tweets | January 09, 2017, to October 10 2017 |
| Bashir et al. (2021) | Not specified | Tweets with hashtags were harvested as a dataset | 13,156 tweets | April 4, 2017 to April 30, 2017 |
| Forati and Ghose (2022) | Twitter API | Twitter Streaming API was used to collect this dataset | 3.5 million tweets and 176,000 images | September 6, 2016 to September 21, 2017 |
| Bhullar et al. (2022) | Using datasets | A public dataset from the Nepal earthquake 2015, a dataset from the Italy earthquake 2016, and a COVID-19 dataset were used | Not specified correctly | October 26 to November 4, 2016, for the Nepal earthquake (2015), April 25 to May 5, 2015, for the Italy earthquake, and June to July 2020 for COVID-19 |

**Table 1** (continued)

| References | Data collection strategy | Dataset description | Dataset size | Temporal scope |
|---|---|---|---|---|
| Paradkar et al. (2022) | Twitter API | Tweets were collected using Twitter PowerTrack API with the authors in the affected area or with the geo-location within a bounding box | 1,121,363 tweets | 5:00 am, August 25, 2017 to 4:49 am, August 31, 2017 |
| Ngamassi et al. (2022) | Not specified | Public tweets with hashtags during the two-day window of the disaster were used | 28,041 tweets | A two-day window |
| Young et al. (2022) | Twitter API | They used publicly accessible data from Twitter and Telegram. For Twitter, they used Twitter PowerTrack API | 1,363,659 tweets | August 1, 2018 to August 23, 2018 |
| Farnaghi et al. (2020) | Twitter API | Geotagged tweets in a geographical area were collected using Twitter Streaming API | Not specified | September 12, 2018 to September 19, 2018 |
| Eachus and Keim (2020) | Not specified | Tweets from specific accounts were collected in two periods | Not specified | February 22, 2016 to February 24, 2016, and August 11, 2016, to August 15, 2016 |
| Cheong and Babcock (2021) | Using datasets | University of North Texas Libraries' Hurricane Harvey Twitter Dataset was used as the data source | 7,041,794 tweets | August 18, 2017 to September 22, 2017 |
| Contreras et al. (2022) | Using datasets | A third-party vendor collected 675 tweets, and the SM department of Newcastle University provided 1001 tweets | 675 tweets from the first source and 1001 tweets from the second source | November 26, 2019 to February 3, 2020 |
| Samuels et al. (2020) | Twitter API | Gathered the tweets using Twitter API | 1.1 million tweets | Not specified |
| Hunt et al. (2020) | Twitter API | Standard Twitter Search API was used to gather the tweets | 2,633 tweets | August 28, 2017 to September 24, 2017 |
| Wang et al. (2021) | Twitter API | Twitter standard search API and advanced search service were used to collect tweets | 983,990 tweets | February 1, 2020 to April 30, 2020 |
| Havas and Resch (2021) | Twitter API | Twitter's Rest streaming API was used to gather georeferenced tweets to gather tweets ain specific areas | Different sizes and different durations | Multiple durations |
| Kusumasari and Prabowo (2020) | Twitter API | Tweepy python library were used to collect the tweets | Different for each disaster | A period of 1-2 weeks following each disaster |
| Dong et al. (2021) | Scraping Twitter | TwitterScraper was used to collect the tweets | 41,993 tweets | Different time frames for each disaster |
| Samuels et al. (2022) | Twitter API | Geolocated Twitter data were collected using Twitter API | Not specified | Different states of the disaster have different periods |

**Table 1** (continued)

| References | Data collection strategy | Dataset description | Dataset size | Temporal scope |
|---|---|---|---|---|
| Yum (2021) | Twitter API | Tweets were collected using Twitter API based on keywords | 1,973,790 tweets | August 8, 2019 to September 21, 2019 |
| Turner-McGrievy et al. (2020) | Using third-party tools and applications; Crimson Hexagon | A third-party tool called Crimson Hexagon was used to gather the related tweets | 506,620 tweets | September 8 to 21 in different segments based on the stage of the emergency |
| Dereli et al. (2021) | Scraping Twitter | The tweets were scraped from Twitter using the Selenium library | 5201 tweets | August 28, 2008 to July 27, 2018 |
| Olynk Widmar et al. (2022) | Using third-party tools and applications; Netbase Platform | Netbase platform was used to collect tweets (Netbase also provides other social media and news data) | Not specified | September 16, 2016 to December 16, 2018 |

- *Larger sample sizes* Datasets collected from previous studies or public repositories often comprise a large volume of tweets, providing a robust foundation for analysis and generalizability.

#### 2.3.2 Limitations

- *Limited scope and applicability* Existing datasets may focus on specific events or topics, which may not align perfectly with the research objectives of a particular study.
- *Potential bias and selection effects* Datasets collected by researchers or third-party vendors may contain inherent biases or limitations based on their collection methods, filters applied, or the target population.

#### 2.3.3 Potential challenges

- *Data relevance and currency* Datasets may become outdated or lose relevance over time, especially if they were collected some time ago or cover events that have evolved significantly.
- *Data availability and access* Researchers may encounter challenges when attempting to access specific datasets due to limitations imposed by data providers, access restrictions, or copyright concerns.

### 2.4 Using third-party tools and applications

There are a number of third-party tools and applications available for researchers to access data. These third-party tools can be used with any of the data collection strategies to collect data, then filter the data, and finally provide others with the filtered data. Here are some of the tools used by researchers:

- Texifter (Discovertext 2022) is a text-mining tool that has multiple data science-related tools. This tool could be used for cleaning and analyzing unstructured texts such as tweets. Some researchers have employed Texifter's search application to collect tweets (Beedasy et al. 2020).
- Crimson Hexagon (Brandwatch 2022b) is a tool that combines a powerful social listening platform with AI (Artificial intelligence) tools that could be used as the data source for research where the social listening platform is a tool that helps companies to track mentions and conversations related to a specific topic. Researchers then use AI methods to gain insights. Marketing researchers mostly use these tools to track their brand on social media for insights into their popularity and use and devise actions that needs to be taken to improve their marketing position. Researchers can also use this platform for specific topics. This tool provides researchers with dedicated

research licenses (Brandwatch 2022a). It has been used by researchers (Wang and Guo 2021; Turner-McGrievy et al. 2020) as a data source for gathering tweets and news related to an emergency.
- Netlytic (Netlytic 2022) is a cloud-based text and social networks analyzer that could be used by researchers to source data (Whittingham et al. 2020).
- Coosto (Coosto 2022) is a content and social media marketing tool focused on Dutch messages. It was previously used (Lauran et al. 2020) to collect tweets related to a specific emergency in the Netherlands.

#### 2.4.1 Advantages

- *Streamlined data collection and analysis* Third-party tools offer integrated functionalities for data collection, filtering, and analysis, providing researchers with a comprehensive solution.
- *Advanced features and insights* These tools often incorporate AI-driven techniques for sentiment analysis, topic modeling, social network analysis, and other analytical approaches, enabling researchers to gain deeper insights from the collected data.

#### 2.4.2 Limitations

- *Cost implications* Some third-party tools may require licensing or subscription fees, which could pose financial constraints for researchers with limited budgets.
- *Dependency on tool availability* Researchers rely on the availability and reliability of third-party tools, and any changes or discontinuation of these tools could affect their research workflow.

#### 2.4.3 Potential challenges

- *Learning curve and tool proficiency* Researchers need to familiarize themselves with the functionalities, interface, and usage of third-party tools, which may require additional training or support.
- *Integration with research workflow* Integrating third-party tools seamlessly into the research process, including data collection, preprocessing, and analysis, may require adjustments and coordination with existing methodologies and tools.

## 3 Pre-processing data

The raw data collected from Twitter generally needs to be pre-processed since it is not directly usable for research purposes. Some third-party scraping tools perform the pre-processing as part of the scraping process and provide researchers with the required cleaned data. However, most researchers use various methods to process their data, regardless of whether third-party tools have pre-processed it. The pre-processing methods used by various papers are outlined in what follows. When preprocessing the data, it may be desirable to count retweets instead of removing them from the dataset. This helps identify trends that help prioritize emergency responses. Rather than excluding non-English tweets, they can instead be translated into English. For this purpose, it is recommended to seek the help of a native speaker of the detected language for an accurate translation rather than relying solely on a translating app. Furthermore, translating emoji to text can provide additional information for analysis. In what follows, we detail the pre-processing steps.

### 3.1 Location and language filter

In emergency management, it is important to filter out tweets that provide no information and tweets that are unimportant. A Location and Language filter is a basic yet important filter that significantly improves the filtering results. Based on the research purpose, some researchers focus on only one language and region since extracting information is easier in this case. The final results could be processed based on related information with the same structure and location (Pont-Sorribes et al. 2020; Bhavaraju et al. 2019; Forati and Ghose 2022; Young et al. 2022). Since tweets can be tagged based on location and language, some researchers (Pont-Sorribes et al. 2020; Forati and Ghose 2022; Young et al. 2022) filter language and location using various tools. Researchers can also do this manually after gathering the data. For example, Bec and Becken (2021) considered a rectangular bounding box to define eligible tweets. However, because not all the tweets were geotagged, they manually filtered the remaining ones. Researchers use other approaches for disasters such as storms that impact different locations on different days and hours. Researchers in Mohanty et al. (2021) used Inverse Distance Weighting (IDW) to filter based on the distance of the tweets from the disaster at the time of the tweet to find near-disaster information based on the timeline of the disaster. Researchers use filtration to consider the people who are directly involved and remove irrelevant tweets. Location Filtration is essential

in methodologies that will be analyzed based on location. This is why a portion of these papers that used location filtration are doing a Geospatial analysis (Forati and Ghose 2022; Young et al. 2022).

### 3.2 Relevance filter

It is important to remove irrelevant tweets before processing the information (Kankanamge et al. 2020). In Whittingham et al. (2020), researchers cleaned the dataset by manually removing off-topic tweets. Removing irrelevant tweets manually is time-consuming and inefficient, and effectively impossible for large datasets. Some researchers, therefore, automate the process using different tools. For example, Paradkar et al. (2022) adopted a weakly supervised fine-grained event recognition method to identify tweets discussing disaster-related topics. Latent Dirichlet Allocation (LDA), a common unsupervised learning technique for document topic discovery, can also be used to find off-topic tweets. Bhavaraju et al. (2019) used LDA to identify a broader list of disaster-related keywords for filtering tweets. They combined their keywords list with another dataset to improve filtering for crises.

### 3.3 Deduplication

Duplicate information in a dataset could result in longer algorithm execution times and impact the precision of machine learning algorithms or cause biases. Therefore, removing duplicated tweets is important. In Contreras et al. (2022), repeated tweets were removed from their database to reduce bias. A more effective method is to remove duplicated tweets and tweets with similar content. Researchers in Young et al. (2022) calculated the word frequency and cosine similarity between tweets and removed duplicates if any of the tweets were very similar to each other.

Regarding the removal of retweets, it is essential to clarify whether retweets are considered duplicates. If retweets are removed, it could potentially generate a bias instead of reducing it, as retweets can provide valuable information about the spread and impact of information during emergencies. Therefore, careful consideration is needed to decide whether to include or exclude retweets in the deduplication process.

### 3.4 Automatic agents or bots

Researchers need to remove automated tweets to find the information posted by real people on Twitter. Kankanamge et al. (2020) removed automated messages in the bot filtering process. Detecting automatic agents or bots is not always an easy task, however. In Samuels et al. (2020), they filtered out bots using keywords they built through testing with the OSoMe's Botometer tool (Davis et al. 2016). In Whittingham et al. (2020), Twitter accounts with repetitive messages, marketing information, or similar names were flagged as bots, and accounts with many tweets in a specific period were reviewed manually. In comparison, Samuels et al. (2020) used a more complex method borrowed from another research paper (Davis et al. 2016) to filter tweets, while Whittingham et al. (2020) used a relatively less complicated method implemented by themselves. Their method defined thresholds for each account's number of tweets published on specific days. After that, they manually checked the account. This method is simpler, but it is not scalable.

### 3.5 Normalizing and removing non-English content

The most common preprocessing step is normalizing the data and removing non-English content (links, hashtags, and other symbols, special characters, and platform-specific terms such as mentions). Several researchers (Gulesan et al. 2021; Li et al. 2021; Farnaghi et al. 2020; Havas and Resch 2021; Dereli et al. 2021) normalized the words by lowercasing them, removing URLs, usernames, hashtag symbols, repeated characters, punctuations, and stop words. In Bhavaraju et al. (2019), researchers removed HTML links, tags, and special characters. Zhai et al. (2020) and Kitazawa and Hale (2021) pre-processed the tweets by normalizing all characters to the lowercase format, removing special characters, spelling out every digit, and removing extra punctuation marks. Devaraj et al. (2020) removed all hashtags and non-word tokens and replaced mentions and URLs with specific words for better performance in their next steps. In Table 2, we show how researchers normalize the contents. As shown in the table, lowercasing and removing repeated characters are very common. This can help prevent detecting the same words as different words; for example, "WaTer" and "water" or "earthquake" and "earhquakeeee" in different studies need to be considered the same. This is also applied for spelling correction, stemming, and lemmatization. Removing stopwords, special characters, and URLs is used because they do not contain valuable information in some studies (Li et al. 2021). Although some studies (Kitazawa and Hale 2021) mentioned that some special characters, such as emojis, contain meaning, they also removed them to reduce complexity. Text normalizing is also helpful in other languages. The methods could be the same with small differences. Kitazawa and Hale (2021) is a research example of normalizing text in non-English content.

**Table 2** Normalization - Y represents yes, R represents replaced and D represents deleted

| Reference | Lowercased | URLs | Usernames | Digits | Hashtag symbols | Repeated characters | Punctuations | Stopwords | Special Characters | Stemming | Lemmatization |
|---|---|---|---|---|---|---|---|---|---|---|---|
| Bhavaraju et al. (2019) | | D | D | D | D | | | D | D | | |
| Zhai et al. (2020) | Y | | | D | | D | D | | D | | |
| Devaraj et al. (2020) | Y | R | R | R | R | | D | D | R | | Y |
| Gulesan et al. (2021) | Y | D | D | | D | D | D | D | | | Y |
| Li et al. (2021) | | D | D | D | | | D | D | D | Y | Y |
| Kitazawa and Hale (2021) | | D | D | D | D | | | | D | | |
| Farnaghi et al. (2020) | Y | D | | D | D | D | D | D | D | | Y |
| Havas and Resch (2021) | Y | D | | D | D | | | D | D | Y | |
| Dereli et al. (2021) | Y | | | | D | | D | D | | Y | |

### 3.6 Preprocessing the data for specific use cases

Data preprocessing in Twitter research is crucial for quality and pattern analysis. The variety of methods and techniques applied highlights its significance.

For network analysis, new data structures must be formed in the preprocessing step. For example, Lauran et al. (2020) preprocessed the data to form an edge list and node list from the author-tweet combinations based on retweets and mentions. Information can also be augmented, such as finding implicit locations from tweets and adding them as location references. In Schempp et al. (2019), researchers transformed their dataset into a geographic dataset using GIS software, adding latitude and longitude attributes to each tweet.

Researchers use tailored NLP preprocessing to prepare data for NLP approaches. This includes tokenization, lemmatization, stemming, and features such as word embedding, n-grams, and POS tags. Stemming reduces words to their word stem (e.g., "connection," "connections," "connected," and "connecting" to "connect") (Bhavaraju et al. 2019). Machine learning models require numeric data, so raw text is converted into tokens, lemmatized, and replaced with word embeddings representing their meanings (Devaraj et al. 2020; Li et al. 2021; Behl et al. 2021). Features like N-grams and POS tags are also used for text analysis. N-grams are sequences of n words in a text, while POS tags count different parts of speech for each tweet. Oversampling addresses class imbalances by duplicating minority class tweets in the training set (Devaraj et al. 2020).

Preprocessing is necessary as raw Twitter data is not suitable for research. Methods include using third-party scraping tools and manual filtering. Filtering by location and language is crucial in emergency management to retrieve relevant information. Manual filtering of irrelevant tweets is impractical for large datasets; automated methods like weakly supervised event recognition and LDA are used to identify disaster-specific keywords. Eliminating duplicate tweets is also essential to prevent bias and improve algorithm performance.

Distinguishing between automated tweets and bots requires techniques like keyword detection or manual account examination. Advanced tools are often necessary for accurate identification (Rodríguez-Ruiz et al. 2020). Data normalization and removing non-English content are common preprocessing steps, including lowercasing text and removing URLs, usernames, special characters, and stop words. Stemming, spell correction, and lemmatization are used to enhance uniformity.

Common NLP data preparation techniques include tokenization, stemming, lemmatization, and word embedding, which create numeric vectors representing word meanings. POS tags and n-grams extract significant text

features, enhancing machine learning models' generalization capabilities.

## 4 Information processing

There have been many innovative approaches for leveraging Twitter data in emergency management. We can categorize the information processing step based on techniques and methods (Saroj and Pal 2020) or based on disaster stage and statistics (Ogie et al. 2022). Previous papers have discussed emergency management based on these two approaches. This paper takes a different approach by clustering the information processing based on the source of the analysis. We classified previous works into three categories based on the resource used: content analysis, network analysis, and metadata analysis. Content-based processing focuses on the text of the tweets by directly measuring approaches such as including specific keywords or through additional text analysis using natural language processing approaches. Network analysis focuses on following the structure of Twitter messages and the user's interactions on the platform based on Twitter features such as retweets, quotes, likes, and comments. Finally, metadata analyses are based on non-content tweet data such as the time posted, location, and other metadata. This section provides examples of how researchers have analyzed Twitter data based on each analysis type by reviewing previous works.

### 4.1 Content analysis

A common content analysis approach is creating a dataset from manually coded Twitter messages (Wang and Guo 2021; Lauran et al. 2020; Wong and Jensen 2020; Bashir et al. 2021; Ngamassi et al. 2022; Young et al. 2022; Wang et al. 2021; Kusumasari and Prabowo 2020), where coding involves reading and tagging the tweets. For example, researchers in Wang and Guo (2021) coded the tweets after reading and classifying them manually. In Lauran et al. (2020), the authors coded the tweets based on a codebook they had established to code the tweets into a structured framework. Similarly, Bashir et al. (2021) and Ngamassi et al. (2022) used a similar approach by defining a coding framework. In Ngamassi et al. (2022), four researchers coded the tweets based on 25 topics and then evaluated the coding process by calculating and analyzing percentage agreements between different coders (inter-coder reliability).

Papers coding the tweets are shown in Table 3. The important limitation of manual coding is that this approach is limited to the available manpower and is therefore only scalable to a limited extent. Furthermore, the reliability and robustness of the coding process could be questioned since it is dependent on the quality of the individual coder. As Table 3 shows, most papers are coding their content for quantitative analysis.

Topic mining is an approach used by multiple researchers for content analysis. In this approach, the topics are extracted from the tweets. In Jin and Spence (2021), the researchers perform a series of topic model analyses using JMP software (JMP 2022). They generated a varimax rotated SVD of the document term matrix (DTM) to produce groups of terms known as topics. They then analyzed the topics that emerged in different emergency phases. In their research, they discovered 20 topics for various phases. In other papers, researchers use Latent Dirichlet Allocation (LDA) to find topics (Bhavaraju et al. 2019; Kitazawa and Hale 2021). For example, in Bhavaraju et al. (2019), they used LDA for topic discovery, and the results of the process were used to identify a broader list of disaster-related keywords in the corpus of documents. In Kitazawa and Hale (2021), researchers used LDA to assign each tweet topic to one of their five predefined categories. Also, LDA could be used in combination with other approaches; for example, Zhai et al. (2020) used LDA and Convolutional Neural Network (CNN) for topic discovery in emergencies. The authors mentioned an improvement in the results when using LDA with CNN.

Statistical analysis can also be performed for content analysis based on different statistics. In Schempp et al. (2019), terms such as clean water were conceptualized using static data and explicit location information to find demands and resources in disasters. They then performed a quantitative analysis based on the contents related to the demands and resources. In Turner-McGrievy et al. (2020), statistical analysis of the number of mentions in different locations was applied to find the demands in different locations. Also, researchers used text mining and word frequency calculation to find behaviors in disasters (Jamali et al. 2020). This paper combines statistical approaches with machine learning techniques to find patterns and behaviors. They also introduced an approach that could identify patterns using statistical and machine learning.

Sentiment and emotion analysis is a common approach in Twitter analysis (Bec and Becken 2021; Wong and Jensen 2020; Kankanamge et al. 2020; Zhai et al. 2020). In Bec and Becken (2021), the authors analyzed emotional stability during and after a disaster to provide insight into risk perceptions and individual responses to the disaster. In their approach, they analyzed the emotions based on manually coding the tweets. Wong and Jensen (2020) examined the interaction between trust in government, risk perceptions, and public compliance in Singapore. In this paper, the authors counted occurrences of keywords evoking negative and positive emotional valence for analysis of tweet content. In Zhai et al. (2020), the authors applied sentiment analysis using the TextBlob tool (Reynard and Shirgaokar 2019), then categorized tweets into three categories based on the

**Table 3** Coding the tweets in previous research

| Reference | Coding method | Coding process | Tweet count | Scalability | Coding usage |
|---|---|---|---|---|---|
| Wang and Guo (2021) | Manually | Two coders coded a random sample of the relevant tweets. They compared and then discussed the coding and then coded the rest | 2.8K | Not scalable | Quantitative analysis |
| Lauran et al. (2020) | Manually | Due to the limitation of manual coding, a random sample was first drawn. Then, based on a condition, some tweets were removed. Finally, tweets were coded manually | 1.3K | Not scalable | Quantitative and qualitative content analysis |
| Wong and Jensen (2020) | Manually | Several tweets were coded manually to identify sub-topics and models of impression | 100 | Not scalable | Sentiment analysis |
| Bashir et al. (2021) | Manually | Several tweets were coded manually to find the categories for sentiment analysis | 100 | Not scalable | Sentiment analysis |
| Ngamassi et al. (2022) | Hybrid; manually and LDA | Extracted topics and words related to each topic from the contents. Four researchers coded tweets independently. Then the results were compared and tested for reliability. They then discussed the results to reconcile the disagreements. Finally, they applied LDA with changed topics and words | Not specified | Scalable | Quantitative analysis |
| Young et al. (2022) | Manually | A human coder categorized the tweets. Then two researchers validated 50 tweets to verify that this coding approach is robust | Not specified | Not scalable | Quantitative analysis |
| Wang et al. (2021) | Manually | Two coders first coded a portion of the tweets. They then discussed the differences in the coding scheme using reliability testing. Finally, they coded the rest of the tweets | 1.2K | Not scalable | Quantitative analysis |
| Kusumasari and Prabowo (2020) | Manually | Tweets were coded manually into categories | Not specified | Not scalable | Quantitative analysis |

sentiment index. They then combined the results with a situational awareness classifier to form a multinomial regression model. In Bashir et al. (2021), researchers showcased the role of Twitter in a crisis by analyzing the nature of tweets and the sentiments expressed by the Twitter-sphere during and after the disaster. The researchers calculated word frequency, and then the content analysis was done manually to extract information from the tweets. Bhavaraju et al. (2019) used sentiment and proximity analysis to analyze the disaster and behavior for different periods. Sentiment analysis is not always a simple task since it depends on the context of the texts. Multiple no-code machine learning platforms provide sentiment analysis APIs, such as MonkeyLearn (Monkeylearn 2022) and Lexalytics (Lexalytics 2022). In Contreras et al. (2022), researchers used disaster data collected from Twitter to test the accuracy of a pre-trained sentiment analysis classification model developed by MonkeyLearn to identify polarity in text data.

natural language processing (NLP), machine learning (ML), and deep learning are Artificial Intelligence

methods used for content analysis. In Gulesan et al. (2021), researchers used K-Nearest Neighbors (KNN), Support Vector Machine (SVM), and the Naive Bayes algorithm to detect anomalies based on the content of the tweets. In this paper, their aim is to find the location of earthquakes as soon as possible based on the tweets, then detect demand requests and volunteer offers. Devaraj et al. (2020) aimed to identify social media-based requests for urgent help during hurricanes. They performed classifier analyses of Twitter posts using different machine learning methods and compared them. They found that the best-performing classifiers are convolutional neural networks (CNN) trained on word embeddings, support vector machines (SVM) trained on average word embeddings, and multilayer perceptrons (MLP) trained on a combination of unigrams and part-of-speech (POS) tags. Li et al. (2021) introduced a social media-based approach that quantifies the pro/anti-lockdown ratio as an indicator of the risk of adverse human interactions. They used multi-class classifiers: decision tree (DT), random forest (RF), Naïve Bayes (NB), logistic regression (LR), support vector machine (SVM), and neural network (NN) to classify the tweets. Classification is used a lot in this context. Some papers (Whittingham et al. 2020) used classification tools such as IBM Watson and mathematical models to classify the tweets. In Kumar and Singh (2019), researchers used word embedding, context-dependent feature extraction, and a deep learning approach to identify location references in the tweets. Researchers in Dereli et al. (2021) used LDA, Bag of Ngrams, sentiment, and ML classification to classify the data based on NLP approaches. Paradkar et al. (2022) also used NLP and a fine-grained analytics approach to classifying information. They used an AI-based algorithm to classify disaster tweets based on keywords within the content to identify the consistency between mentioned and embedded locations. Researchers also use deep learning methods to classify tweets. In Behl et al. (2021), researchers used Multi-Layer Perceptron (MLP) for classification. They also compared it with supervised learning classification approaches.

The volume of streaming data from Twitter makes mining information challenging (Mohanty et al. 2021). Machine learning algorithms could be used to filter information in the data-gathering phase. In Mohanty et al. (2021), the authors explored the effectiveness of data-learned approaches to mine and filter information. They developed four independent models to filter the data with the option of combining them to quickly filter and visualize tweets based on user-defined thresholds for each sub-model. They showed that this type of filtering could be useful as a base model for data capture from noisy sources such as Twitter. Also, some papers used LDA (Ngamassi et al. 2022; Havas and Resch 2021; Dereli et al. 2021) in combination with machine learning approaches to extract information from the tweets. In Ngamassi et al. (2022), authors used NLP and a fine-grained analytics approach to retrieve important information. The study explored the consistency of the locations mentioned in disaster-related tweets with their embedded geo-coordinates using LDA and manual annotations. In Havas and Resch (2021), the authors presented a methodology to identify disaster-impacted areas by combining semantic and geospatial machine learning methods. They used LDA to classify the tweets by their topic, then combined the results with the geospatial information they had to extract about the impacted spots. Chen and Lim (2021) proposed a model to focus on the in-depth understanding of social media texts while involving minimal manual efforts in a semi-supervised manner. This paper introduces a novel method to evaluate the damage extent indicated by social media texts. In the paper, the damage reports are recognized among the social media messages as damage relation categories by a multi-instance multi-class classifier. The damage extent implied by the words in a report is quantified by measuring the word similarity with the known damage-related words. In Bhullar et al. (2022), the authors combined vectorization, ML, deep learning, and time-series sentiment analysis to increase the transparency and understandability of model decisions. In Table 4, the papers using these approaches are shown. As shown in the table, multiple common themes can be identified. Also, you can see in the table that there is a correlation between the methods and goals in the papers. For example, multiple papers use machine learning to detect and identify events and locations (Kumar and Singh 2019; Paradkar et al. 2022; Ngamassi et al. 2022; Havas and Resch 2021).

Spatial analysis is another approach to disaster management. Researchers study disaster locations and impacted areas to track demands, detect events, and calculate damages and risks. Farnaghi et al. (2020) applied spatio-temporal tweet mining as a method for dynamic event extraction. In their paper, geotagged tweets were studied to evaluate the quality of the outputs based on the impact of selected embedding algorithms. In Kankanamge et al. (2020), researchers applied descriptive analysis using user, Twitter, and URL statistics. They then applied content analysis, calculating word frequency and text clustering based on the number of mentions of different locations. Finally, they combined the results to achieve a geolocation classification to generate a spatial analysis and present tables, graphs, and maps so they could better understand a given situation. In their research, they used spatial analysis to detect impacted locations. Yum (2021) calculated word frequency and statistical analysis to show how a hurricane affects spatial reactions and mobility across the US states. In their method, geotagged tweets from Twitter were gathered and analyzed using mapping information. They provided visualizations of their findings on the

**Table 4** AI methods used for information processing

| Reference | Approach | Goal |
|---|---|---|
| Whittingham et al. (2020) | Classification tools; IBM Watson | Analyzing different personal values and personality traits toward a topic |
| Kumar and Singh (2019) | Deep learning | Identify the location of the tweets |
| Devaraj et al. (2020) | Machine learning; CNN, SVM, MLP | Detect urgent help requests |
| Gulesan et al. (2021) | Machine learning; KNN, SVM, NB | Detect anomalies |
| Li et al. (2021) | Machine learning; multi-class classifiers, DT, RF, NB, LR, SVM, NN | Find the ratio of two groups of people |
| Chen and Lim (2021) | Machine learning; multi-class classifier | Evaluate the damage |
| Bhullar et al. (2022) | Machine learning, deep learning, sentiment analysis | Relief operation management |
| Paradkar et al. (2022) | NLP, fine-grained analytics | Identify the consistency between mentioned and embedded location |
| Ngamassi et al. (2022) | LDA and machine learning | Identify the consistency between mentioned and embedded location |
| Havas and Resch (2021) | LDA and machine learning | Identify the impacted spot |
| Dereli et al. (2021) | Machine learning; NB, SVM, DT, random forest, NN, KNN | In this paper, classification methods were compared |

maps for the disasters they considered. In Kumar and Singh (2019), researchers used CNN to detect locations. They used named entity recognition to detect locations based on the tweets. Kumar and Singh (2019), Paradkar et al. (2022), Ngamassi et al. (2022) used machine learning approaches to identify the locations of the tweets. In Kumar and Singh (2019), they only identified the location based on the named entities, but in Paradkar et al. (2022), Ngamassi et al. (2022), they identified the location and then compared it with the location embedded as an attribute of the tweet. Papers using spatial analysis are summarized in Table 5.

For emergency management, there are various approaches and methods used to analyze Twitter data. Content analysis, network analysis, metadata analysis, sentiment and emotion analysis, natural language processing (NLP) and machine learning (ML), as well as spatial analysis, categorize the analysis based on its source.

Content analysis involves manually coding tweets based on a predefined framework or using topic-mining techniques. Researchers developing content analysis used coding frameworks, topic model analysis, and statistical analysis to extract information from tweets. Sentiment and sentiment analysis is also used to analyze the emotional responses of Twitter users during and after disasters. NLP, ML, and deep learning techniques are used for content analysis, including anomaly detection, identification of emergency help requests, and sentiment analysis. Spatial analysis focuses on the geographic aspects of Twitter data, including event extraction, geolocation classification, and the impact of disasters on spatial response and mobility.

Overall, a variety of methods and techniques are used to extract valuable information from Twitter data for emergency management purposes. These methods range from manual coding and statistical analysis to advanced AI-based methods such as NLP, ML, and deep learning. Spatial analysis enables researchers to track disaster locations and estimate the impact of events on specific areas. By using these different approaches, researchers can gain insight into user behavior, emotions, and spatial patterns during emergencies, leading to better decision-making and response planning.

**Table 5** Spatial analysis

| Reference | Metadata used | Approach |
|---|---|---|
| Kumar and Singh (2019) | CNN | Detect locations |
| Kankanamge et al. (2020) | Quantitative analysis and classification | Identify impacted location |
| Paradkar et al. (2022) | NLP, fine-grained analytics | Identify the consistency between mentioned and embedded location |
| Ngamassi et al. (2022) | LDA and machine learning | Identify the consistency between mentioned and embedded location |
| Farnaghi et al. (2020) | Cluster detection and cluster linking | Dynamic event extraction |
| Havas and Resch (2021) | LDA and machine learning | Identify the impacted spot |
| Yum (2021) | Qualitative analysis | Identify the effects on spatial reactions and mobility |

### 4.2 Network analysis

Using various network attributes for social media is another method for analyzing Twitter data. These attributes are analyzed based on graphs that can be generated based on the data from Twitter. On Twitter, tweets are posted, re-posted (retweets), or re-posted with extra information (quotes) by individuals. Individuals can comment on a post using Twitter, and also they can follow other individuals and accounts. The interactions mentioned in the tweets form graphs that can be analyzed with graph analysis techniques.

Lauran et al. (2020) analyzed whether and how different groups of stakeholders frame a crisis and the extent to which they attribute responsibility for the crisis to actors. They preprocessed tweets to generate edge and node lists that can be used for graph analysis. They produced a node list containing authors and their information and an edge list based on retweets, mentions, and author-tweet combinations. They used the OpenOrd algorithm (Martin et al. 2011) to cluster the groups based on the frequency of the interactions. In this paper, they manually coded the tweets to find patterns in different groups and clusters.

In Pourebrahim et al. (2019), following a survey of the general population and exploring communication dynamics on Twitter, researchers investigated Twitter usage during Hurricane Sandy. The investigation was conducted on the graph properties of the social networks used by the general population. Social network analysis uses the mathematics of graph theory to understand social phenomena through the relationship between people, groups, and things (Hansen et al. 2011). A social network is represented by a graph of nodes and links, where nodes are individual actors and links are social ties, relationships, exchanges, or interactions among actors Chatfield and Brajawidagda (2012). They performed a user network analysis based on the metadata of user mentions. They generated a directed graph for each mention of the user from the user to the mentioned user. Finally, they calculated different network metrics, and they also divided the groups of users into clusters.

Beedasy et al. (2020) used a mixed-methods approach, including human coding, machine learning, and social network analysis. This research generated a directed graph from the Twitter data related to the Deepwater Horizon oil spill. They identified nodes' positions and relative importance in the network structure and the subgroups or clusters. They visualized the network data to present the patterns and connections within the dataset in a more readable and interpretable format. They used the results from the analysis and other methods to gain insights into the content and flow of the tweets that had shaped the conversation related to the oil spill.

### 4.3 Metadata analysis

The final method for analyzing Twitter data is by using the metadata and non-content information accompanying a tweet, such as tweet time, location, language, and other Twitter metrics such as likes count, retweets count, and users' follower and following counts. Here we summarize the previous work in emergency management that analyzed the tweets based on the metadata.

Yue et al. (2021) applied spatial analysis based on geotagged tweets. They used tweets with geographical coordinate information and analyzed the locations of the tweets related to the disaster. They presented a proof of concept for assessing wildfire risk using geo-tagged social media data by taking wildfire risk as a function of wildfire hazard and social-ecological vulnerability. In Havas and Resch (2021), researchers defined topics and then used the location data to monitor social media to extract information. They presented an improved version of a recently proposed methodology to identify disaster-impacted areas by combining semantic and geospatial machine learning methods.

Lucas and Landman (2021) and Pont-Sorribes et al. (2020) used Twitter metrics for analysis. Lucas and Landman (2021) presented a social listening framework as a toolkit for revealing and understanding insights into public awareness and engagement building during disasters using hashtags and retweeting metadata. Pont-Sorribes et al. (2020) investigated Twitter communications by the media and institutions regarding two emergencies.

Thekdi and Chatterjee (2019) used a novel method. They analyzed the infrastructure state and network conditions using Twitter in different areas. They developed methodologies for assessing community perceptions of infrastructure system resilience using observable factors and inferring hidden performance measures to facilitate adaptive decision-support systems. Researchers also studied the correlations between Twitter activity and damage in specific disaster locations (Samuels et al. 2020, 2022). These papers used Twitter usage metadata to show the damages.

Some papers used novel approaches to extract information and analyze the data. Some used statistical approaches (Pelfrey 2021; Kusumasari and Prabowo 2020) and other defined metrics related to tweets like attention, persistence, speed, and reach (Cheong and Babcock 2021) and analyzed the data based on these metrics. Also, Hunt et al. (2020) defined metrics such as lifespan, debunking, and cross-platform sourcing for each tweet to extract information based on them. In Yang et al. (2019), the researchers identified the general relationship between followers and tweet content. In Table 6, these papers are shown in detail. The common theme is using quantitative analysis, even using metadata.

**Table 6** Metadata analysis

| Reference | Metadata used | Approach | Goal |
|---|---|---|---|
| Yue et al. (2021) | GeoLocation | Spatial analysis | Analyze the location of the tweets related to disaster |
| Lucas and Landman (2021) | Hashtags and retweeting metadata | Quantitative analysis | Getting insights |
| Thekdi and Chatterjee (2019) | Twitter usage | Quantitative analysis | Finding the correlations between Twitter usage and infrastructure system resilience |
| Yang et al. (2019) | Followers and number of tweets | Machine learning; negative binomial and Random forest regression | Analyzing the emergence of influencers |
| Cheong and Babcock (2021) | Custom metrics; attention, persistence, speed, and reach | Quantitative analysis | Analyzing popular but ambiguous conversations |
| Samuels et al. (2020) | Twitter usage | Spatial and quantitative analysis | Finding the correlations between Twitter activity and damage |
| Hunt et al. (2020) | Custom metrics; lifespan, debunking, and cross-platform sourcing | Quantitative analysis | Analyzing misinformation spread |
| Havas and Resch (2021) | GeoLocation | Machine learning | Identify impacted areas |
| Kusumasari and Prabowo (2020) | Twitter usage | Quantitative analysis | Patterns of Twitter usage and its effects |
| Samuels et al. (2022) | GeoLocation and Twitter usage | Quantitative analysis | Finding the correlations between Twitter activity and damage |

Common metadata information used in papers is Twitter usage and the location of the tweets.

## 5 Discussion

This paper aims to advance emergency management by analyzing previous research on using social media, specifically Twitter, as a source of information during emergencies. By examining the different sources of information used in previous research, this paper provides insights into the potential for using Twitter as a communication and information-sharing platform in times of crisis. This can inform future researchers on how social media can be utilized for emergency management and response, ultimately helping to improve disaster preparedness and response efforts.

One key insight from this analysis is the importance of aligning data collection and processing methodologies with disaster management goals. Different types of data and methods can introduce confusion and complexity. Researchers should leverage this paper to gain insights into various information sources, thus reducing confusion and complexity. Understanding the advantages, limitations, and potential biases of each data source and method is crucial. For instance, the chosen data collection and preprocessing steps can significantly influence the outcomes, potentially causing critical data to be missed due to specific filters or keywords.

This paper also highlights the need for a comprehensive approach to analyzing Twitter data, integrating content, network, and metadata analyses. Each type of analysis offers unique insights that can contribute to a more holistic understanding of emergency situations. For instance, content analysis can reveal public sentiment and key topics of discussion, network analysis can map information dissemination and influential actors, and metadata analysis can provide temporal and spatial context.

By providing a structured review of the methodologies and their applications in disaster management, this paper helps researchers navigate the complex landscape of social media data analysis. It encourages the adoption of best practices and innovative approaches that can enhance the effectiveness of emergency response and management. The findings not only interpret the results but also connect them back to the broader research context, elucidating the implications for both theory and practical application in emergency management.

## 6 Future work

Despite significant progress in using Twitter for emergency management, there are still several areas that require further research and development. Potential future research directions in this field include:

*Auto-encoding* Manually encoding tweets for content analysis can be time-consuming and has limited scalability.

Future work may focus on developing autoencoding techniques using natural language processing (NLP) and machine learning (ML) algorithms. These technologies can help automate the process of categorizing and analyzing tweets based on predefined codes or topics, reducing reliance on human programmers and improving scalability.

*Real-time information processing* Many existing studies focus on analyzing historical Twitter data for emergency management. Future work may explore real-time information processing techniques to extract timely and relevant information from Twitter during ongoing emergencies. This could include developing algorithms and models that can detect and filter tweets about specific events or situations in real-time, providing emergency responders with more immediate and actionable insights.

*Multimodal analysis* While this discussion focuses on content analysis, future work may explore the integration of other modalities such as imagery, video, and geolocation data into emergency management. Combining text analytics with image and video analytics techniques provides a more complete picture and the ability to extract valuable information from a variety of data sources.

*Deep learning methods* The papers discussed briefly mentioned the use of deep learning techniques for content analysis in emergency management. Future work may delve deeper into applying deep learning models such as recurrent neural networks (RNNs) or transformers to tasks such as sentiment analysis, topic modeling, and event detection in Twitter data. These advanced models have achieved promising results in natural language processing, which can improve the accuracy and effectiveness of information processing in emergency management.

*Context analysis* Twitter messages often contain implicit information and contextual clues that may affect their interpretation and relevance in emergency management. Future work may focus on developing methods to capture and analyze contextual information in tweets, such as user profiles, social network connections, and temporal dynamics. Incorporating contextual analysis can help improve the accuracy and precision of information extraction and provide a more nuanced understanding of changing emergencies.

*Evaluation and validation* Evaluating and validating the effectiveness of different information processing techniques and models in business continuity management is crucial. Future work may include comparative studies to evaluate the performance and robustness of different methods, assessing their accuracy, scalability and reliability. In addition, exploring the integration of human-in-the-loop methods, where human experts verify and validate the results of automated analyses, can help ensure the quality and trustworthiness of the extracted information.

*Privacy and ethical considerations* As with any research on social media data, future work should address privacy and ethical concerns. Developing frameworks and policies for handling sensitive user information while respecting the right to privacy is critical. Researchers should also consider potential bias and ethical implications of using social media data and ensure fairness and transparency in data collection, analysis, and reporting.

By exploring these future directions, researchers can further improve the use of Twitter data in emergency management, enabling more effective and efficient response strategies in times of crisis.

## 7 Conclusions

People engage with social media more intensely than ever before. People normally share emotions, opinions, images, audio files, and videos with their circle of people on social media. During an emergency, people use social media even more. They share first-hand experiences through their social networks before traditional media. This information could be used for emergency management. This information contains safety and security states, infrastructure and facilities states in the impacted area, food and beverage needs, help needed in different locations, and other useful information that could be helpful in emergencies. Although the information on Twitter may be inaccurate and false or include rumors, it is over-empowering, especially in emergencies. In some cases, Twitter is the best and only source of information that could be used. In this literature review, we divided the content into three main sections. First, we discussed the different approaches to data collection from Twitter. This section mentioned different approaches researchers used to gather the information. Then in the next section, we discussed previous works' different data pre-processing approaches. The researchers cleaned and changed the raw data for information processing based on the methodology. Finally, we discussed how the information is processed and how the researchers use it. In this section, different approaches and methods are presented. This paper underlines the importance of Twitter in emergency management and presents a basic structure for using Twitter in emergencies. We hope this paper can be used as a reference to the research surrounding disasters.

**Author contributions** A.A. wrote the manuscript and prepared all tables. J.R. and R.A. reviewed the manuscript and applied changes.

**Funding** The authors declare that no funds, grants, or other support were received during the preparation of this manuscript.

## Declarations

**Conflict of interest** The authors have no relevant financial or non-financial interests to disclose.

## References


Bashir S, Bano S, Shueb S et al (2021) Twitter chirps for syrian people: sentiment analysis of tweets related to syria chemical attack. Int J Disaster Risk Reduct 62(102):397

Bec A, Becken S (2021) Risk perceptions and emotional stability in response to cyclone debbie: an analysis of twitter data. J Risk Res 24(6):721–739

Beedasy J, Zuniga AFS, Chandler T et al (2020) Online community discourse during the deepwater horizon oil spill: an analysis of twitter interactions. Int J Disaster Risk Reduct 51(101):870

Behl S, Rao A, Aggarwal S et al (2021) Twitter for disaster relief through sentiment analysis for covid-19 and natural hazard crises. Int J Disaster Risk Reduct 55(102):101

Bhavaraju SKT, Beyney C, Nicholson C (2019) Quantitative analysis of social media sensitivity to natural disasters. Int J Disaster Risk Reduct 39(101):251

Bhullar G, Khullar A, Kumar A et al (2022) Time series sentiment analysis (sa) of relief operations using social media (sm) platform for efficient resource management. Int J Disaster Risk Reduct 75(102):979

Brandwatch (2022a) Brandwatch-for-research. https://www.brandwatch.com/p/brandwatch-for-research/

Brandwatch (2022b) Crimson-hexagon. https://www.brandwatch.com/p/crimson-hexagon/

Chatfield A, Brajawidagda U (2012) Twitter tsunami early warning network: a social network analysis of twitter information flows. University of Wollongong, Tech. rep

Chen Z, Lim S (2021) Social media data-based typhoon disaster assessment. Int J Disaster Risk Reduct 64(102):482

Cheong SM, Babcock M (2021) Attention to misleading and contentious tweets in the case of hurricane harvey. Nat Hazards 105(3):2883–2906

Contreras D, Wilkinson S, Alterman E et al (2022) 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. Nat Hazards 113:1–19

Coosto (2022) Coosto. https://www.coosto.com/en

Davis CA, Varol O, Ferrara E, et al. (2016) Botornot: A system to evaluate social bots. In: Proceedings of the 25th international conference companion on world wide web, pp 273–274

Dereli T, Eligüzel N, Çetinkaya C (2021) Content analyses of the international federation of red cross and red crescent societies (ifrc) based on machine learning techniques through twitter. Nat Hazards 106(3):2025–2045

Devaraj A, Murthy D, Dontula A (2020) Machine-learning methods for identifying social media-based requests for urgent help during hurricanes. Int J Disaster Risk Reduct 51(101):757

Discovertext (2022) Discovertext. https://discovertext.com/

Dong ZS, Meng L, Christenson L et al (2021) Social media information sharing for natural disaster response. Nat Hazards 107(3):2077–2104

Eachus JD, Keim BD (2020) Content driving exposure and attention to tweets during local, high-impact weather events. Nat Hazards 103(2):2207–2229

Farnaghi M, Ghaemi Z, Mansourian A (2020) Dynamic spatio-temporal tweet mining for event detection: a case study of hurricane florence. Int J Disaster Risk Sci 11(3):378–393

Forati AM, Ghose R (2022) Examining community vulnerabilities through multi-scale geospatial analysis of social media activity during hurricane irma. Int J Disaster Risk Reduct 68(102):701

Gaisbauer F, Pournaki A, Banisch S, et al. (2009) How twitter affects the perception of public opinion: two case studies. arXiv preprint arXiv

Gulesan OB, Anil E, Boluk PS (2021) Social media-based emergency management to detect earthquakes and organize civilian volunteers. Int J Disaster Risk Reduct 65(102):543

Hansen DL, Shneiderman B, Smith M, et al. (2011) Social network analysis: measuring, mapping, and modeling collections of connections. Analyzing social media networks with NodeXL: insights from a connected world. Burlington: Elsevier Inc

Havas C, Resch B (2021) Portability of semantic and spatial-temporal machine learning methods to analyse social media for near-real-time disaster monitoring. Nat Hazards 108(3):2939–2969

Huang D, Wang S, Liu Z (2021) A systematic review of prediction methods for emergency management. Int J Disaster Risk Reduct 62(102):412

Hunt K, Wang B, Zhuang J (2020) Misinformation debunking and cross-platform information sharing through twitter during hurricanes harvey and irma: a case study on shelters and id checks. Nat Hazards 103(1):861–883

Internetlivestats (2021) Twitter usage statistics. https://www.internetlivestats.com/twitter-statistics/

Jamali M, Nejat A, Moradi S et al (2020) Social media data and housing recovery following extreme natural hazards. Int J Disaster Risk Reduct 51(101):788

Jin X, Spence PR (2021) Understanding crisis communication on social media with cerc: topic model analysis of tweets about hurricane maria. J Risk Res 24(10):1266–1287

JMP (2022) Jmp. https://www.jmp.com/en_us/software/predictive-analytics-software.html

Kankanamge N, Yigitcanlar T, Goonetilleke A et al (2020) Determining disaster severity through social media analysis: testing the methodology with south east queensland flood tweets. Int J Disaster Risk Reduct 42(101):360

Kitazawa K, Hale SA (2021) Social media and early warning systems for natural disasters: a case study of typhoon etau in japan. Int J Disaster Risk Reduct 52(101):926

Kumar A, Singh JP (2019) Location reference identification from tweets during emergencies: a deep learning approach. Int J Disaster Risk Reduct 33:365–375

Kusumasari B, Prabowo NPA (2020) Scraping social media data for disaster communication: how the pattern of twitter users affects disasters in asia and the pacific. Nat Hazards 103(3):3415–3435

Lauran N, Kunneman F, Van de Wijngaert L (2020) Connecting social media data and crisis communication theory: a case study on the chicken and the egg. J Risk Res 23(10):1259–1277

Lexalytics (2022) Lexalytics. https://www.lexalytics.com/technology/sentiment-analysis/

Li L, Ma Z, Lee H et al (2021) Can social media data be used to evaluate the risk of human interactions during the covid-19 pandemic? Int J Disaster Risk Reduct 56(102):142

Long CR, Stewart MK, Cunningham TV et al (2016) Health research participants' preferences for receiving research results. Clin Trials 13(6):582–591

Lucas B, Landman T (2021) Social listening, modern slavery, and covid-19. J Risk Res 24(3–4):314–334

Martin S, Brown WM, Klavans R, et al. (2011) Openord: an open-source toolbox for large graph layout. In: Visualization and data analysis 2011, SPIE, pp 45–55

Mohanty SD, Biggers B, Sayedahmed S et al (2021) A multi-modal approach towards mining social media data during natural disasters-a case study of hurricane irma. Int J Disaster Risk Reduct 54(102):032

Monkeylearn (2022) Monkeylearn. https://monkeylearn.com/sentiment-analysis-online/

Murakami A, Nasukawa T, Watanabe K et al (2019) Understanding requirements and issues in disaster area using geotemporal visualization of twitter analysis. IBM J Res Dev 64(1/2):10–1

Netlytic (2022) Netlytic. https://netlytic.org/index.php

Ngamassi L, Shahriari H, Ramakrishnan T et al (2022) Text mining hurricane harvey tweet data: lessons learned and policy recommendations. Int J Disaster Risk Reduct 70(102):753
Ogie R, James S, Moore A et al (2022) Social media use in disaster recovery: a systematic literature review. Int J Disaster Risk Reduct 70:102783
Olynk Widmar N, Rash K, Bir C et al (2022) The anatomy of natural disasters on online media: hurricanes and wildfires. Nat Hazards 110(2):961–998
Paradkar AS, Zhang C, Yuan F et al (2022) Examining the consistency between geo-coordinates and content-mentioned locations in tweets for disaster situational awareness: A hurricane harvey study. Int J Disaster Risk Reduct 73(102):878
Pelfrey WV (2021) Emergency manager perceptions of the effectiveness and limitations of mass notification systems: a mixed method study. J Homel Secur Emerg Manage 18(1):49–65
Pewresearch (2021) Usagesocial media use in 2021. https://www.pewresearch.org/internet/2021/04/07/social-media-use-in-2021/
Pont-Sorribes C, Suau-Gomila G, Percastre-Mendizábal S (2020) Twitter as a communication tool in the germanwings and ebola crises in europe: analysis and protocol for effective communication management. Int J Emerg Manage 16(1):22–40
Pourebrahim N, Sultana S, Edwards J et al (2019) Understanding communication dynamics on twitter during natural disasters: a case study of hurricane sandy. Int J Disaster Risk Reduct 37(101):176
Rachunok B, Bennett J, Flage R et al (2021) A path forward for leveraging social media to improve the study of community resilience. Int J Disaster Risk Reduct 59(102):236
Reynard D, Shirgaokar M (2019) Harnessing the power of machine learning: can twitter data be useful in guiding resource allocation decisions during a natural disaster? Transp Res Rart D Transp Environ 77:449–463
Rodríguez-Ruiz J, Mata-Sánchez JI, Monroy R et al (2020) A one-class classification approach for bot detection on twitter. Comput Secur 91(101):715
Samuels R, Taylor JE, Mohammadi N (2020) Silence of the tweets: incorporating social media activity drop-offs into crisis detection. Nat Hazards 103(1):1455–1477
Samuels R, Xie J, Mohammadi N et al (2022) Tipping the scales: how geographical scale affects the interpretation of social media behavior in crisis research. Nat Hazards 112:1–20
Saroj A, Pal S (2020) Use of social media in crisis management: a survey. Int J Disaster Risk Reduct 48(101):584
Schempp T, Zhang H, Schmidt A et al (2019) A framework to integrate social media and authoritative data for disaster relief detection and distribution optimization. Int J Disaster Risk Reduct 39(101):143
Southern MG (2022) Elon Musk's twitter takeover: a timeline of events. Search Engine J
Thekdi SA, Chatterjee S (2019) Toward adaptive decision support for assessing infrastructure system resilience using hidden performance measures. J Risk Res 22(8):1020–1043
Turner-McGrievy G, Karami A, Monroe C et al (2020) Dietary pattern recognition on twitter: a case example of before, during, and after four natural disasters. Nat Hazards 103(1):1035–1049
Twitter (2022) About the twitter api. https://developer.twitter.com/en/docs/twitter-api/getting-started/about-twitter-api
Wang B, Liu B, Zhang Q (2021) An empirical study on twitter's use and crisis retweeting dynamics amid covid-19. Nat. Hazards 107(3):2319–2336
Wang W, Guo L (2021) Benefits and risks of genetically modified mosquitoes: news and twitter framing across issue-attention cycle. J Risk Res 24(9):1086–1100
Whittingham N, Boecker A, Grygorczyk A (2020) Personality traits, basic individual values and gmo risk perception of twitter users. J Risk Res 23(4):522–540
Wong CML, Jensen O (2020) The paradox of trust: perceived risk and public compliance during the covid-19 pandemic in singapore. J Risk Res 23(7–8):1021–1030
Xu Z (2020) How emergency managers engage twitter users during disasters. Online Inf Rev 44(4):933–950
Yang Y, Zhang C, Fan C et al (2019) Exploring the emergence of influential users on social media during natural disasters. Int J Disaster Risk Reduct 38(101):204
Young JC, Arthur R, Spruce M et al (2022) Social sensing of flood impacts in india: a case study of kerala 2018. Int J Disaster Risk Reduct 74(102):908
Yue Y, Dong K, Zhao X et al (2021) Assessing wild fire risk in the united states using social media data. J Risk Res 24(8):972–986
Yum S (2021) The effects of hurricane dorian on spatial reactions and mobility. Nat Hazards 105(3):2481–2497
Zhai W, Peng ZR, Yuan F (2020) Examine the effects of neighborhood equity on disaster situational awareness: harness machine learning and geotagged twitter data. Int J Disaster Risk Reduct 48(101):611