

## Secondary Publication



Zander, K.K.; Nguyen, D.; Mirbabaie, Milad; Garnett, S.T.

### **Aware but not prepared : understanding situational awareness during the century flood in Germany in 2021**

Date of secondary publication: 12.09.2025

Version of Record (Published Version), Article

Persistent identifier: urn:nbn:de:bvb:473-irb-110365x

#### **Primary publication**

Zander, K.K.; Nguyen, D.; Mirbabaie, Milad; Garnett, S.T. (2023): Aware but not prepared : understanding situational awareness during the century flood in Germany in 2021, in: International journal of disaster risk reduction : IJDRR, Amsterdam: Elsevier, Vol. 96, Nr. 103936, pp. 1–14, doi: 10.1016/j.ijdr.2023.103936.

#### **Legal Notice**

This work is protected by copyright and/or the indication of a licence. You are free to use this work in any way permitted by the copyright and/or the licence that applies to your usage. For other uses, you must obtain permission from the rights-holders.

This document is made available under a Creative Commons license.



The license information is available online:

<https://creativecommons.org/licenses/by-nc-nd/4.0/legalcode>

Contents lists available at [ScienceDirect](https://www.sciencedirect.com)

## International Journal of Disaster Risk Reduction

journal homepage: [www.elsevier.com/locate/ijdr](http://www.elsevier.com/locate/ijdr)

# Aware but not prepared: understanding situational awareness during the century flood in Germany in 2021

K.K. Zander<sup>a,\*</sup>, D. Nguyen<sup>a</sup>, Milad Mirbabaie<sup>b</sup>, S.T. Garnett<sup>c</sup>

<sup>a</sup> Northern Institute, Charles Darwin University, Australia

<sup>b</sup> Paderborn University, Germany

<sup>c</sup> Research Institute for the Environment and Livelihoods, Charles Darwin University, Australia

## ARTICLE INFO

### Keywords:

Climate change impact  
Evacuation  
Machine learning  
Pluvial floods  
Social media  
Topic modelling  
Twitter  
Warnings

## ABSTRACT

In July 2021, intense rainfall in parts of Western Europe was followed by unprecedented flash flooding. The flooding killed many people, mostly in Germany, which was subsequently blamed on a lack of preparation by either authorities or the broader populace. Knowledge about people's awareness of a hazard, its impacts and possible adaptation can help improve hazard management. We used social media data to assess situational awareness, sentiments and behaviour before, during and after the flood. We analysed nearly 58,000 German Twitter (now X) messages about the flood using machine learning-based unsupervised topic modelling. We showed that message frequency translated into four phases with message content suggesting sender priorities shifted in each phase of the flood. Besides messages with weather updates, correlated topics included 'Solidarity, recovery and aid' and 'Grief and empathy' (together 22% of all tweets) and a set of four topics about climate change attribution, extreme weather, long-term flood protection measures and politics (together 38% of all tweets). Many topics depended on the phase of the flood and on the distance from the affected areas. For those near the affected areas, tweets about evacuation (in peak phase) and damage assessment (in recovery phase) were particularly prominent. The high engagement on Twitter to seek weather information might indicate awareness of the extreme climatic conditions before the flood, but the severity was unexpected. Social media data provided unprecedented citizen-generated real-time information. We discuss how rapid automated analysis could contribute to disaster communication and risk mitigation.

## 1. Introduction

Floods are among the most devastating of natural hazards, causing death and injury, social disruption and damage to infrastructure and ecosystems [1]. Globally they have accounted for nearly 40% of natural disasters since 2000, affecting nearly 100 million people [2]. Floods are predicted to become more frequent and more severe as rainfall becomes more erratic [3]. Besides coastal and river floods, intense rainfall cause severe floods in urban areas (pluvial or flash floods). Damage from flash floods will increase as urban populations grow exposing, a larger share of people to flood risk [4]. Additional flood management and protection is needed to counteract potential damage [4,5], particularly in highly urbanised countries.

There is a large body of literature on how floods are being managed, often from a top-down perspective, involving physical flood protection (e.g. Ref. [6]), improvements in flood monitoring and prediction (e.g. Ref. [7]) or insurance schemes and policies that aim to

\* Corresponding author.

E-mail addresses: [kerstin.zander@cdu.edu.au](mailto:kerstin.zander@cdu.edu.au) (K.K. Zander), [duy.nguyen@cdu.edu.au](mailto:duy.nguyen@cdu.edu.au) (D. Nguyen), [milad.mirbabaie@uni-paderborn.de](mailto:milad.mirbabaie@uni-paderborn.de) (M. Mirbabaie), [stephen.garnett@cdu.edu.au](mailto:stephen.garnett@cdu.edu.au) (S.T. Garnett).

<https://doi.org/10.1016/j.ijdr.2023.103936>

Received 23 May 2023; Received in revised form 3 August 2023; Accepted 4 August 2023

Available online 7 August 2023

2212-4209/© 2023 The Authors. Published by Elsevier Ltd. This is an open access article under the CC BY-NC-ND license (<http://creativecommons.org/licenses/by-nc-nd/4.0/>).

share the risks [8,9]. However, people lie at the heart of flood management, and flood research often overlooks the social dimensions of floods, such as the public's understanding of risk and the behaviour of people during flood events [10]. An emerging branch of research therefore investigates how people perceive flood risk, under the assumption that risk perception heightens awareness which is crucial for flood risk management and preparedness [11,12].

Many communities and people are not well prepared for floods. As a result, extreme rain events can turn unexpectedly into disasters, as happened in Germany in July 2021. The flood in July 2021 in Western Europe was one of the five costliest natural disasters in Europe in the last half century with the estimated total losses of EUR 32 billion [13]. Germany and the Benelux countries suffered the greatest infrastructure damage and fatalities with over 200 people losing their lives in the floods; 184 in Germany alone [14]. Damages to roads and railways left many places inaccessible for days, compromising emergency responses and evacuations [15,16].

Climate change, which made the heavy summer rainfall between 3% and 19% stronger, and 1.2 to nine times more likely [14,15], only partly explains why the floods had such a disastrous impact. As least as important was the low level of risk awareness and preparedness and a lack of timely warnings [17]. That many people were surprised by the flood in July 2021, despite there having been other severe pluvial floods in parts of Europe, including Germany, in the previous decade [8,18], is shown by the dearth of people in Germany with flood damage insurance [19]. Hubris may also have contributed; increasing dependence of German society on its infrastructure and emergency systems may mean the public expected planning and contingency plans to be of a quality that avoids such impacts [20].

Given the rate of climate change, there is now an urgent need to understand how people perceive flood risk, how they access warnings and communicate them, and how they respond, immediately and in preparation for future flood events if similar catastrophes are to be averted (e.g. Ref. [21]). Previous flood-related research from Germany and neighbouring countries has already investigated how communities and individuals prepare themselves for future flood hazards, both in areas that are prone to floods and in places may flood under exceptional circumstances ([19]22–24). There is also a growing literature on how people in this region respond to flood warnings and emergency plans [18,20]. Many studies have used survey instruments and interview people in flood affected or flood prone areas.

However, surveying people in stress situations and collecting timely information while an emergency is unfolding is challenging [25] and probably unethical. Surveys in disaster situations can also lead to biased responses and are unlikely to reflect the correct overall situation [26]. However, surveys subsequent to an event must contend with the inevitable biases of hindsight [27]. One new and innovative alternative data source, which can provide a much bigger picture than surveys alone, is social media data. The collection of social media data does not involve respondents directly (it is sometimes referred to as passive data collection) so does not require them to self-report their recollections of a disaster such as a flood. Social media can nevertheless provide researchers with real-time data from both individual private users, which are mostly used to understand public reactions, emotions and opinions [28], and from official and non-official group and government channels. These channels have become important means for disseminating information rapidly and in real-time about disasters and are often used both formally by official agencies and informally by private citizens and non-government organisations to warn and advise people at risk [29,30].

In this study, we took advantage of the unprecedented scale of social media conversations during the flood in Germany in July 2021. We collected and analysed Twitter data with the aim of understanding 1) if and when twitter was used to talk about the flood; 2) citizens' perceptions of flood risk in areas in Germany affected by the 2021 flood and across Germany in general; 3) the content of the social media use in this emergency situation; and 4) emotions and sentiments and how these varied as the flood progressed.

Our analysis of Twitter data can provide insights on how people in Germany utilise social media to inform themselves about natural hazards such as the 2021 flood and how they disseminate information. The abundance of citizen-generated social media data is only useful when analysed and digested in a way that is informative for decision-makers. The findings can be used during future floods by different actors to reduce disaster risk, for example, by emergency responders, providers and managers of warning systems and local governments and agencies in adopting long-term flood mitigation adjustments. In Germany, for example, rapidly analysed social media data might have been very helpful to local branches of German Red Cross or THW [Federal Agency of Technical Relief] who were heavily involved during and after the floods in 2021. Policies that might arise from our findings will eventually help flood disaster managers enhance community resilience and reduce the negative consequences to society by improving flood planning and risk mitigation.

Our study complements existing studies using natural language processing about recent German floods, including that of 2021. Studies that did not use machine learning included a network analysis of Twitter data following heavy rain in Germany during the summer of 2019 to identify communication patterns among various Twitter users, such as public institutions, interest groups, media, experts, weather portals and the general public [31] and text mining of digital newspaper articles to understand the 2021 floods thematically and to compare them with historical flood events [32]. In an earlier study [33] researcheremployed topic modelling to analyse approximately 35,000 tweets to assess volunteer networks following severe flooding of the river Elbe in Saxony in 2013, as well as a similar number of posts on Facebook. Most recently [34], used topic modelling to analyse a relatively small sample of tweets (4,842) on the same German flood we examine (of July 2021) in which they identified five themes related to disaster resilience capacity (preventive, anticipative, absorptive, adaptive, and transformative). Our work expands and complements this study by examining a much larger sample of tweets (55,000) collected over a longer period (1–31 July) which allows us to identify a greater diversity of relevant topics, including conversations that occurred before the flood arrived.

## 2. Methods

### 2.1. Flood event and research area

The floods in 2021 affected multiple countries, including Germany, the Netherlands, Luxembourg and Belgium [14]. Our research focused on Germany, particularly the two German states of North-Rhine Westphalia and Rhineland-Palatinate which were particularly badly affected. The flood in Germany occurred on 14–15 July 2021. The summer leading up to the flood had been exceptionally dry, with little rain falling during June or the first two weeks of July. Then a large slow-moving summer storm system discharged up to 150 mm of rain within 48 h, causing the rivers Ahr and Erft to sweep away buildings, streets and bridges, many of them several centuries old. More than 180 people died and the livelihoods of 40,000 people were badly affected [14,20].

### 2.2. Data collection and processing

We used a scraper built for the Python environment to download Twitter data (snsrape). The scraper allows retrieval of historical tweets based on a series of keywords or hashtags. Hashtags are words or terms preceded by a hash symbol (#). Hashtags are used within tweets to draw attention to a certain topic and to help others to find certain topics and keywords [35]. We only downloaded tweets in German that were sent during July 2021.

To capture tweets related to the flood during this 31-day period, we searched the following hashtags and keywords.

Hashtags: #flut, #flutwelle, #hochwasser, #Flutkatastrophe, #Ahrflut, #Ahrhochwasser, #jahrhunderthochwasser

Keywords: starkregen, sturzflut

Tweets were only downloaded if they included at least one of the keywords or hashtags. The keywords and hashtags were chosen based on a rapid assessment of German tweets and hashtags that had been used to converse about the flood. Previous studies also used similar context-specific hashtags [35]. Besides the actual tweet content, we also downloaded the date of the tweet, users' self-declared locations and, if available, the coordinates of the actual locations from which the tweets were posted. To maintain the anonymity of the people, we did not download author's username and display names.

We downloaded 70,639 tweets. After deleting duplicates and tweets from users from outside Germany, 57,877 tweets (82%) were kept and cleaned for further analysis. The text cleaning process included.

- Converting all words to lowercase
- Removing punctuation, symbols, emojis, and numbers
- Removing standard German 'stop words' (words that add no meaning to sentences such as: aber, an, am, der, die, das, mit, und, oder, denn)
- Removing words of fewer than 3 characters
- Stemming all words

Stemming has previously shown to increase the quality of the model performance when analysing German tweets [36] and was also applied here. All words were weighted using term frequency-inverse document frequency (TFIDF) [37]. This reflects how important a word is to a particular document in the collection of documents called a corpus. The importance increases proportionally to the number of times a word appears in the document but is offset by the frequency of the word in the corpus.

Using either the coordinates provided or user location names, 38,142 tweets (66% of the total) could be assigned to a state and 19,735 (34%) to a city. Tweets from users in the two states where the floods caused most damage, North Rhine-Westphalia and Rhineland-Palatinate, were classified as tweets from users in affected areas.

### 2.3. Text mining

Text mining is a form of automated content analysis which is often applied when analysing large amounts of unstructured text which cannot easily be analysed manually [38]. Compared to manual coding and classification, text mining has the benefit of being more objective and is more reliable since it removes the need for multiple researchers to assess the relevance of a document [39]. Instead, text mining uses natural language processing and machine learning algorithms to detect patterns and to summarise and cluster text in a consistent way using quantifiable metrics including word frequency and co-occurrence of words [40]. The output, however, needs the careful interpretation of a researcher in order to deliver robust results.

There are two approaches to text mining, supervised and unsupervised. The differences lie in the purpose of the analysis and the algorithms applied. Supervised approaches are applied when the structure and classifications of the text are known, and predictive models are needed. Here we applied unsupervised text mining for which no prior knowledge about the text is needed and, instead, the aim of the analysis is to discover hidden and previously unknown clusters or themes, referred to as latent, within the entire corpus of text data (see next section).

### 2.4. Topic modelling

Topic modelling is one approach to text mining using unsupervised machine learning. In a topic model, text documents are grouped based on latent clusters or themes, referred to as topics. The grouping is done by probabilistic word distributions with each word and document having a probability of inclusion in a topic [41]. Among the most-commonly used approaches using different algorithms to detect latent topics is Latent Dirichlet Modelling (LDA) which relies upon a Dirichlet probability distribution [42]. This model is commonly used for large texts but has some limitations when used for the analysis of short texts such as Twitter messages (limited to 280-characters, having increased from 140 characters in 2017). We therefore used a structural topic model (stm), which is

an extension of the probabilistic LDA topic models allowing correlated structures [43] which allow inclusion of non-textual covariates [44]. These covariates can explain the proportions of a topic as a function of words found in texts (topical prevalence) and the use of words within a topic (topical content). A tweet can be in multiple topics, with the probability of membership summing up to 100%. The relative shares are referred to as topic prevalence, defined as the degree to which a single response belongs to a particular topic [45]. The covariate of the model can then determine whether some topics are more prevalent among a subgroup of Twitter users. Twitter data only provides limited information about the users and here we include two co-variables, the location of the user and the date when the tweet was posted.

The cleaned corpus with all tweets was converted to stm format to be analysed using the R package ‘stm’ [46]. The decision on the number of topics, which has to be determined by the researcher, was based on four metrics. Two metrics measure the quality of the topics, the average word exclusivity and semantic coherence, while two metrics, the held-out likelihood and residual analysis, provide an understanding of the model fit. All four measures were applied here and the outcomes triangulated. This can be challenging as some trade-offs are required, in particular between the semantic coherence and the exclusivity which tend to be anti-correlated (see Refs. [47,48]). There is no ‘right’ number of topics [41] and the researcher needs to make the decision based on these four criteria while also exercising judgement about the conceptual interpretation of the topics. We applied the function ‘searchK’ to run models with 10–30 topics and to summarise the diagnostic of each topic. The results of this function are plotted in Fig. S1. Exclusivity measure of each topic are presented in Table S1 in the Supplementary Materials.

We used the function ‘estimateEffect’ which employs an ordinary least squares (OLS) regression of the covariates on topic prevalence. Correlations between topics were tested using the function ‘topicCorr’ and visualised to show the relationship between topics (Fig. S2). Topics which show connection lines have a high probability of co-occurrence. We used 0.16 as the correlation coefficient threshold [49]. A qualitative assessment of the results was taken from the plot. Through this plot it became possible to link related topics to one another without relying upon an overlap of top keywords alone.

## 2.5. Sentiment analysis

Sentiment analysis can capture emotions and moods of Twitter users expressed in the text [28,50] and it has been used to capture these emotions during natural hazards (e.g., Refs. [28,51]). Understanding emotions and moods can be used as indicators of well-being and happiness, and, in disaster research, also of resilience and preparedness. Positive thinking people, it has been argued, are happier and more resilient [52,53].

A sentiment score is given to each tweet in the dataset and each tweet is categorised as positive, negative, or neutral. We used a lexicon of words to classify the tweets into these three categories. We used GerVADER [54], a German adaptation of an existing sentiment classification tool named VADER [55]. While there are many lexica for the English language, those for German are limited. At this time, GerVADER is only accessible as an Application Programming Interface (API). We developed a Python script that sent 57,854 individual API calls for all tweets in the dataset and saved the JSON results in a Python dictionary. We applied a cluster of four desktop computers to generate the sentiment scores for each tweet. The process took ~48 h.

## 3. Results

### 3.1. Timeline of tweets

The day with most tweets (12,865; 22% of all tweets) was the July 15, 2021 (Fig. 1). Many tweets were also posted on the 14th (8%), on the evening of which the first severe impacts were becoming apparent, and on the 16th (11%) in the immediate aftermath.

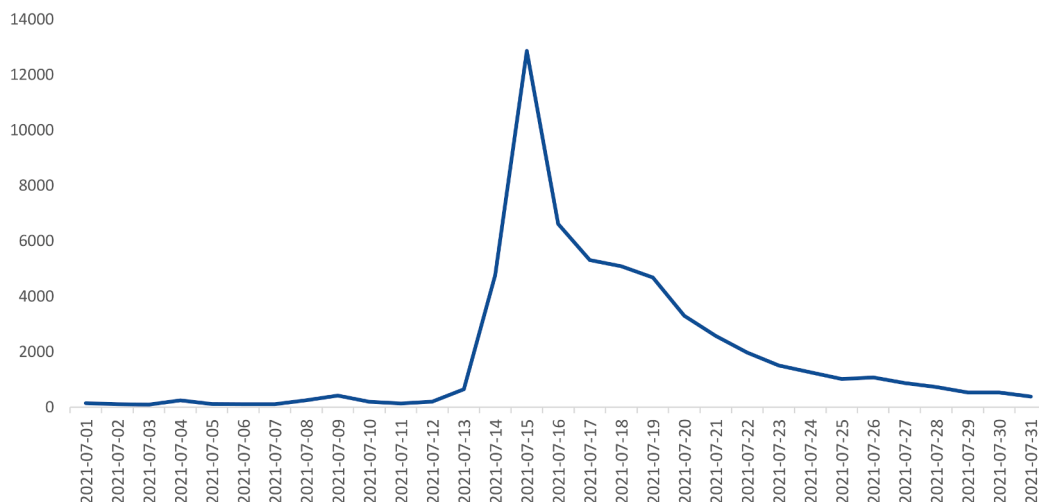


Fig. 1. Number of tweets per day in July 2021.

Based on what happened, we classified the days into five phases. The number of rain and flood related tweets was, as expected, low until first warnings were issued on the 12th July. We named the timeframe before the event between the 1st and 11th of July ‘Before warnings’. The phase between the 12th and 13th July was named ‘First warnings’. This coincided with the start of extreme rainfall (see Ref. [14]). The three days with most tweets were termed the ‘Peak event’ (14th, 15th and 16th July). During the few days after the peak, flooding was still occurring in the region but people had started to recover bodies, assess damage and reduce the risk for further damage. We named this phase ‘First response’. The days subsequent to this were grouped into a ‘Recovery’ phase.

In the 11 days before the event there were 1933 tweets (3.3% of all tweets), 856 occurred during the two days of ‘First warning’ (1.5%). There were 24,249 tweets during the three days of peak impact (41.9%), 18,732 in the four days of first response (31.7%) and 12,467 tweets during the recovery phase (21.5%; 12467; Fig. 1).

### 3.2. Topics

The topic modelling identified ten topics (see Methods for the determination of the number of topics). The most prevalent topic was on weather updates including water levels of affected rivers (Topic 8; 13.6% of all tweets; Table 1). There was a relationship between four of the topics (2, 4, 5 and 7). The four topics within the cluster were about overarching issues which are indirectly related to the flood, such as climate change attribution (Topic 7), politics (Topic 2), extreme weather (Topic 5) and protection measures (Topic 4). Tweets within the topics ‘Solidarity, recovery and aid’ (Topic 9) and ‘Grief and empathy’ (Topic 10) were also correlated and clustered together. The other four topics (‘Evacuation’, ‘Failure and blame’, ‘Damage assessment’, ‘Weather updates’) were not related to any of the others (Fig. S2 in the Supplementary Materials).

### 3.3. Effects of phase and distance to flood on the content of twitter conversations

The prevalence of most topics significantly depended on the phase of the flood and users’ stated location (either in or near the affected area, further away or at an unknown location). For example, those living close to the flood affected areas, in the states of North Rhine Westphalia (NRW) and Rhineland-Palatinate (RP), were, overall, 13% less likely to talk about weather (Topic 8) than those from other locations (Table S2 in the Supplementary Materials). However, this depended heavily on the phase (Fig. 2).

Tweets about weather (Topic 8), particularly extreme weather (Topic 5), were most common before the peak. At the beginning of July, before any warnings, those near the affected areas used Twitter to talk about the weather (Topic 8) with a steep decline subsequently while Twitter conversations about the weather among those living further away peaked later, during the first warnings.

The interactions between phase and location were significant for many topics. For example, while evacuation (Topic 1) was not significantly more discussed during the peak phase than during the warning phases (Table S2), it was so among Twitter users near the affected area. Users nearby were 8% more likely to talk about evacuation during the peak phase than people living further away (Fig. 3). Overall, evacuation was also 5% less likely to be discussed during the first response and recovery phases than during the other three phases.

Other topics that were most common during the peak phase included politics (Topic 2), climate change attribution (Topic 7) and grief and empathy (Topic 10). The topic about grief and empathy was also discussed frequently in the first response and recovery phase, alongside ‘Solidarity, recovery and aid’ (Topic 9), a related topic. While the interest in Topic 10 declined after the peak, Topic 9 gained frequency. Both these topics were more likely to be discussed during the last three phases among users near the affected areas than far away (Fig. 3). During the first response phase, Twitter users near the affected areas were 10% more likely, and during recovery 12% more likely, to tweet about grief and empathy than users further away.

Tweets about protection measures (Topic 4) and about failure and blame (Topic 3) were most common during the first response phase. Topic 3 increased rapidly after the peak phase with a slight decline afterwards, during recovery, while Topic 4 plateaued (Fig.

**Table 1**  
Top distinctive words in each topic, translated from German and with explanations in brackets.

Topic	Label	Top words	%
1	Evacuation	Flood, flooded, evacuate, Wuppertal, Hagen, cellar, house, streets, car, locked, river, mud, stay, dam, full, hopefully, Erft	9.7
2	Politics	Flood, Laschet [German politician who visited the affected area], climate catastrophe, climate protection, CDU [political party], politics, election campaign, Baerbock [German politician], NRW, AFD [political right-wing party], chancellor, lies, welfare	11.4
3	Failure and blame	Flood catastrophe, Merkel [former German chancellor], warned, Germany, warning system, population, fail, sirens, blame, responsibility, ignored, before, in time, system failure	10.5
4	Protection measures	Measures, flood protection, future, money, important, needed, questions, building, government, next, think, money, climate change	5.1
5	Extreme weather	Extreme weather, storm, flooding, frequent, more, heat, dry, trend, soil, heatwave, water, summer	10.1
6	Damage assessment	Flood victims, corona, political, damage, millions, infrastructure, costs, communities, insurance, destroyed, reconstruction, cellar	6.5
7	Climate change attribution	Climate change, flood, climate crises, adaptation, connection, action, time, knowledge, events	11.0
8	Weather updates	Flood, Rhein, level, storm, heavy rain, storm, weekend, warning, office of meteorology, Mosel, expected	13.6
9	Solidarity, recovery and aid	Flood, catastrophe, help, aid, donations, solidarity, affected, Rhineland Palatine, NRW, recovery, support, victims, clean-up, needed, urgent	12.3
10	Grief and empathy	Flood, people, Ahrweiler/Ahrtal, NRW, Eifel, thank you, affected, lost, dead, fatalities, compassion, relatives, Euskirchen, grief, incomprehensible	9.9

NRW = North Rhine Westphalia, one of the affected states.

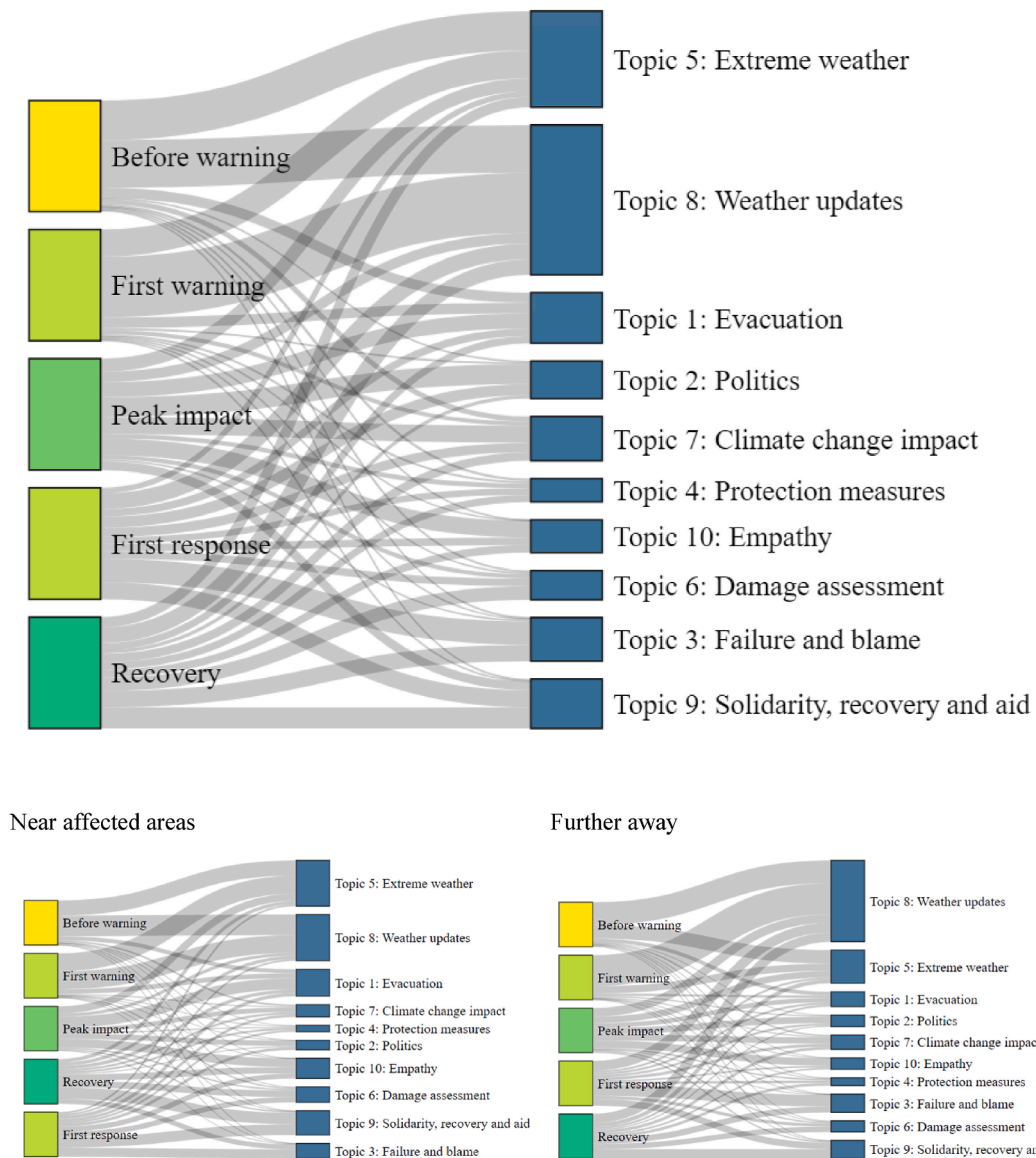


Fig. 2. Distribution of topics across the four phases, for all Twitter conversations and separately for those from Twitter users near the flood affected areas and from users further away.

3). Both topics were slightly more prevalent among users from further away than nearby during these last two phases. The topic about damage assessment (Topic 6) was also most common during the last two phases, particularly among users living near the affected areas.

### 3.4. Sentiments

Overall, there were more tweets with words associated with positive sentiments (43%) than negative (23%) or neutral (34%) sentiments. Tweets about solidarity, recovery and aid (Topic 9) had the highest share of positive words with 59% of all tweets within that topic being positive and only 14% being negative, followed by tweets about weather updates (Topic 8; 55% positive and 16% negative) and about protection measures (Topic 4; 47% positive and 24% negative) (Fig. 4a).

Tweets during the first response phase were most likely to be positive (47%), followed by those during the recovery phase (43%) and during peak impact (42%) (Fig. 4b). Tweets during the first two phases were mostly neutral, being also the least negative, although the share of negative tweets did not differ greatly across the phases.

Overall, users who lived near the affected areas were less likely to post negative tweets (21%), compared to those further away and those with unknown locations (both 23%) (X-squared = 70.30, df = 4, p-value <0.001). At the same time, users nearby were also less likely to post positive tweets (44%) than those further away (46%) and those whose locations were unknown (42%).



Fig. 3. Association between topic prevalence, location and phase, blue colour = Twitter users nearby flood area, red colour = Twitter users further away from affected area.

## 4. Discussion and implications

### 4.1. Situational awareness, information seeking and warning

German Twitter posting behaviour during the 2021 flood, with 41% of tweets occurring over a 48-h period at the flood's peak [17], followed the usual pattern for Twitter users, suggesting that the frequency of social media posts was a good indicator of the sequence of events during the flood (see Refs. [28,56,57]). A large number of social media conversations can be an indicator of high situational awareness and risk perception, which have been linked positively to preparedness and resilience [12,58]. Furthermore, while the Twitter conversations before the first flood warning mainly concerned the extremely hot and dry summer in Germany, the discussion of extreme weather included flooding (Topic 5), and possible river levels once the rain arrived (Topic 8). However, once the rain arrived, the floods were far more severe than anyone had anticipated. While people were aware that flooding was a possibility, no-one living in old houses along the valleys imagined they could be as severe as they turned out to be.

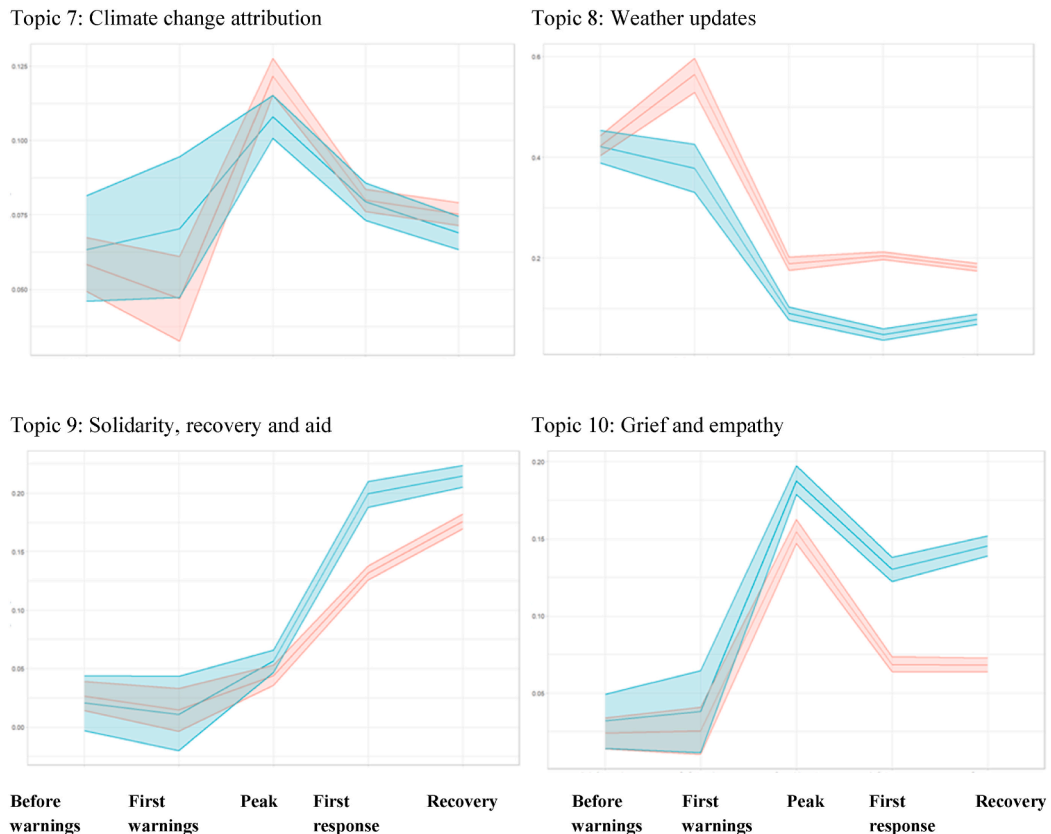


Fig. 3. (continued)

There had been an accurate meteorological forecast by the German Weather Service (Deutscher Wetterdienst; DWD) several days in advance of the storm system, and at least three weather warning apps were available in Germany in 2021 to those who had installed them - 'NiNA',<sup>1</sup> 'Katwarn'<sup>2</sup> and 'WarnWetter'<sup>3</sup> - as well as a specific dissemination platform [17]. However, these warnings were inadequate. Researchers of this previous study [17] estimated that that "30–40% of the residents in areas affected by the floods were not reached by severe weather or flood warnings" and concluded that a failure to warn people contributed to the high number of fatalities and improvements are needed to the whole national flood forecasting, warning and response systems so that risk is communicated rapidly to and among communities likely to be affected. Our analysis did not detect any of the app names as important words in peoples' tweets. While this does not mean that people did not use them, it does suggest that these apps were not frequently advocated or retweeted. In fact, more research is needed about people's knowledge of available apps and the status-quo of their usage across Germany.

Social media should play a central role in any reorganisation of warning systems in Germany and Europe. Access to weather updates is pivotal to resilience, risk identification and rapid response [16]. In other European countries such as France, short texts are automatically delivered to peoples' cell phones using a cell-broadcast approach regardless of whether people have downloaded a weather app. As a result, not only are residents warned of imminent bad weather but, so too, are temporary visitors. Such an approach could readily be adopted across Germany and the European Union, and more broadly around the globe. At the moment, natural hazard warnings in Germany are the responsibility of the 16 federal states, but as the floods of 2021 have shown, hazards can occur simultaneously across multiple states and countries so there is a need to improve coordination and standardise warning protocols.

However, while forecasts and early warning systems are key factors in disaster flood preparedness and response, and their presence and effectiveness can be key to reducing the impacts of hazards [15], such communication is one way, from authorities to people in the area likely to be affected. There are two properties of social media that can augment such warnings. First, it can amplify warnings by disseminating them through social networks which may not only extend their reach but also increase its local impact. Secondly, it can be used to understand what is happening in flood affected areas in real time because the information is communicated by and among affected people [17,29,56]. Our analysis suggests that-rapid automated content analysis of incoming tweets during an emergency could be used by emergency agencies to direct resources to people in need. Warning apps could then be modified to pro-

<sup>1</sup> NiNA (Notfall-Informationen-und Nachrichten-App) = the official warning app of the German government, provided by the Federal Office for Civil Protection (BBK).

<sup>2</sup> Katwarn- = provided by the Fraunhofer-Institut. The app is also linked to the German Weather Service (DWD) and to the NiNA app.

<sup>3</sup> WarnWetter = provided by the German Weather Service (DWD).

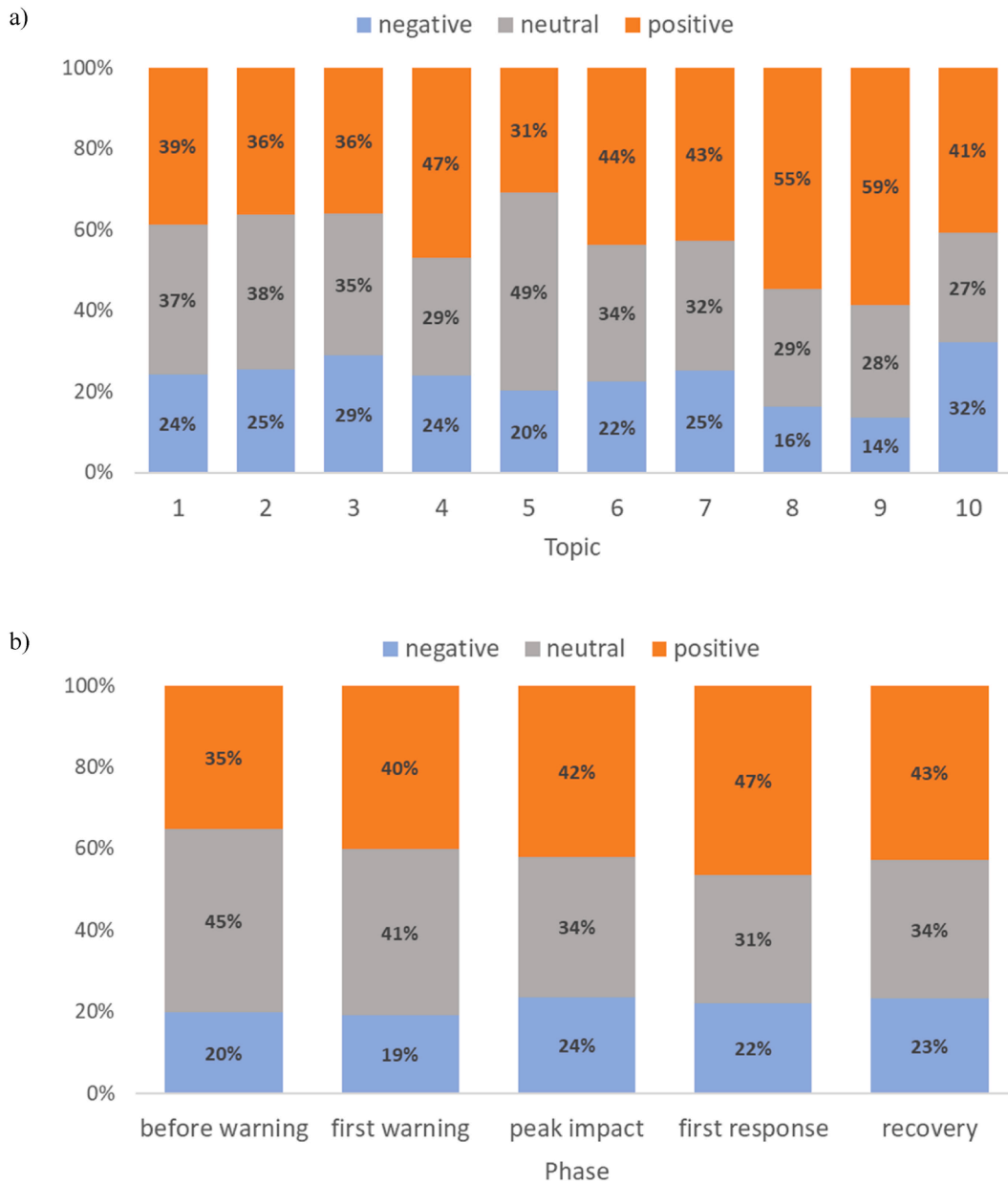


Fig. 4. Sentiments across topics (a) and phases (b), showing the percentage of positive and negative sentiments.

vide specific content with advice, warnings or directions in response to the issues emerging from the Twitter analysis. How this could be done is discussed in the following section about different topics discussed on Twitter during the flood. The two topics about grief and solidarity also suggest that people use Twitter to give mental support, which can be helpful if people are in isolated situations and could also help guide support agencies to places where such social support is most needed.

#### 4.1.1. Thematic analysis

Overall, natural language processing revealed a variety of meaningful and coherent themes. A similar broad range of topics including issues of solidarity, rescue organisations, political and administrative failure have also been found when analysing digital newspapers about the same flood [32]. This suggests it would be valid to employ the topics from our social media analysis to filter, sort and analyse large Twitter datasets concerning floods for future use [33]. Some topics are also similar to those found during the Elbe flood in 2013 in Germany [33]. However, our analysis provided additional thematic details to topics about weather warnings, weather updates and flood aids that emerged from the 2013 study. Here we discuss some of those key themes and potential policy implications.

#### 4.1.2. Topics about flood responses

One topic stood out describing responses to the flood: Topic 1 about evacuation (9.7% of all tweets). Not surprisingly, people further from affected areas communicated less about evacuation than those directly affected. However, what was surprising was that communication about evacuation did not peak before or immediately after official warnings to evacuate but remained similar throughout the study period with a lower likelihood of being communicated during ‘First response’ than during the other phases (see [Table S2](#)). One would have expected a flurry of tweets about evacuation intentions during the ‘First warning’ phase given that timely evacuation is essential to avoid fatalities. Later during the peak phase, it was probably already too late to evacuate because of high water and damage to roads. The result might indicate a lack of evacuation planning or an unwillingness to consider evacuations, even among those near the affected areas. This again shows the extent to which people were surprised by the severity of the flood in some areas, to the extent that they did not contemplate evacuation. In future flood events fatalities could be reduced by warnings by authorities such as emergency management agencies that include explicit directions for early evacuation and warnings about the possible consequences of not evacuating. However, evacuation training is also needed. Many people will not know what to do when they receive advice to evacuate. Lessons could be learned from wildfire emergency management. During wildfires, authorities disseminate urgent explicit advice to leave early since late evacuation can cause more harm than staying. If people feel they have to stay they should take precautionary measures as soon as possible (e.g. Refs. [59,60]. The real time information from social media could be translated into short text messages with, for example, specific directions to evacuate.

For floods, a sensible precautionary measure would be to avoid cellars and make for high ground. During the German flood, many people died in their cellars [61]. In the tweets, the word cellar appears frequently, mostly in relation to damage assessment in the first response phase (Topic 6) and in the topic about evacuation (Topic 1). Many Tweets were just about flooded cellars (e.g. “water in the cellar up to the roof”) but some were also about people going into their cellars to fetch things or to check their belongings such as washing machines, e.g. “I saw the heavy rain coming, from the weather watch, and just gone down into the cellar to get things”. At least some people who tried to rescue their belongings from cellars were surprised by the flash flood and drowned. While the dangers of visiting cellars did not come up in Twitter conversations, social media could be used to warn people living near rivers explicitly to avoid visiting their cellars during heavy rain when river levels are high.

Tweets within Topic 4 (5.1%) indicate some degree of long-term precautionary planning to mitigate the impacts of heavy rain and flash floods. This topic covers measures for future flood protection, mainly the costs and the urgency with which they are needed as the climate changes. This topic was mostly discussed after the peak phase. Unlike conversations about immediate responses, such as in Topic 1, future protection measures were less likely to be discussed by people near the affected area than by those further away or those with unspecified locations ([Table S2](#)). This shows users’ awareness of the high risk of future similar heavy rain and flood events across Germany and how people can be protected. The geographic breadth of Topic 4 indicates that the flood had an impact on the whole of German society, probably raising awareness across the country of the dangers of flash floods as the climate changes, and, hopefully, also investments into country-wide pluvial flood planning and protection, particularly in urban areas.

#### 4.1.3. Topics about damage assessment

Only one topic was mainly about damage assessment (Topic 6; 6.5% of all tweets). Unsurprisingly, this topic was most common during the last two phases. People from near the affected areas were more likely to talk about damage in all phases except the peak phase ([Fig. 3](#)) and this confirms that Twitter conversations can be indicators of the degree and nature of damage [28,51]. Apart from flooded houses, gardens and cellars, people talked about damaged and closed roads, bridges and schools. The topic also reflected the concerns of those affected, such as about insurance cover and payouts, the need for special insurance for natural hazards and also the broader implications of the increasing frequency of extreme weather events on insurance premiums.

If damage data were to be analysed in real-time using machine learning, Twitter and other social media conversations could have been used by government agencies to deploy timely relief, such as for closed roads and schools. Social media activity can indicate more quickly than conventional methods the nature and impact of damage and the need for support and rescue [28], in particular when landlines are damaged and cellular networks overloaded (or cellular towers destroyed) which means that even text messages might take hours to get through. While Internet based applications such as Twitter also rely on a network which might collapse, at least temporarily, these applications have been shown to be less likely than other channels to be overwhelmed during disasters [62]. Moreover, data networks recover faster and perform better than the landline networks, as found during Hurricane Katrina in the US [63].

Crowdsourcing approaches using Twitter for rapid damage assessment after natural hazards has been suggested for earthquakes [64], hurricanes [28], wildfires [29] and floods [65]. Given the high incidence of tweets about the flood and the range of relevant topics, this is something the local (state) and the German Governments could consider as part of a future flood warning strategy (see Ref. [66]).

#### 4.1.4. Topics about politics and blame

One of the most frequently used words in Topic 2 (11.4%) was ‘election’. The flood occurred during the election campaign for the German chancellor (federal election on the September 26, 2021) and many politicians visited to affected area to express their solidarity and to get a picture of the damage. However, these politicians were also accused of using the flood as a platform to boost their election chances [32]. Few tweets on the topic had positive sentiments ([Fig. 4a](#)) and the word ‘lies’ was used frequently in relation to one of the politicians who showed up in the affected area (called Laschet; [Table 1](#)). Many users suggested he step down from his candidature for chancellor. The Twitter response suggests that politicians should consider carefully whether visits to disaster areas benefit either the community they are visiting or their own electoral standing.

Tweets discussing failures that led to the disaster and blaming those thought responsible (Topic 3) were the major social media conversations after the flood. This might indicate that people had been relying on governmental authorities to warn them of the impending flood more quickly and efficiently, and reflects a low level of self-efficacy common among people affected by natural hazards (e.g. Ref. [67]). One user wrote “How dumb does one have to be to warn people 30 min before the flood instead of 24 h before. No wonder many people go downstairs into their cellars to rescue their stuff”. However, many other tweets in Topic 3 talked about how people under-estimated how damaging the flash floods would be, despite being warned. This reflects that, although information was available, warnings were issued and people had been noticing the extreme weather during the summer (Topic 5), anticipation of and preparation for the flood were lacking or inadequate [20].

As discussed, the two related topics on solidarity and grief (9 and 10, totalling 22.2% of all tweets) were separated primarily on the sentiment of the messages, although Topic 9 (‘Solidarity, recovery and aid’) was mostly discussed in the first response and recovery phases, and Topic 10 (‘Grief and empathy’) during the peak phase (Fig. 3). A topic about sharing disaster-related sentiments including empathy, solidarity and help was also identified by Ref. [34]. To some extent, solidarity can contribute to resilience and increase awareness of both the flooding and its impacts [68,69]. The topics also have practical implications. Within Topic 9, some users from Euskirchen, close to the flooded river Ahr, talked about empty shelves in supermarkets, in particular fresh vegetables, fruits and frozen food. Such information could also be used to send needed resources to affected areas and accelerate recovery were rapid analysis of Twitter messages a standard part of emergency response.

#### 4.2. Sentiments

Our results on sentiments echo previous findings that Twitter users are more likely to post positive tweets during the recovery than earlier during a natural hazard [51]. During hurricane Sandy in 2012, positive sentiments expressed in social media were thought to reflect hope during the recovery process, but also that users who were affected by the hurricane were too preoccupied with rebuilding their houses and unlikely to tweet during that time [51]. This could also have explained some of the variation in Twitter use during the flood in Germany.

Our findings partially corroborate previous studies suggesting that sentiments in Twitter reflect spatial–temporal mood variations related to damage suffered during natural disasters [28]. Two correlated topics with similar key words about aid and empathy (Topics 9 and 10) and that were common among people living near the affected areas, differed primarily in their sentiments. Topic 9 (‘Solidarity, recovery and aid’; 12.3%) reflected positive feelings about support given to flood-affected people while Topic 10 (‘Grief and empathy’; 9.9%) expressed the sadness and bewilderment at the sudden flooding. Overall, however, users who lived near the affected areas were less likely to post negative tweets (21%) than those further away and those with unknown locations (both 23%) ( $X^2 = 70.30$ ,  $df = 4$ ,  $p$ -value  $< 0.001$ ), and also less likely to post positive tweets (44%) than those further away (46%) and those whose locations were unknown (42%).

#### 4.3. Study limitations

A few limitations are worth noting, the first related to the data collection and quality, the second to automated text analysis. First, as with any data collection process in social science, our data is not representative of the whole society. We limited our data source to Twitter users. Recent surveys found that 21.4 million people in Germany have access to smartphones [70] and 8 million use Twitter [71], with the highest share of Twitter users being younger people. We could have also included other social media platforms such as Facebook. Facebook is also often used in emergency situations, in particular to share stories, experiences and photos (e.g. Refs. [72, 73], but less so to access real-time information [74]. Twitter is also more often used by news outlets, government officials and emergency services for disaster communication and management than other social media platforms [75,76]. We are therefore confident that the data collected from Twitter depicted a wider picture of conversations by various actors involved in the flood. Another reason, we limited our data collection to Twitter was that Twitter offers access to publicly available data while Facebook data are private (see Ref. [77]; for example, for a discussion on the ethics of social media use).

Social media users are usually considered to be young, with Twitter users considered younger than Facebook but older than Instagram users [78]. However, the age of social media users has been steadily increasing [79]. It has also been shown that, while younger family members seek and disseminate information during emergencies, they do so on behalf of the whole household [30]. This means that Twitter can nevertheless capture the views of the broader society. Nevertheless, the bias in the data can affect the generalisability of our findings and more in-depth research is needed with potentially under-represented groups such as social media users in Germany whose main language is not German and who might communicate on Twitter in another language and older people who do not use smartphones and social media. Dependent older people living alone are among the most vulnerable during natural hazards and many of those who died in the floods were older people. Other more qualitative research methods are needed to understand how they coped during the flood, what barriers they faced and how they could benefit from social media use in future similar events.

Second, topic modelling and sentiment analysis cannot detect nuances of language such as sarcasm, irony or local idioms. Although there are models for sarcasm detection [80], this remains a limitation and highlights the need for manual interpretation and contextualisation of the results derived from machine learning approaches [81]. Despite the many advantages of topic modelling, it is highly sensitive to ‘hyperparameters, parameters that are set before training the model, such as the number of topics, the alpha and beta priors, and the number of iterations. Hyperparameters can significantly affect the output of the model and setting them improperly can lead to poor results. Another limitation of topic modelling is that it does not capture causality. We only identified correlations between words and topics (co-occurrence), not causation. This means that although topic modelling can identify the topics that occur frequently in a document, it cannot explain why these topics are related. Topic-modelling and sentiment analysis results therefore need manual interpretation of potential cause which may not be objective and our findings could be considered speculative.

Our topics all seem to reflect genuine concerns and conversations, but misinformation dissemination is a known disadvantage of social media [82,83]. Again, topic modelling does not easily detect fake messages which further compromises the generated data. While other studies have found little evidence for misinformation spreading during natural hazards (e.g. Refs. [84,85], we did not analyse the potential existence of rumours and fake news. Most studies which do analyse social media data for misinformation use English text and the fact that the information shared on Twitter during the flood was in German might have reduced the threat of fake news generation using bots.

## 5. Conclusion

In July 2021, parts of Western Europe, particularly Germany and the Benelux countries, were hit by intense rainfall after a prolonged period of extremely dry and hot weather, which resulted in unprecedented flash flooding along river systems. Many people died, nearly 200 in Germany alone, while many more people lost their homes. In times of emergency, many people turn to social media and many thousands of flood-related German Twitter messages (tweets) were sent during this difficult time. We downloaded and analysed nearly 58,000 tweets related to the flood for their content and sentiments. Unsurprisingly, most tweets were posted during the peak of the flood (evening of the 14th and during July 15, 2021) with many messages disseminating and seeking information, including weather updates, climate change attribution of the flood and extreme weather in the summer of 2021. While the high level of community engagement even before the flood could be an indicator for awareness and risk perception, most people and local agencies were surprised by the severity of the flood. During the peak and immediately after the flood, the topics were very diverse, shifting from information seeking and discussing climate change impacts to politics, failure, damage, response behaviour, solidarity and grief. Surprisingly, the language used in flood-related tweets depicted more positive than negative emotions. Sentiments of tweets were most positive during the first response and recovery phase and people living near the affected areas were less likely to post negative tweets. This might reflect the sheer abundance of help on the ground by volunteers and professionals of the Red Cross, the THW and other groups.

In the future, social media could be harnessed in several ways that would increase the speed and efficacy of flood preparedness and response. For example, our thematic analysis showed a low level of awareness of a need to evacuate and a lack of evacuation planning, even among those near affected areas. As is happening in other emergency response protocols, standardised explicit messages through Twitter and other social media could rapidly disseminate warnings about the timing and severity of likely floods, and advice on sensible responses, like, evacuation, avoiding cellars and seeking high ground. Such rapid warnings of impending dangerous weather events are already being disseminated in some European countries, regardless of whether particular social media apps have been downloaded.

However, warning apps do not yet respond to changing localised situations. While Twitter communication is perfunctory and diverse, the very volume and immediacy of Twitter responses means that, if messages can be extracted quickly from Twitter feeds and analysed appropriately, they will provide information more rapidly than is likely to be possible using formal media channels and other administrative means. For example, during and immediately after a flood, they could indicate the location and severity of damage, and of the requirements of the affected populace. The automated topic modelling conducted here could help achieve that end. Twitter conversations can also provide an evidence base on whether official visits are actually providing the intended succour to the afflicted or are resented and considered a distraction from recovery efforts. Lastly, Twitter can be used to gauge sentiment and the extent of resilience of affected communities, particularly if multiple Twitter analyses are undertaken after stressful events so that there is a statistically robust baseline against which to compare subsequent individual events.

Our results can contribute to better warnings and more detailed and targeted information based on citizen generated data. When these data are rapidly analysed, the information can also be used by responders on the ground to target resource allocations to people in need, including providing mental support. Social media is now crucial for disaster communication and its use is the type of innovation that will be essential for adaptation to climate change. As a consequence, any tightening of free use of social media data will not only return society to a time when real time information was unavailable but will also inhibit adaptation to climatic conditions the modern world has never before experienced.

## Ethics declarations

We used freely available Twitter data and did not record username specific data. No primary data were collected. The use of social media data was in line with the Charles Darwin University human ethics committee.

## Declaration of competing interest

The authors declare the following financial interests/personal relationships which may be considered as potential competing interests:

Kerstin Zander reports financial support was provided by Charles Darwin University.

## Data availability

Data will be made available on request.

## Appendix A. Supplementary data

Supplementary data to this article can be found online at <https://doi.org/10.1016/j.ijdr.2023.103936>.

## References

- [1] S.N. Willner, C. Otto, A. Levermann, Global economic response to river floods, *Nature Clim Change* 8 (2018) 594–598.
- [2] R.E. Emerton, E.M. Stephens, F. Pappenberger, et al., Continental and Global Scale Flood Forecasting Systems, *Wiley Interdiscip Rev Water* 3, 2016, pp. 391–418.
- [3] IPCC, in: Pörtner, et al. (Ed.), *Climate Change 2022: Impacts, Adaptation and Vulnerability. Contribution of Working Group II to the Sixth Assessment Report of the Intergovernmental Panel on Climate Change*, Cambridge University Press, 2022.
- [4] M. Tanoue, R. Taguchi, H. Alifu, et al., Residual flood damage under intensive adaptation, *Nature Clim Change* 11 (2021) 823–826.
- [5] B. Jongman, Effective adaptation to rising flood risk, *Nature Commun* 9 (2018) 1986.
- [6] P.J. Ward, B. Jongman, J.C.J.H. Aerts, et al., A global framework for future costs and benefits of river-flood protection in urban areas, *Nature Clim Change* 17 (2017) 642–646.
- [7] O.E.J. Wing, W. Lehman, P.D. Bates, et al., Inequitable patterns of US flood risk in the Anthropocene, *Nature Clim Change* 12 (2022) 156–162.
- [8] P. Hudson, A.H. Thielen, The presence of moral hazard regarding food insurance and German private businesses, *Nat Hazards* 112 (2022) 1295–1319.
- [9] L.T. de Ruig, T. Haer, H. de Moel, et al., How the USA can benefit from risk-based premiums combined with flood protection, *Nature Clim Change* 12 (2022) 995–998.
- [10] E. Lechowska, Approaches in research on flood risk perception and their importance in flood risk management: a review, *Nat Hazards* 111 (2022) 2343–2378.
- [11] T. Grothmann, F. Reusswig, People at risk of flooding: why some residents take precautionary action while others do not, *Nat Hazards* 38 (2006) 101–120.
- [12] R. Raaijmakers, J. Krywkow, A. van der Veen, Flood risk perceptions and spatial multi-criteria analysis: an exploratory research for hazard mitigation, *Nat Hazards* 46 (2008) 307–322.
- [13] S. Mohr, U. Ehret, M. Kunz, et al., A multi-disciplinary analysis of the exceptional flood event of July 2021 in central Europe—Part 1: event description and analysis, *Nat Hazards Earth Syst Sci* 23 (2023) 525–551.
- [14] F. Kreienkamp, **Rapid Attribution of Heavy Rainfall Events Leading to the Severe Flooding in Western Europe during July 2021**, *World Weather Attribution*, <https://www.worldweatherattribution.org/wp-content/uploads/Scientific-report-Western-Europe-floods-2021-attribution.pdf>.
- [15] J.S. Tradowsky, S.Y. Philip, F. Kreienkamp, et al., Attribution of the heavy rainfall events leading to severe flooding in Western Europe during July 2021, *Clim Change* 176 (2023) 90.
- [16] B. Manandhar, S. Cui, L. Wang, S. Shrestha, Post-flood resilience assessment of July 2021 flood in Western Germany and Henan, China, *Land* 12 (2023) 625.
- [17] A.H. Thielen, P. Bubeck, A. Heidenreich, J. von Keyserlingk, L. Dillenaar, A. Otto, Performance of the flood warning system in Germany in July 2021 – insights from affected residents, *Nat Hazards Earth Syst Sci* 23 (2023) 973–990.
- [18] L.M. Netzel, S. Heldt, S. Engler, M. Denecke, The importance of public risk perception for the effective management of pluvial floods in urban areas: a case study from Germany, *J Flood Risk Management* 14 (2021) e12688.
- [19] D. Osberghaus, The determinants of private flood mitigation measures in Germany — evidence from a nationwide survey, *Ecol Econ* 110 (2015) 36–50.
- [20] A. Fekete, S. Sandholz, Here comes the flood, but not failure? lessons to learn after the heavy rain and pluvial floods in Germany 2021, *Water* 13 (2021) 3016.
- [21] B. Noll, T. Filatova, A. Need, et al., Contextualizing cross-national patterns in household climate change adaptation, *Nature Clim Change* 12 (2022) 30–35.
- [22] P. Babczyk, S. Seebauer, Unpacking protection motivation theory: evidence for a separate protective and non-protective route in private flood mitigation behavior, *J Risk Res* 22 (2019) 1503–1521.
- [23] A. Heidenreich, T. Masson, S. Bamberg, Let's talk about flood risk – evaluating a series of workshops on private flood protection, *Int J Disaster Risk Reduct* 50 (2020) 101880.
- [24] L. Dillenaar, P. Hudson, A.H. Thielen, Urban pluvial flood adaptation: results of a household survey across four German municipalities, *J Flood Risk Manag* 15 (2022) e1274.
- [25] F. Yuan, R. Liu, Feasibility study of using crowdsourcing to identify critical affected areas for rapid damage assessment: hurricane Matthew case study, *Int J Disaster Risk Reduct* 28 (2018) 758–767.
- [26] P.R. Spence, K.A. Lachlan, D.R. Griffin, Crisis communication, race, and natural disasters, *J Black Stud* 37 (2007) 539–554.
- [27] P. Hudson, A.H. Thielen, P. Bubeck, The challenges of longitudinal surveys in the flood risk domain, *J Risk Res* 23 (2020) 642–663.
- [28] Y. Kryvashyeyu, H. Chen, N. Obradovich, et al., Rapid assessment of disaster damage using social media activity, *Sci Adv* 2 (2016) e1500779.
- [29] B. Abedin, A. Babar, Institutional vs. non-institutional use of social media during emergency response: a case of Twitter in 2014 Australian bush fire, *Inf Syst Front* 20 (2018) 729–740.
- [30] K.K. Zander, R. Sibarani, D. Nguyen, J. Lassa, A. Dimmock, How do Australians use social media during natural hazards? A survey, *Int J Disaster Risk Reduct* 81 (2022) 103207.
- [31] L. Netzel, S. Heldt, M. Denecke, Analyzing Twitter communication about heavy precipitation events to improve future risk communication and disaster reduction in Germany, *Urban Water J* 18 (2021) 310–313.
- [32] M. Kahle, M. Kempf, B. Martin, R. Glaser, Classifying the 2021 'Ahrtal' flood event using hermeneutic interpretation, natural language processing, and instrumental data analyses, *Environ Res Commun* 4 (2022) 051002.
- [33] S. Gründer-Fahrer, C. Berger, A. Schlaf, G. Heyer, Computational, communicative, and legal conditions for using social media in disaster management in Germany, in: 11<sup>th</sup> International Conference on Availability, Reliability and Security, 2016, <https://doi.org/10.1109/ARES.2016.68>.
- [34] M. Moghadas, A. Fekete, A. Rajabifard, T. Kötter, The Wisdom of Crowds for Improved Disaster Resilience: a Near-Real-Time Analysis of Crowdsourced Social Media Data on the 2021 Flood in Germany, *GeoJournal* 88, 2023, pp. 4215–4241.
- [35] M.R. Nair, G.R. Ramya, P.B. Sivakumar, Usage and analysis of twitter during 2015 Chennai flood towards disaster management, *Procedia Comput* 115 (2017) 350–358.
- [36] O. Abramova, K. Batzel, D. Modesti, Collective response to the health crisis among German Twitter users: a structural topic modeling approach, *International Int J Inf Manag Data Insights* 2 (2022) 100126.
- [37] K. Jones, A statistical interpretation of term specificity and its application in retrieval, *J Doc* 60 (2004) 493–502.
- [38] S.A. Salloum, M. Al-Emran, A.A. Monem, K. Shaalan, Using text mining techniques for extracting information from research articles, in: K. Shaalan, A.E. Hassanien, F. Tolba (Eds.), *Intelligent Natural Language Processing: Trends and Applications*, Springer International Publishing, 2018, pp. 373–397.
- [39] J.A. Evans, P. Aceves, Machine translation: mining text for social theory, *Annu Rev Sociol* 42 (2016) 21–50.
- [40] D. Bzdok, N. Altman, M. Krzywinski, Statistics versus machine learning, *Nat Methods* 15 (2018) 233234.
- [41] J. Grimmer, B.M. Stewart, Text as data: the promise and pitfalls of automatic content analysis methods for political texts, *Polit Anal* 21 (2013) 267–297.
- [42] D.M. Blei, A. Ng, M. Jordan, Latent dirichlet allocation, *J Mach Learn Res* 3 (2003) 993–1022.
- [43] D.M. Blei, J.D. Lafferty, A correlated topic model of science, *Ann Appl Stat* 1 (2007) 17–35.
- [44] M.E. Roberts, B.M. Stewart, E.M. Airoldi, A model of text for experimentation in the social sciences, *J Am Stat Assoc* 111 (2016) 988–1003.
- [45] E. Tvinnereim, K. Fløttum, Ø. Gjerstad, M.P. Johannesson, Å.D. Nordø, Citizens' preferences for tackling climate change. Quantitative and qualitative analyses of their freely formulated solutions, *Glob Environ Change* 46 (2017) 34–41.
- [46] M.E. Roberts, B.M. Stewart, D. Tingley, stm: an R package for structural topic model, *J Stat Softw* 91 (2019) 1–40.
- [47] D. Mimno, H.M. Wallach, E. Talley, M. Leenders, A. McCallum, Optimizing semantic coherence in topic models, in: *Proceedings of the Conference on Empirical Methods in Natural Language Processing EMNLP'11*, Association for Computational Linguistics, 2011, pp. 262–272.
- [48] M.E. Roberts, B.M. Steart, D. Tingley, et al., Structural topic models for open-ended survey responses, *Am J Pol Sci* 58 (2014) 1064–1082.
- [49] S. Tonidandel, K.M. Summerville, W.A. Gentry, S.F. Young, Using structural topic modeling to gain insight into challenges faced by leaders, *Leadersh Q* 33

- (2022) 101576.
- [50] B. Dahal, S.A.P. Kumar, Z. Li, Topic modelling and sentiment analysis of global climate change tweets, *Soc Netw Anal Min* 9 (2019) 24.
- [51] L. Zou, N.S.N. Lam, H. Cai, Y. Qiang, Mining Twitter data for improved understanding of disaster resilience, *Ann Am Assoc Geogr* 108 (2018) 1422–1441.
- [52] W.W.S. Mak, I.S.W. Ng, C.C.Y. Wong, Resilience: enhancing well-being through the positive cognitive triad, *J Couns* 58 (2011) 610–617.
- [53] M.E. Gowan, R.C. Kirk, J.A. Sloan, Building resiliency: a cross-sectional study examining relationships among health-related quality of life, well-being, and disaster preparedness, *Health Qual Life Outcomes* 12 (2014) 85.
- [54] K. Tymann, M. Lutz, P. Palsbröker, C. Gips, GerVADER-A German adaptation of the VADER sentiment analysis tool for social media texts, in: *LWDA*, 2019, pp. 178–189.
- [55] C.J. Hutto, E.E. Gilbert, VADER: a parsimonious rule-based model for sentiment analysis of social media text, *Proceedings of the International AAAI Conference on Web and Social Media* 8 (2014) 216–225.
- [56] Z. Wang, X. Ye, M.H. Tsou, Spatial, temporal, and content analysis of Twitter for wildfire hazards, *Nat Hazards* 83 (2016) 523–540.
- [57] A. Silver, J. Andrey, Public attention to extreme weather as reflected by social media activity, *J Contingencies and Crisis Management* 27 (2019) 346–358.
- [58] J. Rexiline Ragini, P.M. Rubesh Anand, V. Bhaskar, Mining crisis information: a strategic approach for detection of people at risk through social media analysis, *Int J Disaster Risk Reduct* 27 (2018) 556–566.
- [59] J. Handmer, M. Van der Merwe, S. O'Neill, The risk of dying in bushfires: a comparative analysis of fatalities and survivors, *Prog Disaster Sci* 1 (2019) 100015.
- [60] A.M. Stasiewicz, T.B. Paveglio, Preparing for wildfire evacuation and alternatives: exploring influences on residents' intended evacuation behaviors and mitigations, *Int J Disaster Risk Reduct* 58 (2021) 102177.
- [61] H. Nyarsik, Der Keller wird zur Todesfalle: Wie die Menschen in den Fluten sterben, *N-TV* 16.07, 2021. <https://www.n-tv.de/panorama/Wie-die-Menschen-in-den-Fluten-sterben-article22686470.html>. accessed April 2023.
- [62] T. Simon, A. Goldberg, B. Adini, Socializing in emergencies—a review of the use of social media in emergency situations, *Int J Inf Manag* 35 (2015) 609–619.
- [63] A. Banipal, Strategic approach to disaster management: lessons learned from Hurricane Katrina, *Disaster Prev Manag* 15 (2006) 484–494.
- [64] L. Li, M. Bensi, Q. Cui, G.B. Baecher, Y. Huang, Social media crowdsourcing for rapid damage assessment following a sudden-onset natural hazard event, *Int J Inf Manag* 60 (2021) 102378.
- [65] G. Cervone, E. Sava, Q. Huang, E. Schnebele, J. Harrison, N. Waters, Using Twitter for tasking remote-sensing data collection and damage assessment: 2013 Boulder flood case study, *Int J Remote Sens* 37 (2016) 100–124.
- [66] Bundesregierung, Miteinander Sicherheit schaffen, 2022. <https://www.bundesregierung.de/breg-de/aktuelles/hochwasserschutz-1944386>. accessed April 2023.
- [67] K.K. Zander, J. Rieskamp, M. Mirabaia, A. Alazab, D. Nguyen, Responses to heat waves: what can Twitter data tell us? *Nat Hazards* 116 (2023) 3547–3564.
- [68] A. Fekete, J. Rhyner, Sustainable digital transformation of disaster risk—integrating new types of digital social vulnerability and interdependencies with critical infrastructure, *Sustainability* 12 (2020) 1–18.
- [69] T. Yabe, S.C.P. Rao, S.V. Ukkusuri, S.L. Cutter, Toward data-driven, dynamical complex systems approaches to disaster resilience, *Proce Natl Acad Sci USA* 119 (2022) e2111997119.
- [70] Statista Statistiken, zur Smartphone-Nutzung in Deutschland, 2023. <https://de.statista.com/themen/6137/smartphone-nutzung-in-deutschland/#:~:text=Die%20Anzahl%20der%20Smartphone%2DNutzer,bis%2039%2DJ%C3%A4hrigen%2096%20Prozent.> (accessed April 2023).
- [71] Nutzerzahlen in Deutschland und weltweit, 2022. <https://www.blog2social.com/de/blog/social-media-nutzer/#:~:text=Monatlich%20aktive%20Social%20Media%20Nutzer%20in%20Deutschland&text=Darauf%20folgen%20Linkedin%20mit%2014,Twitter%20mit%208%20Millionen%20Nutzern.> accessed April 2023.
- [72] J. Kim, M. Hastak, Social network analysis: characteristics of online social networks after a disaster. *International, J Inf Manag* 38 (2018) 86–96.
- [73] J. Stephenson, M. Vaganay, D. Coon, R. Cameron, N. Hewitt, The role of Facebook and Twitter as organisational communication platforms in relation to flood events in Northern Ireland, *J Flood Risk Manag* 11 (2018) 339–350.
- [74] O. Anikeeva, M. Steenkamp, P. Arbon, The future of social media use during emergencies in Australia: insights from the 2014 Australian and New Zealand Disaster and Emergency Management Conference social media workshop, *Aust J Emerg Manag* 30 (2015) 22–26.
- [75] K.M. Carley, M. Malik, P.M. Landwehr, J. Pfeffer, M. Kowalchuck, Crowd sourcing disaster management: the complex nature of Twitter usage in Padang Indonesia, *Saf Sci* 90 (2016) 48–61.
- [76] B. Kusumasari, N.P.A. Prabowo, Scraping social media data for disaster communication: how the pattern of Twitter users affects disasters in Asia and the Pacific, *Nat Hazards* 103 (2020) 3415–3435.
- [77] J. Mahoney, K. Le Louvier, S. Lawson, D. Bertel, E. Ambrosetti, Ethical considerations in social media analytics in the context of migration: lessons learned from a Horizon 2020 project, *Res Ethics* 18 (2022) 226–240.
- [78] T. Laor, My social network: group differences in frequency of use, active use, and interactive use on Facebook, Instagram and Twitter, *Technol Soc* 68 (2022) 101922.
- [79] S.R. Cotton, A.M. Schuster, A. Seifert, Social media use and well-being among older adults, *Curr Opin Psychol* 45 (2022) 101293.
- [80] F.B. Kader, N.H. Nujat, T.B. Sogir, M. Kabir, H. Mahmud, K. Hasan, Computational Sarcasm Analysis on Social Media: a Systematic Review, 2022 arXiv preprint arXiv:2209.06170.
- [81] K.K. Zander, S.T. Garnett, H. Sterly, S. Ayebe-Karlsson, B. Šedová, H. Lotze-Campen, C. Richerzhagen, H.S. Baggen, Topic modelling exposes disciplinary divergence in research on the nexus between human mobility and the environment, *Humanit Soc Sci Commun* 9 (2022) 34.
- [82] D. Alexander, Social media in disaster risk reduction and crisis management, *Sci Eng Ethics* 20 (2014) 717–733.
- [83] B. Wang, Zhuang J. Rumor response, debunking response, and decision makings of misinformed Twitter users during disasters, *Nat Hazards* 93 (2018) 1145–1162.
- [84] M. Bednarek, A.S. Ross, O. Boichak, Y.J. Doran, G. Carr, E.G. Altmann, T.J. Alexander, Winning the Discursive Struggle? the Impact of a Significant Environmental Crisis Event on Dominant Climate Discourses on Twitter, *Discourse Context Media* 45, 2022 100564.
- [85] K.K. Zander, S.T. Garnett, R. Ogie, M. Alazab, D. Nguyen, Trends in bushfire related tweets during the Australian 'Black Summer' of 2019/20, *For Ecol Manag* 545 (2023) 121274.