Real Time Sentiment Analysis of Online Information for Fast Emergency Response

Michael Lewis
Volgenau School of Engineering

Munira Tabassum
Volgenau School of Engineering

Reeti Bibhuti
Volgenau School of Engineering

***Abstract*—Semantic analysis has been widely researched in the domain of online review sites with the aim of generating summarized opinions of users about different aspects of products. Analyzing such semantics from online social networking sites can help emergency responders understand the dynamics of the network. In this paper, we perform an analysis of tweets posted on Twitter during the disastrous Hurricane Ida and create dashboards based on extracted semantic metadata. The research and development of this product seeks to address the lag times between disaster and disaster response.**

***Index Terms*—Machine learning, Social Media, NLP, Emergency response**

1. Introduction

Social media is the widest used form of open communication today. The speed at which users post and respond to postings is nearly instantaneous. This produces a huge amount of data which includes a wealth of potential information. This data could be used to correlate a specific posting to a specific event. Twitter has been on the rise in recent years as one of the platforms which receive nearly a constant stream of postings. These twitter postings come from all sorts of sources. They come from individuals as well as businesses.

Disaster response to natural disasters has traditionally been centered around population distribution. This does not consider the number of people that evacuated. As a result, humanitarian aid can be improperly distributed over a geographical area. Hurricane Ida struck Louisiana as a category 4 hurricane. This resulted in a death toll of 82 in Louisiana alone. [1] Several of these deaths were indirectly caused by the storm. We believe that leveraging social media data in relation to these events could be used to severely reduce the death toll.

Hurricane Ida made landfall as a major hurricane on the Louisiana coastline on 29 August 2021. It generated damaging winds and storm surge causing widespread damage to structures and to power and telecommunication infrastructure throughout Louisiana. The damage to the housing stock in Louisiana was widespread, variable and in places catastrophic. Structures in coastal communities suffered most from the combined impacts of storm surge and the high wind speeds. Water leaks through failed roofing systems and damaged structures also exacerbated losses by destroying the interior contents of many homes as the water ruined interior wall finishes, flooring, and personal property. The extensive damage to residential structures appears to date to have a disproportionately large impact on rural populations of limited means, a recurring theme from previous post-event studies. [2]

Social media are important channels for information exchange, and we are using the big data from twitter to help better people in need based on twitter data.

2. Background

2.1. background

Tropical storms and hurricanes contain some of the most devastating forces in the world. The impacts of climate change have increased both the frequency and the strength of these storms. A very large portion of the population lives near bodies of water which incidentally is the very location most likely to be hit with one of these storms. The last five hurricane seasons have set a record of being the most active and destructive season. This troubling trend of increasing frequency and strength of hurricanes is a direct threat to all coastal living populations. The table below represents the last five years of hurricanes that have struck from the Atlantic. NOAA classifies a major hurricane as a category 3 or higher (shaded area of table). [3]

TABLE 1. NAME OF HURRICANE

	Cat 1	Cat 2	Cat 3	Cat 4	Cat 5
2021				Ida	
2020	Isaias	sally		Laura	
2019	Barry				Michael
2018	Florence				
2017				Harvy, Irma, Maria	

Ida started forming August 26, 2021, and was in effect till September 5, 2021, Ida intensified early Sunday 29th to a Category 4 hurricane, with maximum sustained winds of 150 miles an hour. Areas affected by Ida were Venezuela, Colombia, Jamaica, Cayman Islands, Cuba, Gulf Coast of the United States (especially Louisiana), East Coast of the United States (especially the Northeastern United States), Atlantic Canada. [4]

On August 23 the National Hurricane Center (NHC) first noted the potential for tropical cyclone development in the

TABLE 2. COSTLIEST HURRICANES TO HIT THE US SINCE 2000
Source: Accuweather

Hurricane	Economic Impact
Katrina	$320B
Maria	$215B
Sandy	$210B
Harvey	$210B
Ivan	$115-125B
Rita	$105-115B
Ida	$95B
Irma	$80B
Ike	$75B
Florence	$50-60B

southwestern Caribbean Sea, related to a tropical wave that entered the eastern Caribbean Sea on the same day. By August 25, the NHC assessed a high likelihood of development as the wave moved westward through the Caribbean. Late on August 26, a Hurricane Hunters flight indicated that the depression intensified into Tropical Storm Ida 130 mi (209 km) Southwest of Grand Cayman, Cayman Islands. Ida then rapidly intensified, with its winds increasing by 35 mph (55 km/h) in just over 11 hours. Late on August 27, the NHC upgraded Ida to Category 1 hurricane status, soon afterward, Ida moved over the warm waters of the Loop Current in the Gulf of Mexico, and the storm intensified into a Category 3 on August 29. Shortly after being upgraded to a major hurricane, Ida began intensifying even more quickly, with the system's minimum central pressure dropping from 955 mbar (28.2 inHg) to 948 mbar (28.0 inHg)in an hour, quickly Ida had further intensified into a Category 4 hurricane, with the storm's sustained winds reaching 130 mph. [4]

As of September 15, a total of 112 deaths were confirmed in relation to Ida, including 96 in the United States and 20 in Venezuela. In the United States, 30 deaths were in Louisiana, 30 in New Jersey, 18 in New York, 5 in Pennsylvania, 3 in Mississippi, 2 in Alabama, 2 in Maryland, 1 in Virginia, and 1 in Connecticut. The storm has caused 43 indirect deaths, including 20 deaths in Venezuela caused by flooding from Ida's precursor. [1] Power outages in the most heavily affected areas were expected to last for up to a month, more than 1 million Louisiana residents were without power. [5] Monday morning. States of emergency were declared for Louisiana and portions of the Northeast and appx $50 billion damage happened because of Ida. [5]

2.2. Problem Description

To look into the aftermath of Hurricane Ida we have to calculate the loss and how much resources we need to recover the loss. Economic impact of hurricane Ida is estimated to be about 95 billion which was calculated beyond the damage to houses and cars. This estimate incorporates variables like the impact of people unable to get to work, travel disruption and a halt to tourism, as well as the cost of clean-up crews, and this makes Ida one of the costliest disasters. [6]

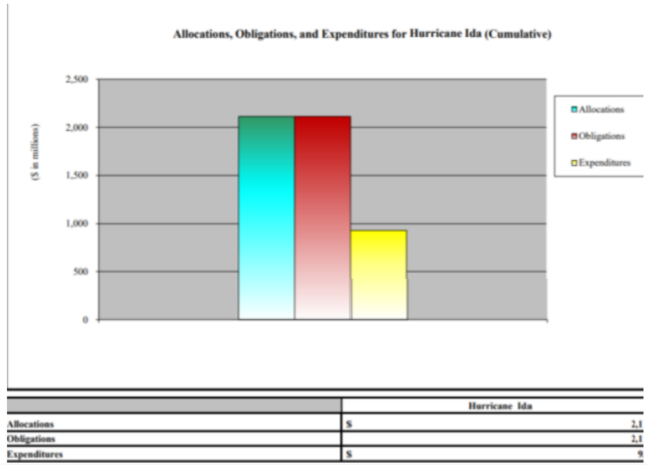

Figure 1. Allocations, Obligations & Expenditures for Hurricane Ida (Cumulative) report by FEMA

According to FEMA report published in October 2021, it already allocated more than 2000 million as disaster management process and has 2000 million more obligations to allocate. [7]

The Federal Emergency Management Agency (FEMA) and the American Red Cross try to provide immediate disaster relief when natural disasters strike. These groups typically respond at the minimum safe location from the disaster. Ground crew efforts are very labor intensive and often rely on the senses of the first responders to detect people in need. Also, studies have shown that increases in mortality and adverse health outcomes due to differentiated disaster response and recovery efforts may promote inequity among populations that receive less aid. [8] A modern solution that can identify impacted areas that have unevacuated people is desired.

3. Reason for Big data solution

Social media is prolific today. The abundance of people posting to each platform provides a large pool of data. This data could be leveraged by a capable group to facilitate several missions for humanitarian projects. This is especially useful in situations where there may be a lack of infrastructure after a disaster. This processed data can help indicate where larger population densities of peoplethose who were unable to evacuate in time may be stranded. This information is very time critical as the difference, in a medical emergency [9], [10], [11], [12], between a few hours and a few days is a matter of life and death. This information can be used to get resources to the people in the event of a larger disaster where people may not be able to be evacuated. The volume and variety of these Twitter posts makes this project an excellent use case scenario for a big data solution.

emotion and sentiment analysis of comments in the social media can be used for fast emergency response [13], [14], [15], [16], [17], [18] . In this work, we used machine learning and NLP models for analytic of online texts based on the NLP models developed by our research mentors at George Mason University and other researchers to tweets with the context of emergency. [19], [20], [21], [22], [23], [24], [25], [26], [27], [28], [29], [30], [31]

However we use applications of machine learning in natural language processing [25], [32], [33], [34], [35], [36], [37], [38], [39], [40], [41], [42], [43], [44], [45], [46], [47] in other domains such as health, security and business, and apply transfer learning models to analyse the online comments for creating a model for emergency response [**?**], [48], [49], [50], [51], [52], [53], [54], [55], [56], [57], [58], [59], [60], [61], [62].

3.1. Problem Statement

Is it possible to leverage twitter posts to accurately determine where humanitarian aid will be required most urgently?

4. Project

4.1. Focus

Twitter data will be analyzed for following points:

- Is this a post indicating a desire to try to weather the storm?
- Is there an indication of urgent need?

4.2. Phases

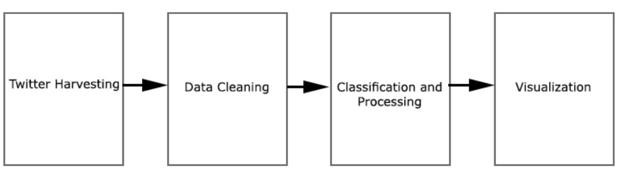

Figure 2. Phases

Python was leveraged to scrape the Twitter API. Since the topic of this project is a past event, a premium access was required to utilize the Twitter archive. Two versions of the Twitter harvester were created. Version one is for scraping a past event from twitter archives. Version two is for real time scraping. This real time scraper is the intended product as it will allow a live stream of Twitter data to be automatically harvested at a regular interval. This live stream is adjustable to handle differing velocities of Twitter data. The rate limit for this is 500 records in any single request. Testing of version two results in a reliable return as refined as 500 records for every 30 second interval. The twitter raw data is stored in a json file with the tweet data, place data (when available), and some statistics about the query. The version one harvester has scraped approximately

Figure 3. Tweet processor

Figure 4. Location Processor

200,000 tweets relating to Hurricane Ida as of September 4th, 2021.

These tweets are held in individual json files which are then picked up by a data cleaning python script. In this script, the raw json file is analyzed tweet by tweet. When place data does not exist for a tweet, a null value is created for the geo value. The keywords from the message body are tokenized into individual keywords and held in a nested json object. Additionally, we leverage SpaCy libraries in python to extract microevents from the tweet as a form of semantic extraction. These microevents are also tokenized and stored as a nested json object. The places object in the raw data is also analyzed and saved to a processed places json file. This places file is used as a places dictionary. All raw data is moved to a raw archive folder upon being processed. The processed json data gets placed into a directory for Elasticsearch ingestion.

These tweets are held in individual json files which are then picked up by a data cleaning python script. In this script, the raw json file is analyzed tweet by tweet. When place data does not exist for a tweet, a null value is created for the geo value. The keywords from the message body are tokenized into individual keywords and held in a nested json object. The places object in the raw data is also analyzed and saved to a processed places json file. This places file is used as a places dictionary. All raw data is moved to a raw archive folder upon being processed. The processed json data gets placed into a directory for Elasticsearch ingestion.

The third step in our pipeline is to push the tweet data directly into an Elasticsearch cluster. Every tweet currently

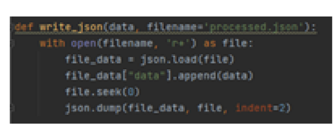

Figure 5. Processed JSON writer

Extracted tweet structure

@type – Social media type (only tweet as of December 2021)

@id – UUID of the posting

@geo – UUID of the place of the posting (see Place structure for Places data)

@rt – is a retweet/repost

@author – Authors UUID

@content – Raw textual content of the posting

@posted – Datetime of the posting

@keywords – Extracted Tags of the post

@mentions – Extracted mentions in the post

@extracts – Extracted semantic information from the post (using spacy to pull out events like power failures and flooding)

Place Structure

@id – UUID of the place

@place – Textual location of the place ex. Baton Rouge, LA

Figure 6. Extracted Tweet Structure

keys off a unique identifier that is included in the original raw data. This prevents the cluster from receiving duplicate data (since the harvester has run several queries that included some overlapping times). Since we are using an automated process for harvesting and processing, there is not a need to worry about data integrity issues with the duplicate tweets. Our mappings for the Elasticsearch index were dynamically created based on the detected datatype of the json fields.

In addition to utilizing Elasticsearch, the cluster also hosts Kibana. Kibana is an excellent tool that allows for the rapid visualization of the data held within the Elasticsearch indexes. Since we have the posted date of each tweet, it is possible to visualize narrow timeframes extremely quickly with a low amount of effort. It is possible to create dashboards that are specially filtered for the time critical tweets. Kibana is the front-end visualization tool of our twitter metadata project. It is an open-source piece of software that is commonly packaged as part of an Elasticsearch stack. Since it is native to the Elasticsearch stack, no additional coding was required to interface with our data. Kibana leverages defined queries to an Elasticsearch index to provide dynamic dashboards.

Since the focal point of this project was Hurricane Ida, we had to test the Kibana suite using absolute time windows. The increase in trends of mentions of the hurricane, along with semantic event metadata that was extracted from the tweets, we were able to show an increase in certain events (like power grid issues and flooding) as the hurricane made landfall in the United States. There are several other semantic events that could be explored and visualized in event specific dashboards. Our semantic extractions revealed that the most common tweets involved flooding. wind, and power failures. We had several extractions indicating that the users were safe or in a safe area. These extractions are easily filtered within Kibana. This allows custom dashboards to be made for specific semantic events and should help streamline the response of first responders.

5. Related Works

Innumerable research projects concerning the analysis of twitter data have been published since this particular social media started to generate valuable data which is easily accessible for developers and analyst. Different analyses approaches have been applied to find solutions to many problems existing in various communities and with existing vast data range it is becoming a great tool to assess patterns of solution for any problem. For determining a better solution to disaster management twitter data scraping has become popular in recent times as twitter was created only 15 years ago. As the popularity of social media, especially twitter grows, the dataset and our ability to analyze the vast data grows. A lot of research has been done about natural disaster management from twitter data in recent times which has provided valuable inputs to our own research. The very first paper we came across is “An Evaluation of Geotagged Twitter Data during Hurricane Irma Using Sentiment Analysis and Topic Modeling for Disaster Resilience” [63]. This paper focused on twitter users’ sentiment among Irma affected victims with keywords such as “hurricaneirma”, “Irma”. A sentiment library has been used to calculate sentiment for this study where words are assigned a score representing their relative polarity and subjectivity. The scores of words the has been averaged for the total of each document. The polarity and subjectivity of each tweet has been calculated for each tweet and a string-based value has been set as sentiment. So, the authors use tokenization and sort raw data accordingly. For example, they consider a tweet’s sentiment as “positive” if the value for polarity is greater than zero. The polarity is returned as “negative” when polarity is less than zero. Data with a polarity value close to zero is considered “neutral”, because no emotion or opinion is conveyed. Then they use different topic modelling to find the clusters/ group from the data. Based on this cluster they have provided visualization for the results. Basically, the study shows a representation of people’s emotional response

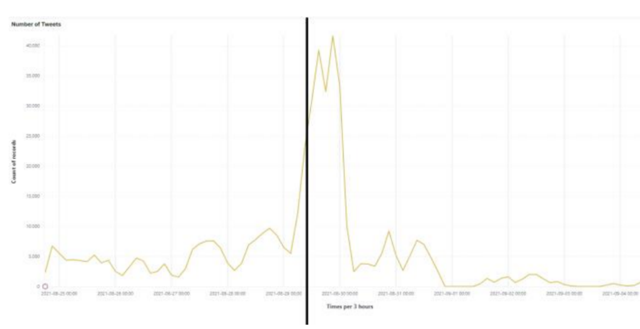

Figure 7. Hurricane Ida tweets during the storm (black line is US landfall)

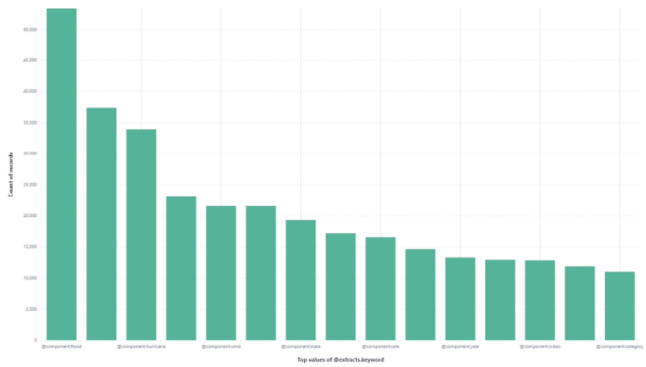

Figure 8. Top semantic extractions for data exploration

during hurricane Irma but didn't show the possibility of using this data to improve disaster response.

Author Catherine M "Using Twitter for crisis communications in a natural disaster: Hurricane Harvey" [64], this paper examines Mayesr Tuner communication through his twitter account during Hurricane Harvey crisis to improve understanding of the restorative rhetoric with natural disasters. This research paper focuses on understanding how social media can be used as a communication tool during a natural disaster.

One of the earliest researches on twitter and disaster data is based on Hurricane Sandy. We analyze the paper "The Hurricane Sandy Twitter Corpus" [65] and again we see the authoris using twitter to identify the most effective way of communication as it collaborates as a widely accessible public news outlet as well as the fastest information dispersion mechanisms. Twitter allows post-hoc data collection. So, this is used to gather knowledge about the disaster affected area and people. This means it also enables moment-by-moment, geo-specific and participant-oriented situation analysis which can be used to further improve response management in the future.

However, one of the most helpful and organized articles we found is "Social and geographical disparities in Twitter use during Hurricane Harvey" [66]. The authors use Hadoop Distributed File System to store raw twitter data in a MongoDB database then use Apache Spark to process and analyze this large dataset. They also analyze papers that use twitter data of Hurricane Irma affected to implement their own data and they have calculated two databases where they have removed the keyword "Irma" with "Harvey". Then they calculate Ratio and Sentiment indexes from these two Harvey databases. To calculate Ratio index, they have divided the number of disaster-related tweets by the number of background tweets and for the Sentiment index, they calculate average sentiment scores. This paper has helped us navigate our own project and has provided ways to navigate our raw data. These indexes also help in estimating damage, flood mapping, and assessing citizen awareness on climate change. This gives them four phases of emergency management, Preparedness, Response, Mitigation, Recovery but it also shows affected people's socio-economic condition plays a vital role in receiving aid.

Both papers "Social media and disasters: a functional framework for social media use in disaster planning, response, and research" [67] and "Mining Twitter Data to Understand the Human Sentiment on Hurricane Florence" [68] researches to understand disaster affected peoples' sentiment and effective ways to quickly disburse aid. The study of Hurricane Florence has given us another view into collecting and analyzing twitter data. We are planning to improve and implement a disaster planning framework from Houston J and Hawthorne. [68] They point out many stages of aid via social media such as providing and receiving disaster warnings, detecting affected areas with traffic data, sending and receiving emergency response requests, and most importantly the aftermath of disaster recovery.

As we are working to find better ways to respond to hurricane Ida affected people, we investigate the main challenges of response and aid after a disaster. In the book "Facing Hazards and Disasters: Understanding Human Dimensions" it is discussed that paralleling preparedness measures, disaster response activities take place at various units of analysis, from individuals and households, to organizations, communities, and intergovernmental systems is the most important aspect of disaster management and it highlights key themes in the literature, with an emphasis on NEHRP-based findings. [69] While looking at geotagging we have come across a book that discusses Geographic Information Systems (GIS), Remote Sensing (RS), and Global Positioning Systems (GPS) applications and utilization in the entire disaster management cycle as a tool to support decision making. This includes risk assessment, monitoring and detection with Geographic Information Systems/Remote Sensing, Information Sharing for Decision-Support and Risk Communication, Vulnerability Mapping, Technological Hazards and Security. [70] Other keys to addressing any of these challenges while conducting disaster research we need foresight and preparation. It is shown in another study that very few researchers seek to empirically tackle ethical or methodological considerations for this type of research, instead broaching the topic as anecdotes in their studies. [71]

In the paper "Twitter Analytics for Disaster Relevance and Disaster Phase Discovery" the authors propose a general framework for a cloud-based Twitter analytics platform for identification relevant to disaster phase discovery. This paper focuses on three major hurricanes and specially focused on studying three main disaster phases: disaster preparedness, disaster response, and disaster recovery. The way this paper classifies the tweets is to create a system that consists of three main components of Twitter analytics which is relevance classification, disaster phase classification and knowledge extraction. The authors demonstrate a general classifier with good accuracy around 86% to classify relevant tweets from a hurricane disaster which we use to classify the tweets from hurricane Ida. From their perspective, disaster phase discovery using multi-class text classification is a better choice to classify tweets. [72]

In "Using geotagged tweets to track population movements to and from Puerto Rico after Hurricane Maria" the authors examine the suitability of Twitter data for measuring post-disaster population mobility using the case of Hurricane Maria in Puerto Rico. [73] This paper is very important in our own research as we are also trying to find way to find geotags and location of hurricane Ida affected population through twitter scraping. This paper also merges twitter data and data from other resources to cross check their findings from tweets about displaced populations in the aftermath of hurricane Maria. The results presented in this paper confirm the potential of Twitter data to estimate the magnitude, timing, destination, and return of the displaced, as well as the number of non-residents arriving in Puerto Rico. Though the authors express the limitation of their data as male biased as the twitter dataset they use consists mostly of male users, they calculate that the hurricane results in 8.3% off-island displacement, which was 4% of their Twitter sample dataset.

It is imperative for our research that we can clean out any rumors and process the tweet in a way that it does not affect our data and corrupt our results. According to Wang and Zhuang (2018), over 85% of Twitter users who were exposed to false information during disasters responded by spreading the falsehoods due to the lack of debunking information at the time of their post. [74] Even if the users receive debunking information, between 78 and 97% of the spreaders do not delete nor clarify their tweets and confirm that their previous tweet is a result of false news or rumors which make it very hard to analyze any disaster affected area through tweets. Another research paper "Attention to misleading and contentious tweets in the case of Hurricane Harvey" focuses on popular yet ambiguous and contentious narratives transmitted via Twitter during Hurricane Harvey and it focuses on how misleading those tweets can be. The authors use Hurricane Harvey Twitter Dataset, available from the University of North Texas Libraries and cross check with 44 online news and fact-checking articles related to rumors and false claims regarding Hurricane Harvey to determine credibility and authenticity. [75]

"A Twitter Tale of Three Hurricanes: Harvey, Irma, and Maria" By Alam F et al. focuses on multimedia content analysis as well as textual analysis with topic modeling and sentiment analysis of millions of twitter data. This is one of the most advanced and detailed visualization research papers we have come across while reviewing related works by our peers. They employ various Artificial Intelligence techniques from Natural Language Processing and Computer Vision fields, which exploit different machine learning algorithms to process the data generated during the disaster events of three hurricanes of 2017. They classify images posted in social media into different categories to determine damage. Another crucial part of their classification is that they state that we can get damage information related to power lines, roads, etc. from images whereas the corresponding text reports a dead person, questions why hurricanes are named, or mentions the path of the hurricane. Even though many picture/media contents attached to tweets do not show any damage content, they provide critical information for other humanitarian categories taxonomy such as valuable insight and illustration for the quality of shelter, evacuation, and displaced people. [76]

Reviewing textual twitter content to be trustworthy is a challenge as the vast dataset can have untrue and false information. To ensure the accountability of twitter as a data source, Venkata K. et al analyzes textual and user account profile to determine authenticity of dataset in their research paper "Retweetability Analysis and Prediction during Hurricane Sandy". They extract features from tweets' content and user account information and perform experiments to develop models that automatically predict the retweetability of a tweet in the context of the Hurricane Sandy. The results of their experiments using different threshold values for labeling a tweet as highly retweetable show improved performance for classifiers trained using the combination of tweet content features and user details features. This features over classifiers that are trained on each feature type independently to identify which tweet is trustworthy and which tweet seems to be untrue which helps promoting good information and could be used to stop spread of misinformation. [77]

A great advantage of using twitter or any social media tool to determine damage of any natural disaster is that the data is free and accessible. This calculation or estimation of damage using twitter is very cheap since the data is free and it was proven that twitter's damage mapping was better than FEMA's damage modelling in predicting the location and severity of damage. [78]

Another real time data modelling approach is by Bica M. and her approach is identifying and studying uncertainty around hurricane risk information is bottom-up, driven by people and the data they produce in social media by collecting contextual data, applying context-sensitive methods, and iterating between the micro scale of individual activity and the macro scale of social dynamics and implementation of these agendas based on a dataset of tweets during the 2017 Atlantic hurricane season. [79]

Sit A. et al in their research paper "Identifying disaster-related tweets and their semantic, spatial and temporal context using deep learning, natural language processing and spatial analysis: a case study of Hurricane Irma" define their specific contributions to the studies that focus on the classification of crisis related information from social media data by comparing methodology and results with the previous work done by Burel, Saif, and Alani (2017), Nguyen et al. (2016), Olteanu, Vieweg, and Castillo (2015) and others.They introduce an analytical framework for analyzing tweets to identify and categorize fine-grained details about a disaster. They also categorize affected individuals, damaged infrastructure, and disrupted services as well as impacted areas and time periods, and relative prominence of each category of disaster-related information using Hurricane Irma twitter data. Their twitter dataset is consisted of over 500 million keyword-based and geo-located collection of tweets before, during and after the disaster and to train this dataset, they performed classification experiments using Logistic Regression, Support Vector Machines (SVM), Convolutional Neural Networks (CNNs) and Long Short-Term Memory

(LSTM) Networks. Their results highlight potential areas with high density of affected individuals and infrastructure damage throughout the temporal progression of the disaster. [80], [81], [82]

6. Project Challenges

The main challenge was finding the adequate data to analyze and research from Twitter, initially we were able to get a few thousands records but we were able to get more than 1 million Twitter data related to hurricane Ida. The other challenges we faced to extract exact geolocation of the users.

A major setback occurred during the semantic extraction development. The natural language processing that was initially used revealed a bug in the code which resulted in a 25% error rate in the semantic extraction of microevents. This was not caught during the testing of the code owing to the small batch of test data. As a result, the pipeline had to rerun the processing of the raw tweets again resulting in a 9-day setback.

The current code baseline does not attempt to validate any information. As a result, misinformation may impact the behaviors of the visualizations. This product would not be useful for a category that is prone to misinformation (ex. Covid-19).

Geospatial data was not as widespread as initially hoped. As a result, development of a geospatial tool for data visualization was scrapped. Our extractions proved to have less than one percent of posts to include geospatial information.

7. Discussion

The intent of our research paper is to provide help to hurricane affected people in different regions, the people who need humanitarian aid urgently like the need of medical aid, food supply and need evacuation from certain areas. Twitter is a free source of the data and being used by 73 million users in the USA itself including individuals, communities, governments and news media, it provides a large volume of the data that we can leverage to analyze the twitter data to find the location and provide help to people who are in need. It can help organizations who want to help users in need to be able to provide help in a timely manner. Our metadata analysis application is a promising start to a mission application for a customer who requires analysis of smaller events within a larger event tag. One of the shortfalls was the lack of geospatial data included in the tweets. Since less than one percent of the tweets contained place information, the geospatial visualizations were scrapped.

8. Future Works

A future improvement of this platform will include the geospatial visualization with the caveat that it may only contain one percent or less of the total tweet data. This test case focused on Hurricane Ida tweet data. In the process of conducting this project, additional future improvements were detected. The future of this application includes creating a collections service which would be used for users to define a major event that they would like to collect tweet data on. Future data sources will include other social media platforms like Facebook and Instagram but will require additional ingestion engineering to accomplish. There is also a need for a search function where a user would be able to search what events we are currently collecting. A service for visualizations will also be needed as Kibana cannot be expected to be the user facing product. It is possible to replicate Kibana dashboards out to users via a web application. This product will likely be useful for agencies that would benefit from a faster response to events. Two likely groups that would benefit from this application are FEMA and the American Red Cross.

Acknowledgments

This research is the result of the support and guidance provided by Dr. Ebrima Ceesay of George Mason University. His professionalism, mentorship and expertise proved critical throughout all stages of this project. We also would like to thank his detailed and constructive comments, and for helping us with the data collection task.

References

[1] S. Reardon. (September 9, 2021) “cbs news : Hurricane ida death toll jumps to 82”. [Accessed October 9, 2021]. [Online]. Available: https://www.cbsnews.com/news/hurricane-ida-death-toll-update-82-louisiana-northeast/

[2] D. Prevatt, S. Kameshwar, D. Roueche, B. Rittelmeyer, T. Duarte, T. Heo, H. Ibrahim, S. Klepac, O. Lafontaine, T. Lin, L. Manuel, S. Pilkington, Y. Pinyochotiwong, J. Santiago-Hernandez, S. Strader, K. Gurley, T. Kijewski-Correa, K. Mosalam, and I. Robertson, “Steer: Hurricane ida joint preliminary virtual reconnaissance report-early access reconnaissance report (pvrr-earr),” 2021. [Online]. Available: https://www.designsafe-ci.org/data/browser/public/designsafe.storage.published/PRJ-3268/\#details-56471329073917460-242ac114-0001-012

[3] National Oceanic and Atmopspheric Administration U.S. Department of Commerce. (November 22, 2019) “noaa’s national ocean service: High tide bulletin archive. noaa’s national ocean service: Recent hurricane archive.”. [Accessed October 9, 2021]. [Online]. Available: https://oceanservice.noaa.gov/hazards/hurricanes/hurricane-archive.html

[4] Brown and Stewart. “tropical weather outlook (august 23, 2021)” (txt). [Accessed October 9, 2021]. [Online]. Available: https://www.nhc.noaa.gov/archive/2021/MARTY.shtml

[5] R. Santana, M. Deslatte, and J. McConaughey. (September 2, 2021) “after ida, small recovery signs amid daunting destruction”. [Accessed October 9, 2021]. [Online]. Available: https://apnews.com/article/business-environment-and-nature-louisiana-392d8001a9c69df607d9a7b5b84474fa

[6] P. Stevens. (September 8, 2021) “cnbc : Hurricane ida’s damage tally could top $95 billion, making it 7th costliest hurricane since 2000.”. [Accessed November 7, 2021]. [Online]. Available: https://www.cnbc.com/2021/09/08/hurricane-idas-damage-tally-could-top-95-billion-making-it-7th-costliest-hurricane-since-2000-.html

[7] F. E. M. A. U. Gov. (October 7, 2021) "disaster relief fund: Monthly reports". [Accessed November 7, 2021]. [Online]. Available: \url{https://www.fema.gov/sites/default/files/documents/fema_disaster-relief-fund-report_102021.pdf}

[8] C. E. Willison, P. M. Singer, M. S. Creary, and S. L. Greer, "Quantifying inequities in US federal response to hurricane disaster in texas and florida compared with puerto rico," *BMJ Global Health*, vol. 4, no. 1, p. e001191, Jan. 2019. [Online]. Available: https://doi.org/10.1136/bmjgh-2018-001191

[9] P. Khoshpouri, M. Ghadimi, R. Rezvani Habibabadi, M. Motaghi, B. A. Venkatesh, M. Shaghaghi, A. Pandey, B. Hazhirkarzar, S. Ameli, M. A. Ghasabeh, P. Pandey, and I. R. Kamel, "Cross-sectional imaging in patients with primary sclerosing cholangitis: Single time-point liver or spleen volume is associated with survival," *European Journal of Radiology*, vol. 132, p. 109331, 2020. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0720048X20305209

[10] A. Baghdadi, M. Ghadimi, S. Mirpour, B. Hazhirkarzar, M. Motaghi, T. M. Pawlik, and I. R. Kamel, "Imaging neuroendocrine tumors: Characterizing the spectrum of radiographic findings," *Surgical Oncology*, vol. 37, p. 101529, 2021. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0960740421000189

[11] R. R. H. M. M. M. S. P. K. A. P. B. H. S. A. M. A. G. A. B. . I. R. K. Maryam Ghadimi, Pallavi Pandey, "Role of volumetric multiparametric mri in distinguishing between intraductal papillary mucinous neoplasms and serous cystadenoma," *Abdominal Radiology*, 2020.

[12] R. W. England, M. Motaghi, B. Kohler, T. E. Hemmingson, A. W. Wu, and C. R. Weiss, "Development and preliminary validation of the patient-reported outcome measure for vascular malformation questionnaire: A prospective cohort study," *Journal of Vascular and Interventional Radiology*, vol. 32, no. 5, pp. 683–690.e4, 2021. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S1051044321002827

[13] S. Zad, M. Heidari, H. James Jr, and O. Uzuner, "Emotion detection of textual data: An interdisciplinary survey," in *2021 IEEE World AI IoT Congress (AIIoT)*. IEEE, 2021, pp. 0255–0261.

[14] L. Cui and D. Lee, "Coaid: COVID-19 healthcare misinformation dataset," *CoRR*, vol. abs/2006.00885, 2020. [Online]. Available: https://arxiv.org/abs/2006.00885

[15] D. M. J. Lazer, M. A. Baum, Y. Benkler, A. J. Berinsky, K. M. Greenhill, F. Menczer, M. J. Metzger, B. Nyhan, G. Pennycook, D. Rothschild, M. Schudson, S. A. Sloman, C. R. Sunstein, E. A. Thorson, D. J. Watts, and J. L. Zittrain, "The science of fake news," vol. 359, no. 6380, pp. 1094–1096, Mar. 2018. [Online]. Available: https://doi.org/10.1126/science.aao2998

[16] Q. Su, M. Wan, X. Liu, and C.-R. Huang, "Motivations, methods and metrics of misinformation detection: An NLP perspective," vol. 1, no. 1-2, p. 1, 2020. [Online]. Available: https://doi.org/10.2991/nlpr.d.200522.001

[17] S. Cresci, R. D. Pietro, M. Petrocchi, A. Spognardi, and M. Tesconi, "The paradigm-shift of social spambots: Evidence, theories, and tools for the arms race," in *Proceedings of the 26th International Conference on World Wide Web Companion, Perth, Australia, April 3-7, 2017*, 2017, pp. 963–972. [Online]. Available: https://doi.org/10.1145/3041021.3055135

[18] S. Zad, M. Heidari, J. H. Jones, and O. Uzuner, "A survey on concept-level sentiment analysis techniques of textual data," in *2021 IEEE World AI IoT Congress (AIIoT)*. IEEE, 2021, pp. 0285–0291.

[19] M. Heidari, H. James Jr, and O. Uzuner, "An empirical study of machine learning algorithms for social media bot detection," in *2021 IEEE International IOT, Electronics and Mechatronics Conference (IEMTRONICS)*. IEEE, 2021, pp. 1–5.

[20] C. Yang, R. C. Harkreader, and G. Gu, "Empirical evaluation and new design for fighting evolving twitter spammers," *IEEE Trans. Information Forensics and Security*, vol. 8, no. 8, pp. 1280–1293, 2013. [Online]. Available: https://doi.org/10.1109/TIFS.2013.2267732

[21] J. Devlin, M. Chang, K. Lee, and K. Toutanova, "BERT: pre-training of deep bidirectional transformers for language understanding," in *Proceedings of the 2019 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies, NAACL-HLT 2019, Minneapolis, MN, USA, June 2-7, 2019, Volume 1 (Long and Short Papers)*, J. Burstein, C. Doran, and T. Solorio, Eds. Association for Computational Linguistics, 2019, pp. 4171–4186. [Online]. Available: https://doi.org/10.18653/v1/n19-1423

[22] D. S. Khoury, D. Cromer, A. Reynaldi, T. E. Schlub, A. K. Wheatley, J. A. Juno, K. Subbarao, S. J. Kent, J. A. Triccas, and M. P. Davenport, "Neutralizing antibody levels are highly predictive of immune protection from symptomatic SARS-CoV-2 infection," *Nature Medicine*, vol. 27, no. 7, pp. 1205–1211, May 2021. [Online]. Available: https://doi.org/10.1038/s41591-021-01377-8

[23] J. Havey. "pharma research progress hope,". [Online]. Available: https://catalyst.phrma.org/a-year-and-a-half-later-the-biopharmaceutical-industry-remains-committed-to-beating-covid-19?utm_campaign=2021-q3-cov-inn&utm_medium=pai_srh_cpc-ggl-adf&utm_source=ggl&utm_content=clk-pol-tpv_scl-geo_std-usa-dca-pai_srh_cpc-ggl-

[24] M. Heidari, S. Zad, B. Berlin, and S. Rafatirad, "Ontology creation model based on attention mechanism for a specific business domain," in *2021 IEEE International IOT, Electronics and Mechatronics Conference (IEMTRONICS)*. IEEE, 2021, pp. 1–5.

[25] N. Sallahi, H. Park, F. E. Mellouhi, M. Rachdi, I. Ouassou, S. Belhaouari, A. Arredouani, and H. Bensmail, "Using unstated cases to correct for COVID-19 pandemic outbreak and its impact on easing the intervention for qatar," *Biology*, vol. 10, no. 6, p. 463, May 2021. [Online]. Available: https://doi.org/10.3390/biology10060463

[26] M. El-Harbawi, B. B. Samir, M.-R. Babaa, and M. I. A. Mutalib, "A new QSPR model for predicting the densities of ionic liquids," *Arabian Journal for Science and Engineering*, vol. 39, no. 9, pp. 6767–6775, Jun. 2014. [Online]. Available: https://doi.org/10.1007/s13369-014-1223-3

[27] M. Heidari and S. Rafatirad, "Semantic convolutional neural network model for safe business investment by using bert," in *2020 Seventh International Conference on Social Networks Analysis, Management and Security (SNAMS)*. IEEE, 2020, pp. 1–6.

[28] P. R. Krause, T. R. Fleming, R. Peto, I. M. Longini, J. P. Figueroa, J. A. C. Sterne, A. Cravioto, H. Rees, J. P. T. Higgins, I. Boutron, H. Pan, M. F. Gruber, N. Arora, F. Kazi, R. Gaspar, S. Swaminathan, M. J. Ryan, and A.-M. Henao-Restrepo, "Considerations in boosting COVID-19 vaccine immune responses," *The Lancet*, vol. 398, no. 10308, pp. 1377–1380, Oct. 2021. [Online]. Available: https://doi.org/10.1016/s0140-6736(21)02046-8

[29] J. H. Kim, F. Marks, and J. D. Clemens, "Looking beyond COVID-19 vaccine phase 3 trials," *Nature Medicine*, vol. 27, no. 2, pp. 205–211, Jan. 2021. [Online]. Available: https://doi.org/10.1038/s41591-021-01230-y

[30] E. C. Fernández and L. Y. Zhu, "Racing to immunity: Journey to a COVID-19 vaccine and lessons for the future," *British Journal of Clinical Pharmacology*, vol. 87, no. 9, pp. 3408–3424, Jan. 2021. [Online]. Available: https://doi.org/10.1111/bcp.14686

[31] M. Heidari, S. Zad, and S. Rafatirad, "Ensemble of supervised and unsupervised learning models to predict a profitable business decision," in *2021 IEEE International IOT, Electronics and Mechatronics Conference (IEMTRONICS)*. IEEE, 2021, pp. 1–6.

[32] Code Academy. "natural language processing/text preprocessing". [Online]. Available: https://www.codecademy.com/learn/dscp-natural-language-processing/modules/dscp-text-preprocessing/cheatsheet

[33] DeepAI. "named-entity recognition". [Online]. Available: https://deepai.org/machine-learning-glossary-and-terms/named-entity-recognition

[34] C. on foreign relations. "the u.s. war in afghanistan". [Accessed December 9, 2021]. [Online]. Available: https://www.cfr.org/timeline/us-war-afghanistan

[35] P. Hajibabaee, M. Malekzadeh, , M. Heidari, S. Zad, O. Uzuner, and J. H. Jones, "An empirical study of the graphsage and word2vec algorithms for graph multiclass classification," in *2021 IEEE 12th Annual Information Technology, Electronics and Mobile Communication Conference (IEMCON)*. IEEE, 2021.

[36] "covid-19 third doses — health". [Accessed December 5, 2021]. [Online]. Available: https://www.fairfaxcounty.gov/health/novel-coronavirus/vaccine/third-doses

[37] S. Zad, M. Heidari, P. Hajibabaee, and M. Malekzadeh, "A survey of deep learning methods on semantic similarity and sentence modeling," in *2021 IEEE 12th Annual Information Technology, Electronics and Mobile Communication Conference (IEMCON)*. IEEE, 2021.

[38] S. Akon and A. Bhuiyan, "Covid-19: Rumors and youth vulnerabilities in bangladesh," 07 2020.

[39] M. Heidari, S. Zad, M. Malekzadeh, P. Hajibabaee, , S. HekmatiAthar, O. Uzuner, and J. H. J. Jones, "Bert model for fake news detection based on social bot activities in the covid-19 pandemic," in *2021 12th IEEE Annual Ubiquitous Computing, Electronics Mobile Communication Conference (UEMCON)*. IEEE, 2021.

[40] "suicide statistics". [Accessed December 9, 2021]. [Online]. Available: https://www.befrienders.org/suicide-statistics

[41] T. V. Green and C. Doherty. "majority of u.s. public favors afghanistan troop withdrawal; biden criticized for his handling of situation". [Online]. Available: https://www.pewresearch.org/fact-tank/2021/08/31/majority-of-u-s-public-favors-afghanistan-troop-withdrawal-biden-criticized-for-his-handling-of-situation/

[42] M. U. Islam, F. B. Ashraf, A. I. Abir, and M. A. Mottalib, "Polarity detection of online news articles based on sentence structure and dynamic dictionary," in *2017 20th International Conference of Computer and Information Technology (ICCIT)*. IEEE, Dec. 2017. [Online]. Available: https://doi.org/10.1109/iccitechn.2017.8281777

[43] M. Malekzadeh, P. Hajibabaee, M. Heidari, S. Zad, O. Uzuner, and J. H. Jones, "Review of graph neural network in text classification," in *2021 12th IEEE Annual Ubiquitous Computing, Electronics Mobile Communication Conference (UEMCON)*. IEEE, 2021.

[44] E. Kiely and R. Farley. "timeline of u.s. withdrawal from afghanistan". [Online]. Available: https://www.factcheck.org/2021/08/timeline-of-u-s-withdrawal-from-afghanistan/

[45] Real Python. "tokenization in spacy". [Online]. Available: https://realpython.com/natural-language-processing-spacy-python/\#tokenization-in-spacy

[46] D. Smeltz and E. Sullivan. (August 9, 2021) "us public supports withdrawal from afghanistan". [Online]. Available: https://www.thechicagocouncil.org/commentary-and-analysis/blogs/us-public-supports-withdrawal-afghanistan

[47] spaCy. "spacy 101: Everything you need to know". [Online]. Available: https://spacy.io/usage/spacy-101

[48] Code Academy. "text preprocessing". [Online]. Available: https://www.codecademy.com/courses/text-preprocessing/lessons/text-preprocessing/exercises/introduction

[49] J. Howard. "could covid-19 vaccine boosters be necessary? here's what experts are saying". [Accessed October 2021]. [Online]. Available: https://www.wesh.com/article/could-covid-19-vaccine-boosters-be-necessary-heres-what-experts-are-saying/36519793

[50] A. Ain, "The WHO is right to call a temporary halt to COVID vaccine boosters," *Nature*, vol. 596, no. 7872, pp. 317–317, Aug. 2021. [Online]. Available: https://doi.org/10.1038/d41586-021-02219-w

[51] E. Callaway, "COVID vaccine boosters: the most important questions," *Nature*, vol. 596, no. 7871, pp. 178–180, Aug. 2021. [Online]. Available: https://doi.org/10.1038/d41586-021-02158-6

[52] A. Weatherton. "health expert says booster shot could be needed after getting covid-19 vaccine". [Accessed June 8, 2021]. [Online]. Available: https://www.13newsnow.com/article/life/booster-shot-may-be-needed-after-covid-19-vaccine/291-49a8966c-3d91-48ad-99a0-02905c5593cc

[53] P. Naaber, L. Tserel, K. Kangro, E. Sepp, V. Jürjenson, A. Adamson, L. Haljasmägi, A. P. Rumm, R. Maruste, J. Kärner, J. M. Gerhold, A. Planken, M. Ustav, K. Kisand, and P. Peterson, "Dynamics of antibody response to BNT162b2 vaccine after six months: a longitudinal prospective study," *The Lancet Regional Health - Europe*, vol. 10, p. 100208, Nov. 2021. [Online]. Available: https://doi.org/10.1016/j.lanepe.2021.100208

[54] M. Heidari, J. H. Jones, and O. Uzuner, "Deep contextualized word embedding for text-based online user profiling to detect social bots on twitter," in *2020 International Conference on Data Mining Workshops (ICDMW)*. IEEE, 2020, pp. 480–487.

[55] S. J. Thomas, E. D. Moreira, N. Kitchin, J. Absalon, A. Gurtman, S. Lockhart, J. L. Perez, G. P. Marc, F. P. Polack, C. Zerbini, R. Bailey, K. A. Swanson, X. Xu, S. Roychoudhury, K. Koury, S. Bouguermouh, W. V. Kalina, D. Cooper, R. W. Frenck, L. L. Hammitt, Özlem Türeci, H. Nell, A. Schaefer, S. Ünal, Q. Yang, P. Liberator, D. B. Tresnan, S. Mather, P. R. Dormitzer, U. Şahin, W. C. Gruber, and K. U. Jansen, "Safety and efficacy of the BNT162b2 mRNA covid-19 vaccine through 6 months," *New England Journal of Medicine*, vol. 385, no. 19, pp. 1761–1773, Nov. 2021. [Online]. Available: https://doi.org/10.1056/nejmoa2110345

[56] E. Dolgin, "COVID vaccine immunity is waning — how much does that matter?" *Nature*, vol. 597, no. 7878, pp. 606–607, Sep. 2021. [Online]. Available: https://doi.org/10.1038/d41586-021-02532-4

[57] "virginia open data portal". [Accessed November 6, 2021]. [Online]. Available: https://data.virginia.gov/Government/VDH-COVID-19-PublicUseDataset-Vaccines-DosesAdmini/28k2-x2rj

[58] U.S. Food and Drug Administration. "covid-19 vaccines". [Accessed November 6, 2021]. [Online]. Available: https://www.fda.gov/emergency-preparedness-and-response/coronavirus-disease-2019-covid-19/covid-19-vaccines

[59] M. Heidari and S. Rafatirad, "Bidirectional transformer based on online text-based information to implement convolutional neural network model for secure business investment," in *2020 IEEE International Symposium on Technology and Society (ISTAS)*. IEEE, 2020, pp. 322–329.

[60] populationU. "populationU". [Accessed October 10, 2021]. [Online]. Available: https://www.populationu.com/us/virginia-population

[61] V. D. of Health. "covid-19 vaccine summary – coronavirus". [Accessed October 10, 2021]. [Online]. Available: https://www.vdh.virginia.gov/coronavirus/covid-19-in-virginia/covid-19-vaccine-summary/

[62] M. S. Berg. "what doctors wish patients knew about covid-19 herd immunity". [Accessed October 10, 2021]. [Online]. Available: https://www.ama-assn.org/delivering-care/public-health/what-doctors-wish-patients-knew-about-covid-19-herd-immunity

[63] I. Vayansky, S. A. Kumar, and Z. Li, "An evaluation of geotagged twitter data during hurricane irma using sentiment analysis and topic modeling for disaster resilience," in *2019 IEEE International Symposium on Technology and Society (ISTAS)*. IEEE, Nov. 2019. [Online]. Available: https://doi.org/10.1109/istas48451.2019.8937859

[64] C. M. Vera-Burgos and D. R. G. Padgett, "Using twitter for crisis communications in a natural disaster: Hurricane harvey," *Heliyon*, vol. 6, no. 9, p. e04804, Sep. 2020. [Online]. Available: https://doi.org/10.1016/j.heliyon.2020.e04804

[65] H. Wang, E. H. Hovy, and M. Dredze, "The hurricane sandy twitter corpus," in *AAAI Workshop: WWW and Public Health Intelligence*, 2015. [Online]. Available: https://www.semanticscholar.org/paper/The-Hurricane-Sandy-Twitter-Corpus-Wang-Hovy/bf7f220a2908268a7a1c432a2c2efd81afed6808

[66] L. Zou, N. S. N. Lam, S. Shams, H. Cai, M. A. Meyer, S. Yang, K. Lee, S.-J. Park, and M. A. Reams, "Social and geographical disparities in twitter use during hurricane harvey," *International Journal of Digital Earth*, vol. 12, no. 11, pp. 1300–1318, Dec. 2018. [Online]. Available: https://doi.org/10.1080/17538947.2018.1545878

[67] J. B. Houston, J. Hawthorne, M. F. Perreault, E. H. Park, M. G. Hode, M. R. Halliwell, S. E. T. McGowen, R. Davis, S. Vaid, J. A. McElderry, and S. A. Griffith, "Social media and disasters: a functional framework for social media use in disaster planning, response, and research," *Disasters*, vol. 39, no. 1, pp. 1–22, Sep. 2014. [Online]. Available: https://doi.org/10.1111/disa.12092

[68] S. Yum, "Mining twitter data to understand the human sentiment on hurricane florence," *Journal of Disaster and Emergency Research*, Aug. 2020. [Online]. Available: https://doi.org/10.18502/jder.4069

[69] *Facing Hazards and Disasters*. National Academies Press, Aug. 2006. [Online]. Available: https://doi.org/10.17226/11671

[70] D. S. Thomas, K. Ertuˆgay, and S. Kemec¸, "The role of geographic information systems/remote sensing in disaster management," in *Handbook of Disaster Research*. Springer New York, 2007, pp. 83–96. [Online]. Available: https://doi.org/10.1007/978-0-387-32353-4_5

[71] SAMHSA Disaster Technical Assistance Center Supplemental Research Bulletin. (August 2016) "stronger together: An in-depth look at selected community-level approaches to disaster behavioral health". [Accessed November 7, 2021]. [Online]. Available: \url{https://archive.org/download/www.samhsa.gov/www.samhsa.gov/sites/default/files/programs_campaigns/dtac/srb-community-approaches.pdf}

[72] A. A. Khaleq and I. Ra, "Twitter analytics for disaster relevance and disaster phase discovery," in *Proceedings of the Future Technologies Conference (FTC) 2018*. Springer International Publishing, Oct. 2018, pp. 401–417. [Online]. Available: https://doi.org/10.1007/978-3-030-02686-8_31

[73] Y. Martín, S. L. Cutter, Z. Li, C. T. Emrich, and J. T. Mitchell, "Using geotagged tweets to track population movements to and from puerto rico after hurricane maria," *Population and Environment*, vol. 42, no. 1, pp. 4–27, Feb. 2020. [Online]. Available: https://doi.org/10.1007/s11111-020-00338-6

[74] K. Hunt, B. Wang, and J. Zhuang, "Misinformation debunking and cross-platform information sharing through twitter during hurricanes harvey and irma: a case study on shelters and ID checks," *Natural Hazards*, vol. 103, no. 1, pp. 861–883, May 2020. [Online]. Available: https://doi.org/10.1007/s11069-020-04016-6

[75] S.-M. Cheong and M. Babcock, "Attention to misleading and contentious tweets in the case of hurricane harvey," *Natural Hazards*, vol. 105, no. 3, pp. 2883–2906, Nov. 2020. [Online]. Available: https://doi.org/10.1007/s11069-020-04430-w

[76] F. Alam, F. Ofli, M. Imran, and M. Aupetit, "A twitter tale of three hurricanes: Harvey, irma, and maria," 2018. [Online]. Available: \url{https://arxiv.org/pdf/1805.05144.pdf}

[77] V. K. Neppalli, M. C. Medeiros, C. Caragea, D. Caragea, A. H. Tapia, and S. E. Halse, "Retweetability analysis and prediction during hurricane sandy," in *ISCRAM*, 2016. [Online]. Available: \url{https://www.semanticscholar.org/paper/Retweetability-Analysis-and-Prediction-during-Sandy-Neppalli-Medeiros/52591b302291d7cfb5f0d72a3c310b60724c8e7e}

[78] J. Bohannon, "Twitter can predict hurricane damage as well as emergency agencies," *Science*, Mar. 2016. [Online]. Available: https://doi.org/10.1126/science.aaf4182

[79] M. Bica. (August 6, 2019) "twitter in the eye of the storm: Assessing the public's risk perceptions". [Accessed December 5, 2021]. [Online]. Available: \url{https://items.ssrc.org/chancing-the-storm/twitter-in-the-eye-of-the-storm-assessing-the-publics-risk-perceptions/}

[80] A. Olteanu, C. Castillo, F. Diaz, and S. Vieweg, "Crisislex: A lexicon for collecting and filtering microblogged communications in crises," *Proceedings of the 8th International Conference on Weblogs and Social Media, ICWSM 2014*, pp. 376–385, 01 2014.

[81] D. M. Blei, A. Y. Ng, and M. I. Jordan, "Latent dirichlet allocation," *J. Mach. Learn. Res.*, vol. 3, pp. 993–1022, 2003. [Online]. Available: http://jmlr.org/papers/v3/blei03a.html

[82] G. Burel, H. Saif, and H. Alani, "Semantic wide and deep learning for detecting crisis-information categories on social media," in *Lecture Notes in Computer Science*. Springer International Publishing, 2017, pp. 138–155. [Online]. Available: https://doi.org/10.1007/978-3-319-68288-4_9