Title	Design and evaluation of a SNS-based sentiment analyzer to support effective disaster planning for foreigners
Sub Title	
Author	Detera, Bernadette Joy(Kōtake, Naohiko) 神武, 直彦
Publisher	慶應義塾大学大学院システムデザイン・マネジメント研究科
Publication year	2021
Jtitle	
JaLC DOI	
Abstract	
Notes	修士学位論文. 2021年度システムエンジニアリング学 第335号
Genre	Thesis or Dissertation
URL	https://koara.lib.keio.ac.jp/xoonips/modules/xoonips/detail.php?koara_id=KO40002001-00002021-0007

Master’s Dissertation 2021

Design and Evaluation of a SNS-based Sentiment Analyzer to Support Effective Disaster Planning for Foreigners

Bernadette Joy Detera
（Student ID Number：81934572）

Supervisor: Professor Naohiko Kohtake

September 2021

Graduate School of System Design and Management,
Keio University
Major in System Design and Management

ABSTRACT

Student Identification Number	81934572	Name	Detera, Bernadette Joy

Title

Design and Evaluation of a SNS-based Sentiment Analyzer to Support Effective Disaster Planning for Foreigners

Abstract

Having proper situational awareness during disaster situations is crucial in planning and mitigation. Knowing people's perception, needs and behavior during disasters is critical in developing the right management strategies. However, cities with multilingual and diverse international population may react differently to disasters and gaps still exist in understanding this issue. Microblogging using social media has become a prevalent tool during emergencies and disasters. In this paper, we present a system to analyzing the sentiment of both the local native population and foreigners during earthquake and typhoon. The goal of this research is to get insights on the sentiment of foreigners during disasters which can be helpful for disaster planners to develop more effective disaster management plans considering the social vulnerability of foreigners. Through the use of Social Networking Sites (SNS) such as Twitter, we retrieve individual tweets specifically on the onset of the disaster both in Japanese and English. Through this, a granular approach to data collection was realized with raw data collected only on the specific geographical area of interest and specific timeframe. After preprocessing of raw data, we develop supervised Machine Learning algorithms based on Support Vector Machine (SVM) and Extreme Gradient Boosting (XGBoost) to predict and classify tweets as positive, neutral, or negative in sentiment. The sentiment analysis models obtain high accuracy and could be used for identification and classification of sentiment during disaster scenarios. In addition, since our model is trained specifically on disaster tweets, it could yield a more accurate and contextual result when applied to future disasters. Furthermore, we performed Sentiment Analysis techniques through Word Cloud, keyword analysis and time series analysis to generate direct comparison of sentiment between Japanese and foreigners. We deduce that Japanese show a more positive sentiment than foreigners at times of disaster and foreigners experience disasters significantly more negatively. These results confirm that foreigners face more difficulty, stress, and anxiety than the native population.

Additionally, we observe that negative sentiment of both groups is higher in earthquakes. Through this study, a deeper understanding of foreigners' social vulnerability is achieved. Moreover, the proposed system generated insights specific to the type of natural disaster (i.e., typhoon and earthquake). Several recommendations were proposed to support disaster planning specifically on the timing of disaster communication as well as content and type of pre-disaster support and early warning message to be given to foreigners. Lastly, we validate that using this approach could provide insights to elicit requirements for disaster information or warning systems catered towards foreigners in the area which could be useful in improving disaster planning.

Key Words (5 words)
Disasters, Sentiment Analysis, foreigners, SNS, Twitter

ACKNOWLEDGEMENT

Before anyone else, I would like to express my sincerest gratitude to Professor Naohiko Kohtake, my primary supervisor, who has been very supportive in all aspects of my conducting this research, and whose guidance propelled my research forward in the right direction. I also like to express my thanks for his passion which has inspired me to join Keio SDM and pursue research on this area. I also thank him for giving me the opportunity to participate in international research projects such as the Belmont Forum which supported my data collection and analysis.

I would also like to express my earnest thanks to Assistant Professor Akira Kodaka, who provided invaluable support especially in data analysis and validation activities. Furthermore, I would like to extend my thanks to the members of Kohtake Laboratory for providing feedback to improve my research. In particular, my sincere thanks to Mr. Akihiko Nishino and Ms. Kaya Onda who have provided meaningful comments and advice on my research and conference paper submission. This research has been conducted as part of the Data-Driven Disaster Response Systems Dependent on Time of Day, Season and Location for Megacities (DUAL) project.

In addition, my sincere thanks to Associate Professor Makoto Ioki, my secondary supervisor, for providing constructive comments and helpful advice on my research and thesis. I am also indebted to all professors and faculty of Keio SDM who have all been a great source of learning.

I would also like to thank Mitsubishi Estate, especially Mr. Kotaro Sawabe and Mr. Toshitaka Nakajima, for their kind feedback and opportunity to interview them.

I also extend my thanks to Mr. Akihiro Sato and Mr. Yu Ohno for providing guidance on doing research and extending their knowledge.

Finally, I like to extend my heartfelt gratitude to my family and friends who have been very supportive and encouraging throughout my journey especially amidst these trying times of doing research during the COVID-19 pandemic.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

DEFINITION OF KEY TERMS

Disaster – a serious disruption of the functioning of the community or a society causing widespread human, material, economic or environmental losses that exceed the ability of the affected community to cope using its own resources.

Social Vulnerability – inability of people to withstand adverse impacts of external stressors such as disasters.

Sentiment Analysis – use of Natural Language Processing, text analysis and computational linguistics to identify, extract, quantify and study subjective information.

Machine Learning – the study of computer algorithms that improve automatically through experience and by the use of data.

I. INTRODUCTION

1. Background

1.1 Disaster Management

Natural disasters are ubiquitous and unpredictable. The most common regional-scale natural disasters in the Asia-Pacific are typhoons, earthquakes, volcanic eruptions, tsunamis, and flooding. Globally, natural disasters kill 60,000 people per year and are responsible for around 0.1% of deaths in the past decade [1]. Aside from casualties, millions of people are affected by natural disasters. Moreover, natural disasters heavily affect the economy and incur significant economic losses that last for decades. In the United States, Hurricane Katrina in 2005, one of the most devastating disaster to hit the US has incurred an estimated value of $160 Billion [2]. The 2011 Great East Japan Earthquake which was considered a triple disaster (earthquake, tsunami and nuclear power plant accident) has cost an estimate of $220 Billion [3] with around 20,000 deaths and 500,000 people displaced. Because of the sever impact of natural disasters in the economy and human life, governments and institutions have been conducting extensive efforts to reduce its impact and hazard exposure. To do this, disaster management cycle is widely used to understand the disasters in its lifecycle so that we can adopt better and become more resilient. Although there are different versions of the cycle, the figure below shows a typical disaster life cycle.

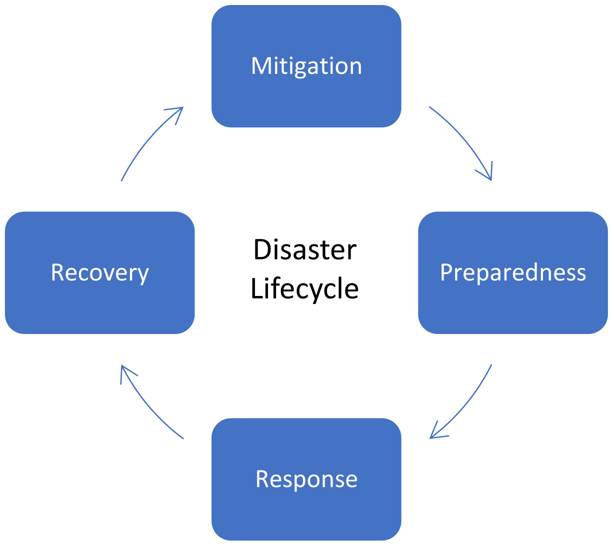


Figure 1 Disaster Life Cycle Stages

As shown in Figure 1, the disaster lifecycle includes four phases namely:

- Mitigation – disaster mitigation involves efforts to prevent future emergencies or to minimize the negative effects of disaster. This includes risk identification, hazard analysis, identifying areas prone to specific disasters, flood-proofing infrastructures, etc.
- Preparedness – disaster preparedness involves making plans or preparation activities in advance to help the society become prepared for upcoming natural disasters. Efforts for this phase include stock piling emergency supplies, food, and water, and preparing humanitarian teams or volunteers for post-disaster relief missions.
- Response – disaster response involves activities done during or immediately after a disaster, to save or prevent loss of life and prevent further damage to property infrastructure. Most efforts in disaster management are centered in this phase.
- Recovery – disaster recovery involves actions to help communities return to pre-disaster states, which include restoration of infrastructure, provision of trainings or more awareness campaigns to make the community less vulnerable to hazards in the future.

By understanding all phases of disasters, government and institutions can effectively prepare management plans and allocate resources to prevent and minimize its impact. The impact of natural disasters is more aggravated in cities because of their large population and prospering economy which incur many deaths, damage to infrastructure and interruption of business activities. For this reason, disaster management is not only conducted on a national scale but also on a regional or city level.

1.2 Social Vulnerability

Megacities, cities having populations of more than 10 million, are extremely vulnerable to this which is why most megacities in the world have been putting a lot of resources to better prepare themselves through thorough disaster management planning. However, not all gets affected by disaster when they occur. Some people experience more difficulty and are less capable of responding to these emergencies. These vulnerable populations are often the hardest hit in disasters. Particularly in megacities having dense populations, there are several vulnerable social groups – elderly, persons with disability, etc., - that gets affected. One vulnerable group are migrants – foreigners who are in the city. According to United Nations data, approximately 244 million people are living and working abroad [4]. Foreigners have become part of today's societies – labor migrants, international students, tourists, and expatriates. Having foreigners in the city pose a great challenge to deal with diversity when emergencies happen. Expanding megacity populations also come with an increase in population diversity as, primarily, more people immigrate for the industry. Some groups of people consequently become especially vulnerable because they (1) are different from the locals, (2) are limited in the knowledge of the area, (3) may have difficulties expressing their concerns or asking for information due to language, social, and cultural barriers. According to a publication of the United Nations Development Programme on social vulnerability indicators [5], population density, immigrants and percentage of urban population and vulnerable minorities all increase social vulnerability. Thus, megacities like Tokyo, Japan, and New York in the US face higher social vulnerability. Savage and Trogler [6] noted that recent disaster management studies still lack factors and approaches to disaster management.

Diverse population elicits different needs and capabilities which emergency actors need to take note. Cultural diversity has many aspects, and these affects how people perceive disaster risk, prepare for disaster, and react to them. Figure 2 shows the different elements of diversity.

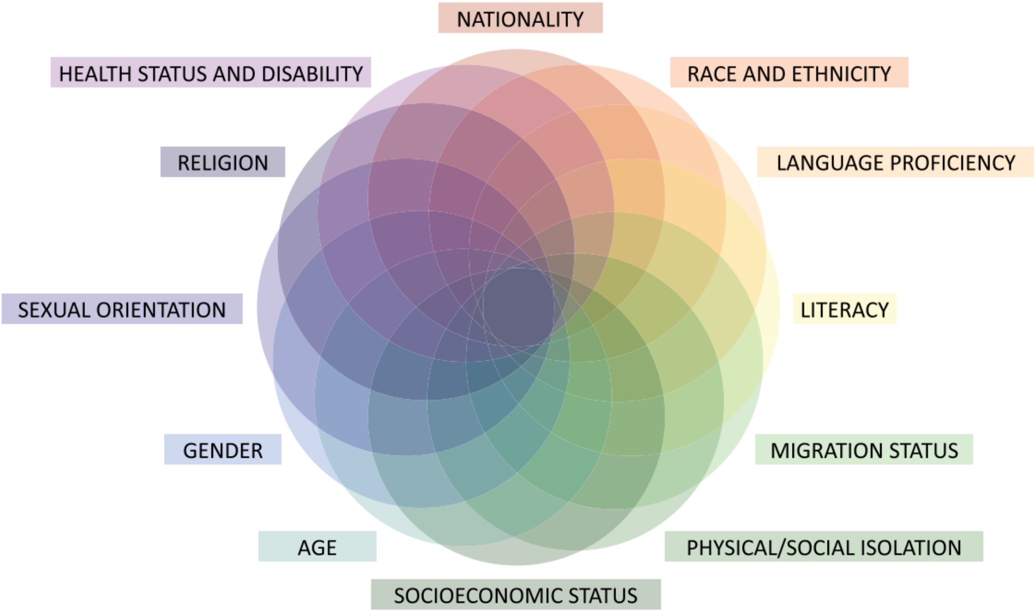
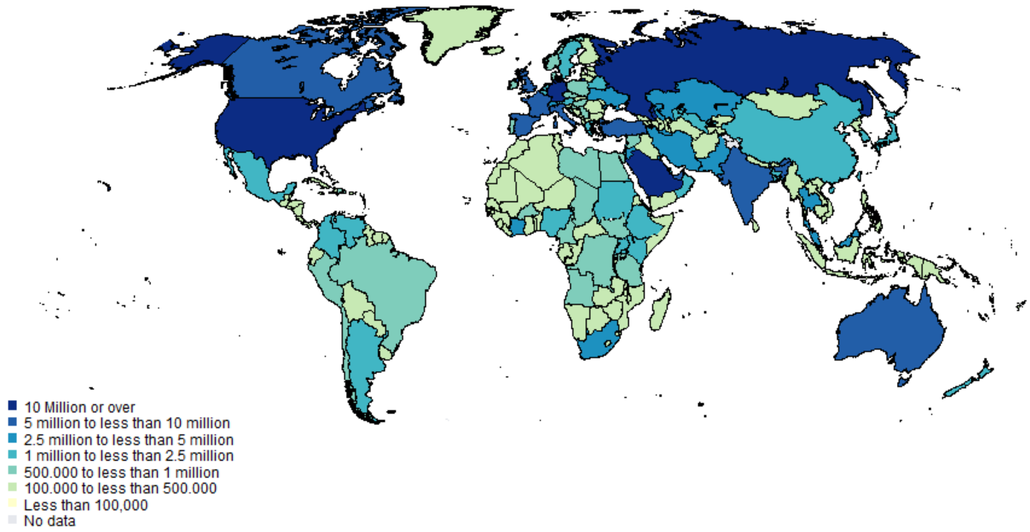


Figure 2 Elements of Cultural Diversity [7]

Emergency actors and disaster management planners need to understand these elements and adopt existing practices to consider these specificities across all phases of the disaster lifecycle. There are several examples of disaster when foreigners were heavily affected. These include the 2004 Indian Ocean tsunami, the 2011 floods in Thailand, the Japanese earthquake, tsunami and nuclear meltdown in 2011, and Hurricane Sandy in 2012 [4].

Looking at the global statistics, the rate of migration has been increasing up until the COVID-19 pandemic. The number of migrants all over the world can be seen in the map below:



Source: UN Department of Economic and Social Affairs

Figure 3 Number of international migrants in 2019 [8]

As shown in Figure 3, the United States, Russia, and Australia has the greatest number of foreign residents [8]. According to population statistics in Japan[9], there are over 2.6 million foreign nationals in the country in 2017. This is about 2% of the population at that time.

To address these issues, certain legal and policy frameworks are already in place. Firstly, it is stipulated in the 1963 Vienna Convention on Consular Relations [10] that the host country is responsible for ensuring safety and welfare of all individuals during crisis or disasters, irrelative of their status (tourists, migrant workers, international students, etc.). More recent policies such as the 2005-2015 Hyogo Framework Agreement [11] and Sendai Framework for Disaster Risk Reduction (SFDRR) [12] call for action to take into account vulnerable groups when planning for disasters, and the inclusion of migrants in implementing disaster risk reduction strategies.
Three provisions [13] from the SFDRR explicitly articulate this notion:

- Paragraph 7: Governments should engage with relevant stakeholders, including [...] migrants [...] in the design and implementation of policies, plans and standards.
- Paragraph 27(h): Empower local authorities, as appropriate, through regulatory and financial means to work and coordinate with [...] migrants in disaster risk management at local level.

- Paragraph 36(a)(vi): Migrants contribute to the resilience of communities and societies and their knowledge, skills and capacities can be useful in the design and implementation of disaster risk reduction.

2. Problem Analysis

To understand the problem of social vulnerability of foreigners more deeply, it is important to understand the context of the problem. Impact of natural disasters is heavily determined by key characteristics of cultural diversity as outlined in Figure 2. These include class, ethnicity, gender, and age, among others [14]. Foreigners are specifically vulnerable and face the impact of disasters differently. Figure 4 shows the vulnerability cycle experienced by foreigners. Foreigners are considered inherently vulnerable to the impact of disaster. Lack of access to timely, accurate and easily understandable information engenders low disaster awareness which makes them insufficiently prepared for disasters. According to initial interviews conducted with foreigners who have been to and are currently visiting Japan, all of them do not have any preparation to disasters before or during their stay in Japan. Being unprepared for disasters could cause more anxiety. According to the interviewees, they would not know what to do in case of disasters in

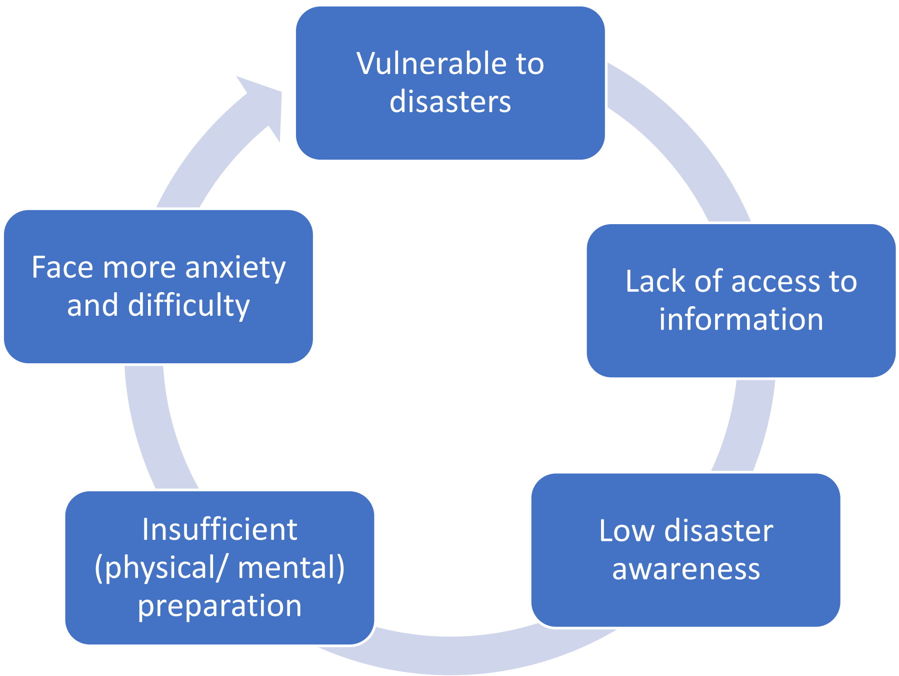


Figure 4 Vulnerability Cycle of Foreigners on Disaster Anxiety

Tokyo but are confident there would be information available in major Tokyo train stations and so on.

One important factor for vulnerability is the language barrier when accessing disaster-related information which in turn makes it more difficult to cope with disasters. In a study by Kawasaki et al [15] on the information needs of foreigners during the 2011 Great East Japan Earthquake [3], it was concluded that the lack of proficiency in the Japanese language should be considered in disaster prevention planning.

Several studies [16][17] have already noted the vulnerability of foreigners during disasters and even the Japanese government has noted the struggle to provide adequate support for foreigners [18] and that the existing disaster plans in place are still focused only on the local native population. Thus, we must ask: **How might we understand the different needs of foreigners to improve current one-size-fits-all disaster management planning?**

To further understand the issue of understanding the needs of foreigners, it is important to look at the factors affecting disaster planners to develop appropriate strategies accounting for a culturally diverse population. Analysis of the problems faced by disaster management institutions is shown in Figure 5. Disaster planners require quality data as inputs for disaster management plans. By having better quality data, a more accurate risk assessment can be conducted which provides a more suitable and appropriate disaster plans. However, data quality is affected by many factors. Firstly, disasters do not happen everyday nor on a regular schedule. Having only a few disaster events in a year leads to poor quality data. Moreover, severe disasters happen on a rarer occasion. Because of this, it is also difficult to gather enough data on each type of natural disaster that could possibly occur. Amount of data coming from vulnerable groups such as foreigners is also an important factor. Because foreigners only stay temporarily in the host country (for example, as a tourist or international student), the amount of data generated by them is scarce and difficult to access. Another important factor is the accessibility of data. In disaster research, access of personal data can be very limited because of data privacy. For example, camera footages during evacuation in public places cannot be easily obtained because of data privacy. Having easier access to such data improves the quality of data collected for disaster planning. Access to data from the people or victims of disasters typically utilize surveys or in-person interviews which can be costly and difficult to obtain in most cases. Finally, having a more diverse population pose greater challenge in developing appropriate disaster management plans because there would be more aspects of the cultural diversity shown in Figure 2 that should be considered.

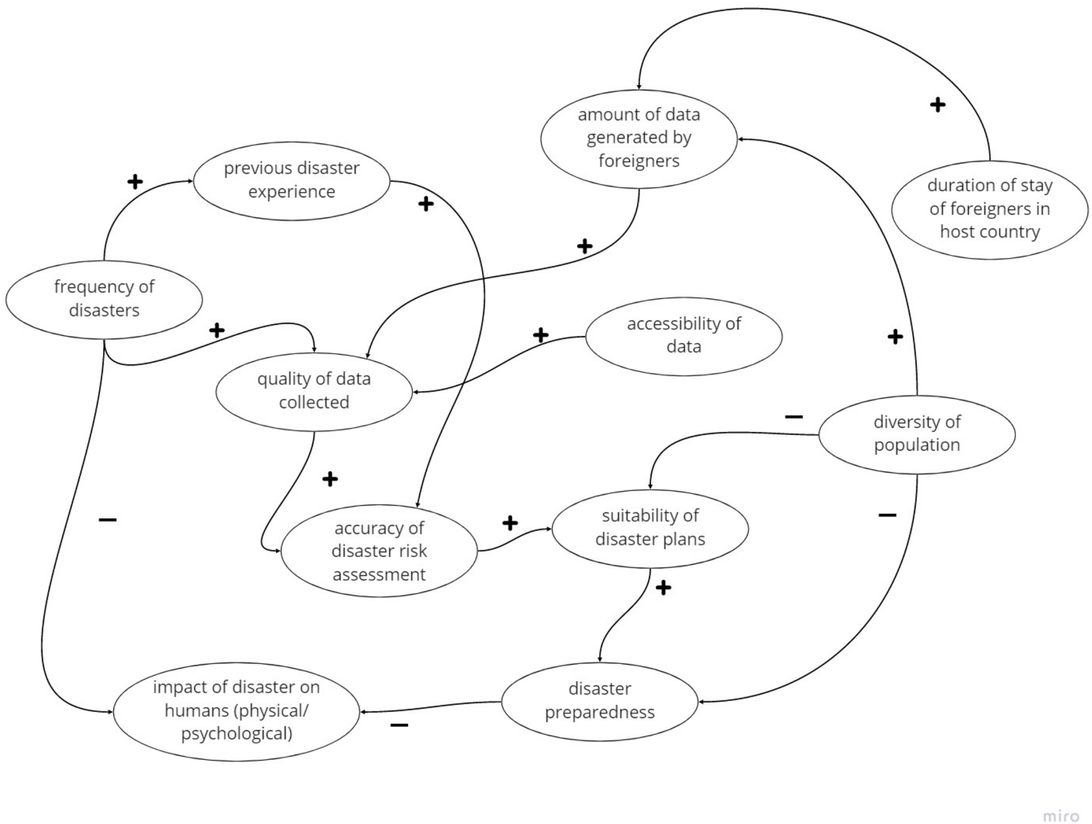


Figure 5 Causal Loop Analysis for Disaster Planning

To address the question on understanding the needs of foreigners to improve existing one-size-fits-all approach to disaster planning, a possible solution is to have a new system for data collection on how foreigners reacted to previous disasters by comparing them with the local native population. To achieve this, the following problems must be solved by the system:

Table 1 Summary of Problems

Problems
Data collection is typically limited and difficult to get leading to poor quality data.
Clear differences between how foreigners and the native population react during disasters are not known.
Difference in disaster behavior and anxiety specific depending on the type of natural disaster is still unknown.

3. Research Goal

The goal of this research is to design a Disaster Sentiment Analyzer System which can be used by institutions and disaster planners to gather useful insights on requirements of foreigners and design more effective disaster management plans.

4. Research Scope

The scope of this research is limited to the social vulnerability of foreigners in Tokyo, Japan. There are approximately 3 million foreign residents registered in Japan in 2019, which comprise 2.3% of the population with Tokyo having the greatest number of international residents and tourists alike. The map of Tokyo is shown in the figure below. Moreover, this research is only limited to two types of natural disaster: typhoons and earthquake which are the most frequent type of disasters faced in Japan.

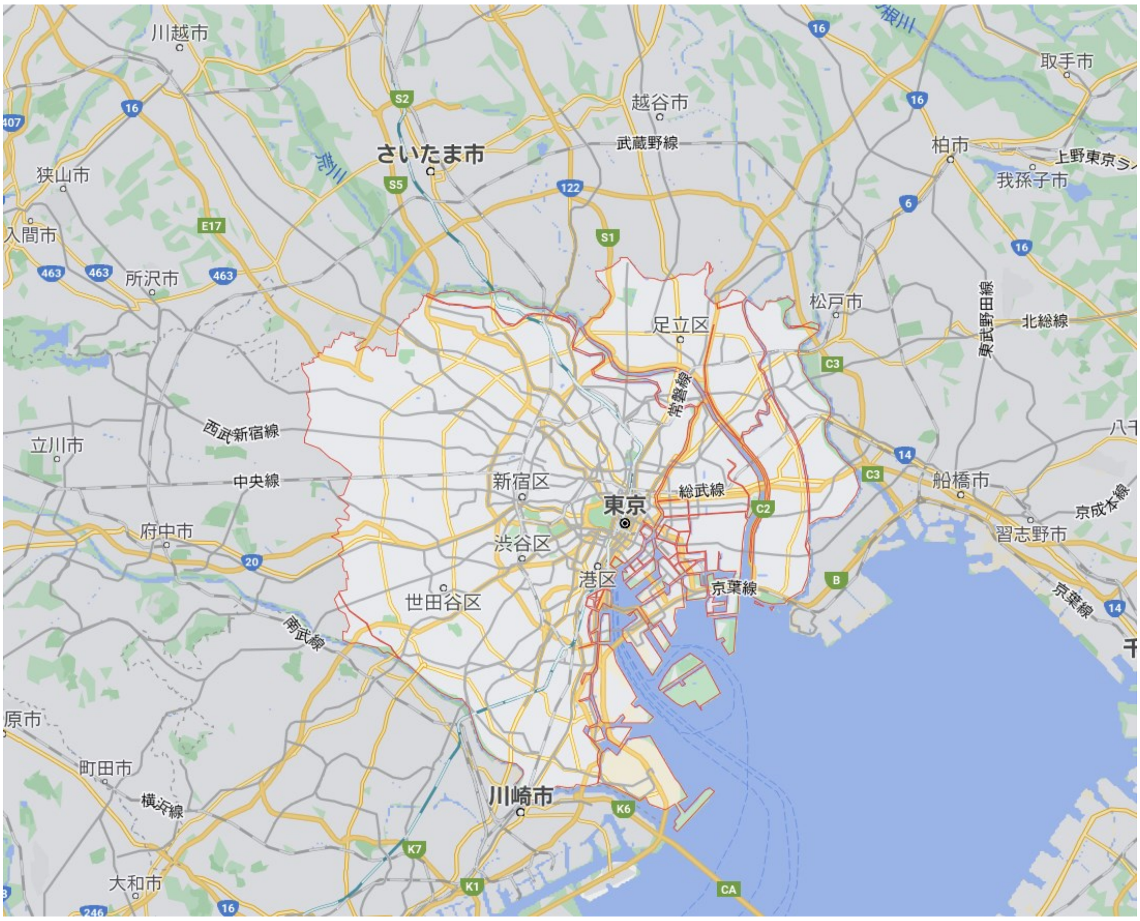


Figure 6 Map of Tokyo, Japan

Figure 6 shows a map of Tokyo outlining its boundary colored in red. The area within this geographical consists of Tokyo's 23 wards, which are considered to be the country's capital where major business districts thrive. This is chosen as the area of study because of the megacity nature of Tokyo, having dense population with a relatively high percentage of international population, as well as receiving a huge volume of international tourists and business travelers every year.

5. Research Objectives

As outlined in Table 1, key problems to be solved by this research involves addressing limited data collection method and having insights of disaster perception and behavior specific to foreigners and type of natural disaster. To address each problem, objectives of the system are outlined in Table 2.

Table 2 Summary of Research Objectives

Problems	Objectives
Data collection is typically limited and difficult to get leading to poor quality data.	Identify another source of data coming from foreigners and/or disaster victims
Clear **differences between how foreigners and the native population** react during disasters are not known.	Understand and assess vulnerability of foreigners in terms of anxiety or distress during disasters. Compare disaster behavior of foreigners against locals
Difference in disaster behavior and anxiety specific depending on the type of natural disaster is still unknown.	Assess vulnerability during at least two types of natural disaster

First, another data source should be identified by the system aside from the existing interview and survey methodologies which are typically conducted after the disaster. Second, the system aims to understand how the foreigners are feeling and reacting during previous disasters. Subsequently, the objective is to clearly delineate the sentiment of the foreigners by comparing it to that of the local native population. Finally, the system aims to conduct these assessments for specific type of natural disasters instead of generalizing natural disasters into one category. For this system, typhoons and earthquakes are chosen as the disasters of interest.

6. Originality

To achieve the objectives of this research, a potential solution is proposed involving the use of Twitter, a social networking site, together with big data analytics to generate insights on the requirements of foreigners during disasters. To highlight the originality of this research, the approach towards achieving the objectives is outline in Table 3.

Table 3 Research Approaches to Achieve Objectives

Problems	Objectives	Approach
Data collection is typically limited and difficult to get leading to poor quality data.	Identify another source of data coming from foreigners and/or disaster victims	Use Social Networking Site (SNS) as the data source
Clear **differences between how foreigners and the native population** react during disasters are not known.	Assess vulnerability of foreigners in terms of anxiety or distress during disasters	Develop Machine Learning algorithms to automatically classify people's sentiment. Use sentiment intensity of both English and Japanese to directly compare results
Difference in disaster behavior and anxiety specific depending on the type of natural disaster is still unknown.	Assess vulnerability during at least two types of natural disaster	Gather historical data on previous disasters specifically: severe typhoon and earthquake

As shown in Table 3, SNS data will be used instead of traditional means of getting information through surveys and in-person interviews. Specifically, Twitter which is a popular microblogging site with over 199 million daily active users [19], is utilized in this research. Machine Learning algorithm will be used to automatically classify what kind of sentiment do people experience. While most studies have only focused on single language, this research contributes by developing algorithms for both English and Japanese. This bilingual nature is one of the main contributions of this research. Moreover, the direct comparison between foreigners and native population in terms of disaster sentiment has not been studied yet in the past. A previous study [20] compared the behavior of Japanese and foreigners in terms of their information gathering behavior but not on disaster sentiment. With the proposed approach, an analysis of disaster sentiment specific to certain type of disaster is possible in this system. Specifically, this research provides analysis on typhoon and earthquake. Finally, the most important contribution of this research is the use case of utilizing Twitter for sentiment analysis. Some studies have previously utilized Twitter to aid disaster management [21] [22] [23] specifically on emergency response. The main contribution of this research is the utilization of this approach for disaster planning phase, instead of the typical use case on disaster response.

7. Structure of Thesis

The thesis is organized into the following chapters:

Chapter 2 presents the review of related literature on social vulnerability of foreigners and existing scholarship on the use of Twitter for sentiment analysis. The author briefly discusses studies on social vulnerability of foreigners, microblogging during disasters, and sentiment analysis methodologies, and explains what have not yen been achieved in each area thus far.

Chapter 3 presents the proposed design and methodology of the system. The author presents the system requirements derived from stakeholder needs and subsequently present the end-to-end methodology as well as sub-systems using diagrams.

Chapter 4 presents the evaluation of the system through verification and validation activities. This chapter utilizes two case studies on typhoon and earthquake and presents results of evaluation upon using the system. The author presents results of prototyping, simulation and interviews conducted.

Chapter 5 presents a thorough discussion on the results of evaluation. Specifically, this chapter analyzes the results of sentiment statistics and time series analyses for both Japanese and English tweets, interpreting its implications to reflect actual behavior of native and foreign populations in Tokyo. The author also discusses the limitations of the study.

Chapter 6 presents the conclusion of the thesis as well as future research directions for this kind of study.

II. RELATED WORK

1. Social Vulnerability and Disaster Behavior of Foreigners

Studies highlighting disaster behavior of foreigners such as tourists or international residents have made contributions to deepen the understanding of the social vulnerability of foreigners. Foreigners and migrants in a country are more vulnerable to disasters. In the past, studies have shown insights on the preparedness and post-disaster behavior of foreigners in Japan during the Great East Japan Tohoku earthquake in 2011 [15] [24] [25]. Kawasaki et al. [15] highlight the language barrier as one of the main factors contributing to the social vulnerability of foreigners in the eastern region of Japan during the 11 March 2011 magnitude 9.0 earthquake which hit the region. In addition to language, physical factors such as power outage, mobile congestion, etc. also were also concerns shown in the results.

Another study by Kawasaki et al. [20] compared the disaster information gathering behavior of foreigners with Japanese people through a survey conducted after the 2011 Great East Japan Earthquake. The comparison on their information gathering behavior is shown below in Figure 7.

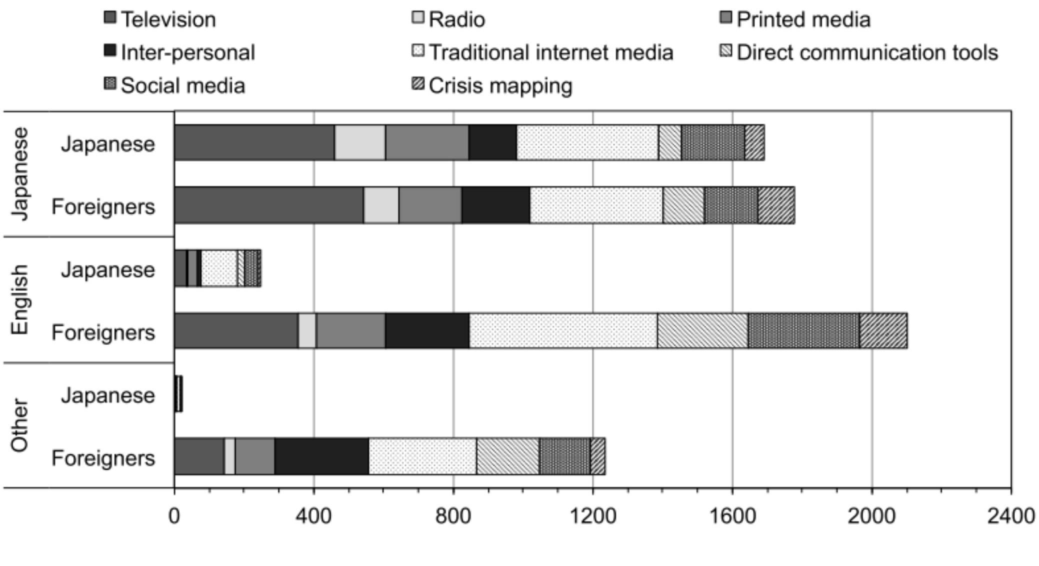


Figure 7 Media and language in information acquisition of Japanese and Foreigners [20]

Understandably, Japanese people overwhelmingly used Japanese as their main source language of choice in traditional media (television and radio). Foreigners turned to internet-based forms of media when using English. It is well noting that other languages, perhaps mother languages of foreigners), were also utilized especially for interpersonal communication like face-to-face conversations. In terms of relative importance of information over time, the result is shown in the figure below:

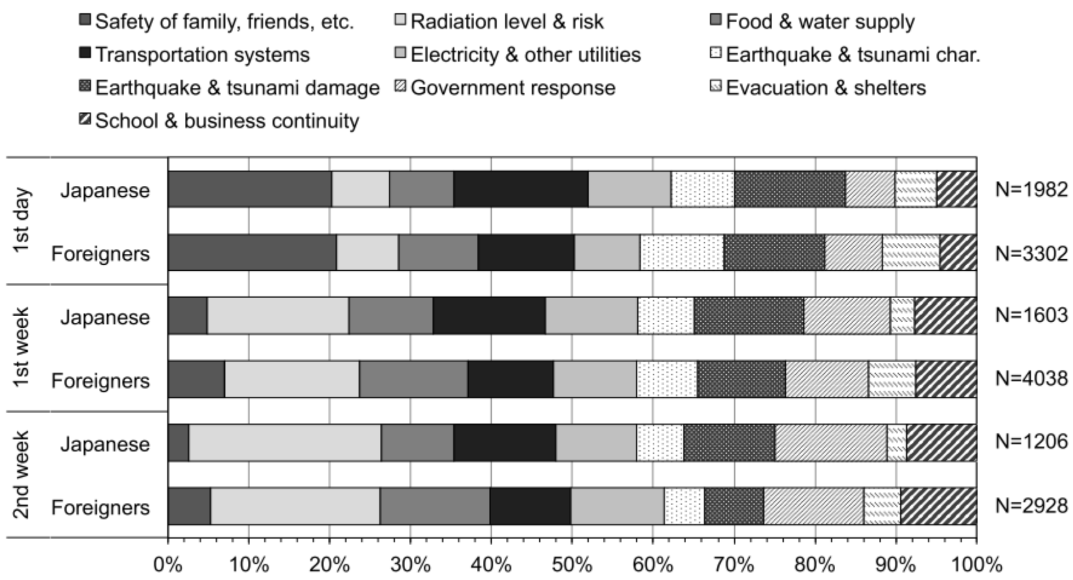


Figure 8 Importance of information over time during disasters [20]

For both Japanese and foreigners, safety of family and friends is the most important information on the 1st day or onset of disaster. Information on the damage from earthquake, tsunami and radiation was also important for both groups. However, it is worth noting that Japanese worry more about the transportation system than foreigners. In addition, foreigners worry more about the food and water supply.

Lastly, through Kawasaki et al.'s study, it was also confirmed that a major reason why information was unavailable to or understood by foreigners was that they could not understand information due to lack of language comprehension. On the other hand, both Japanese and foreigners experienced confusing or differing information from media sources. These results are shown in the figure below:

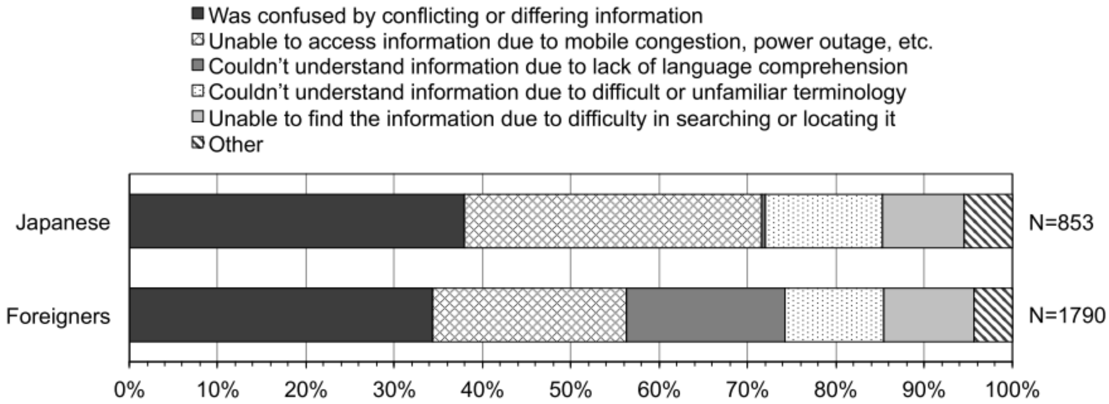


Figure 9 Reasons why disaster information are unavailable or unclear[20]

Apart from the aspect of disaster information gathering behavior, disaster preparedness of foreigners as compared to the Japanese population has been revealed by a study by Dasgupta et al [26]. In this study, the authors used six verifiers to understand foreigner residents' disaster preparedness through self-evaluation. Results are shown in Table 4 below.

Table 4 Disaster preparedness of Foreigners compared to Japanese population [26]

Statements (when compared to the native Japanese population)	% Strongly Disagree	Disagree	Neutral	Agree	Strongly agree	Mean	Standard Deviation
P1. I have the same level of disaster information and knowledge regarding critical disaster facilities	29.32	36.84	16.54	15.04	2.26	2.24	1.10
P2. I work and participate closely in community organizations for disaster preparedness	42.86	42.86	9.77	3.76	0.75	1.77	0.83
P3. I can seek help from my neighbors/neighborhood association in the time of need	25.56	29.32	28.57	12.03	4.51	2.41	1.13
P4. I can follow the local news, FM stations etc.	29.32	22.56	24.06	19.55	4.51	2.47	1.23
P5. I understand the instruction and technical words of the local community representative or the city counselor	33.08	27.07	20.30	16.54	3.01	2.29	1.18
P6. I am comfortable moving and staying in an evacuation center	11.28	36.84	29.32	18.80	3.76	2.67	1.03

As shown in the table, all indicators and statements show strong to moderate disagreement with the strongest on "I work and participate closely in community organizations for disaster preparedness" followed by "I have the same level of disaster information and knowledge regarding critical disaster facilities". This shows that respondents did not have comparable level of preparedness with that of Japanese people. This result revealed that foreigners believed they were significantly lacking in terms of disaster preparedness as compared to the native Japanese population.

2. Microblogging in Disasters

Twitter has evolved from just being a social media platform to a source of news information despite being unchecked in credibility [27]. During disasters, twitter has already been used as a communication platform, either directed or opportunistic [28], for broadcasting information abundantly first-hand (i.e., those who experience the disasters themselves [29], by professional journalists on social media to disseminate information), etc. It has also been used for disaster preparedness.

This realized functionality of twitter owes it to the speed of information relay (i.e., retweeted) between users even way back in 2009 [30]. Twitter, and other social media sources, has been proven to be vital sources of information during disasters [31] [32] and a potential tool for risk reduction [33]. Recently, it is also being utilized to analyze the impact of COVID-19 pandemic [34].

Disaster management consists of four phases: prevention and mitigation, preparedness, response, and recovery. Khosla et al [35] proposed the use of social media to aid response and recovery by categorizing Twitter data into several post-disaster items (water, electricity, etc.) which are useful for emergency responders. By defining certain keywords for each category as shown in Table 5 below, tweets with the defined keywords were extracted and presented in Figure 10.

Table 5 Category List and Associated Keywords in Disaster Tweet Extraction [33]

Category	Keywords
Water	Water, drink, drinking, thirsty, thirst, dehydration
Food	Food, starve, hungry, milk, bread, formula, eat, foodstuff
Shelter	Shelter, house, living place, sleep, rest, accommodation
Medical emergency	Medicine, clinic, hospital, medicine, doctor, nurse, syrup, first aid, tonic.
Electricity	Electricity, power, electricity, light, fan, energy, current, charge

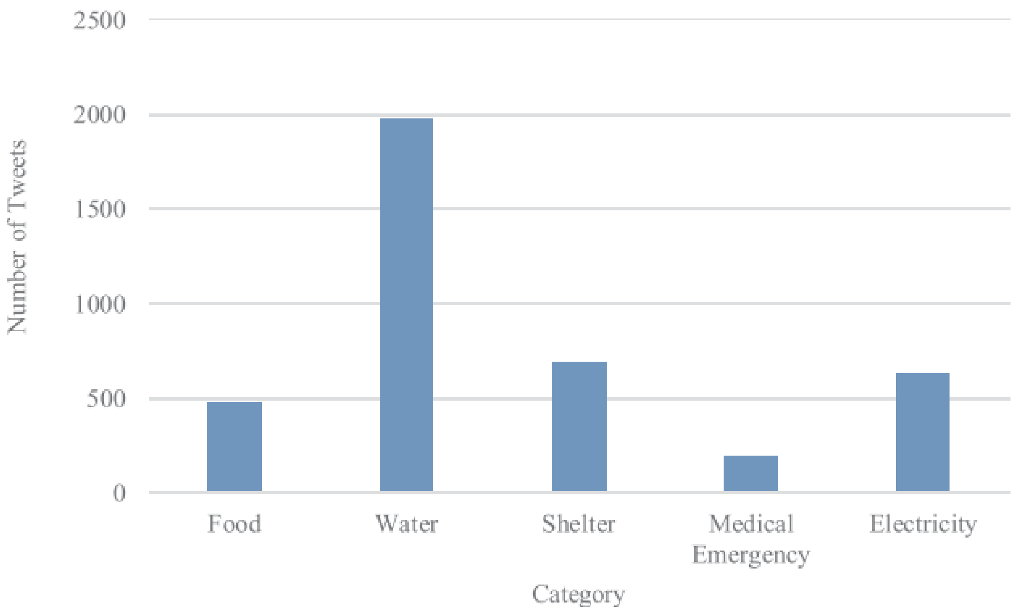


Figure 10 Number of disaster-related tweets in various categories about resources [33]

By utilizing Twitter data and counting the number of tweets relating to a certain category of relief resource, emergency responders could better prioritize and allocate resources. Another study utilized sentiment analysis to categorize tweets into those related to people's needs and availability of resources after disaster [35] [23]. By extracting tweets from the case of a strong earthquake in Nepal, the authors used a dataset of samples of tweets which refers to needed resources of people at that time of disaster as well as the availability of those resources; a simple method to match them is proposed by Khosla et al. Samples are shown in Table 6:

Table 6 Excerpts of needs and availability tweets in Nepal earthquake

Need-tweet (excerpts)	**Availability-tweet (excerpts)**
Mobile phones are not working, no electricity, no **water** in #Thamel, #Nepalquake	Please contact for drinking free service **water** specially for Earthquake Victim. Sanjay Limbu [mobile num]
Over 1400 killed. Many Trapped. **Medical Supplies** Requested.	20,000 RSS personnel with **medical supplies** and other help the first to reach earthquake damaged zones in #Nepal
Nepal earthquake: thousands in need of **shelter** in country little able to cope [url]	can anyone we know pick the 2000 second hand **tents** from Sunauli and distribute it to the people in need in Nepal?

As shown in Figure 11, a study has already endeavored to utilize Twitter to extract relevant tweets pertaining to two categories: needs- which informs about needs or required resources after the earthquake and availability- tweets that provide information on the availability of resources in the affected area.

Moreover, Twitter is also used to help emergency responders during Hurricane Sandy [36]. Although these studies reveal the usefulness of social media in the response and recovery phases, research on its application to disaster planning and preparedness is still understudied.

In terms of studies pertaining to the case of Japan, a study by Doan et al [37] analyzed the frequency and volume of tweets during the 2011 Great East Japan Earthquake which was one of the strongest disaster to hit Japan on March 11, 2011. Tweets were extracted for three categories of event: first is the earthquake and tsunami event, second is the radiation event which followed shortly after the earthquake, and third is categorized as anxiety event. To filter tweets that fit these three categories, certain terms were used as shown in Table 7:

Table 7 Keywords for earthquake and tsunami, radiation, and anxiety events during the 2011 Great East Japan Earthquake

English terms	Japanese terms
Earthquake and Tsunami event	
earthquake, quake, quaking, post-quake,shake, shaking,shock, aftershock,temblor, tremor, movement, sway, landslide seismic, seismography, seismometer, seismology, tsunami, wave	大地震，大震災 震災，地震，余震，揺れ 震度，震源，マグニチュード 津波
Radiation event	
radiation, nuclear, reactor, radioactivity, radioactive, iodine, TEPCO, meltdown, explosion, power plant, micro sievert	放射，放射線，放射能，放射性物質，原発，マイクロシーベルト，ヨウ素，イソジン，ヨウ化カリウム，炉心溶融，メルトダウン
Anxiety event	
die, scary, scared, incredible, worrying, worried, anxious, annoying	死亡，死ぬ，やられてる，やばい，やばかった，ヤバい，やばっ，やべ，怖い，怖かった，怖っ，すごい，すげえ，すげー，すっげー，びびる，びびった，混乱，微妙，避難，助けて，わかり辛い，連絡とれない，大変，心配，恐れ，船酔いしそう

After categorizing tweets intro earthquake and tsunami, radiation, and anxiety, the number of tweets in each category for English and Japanese tweets was compared. For the anxiety event, the comparison is shown in the figure below.

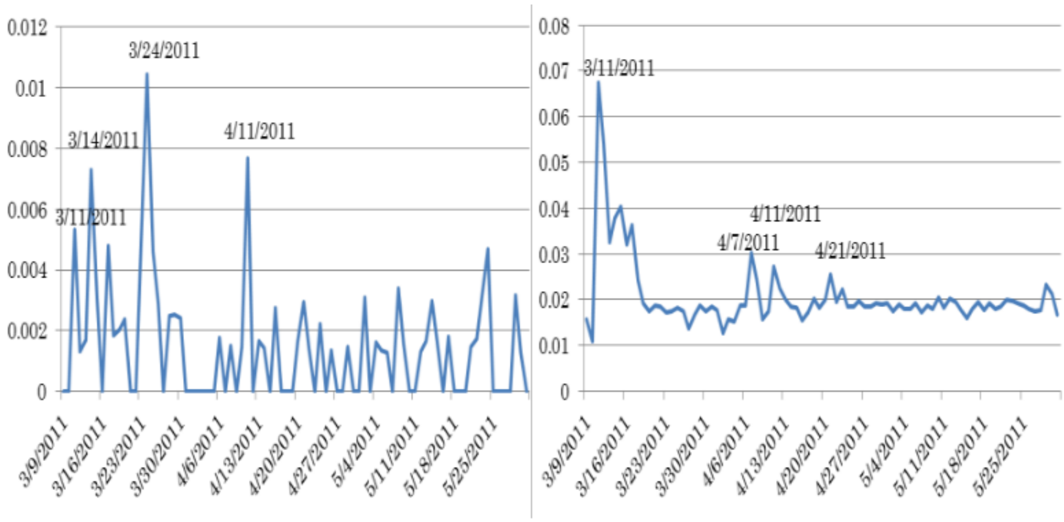


Figure 11 Frequencies of the anxiety event by dates during the 2011 Great East Japan Earthquake

As shown in Figure 12, the left chart shows that of Japanese and the right shows that of English tweets. The percentage of anxiety tweets in the total tweet dataset is calculated and shown in the figure. As shown in the left, the mean percentage is only around 0.2% for Japanese tweets but for English, the percentage of anxiety tweets is around 2% with a peak at 7% on the day of the earthquake itself. This suggested that foreigners experienced significantly more anxiety than Japanese.

Although it revealed that Japanese may have been calmer during disasters, the study only utilized empirical research using the volume of tweets. A more detailed analysis of this comparison is needed.

3. Sentiment Analysis Methods

Sentiment analysis has been widely used as a tool to analyze data on the internet and social media particularly to monitor a company's reputation or assess political campaigns [38]. In the context of disaster, sentiment analysis of Twitter data has become a popular tool in analyzing people's behavior during disasters.

Support Vector Machine (SVM) has become a popular algorithm in machine learning. It is widely used in sentiment analysis studies, and among many other applications (e.g., evaluating the sentiments of online reviews with the most impact [39], assist in selecting better tourist destination sites by evaluating travel blogs [40], facial recognition [41]). Other studies considered using SVM in hybrid with other machine learning algorithms, e.g., Random Forests [42], Convolutional and Artificial Neural Networks [43] in the hopes of extending the performance of the total model. An ample number of studies have also applied the algorithm in languages like Bangla [43], However, fewer studies exist which apply this same algorithm to two languages for a single goal (i.e., sentiment analysis).

Another promising machine learning method is gradient tree boosting [44] which has shown state-of-the-art results in many applications. In particular, the Extreme Gradient Boosting (XGBoost) algorithm [45] showed great potential in recent years. Studies show great potential for these algorithms. However, fewer studies exist which apply this same algorithm to two languages for a single goal (i.e., sentiment analysis). Most sentiment analysis studies using Twitter data only utilize one language (i.e., English). There is still a need to conduct methods to analyze the sentiment of groups of people utilizing multiple languages.

4. Twitter for Disaster Management

As evident in section 2.2, Twitter has been widely used and accepted as an invaluable source of information in disasters. In a comprehensive literature review of the use of social media data in situational awareness during disasters by Vongkusolit and Huang [46], it was found that Twitter was utilized for 62% of the studies reviewed. On the type of natural disaster, most studies have been conducted on the cases of hurricanes, earthquakes, and flooding. The result is shown in the figure below:

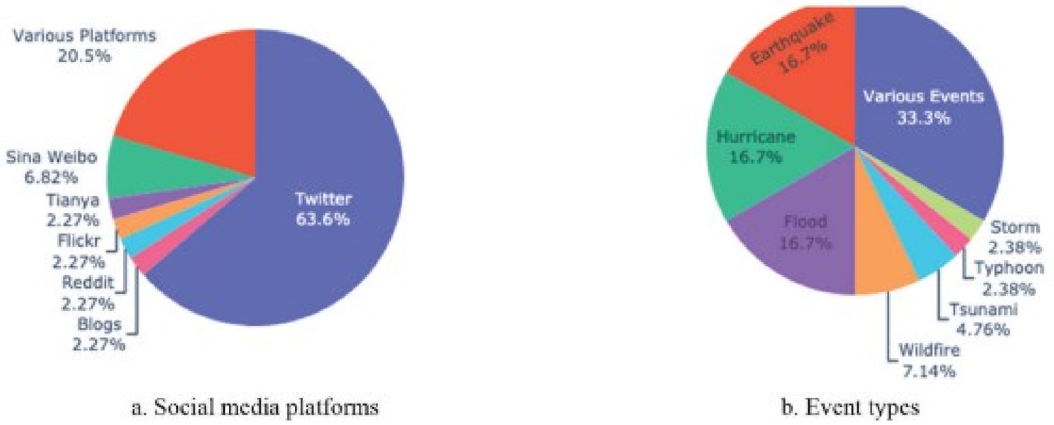


Figure 12 Percentages of different social media platforms and event type for content classification [47]

Studies which conducted content classification, that is, classifying social media messages by their content and keywords, use Twitter and other social networking sites (SNS) such as Sina Weibo and Flickr. Most commonly, floods and hurricanes are being studied especially in the United States [47] where those are the most common natural disasters occurring.

In terms of applications and usages of Twitter in disasters, a study by Seddighi et al [48] provides a summary on how Twitter utilization could save lives in times of disaster. The several functions of Twitter based on systematic review of existing literature is shown in the figure below:

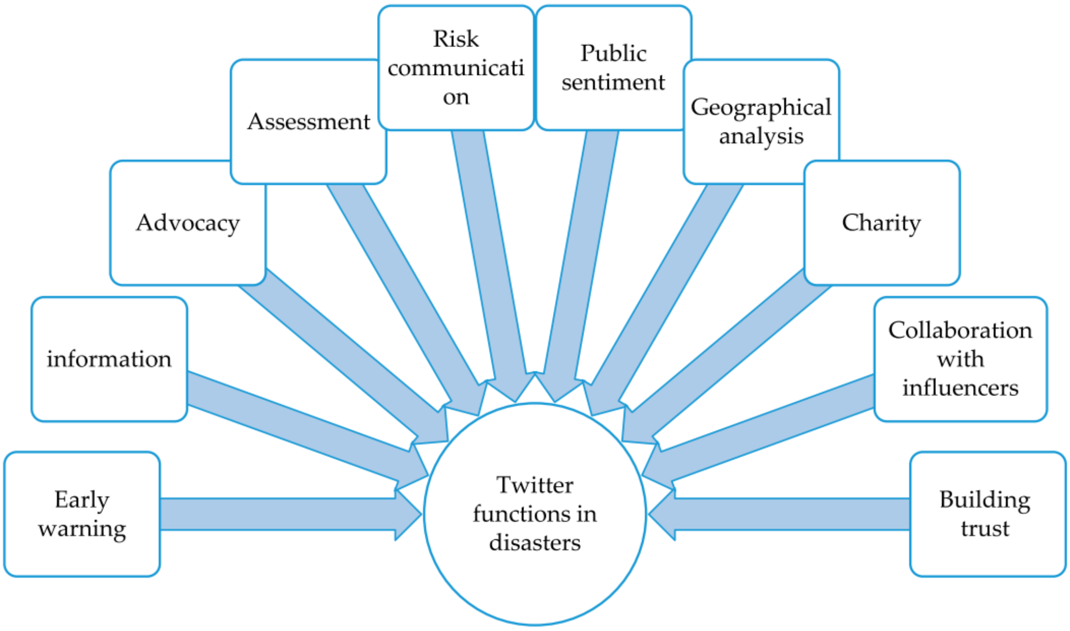


Figure 13 Functions of Twitter in disasters

These functions shown in Figure 13 is the most up to date thus far and shows how Twitter has significantly gained wide acceptability and applications besides the traditional use cases such as medium for early warning and information dissemination but also on assessing public sentiment and building trust with others. Based on the findings of Seddighi et al[48], a framework on the use of Twitter for disaster management is proposed and shown in Table 8.

Table 8 A framework for using Twitter during disasters

Stakeholders	**Type of Hazards**	**Twitter Functions**	**Tweets Characteristics**	**Disaster Risk Management**
Governments, Emergency organizations, Celebrities, People, News agencies, Donors, Affected people	Disasters triggered by natural and technological hazards, pandemics and complex disasters	Early warning, disseminating information, advocacy, assessment, risk communication, public sentiment, geographical analysis, charity, collaboration with influencers and building trust	Transparency, on-time messages, using different local languages, Using different media (text, video, photo)	Using Twitter during different phases including mitigation, preparedness, response and recovery

III. PROPOSED SYSTEM DESIGN

In this chapter, we design a system for gathering insights on the sentiment of foreigners on the onset of disasters which can be useful in developing requirements for future disaster management plans geared towards addressing the needs of foreigners. The system is called "**S**NS-**b**ased **D**isaster **S**entiment **A**nalyzer (**SbaDSA**)". This chapter examines the stakeholders involved in the system and shows the proposed methodology and detailed design of the system.

1. Concept of Operations

The Concept of Operations (ConOps) of the system shows the overview of how the system will operate and on what scenario. The ConOps is shown in the figure below. The proposed disaster sentiment analyzer system SbaDSA begins with people experiencing natural disasters. In these natural disasters, people use social media to generate microblogs about the disaster. These social media data comes from anyone, foreigners, local population, and other social groups.

Next, the user accesses the SbaDSA system and requests an analysis on a certain disaster. This is conducted by providing input parameters to the system such as the dates of the disaster of interest as well as the location (latitude and longitude) of the area of interest. The system then accesses the data from Social Networking Sites (SNS) such as Twitter and extracts raw data based on the User's input parameters. Based on these inputs, the SbaDSA system filters and cleans the data and analyze disaster sentiment of foreigners. The system then outputs statistics, scores, and analysis results of the disaster sentiment of foreigners as compared with the native population. This includes figures and graphs on disaster sentiment over time which provides insights on how foreigners perceived disaster risk as well as their behavior (panicked, scared, etc.) as compared with locals. These outputs are then used by the User to develop requirements when designing new Disaster Management Plans and programs that are geared towards foreigners.

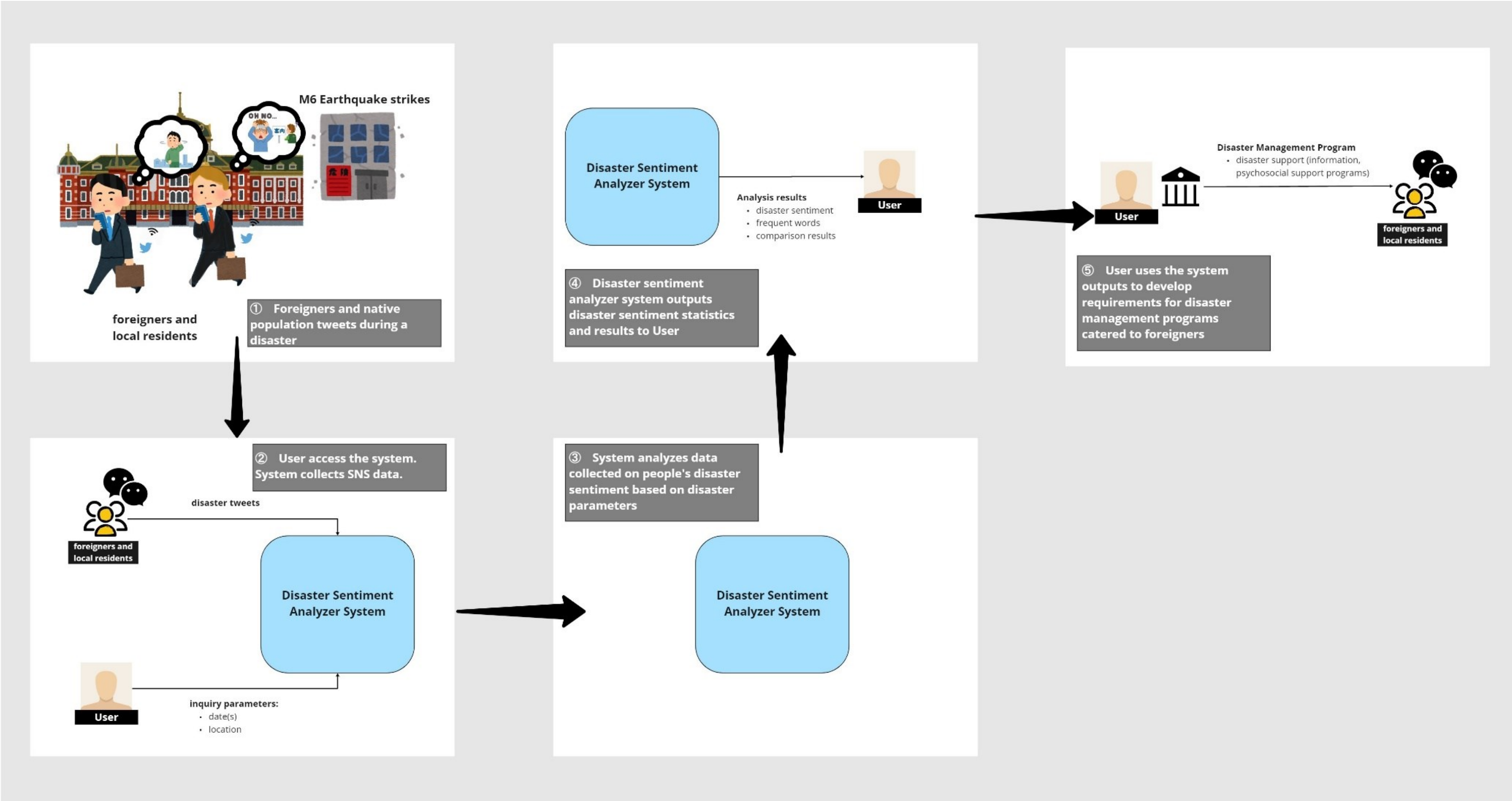


Figure 14 Concept of Operations of the Disaster Sentiment Analyzer

2. Stakeholders

There are several stakeholders interacting and affecting how the system works. As shown in the Concept of Operations, foreigners and native population are major stakeholders of the system. Foreigners are the main stakeholder affected by natural disasters and would be the end beneficiary of disaster management programs following the native residents or local population. The most important stakeholder of this system are disaster management institutions who design and implement strategies and programs on disaster risk reduction. These include the city government, national government, disaster management specialists and institutions conducting disaster research. Another important stakeholder is the SNS data company which provides the data. Other stakeholders can also affect or get affected by the system but are only considered minor. The full list of stakeholders identified are outlined here:

Table 9 List of Stakeholders of the SbaDSA System

Stakeholder	**Description**
Major	
Foreigners	Migrants, tourists, or international students in the area who are more vulnerable during disasters
Disaster Management Institutions	Local government or government institutes in charge of developing and implementing disaster risk reduction strategies on areas (cities, states) which has international and diverse population.
SNS Data Provider	Company (e.g., Twitter) which provides the open-source data on people's microblogs during disasters
Minor	
Locals or Native Population	The native population living in the area of interest (In the case of Japan, these are the Japanese people and residents at large)
Non-governmental organizations (NGO)/ International Organization for	Affiliated with the United Nations, the IMO provides guidance and advice to governments and migrants [49]

Stakeholder	Description
Migration (IMO)	
Disaster Management Specialists	Other researchers, or institutions specializing in Disaster Research

As mentioned above, the most important stakeholder are the disaster management institutions who will develop disaster management plans for its citizens or constituents. Their main needs are:

- To get better understanding of foreigners' needs and concerns during disasters, and
- To improve existing disaster management plans and programs to cater the needs of foreigners (not only the native population)

As for the foreigners, their main need is to receive better support during disasters that would enable them to increase their capability to react appropriately to disaster or minimize the impact of the disaster on their lives and well-being.

3. Requirements Analysis

Based on this, the Use Case Diagram for the system is shown in Figure 15, with the main User of the system being the Disaster Management Institutions.

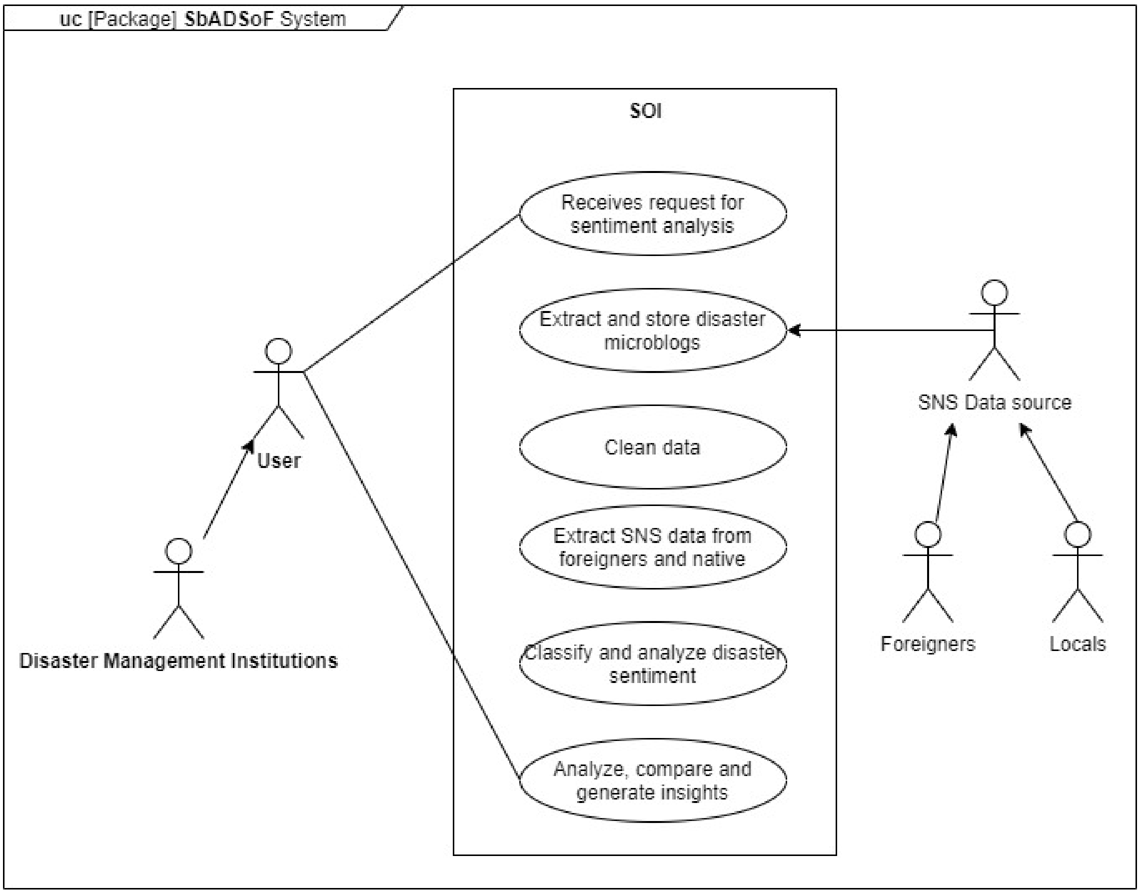


Figure 15 Use Case Diagram of the Disaster Sentiment Analyzer System

To provide a better understanding on how the system is used in the diagram in Figure 15, use case descriptions are outlined below in Table 10.

Table 10 Use Case Description

No.	Use Case Description
1	The system receives request to analyze people's sentiment for a certain disaster
2	The system collects and stores data from SNS data provider based on the disaster parameters
3	The system extract SNS data coming from foreigners as well as those coming from the native population.
4	The system prepares the data for analysis
5	The system classifies and predict the sentiment of the disaster data (microblog)

No.	Use Case Description
6	The system analyzes the data and compare sentiments
7	The system generates comparisons and insights on foreigners versus the native population

It is important to note that the Use Case Diagram shown above only focused on the disaster management institution's needs. Their needs are prioritized since they are the most important stakeholder of the system. Thus, when developing the technical requirements of the system, only their needs are considered.

4. System Requirements

Based on the Use Case Diagram shown in Figure 15, the system requirements were developed as shown below in Table 11.

Table 11 System Requirements

No.	System Requirement
1	The system shall receive request for analysis of disaster sentiment for a disaster event.
2	The system shall collect and store data
2.1	The system shall collect raw data from Social Networking Sites (SNS) based on input parameters from the User
2.2	The system shall store collected data from SNS
3	The system shall classify English and Japanese microblogs
3.1	The system shall classify data utilizing English language to represent the blogs of foreigners
3.2	The system shall classify data utilizing Japanese language to represent the blogs coming from locals or the native population
4	The system shall clean raw datasets to prepare them for data analysis
5	The system shall predict the sentiment of each disaster microblog (e.g., tweet) into positive, neutral, or negative.

No.	System Requirement
6	The system shall analyze the data and store analysis results
6.1	The system shall analyze frequent words used in the microblogs
6.2	The system shall analyze disaster sentiment through time
6.3	The system shall compare English and Japanese microblogs
6.4	The system shall store the generated charts and figures
7	The system shall generate insights on differences between foreigners and native population that could be useful for disaster planning

These system requirements 1 to 7 can be traced to the Use Case descriptions shown in Table 10. The system requirements 2, 3 and 6 are broken down into more specific requirements based on the use case. For the use case description number 5, the system analyzes the disaster sentiment. To reduce ambiguity of analysis, requirements on analysis of frequent words, analysis on disaster sentiment over time and so on are defined in the system requirements.

5. System Architecture

Upon defining the system requirements, the functional flow block diagram is presented in Figure 16. The system flow starts from receiving request from the User to analyze people's sentiment for a certain disaster event, followed by data extraction for both English and Japanese microblogs which are conducted simultaneously. The flow follows the system requirements outlined in Table 11 and is sequenced from top to bottom.

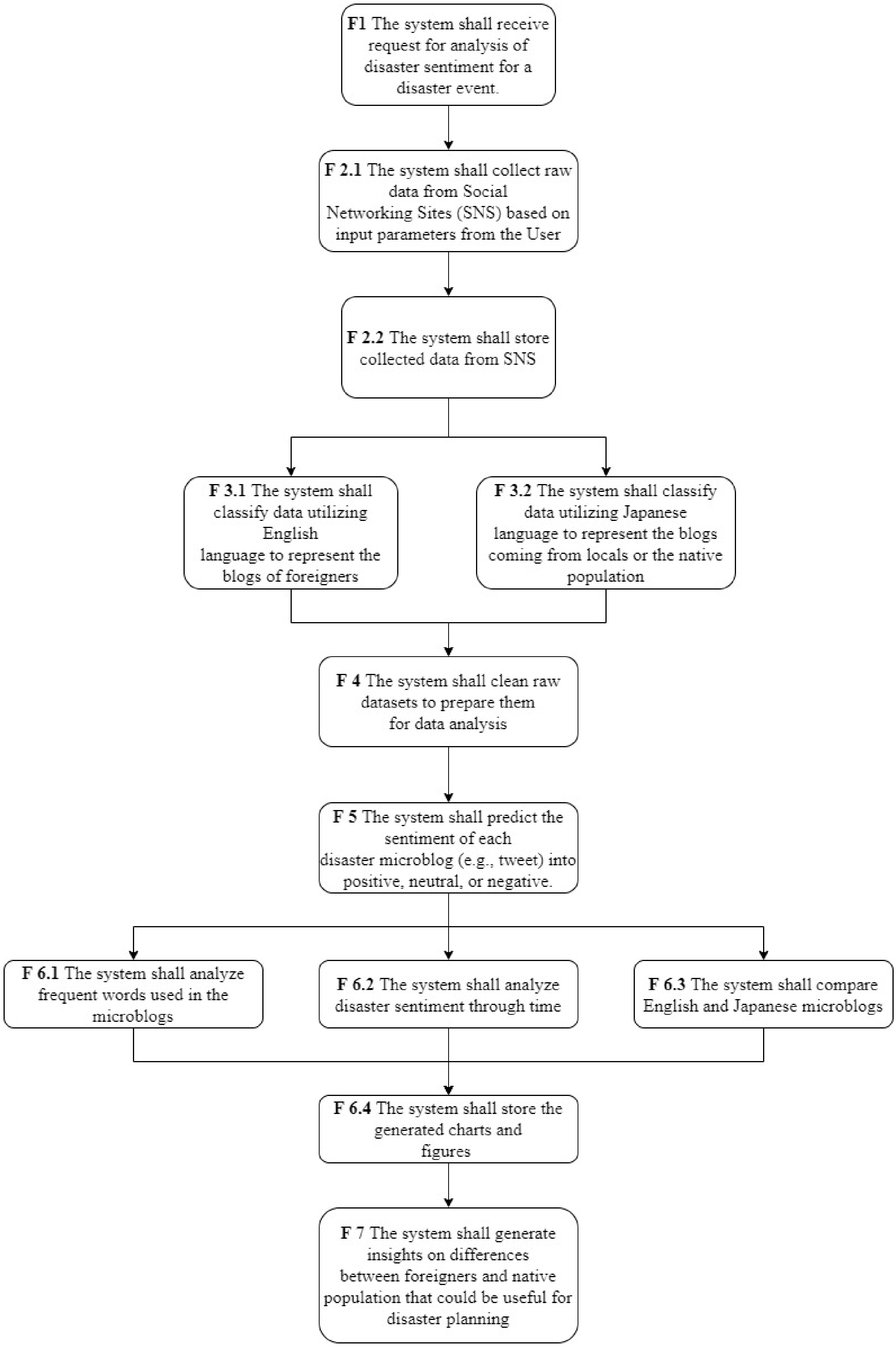


Figure 16 Functional Flow Block Diagram of SbaDSA

Based on the flow diagram, we propose an overall methodology to conduct the analysis and comparison between the two groups, foreigners, and locals. The proposed methodology is shown in Figure 17 in the case of typhoons and earthquakes using Twitter data.

First, raw data generated at times of natural disaster is extracted from the Twitter API to scrape historical data for the period of interest during the length of time (duration) of the disaster(s) of interest. Since the raw data is very noisy, preprocessing is required. Upon extraction, Twitter data, in the form of tweets, are then prepared for analysis. Preprocessing techniques are employed for both Japanese and English datasets. A sentiment classification model is then developed using Support Vector Machine (SVM) and Extreme Gradient Boosting (XGBoost) algorithms. Results of the disaster sentiment classification are stored separately in databases. Lastly, data is further analyzed using Word Cloud visualization technique. In addition, frequent words and time series analysis is presented. Insights from these analyses are presented and could be used as input in future disaster planning strategies.

Consistent with the system functions that it must perform each function is allocated to a subsystem through which the system function will be carried out. The Physical Design of the system is presented in Figure 18. The Data Extraction subsystem collected data based on the User's request depending on the scale and granularity of the scope of the User's request. Raw data can be collected from a small area and a specific time period or could span the entire country's geographical area and collect data for longer timeframes. It is important to note that there are segments of the Sentiment Classification subsystem. First is the Machine Learning algorithm which uses Artificial Intelligence to predict and automatically classify the sentiment category of each data point. On the other hand, the lexicon-based classifier calculates the intensity of sentiment (i.e., how positive or negative the sentiment is). Through the combination of these two segments, the functions performed by the Analysis and Visualization subsystem could be performed. This combination of two techniques is a key aspect of this system's architecture which is not observed in previous studies.

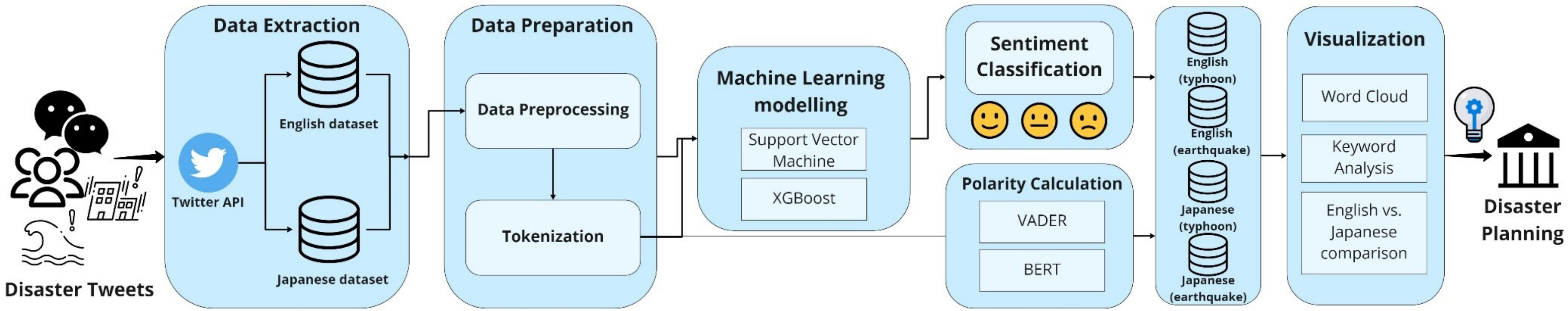


Figure 17 Proposed Methodology for Analyzing Disaster Sentiment

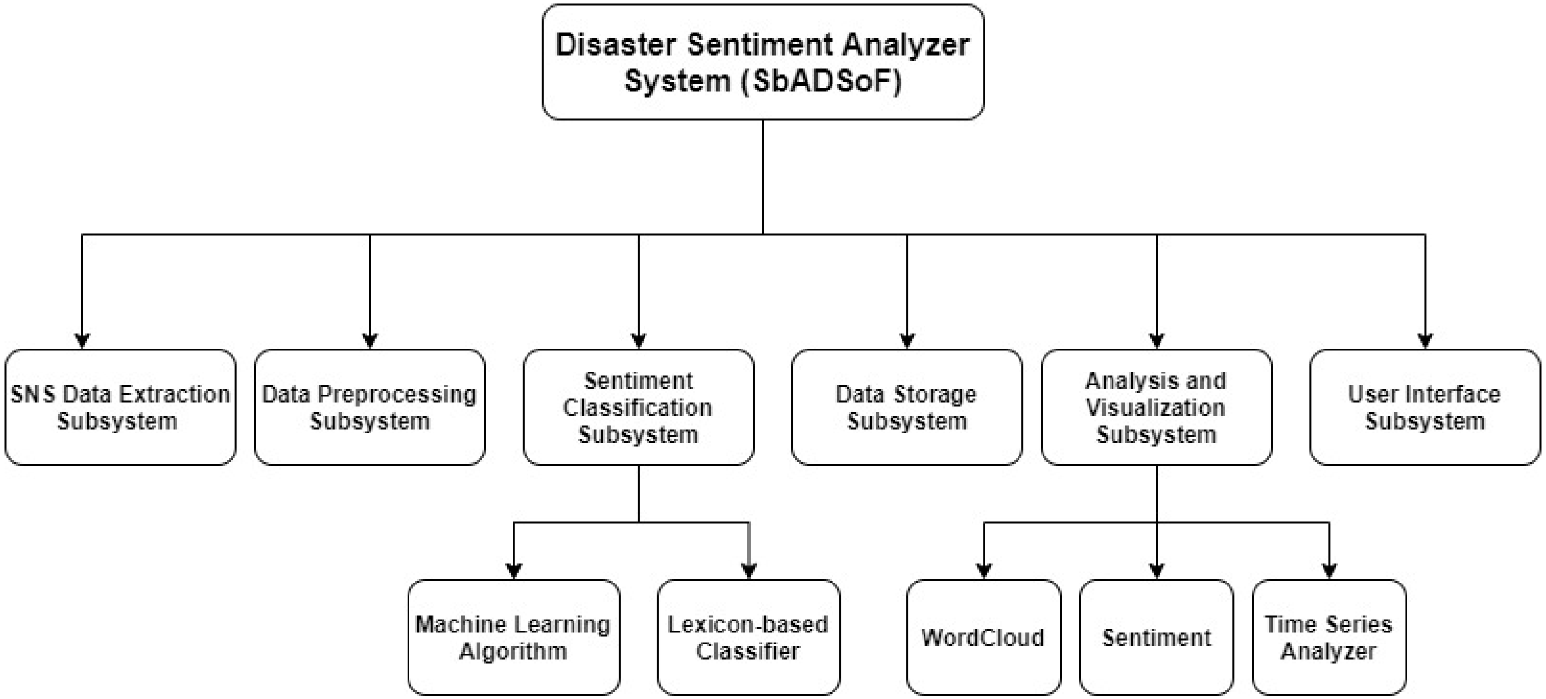


Figure 18 Physical Design of the SbaDSA System

Moreover, previous studies on Sentiment Analysis have not employed time series analysis as well as multiple language(s) as input in the data extraction subsystem.

The subsystems shown in Figure 18 is discussed below.

5.1 Data Extraction Subsystem

Raw data is collected from Twitter using its v2 API (https://developer.twitter.com/en/docs/api-reference-index). As shown in Figure 8, access to raw tweets is requested to the API. Upon generating access keys and tokens, the User is authorized to scrape raw tweets using an application.

To configure the system, a Dell XPS 9500 laptop computer with Intel core i7 10th generation processor is used to connect with the application and Twitter API and to perform all computing and data processing work. Python 3.9 was used as the programming language. To develop the models, existing libraries, specifically Scikit-Learn were used [50].

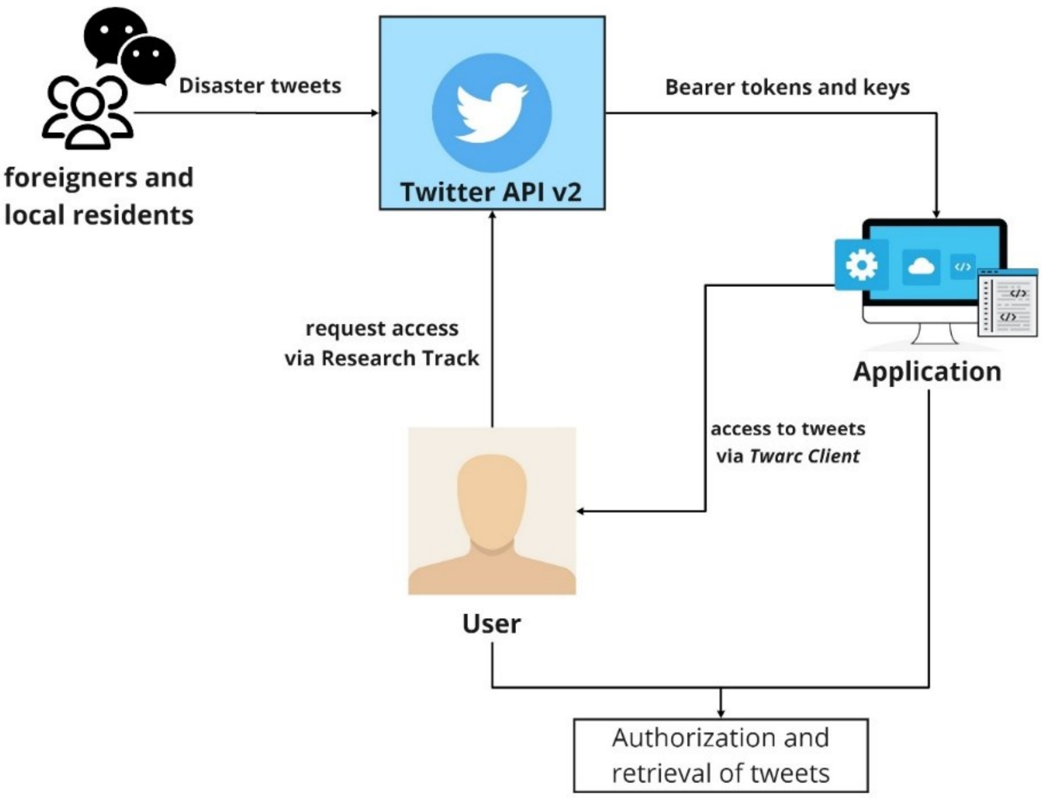


Figure 19 Tweet Extraction Architecture

Once gaining access to the historical data from the API, tweets are collected based on the parameters shown in Table 1. For each case, both Japanese and English tweets were collected. It is assumed that local residents use the Japanese language while foreigners use English or other languages. To extract tweets from each type of disaster, certain keywords were used. For the typhoon case, keywords such as {typhoon, #[typhoon name], #typhoon, #taifu} and {台風, #台風, #タイフーン} were used for English and Japanese, respectively. For the case of earthquake, {quake, #quake, earthquake, #earthquake, 地震, #地震, #jishin, jishin} were used.

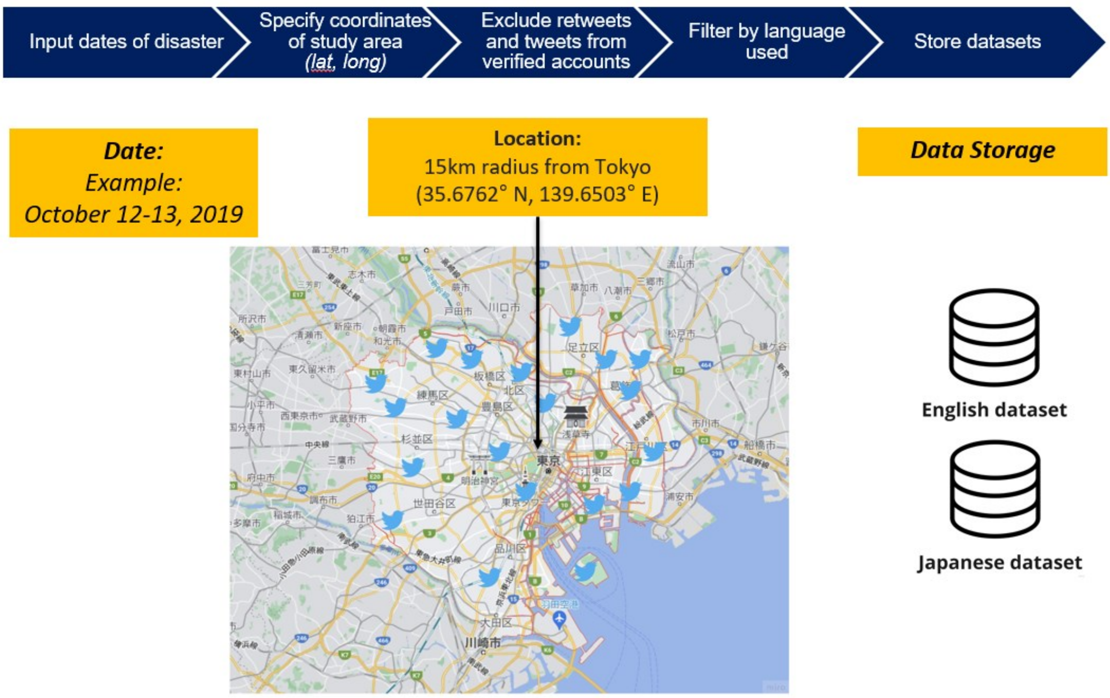


Figure 20 Tweet Extraction Method

The method of obtaining the required data from SNS such as Twitter is shown in Figure 20. First, the User provides the inclusive date(s) of the disaster event of interest and will specify the coordinates in longitude and latitude of the area of study. The radius (in kilometers) could be specified to narrow the inquiry parameters for a specific geographic area.

The system filter raw data following the data points (SNS microblog) which has any of the keywords associated with the type of disaster. For the SNS provider Twitter, the system excludes retweets and those coming from verified accounts. The system then detects the language used in the microblog and store the raw dataset of English and Japanese disaster tweets separately.

Sample tweets in Japanese and English are shown in the tables below:

Table 12 Disaster Tweets in English

Sample Tweet
Yesterdays earthquake was the most huge one ever before, just scaring me, I thought that the end of life would come close.
The damn earthquake took down my frosties! #japan #earthquake #northeast #地震 #宮城 #東北 #frosties #frostedflakes
I thought I was just tired and anemic but it was a hell of an earthquake. I should have known. 🙏 🙏 🙏
It was a magnitude 7.3 earthquake struck northeastern Japan on last night. about 11:08pm. More than 120 people have been reported injured. #Fukushima #地震 #Japan #M7.1 #earthquake
@hlcoatesmusic The earthquake happens suddenly.. I am a victim of the Great East Japan Earthquake. No matter how many times I experience the earthquake, I'm scared.. Aftershocks are still going on.. Thanks for your concern dear Harry🙏 Have a happy Valentine's Day🙏🍀🕊*·∘
@kevinwoo91 Thank you Kevin ~ 💖The big earthquake after a long time 💦 I was scared because the shaking did not stop for a few minutes 😰😰😰
Hard to believe we had a major earthquake last night. #sunset #nature #cokoguri @ Umedamachi, Sendai https://t.co/XjFvDRdHul
But fr I'm glad David and I were in a place of comfort and security when the earthquake hit. Prayers to anyone heavily effected last night. 🙏🏻
Earthquake again 😖

Sample of Japanese tweets are shown below:

Table 13 Disaster Tweets in Japanese

Sample Tweet
揺れた！地震？！って 思ったら 胎動でした。双子の威力は半端ない
地震起こる度に地球が怒ってるとか悪政への罰だとか安倍のせいとか言うのいい加減にして欲しい。日本は古来から地震大国なの。五輪マラソン中でも紅白歌合戦の最中でも、天皇即位の瞬間でも崩御の瞬間でも地震は平等に起こるんだよ。陰謀論なんか大嫌い。
地震に備えてるらしい😓 https://t.co/TzezI1dPI0
うーむ。週末の地震でどうやら壁にヒビが入り…雨漏りが酷い。参った。
息子が、地震がきて地球が割れて おうちの牛乳がこぼれたらどうする？ って心配してる。
@bakubakuiku 👹 姐様❗ １０年前も今回も余(予兆)震で、今後更に大きな本震が起こるのかもしれません😨 出来る備えをするしかありませんけどね😃 最近世界中で大きな地震が増えてるような・・・🤔 #あいうえお時事川柳 煮えたぎる マグマを加熱 温暖化 酔爺 https://t.co/V3XX9ulAF7
@ai_zome 地震も怖いですけど、一昨年の台風の被害も酷かったです(>_<)家の近所です💦自然災害は恐ろしいです。スーパーも買い占めが凄くてしばらくガランでした(๑•̀ㅁ•́๑) https://t.co/6iEnKnQwlv
熊谷市長は土曜日深夜の地震発生からすぐに状況や市の対応をツイートし非常に積極的に動いています。国全体でもこのような動きがあると安心できますね。 [url]
週末の地震は凄かった💦 震災の時を思い出して怖くなったよ😵

5.2 Data Preprocessing Subsystem

For each dataset, data were collected for a period of 48 hours to extract data during the onset of the disaster. The extraction flow is shown in Figure 21.

We collected tweets which have geo location and only geo-tagged tweets coming from Tokyo were collected. Moreover, tweets coming from verified accounts such as news organizations or official Twitter celebrities were filtered. Retweets were also removed from the dataset.

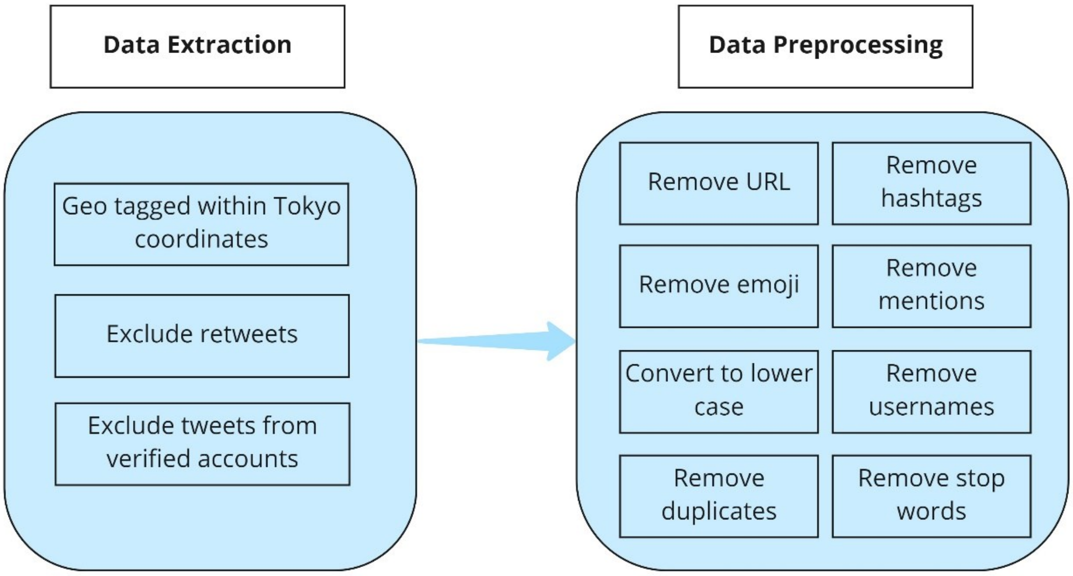


Figure 21 Data Extraction and Preprocessing Techniques

Raw twitter data needs to be cleaned of uninformative noise terms as shown in Figure 3. The packages used here have been commonly used in the literature. Some, however, were specifically used for the reasons stated. First, raw tweets were stripped of twitter mentions (i.e., @twitter_username) because the specific interconnections were not relevant to this study. Hashtag (e.g., #earthquake or #typhoon) characters (i.e., the pound sign) were also removed but not the associated words, as well as punctuation, and emojis or their Japanese counterpart, kaomoji. These were done with the python packages tweet-preprocessor and nltk . Next, uninformative words known as stopwords (e.g. articles, conjunctions, linking verbs, etc.) were also filtered out using the Natural Language Toolkit *nltk.stopwords* package and an additional custom list defined by the author.

The outputs were then lemmatized with *nltk.WordNetLemmatizer* to correct for inflections. These steps are necessary as they would only be noise if they were to add to the dimensions of the data, at least in this context. This would confuse the machine learning algorithms to be employed.

Only then were the preprocessed tweet strings converted into text vectors, i.e., tokenization, using sklearn's *CountVectorizer* and *TfidfVectorizer*. The output is finally then for input to the machine learning algorithms.
For tokenization of Japanese tweets, the pre-trained bert-japanese model [51] from Inui Laboratory of Tohoku University was used.

5.3 Sentiment Classification Model

Sentiment analysis is usually conducted using either a lexicon-based analyzer or using Machine Learning (ML) algorithms. In this study, two supervised ML algorithms were implemented: Support Vector Machine (SVM) and Extreme Gradient Boosting (XGBoost).

To automatically classify an unlabeled tweet's sentiment (positive or negative), both algorithms first need to make a classification model based on (i.e., learning from) pre-labelled tweets (i.e., the output of tweet preprocessing) where each of the tweets' sentiment is known. Each model respective to an ML algorithm would then be the basis of classification of new tweets of unknown sentiment labels as shown in Figure 22.

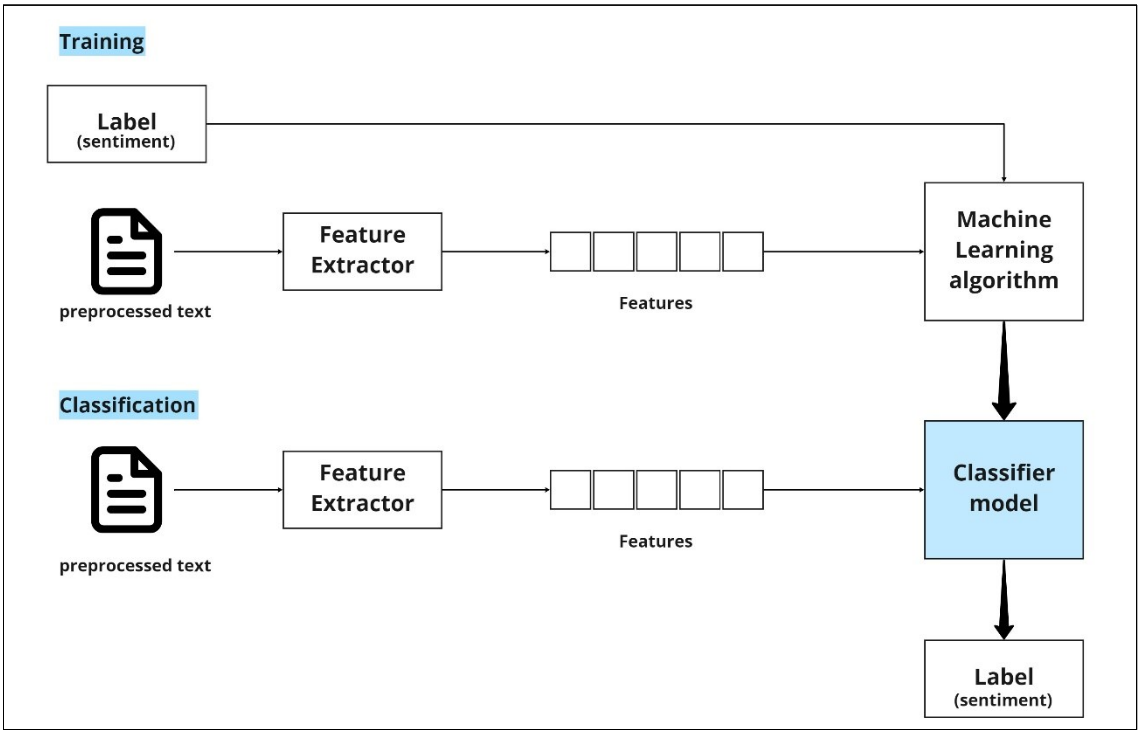


Figure 22 Supervised Machine Learning model

To develop the Machine Learning algorithm, first, a training dataset is prepared to be used as an input for the supervised classification model. A number of tweets is randomly selected from the original dataset. Since the Japanese dataset is large, only 1000 tweets were randomly selected. The tweets were manually labelled with positive, neutral, or negative sentiment. Since the author is a native English speaker, sentiment labels for the English tweets are considered accurate. For the Japanese dataset, the tweets were first translated to English using the online translator *Deepl* (https://www.deepl.com/translator) and were then interpreted by the author.

The data preprocessing techniques discussed previously yield high-dimensional numerical data for the ML algorithms to classify. In SVM, the number of dimensions in the data is the number of tokens (components) generated by the vectorizer. During learning, SVM seeks to find the best surface in all dimensions of the data to separate them according to their own pre-labelled classes [52].

Support Vector Machine

Support Vector Machine (SVM) is a supervised machine learning algorithm which is commonly used for classification or regressions challenges. SVM performs classification by finding the hyper-plane that differentiate the classes in n-dimensional space as shown in Figure 23. The hyperplane is drawn through the use of mathematical functions called kernels. Support Vectors are data points close to the hyperplane which help build the SVM model [53].

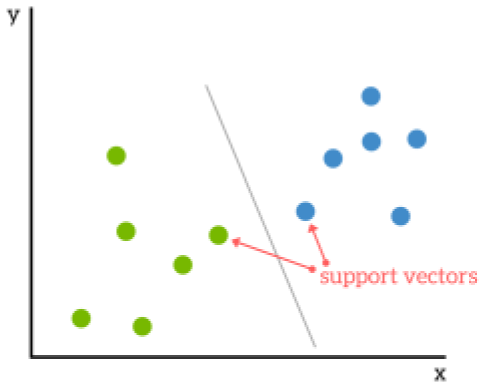


Figure 23 Support Vector Machine algorithm

Extreme Gradient Boosting (XGBoost)

XGBoost, on the other hand, learns by aiming to approximate the training dataset in terms of additive functions. The direction of learning is controlled by minimizing both the loss (prediction error) and the smoothing parameter (which is also to avoid model overfitting) [30].

As shown Figure 24, XGBoost utilize gradient boosting technique as a machine learning algorithm and applies weights in datasets to account for overfitting or underfitting of datasets. For each classifier, the algorithm is “boosted” by applying and updating weights which serve as "punishments" for the system. This technique boosts accuracy of the classifier model.

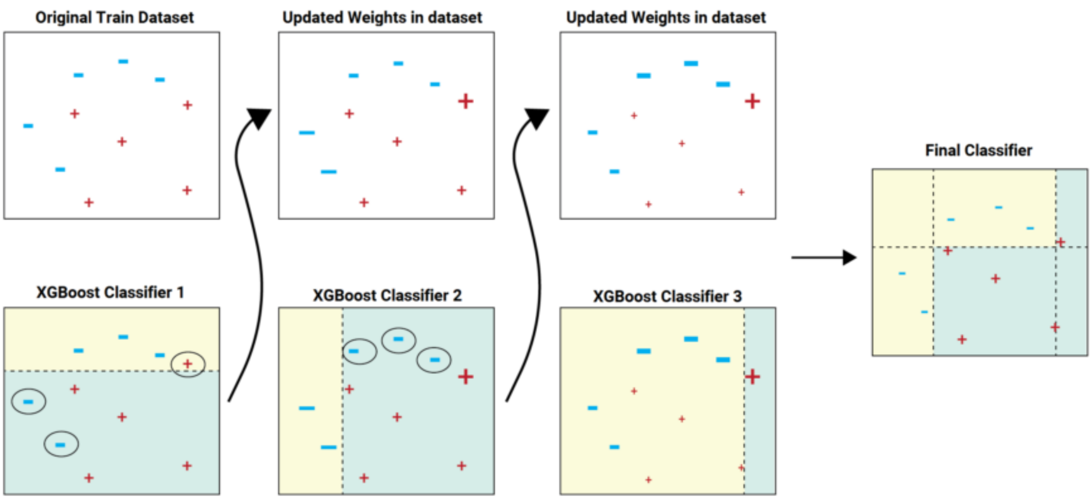


Figure 24 Extreme Gradient Boosting (XGBoost) Algorithm [54]

Sentiment Classification

To complete the supervised Machine Learning model to classify each microblog (e.g., tweet), it is necessary to allocate a portion of the data as training data. The training data is manually labelled as positive, neutral, or negative. To avoid ambiguity, the following definition is proposed in this system:

Table 14 Sentiment Classification Definition

Sentiment Classification	**Definition**	**Example**	**Polarity Score**
Positive (pos)	• the microblog/ tweet expresses positive feeling, calmness, attitude, or concern	"I feel so relieved ease after the typhoon"	1
Neutral (neu)	• the microblog/ tweet	"The weather suddenly	0

Sentiment Classification	Definition	Example	Polarity Score
	no inclination towards positive/ negative; or merely states facts	changed after the typhoon”	
Negative (neg)	• the microblog/ tweet shows distress or negative emotion	“This is scariest earthquake I’ve ever experienced”	-1

5.4 Data Visualization and Comparison

This subsystem includes three subsystems: Word Cloud, Disaster Sentiment calculation, and time series analysis. These are major data visualization techniques to be employed by the system. System outputs will comprise of those charts and diagrams for the disaster event of interest.

6. Implementation

The proposed system design most consist of software and internet-based interfaces to be implemented. The most important functions of the system are Functions 5 and 6 which are to classify and predict the sentiment of each microblog and analyze them in the form of data analysis and visualization techniques.

IV. SYSTEM EVALUATION

1. Test Area

To understand the sentiment during disasters, two cases were selected for this study. Since the paper highlights comparison between sentiment of local residents and foreigners, only the city of Tokyo is considered for this study. Tokyo currently has around 4% of the population to be international. Hereby making it a good area of choice. Earthquakes and typhoons are the most common natural disasters in Japan and were thus selected. The two case studies are as follows:

Case1: Typhoon Hagibis

Typhoon Hagibis, also known as Reiwa 1 East Japan Typhoon (令和元年東日本台風) is considered one of the strongest tropical cyclones to hit Japan when it made its landfall on 12 October 2019 around 7 p.m. (Japanese Standard Time) in the Greater Tokyo Area [55]. Around this period, a magnitude 5 earthquake also struck the coast of Chiba Prefecture, east of Tokyo.

Figure 25 Excessing flooding damage in Tokyo during Typhoon Hagibis

Case2: 2021 Fukushima Earthquake

The Fukushima Prefecture Offshore earthquake, also known as the *2021 Fukushima Earthquake*, struck the shores of Fukushima, 135 miles north of Tokyo, with a magnitude of 7.3 on 13 February 2021 at 11:07 P.M. It measured a strong 6 on the Japanese seismic intensity scale which is the second-highest level.

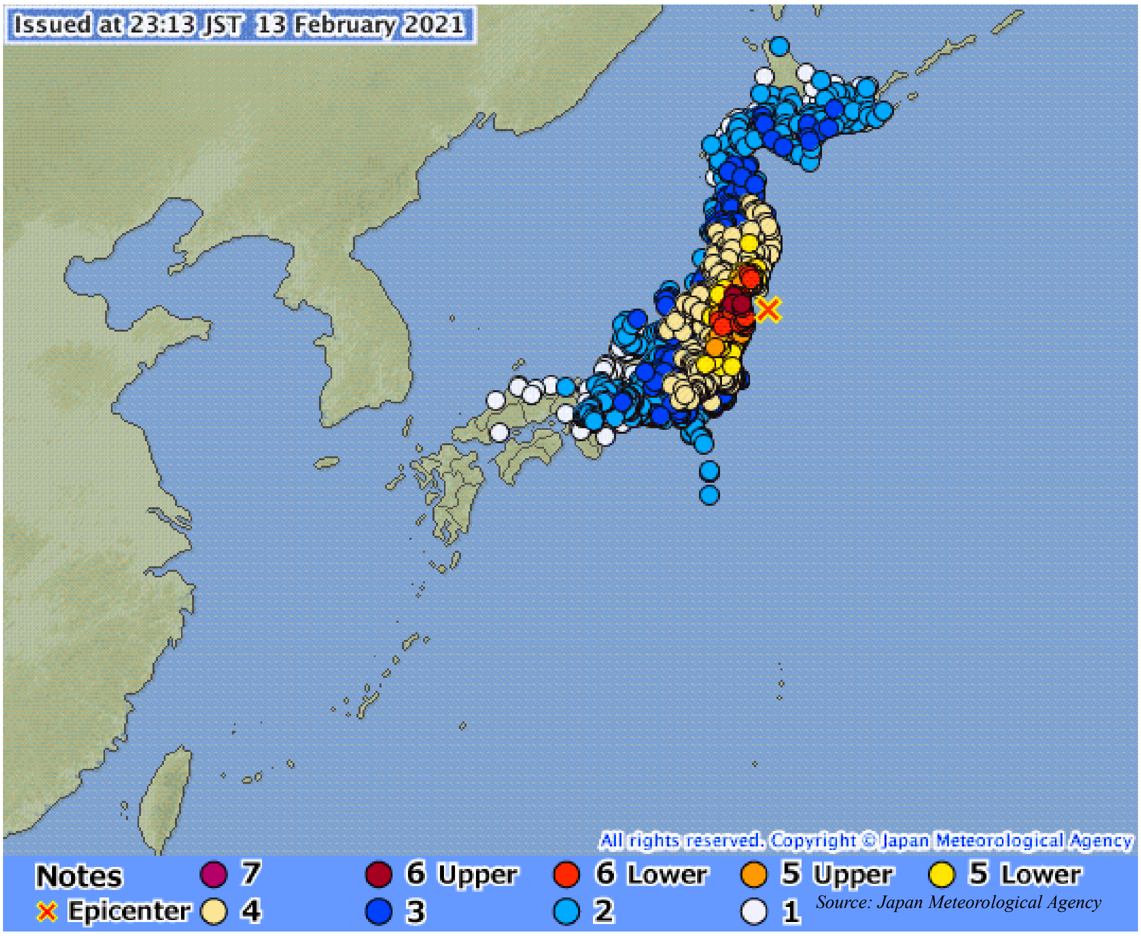


Figure 26 2021 Fukushima Earthquake intensity

Evaluation of the system in terms of verification and validation are focused on these two specific disaster events and will be the main disasters of interest throughout the rest of the thesis.

2. Verification

To verify if the system performs its functions and fulfills the requirements, several activities are performed which are simulation and demonstration of the system. The traceability of requirements and verification method used is shown in Table 15:

Table 15 Requirements Verification Traceability Matrix

No.	System Requirement	Verification Method
1	The system shall receive request for analysis of disaster sentiment for a disaster event.	Demonstration/ Test • Test if system could process request for both disaster and non-disaster events
2	The system shall collect and store data	Prototyping
2.1	The system shall collect raw data from Social Networking Sites (SNS) based on input parameters from the User	Prototyping/ Test • Test system output (raw data) can be extracted based on parameters
2.2	The system shall store collected data from SNS	Demonstration
3	The system shall classify English and Japanese microblogs	Simulation/ Inspection
3.1	The system shall classify data utilizing English language to represent the blogs of foreigners	Simulation/ Inspection • Inspect if the microblogs classified are in English during simulations
3.2	The system shall classify data utilizing Japanese language to represent the blogs coming from locals or the native population	Simulation/ Inspection • Inspect if the microblogs classified are in Japanese during simulations
4	The system shall clean raw datasets to prepare them for data analysis	Simulation/ Prototyping

No.	System Requirement	Verification Method
5	The system shall predict the sentiment of each disaster microblog (e.g., tweet) into positive, neutral, or negative.	Prototyping/ Test • Test if microblogs are classified into the three categories. • Test algorithm accuracy if it is more than 50%
6	The system shall analyze the data and store analysis results	Demonstration/ Prototyping
6.1	The system shall analyze frequent words used in the microblogs	Demonstration • Demonstrate if frequency of words is analyzed
6.2	The system shall analyze disaster sentiment through time	Demonstration • Demonstrate through earthquake and typhoon scenarios
6.3	The system shall compare English and Japanese microblogs	Demonstration • Demonstration through earthquake and typhoon scenarios
6.4	The system shall store the generated charts and figures	Demonstration • Demonstrate through earthquake and typhoon scenarios
7	The system shall generate insights on differences between foreigners and native population that could be useful for disaster planning	Demonstration • Demonstrate through earthquake and typhoon scenarios

As could be observed, the main verification used is through demonstration that the system works. To achieve this, prototyping is conducted through various scenarios, including events that are not natural disasters.

2.1 Prototyping

Prototyping of various algorithms and disaster events are conducted. One simulation was on an earthquake which happened last May 1, 2021, near Fukushima, Japan. A sample of the extracted data from Twitter for both English and Japanese is shown below in Figure 27.

```
In [45]:  print(df1['Tweet'].head(20))

0     @wanimaruwani1 おはﾜﾝ♪昨日の地震は大きかったですよね？びっくりされたでしょ...
1     @5Q4YJCa9ECEX7sJ おはようございます☀本日も1日宜しくお願いします😁最近地震...
2     RT @hiyoko5stars: みんな忘れてるよ。新型コロナウィルス感染拡大初期の頃、危...
3     RT @katukawa: 処理済みのトリチウム水は、基準値以下に希釈した上で排出して、量を...
4     5/1  【まとめPM】②#地震予測  #遊動コンパス  #ユーコン https://t.co/...
5     RT @VkdsqnFrymIumLT: @matsuzaka_1021 アモーレ💖地震や天...
6     5/1  【まとめPM】①#地震予測  #遊動コンパス  #ユーコン https://t.co/...
7     RT @matsuzaka_1021: 昼間は晴れて地震。夜は雷⚡目まぐるしい一日でしたが...
8     @President_mie27 昨日は凄かったね⚡地響きもした地震もあったしね💦💦💦コロコ...
9     RT @asanonami: これだけ地震が頻発しても、築４０年以上たつ原発を再稼働させる。...
10    @f1p6pUmereinyan おはようございます！\n\n地震、雷雨と💦なことばかりです...
11    気象庁「地震発生の2、3日程度 強い揺れをもたらすこと多い」 | NHKニュース  htt...
12    RT @arakencloud: 大規模な地震の発災時・避難時の注意点等は「東京防災」にまと...
13    RT @MY_MURMUR: さっきの地震、津波の心配はないようですね。火の元など大丈夫でし...
14    RT @hiyoko5stars: みんな忘れてるよ。新型コロナウィルス感染拡大初期の頃、危...
15    @himawari986 地震多いですよね！\n歪みを震度4程度で解消させていてくれたら巨大...
16    @kinironotsubasa おはようございます。😃\nまた仙台? 😱\n地震の度にドキド...
17    RT @kentaro666: このままだと首都圏直下型の大地震が起きて都民が十万人くらい焼...
18    @naniwanomachi77 おはようございます！\n\n地震やら、雷雨やら。。。\n\...
19    RT @hiro0725: 🎙トランプ：5/1声明🎙\n【おはようございます】\n今日も...
Name: Tweet, dtype: object
```

Figure 27 Prototype Sample of Japanese Tweet Extraction

As can be observed in Figure 27, the system was able to extract Twitter data in Japanese language accurately. Similarly, accurate results were also achieved in the English dataset as shown in the sample below:

```
In [34]:  print(df0['Tweet'].head(20))

0     RT NERV Earthquake Detailed Report 5 1 At arou...
1     Eh Just found out that yesterday s earthquake ...
2     RT Strong earthquake hits northeastern Japan h...
3     RT Magnitude 6 6 earthquake jolts northern Jap...
4     I woke up to a strong earthquake this morning ...
5     ICYMI There was a strong earthquake earlier to...
6     Japan A strong earthquake with a magnitude of ...
7     RT NERV Earthquake Detailed Report 5 1 At arou...
8                        Thank you for the earthquake
9     RT UPDATE M6 6 earthquake strikes off northeas...
10    RT NERV Earthquake Detailed Report 5 1 At arou...
11    RT NERV Earthquake Detailed Report 5 1 At arou...
12    RT Earthquake shaking Harajuku No worries here...
13              Earthquake and a thunderstorm Busy day
14    Earthquake Detailed Report 5 1 At around 10 06...
15    What a day Started off with a serious earthqua...
16    First time experiencing an earthquake and a th...
17                RT holy shit earthquake make it stop
18    RT Japan News A magnitude 6 6 earthquake struc...
19    RT NERV Earthquake Detailed Report 5 1 At arou...
Name: Tweet, dtype: object
```

Figure 28 Prototype Sample of English Tweet Extraction

For detailed information on the results of this prototyping activity, see Appendix.

To create prototype of the data visualization subsystem of the SbaDSA system, a simulation is also conducted on a non-disaster event to perform early verification. A simulation is conducted on the issue of former Tokyo Olympic and Paralympic Committee Chief, Yoshihiro Mori, which gained international media coverage over sexist public comments [56]. A prototype is made to evaluate the data visualization subsystem for both English and Japanese.

Sample of the result is shown in the figures below. It can easily be observed that the data visualization for Japanese is too messy. Furthermore, the text is not segregated into the list of words compromising it. Although the prototype of the Word Cloud visualization subsystem for Japanese failed, the result for English dataset in Figure 30 is a success. This is a challenge with Natural Language Processing (NLP) techniques because most data preprocessing techniques are only for English language.

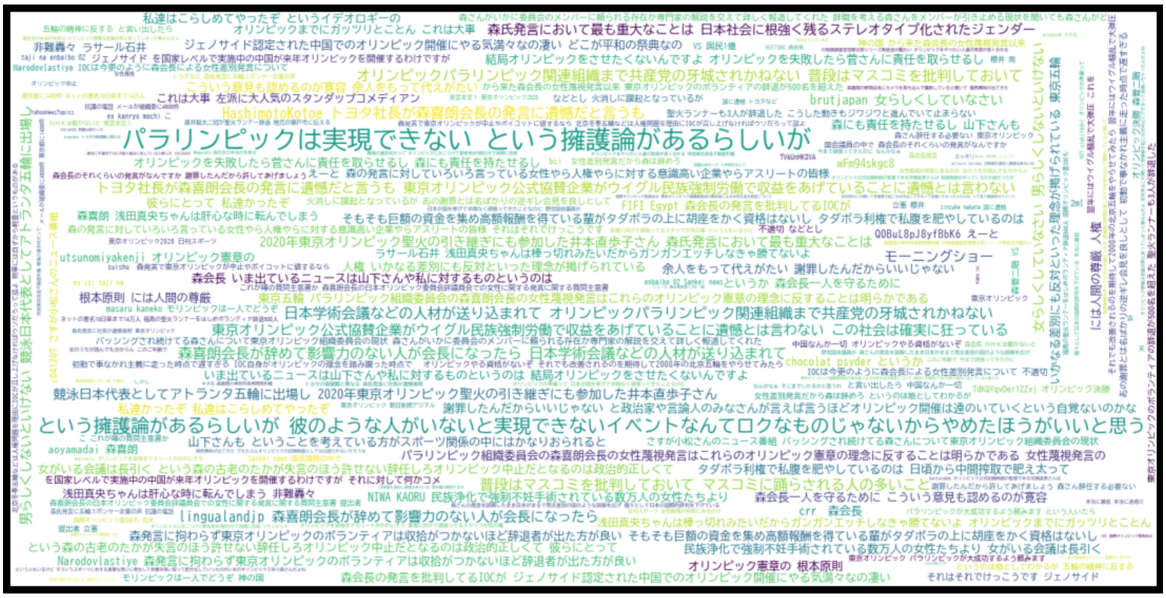


Figure 29 Prototype Sample of Japanese Word Cloud

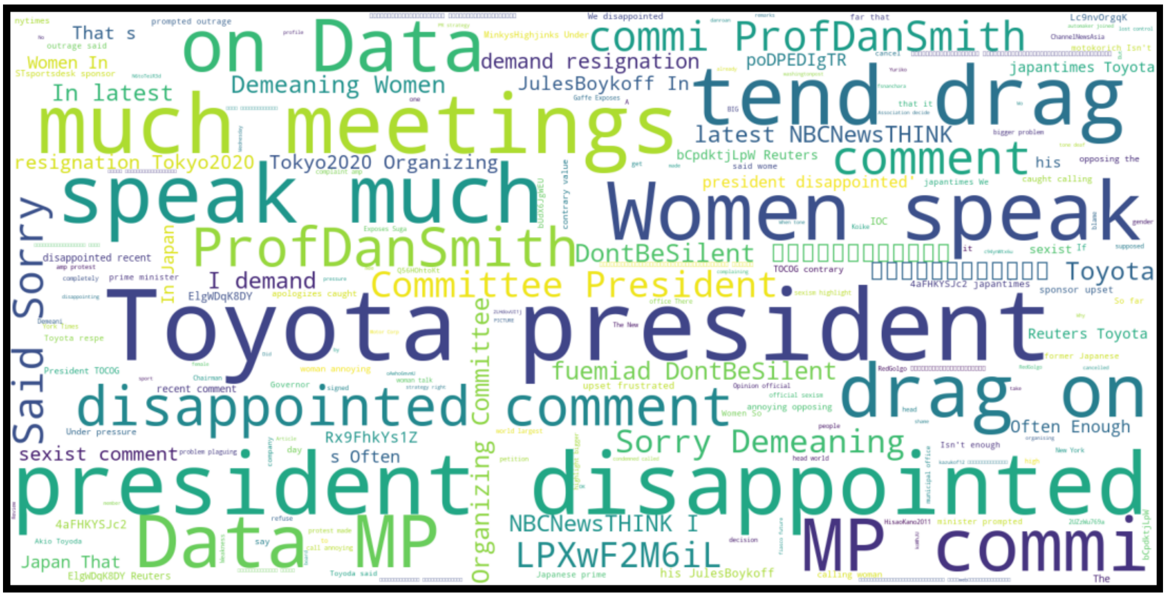


Figure 30 Prototype Sample of English Word Cloud

Because of the failure of the visualization subsystem to accurately visualize the frequent words in the Japanese dataset, prototyping was continued to develop a better algorithm for cleaning the data and generating the Word Cloud. For more details of the prototyping results, refer to the Appendix.

2.2 Case Studies

As mentioned above, the most prominent method to evaluate the system is through actual demonstration and testing. To achieve this, two case studies are presented in Section 4.1. The following shows the result of these case studies.

For each case, both Japanese and English tweets were collected. It is assumed that local population use the Japanese language while foreigners use English or other languages. To extract tweets from each type of disaster, certain keywords were used. For the typhoon case, keywords such as {typhoon, #Hagibis #hagibis, #typhoon, #TyphoonHagibis} and {台風, #台風, 令和元年東日本台風} were used for English and Japanese, respectively. For the case of earthquake, {earthquake, #earthquake, 地震, #地震} were used.

The input parameters for these disaster events are shown in Table 16. The location of interest is Tokyo city, with geographic coordinates of 35.6762° N, 139.6503° E.

The resulting datasets after filtering were shown in Table 16.

Table 16 Collected Tweets After Filtering

Dataset	***Timeframe***	***# of Tweets***	
		Japanese	***English***
Typhoon	October 12-13, 2019	12440	858
Earthquake	February 13-14, 2021	5971	201

As shown in Table 15, the timeframe only covered two days which is equivalent to 48 hours. This is to capture people's sentiment on the onset of disasters which are considered to be the most critical period. In disaster response, the first 72 hours after the disaster is mist crucial especially for rescue operations. In these case studies, this timeframe is further shortened. For the case of typhoon, the 48-hr timeframe started the day before the landfall up to the period right after landfall.

2.2.1 Word Cloud and Frequent Words

To deepen understanding of the issues associated with positive sentiment as well as negative sentiment, we generate a Word Cloud to visually see the frequent words used in the text. A Word Cloud is an image showing a list of words. The size of the word indicates the frequency of that word in the text. Thus, by looking at the Word Cloud, frequent words which are the most important words can be identified and easily analyzed. After classifying the tweet's sentiment category (i.e., positive, neutral, or negative), Word Clouds generate a much deeper analysis on the content of the tweets. For the purpose of evaluating the analyzer system, only positive and negative sentiments are considered. Tweets under the Neutral sentiment category are excluded.

Based on the preprocessed tweets, we tokenize all words and use the text data to plot word clouds of the positive and negative sentiments from both English and Japanese datasets. Figure 31 exhibits the word cloud of positive English tweets. Words like 'thank', 'safe', 'today', and 'everyone' dominate the cloud. Moreover, the word rugby is mentioned for the typhoon case. This is because at the time of typhoon Hagibis, the World Rugby Cup was being held as well. Negative tweets Word Cloud is shown in Figure 6. Words like 'scary', and 'worst' dominate the Word Cloud. People clearly expressed their sense of fear and intense emotion. Curse words like 'shit' can also be seen. It is worth noting that the negative sentiment during typhoons also include 'defense minister' indicating their concern on the government's disaster relief efforts. For the negative earthquake sentiment, the word 'phone' and 'alarm' are one of the most frequently used. Interestingly, the tweets air concerns that the earthquake alarms on their mobile phones did not work.

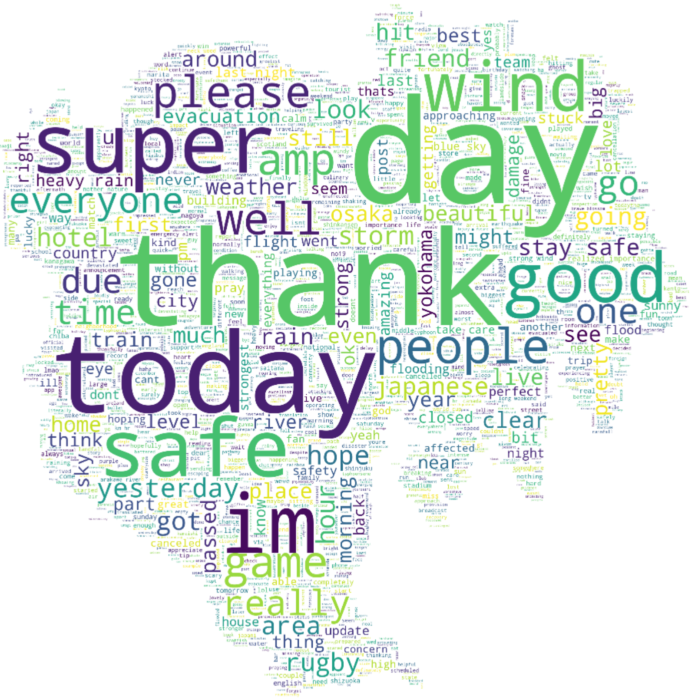


Figure 31 Word Cloud of Positive Typhoon Sentiment of Foreigners

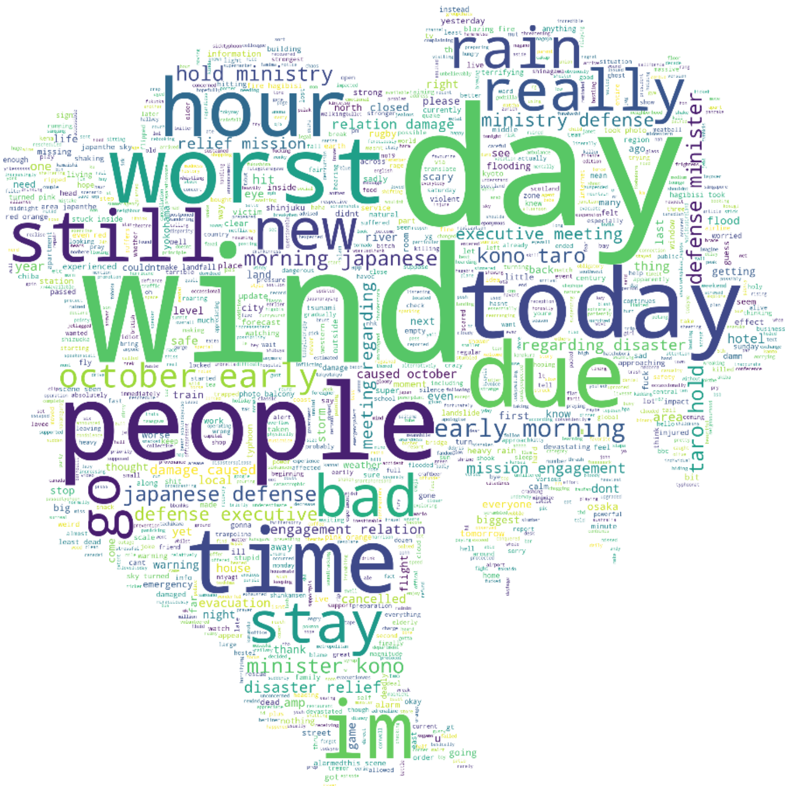


Figure 32 Word Cloud of Negative Typhoon Sentiment of Foreigners

Figure 33 Word Cloud of Positive Typhoon Sentiment of Japanese

Figure 34 Word Cloud of Negative Typhoon Sentiment of Japanese

For the Japanese tweets, the Word Cloud is exhibited in Figure 33 and 34 for the positive and negative sentiments, respectively. We observe that for three scenarios, the word ‘大

丈夫’ which means okay or alright is present.

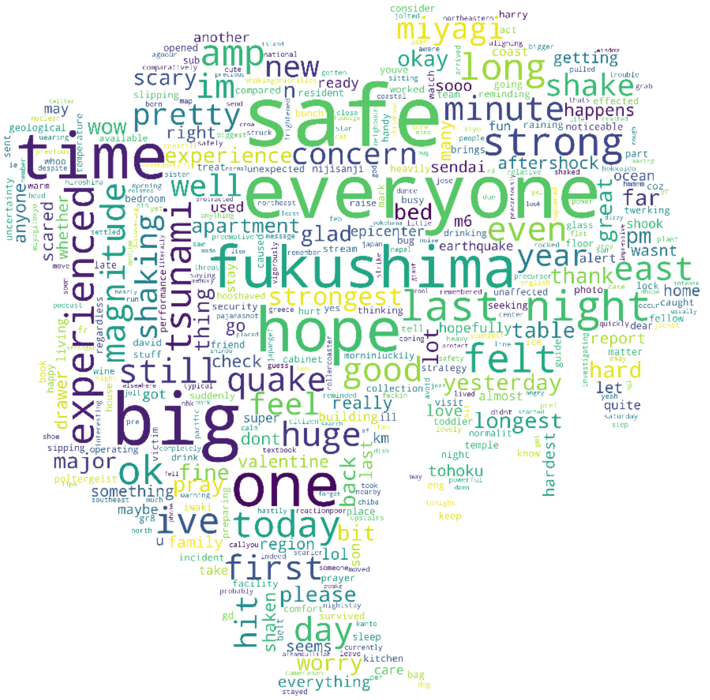


Figure 35 Word Cloud of Positive Earthquake Sentiment of Foreigners

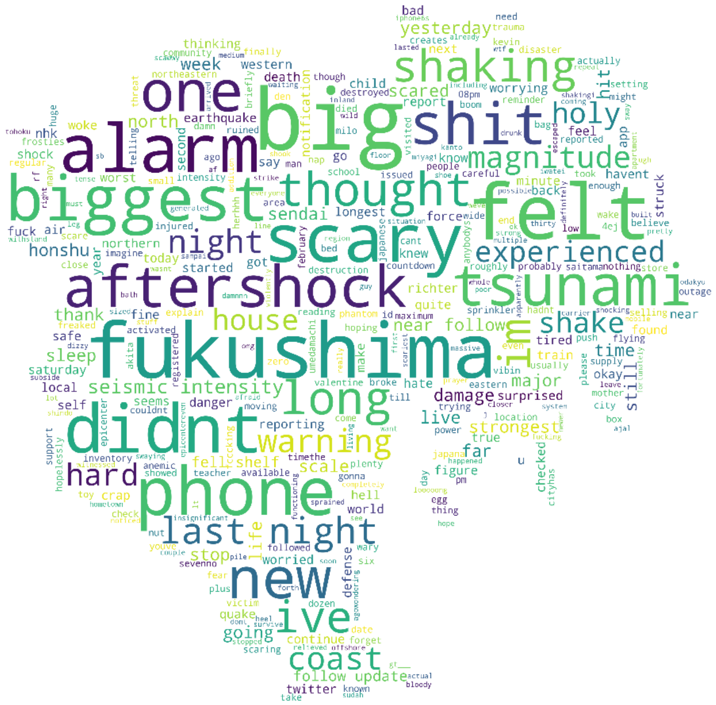


Figure 36 Word Cloud of Negative Earthquake Sentiment of Foreigners

Figure 37 Word Cloud of Positive Earthquake Sentiment of Japanese

Figure 38 Word Cloud of Negative Earthquake Sentiment of Japanese

It should be noted that the word ‘揺れる’ which means ‘to shake’ is a frequent word in all scenarios even during the typhoon. These tweets exhibited their sentiment on their houses being shaken, not just in earthquakes but during typhoons as well. Moreover, an earthquake also strikes near Tokyo during the typhoon Hagibis and showed people’s sentiment on having two natural disasters at the same time. This insight reveals the need for multi-hazard analysis of sentiment and people’s needs during disasters, too.

2.2.2 Disaster Sentiment

In order to build the classifier, it is essential to have training datasets. For these two cases, samples of tweets are manually labelled as shown below:

Table 17 Training Dataset Extracted from Japanese and English tweets

Tweet	Sentiment	Polarity
おはようございます😃 皆さん台風🌀大丈夫でしたか？ 甚大な被害の地域のトレーナーさん達も中にはいらっしゃると思いますが、力を合わせて復興頑張って下さい‼ライダーも支援物資運びます🚚💨💨💨💨💨 *Translation:* Good morning 😃. I hope everyone is ok with Typhoon 🌀. I'm sure there are some trainers in the areas that were severely affected by the typhoon, but please work together to rebuild the area Riders are also carrying relief supplies 🚚💨💨💨💨💨	pos	1

Tweet	Sentiment	Polarity
台風一過…7 ホッピー通りは普段から、店じまいは早いのですが、やはり空気感が違いました。 *Translation:* After the typhoon...7 Hoppy Street usually closes early, but the atmosphere was still different.	neu	0
台風が去ったのに秋晴れは昨日たったの1日だけでしたね💦 今日も明日もお天気が悪くて憂鬱です😢 *Translation:* The typhoon went away and we only had one day of clear autumn weather yesterday💦. Today and tomorrow the weather will be bad and depressing😢.	neg	-1
Catching up with my friend, feeling so ease after the typhoon Hagibis 😊	pos	1
Off to see if the local conbinis have food after the typhoon. I'm hungry but lazy	neu	0
Like the typhoon is not scary enough. A strong earthquake has to occur too 😔	neg	-1

The algorithms developed are now tested on the raw dataset. Disaster sentiment of each tweet is predicted by the system for both English and Japanese datasets on both typhoon and earthquake cases. Of the four models trained, the most accurate was used to classify the sentiment of all tweets in the English and Japanese datasets for the typhoon and earthquake datasets (Figure 39-42).

The results are shown in the figures below:

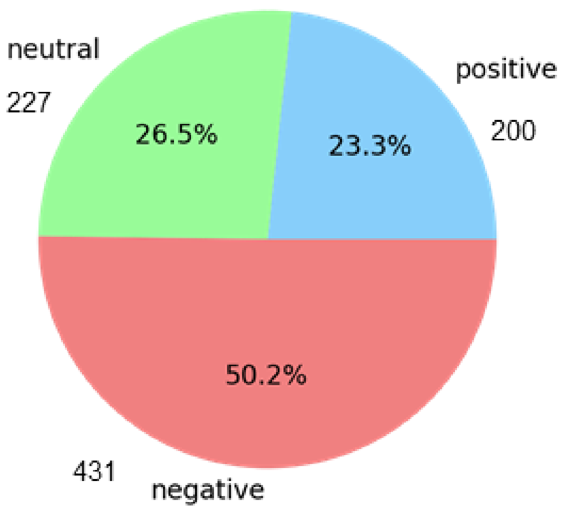


Figure 39 Disaster Sentiment of Foreigners during Typhoon

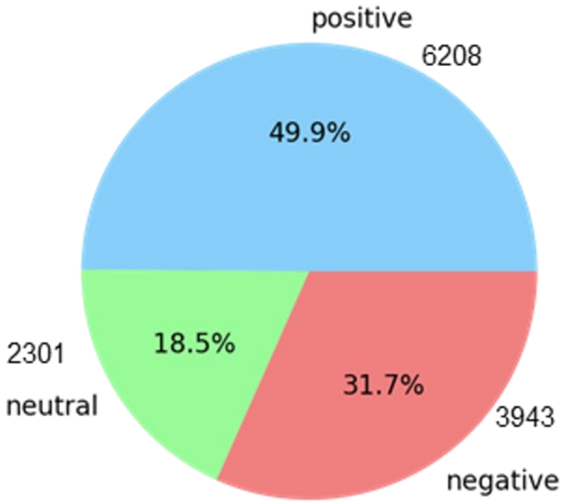


Figure 40 Disaster Sentiment of Japanese during Typhoon

For the case of typhoon, it can be seen in Figure 40 that almost half of the Japanese tweets showed positive sentiment while in Figure 39, only 23% of English tweets are positive. Looking at the total number of data points in Table 16, there is a total of 12,440 Japanese tweets and 858 English tweets. Half of disaster tweets in Japanese are positive but for English, half of it is of negative sentiment.

For the case of earthquake, the result of sentiment classification is shown in Figure 41 and Figure 42, for English and Japanese tweets, respectively. The total number of English tweets is 201 while for Japanese 5,971 tweets. These numbers are much lower than that of typhoon datasets. A large majority of English tweets are negative, with an overwhelming 78.1%. For the Japanese tweets, however, only 32% are negative. In addition, the Japanese tweets during earthquake showed a balanced sentiment distribution as shown in Figure 42.

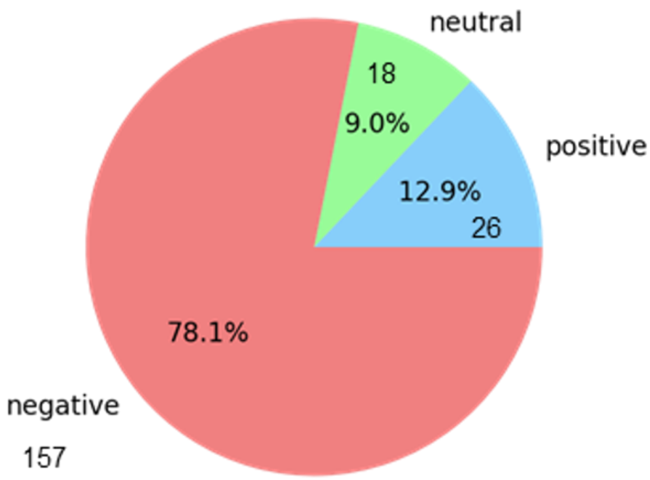


Figure 41 Disaster Sentiment of Foreigners during Earthquake

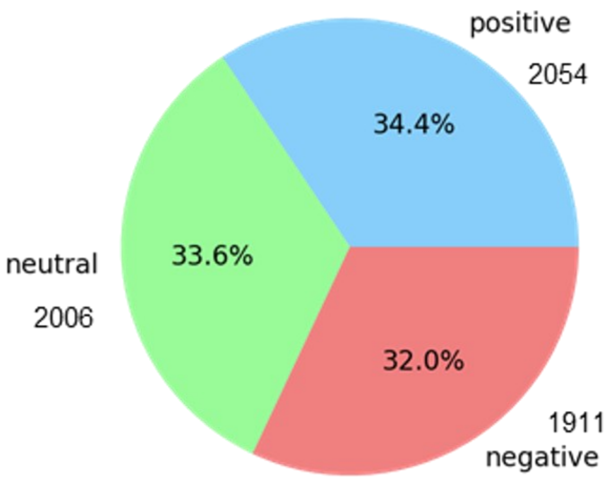


Figure 42 Disaster Sentiment of Japanese during Earthquake

For both typhoon and earthquake cases, tweets in English show inclination towards negative sentiment whereas tweets in Japanese display balanced to more positive sentiment. Negative sentiment is much more prominent in the case of the earthquake than the typhoon.

The system was able to demonstrate the data visualization and analysis subsystems through the demonstration in these two case studies.

2.2.3 Time Series Analysis

To gauge the intensity of the tweet's sentiment, the polarity (how much it aligns to a positive or negative sentiment) of each tweet is calculated using VADER [57] and pre-trained BERT algorithms for the English and Japanese datasets, respectively. Their polarities over time are plotted in the figures below.

Figures 8 and 9 shows time series (timestamp in UTC) plots in sentiment polarities of tweets on the day of the typhoon's landfall (around 9:00 UTC) while Figures 10 and 11 show the polarities on the day of the earthquake which struck around 14:00 UTC. During the earthquake, both English and Japanese tweets showed a decline from an intense period of tweeting activity in the first 2 h after, followed by a decrease over the following 4 h. In spite of the blind area in the Japanese tweet sentiment polarity index (i.e. from 0.5 to -0.5) dataset, it could still be seen that within 6 h after the earthquake, positive sentiment points are denser than negative sentiments. It could be that this decline was due to the exhaustion after such a disturbance, considering the local time which is past midnight.

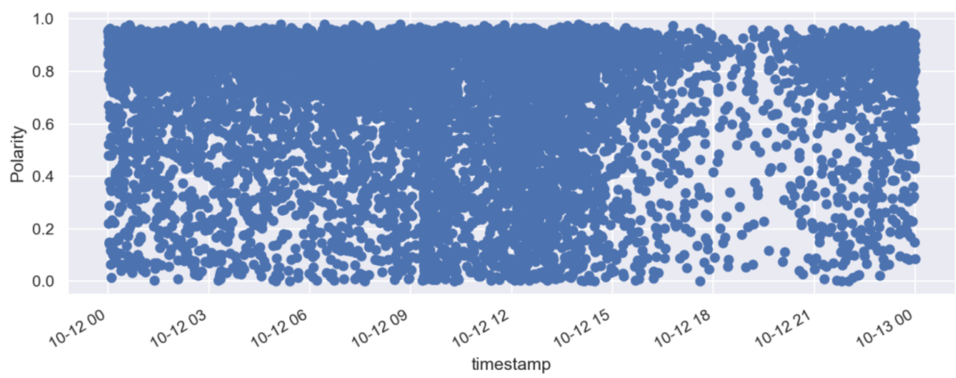


Figure 43 Sentiment Polarity during Typhoon (Japanese)

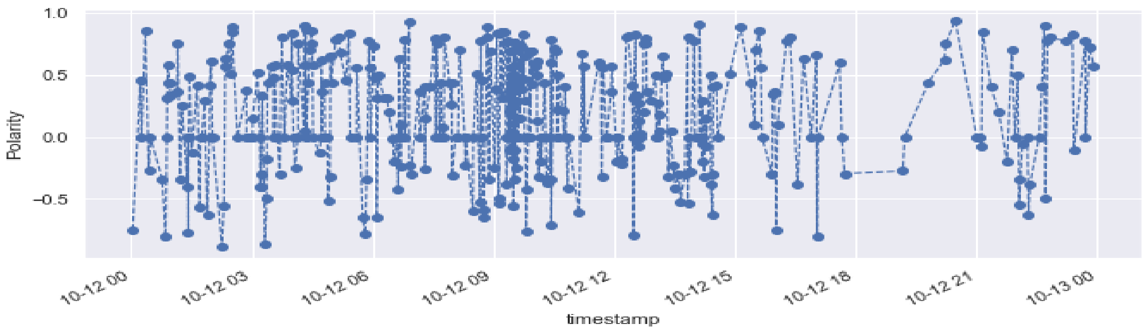


Figure 44 Sentiment Polarity during Typhoon (English)

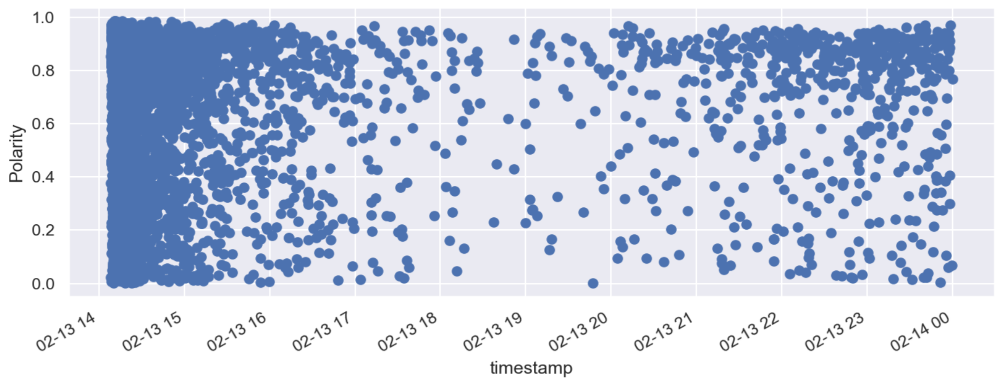


Figure 45 Sentiment Polarity during Earthquake (Japanese)

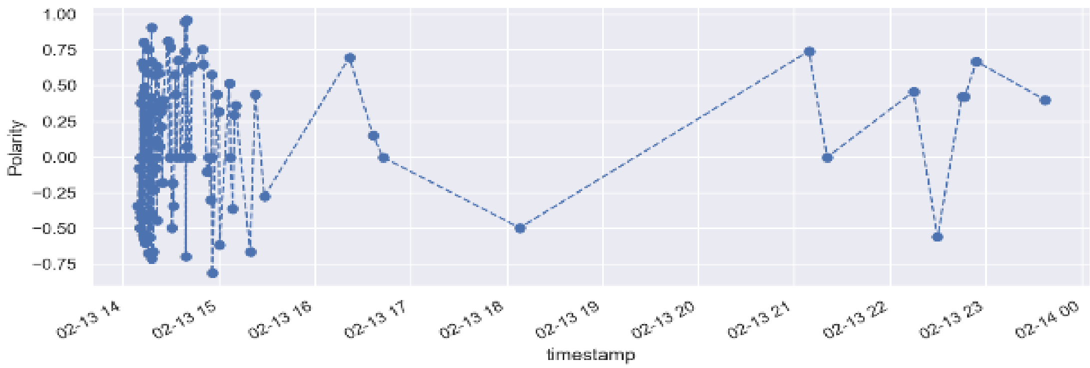


Figure 46 Sentiment Polarity during Earthquake (English)

The most common negative words (features) in the Japanese earthquake dataset are 揺れ and 揺れる, which means "shakes" as a noun or a verb respectively, this would not be surprising as it is common with all other earthquakes. However, words like 大きい (huge/large) and 長い (long) may indicate that this earthquake is unusually strong and long which is supported by the frequent appearance of words like 怖い (scary). Also, words like 来る or くる (coming) could indicate people's speculation or anticipation of a subsequent tsunami or aftershocks.

On the other hand, positive-sentiment words most common are 大丈夫 which could either mean "I'm okay" or "Are you okay?". To lesser extent, words such as 皆さん or 皆様 (everyone) also appear, for which the "okay" word may be extended not only to mark oneself safe after the earthquake or ask a familiar person if he or she is okay, but also to ask if (or even hope that) everyone is okay. Additionally, the appearance of 昨夜 (last night) and 昨日 (yesterday) may indicate relief from the event, that which occurred in the late night of 13 February 2021.

2.3 Algorithm performance

In order to build an optimal classifier, it is important to evaluate the model. Evaluation parameters used were precision, recall and F1 score. Precision indicates the proportion of predicted values that is truly positive and is calculated by (1).

$$Precision = \frac{True\ Positive}{True\ Positives + False\ Positives} \quad (1)$$

Recall shows the proportion of true positives that is correctly classified and is calculated using (2)

$$Recall = \frac{True\ Positive}{True\ Positives + False\ Negatives} \quad (2)$$

The F1 score is the weighted mean of precision and recall to capture how well it manages the tradeoff between precision and recall. It is calculated by using (3).

$$F1\ score = \frac{2\ x\ precision\ x\ recall}{precision + recall} \quad (3)$$

Table 18 shows the evaluation results for the SVM Classifier and XGBoost models for both the English and Japanese models for both types of disasters.

Table 18 Classifier Model Performance Evaluation

			Precision	Recall	F1-score	Accuracy
SVM	En	Ty	0.59	0.60	0.60	0.60
		Eq	0.64	0.75	0.68	0.75
	Ja	Ty	0.62	0.61	0.61	0.61
		Eq	0.56	0.56	0.56	0.56
XGB	En	Ty	0.54	0.55	0.50	0.55
		Eq	0.65	0.75	0.70	0.75
	Ja	Ty	0.56	0.62	0.58	0.62
		Eq	0.55	0.53	0.47	0.53

Overall, the models received an accuracy between 53 to 75%, which is relatively high considering the small dataset especially for English datasets. Optimal algorithms for each type show mixed results.

3. Validation

To validate the system, an interview with key stakeholders is conducted. Specifically, the interview is held with Mitsubishi Real Estate who has been working on developing more adaptive disaster management plans for large business districts in Tokyo wherein a diverse population, foreigners, and locals, populate the area for business travels, activities and so on.

DISASTER MANAGEMENT INSTITUTIONS

The intended user of the system output is the institution(s) who develop disaster management plans and risk reduction strategies in areas of diverse international population. Thus, to determine whether the system output could be used by the disaster management institution to make decisions on disaster management, the value of the system output to such user should be validated.

To conduct the validation, an interview is conducted online as shown in the photo below:

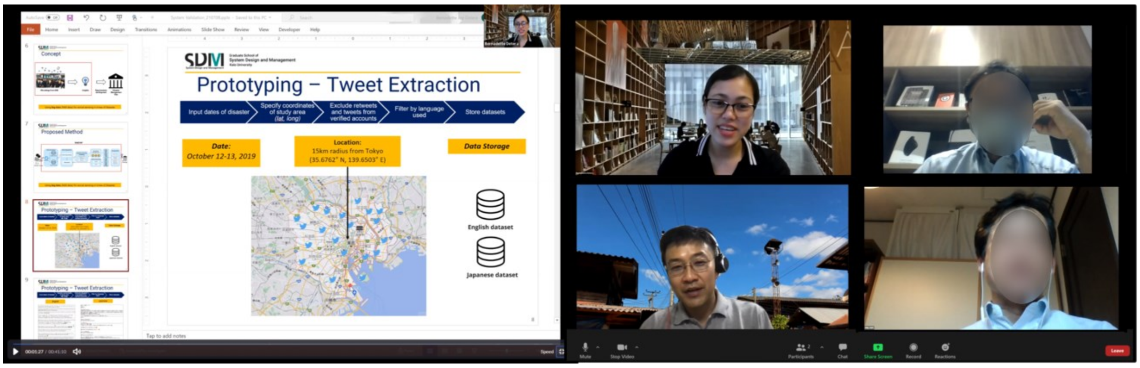


Figure 47 Interview with Stakeholders (Disaster Planning Institution)

Date: July 8, 2021

Interviewee:

Mitsubishi Estate Co., Ltd
Urban Planning Department
Development Promotion Division

Location: Online (Zoom)

Purpose: to receive validation from the perspective of disaster management specialists

Details: The interviewees were shown the concept of operations and the proposed methodology of analyzing disaster sentiment. Results from the system output were shown. Specifically, 1) result of sentiment score and intensity, 2) result of time series analysis and 3) recommendations on future disaster management plans based on 1 & 2.

By doing these, the system was validated if it can provide insights in understanding foreigners (based on their sentiment) which can be useful to improve existing disaster management strategies.

Key questions:

1. Is this approach useful for supporting disaster planning?
2. Does this approach provide insights on the needs of foreigners during disasters?

3. What kind of improvements are needed to make this more valid? Any other suggestion for improvement?

Results:

The interviewees expressed that this approach is useful for disaster planning and is quite interesting since they have never thought about using Twitter for disaster management. Focusing on SNS data may have false information but they expressed that this approach is interesting as it focuses on people's emotions. The extraction of tweets within a specific area (in this case, within 15 km radius from Tokyo) is very interesting and can be very useful in mapping the sentiment of people. Mr. A suggested, for example, that this could be used to map sentiment and the analysis or category within the hotel vicinity, public area and so on. A digital sentiment map with sentiment scores would be very interesting if implemented. The interviewees concur that Japanese may have the sense of disaster culture – living with disasters, and so may not feel much negative sentiment during disasters. They also stated that this could be broadcasted through NHK, Japan's public media organization, in their disaster dashboard by having a map showing the category and level of sentiment.

The interviewees agree that this generated useful insights about the needs of foreigners. They stated that they gained more knowledge that will lead to the next stage of disaster prevention that would be needed to incorporate the needs of foreigners. They also commented that time series analysis is insightful as it gave insights on the timing and how people's feelings changed over time. For the choice of study area, it is recommended to expand the geographic location for analysis of earthquake scenario (more than 15 km) from Tokyo. Also, they said this approach is helpful as they are also thinking about providing community support to foreigners. Furthermore, they mentioned that this could be further zoomed in to specific nationalities. For example, Chinese people also experience earthquakes in their own country so their sentiment may be more neutral or positive (calm). This could be useful when targeting certain social groups like non-Japanese and non-English native speakers.

FOREIGNERS

The system proposes a method to have a deeper understanding of how foreigners react, and experience disasters differently in terms of sentiment compared to the local population in order for disaster planning institutions to provide them more adequate support in future disaster scenarios. To achieve this, we calculated and analyzed people's sentiment in Tokyo during recent disasters and was able to get results and insights. To validate the results of the analysis, interviews with foreign residents living in Tokyo were conducted. The interview questions are shown in the Appendix.

The interviews were conducted online via Zoom meetings as shown in the photo below:

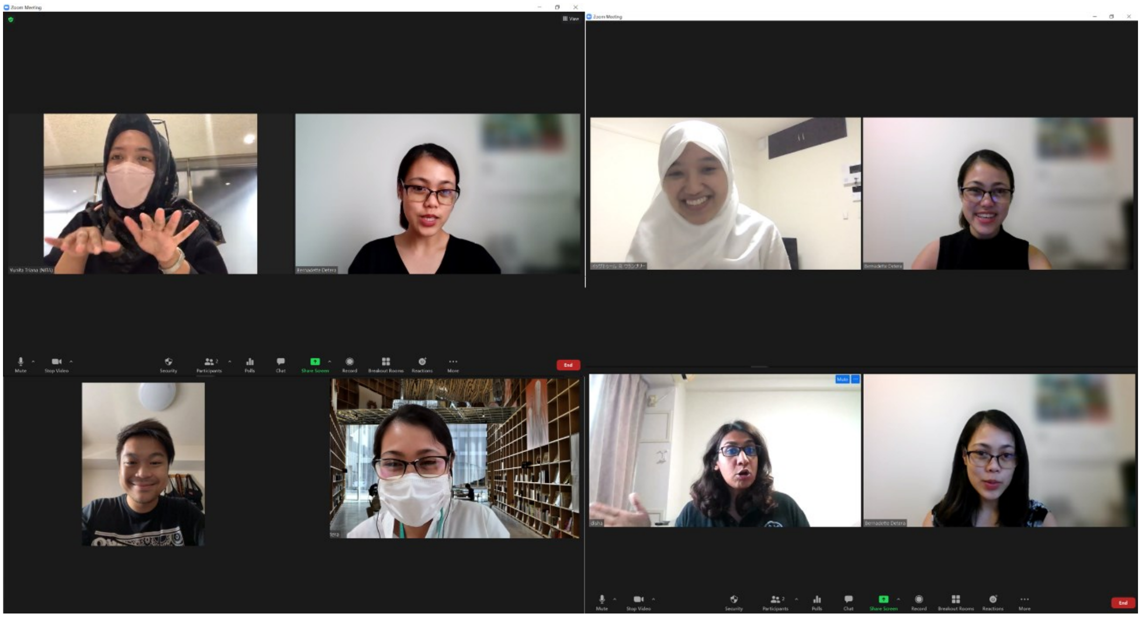


Figure 48 Interview with Stakeholders (Foreigners living in Japan)

Results of the validation interview are shown below:

Date(s):

July 1, 2021/ July 2, 2021/ July 5, 2021

Interviewee: Foreigners living in Tokyo

Location: Online (Zoom)

Purpose: to validate the sentiment results that the system generated based on the model

Details: The interviewees targeted are foreigners living in Tokyo, whether for work or study, who have experienced the case studies used in the study, Typhoon Hagibis and the 2021 Fukushima Earthquake. The interviewees shared how they behaved during the said disasters and how they felt. Their sentiments generally agree with the systems output – feeling mixed negative emotions of nervousness, unease, distress, panic and worry. However, one interviewee demonstrated very limited negative sentiment and generally considered herself to be calm or in a neutral sentiment during disasters. This may be attributed to her strong Japanese language ability and existing knowledge on disaster preparedness. The details of the interview results are shown below. Do note that interviewees' personal information such as their names and ages are not disclosed to protect their individual identities.

Table 19 Validation Interview with Foreigners

Interviewee	**Date**	**Details**
A	July 1, 2021	• Nationality: Indonesia • PhD Student • Living in Japan for 2.5 years • Basic/ limited Japanese language proficiency (JLPT N4 level) • No experience with severe cases of typhoon nor earthquake • Prior coming to Japan, has limited knowledge on disaster preparedness. • Mainly relied on her professor during Typhoon Hagibis for information on what to do and how to prepare for the landfall. • Stated that main issue was everything is in Japanese kanji so it was hard to understand early warnings. • Relied on Google Translate camera function to translate screenshots of

Interviewee	Date	Details

		Japanese text about the warning. However, this takes a lot of time.
		• Strongly agree that she experienced strongly negative emotions and panicked during disaster.
		• Strongly agree that earthquake scenario was much scarier than typhoons.
		• Strongly agree that the most crucial period is preparation for disasters, right before the onset
		• Suggested to receive *only* the following information from disaster management (mobile) app or other platforms:
		  ○ 1st: message saying "Don't Panic. Stay Calm."
		  ○ 2nd: information about the disaster
		  ○ 3rd: Where I need to go (evacuation area with map of local area) or what I need to do

B	July 2, 2021	• Nationality: Indonesia • Working at a Japanese company • Living in Japan for 5 years • Proficient in Japanese language (JLPT N2) • No experience with severe cases of typhoon nor earthquake prior to living in Japan. • Participated in several disaster preparedness interventions in Japan such as visiting Disaster Prevention Center in Yokohama

Interviewee	Date	Details
		• Did not experience panic during Typhoon Hagibis nor face significant difficulty. • Got very used to disasters here in Japan. • Relied on Twitter and on the internet very much for disaster information; especially internet personalities whose Twitter accounts can be trusted. • Strongly agree that Japanese people react calmly and are not so much bothered by disasters. • Agree that earthquake scenario was much scarier than typhoons. • Stated that the hardest challenge for foreigners is the expectation on how it feels to *actually* experience the disaster. Stated that it is hard for those who have zero prior experience with disasters. • Strongly agree that there should be more disaster support for foreigners in English.
C	July 5, 2021	• Nationality: Philippines • Working at a Japanese company • Living in Japan for 2 years • Intermediate Japanese language proficiency (JLPT N3) • Has one experience with a strong earthquake prior to living in Japan. • No disaster preparedness activity or intervention or prevention effort

Interviewee	Date	Details
		• Agree that he experience negative sentiment especially during the Fukushima earthquake. • Experienced a sense of confusion, worry and a bit of panic; stated that he suddenly opened his television which he never used until that point, to know and confirm what is happening. • Strongly agree that Japanese people react calmly. • Agree that earthquake scenario was much scarier than typhoons. • Stated that the existing disaster warnings are not effective especially because they are in Japanese. Stated that for people who cannot read it, those warnings would not induce the 'necessary panic' or alertness that you should feel when there is a disaster. • Strongly agree that there should be more disaster support for foreigners in English.

V. DISCUSSION

1. Tweeting Behavior during Disasters

Overall, we observed that the number of tweets during typhoons is much higher than that of earthquakes as seen in Table 1. Also, the number of tweets in English is much less than Japanese as expected because generally, only non-Japanese people tweet in English. The English tweets comprise about 2% and 7% of tweets during the earthquake and typhoon, respectively. Additionally, other languages such as Chinese, Vietnamese, Hindi and Filipino were detected from Twitter data during the disasters. Although they were not considered in this study, it is important to note the presence of these languages for future research. The number of data points from non-English and non-Japanese tweets is very low in the study area (i.e., Tokyo). However, the fact that other microblogs in other languages were detected engenders the need to conduct expansion of this analysis to other languages. The objective of this thesis is geared towards fostering understanding of the issues faced by foreigners at large during disasters. However, there is great potential in intensifying the research for the sentiment analysis of foreigners from certain nationalities especially if foreigners in an area are dominated by one or two nationalities.

2. Classification Model Evaluation

For classification of English typhoon tweets, SVM showed higher accuracy while XGBoost is better for the earthquake dataset. On the other hand, classification of Japanese typhoon dataset can achieve higher accuracy with XGBoost while SVM for the earthquake dataset. In general, larger datasets lead to higher accuracy and this is consistent with all datasets except for the case of English earthquake tweets. Instead, it achieved a much higher accuracy.

This may be due to the extremely low number of data points and skewed distribution to negative sentiment. While there is no stark difference between using the two algorithms, this reveals the great potential of using the relatively new XGBoost algorithm in addition to or instead of the traditional SVM algorithm.

3. Concerns during Disasters

For the negative earthquake sentiment, the word ‘phone’ and ‘alarm’ are one of the most frequently used. Interestingly, the tweets air concerns that the earthquake alarms on their mobile phones did not work.

For the Japanese tweets, the Word Cloud is exhibited in Figure 39 and 40 for the positive and negative sentiments, respectively. We observe that for three scenarios, the word ‘大丈夫’ which means okay or alright is present. For Japanese, it was observed that Twitter was widely used to check and update safety statuses of friends and family. Understandably, for foreigners this might not be the case since their families are living overseas.

The word ‘揺れる’ which means ‘to shake’ is a frequent word in all scenarios even during the typhoon. These tweets exhibited their sentiment on their houses being shaken, not just in earthquakes but during typhoons as well. Moreover, an earthquake also striked near Tokyo during the typhoon Hagibis and showed people’s sentiment on having two natural disasters at the same time. This insight reveals the need for multi-hazard analysis of sentiment and people’s needs during disasters, too. One of this research’s objectives are to gain insights specific to the type of disaster (i.e., typhoon and earthquake). The results showed some people were concerned of having another disaster (of different type). Disaster behavior during multi-hazard disaster events is also an interesting insight. This is especially observed in Japanese tweets which suggests their strong awareness of different types of disaster and how disasters relate to each other. For example, during the earthquake, a lot of people are already mentioning tsunami, showing their knowledge that a tsunami could possibly occur after earthquakes. Foreigners, on the other hand, may not have the same level of understanding of disasters and their interrelations because of their lack of experience with natural disasters. This finding is also confirmed during the interview with foreigners living in Tokyo who experienced the two case studies (Typhoon Hagibis and the 2021 Fukushima Earthquake). Detailed results of the interview with foreigners can be seen in the Appendix.

There is still a challenge in having deep understanding of the needs of foreigners during disasters. The frequent words and Word Cloud revealed that foreigners' blogs are much more diverse than Japanese whose tweets contained very similar words implying a more homogenous and similar concerns with each other. To better understand foreigners' tweets, it could be further improved by utilizing other techniques such as topic modelling algorithms.

4. Sentiment Analysis

It is interesting to note that despite the current matter, i.e., Typhoon Hagibis, tweets regarding earthquake aftershocks and intensity, as well as the 2011 Great East Japan Earthquake, still appeared in both positive- and negative-sentiment tweets. It could be that such an extreme event, i.e., one of the worst typhoons to hit Japan, may have conjured memories of similar feeling (sentiment) and magnitude. Alternatively, such an event may have had incited speculation about other disasters that could come.

For the typhoon dataset (both Japanese and English) it can be seen that tweet count increased at the start of landfall. A decline in tweets was observed after the peak of the event. Three hours after, tweet count increase could be seen again in the Japanese dataset but not in the English dataset. This feature may be inconclusive since the English typhoon dataset is way scarcer than the Japanese dataset. A typhoon passes through town or city in around 4 h. This could be that people are more focused on the event, which is longer in duration than the earthquake, and do not use twitter. Some of the activities could include securing everyone's own houses and everyone in it, evacuation. Alternatively, people might have been conserving resources such as power or could be using their phones for purposes other than tweeting, e.g., waiting for announcements, disaster information, other platforms of communication, seeking assistance, etc.

These results agree with that of Lu and Brelsford [58] who found out that Japanese user tweet topics focused on the disaster at hand more than non-Japanese user tweets. Their study also mentions that both Japanese and non-Japanese users expanded their social community (i.e., virtual), but remained there. They did not shift to another virtual community in social media. This could support the top Japanese tweets of “okay” and “everyone” where Japanese-speakers ask members of their virtual social media community, even those who are in their extended network. The authors’ analysis was done on Twitter and Topsy social media data shortly before, during, and after the 2011 Great East Japan Earthquake.

Sentiments on typhoons are more varied for both English and Japanese tweets, but rather skewed to the negative in both languages during earthquakes. It could be because (1) strong typhoons are more frequent than earthquakes of comparable magnitudes. Earthquake-related tweets are relatively fewer than typhoon which could be because people have time to tweet about a typhoon as it approaches, in contrast to an earthquake whose occurrence is unpredictable.

5. Applications to Disaster Planning

Type of support and content in disaster communication

Perhaps the most useful finding from this research is the revelation that foreigners experience negative sentiment during disasters significantly more intensely than the local Japanese people. This is also validated through interviews conducted to foreigners in Tokyo. This stark difference in disaster sentiment between foreigners and Japanese was observed in both typhoon and earthquake scenarios, with the case of earthquake having more negative sentiment for both groups. One application of this is disaster planning is in disaster communication such as early warning systems. It is recommended for disaster management institutions or the local government to include the provision of psychosocial support during disasters in addition to providing basic information about the disaster. Including support on such areas (anxiety, stress, etc., during disasters) could reduce post-disaster stress for foreigners and reduce the impact of disasters on them. When delivering early warnings to foreigners, including a message on this point, could be helpful. It the interview with foreigners, one person suggested to include messages like “Do not panic. Please stay calm.” To early warning messages on mobile phones would very helpful.

Timing of disaster communication

Another important finding from this research is the observation that tweets of foreigners are most condensed before the typhoon landfall as compared to Japanese whose tweets gravitated during the landfall. This indicates the level of vulnerability or difficulty foreigners are facing is strongest pre-disaster. The implication of this is the need for more pre-disaster support for foreigners. This is aligned with the findings from previous studies of Kawasaki et al [16] where it was observed that foreigners faced lack of reliable information, and faced conflicting news during the 2011 Great East Japan Earthquake. Although early warning messages are in place in Japan, most of them are only communicated in Japan. All foreigner interviewees (see Appendix) also stated that they experienced a lot of difficulty in understanding disaster messages during disasters here in Japan because they are only in Japanese. One interviewee mentioned that she had to rely on Twitter to ask other public what does one disaster message on her phone means.

One recommendation for disaster planning based on the findings of this research is to increase efforts in communicating relevant and necessary information about the disaster before the onset of disaster (pre-disaster). Having more (in number) of more frequent messages before disaster as part of preparedness could be helpful. Specifically, as an example, disaster communication and support could be prioritized to foreigners a few hours (3-4 hours) before the typhoon landfall. It is recommended to reconsider timing or volume of disaster messages.

6. Limitations

6.1 Data Availability

The system relies heavily on the available data from Social Networking Sites (SNS). Although this proposed system could function for any of the open-data sourced SNS data provider, the study only utilized Twitter at this stage. More importantly, one limitation is that Twitter data availability only starts from the year 2006 since the advent of popularity of SNS such as Facebook in that decade. Because of this, historical data from previous disasters before 2006 cannot be accessed.

In addition, one important limitation is the volume of tweets of foreigners during disaster scenarios. Because of the nature of this study is to compare foreigners with locals, the volume of data available greatly depends on the percentage of foreigners in the total population. Moreover, this is also affected by the market adoption rate or utilization rate of Twitter in the country or area of study. Since this research involved a very granular approach, specific only to Tokyo area and case studies involved only the onset (48 hours) of the disaster, raw data extracted for English tweets was very limited on the case of earthquakes.

6.2 Historical Data on Disasters

The system also relies heavily on the occurrence of natural disasters. The frequency of natural disasters occurring vary from country to country. Since the system requires historical data for the type of disaster event that the User requests to the system, if there is little to no occurrence of that type of disaster within the data availability period (e.g., in the case of Twitter, from 2006 to present), the system would not be able to generate outputs.

6.3 Personal Characteristics of Foreigners

Although this study was able to extract data coming from foreigners using SNS data, the data collected is limited to treating all data points under one category of foreigners, regardless of factors that may influence how they behave and react to natural disasters. Due to data privacy restrictions of the publicly available Twitter data as well as data from other SNS, personal characteristics such as age, nationality and gender of the users cannot be determined nor investigated. In addition, the data is also limited to the market penetration of the Social Networking Service in the area of interest. Twitter is widely used all over the world. However, the demographics of its users (e.g., young adults, elderly, etc.) can greatly affect the analysis.

6.4 Algorithm Accuracy

The system utilizes supervised machine learning to develop the classifier models. Using this technique requires training data which you must label and feed into the system. In this thesis, the training datasets are manually labelled and are based on the author's judgement following the sentiment definition presented in Table 14. Since the author is not a native speaker of Japanese and only has an intermediate Japanese language proficiency, some tweets may be misunderstood. This could affect the accuracy of the training dataset which greatly affects the performance of the algorithm developed for Japanese datasets. Moreover, since the nature of disaster experience is generally negative, tweets or any microblog from other SNS sources would be understandably more inclined towards negative sentiment. This could play a great role in the model accuracy and could lead to overfitting or underfitting of the algorithms.

6.5 Prototyping on Other Disaster Events

The system was able to successfully demonstrate its application in the case of two types of disasters: typhoon and earthquake. Although insights specific to type of disaster were generated by the system, it might provide more useful insights if the system is run and iterated through other instances of typhoon and earthquake scenarios. This research only utilized one case of typhoon and one case of earthquake. However, this is again subject to the limitation of the occurrence of that type of disaster in your study area which is mentioned in Section 5.5.2. Additionally, other types of disaster (e.g., forest fires, flooding, landslide) could be covered by the system.

7. Future Work

There are still gaps and opportunities for future research on this topic. As mentioned above, one limitation is the data availability. Future research could utilize other SNS sources such as Sino Weibo to capture insights specific to certain demographics or nationalities. Moreover, in the discussion on frequent words and Word Cloud analysis on the case of earthquake, it was observed that many people were also concerned about the occurrence of tsunami after the earthquake. This raises an important aspect in Disaster Research of looking at multi-hazard perspective.

In terms of developing the sentiment classifier model, other machine learning techniques still need to be explored when working on Natural Language Processing (NLP) on multiple languages. The increasing scholarship on sentiment analysis algorithms on languages other than English is also expected to accelerate research on this topic. An example is the use of multi-class BERT algorithm in Swahili language [59].

Finally, research on other information or type of disaster support besides basic disaster information and psychosocial support that foreigners need in times of disaster is very important. Aspects such as topic modelling which could compliment research on disaster sentiment is also recommended.

8. System Implementation and Value

If implemented, the system could generate insights on foreigners' social vulnerability specifically on their concerns, sentiment, and intensity of these sentiments during natural disasters. If realized, this system may inform policy makers and disaster planners with better requirements as inputs in designing disaster management plans. To highlight the value of the system upon implementation, a Customer Value Chain Analysis (CVCA) diagram is presented in Figure 49.

In terms of implementing the findings from this research, there are several avenues where such recommendations can be integrated. One example is through the Tokyo Metropolitan Government's Urban Regeneration Safety Assurance Plan of the Otemachi-Marunouchi-Yurakucho (OMY) District which caters to approximately 330,000 residents in the area, 32,000 business visitors and 33,000 general visitors [60]. Although the exact volume of foreign residents and visitors within the area cannot be determined accurately, the OMY District is no stranger to international visitors as many international meetings and events occur in these business districts. The latest safety assurance plan involves creating an information gathering and dissemination HUB which will be implemented during disasters in the OMY District to achieve situational awareness (e.g., current damage to facilities, etc.) Through this HUB, information designed to be conveyed to foreigners in the area could be included; not only the provision of information in English but also the inclusion of psychosocial messages such as “Do not panic” and other messages to provide assurance to foreigners who may not know what to expect and has never experienced a natural disaster before. In addition, some wards (local cities) in Tokyo offer Disaster Response Manual in English through their online websites. An example of which is the manual [61] from Chiyoda Ward at the center of Tokyo. Although the manual already included “Don’t panic and rush home”, more messages to strengthen foreigner’s assurance and reduce disaster anxiety can be included. Most importantly, it is recommended to prioritize disaster risk communication prior to the disaster happening, rather than disaster response after the disaster. For the findings of this research to be of more value, another potential avenue is through integrating more information and support targeted for foreigners using the city’s public information dissemination and early warning systems. It is highly recommended to broadcast foreigner-targeted information (and psychosocial messages) more frequently especially before the onset of the disaster to aid foreigners become more prepared and able to appropriately react to disasters.

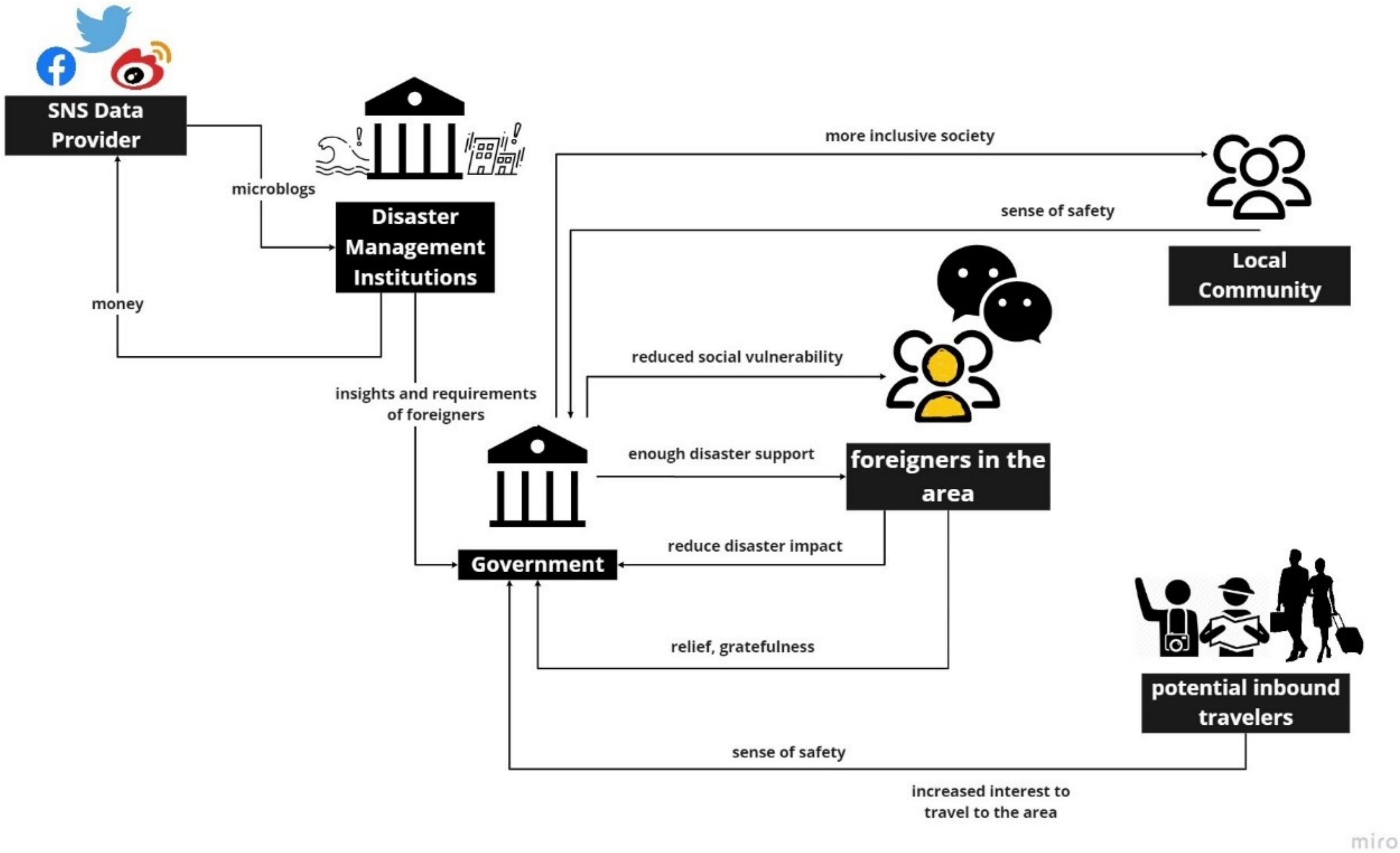


Figure 49 Customer Value Chain Analysis (CVCA) Diagram

VI. CONCLUSION

Foreigners are more socially vulnerable to disasters because they experience emergencies and natural disasters differently than that of the local population. To help institutions and disaster planners develop more effective disaster planning strategies to account the needs of foreigners, this study designed a Disaster Sentiment Analyzer system based on microblogs from Social Networking Services (SNS). To analyze how people react to disasters from the perspective of multilingual population, we looked at two case studies in Tokyo where there is a considerable number of international populations who experience the same natural disasters with the locals. The cases of Typhoon Hagibis in 2019 and the 2021 Fukushima Earthquake were used for typhoon and earthquake scenarios, respectively. We used Twitter data to extract English and Japanese tweets generated in central Tokyo during the onset of disaster. On a 48-hour period during the onset of disaster, a total of 13, 298 tweets were collected during the typhoon and 6,172 tweets during the earthquake. To understand the sentiment (i.e., positive, neutral, negative), we developed a machine learning Support Vector Machine and Extreme Gradient Boosting (XGBoost) model of fair but promising accuracy for both English and Japanese languages. The model was trained specifically on tweets during disasters and could yield a more contextual approach to other disasters as well.

Differences in English and Japanese sentiment were revealed (foreigner vs. locals). Both groups face more anxiety and negative sentiment during earthquakes than typhoons. This could be due to the available time in preparing for these disasters: Days for typhoons and none at all for earthquakes. Still, Japanese locals in the study area showed a significantly more positive attitude and sentiment during disasters. The study showed foreigners experienced significantly more negative sentiment, anxiety, and stress than Japanese. It was recommended to include psychosocial support in addition to early warnings when communicating disasters to foreigners.

The study also revealed insights on how sentiment changed over time. It was observed that English tweets concentrated most a few hours before the typhoon's landfall, indicating the most critical point in the vulnerability of foreigners is right before the disaster's onset. It was recommended to reconsider timing of provision of disaster information and provide them earlier than usual during pre-disaster information dissemination.

This method is validated and can be useful in designing better disaster management strategies especially for cities with various social groups and foreigners. For future work, the models used here could be enhanced by using larger datasets and be designed to collect near-real time Twitter data and generate sentiment analysis results immediately. In addition, languages besides English which are used by foreigners could be explored further as well as other sources of SNS data. Furthermore, analysis of cases of multi-hazard disasters could be realized.

REFERENCES

[1] "Natural Disasters - Our World in Data." https://ourworldindata.org/natural-disasters (accessed Jun. 27, 2021).

[2] N. H. Center, "Costliest U.S. tropical cyclones tables updated." https://www.nhc.noaa.gov/news/UpdatedCostliest.pdf (accessed Jun. 27, 2021).

[3] NOAA, "March 11, 2011 Japan Earthquake and Tsunami," 2015. Accessed: Jun. 27, 2021. [Online]. Available: http://www.ngdc.noaa.gov/hazard/honshu_11mar2011.shtml.

[4] L. Guadagno, "Integrating migrants in emergency preparedness , in their host countries," 2017. doi: 10.13140/RG.2.2.35968.02562.

[5] S. Manzuik, A. Gold, and C. Gatford, *Social Vulnerability Assessment Tools for Climate Change and DRR Programming*. United Nations Development Programme, 2017.

[6] D. A. Savage and B. Torgler, "Methods and Insights on How to Explore Human Behavior in the Disaster Environment," in *Economic Effects of Natural Disasters*, Elsevier, 2021, pp. 191–209.

[7] "Integrating migrants in emergency preparedness, response and recovery in their host countries Reference handbook," 2016. Accessed: Oct. 05, 2020. [Online]. Available: www.iom.int.

[8] "United Nations Population Division | Department of Economic and Social Affairs." https://www.un.org/en/development/desa/population/migration/data/estimates2/estimatesmaps.asp?0t0 (accessed Jun. 29, 2021).

[9] "Statistics Bureau Home Page/JAPAN STATISTICAL YEARBOOK 2019 - Chapter 2 Population and Households." https://www.stat.go.jp/english/data/nenkan/68nenkan/1431-02.html (accessed Jun. 29, 2021).

[10] U. Nations, "Vienna Convention on Consular Relations, 1963."

[11] U. Nations, "Hyogo framework for action 2005–2015," 2013. doi: 10.1007/978-1-4020-4399-4_180.

[12] "Sendai Framework for Disaster Risk Reduction 2015-2030 | UNDRR." https://www.undrr.org/publication/sendai-framework-disaster-risk-reduction-2015-2030 (accessed Sep. 05, 2020).

[13] L. Guadagno, "Reducing Migrants' Vulnerability to Natural Disasters through Disaster Risk Reduction Measures," no. October 2015, 2015.

[14] L. Robles, "Linkages in disasters: Perspective on a migrant inclusive disaster risk reduction and recovery," 2016, doi: 10.13140/RG.2.2.33723.41760.

[15] A. Kawasaki, M. Henry, and K. Meguro, "Media Preference, Information Needs, and the Language Proficiency of Foreigners in Japan after the 2011 Great East

Japan Earthquake," *Int. J. Disaster Risk Sci.*, vol. 9, no. 1, pp. 1–15, Mar. 2018, doi: 10.1007/s13753-018-0159-8.

[16] M. Henry, A. KAWASAKI, and K. MEGURO, "Foreigners' disaster information gathering behavior after the 2011 Tohoku Earthquake part 3: analysis of foreign students considering theirpost-disaster action," *SEISAN KENKYU*, vol. 64, no. 4, pp. 497–503, Jul. 2012, doi: 10.11188/seisankenkyu.64.497.

[17] S. Uekusa and S. Matthewman, "Vulnerable and resilient? Immigrants and refugees in the 2010–2011 Canterbury and Tohoku disasters," *Int. J. Disaster Risk Reduct.*, vol. 22, pp. 355–361, Jun. 2017, doi: 10.1016/j.ijdrr.2017.02.006.

[18] "Disaster Message Not Getting Through to Foreign Residents | Nippon.com." https://www.nippon.com/en/news/fnn20190930001/disaster-message-not-getting-through-to-foreign-residents.html (accessed Jun. 30, 2021).

[19] "• Twitter: most users by country | Statista." https://www.statista.com/statistics/242606/number-of-active-twitter-users-in-selected-countries/ (accessed Jul. 05, 2021).

[20] A. Kawasaki, K. Meguro, and M. Henry, "Comparing the disaster information gathering behavior and post-disaster actions of Japanese and foreigners in the Kanto area after the 2011 Tohoku Earthquake."

[21] B. Jongman, J. Wagemaker, B. Revilla Romero, and E. Coughlan De Perez, "Early flood detection for rapid humanitarian response: Harnessing near real-time satellite and twitter signals," *ISPRS Int. J. Geo-Information*, vol. 4, no. 4, pp. 2246–2266, Dec. 2015, doi: 10.3390/ijgi4042246.

[22] D. Grandoni *et al.*, "Space-based Technology for Emergency Management: The COSMO-SkyMed Constellation Contribution," *Procedia Technol.*, vol. 16, pp. 858–866, 2014, doi: 10.1016/j.protcy.2014.10.036.

[23] M. Basu, A. Shandilya, P. Khosla, K. Ghosh, and S. Ghosh, "Extracting Resource Needs and Availabilities from Microblogs for Aiding Post-Disaster Relief Operations," in *IEEE Transactions on Computational Social Systems*, Jun. 2019, vol. 6, no. 3, pp. 604–618, doi: 10.1109/TCSS.2019.2914179.

[24] M. F. Shah and O. Murao, "Foreigners' Experience and Behavior in the Great East Japan Disasters 2011," *10th Int. Conf. Urban Earthq. Eng.*, no. August 2016, 2013, [Online]. Available: https://www.researchgate.net/publication/306094610_FOREIGNERS'_EXPERIENCE_AND_BEHAVIOR_IN_THE_GREAT_EAST_JAPAN_DISASTERS_2011.

[25] S. Doan, B. K. H. Vo, and N. Collier, "An analysis of twitter messages in the 2011 Tohoku Earthquake," in *Lecture Notes of the Institute for Computer Sciences, Social-Informatics and Telecommunications Engineering*, 2012, vol. 91 LNICST, pp. 58–66, doi: 10.1007/978-3-642-29262-0_8.

[26] R. Dasgupta *et al.*, "A rapid indicator-based assessment of foreign resident preparedness in Japan during Typhoon Hagibis," *Int. J. Disaster Risk Reduct.*, vol. 51, p. 101849, Dec. 2020, doi: 10.1016/j.ijdrr.2020.101849.

[27] S. Valenzuela, S. Puente, and P. M. Flores, "Comparing Disaster News on Twitter

and Television: an Intermedia Agenda Setting Perspective," *J. Broadcast. Electron. Media*, vol. 61, no. 4, pp. 615–637, Oct. 2017, doi: 10.1080/08838151.2017.1344673.

[28] T. Hossmann, P. Carta, D. Schatzmann, F. Legendre, P. Gunningberg, and C. Rohner, "Twitter in disaster mode: Security architecture," *Proc. Spec. Work. Internet Disasters, SWID'11*, 2011, doi: 10.1145/2079360.2079367.

[29] B. Takahashi, E. C. Tandoc, and C. Carmichael, "Communicating on Twitter during a disaster: An analysis of tweets during Typhoon Haiyan in the Philippines," *Comput. Human Behav.*, vol. 50, pp. 392–398, 2015, doi: 10.1016/j.chb.2015.04.020.

[30] H. Kwak, C. Lee, H. Park, and S. Moon, "What is Twitter, a social network or a news media?," in *Proceedings of the 19th International Conference on World Wide Web, WWW '10*, 2010, pp. 591–600, doi: 10.1145/1772690.1772751.

[31] J. K. Joseph, K. A. Dev, A. P. Pradeepkumar, and M. Mohan, "Big data analytics and social media in disaster management," in *Integrating Disaster Science and Management: Global Case Studies in Mitigation and Recovery*, no. May, Elsevier Inc., 2018, pp. 287–294.

[32] G. P. Cooper, V. Yeager, F. M. Burkle, and I. Subbarao, "Twitter as a potential disaster risk reduction tool. part i: Introduction, terminology, research and operational applications," *PLoS Curr.*, vol. 7, no. DISASTERS, Jun. 2015, doi: 10.1371/currents.dis.a7657429d6f25f02bb5253e551015f0f.

[33] J. R. Ragini, P. M. R. Anand, and V. Bhaskar, "Big data analytics for disaster response and recovery through sentiment analysis," *Int. J. Inf. Manage.*, vol. 42, no. May, pp. 13–24, 2018, doi: 10.1016/j.ijinfomgt.2018.05.004.

[34] M. Singh, A. K. Jakhar, and S. Pandey, "Sentiment analysis on the impact of coronavirus in social life using the BERT model," *Soc. Netw. Anal. Min.*, vol. 11, no. 1, pp. 1–11, 2021, doi: 10.1007/s13278-021-00737-z.

[35] P. Khosla, M. Basu, K. Ghosh, and S. Ghosh, "Automatic Matching of Resource Needs and Availabilities in Microblogs for Post-Disaster Relief," *arXiv*, pp. 25–26, 2018.

[36] A. Squicciarini, A. Tapia, and S. Stehle, "Sentiment analysis during Hurricane Sandy in emergency response," *Int. J. Disaster Risk Reduct.*, vol. 21, no. May 2016, pp. 213–222, 2017, doi: 10.1016/j.ijdrr.2016.12.011.

[37] S. Doan, B. K. H. Vo, and N. Collier, "An analysis of twitter messages in the 2011 Tohoku Earthquake," *Lect. Notes Inst. Comput. Sci. Soc. Telecommun. Eng.*, vol. 91 LNICST, no. March 2011, pp. 58–66, 2012, doi: 10.1007/978-3-642-29262-0_8.

[38] F. N. Ribeiro, M. Araújo, P. Gonçalves, M. André Gonçalves, and F. Benevenuto, "SentiBench - a benchmark comparison of state-of-the-practice sentiment analysis methods," *EPJ Data Sci.*, vol. 5, no. 1, Dec. 2016, doi: 10.1140/epjds/s13688-016-0085-1.

[39] W. Zhang, S. xi Kong, Y. chun Zhu, and X. le Wang, "Sentiment classification and computing for online reviews by a hybrid SVM and LSA based approach," *Cluster*

Comput., vol. 22, pp. 12619–12632, Sep. 2019, doi: 10.1007/s10586-017-1693-7.

[40] Q. Ye, Z. Zhang, and R. Law, "Sentiment classification of online reviews to travel destinations by supervised machine learning approaches," *Expert Syst. Appl.*, vol. 36, no. 3 PART 2, pp. 6527–6535, Apr. 2009, doi: 10.1016/j.eswa.2008.07.035.

[41] S. Guo, S. Chen, and Y. Li, "Face recognition based on convolutional neural network & support vector machine," in *2016 IEEE International Conference on Information and Automation, IEEE ICIA 2016*, Jan. 2017, pp. 1787–1792, doi: 10.1109/ICInfA.2016.7832107.

[42] Y. Al Amrani, M. Lazaar, and K. E. El Kadirp, "Random forest and support vector machine based hybrid approach to sentiment analysis," in *Procedia Computer Science*, Jan. 2018, vol. 127, pp. 511–520, doi: 10.1016/j.procs.2018.01.150.

[43] M. A. Faruque, S. Rahman, P. Chakraborty, T. Choudhury, J. S. Um, and T. P. Singh, "Ascertaining polarity of public opinions on Bangladesh cricket using machine learning techniques," *Spat. Inf. Res.*, 2021, doi: 10.1007/s41324-021-00403-8.

[44] "Greedy Function Approximation: A Gradient Boosting Machine on JSTOR." https://www.jstor.org/stable/2699986?origin=JSTOR-pdf&seq=1#metadata_info_tab_contents (accessed Jul. 06, 2021).

[45] T. Chen and C. Guestrin, "XGBoost: A scalable tree boosting system," in *Proceedings of the ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, Aug. 2016, vol. 13-17-August-2016, pp. 785–794, doi: 10.1145/2939672.2939785.

[46] J. Vongkusolkit and Q. Huang, "Situational awareness extraction: a comprehensive review of social media data classification during natural hazards," *Ann. GIS*, pp. 1–24, Oct. 2020, doi: 10.1080/19475683.2020.1817146.

[47] J. Vongkusolkit and Q. Huang, "Situational awareness extraction: a comprehensive review of social media data classification during natural hazards," *Ann. GIS*, pp. 1–24, Oct. 2020, doi: 10.1080/19475683.2020.1817146.

[48] H. Seddighi, I. Salmani, and S. Seddighi, "Saving Lives and Changing Minds with Twitter in Disasters and Pandemics: A Literature Review," *Journal. Media*, vol. 1, no. 1, pp. 59–77, 2020, doi: 10.3390/journalmedia1010005.

[49] "International Organization for Migration." https://www.iom.int/ (accessed Jul. 07, 2021).

[50] S. Raschka and V. Mirjalili, *Python Machine Learning: Machine Learning and Deep Learning With Python, Scikit-Learn,* vol. 11, no. 1. Packt Publishing Ltd, 2021.

[51] "cl-tohoku/bert-japanese: BERT models for Japanese text." https://github.com/cl-tohoku/bert-japanese (accessed Jul. 07, 2021).

[52] M. E. Mavroforakis and S. Theodoridis, "A geometric approach to support vector machine (SVM) classification," *IEEE Trans. Neural Networks*, vol. 17, no. 3, pp. 671–682, May 2006, doi: 10.1109/TNN.2006.873281.

[53] "Support Vector Machine — Introduction to Machine Learning Algorithms | by

Rohith Gandhi | Towards Data Science." https://towardsdatascience.com/support-vector-machine-introduction-to-machine-learning-algorithms-934a444fca47 (accessed Jul. 07, 2021).

[54] "XG-Boost-FINAL-01 | Flirting with Models." https://blog.thinknewfound.com/2020/05/defensive-equity-with-machine-learning/xg-boost-final-01/ (accessed Jul. 07, 2021).

[55] "Typhoon Hagibis: death toll rises in Japan as 'worst storm in 60 years' roars through | World news | The Guardian." https://www.theguardian.com/world/2019/oct/12/japan-typhoon-hagibis-tokyo-earthquake-rugby-flood-rain (accessed Sep. 16, 2020).

[56] "Sexist comments made Feb. 3 by Tokyo Olympic chief Mori." https://english.kyodonews.net/news/2021/02/41776ff93754-sexist-comments-made-feb-3-by-tokyo-olympic-chief-mori.html (accessed Jul. 07, 2021).

[57] E. Hutto, C.J. and Gilbert, "VADER: A Parsimonious Rule-based Model for," *Eighth Int. AAAI Conf. Weblogs Soc. Media*, p. 18, 2014, [Online]. Available: https://www.aaai.org/ocs/index.php/ICWSM/ICWSM14/paper/viewPaper/8109.

[58] X. Lu and C. Brelsford, "Network structure and community evolution on Twitter: Human behavior change in response to the 2011 Japanese earthquake and tsunami," *Sci. Rep.*, vol. 4, no. 1, pp. 1–11, Oct. 2014, doi: 10.1038/srep06773.

[59] G. L. Martin, M. E. Mswahili, and Y.-S. Jeong, "Sentiment Classification in Swahili Language Using Multilingual BERT," Apr. 2021, Accessed: Jul. 08, 2021. [Online]. Available: https://arxiv.org/abs/2104.09006v1.

[60] B. of U. D. T. M. Government, "大手町・丸の内・有楽町地区都市再生安全確保計画," Tokyo, 2021. [Online]. Available: https://www.toshiseibi.metro.tokyo.lg.jp/seisaku/toshisaisei/pdf/toshisaisei_keikaku_09.pdf.

[61] C. W. Office, "Chiyoda City Disaster Response Manual," Chiyoda, 2018. [Online]. Available: https://www.city.chiyoda.lg.jp/documents/2094/saigaitaio-en.pdf.

APPENDICES

1. Initial Interview with Foreign Visitors to Tokyo

Objective:

To understand foreigners' disaster preparedness, risk perception and experience with disasters (if any) while in Japan

Table 20 Initial Interview with Foreign Visitors to Japan

Interviewee	Date	Details
A	November 14, 2020 Venue: Online	• Nationality: American • Frequent business visitor to Japan (more than 20 times) • No experience on disasters while staying in Tokyo. • Stated that he does not think about disasters happening when visiting Japan. • In case of disaster happening, he believes his first instinct is to ask other people/ Japanese in the vicinity; or ask the counter at the train station. • Stated that he is confident about Japan's disaster preparedness and would be able to provide him with needed information. • No installed mobile application related to disaster safety. • Stated that he is quite confident Japanese would help him in time of crisis so he believes he would not panic
B	November 26,	• Nationality: Filipino

Interviewee	Date	Details
	2020 Venue: Online	• Assigned to work in Japan for one year. • Stated that Japanese language proficiency is an issue. • No disaster preparedness mobile application. • Stated that the hotel where he usually stays for business trip in Tokyo also do not have any guidance on disaster what-to-dos. Surprisingly, however, there is a pamphlet on what to do in case of nuclear/ radiation leakage. • Have not experienced any severe disaster while staying in Japan. • Stated that he is confident about the integrity of buildings in Japan. • One main concern is how to confirm what kind of disaster has happened (for example, if people in the train station start panicking or evacuating) • Stated that stations in Tokyo provide English guidance/signages so he believes he would not panic so much. However, his concern is if train station staff are capable of providing English instructions right after the disaster (i.e., when a strong earthquake occur, the tarin driver make announcement only in Japan since this is not a pre-prepared message)

2. Prototyping Results

To conduct simulation on the English sentiment analyzer subsystem, a prototype is made on the riot at the US Capitol.

Disaster Event: Storming of the US Capitol

Data Collection: January 6-12, 2021

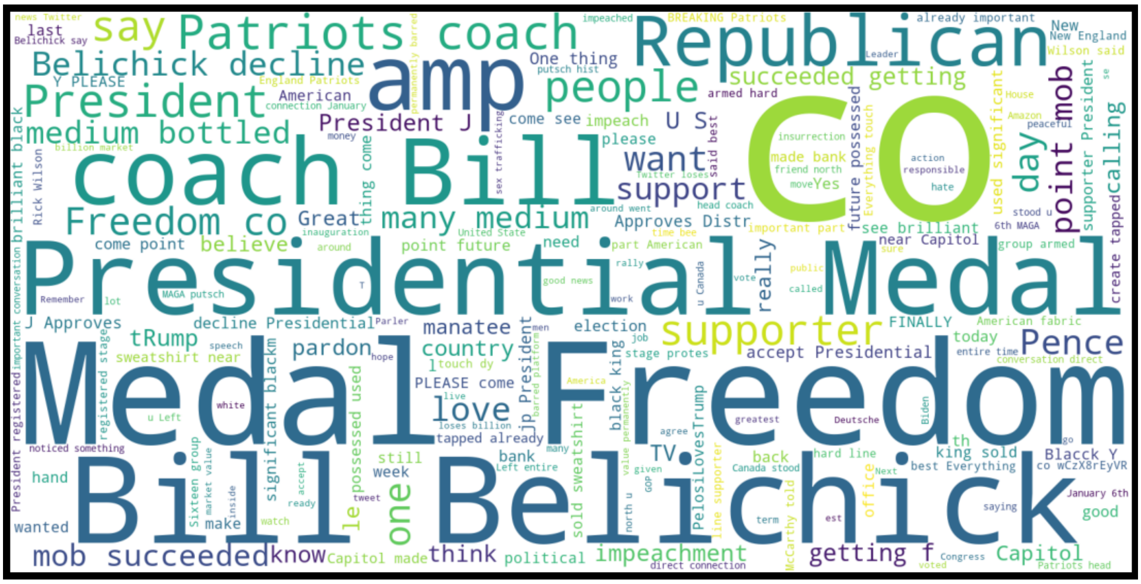
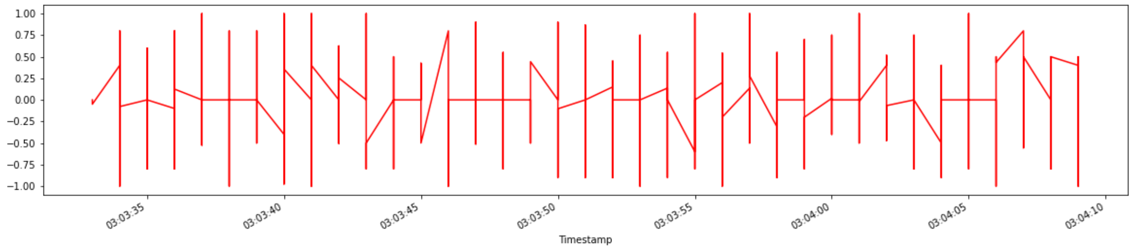


Figure 50 Positive Word Cloud Prototype (English)

Figure 51 Negative Word Cloud Prototype (English)

Prototype of time series analyzer can be seen here:



3. Tweet preprocessing code snippets

```
In [ ]: import pandas as pd
        import re
        import preprocessor as p
        import nltk
        nltk.download('wordnet')
        from sklearn.svm import LinearSVC
        from sklearn.model_selection import train_test_split
        from sklearn.feature_extraction.text import CountVectorizer, TfidfVectorizer
        import sklearn.metrics
        import numpy as np

        [nltk_data] Downloading package wordnet to /root/nltk_data...
        [nltk_data]   Unzipping corpora/wordnet.zip.
```

```
In [ ]: from google.colab import drive
        drive.mount('/content/drive')

        Drive already mounted at /content/drive; to attempt to forcibly remount, call drive.mount("/content/drive", force_remount=True).
```

```
In [ ]: # regexes

        txt='''Chinese Loess Plateau, Upper Continental Crust, delicacy, delicately intricately grooved
         loosed woven Burgess shale loves laden linen lady readily golden.
        The quick brown fox jumps over the lazy dog. Three blind mice followed. beloved'''
        adv=r'(ly|)\b'
        doubles=r'([^s]{1}s|e[nd]|ly)( |[.!?]|$)'
        print(re.sub(doubles,'* ',txt))

        Chinese Loess Plateau, Upper Continental Crust, delicacy, delicate* intricate* grooved
         loos* wov* Burgess shale lov* lad* lin* lady readi* gold*
        The quick brown fox jum* over the lazy dog. Three blind mice follow*  belov*
```

```
In [ ]: # removed emojis
        # removed URLs
        # removed mentions
        # removed numbers

        def pp1(xtr):
         # for i in addtlre:
         #   oustr=re.sub(i,'',oustr)
          xtr=re.sub(r'(([f]?[h]?[t]?tp[s]?[:][/][/])?(\w+)?@?\w+\.\w+[.]?(\w+)?/?([^\s]*)?)|(@\w+)|[#]','',xtr)
          # remove URLs, e-mails, mentions, and hashtags
          xtr=xtr.strip().lower()
          # flatten character case
          nts=r'(^| )(((would|should|could|do|did|have|had|has|is|are)((\'ve)|(n(\')?t))?)|(((wo|sha|ca)n(\')?t)|will|shall|can))( |$)'
          xtr=re.sub(nts,r'\1 \13',xtr)
          # remove because they are noninformative
          xtr=re.sub(r'\s+',' ',xtr)
          # remove repeated whitespaces
          xtr=re.split(r'\s',xtr)
          # split string to list at whitespaces
          xtr=[re.sub(r'\W|\d','',token) for token in xtr]
          # remove non alphanumeric characters
          xtr=str(' '.join(xtr))
          xtr=re.sub(r'([a-zA-Z])\1{3,}',r'\1',xtr)
          # remove extended words
          xtr=re.sub(r'(([bcdfghjklmnpqrstvwxyz]{5,})([aeiou]{4,})?)+','',xtr)
          # remove gibberish
          xtr=re.sub(r'( |^)(not|no)( )?(very|so)?( )?(\w+)( |$)',r'\1\2\4\6\7',xtr)
          # lump negations
          # xtr=re.sub(r'(^| ).?(\W|\d)*.?( |$)','',xtr)
          # remove non-alphabetic characters
          xtr=re.sub(r'\s+',' ',xtr)
          # remove repeated whitespaces, again
          wnl=nltk.stem.wordnet.WordNetLemmatizer()
          xtr=wnl.lemmatize(xtr)
          return xtr
        #negation
        #lemmatization -done
        #stemming - no stemming because it will mutilate text beyond recognition

        # for vectorizer later
        '''stopws=['and','or','but','yet','for','because','since','nevertheless','nonetheless','even','although','however',
                  'a','an','the','you','i','me','us','myself','my','mine','we','your','yourself','them','they','their',
                  'his','her','our','themselves','yours','hers','ours','theirs']
          xtr=[token for token in xtr if token not in stopws]'''

Out[ ]: "stopws=['and','or','but','yet','for','because','since','nevertheless','nonetheless','even','although','however',\n          'a','an','th
        e','you','i','me','us','myself','my','mine','we','your','yourself','them','they','their',\n          'his','her','our','themselves','your
        s','hers','ours','theirs']\n  xtr=[token for token in xtr if token not in stopws]"

In [ ]: testr='''The quick 23 quick heeeeeeelp meeeeeeee I\'m falling for uuuuuuuu
        uuuuuuu brown foxes jumped over the 45 lazy dogs and 2 pigs :D @chiscurl #80234ufdhui $5@0
         dg8nw830#@#$59n jdelacruz@up.edu.ph isdjfjjd@!#$ and somewhere over the rainbow blue birds fly for when all
         is lost then all is found.not so beautiful not so true :D shouldn't'''
        testr=p.clean(testr)
```

4. Validation Interview with Foreigners

To validate the results of the sentiment analysis from the algorithm that the system utilized, interviews were conducted with foreigners living in Tokyo. The following key questions were asked to all interviewees:

1. Have you experienced sever natural disasters before in your home country or overseas? If so, how were they?
2. How do you prepare yourself from disasters?
3. How was your experience with Typhoon Hagibis in 2019? How about the strong Fukushima earthquake last February 2021?
4. How did you feel during those times? Were you calm? Did you feel any difficulty, stress, or negative sentiment?
5. Do you think the existing disaster communication and support for foreigners here in Tokyo (or Japan, in general) is enough? If not, how could we improve it?