International Journal of Information Management 44 (2019) 25–37

Contents lists available at ScienceDirect

International Journal of Information Management

journal homepage: www.elsevier.com/locate/ijinfomgt

Social media data and post-disaster recovery

Mehdi Jamali[a,*], Ali Nejat[b], Souparno Ghosh[c], Fang Jin[d], Guofeng Cao[e]

[a] Department of Civil, Environmental, and Construction Engineering, Texas Tech University, Office: 104, Lubbock, TX 79409, United States
[b] Department of Civil, Environmental, and Construction Engineering, Texas Tech University, Office: 123, Lubbock, TX 79409, United States
[c] Department of Mathematics & Statistics, Texas Tech University, Office: 234, MATH, Lubbock, TX 79409, United States
[d] Department of Computer Science, Texas Tech University, Box 43104, Lubbock, TX 79409-3104, United States
[e] Department of Geosciences, Texas Tech University, MS 1053, Lubbock, TX 79409-1053, United States

ARTICLE INFO

Keywords:
Temporal–spatial patterns
Post-disaster recovery
Social media
Twitter

ABSTRACT

This study introduces a multi-step methodology for analyzing social media data during the post-disaster recovery phase of Hurricane Sandy. Its outputs include identification of the people who experienced the disaster, estimates of their physical location, assessments of the topics they discussed post-disaster, analysis of the tract-level relationships between the topics people discussed and tract-level internal attributes, and a comparison of these outputs to those of people who did not experience the disaster. *Faith-based*, *community*, *assets*, and *financial* topics emerged as major topics of discussion within the context of the disaster experience. The differences between predictors of these topics compared to those of people who did not experience the disaster were investigated in depth, revealing considerable differences among vulnerable populations. The use of this methodology as a new Machine Learning Algorithm to analyze large volumes of social media data is advocated in the conclusion.

1. Introduction

A natural disaster negatively impacts all aspects of one's life. It can not only devastate the physical settings of a community by destroying infrastructure, the landscape, residential and businesses properties, it can also affect one's emotional well being after witnessing loss of life and suffering the disruption of established social interactions. Aside from such immediate mental and physical harm, disasters also have long-term consequences, such as job losses, financially insurmountable property damage, and post-traumatic stress disorder (PTSD). Since the routines of daily life are tightly interwoven with the stability of both physical settings and social interactions, disasters upend the tranquility of people's lives for short, and sometimes long periods of time.

A return to normalcy is the ultimate goal of post-disaster recovery policies. A robust understanding of the patterns and types of damage common to disasters in general is crucial in the process of formulating effective post-disaster recovery policies and programs. Disasters are complex. They impact survivors' quality of life through the damage they inflict on natural and manmade landscapes. The existing literature distinguishes five categories of impacts (Lindell & Prater, 2003): *social impacts*, such as the appearance of conflicts and the loss of social capital (Lindell & Prater, 2003); *psychosocial impacts*, such as post-traumatic stress disorder (PTSD) (Gleser, Green, & Winget, 2013; Steinglass & Gerrity, 1990); *demographic impacts*, such as changes in population distribution (Kaniasty & Norris, 1993; Smith & McCarty, 1996); *socio-economic impacts*, such as job loss and business closures (Okuyama & Chang, 2013); and *political impacts* (Drury & Olson, 1998; Toya & Skidmore, 2014). These impacts are obviously interconnected. For instance, improvements in the socioeconomic condition of a community will influence psychosocial attributes (generally in positive ways), or, abrupt disruption of pre-established social and communal interactions may bring about adverse psychosocial and political reactions. Moreover, the relative importance of these categories can differ among communities and even individuals in the same community who experience the same disaster event, based on the innate characteristics of that community or individual. These characteristics can comprise a many different parameters, including job/income (Fothergill & Peek, 2004; Masozera, Bailey, & Kerchner, 2007), ethnicity (Bolin & Bolton, 1986), and age/gender (Nakagawa & Shaw, 2004), among others.

The recovery priorities of disaster survivors (Nejat, Brokopp Binder, Greer, & Jamali, 2018; Quarantelli, 1999; Wold, 2006) play a major role in the success of disaster recovery policies and programs (Ragini, Anand, & Bhaskar, 2018). Individual priorities are strongly influenced by income, age, and social capital. Hence, the design of a post-disaster recovery plan is a dilemma complicated by a diversity of parameters, internal attributes, unique impacts of a given disaster, and survivors'

* Corresponding author.
E-mail addresses: mehdi29j@gmail.com (M. Jamali), ali.nejat@ttu.edu (A. Nejat), souparno.ghosh@ttu.edu (S. Ghosh), fang.jin@ttu.edu (F. Jin), guofeng.cao@ttu.edu (G. Cao).

https://doi.org/10.1016/j.ijinfomgt.2018.09.005
Received 13 December 2017; Received in revised form 15 September 2018; Accepted 15 September 2018
0268-4012/

priorities. A major objective of recovery plans is the swift return of impacted communities to normalcy. Researchers and policymakers must therefore be able to quickly arrive at an accurate understanding of the relationship between community characteristics, individual personal internal attributes, and survivors' post-disaster priorities design the most effective recovery plans.

Modern social media applications, which have achieved considerable penetration into the everyday life of many users, provide an invaluable source of data regarding user thoughts, beliefs, and opinions. Social media consists of users with diverse backgrounds, and have ability to encourage aggregation of the users, can provide unique substrate for researchers to understand behavioral patters of communities (Dhir, Kaur, & Rajala, 2018; Kapoor et al., 2018; Liu, Lee, Liu, & Chen, 2018; Liu, Shao, & Fan, 2018). Research shows that heavy social media users seek out contacts, content boosts, favorable information, requirement inquiries, stress discharge, "emotional support", and sense of belonging (Gilbert & Karahalios, 2009; Kaplan & Haenlein, 2010; Liu, Shao et al., 2018). As an example, Grover, Kar, and Davies (2018) showed that providing emotional support, creating awareness, and sharing information are important users' motivations that health related industries seeking in social media platforms. Disaster cause severe, instantaneous distress on individuals, who as a result seek to mend their emotional traumas through social media outlets (Gao, Barbier, & Goolsby, 2011; Hughes, Palen, Sutton, Liu, & Vieweg, 2008). Therefore, these applications can provide invaluable data with which to study peoples' priorities after disaster strikes (Li, Zhang, Tian, & Wang, 2018). Currently, Facebook, Twitter and Instagram are major examples of worldwide social media applications with millions of everyday users scattered around the world. Fortunately, their data are relatively publicly available for research purposes. However, as discussed by Stieglitz, Mirbabaie, Ross, and Neuberger (2018), volume and variety of the social media data is the most common cited challenge in social media studies. Unfortunately, besides the volume and variety of data, the complex nature of a given disaster's consequences, makes said data less applicable in post-disaster recovery studies. Also, due to variety of users with wide range of purposes it is necessary to better understand the role of social media data in emergency management (Kim, Bae, & Hastak, 2018; Martínez-Rojas, del Carmen Pardo-Ferreira, & Rubio-Romero, 2018). To better utilize social media data in post-disaster recovery studies, we introduce a new methodology to analyze social media data (specifically Twitter data) and scrutinize community reactions in the aftermath of a disaster, based on social media statements. The methodology answers three questions: 1) What are the priorities of people who have experienced a disaster? 2) What are the tract-level relationships between these priorities and survivors' internal attributes? 3) How do these priorities differ between people who experienced the disaster and those who did not?

Definitely, understanding the priorities of people who have experienced the disaster is an indispensible part of designing an effective post-disaster recovery policy. This perception can assist policy makers to optimize the distribution of federal resources, and enhance the planning for reconstruction process. In order to design an effective post-disaster recovery policy it is not only important to understand the overall priorities of disaster victims, but also it is important to figure out regional priorities of victims which may be differing based on communal internal attributes. Therefore, understating the tract-level relationship between these priorities and survivors' internal attributes can assist policy makers to figure out regional priorities of disaster victims. Finally, in this study we will not only reveal overall and regional variations of priorities, but also we will study how these priorities may affect for people who did not experience the disaster. This outcome may again help policy makers to design better policies for affected and non-affected zones.

Twitter, created in 2006, is a micro-blog social media tool where registered users read and write messages, called "tweets," of up to 140 characters and unregistered users can read such messages (Twitter, 2016). Twitter is available via website, short message services (SMS), and mobile app (Twitter, 2016). As of December 2016, Twitter had more than 300 million monthly active users, with more than 1 billion monthly visits to the website (Twitter). According to estimates by InternetLiveStats (2017), Twitter publishes around 200 billion tweets each year, or approximately 6500 tweets per second. Additionally, as of December 2016, Twitter had more than 67 million active users in the United States, which made up about 20% of the service's active users (InternetLiveStats, 2017). Fortunately, when Twitter, Inc. promoted the Twitter API, tweet data became available for research purposes. This data found wide applications in several research studies in political science where (Tumasjan, Sprenger, Sandner, & Welpe, 2010) analyzed tweet contents for sentiments like anxiety, anger, sadness, and compared their correspondence with subjects' political parties; in human studies where (Bakshy, Hofman, Mason, & Watts, 2011) used graph analysis of followers among 1.6 M Twitter users to define the impact of users on the whole community of users; in human mobility and geography where (Leetaru, Wang, Cao, Padmanabhan, & Shook, 2013) illustrated the year to year growth of social media and visualized the impact of social media on human communication by mapping the available world-wide geo-tagged tweets; in public health where (Cao et al., 2015) developed a spatiotemporal model to understand the movements of Twitter users within specific geographic boundaries; in economics where (Bollen, Mao, & Zeng, 2011) defined public mood as the percetnage of positive tweets and the time series analysis of Dow Jones Industiral Avearage (DJIA) was shown to be predictable by the public mood index; and in education where (Junco, Heiberger, & Loken, 2011) showed engagment in social media can lead to an increase in student and faculty communication and improve the performance of the education process. As another interesting example, Nisar, Prabhakar, and Patil (2018) discussed how sports clubs absorb fans' attention by strategizing their activity in social media applications. As such, social media provides unique substrate to study human behavior, their sentiments and their decision making in everyday situations of life (Dong & Wang, 2018; Jeong, Yoon, & Lee, 2017; Lee & Hong, 2016).

Although the methodology discussed in this study can be generalized to all disasters, Hurricane Sandy was chosen for this case study due to the abundance of available data and the totality of damage it inflicted. Hurricane Sandy formed on October 22, 2012 and faded on November 1. It affected 24 states in all, including those of the United States Eastern Seaboard from Florida to Maine (Force, 2013). Hurricane Sandy was the second most destructive natural disaster in United States history. It caused more than $85 billion in damage and more than 200 fatalities (Force, 2013). The hurricane and its associated floods and fires affected millions of people in both urban and suburban communities, caused power outages, impeded transportation systems, destroyed residential properties, and produced more than an estimated $32 billion in economic losses (Bloomberg, 2013).

In the remaining sections of this manuscript, we will discuss a literature review on manuscripts related to post-disaster recovery indicators, post-disaster studies on Twitter data, suitable text mining algorithms and statistical descriptive method. Following literature review we will discuss theoretical basis of our study in which we will provide a framework for our expected results. And after that, we will have methodology in which will discuss in detail the thirteen steps of the algorithm. Following the methodology we will discuss the findings and at the end we have conclusion and future work.

2. Literature review

2.1. Disaster recovery and related indicators

According to Chang (2010), post-disaster recovery ought to be judged in terms of either returning environments to their pre-disaster conditions, building communities up to where they would have progressed if disaster had not struck, or finding a middle ground between the two. Chang's study suggests a framework for measuring the success of disaster recovery based on many parameters, including Gross Regional Product (GRP), the number of businesses, and population. It should be noted that long-term losses from disasters, for instance disruptions to small businesses, are hard to detect during the first stages of post-disaster recovery.

Several demographic and socioeconomic indicators may influence the disaster recovery process, including ethnicity and income. For example, research on victims six months after Hurricane Andrew shows that ethnicity strongly influences levels of PTSD (Perilla, Norris, & Lavizzo, 2002). Caucasians, African-Americans, and Latinos displayed, respectively, the lowest to highest occurrences of PTSD symptoms (Perilla et al., 2002). Income is also a significant factor in PTSD (Rhodes et al., 2010). According to a study of 392 parents of low-income households hit by Hurricane Katrina, their probability of suffering severe psychological distress was roughly twice that of other people (Rhodes et al., 2010). Overall, as Jamali and Nejat (2016) found in their comprehensive systemic review of approximately 40 studies on determinants of post-disaster behaviors, these determinants can be grouped into four inclusive categories: *demographics* (age, gender, ethnicity, religion); *socioeconomics* (job, income, education, homeownership); *spatial* (home, neighborhood, city, location, rural, urban) and *psychosocial* (social capital, mental health) variables.

2.2. Post-disaster studies on Twitter data

Although applications of social media in the preparedness, response, and mitigation phases of natural disasters have been widely studied (Fraustino, Liu, & Jin, 2012; Kim & Hastak, 2018; Lindsay, 2011; Sutton et al., 2014; Yates & Paquette, 2010), there have been only a few studies using social media data as a predictor for disaster recovery policies. Guan and Chen (2014) produced one work that applied Twitter data to recovery plans. They filtered disaster-related tweets from Hurricane Sandy and found correlations between recovery and the number of disaster-related tweets. Based on their findings, disaster-related tweets were higher in coastline areas with a positive correlation between the level of damage and ratio of disaster-related tweets (Guan & Chen, 2014). Another study by Glasgow, Vitak, Tausczik, and Fink (2016) introduced a different approach to evaluate social support after disasters. They utilized a method to find positive tweets, categorized as "*gratitude*" for social support after a disaster. The authors explained that higher levels of damage lead to lower rates of gratitude after a disaster. Additionally, by scrutinizing tweets, the study found social media to be a powerful tool in assessing "*resilience and healing*" after disasters (Glasgow et al., 2016). Finally, Hughes et al. (2008) demonstrated how advancing mass communication tools can augment "*social convergence*" of people in the recovery phase of disasters. They assert that mass communication tools provide a unique opportunity to transform individual mourning behavior into social cohesion attitudes, a feature that can find extensive applications in post-disaster recovery studies.

Aside from aforementioned studies, some disaster information system researchers have tried to apply Natural Language Processing and Artificial Intelligence methods to the analysis of social media data. For instance, Maldonado, Alulema, Morocho, and Proaño (2016) introduced a combined topic modeling and text filtering algorithm to detect natural disaster events and measure how often Twitter users are interested in natural disasters rather than social and political events. Additionally, Verma et al. (2011) employed Naïve Bayes and MaxEnt topic modeling algorithms to detect disaster victims' situational awareness and compared the algorithms' outcomes with qualitative classifiers. Surprisingly, they found that Machine Learning Algorithms overlap about 80% with qualitative classifiers.

Therefore, the main contribution of this research is to bridge the current gap by developing a model that can account for the complexities of post-disaster recovery by linking tract-level demographic and socioeconomic attributes to tract-level distribution of opinions that can be used to draft tract-level recovery policies to better address the needs.

2.3. Latent dirichlet allocation versus dynamic query expansion (DQE)

Several algorithms have been utilized to analyze Twitter data. Latent Dirichlet Allocation (LDA), first introduced by Blei, Ng, and Jordan (2003) is one of the most common methods of topic modeling. It utilizes Bayesian statistical procedures to find topics hidden in text data (Blei et al., 2003). As mentioned by Mehrotra, Sanner, Buntine, and Xie (2013), due to Twitter's restricted number of characters (maximum 140) and vast number of documents (millions of tweets in each area), the application of LDA to Twitter requires some adaptations and accurate data cleaning. Although LDA has been used widely in text mining analysis of Twitter data, our study required a very specific and accurate topic selection methodology, so we turned to Dynamic Query Expansion.

Dynamic Query Expansion (DQE) is a frequency-based algorithm. While some words that appear in many documents, such as "the," "if," "is," etc., are not useful for topic identification, the bag of words containing the most frequently used words with related meanings is valuable for detecting topics. This bag of words is the steady state of the most frequent words in related tweets, filtered by the previous step's bag of words. This methodology was introduced by Zhao et al. (2014) as an advanced text-mining approach, which is an efficient way of topic modeling micro-blog documents. For instance, when one looks for topics related to "election" for the first iteration, one filters tweets that contain the word "election." The most frequent words of the filtered tweets can then be used to create an updated bag of words. For the second iteration, we filter the tweets by the first step's bag of words and create the second-step bag of words. These steps are repeated until arriving at the steady state of bag of words, wherein the words of n^{th} iteration, have major overlap with $n\text{-}1^{th}$ bag of words. If one assumes that "president," "poll," "candidate," etc. constitutes the final bag of words, one can be confident that these are valuable words to keep track of among the tweets.

2.4. Dirichlet regression

Compositional data is the vector of non-negative percentage proportions with unit-sum. Aitchison (1982) introduced one of the earliest attempts to analyze compositional data by Log-ratio approaches. Based on Aitchison's (1982) method, the trends of compositional data are expressed through known response variables, a procedure that is problematic for datasets with low or zero percentage components (Baxter, Beardah, Cool, & Jackson, 2005). The Dirichlet regression is a flexible model in detecting trends of compositional data (Hijazi & Jernigan, 2009). The data in Dirichlet regression consists of two data-frames, one for dependent variables (which is the unit-sum vectors of responses) and a second for covariates. The goal of Dirichlet regression is to predict responses as a linear function of covariates. Maximum likelihood is the method of parameter estimation in Dirichlet regression (Hijazi & Jernigan, 2009).

Dirichlet regression is broadly applied in psychiatry (Gueorguieva, Rosenheck, & Zelterman, 2008), agricultural sciences (Hickey, Kelly, Carroll, & O'Connor, 2015) and even social science (Ivanova, Maier, & Meyer, 2016). For instance, Gueorguieva et al. (2008) studied the effects of various covariates, such as depression, substance abuse, length of illness, and age on five components of schizophrenia: positive, negative, cognitive, emotional and hostility. In this study, the five components of depression were a set of proportionate compositional data with summed to one. These researchers found a negative contribution for history of depression as a predictor of cognitive schizophrenia (Gueorguieva et al., 2008). Hickey et al. (2015) showed that Dirichlet regression outperformed Neural Network and Multivariate Multiple Regression in predicting models with multiple response variables. The authors tried to predict which forest crop (a set of nine categorical and numerical attributes) was significant in predicting forest compartment proportions (sawlog, pallet, stake, and pulp) and found that the proportion by which the value of sawlog increased approximately doubled

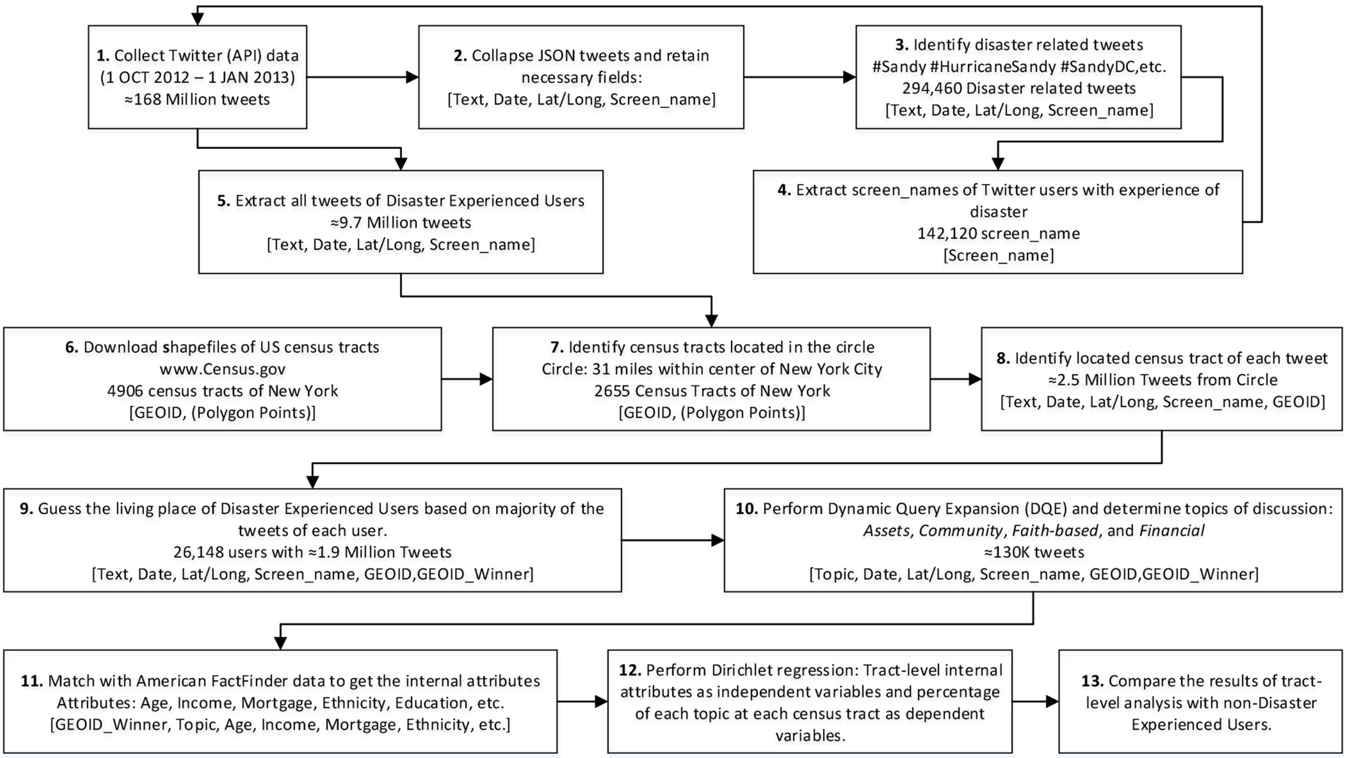

Fig. 1. Flowchart of the presented steps in methodology section.

for the case in which the species of the tree (the predictor) was Norway spruce (Hickey et al., 2015). Last but not least is the research of Ivanova et al. (2016) regarding connections between private organizations and political transitions in Russia. This study measured size, type of agency, type of membership, and presence in social media, and three dependent variables of the prediction. They found that the size and specific type of private organizations are the most influential factors in community functions during political transitions.

3. Theoretical basis

Social media data can provide invaluable opportunities for understanding public opinion in natural disaster studies (Kim & Hastak, 2018); however, solid knowledge of the geographical distribution of users is indispensible when using such data. Motivations for using social media can include the desire to express current status, describe daily activities, share information, and seek social and emotional support, among others (De Choudhury, Gamon, Counts, & Horvitz, 2013; Java, Song, Finin, & Tseng, 2009). In fact, the availability of social media data, in conjunction with users' intentions, makes these applications a rich source of information to understand peoples' behavior and cast light on their sentiments and opinions (Lipizzi, Iandoli, & Marquez, 2015; Pak & Paroubek, 2010). In a basic sense, Twitter's micro-blog data reflects public opinion about economic, social, and political issues, in which, for example, topic modeling of tweets can reveal public opinion toward political campaigns (Tumasjan et al., 2010). While Twitter posts can be perceived as a unique source of information to understand public opinion, the geographical distribution of public opinion is an important aspect that should not be overlooked. As discussed by Han, Cook, and Baldwin (2014), the "geolocatability" of Twitter data has affected the temporal variance, feature selection, and even the user behavior of social media users. In particular, due to the multidimensional impact of disasters, geographical distribution and location prediction are crucial aspects of social media studies in emergency management (Singh, Dwivedi, Rana, Kumar, & Kapoor, 2017). Thus, this study, utilizing social media data, will analyze public opinion in close association with temporal–spatial patterns as two crucial aspects of post-disaster recovery policies.

4. Methodology

The methodology in the present study has a number of steps. To make perusing these steps easier, the flowchart presented in Fig. 1 is included.

4.1. Step 1 & 2 – collecting Twitter data and retaining necessary fields

Twitter Streaming API (https://developer.twitter.com/en/docs) provides tweets available for download and was utilized to collect the data for the present study. The API outputs are in JavaScript Object Notation (JSON) format, within which, in addition to the texts of the tweets, many other attributes of Twitter users are presented. For the purposes of this study, four fields of JSON were retained: "text" - the message a user sends with Twitter; "coordinates" - the latitude and longitude of the user at the moment of sending the tweet; "created_at" - the date and time of the tweet's creation; and "screen_name" - the unique pseudonym of each individual Twitter user. The study was initiated using 167,448,932 geotagged tweets by US users from between October 1, 2012 and December 31, 2012 (the timeframe of 92 days for which the Twitter data was provided for this study).

4.2. Step 3 & 4 – disaster related tweets and their screen names

In order to identify disaster-related tweets, a bag of keywords introduced by (Guan & Chen, 2014) was used. Based on their study, it is possible to consider the tweets sent after October 26, 2012 (the date of first warning issuance) and have at least one of these words or hashtags as tweets related to Hurricane Sandy: 'sandy,' 'hurricane,' 'storm,' #Sandy, #HurricaneSandy, #njsandy, #MASandy, #StormDE, #SandyDC and #rigov. Among the whole data-set (around 168 million tweets), 294,460 tweets were identified as disaster-related. Since Guan and Chen (2014) suggested that the density of disaster related tweets is

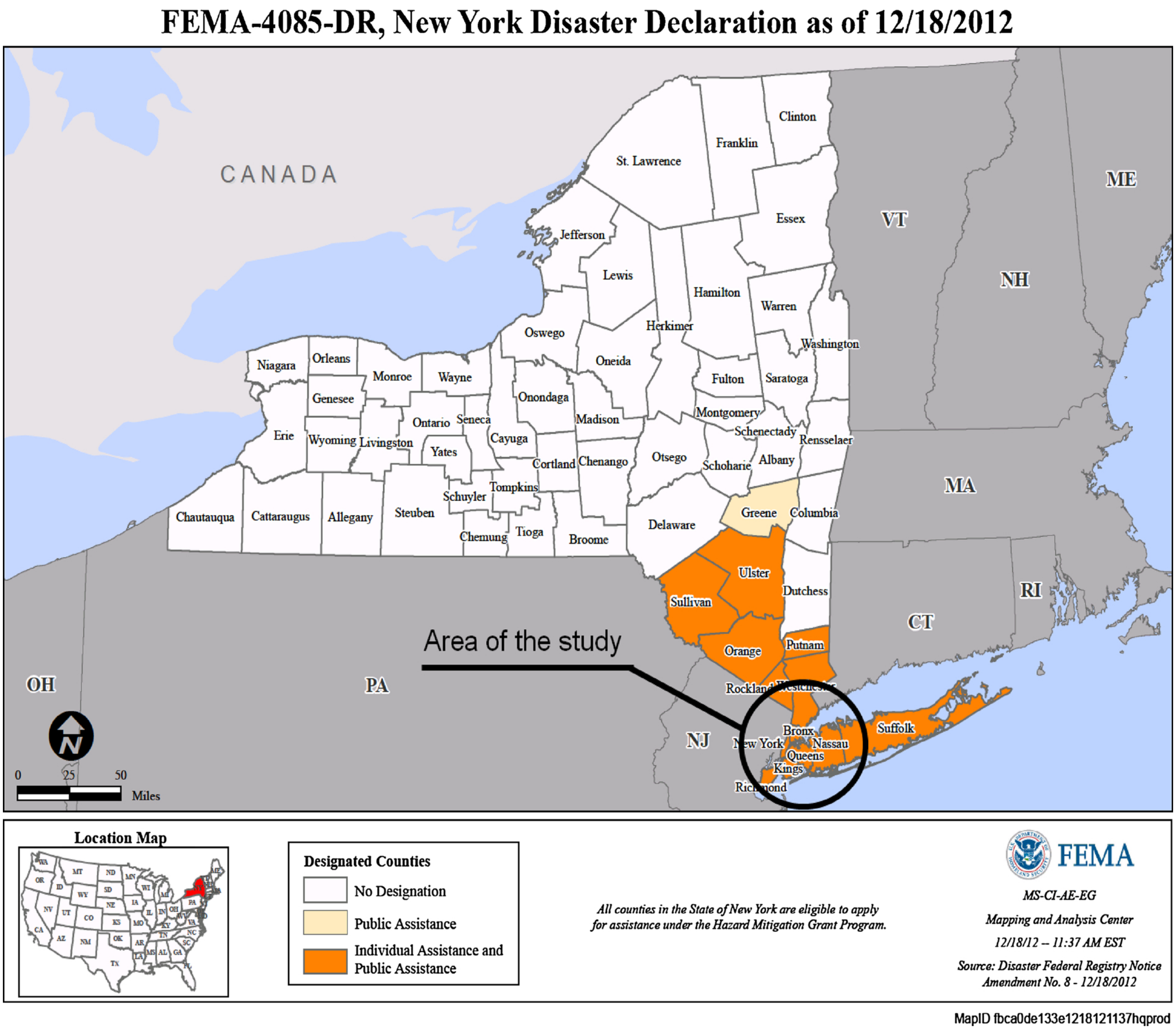

Fig. 2. overlapped map of FEMA disaster declaration zone for hurricane Sandy and area of the study (Circle).

positively correlated with non-recovered outage and level of damage, we assumed that these tweets were sent by the people who felt that the disaster had affected their normal lives. Therefore, 142,120 unique users who sent disaster related tweets were collectively classified as “Disaster Experienced Users.”

4.3. Step 5 – all tweets of disaster experienced users

For this step, all tweets of Disaster Experienced Users were extracted from the whole dataset (~168 million tweets), resulting in 9,710,288 tweets within the abovementioned 92 days. These tweets were to serve two subsequent purposes: 1) Estimation of the living place of Disaster Experienced Users, discussed in Step 9; and 2) Determination of the topics discussed by these users, explained in Step 10.

4.4. Steps 6, 7 & 8 – located census tract of tweets from Step 5

This study focused on New York City, a metropolitan area with a vast number of Twitter users. The study area is a circle centered at (40.7127837, −74.0059413) with a radius of 31 miles (50 km). Hereafter, this area is referred to as the Circle, where the Circle overlaps the FEMA map of the Hurricane Sandy affected zone shown in Fig. 2 (FEMA map available at https://www.fema.gov/disaster/4085). The aim of Steps 6, 7 & 8 is to identify the census tract from which each tweet was sent. Therefore, at Step 6 the shape-file of New York State’s 4906 census tracts was downloaded from available databases (www.census.gov). In Step 7, the 2655 census tracts located fully within the Circle were kept. Afterwards, the census tracts where ~9.7 million tweets from Step 5 originated were identified, showing that 2,545,991 tweets were sent from inside the Circle.

4.5. Step 9 – living quarters locations of disaster experienced users

As this study analyzes the topics discussed by people based on their geographic location, it became necessary to estimate the living quarters locations of Disaster Experienced Users. Therefore, the census tract from which a user had sent most of his/her tweets was considered a possible location of that user’s living quarters. For instance, if one user had sent 10 tweets, 4 of which came from one census tract, which represented the majority of the census tracts from which the user’s tweets originated, it was assumed that user was living in that specific census tract. There were two criteria for eliminating a user from the dataset: 1) if the majority of the user’s tweets were sent from outside the Circle; and 2) if a user had equal numbers of tweets in more than one census tract. For instance, if a user had 7 tweets in total—3 tweets in 2 census tracts and 1 tweet in another—it could not be determined where his/her actual living quarters were and hence the user was excluded. Once 26,148 Twitter users were included, with a total of 1,911,733 tweets, and at least one tweet per user related to Hurricane Sandy, there was enough data to estimate the location of their living quarters.

Table 1
Dynamic Query Expansion results of Disaster Experienced and Non-Disaster Experienced Users.

Topic	Seed words	Selected words	Number of tweets	
			Disaster Experienced	Non-disaster Experienced
Financial	money, financial, insurance, work, budget, making, pay,spending,#money,spent,waste,finance,home,bank,rich,loan,shopping,donate,cash,blow,pocket,paying,bill,dollar,sell,dollars,account,tax,losing,bought,rockefeller,spending,spent,bank,afford,buying,	money, pay, bought, Rockefeller, spend, bank, losing, rich, spent, buying, account, spending,bill,donate,sell,dollar,dollars,cash,paying,tax,financial,afford,insurance,taxes,loan,mortgage,paid, credit, buy, price, fund, expenses, expenses, expense	35,094	20,837
Assets	food,gas,subway,shopping,restaurant,hotel,shop,airport,plaza,macys,flight,downtown,commercial,chipotle,park,school,store,mall,grocery,laundry,theatre,mcdonalds,university,college,starbucks,mcdonalds,	park,restaurant,shopping,hotel,airport,plaza,jfk,bryant,#centralpark,grocery,hospital	33,009	23,708
Community	life,friends,family,friend,halloween,thanksgiving,birthday,sister,loves,social,boyfriend,loving,relationship,cousin,children,holidays,girlfriend,christmas,dad,loving,neighbor,neighbors,luv,love,mother,	friends, family, friend, dad, sister, mother	32,721	18,948
Faith-based	god,scared,bless,sin,lord,prayers,pray,angel,hell,jesus,spirit,heaven,amen,soul,church,	god,hell,jesus,sacred,bless,sin,lord,soul,pray,heaven,amen,angel,prayers,spirit,church	34,047	19,025

4.6. Step 10 – dynamic query expansion (DQE)

DQE was utilized to filter the topics discussed in the tweets from Step 9 (i.e. ~1.9 million Disaster Experienced User tweets). *Faith-based* (i.e. tweets containing faith-based motivations), *assets* (i.e. tweets related to community assets such as substantial infrastructure, public facilities, commercial offerings, etc.), *community* (i.e. tweets related to communal interactions and relationships) and *financial* (i.e. tweets related to financial issues) were four common topics identified in statements made by individuals. Table 1 represents the initial bag of words (seed words), the selected bag of words, and the number of tweets for each topic. Since the impact of the number of tweets per topic on the final results was unclear, the bags of words for the topics were selected in a manner such that all topics are represented by almost equal numbers of tweets (i.e. about 34,000 Disaster Experienced User tweets for each topic). The outcomes from utilizing DQE method were validated for internal consistency using multiple trained coders consisting of one undergraduate and one graduate student. The results from this iterative validation process were in harmony with the original results with *Faith-based*, *Community*, *Assets*, and *Financial* topics common in both.

4.7. Step 11 – American Community Survey data (internal attributes)

The United States Census Bureau (https://www.census.gov) surveys an array of socioeconomic attributes of geographic units across multiple geographic resolutions that together constitute the American Community Survey. Estimates generated by the ACS for 2012, based on census tracts, were downloaded for further processes. Although ACS provides information on hundreds of parameters, 65 attributes deemed most likely to affect individuals were selected based on a schematic overview of existing post-disaster studies. The 65 attributes are sub-attributes of 14 major attributes that have been extracted from census based on the categorization proposed by (Jamali & Nejat, 2016) and are shown in Table 2. As shown in the table, the only category for which data could not be extracted from census was the *psychosocial* category. Table 3 presents a full description of all the 65 sub-attributes.

4.8. Step 12 – Dirichlet regression of disaster experienced users

A Dirichlet regression model was utilized to identify significant internal attributes and study the tract-level correlation between these attributes and identified topics of discussion, The R package DirichletReg asks the user to provide a data-frame of predictors and response variables, in which the response variables must be compositional data that sum to one (e.g. four response variables A:35%, B:25%, C:15% and D:25%). Since this study had 65 predictors characterized in several cases by highly correlated internal attributes (e.g. percent below poverty and income), the Variance Inflation Factor (VIF) based variable selection process was utilized to avoid multicollinearity of predictors. Therefore, all 65 internal attributes were placed into one data-frame. For each iteration of the process the variable with the highest value of VIF was removed from the data-set, until all calculated VIFs fell below the pre-specified cutoff value (i.e. VIF = 10). Of the 65 initial variables, 46 passed the VIF-based variable selection process (see Table 3). Afterwards, in order to study the relationships between internal attributes and topics discussed on Twitter, the percentage of topics at each

Table 2
List of the main attributes and their literature support.

Main Attribute	Sub Attribute	Supporting Literature	
Demographics	Age	(Rubinstein & Parmelee, 1992)	(Hidalgo & Hernandez, 2001)
	Ethnicity	(Perilla et al., 2002)	
Socioeconomics	Homeownership	(Abramson, Stehling-Ariza, Park, Walsh, & Culp, 2010)	
	Level of education	(Lewicka, 2005)	
	Field of study	(Sapat & Esnard, 2012)	
	Poverty	(Nakagawa & Shaw, 2004)	
	Income	(Fothergill & Peek, 2004)	
	Unemployment	(Tierney & Oliver-Smith, 2012)	
	Work status	(Tierney & Oliver-Smith, 2012)	
	Income per capita	(Fothergill & Peek, 2004	(Chang, 2010)
	Job industry	(Badri, Asgary, Eftekhari, & Levy, 2006)	
	Mortgage	(Brown & Perkins, 1992)	(Nejat et al., 2018)
Spatial	Transportation	(Kaufman, Qing, Levenson, & Hanson, 2012)	
	Mobility	(Cutter, Boruff, & Shirley, 2003)	

Table 3
Description of internal attributes and the results of VIF based variable selection.

No.	Major Attribute	Sub-Attribute	Description	VIF (passed ✓)
1	Age	Age_0_14	0–14 years old	✓
		Age_15_24	15–24 years old	✓
		Age_25_34	25–34 years old	✗
		Age_35_44	35–44 years old	✓
		Age_45_54	45–54 years old	✓
		Age_55_69	55–69 years old	✓
		Age_70_100	over 70 years old	✓
2	Transportation	Trans_to_work	Car, truck or van including: Drove alone, Carpooled and excluding: Public transportation, walked, bicycle, taxicab and worked at home.	✓
3	Housholders	Holder_18	Households with one or more people under 18 years	✗
		Holder_60	Households with one or more people 60 years and over	✗
		Holder_alone	Householder living alone	✓
		Owner_occup	Owner-occupied housing units	✗
4	Level of education	Edu_24_college	Population 18–24 years: Some college, bachelor's degree or higher	✓
		Edu_25_school	Population 25 years and over: High school graduate or less	✗
		Edu_25_college	Population 25 years and over: Some college or associate's degree	✓
		Edu_25_grad	Population 25 years and over: Bachelor's degree, graduate or professional degree	✗
5	Education field of study	Edu_25_STEM	Total population 25 years and over with a Bachelor's degree or higher in science and engineering and related fields	✓
6	Poverty	Pct_blw_pvt_18	Percent below poverty level: Under 18 years	✓
		Pct_blw_pvt_64	Percent below poverty level: 18–64 years	✗
		Pct_blw_pvt_65	Percent below poverty level: 65 years and over	✓
		Pct_blw_pvt_whites	Percent below poverty level: Whites	✓
		Pct_blw_pvt_school	Percent below poverty level: High school graduate and less educated (Population 25 years and over)	✗
		Pct_blw_pvt_college	Percent below poverty level: Some college, associate's degree (Population 25 years and over)	✓
		Pct_blw_pvt_bachelor	Percent below poverty level: Bachelor's degree or higher (Population 25 years and over)	✓
		Pct_blw_pvt_empld	Percent below poverty level: Employed (Civilian labor force 16 years and over)	✗
		Pct_blw_pvt_full	Percent below poverty level: Worked full-time, year-round in the past 12 months (Population 16 years and over)	✓
		Pct_blw_pvt_part	Percent below poverty level: Worked part-time or part-year in the past 12 months (Population 16 years and over)	✓
		Pct_blw_pvt_unempld	Percent below poverty level: Did not work (Population 16 years and over)	✗
7	Income	Income_0_24	Less than $24,999	✗
		Income_25_49	$25,000 to $49,999	✓
		Income_50_74	$50,000 to $74,999	✓
		Income_75_99	$75,000 to $99,999	✓
		Income_100	More than $100,000	✗
		Income_Median	–	✗
		Income_Mean	–	✗
8	Unemployment	Unemply_16_24	Unemployment rate: 16–24 years old	✓
		Unemply_25_54	Unemployment rate: 25–54 years old	✓
		Unemply_55_99	Unemployment rate: Over 55 years old	✓
		Unemply_whites	Unemployment rate: Whites	✓
		Unemply_school	Unemployment rate: High school graduate and less educated (Population 25 years to 64 years)	✓
		Unemply_college	Unemployment rate: Some college, associate's degree (Population 25 years to 64 years)	✓
		Unemply_bachelor	Unemployment rate: Bachelor's degree or higher (Population 25 years to 64 years)	✓
9	Ethnicity	Ethincity_Whites	Percentage White	✗
		Ethincity_African	Percentage Black or African American	✓
10	Per capita income	Per_Capita_Income	Per capita income in the past 12 months (in 2013 inflation-adjusted dollars)	✓
11	Work status	Wrk_50_52	Population 16–64 years: 50–52 weeks of employment in past 12 months	✓
		Wrk_27_49	Population 16–64 years: 27–49 weeks of employment in past 12 months	✓
		Wrk_1_26	Population 16–64 years: 1–26 weeks of employment in past 12 months	✓
		Wrk_not	Population 16–64 years: Did not work	✗
		Wrk_hrs_35	USUAL HOURS WORKED: Usually worked 35 or more hours per week	✗
		Wrk_hrs_15_34	USUAL HOURS WORKED: Usually worked 15 to 34 h per week	✓
		Wrk_hrs_1_14	USUAL HOURS WORKED: Usually worked 1 to 14 h per week	✓
12	Job industry	Job_isty_Agric	Civilian employed population 16 years and over: Agriculture, forestry, fishing and hunting, and mining	✓
		Job_isty_Const	Civilian employed population 16 years and over: Construction and manufacturing	✓
		Job_isty_Trade	Civilian employed population 16 years and over: Wholesale trade and retail trade	✓
		Job_isty_Trans	Civilian employed population 16 years and over: Transportation and warehousing, and utilities	✓
		Job_isty_Finan	Civilian employed population 16 years and over: Finance and insurance, and real estate and rental and leasing	✓
		Job_isty_Info	Civilian employed population 16 years and over: Information	✓
		Job_isty_Prof	Civilian employed population 16 years and over: Professional, scientific, and management, and administrative and waste management services	✓
		Job_isty_Edu	Civilian employed population 16 years and over: Educational services, and health care and social assistance	✗
		Job_isty_Arts	Civilian employed population 16 years and over: Arts, entertainment, and recreation, and accommodation and food services	✓
		Job_isty_Public	Civilian employed population 16 years and over: Public administration	✓
13	Mobility	Mobility_same	Percentage of population lived in same house 1 year ago (2013) excluding: Moved within same county, Moved from different county within same state and Moved from different state.	✗
		Mobility_moved	Moved within same county	✓

(*continued on next page*)

Table 3 (*continued*)

No.	Major Attribute	Sub-Attribute	Description	VIF (passed ✓)
14	Mortgage	Mortgage	Housing units with a mortgage, contract to purchase, or similar debt including: With either a second mortgage or home equity loan, but not both; No second mortgage and no home equity loan and Both second mortgage and home equity loan.	✓

census tract was calculated. For instance, if it was observed that the users in one census tract had 3 *faith-based* tweets, 4 *community* tweets, and 3 *assets* tweets, that census tract yielded 30 percent *faith-based*, 40 percent *community*, and 30 percent *assets*. These percentages were considered the "Dependent Variables" for Dirichlet's regression model, whereas internal attributes were considered "Independent Variables." Therefore, the Dirichlet regression model's initial data-frame had 1638 rows (one for each census tract from which Disaster Experienced Users had tweeted) and 50 columns (4 columns for dependent variables, and 46 columns for independent variables).

4.9. Step 13 – Dirichlet regression non-disaster experienced users

By the beginning of Step 13, Disaster Experienced Users had been identified, their living quarters locations estimated, their topics of discussion evaluated and the relationships of their topics to regional internal attributes assessed (Step 12). But major questions remained: Did disaster impact the population of non-Disaster Experienced Users, and if so, in which ways did it affect them? These are significant questions to ask, since their answers were unclear based on identified trends, especially in terms of individual direct experience or not with Hurricane Sandy. Hence, it was necessary to compare the Disaster Experienced Users—that is, those whose everyday lives were disrupted by the hurricane—to those not affected by the disaster. Therefore, similar steps were implemented to identify Twitter users who lived in the Circle but did not send any disaster-related tweets. In another words, we extracted the list of users who had at least one tweet from inside the Circle, excluding the list of Disaster Experienced Users. Then, we predicted the location of their living quarters by executing steps 4 through 10. The resulting 62,571 non-Disaster Experienced Users sent a total of 1,147,945 tweets, distributed among 1875 census tracts. Table 1 presents the number of tweets of non-Disaster Experienced Users.

5. Discussion

One interesting finding of this study can be inferred from comparison of the total numbers of tweets sent by Disaster vs. non-Disaster Experienced Users. Indeed, there were around 26,000 Disaster Experienced Users identified with approximately 1.9 million tweets. However, the population of non-Disaster Experienced Users was larger (around 62,000 users) and produced fewer numbers of tweets (around 1.1 million from October 26 2012 to January 1, 2013). It appears that Disaster Experienced Users thought that Hurricane Sandy had somehow affected their everyday lives. They were living in the Circle and they had sent at least one tweet related to Hurricane Sandy. On the other hand, it seems that non-Disaster Experienced Users thought that Hurricane Sandy had not considerably affected their daily lives. Although they were living in the Circle and they sent on average one tweet per three days (during the 65 days following the hurricane), none of their tweets were disaster related. The disproportionate numbers of tweets and people comprising these two populations of Twitter users can be attributed to the fact that people who directly experienced the disaster are deemed to seek more "emotional support" through increased social media activity. Of course, the diverse characteristics of social media users mean that this finding requires more scrutiny due to the populations' disproportionality and the number of tweets needed to authenticate our approach to recognizing Disaster Experienced Users with the keywords introduced by (Guan & Chen, 2014). This higher level of emotional support which is thought to be sought by Disaster Experienced Users moreover conforms to the finding by (Gilbert & Karahalios, 2009) that "emotional support" is the foremost motivation among all social media users.

Damages to physical community *assets* such as local businesses, transportation infrastructure, landscape values, recreational centers, etc. are inevitable consequences of a natural disaster's impact. Even though damages to the physical settings of a community may be the source of either short or long-term distress for all community members (Silove, Steel, & Psychol, 2006), the findings of (Guimaraes, Hefner, & Woodward, 1993; Morris, Grattan, Mayer, & Blackburn, 2013) illustrate how various job sectors react differently to their region's dependencies and level of damage. The immediate flow of billions of dollars into an area for reconstruction, resources, and social capital creates some of the differences in reactions among job sectors. In this regard, our study revealed considerable differences among the reactions of various job sectors with respect to community *assets* damages. As seen in Table 4, in the case of Disaster Experienced Users, higher percentages of tract-level agriculture related career were positively correlated with *assets*. The trend which is reversed for their non-Disaster Experienced counterparts. To a lesser extent, a similar pattern was observed for people involved in public administration-related careers denoting a positive correlation with *assets* among Disaster Experienced Users and no significant correlation among their Non-Disaster Experienced counterparts.

Table 5 presents significant predictors of *community* for Disaster Experienced and non-Disaster Experienced Users. Social capital, social interactions and *community* relationships play a major role in self-efficacy and psychosocial recovery from disasters (Aldrich, 2010; Cox & Perry, 2011). Reestablishment of personal connections with the exterior world, sharing unpleasant memories of an event, and seeking compensation for psychological damages could be motivating social media users after being hit by a natural disaster. One considerable difference,

Table 4
Dirichlet regression results of Assets for Disaster Experienced Users and Non-Disaster Experienced Users.

Internal Attribute	Disaster Experienced Users		Non-Disaster Experienced Users	
	Estimate	Std. Error	Estimate	Std. Error
Age_0_14	−4.05 ***	0.87	−4.37 ***	0.77
Age_45_54	−4.29 ***	0.96	−2.47 ***	0.91
Age_55_69	−3.08 ***	0.86	−3.86 ***	0.88
Age_70_100	−3.02 ***	0.88	−4.11 ***	0.88
Holder_alone	1.06 ***	0.31	1.07 ***	0.31
Unemply_55_99	0.31 *	0.13	0.00	0.12
Unemply_bachelor	−1.11 *	0.53	−0.95 *	0.47
Per_Capita_Income	2.34 ***	0.53	2.45 ***	0.50
Income_50_74	−2.27 ***	0.55	−1.08 *	0.49
Wrk_1_26	−2.57 *	1.09	0.56	0.97
Wrk_27_49	0.22	0.94	−1.83 *	0.87
Wrk_hrs_15_34	0.58	0.85	2.89 ***	0.81
Job_isty_Agric	16.13 **	5.12	−6.76	4.96
Job_isty_Const	1.42 *	0.67	0.94	0.67
Job_isty_Finan	0.05	0.71	1.89 **	0.69
Job_isty_Info	6.97 ***	1.16	4.71 ***	1.12
Job_isty_Prof	2.11 **	0.71	−0.19	0.64
Job_isty_Public	5.26 ***	1.03	1.61	0.97
Mobility_moved	−0.33	0.17	1.53 *	0.68

Significance codes: 0 '***' 0.001 '**' 0.01 '*' 0.05.

Table 5
Dirichlet regression results of Community for Disaster Experienced Users and Non-Disaster Experienced.

Internal Attribute	Disaster Experienced Users		Non-Disaster Experienced Users	
	Estimate	Std. Error	Estimate	Std. Error
Age_0_14	−3.29 ***	0.76	−5.02 ***	0.78
Age_45_54	−3.40 ***	0.93	−2.24 *	0.91
Age_55_69	−1.53	0.85	−2.86 ***	0.86
Age_70_100	−2.79 ***	0.84	−3.69 ***	0.89
Edu_24_college	0.47 *	0.19	−0.43 *	0.18
Holder_alone	1.41 ***	0.33	1.41 ***	0.33
Unemply_16_24	−0.02	0.07	−0.21 **	0.67
Per_Capita_Income	1.27 *	0.57	0.23	0.56
Trans_to_work	1.18 ***	0.19	1.12 ***	0.19
Income_25_49	0.26	0.50	−1.12 *	0.49
Income_50_74	−2.43 ***	0.55	−2.10 ***	0.53
Income_75_99	−1.24	0.64	−2.05 ***	0.61
Wrk_1_26	−2.21 *	1.03	0.33	1.02
Wrk_hrs_15_34	1.16	0.86	3.16 ***	0.80
Wrk_hrs_1_14	3.80 *	1.75	4.75 **	1.84
Job_isty_Const	2.24 ***	0.65	1.32 *	0.65
Job_isty_Finan	0.63	0.73	2.26 **	0.71
Job_isty_Info	7.29 ***	1.16	0.74	1.13
Job_isty_Prof	1.56 *	0.67	0.15	0.65
Job_isty_Public	5.68 ***	1.06	3.79 ***	1.00

Significance codes: 0 '***' 0.001 '**' 0.01 '*' 0.05.

observed in Table 5, is that in the case of Disaster Experienced Users, higher percentages of careers related to information and technology was positively associated with *community*. Meanwhile, the same did not hold true in the case of non-Disaster Experienced Users. This difference can be attributed to the way that IT specialists' everyday lives are more wrapped up in social media applications and they were more captivated by tragic situations presented on social media applications. A reverse pattern for Disaster Experience Users is seen in the sub-attribute of 1 to 26 weeks of employment. It appears that Disaster Experienced counterparts were less likely to tweet words representing *community* however no significant pattern was observed for non-Disaster Experienced Users. Therefore, our findings on Disaster Experienced Users seem to corroborate the unwillingness of unemployed people to engage their surroundings through social interactions observed by others (Rodriguez, Lasch, Chandra, & Lee, 2001).

Faith-based motivations can play a salient role in both individual and communal disaster recovery (Alawiyah, Bell, Pyles, & Runnels, 2011; Cherry et al., 2011). Alawiyah et al. (2011) argue that religious beliefs and faith-based local communities can expedite the emotional healing after a disaster strikes, and so secular service providers need to understand the contribution of religious beliefs to survivors coping strategies. As Alawiyah et al. (2011) observe, religious beliefs and attitudes vary among communities steeped in their own religious beliefs and cultural practices. King, Elder, and Whitbeck (1997) and Lichter and Carmalt (2009) contend, using similar logic, that age and income are just as significant as faith-based motivations. Table 6 presents the differences between tract-level Disaster Experienced and non-Disaster Experienced Users regarding *faith-based* issues. While similar patterns can be observed among the two populations for internal attributes related to unemployment, lonesome householders, per-capita income, transportation and mortgage for *faith-based* issues, contradicting trends are observed for the rest including age, income, work hours and job sectors. More specifically, unemployment appears to have a significant negative impact on *faith-based* issues for Disaster Experienced Users while being insignificant for non-Disaster Experienced Users. This difference is explained by the fact that less vulnerable populations (i.e. high- income, employed, etc.) deliberately express *faith-based* attitudes to bolster the purposefulness of their activities (Sullivan, 2006). Conversely, vulnerable populations (i.e. low-income, unemployed, etc.) may get involved in *faith-based* initiatives simply to obtain social

Table 6
Dirichlet regression results of Faith-based for Disaster Experienced Users and Non-Disaster Experienced.

Internal Attribute	Disaster Experienced Users		Non-Disaster Experienced Users	
	Estimate	Std. Error	Estimate	Std. Error
Age_0_14	−2.77 ***	0.81	−4.24 ***	0.74
Age_45_54	−4.01 ***	0.98	−3.63 ***	0.86
Age_55_69	1.98 *	0.88	−3.01 ***	0.81
Age_70_100	−0.97	0.91	−3.81 ***	0.97
Holder_alone	1.41 ***	0.33	0.95**	0.32
Unemply_bachelor	−1.52 **	0.53	−0.21	0.48
Per_Capita_Income	1.42 *	0.61	1.40 **	0.53
Trans_to_work	0.68 **	0.21	0.68 ***	0.20
Income_50_74	−3.11 ***	0.57	−1.58 **	0.54
Income_75_99	0.20	0.66	−1.91 **	0.64
Wrk_27_49	2.75 **	0.89	−1.93 *	0.89
Wrk_hrs_15_34	1.58	0.88	3.28 ***	0.87
Wrk_hrs_1_14	2.10 **	0.66	2.13	1.70
Job_isty_Const	2.10 **	0.66	1.94**	0.66
Job_isty_Finan	0.07	0.77	1.91 **	0.71
Job_isty_Info	3.45 **	1.15	1.48	1.11
Job_isty_Arts	0.53	0.60	1.49**	0.55
Job_isty_Public	3.91 ***	1.06	4.96***	0.97
Mortgage	−0.54 **	0.18	−0.17	0.17

Significance codes: 0 '***' 0.001 '**' 0.01 '*' 0.05.

assistance (Sullivan, 2006). Meanwhile, findings indicate that in the course of recovery vulnerable populations with disaster experience less commonly resort to *faith-based* values.

Table 7 compares the predictors of disaster between experienced and non-experienced populations regarding financial-related issues. Indeed, *financial* issues can be considered one of the most important determinants of post-disaster recovery of communities. In the case of Disaster Experienced Users, higher percentages of educated people were positively correlated with *financial* related issues. This finding can be interpreted in light of the fact that educated people are more financially knowledgeable (Lusardi & Mitchell, 2011) and following the disaster they become more concerned about their *financial* affairs.

5.1. Contribution to current literature

The main goal of this tract-level analysis is to identify different

Table 7
Dirichlet regression results of Financial for Disaster Experienced Users and Non-Disaster Experienced.

Internal Attribute	Disaster Experienced Users		Non-Disaster Experienced Users	
	Estimate	Std. Error	Estimate	Std. Error
Age_0_14	−3.88 ***	0.76	−4.29 ***	0.79
Age_45_54	−3.16 ***	0.95	−3.42 ***	0.90
Age_55_69	−1.87 *	0.84	−4.48 ***	0.87
Age_70_100	−2.72 **	0.88	−4.08 ***	0.87
Edu_24_college	0.76 ***	0.19	−0.36	0.18
Edu_25_college	1.59 **	0.57	0.18	0.54
Holder_alone	1.71 ***	0.34	1.05 **	0.35
Unemply_55_99	0.11	0.14	−0.25 *	0.13
Trans_to_work	0.57 **	0.21	0.80 ***	0.19
Income_25_49	−1.01 *	0.51	−0.60	0.48
Income_50_74	−2.68 ***	0.55	−1.14 *	0.53
Wrk_27_49	1.20	0.92	−2.61 **	0.93
Wrk_hrs_15_34	0.60	0.89	1.78 *	0.8
Wrk_hrs_1_14	0.38	1.87	4.76 **	1.77
Job_isty_Agric	11.90 *	5.58	−9.92 *	4.69
Job_isty_Finan	−1.17	0.75	3.07 ***	0.73
Job_isty_Info	6.57 ***	1.14	2.99 **	1.15
Job_isty_Public	3.59 ***	1.05	2.36 *	1.00
Mortgage	−0.39 *	0.18	−0.17	0.16

Significance codes: 0 '***' 0.001 '**' 0.01 '*' 0.05.

populations within the community, understand the topics of discussion these populations were engaged in, explore the variations within these topics of discussion based on tract-level internal attributes, and compare the similarities and differences among different populations. Indeed, the proposed multi-step algorithm of this study connected Twitter and American Community Survey databases based on four components: 1) Text mining to distinguish sub-populations (i.e. Disaster Experienced and non-Disaster Experienced Users); 2) location prediction to determine living quarters locations; 3) topic modeling to identify topics of discussion; and 4) statistical modeling to describe and compare population characteristics. However, each of these components of our algorithm has been studied extensively by other researchers, who have suggested several advanced and practical methods in the literature. Basically, sentiment analysis methods have an extensive application in user classification as a means to distinguish sub-populations of Twitter users (Ragini et al., 2018). For instance, Pennacchiotti and Popescu (2011) developed a machine learning algorithm to classify user profiles based on expressed sentiments toward political parties. While this study has focused primarily on analyzing geo-tagged tweets, a large portion of currently available Twitter data consists of non-geo-tagged tweets. Therefore, geo-locating of non-geo-tagged tweets becomes an important issue in analyzing social media data. Singh et al. (2017) introduced an advanced Markov model for location prediction of non-geo-tagged tweets given the historical locations of a user. Beyond sentiment analysis and geo-locating efforts, topic modeling methods have been extensively discussed in the literature. For instance, K. Lee et al. (2011) introduced a high-precision vector-based topic modeling algorithm with the capability of classifying real-time data. While in this study we used Dirichlet regression as the statistical method to describe population characteristics, Nguyen et al. (2016) performed a sentiment analysis on geo-tagged tweets, grouped the tweets based on their located census tracts, and performed a linear regression model to understand the tract-level indicators of psychosocial parameters (e.g. happiness, and diet). As such, this study can be considered as the basic connecting point of four major branches of social media and Twitter studies: user classification, geo-locating, topic modeling, and statistical analysis. Future advances may improve each component of the algorithm.

5.2. Implications for practice

Post-disaster recovery is a complex social process. Definitely, a successful recovery policy requires a deep understanding of multiple social, regional, and global parameters such as level of damage, demographics, communal relationships, economic interactions, as well as several other parameters. However, due to complicated nature of these parameters and their intricate relationships, this understanding becomes highly difficult task especially in post-disaster context. Fortunately, in last decades, post-disaster recovery and its contributing factors have been studied by hundreds of researchers from different perspectives where their endeavors provided prominent outcomes. For instance, Fothergill and Peek (2004) had exceptionally synthesized the current literature about impact of disaster on vulnerable populations and discussed how low-income and low-educated people are susceptible to natural disasters regarding their place of living, social segregation, etc. Moreover, Duval-Diop, Curtis, and Clark (2010) and Cheng, Fu, and de Vreede (2017) discussed how faith-based organizations can provide a unique substrate to enhance political trust and improve public policies which can subsequent the social equity in disaster victims.

On the other hand, day to day growth of social media applications provides an exceptional opportunity for users to seek institutional, functional, and emotional support (Pogrebnyakov & Maldonado, 2018; Qu, Huang, Zhang, & Zhang, 2011), as well as express their opinions, satisfactions, and objections (Taylor, Wells, Howell, & Raphael, 2012). Nowadays, putting together the complexities of post-disaster recovery and superiorities of social media applications encouraged emergency management researchers to understand public opinion via analyzing the social media data. Houston et al. (2015) introduced a framework to better understand the different kinds of social media content producers in the case of disasters. As another example, Shklovski, Palen, and Sutton (2008) and Simon, Goldberg, and Adini (2015) discussed different ways that social media would foster access to information for geographically dispersed disaster impacted communities. Therefore, our study shows the possibilities to understand users' behavior and illuminate impact of disaster on their behavioral pattern.

The most important implications for practice in our study are the identified topic of discussion of Twitter users. Assets, community, faith-based and financial are most common topic of discussion that had widely scrutinized in the literature. Knowledge about perception of damage to community assets can improve the policies to distribute public relief fund (Morris & Wodon, 2003). Also, understating the community and its initiatives can lead to perception of how people support each other in a face of outbreaks (Wright, Ursano, Bartone, & Ingraham, 1990). Moreover, this study highlighted the impact of faith-based related issues on experience of disasters where faith-based related issues are reasonable initiatives to boost social support following the disasters (Phillips & Jenkins, 2010). Chang (2010) raised financial recovery as a most important aspect of disaster recovery which may be dominated by several parameters and may take several years to recover. As, in this study level of education has identified as one of the prominent parameters in concerns about financial recovery.

However, these are many other important factors play role in post-disaster recovery that while we did not discussed in this manuscript, it is possible to identify and scrutinize by our methodology. Housing related issues (Comerio, 1998) and role of insurance coverage (Kunreuther, 1996) as the leading factors in mitigating the impacts of disasters. Moreover, political trust (Han, Hu, & Nigg, 2011) as the factor to measure the effectiveness of public policies following the disasters can be studied by our suggested methodology.

6. Conclusion and future work

Understanding the priorities of people impacted by natural disasters is a vital part of designing efficient post-disaster recovery policies. Complex interrelationships among community features and dependencies of personal preferences on internal attributes of people, as well as the complexity of the consequences of disasters on a community make understating difficult for researchers and policymakers. Fortunately, the diversity of users, daily growth of social media applications and public access provide a unique opportunity for researchers to study the thoughts, opinions, and attitudes of people regarding various subjects. Hence, the extent of mental and physical damage wrought by disasters can been seen reflected in the activities of social media users. The prevalence of seeking "emotional support" as a motive among social media users adds additional importance to this data for post-disaster recovery scholars. Due to the random appearance and vast extent of users and intentions, social media data has been under-utilized in post-disaster recovery studies. However, this study introduces a methodology to systematically analyze Twitter data in post-disaster recovery studies. The output of this methodology can be summarized as identifying Disaster Experienced Users, predicting their living quarters locations, assessing the topics they discuss, evaluating the relationships between their tract-level internal attributes and their tract-level topics of discussion, and comparing the results with non-Disaster Experienced Users.

Interestingly, Disaster Experienced Users were a smaller population, but they were more active and sent more tweets than their non-Disaster Experienced User counterparts. This difference may be due to the greater extent of emotional traumas felt by Disaster Experienced Users, which they unconsciously attempt to heal through social media activities. Though this deduction seems superficial, to a certain extent, it is a topic for further investigation to elucidate the actual intentions behind disaster-influenced activities of social media users.

The analysis of relationships between tract-level internal attributes and the tract-level distribution of the four topics of discussion (*assets*, *community*, *faith-based* and *financial*) revealed various significant differences between Disaster and non-Disaster Experienced Twitter Users. The observed patterns of *assets* supported the findings of (Guimaraes et al., 1993), and (Morris Jr et al., 2013), where post-disaster phase reactions can vary widely based on tract-level percentages of professions and occupations. Additionally, the study highlighted how disaster experienced population are a vulnerable population due to their perspectives about *community* and *faith-based* motivations which differ systematically from their non-disaster experienced counterpart. Analysis of *community* topics of discussion revealed its negative association with unemployment among Disaster Experiences Users (Finch, Emrich, & Cutter, 2010). High and low-income disaster experienced individuals demonstrate distinct perspectives with respect to communal *faith-based* motivations (Sullivan, 2006) where lower income group are less likely to tweet about *faith-based* issues compare to their higher income counterparts. This study finds that the disaster experience may discourage vulnerable populations from expressing *faith-based* motivations in their statements. The comparisons of *faith-based* and *community* topics discussed here show how certain vulnerable populations may react differently after being hit by disasters. Moreover, evaluations of *financial* asset-related tweets reveal that in the wake of a disaster educated people are more concerned about their *financial* affairs.

Although, prior research on geo-tagged Twitter data has produced several valuable methodologies, the multistep methodology introduced in this study can be considered as a new Machine Learning Algorithm in order to detect, compare, and predict the attitudes of different populations of users. For instance, in a political science study, the application of this methodology can reveal advocates of various political views, predict where they live, and foresee their topics of interest by their tract-level internal attributes.

6.1. Limitations and future research directions

The methodology introduced in this paper represents a new approach to the analysis of social media data in post-disaster recovery studies. For the sake of simplicity, this study was confined to four major topics (*assets*, *community*, *faith-based* and *financial*). More detailed disaster related topics are available for future research using the methodology we have presented. For instance, detailed topics related to housing issues, the impact of local rumors, political trust and mental illnesses can be recognized and studied based on this methodology. Also, directionality of opinions can be evaluated by sentiment analysis procedures, which can provide valuable information about how people were negative or positive about discussed topics. Moreover, as it has been discussed in many Geographic Information Systems (GIS) manuals, spatial dependencies of the parameters should be considered in spatial regression models (Anselin, 2009; LeSage, 2008; Ward & Gleditsch, 2008). Hence, developing the Spatial Dirichlet Regression, the model for analyzing spatial compositional data, offers interesting opportunities for future study. Additionally, effects of the Modifiable Areal Unit Problem (MAUP) on the spatial regression models should be evaluated (Openshaw, 1979). As explained by Yang (2005), MAUP is related to grouping actual points using imaginary boundaries which can produce important inaccuracies in spatial regression models and should be accurately evaluated for future studies. Furthermore, the effects of spatial resolution (i.e. block, block group, census tract, county, etc.) and comparisons of the New York City results to other metropolitan areas can be productive topics for further studies. Finally, longitudinal study of topics alteration for pre- and post-disaster periods can shed a light on unknown aspects of social media data and disaster recovery policies.

In conclusion, one must exercise caution in interpreting the results of this study. The first significant variable found is not necessarily the actual dominant variable in the community. Given that we are working with a very large and messy social media dataset with hundreds of millions of lines and in the chaotic post-disaster period, it was a super difficult challenge to find defendable patterns relevant to our research questions. While we are not hopeless regarding possible advancements in the application of social media in post-disaster recovery, the preliminary findings of this study warn of the difficulties ahead.

Disclaimers

Publication of this paper does not necessarily indicate acceptance by the funding entities of its contents, either inferred or specially expressed herein.

Acknowledgements

This research was supported in part by the National Science Foundation award #1454650, for which the authors express their appreciation. Publication of this paper does not necessarily indicate acceptance by the funding entities of its contents, either inferred or specifically expressed herein. Also, the authors would like to thank Mr. Da Hu and Ms. Kellsy Jeanine Macdonald, graduate and undergraduate research assistants at Texas Tech University, for their important contributions to analyzing the data.

References

Abramson, D. M., Stehling-Ariza, T., Park, Y. S., Walsh, L., & Culp, D. (2010). Measuring individual disaster recovery: A socioecological framework. *Disaster Medicine and Public Health Preparedness, 4*(S1), S46–S54.

Aitchison, J. (1982). The statistical analysis of compositional data. *Journal of the Royal Statistical Society Series B (Methodological),* 139–177.

Alawiyah, T., Bell, H., Pyles, L., & Runnels, R. C. (2011). Spirituality and faith-based interventions: Pathways to disaster resilience for African American Hurricane Katrina survivors. *Journal of Religion & Spirituality in Social Work: Social Thought, 30*(3), 294–319.

Aldrich, D. P. (2010). *Fixing recovery: Social capital in post-crisis resilience*.

Anselin, L. (2009). *Spatial regression. The SAGE handbook of spatial analysis, Vol. 1*, 255–276.

Badri, S. A., Asgary, A., Eftekhari, A., & Levy, J. (2006). Post-disaster resettlement, development and change: A case study of the 1990 Manjil earthquake in Iran. *Disasters, 30*(4), 451–468.

Bakshy, E., Hofman, J. M., Mason, W. A., & Watts, D. J. (2011). *Everyone's an influencer: Quantifying influence on twitter. Paper Presented at the proceedings of the fourth ACM international conference on web search and data mining*.

Baxter, M., Beardah, C., Cool, H., & Jackson, C. (2005). Compositional data analysis of some alkaline glasses. *Mathematical Geology, 37*(2), 183–196.

Blei, D. M., Ng, A. Y., & Jordan, M. I. (2003). Latent dirichlet allocation. *Journal of Machine Learning Research, 3*(January), 993–1022.

Bloomberg, M. (2013). *A stronger, more resilient New York*City of New York: PlaNYC Report.

Bolin, R. C., & Bolton, P. A. (1986). *Race, religion, and ethnicity in disaster recovery*.

Bollen, J., Mao, H., & Zeng, X. (2011). Twitter mood predicts the stock market. *Journal of Computational Science, 2*(1), 1–8.

Brown, B. B., & Perkins, D. D. (1992). *Disruptions in place attachment. Place attachment.* Springer279–304.

Cao, G., Wang, S., Hwang, M., Padmanabhan, A., Zhang, Z., & Soltani, K. (2015). A scalable framework for spatiotemporal analysis of location-based social media data. *Computers, Environment and Urban Systems, 51*, 70–82.

Chang, S. E. (2010). Urban disaster recovery: A measurement framework and its application to the 1995 Kobe earthquake. *Disasters, 34*(2), 303–327.

Cheng, X., Fu, S., & de Vreede, G.-J. (2017). Understanding trust influencing factors in social media communication: A qualitative study. *International Journal of Information Management, 37*(2), 25–35.

Cherry, K. E., Brown, J. S., Marks, L. D., Galea, S., Volaufova, J., Lefante, C., ... Jazwinski, S. M. (2011). Longitudinal assessment of cognitive and psychosocial functioning after Hurricanes Katrina and Rita: Exploring disaster impact on middle-aged, older, and oldest-old adults. *Journal of Applied Biobehavioral Research, 16*(3–4), 187–211.

Comerio, M. C. (1998). *Disaster hits home: New policy for urban housing recovery*. University of California Press.

Cox, R. S., & Perry, K.-M. E. (2011). Like a fish out of water: Reconsidering disaster recovery and the role of place and social capital in community disaster resilience. *American Journal of Community Psychology, 48*(3–4), 395–411.

Cutter, S. L., Boruff, B. J., & Shirley, W. L. (2003). Social vulnerability to environmental hazards. *Social Science Quarterly, 84*(2), 242–261.

De Choudhury, M., Gamon, M., Counts, S., & Horvitz, E. (2013). Predicting depression via social media. *ICWSM, 13*, 1–10.

Dhir, A., Kaur, P., & Rajala, R. (2018). Why do young people tag photos on social networking sites? Explaining user intentions. *International Journal of Information Management, 38*(1), 117–127.

Dong, X., & Wang, T. (2018). Social tie formation in Chinese online social commerce: The role of IT affordances. *International Journal of Information Management, 42*, 49–64.
Drury, A. C., & Olson, R. S. (1998). Disasters and political unrest: An empirical investigation. *Journal of Contingencies and Crisis Management, 6*(3), 153–161.
Duval-Diop, D., Curtis, A., & Clark, A. (2010). Enhancing equity with public participatory GIS in hurricane rebuilding: Faith based organizations, community mapping, and policy advocacy. *Community Development, 41*(1), 32–49.
Finch, C., Emrich, C. T., & Cutter, S. L. (2010). Disaster disparities and differential recovery in New Orleans. *Population and Environment, 31*(4), 179–202.
Force, H. S. R. T. (2013). *Hurricane sandy rebuilding strategy.* Washington DC: US Department of Housing and Urban Development.
Fothergill, A., & Peek, L. A. (2004). Poverty and disasters in the United States: A review of recent sociological findings. *Natural Hazards, 32*(1), 89–110.
Fraustino, J. D., Liu, B., & Jin, Y. (2012). *Social media use during disasters: A review of the knowledge base and gaps*.
Gao, H., Barbier, G., & Goolsby, R. (2011). Harnessing the crowdsourcing power of social media for disaster relief. *IEEE Intelligent Systems, 26*(3), 10–14.
Gilbert, E., & Karahalios, K. (2009). *Predicting tie strength with social media. Paper presented at the proceedings of the SIGCHI conference on human factors in computing systems*.
Glasgow, K., Vitak, J., Tausczik, Y., & Fink, C. (2016). With your help… We begin to heal. *Paper presented at the social, cultural, and behavioral modeling: 9th international conference, SBP-BRiMS 2016*.
Gleser, G. C., Green, B. L., & Winget, C. (2013). *Prolonged psychosocial effects of disaster: A study of Buffalo Creek.* Elsevier.
Grover, P., Kar, A. K., & Davies, G. (2018). Technology enabled Health"—Insights from twitter analytics with a socio-technical perspective. *International Journal of Information Management, 43*, 85–97.
Guan, X., & Chen, C. (2014). Using social media data to understand and assess disasters. *Natural Hazards, 74*(2), 837–850.
Gueorguieva, R., Rosenheck, R., & Zelterman, D. (2008). Dirichlet component regression and its applications to psychiatric data. *Computational Statistics & Data Analysis, 52*(12), 5344–5355.
Guimaraes, P., Hefner, F. L., & Woodward, D. P. (1993). Wealth and income effects of natural disasters: An econometric analysis of Hurricane Hugo. *The Review of Regional Studies, 23*(2), 97.
Han, B., Cook, P., & Baldwin, T. (2014). Text-based twitter user geolocation prediction. *Journal of Artificial Intelligence Research, 49*, 451–500.
Han, Z., Hu, X., & Nigg, J. (2011). How does disaster relief works affect the trust in local government? A study of the Wenchuan earthquake. *Risk, Hazards & Crisis in Public Policy, 2*(4), 1–20.
Hickey, C., Kelly, S., Carroll, P., & O'Connor, J. (2015). Prediction of forestry planned end products using dirichlet regression and neural networks. *Forest Science, 61*(2), 289–297.
Hidalgo, M. C., & Hernandez, B. (2001). Place attachment: Conceptual and empirical questions. *Journal of Environmental Psychology, 21*(3), 273–281.
Hijazi, R. H., & Jernigan, R. W. (2009). Modelling compositional data using Dirichlet regression models. *Journal of Applied Probability and Statistics, 4*(1), 77–91.
Houston, J. B., Hawthorne, J., Perreault, M. F., Park, E. H., Goldstein Hode, M., Halliwell, M. R., ... McElderry, J. A. (2015). Social media and disasters: A functional framework for social media use in disaster planning, response, and research. *Disasters, 39*(1), 1–22.
Hughes, A. L., Palen, L., Sutton, J., Liu, S. B., & Vieweg, S. (2008). *Site-seeing in disaster: An examination of on-line social convergence. Paper presented at the proceedings of the 5th international ISCRAM conference*.
InternetLiveStats (2017). *Twitter usage statistics.* Retrieved fromhttp://www.internetlivestats.com/twitter-statistics/.
Ivanova, E., Maier, M., & Meyer, M. (2016). *Associations in transition: The business of Russian Civil Society*.
Jamali, M., & Nejat, A. (2016). Place attachment and disasters: Knowns and unknowns. *Journal of Emergency Management (Weston, Mass.), 14*(5), 349–364.
Java, A., Song, X., Finin, T., & Tseng, B. (2009). *Why we twitter: An analysis of a microblogging community. Advances in web mining and web usage analysis*. Springer118–138.
Jeong, B., Yoon, J., & Lee, J.-M. (2017). Social media mining for product planning: A product opportunity mining approach based on topic modeling and sentiment analysis. *International Journal of Information Management*. https://doi.org/10.1016/j.ijinfomgt.2017.09.009 in-press.
Junco, R., Heiberger, G., & Loken, E. (2011). The effect of Twitter on college student engagement and grades. *Journal of Computer Assisted Learning, 27*(2), 119–132.
Kaniasty, K., & Norris, F. H. (1993). A test of the social support deterioration model in the context of natural disaster. *Journal of Personality and Social Psychology, 64*(3), 395.
Kaplan, A. M., & Haenlein, M. (2010). Users of the world, unite! The challenges and opportunities of social media. *Business Horizons, 53*(1), 59–68.
Kapoor, K. K., Tamilmani, K., Rana, N. P., Patil, P., Dwivedi, Y. K., & Nerur, S. (2018). Advances in social media research: Past, present and future. *Information Systems Frontiers, 20*(3), 531–558.
Kaufman, S. M., Qing, C., Levenson, N., & Hanson, M. (2012). *Transportation during and after hurricane sandy*.
Kim, J., & Hastak, M. (2018). Social network analysis: Characteristics of online social networks after a disaster. *International Journal of Information Management, 38*(1), 86–96.
Kim, J., Bae, J., & Hastak, M. (2018). Emergency information diffusion on online social media during storm Cindy in US. *International Journal of Information Management, 40*, 153–165.
King, V., Elder, G. H., Jr., & Whitbeck, L. B. (1997). Religious involvement among rural youth: An ecological and life-course perspective. *Journal of Research on Adolescence, 7*(4), 431–456.
Kunreuther, H. (1996). Mitigating disaster losses through insurance. *Journal of Risk and Uncertainty, 12*(2-3), 171–187.
Lee, J., & Hong, I. B. (2016). Predicting positive user responses to social media advertising: The roles of emotional appeal, informativeness, and creativity. *International Journal of Information Management, 36*(3), 360–373.
Lee, K., Palsetia, D., Narayanan, R., Patwary, M. M. A., Agrawal, A., & Choudhary, A. (2011). *Twitter trending topic classification. paper presented at the data mining workshops (ICDMW), 2011 IEEE 11th international conference on*.
Leetaru, K., Wang, S., Cao, G., Padmanabhan, A., & Shook, E. (2013). Mapping the global Twitter heartbeat: The geography of Twitter. *First Monday, 18*(5).
LeSage, J. P. (2008). An introduction to spatial econometrics. *Revue d'économie industrielle,* (3), 19–44.
Lewicka, M. (2005). Ways to make people active: The role of place attachment, cultural capital, and neighborhood ties. *Journal of Environmental Psychology, 25*(4), 381–395.
Li, L., Zhang, Q., Tian, J., & Wang, H. (2018). Characterizing information propagation patterns in emergencies: A case study with Yiliang earthquake. *International Journal of Information Management, 38*(1), 34–41.
Lichter, D. T., & Carmalt, J. H. (2009). Religion and marital quality among low-income couples. *Social Science Research, 38*(1), 168–187.
Lindell, M. K., & Prater, C. S. (2003). Assessing community impacts of natural disasters. *Natural Hazards Review, 4*(4), 176–185.
Lindsay, B. R. (2011). *Social media and disasters: Current uses, future options, and policy considerations.* Washington, DC: Congressional Research Service.
Lipizzi, C., Iandoli, L., & Marquez, J. E. R. (2015). Extracting and evaluating conversational patterns in social media: A socio-semantic analysis of customers' reactions to the launch of new products using Twitter streams. *International Journal of Information Management, 35*(4), 490–503.
Liu, L., Lee, M. K., Liu, R., & Chen, J. (2018). Trust transfer in social media brand communities: The role of consumer engagement. *International Journal of Information Management, 41*, 1–13.
Liu, Q., Shao, Z., & Fan, W. (2018). The impact of users' sense of belonging on social media habit formation: Empirical evidence from social networking and microblogging websites in China. *International Journal of Information Management, 43*, 209–223.
Lusardi, A., & Mitchell, O. S. (2011). Financial literacy around the world: An overview. *Journal of Pension Economics & Finance, 10*(4), 497–508.
Maldonado, M., Alulema, D., Morocho, D., & Proaño, M. (2016). *System for monitoring natural disasters using natural language processing in the social network twitter. Paper presented at the security technology (ICCST), 2016 IEEE international carnahan conference on*.
Martínez-Rojas, M., del Carmen Pardo-Ferreira, M., & Rubio-Romero, J. C. (2018). Twitter as a tool for the management and analysis of emergency situations: A systematic literature review. *International Journal of Information Management, 43*, 196–208.
Masozera, M., Bailey, M., & Kerchner, C. (2007). Distribution of impacts of natural disasters across income groups: A case study of New Orleans. *Ecological Economics, 63*(2), 299–306.
Mehrotra, R., Sanner, S., Buntine, W., & Xie, L. (2013). *Improving lda topic models for microblogs via tweet pooling and automatic labeling. Paper presented at the proceedings of the 36th international ACM SIGIR conference on research and development in information retrieval*.
Morris, S. S., & Wodon, Q. (2003). The allocation of natural disaster relief funds: Hurricane Mitch in Honduras. *World Development, 31*(7), 1279–1289.
Morris, J. G., Jr., Grattan, L. M., Mayer, B. M., & Blackburn, J. K. (2013). Psychological responses and resilience of people and communities impacted by the deepwater horizon oil spill. *Transactions of the American Clinical and Climatological Association, 124*, 191.
Nakagawa, Y., & Shaw, R. (2004). Social capital: A missing link to disaster recovery. *International Journal of Mass Emergencies and Disasters, 22*(1), 5–34.
Nejat, A., Brokopp Binder, S., Greer, A., & Jamali, M. (2018). Demographics and the dynamics of recovery: A latent class analysis of disaster recovery priorities after the 2013 Moore, Oklahoma Tornado. *International Journal of Mass Emergencies and Disasters, 36*(1).
Nguyen, Q. C., Kath, S., Meng, H.-W., Li, D., Smith, K. R., VanDerslice, J. A., ... Li, F. (2016). Leveraging geotagged Twitter data to examine neighborhood happiness, diet, and physical activity. *Applied Geography, 73*, 77–88.
Nisar, T. M., Prabhakar, G., & Patil, P. P. (2018). Sports clubs' use of social media to increase spectator interest. *International Journal of Information Management, 43*, 188–195.
Okuyama, Y., & Chang, S. E. (2013). *Modeling spatial and economic impacts of disasters.* Springer Science & Business Media.
Openshaw, S. (1979). *A million or so correlation coefficients: Three experiments on the modifiable areal unit problem. Spatistica applications in the spatial sciences*127–144.
Pak, A., & Paroubek, P. (2010). *Twitter as a corpus for sentiment analysis and opinion mining. Paper presented at the LREc*.
Pennacchiotti, M., & Popescu, A.-M. (2011). *Democrats, republicans and starbucks afficionados: User classification in twitter. Paper presented at the proceedings of the 17th ACM SIGKDD international conference on knowledge discovery and data mining*.
Perilla, J. L., Norris, F. H., & Lavizzo, E. A. (2002). Ethnicity, culture, and disaster response: Identifying and explaining ethnic differences in PTSD six months after Hurricane Andrew. *Journal of Social and Clinical Psychology, 21*(1), 20–45.
Phillips, B., & Jenkins, P. (2010). *The roles of faith-based organizations after Hurricane Katrina*.
Pogrebnyakov, N., & Maldonado, E. (2018). Didn't roger that: Social media message complexity and situational awareness of emergency responders. *International Journal of Information Management, 40*, 166–174.
Qu, Y., Huang, C., Zhang, P., & Zhang, J. (2011). *Microblogging after a major disaster in*

China: A case study of the 2010 Yushu earthquake. Paper presented at the proceedings of the ACM 2011 conference on computer supported cooperative work.

Quarantelli, E. L. (1999). *The disaster recovery process: What we know and do not know from research*.

Ragini, J. R., Anand, P. R., & Bhaskar, V. (2018). Big data analytics for disaster response and recovery through sentiment analysis. *International Journal of Information Management, 42*, 13–24.

Rhodes, J., Chan, C., Paxson, C., Rouse, C. E., Waters, M., & Fussell, E. (2010). The impact of hurricane Katrina on the mental and physical health of low-income parents in New Orleans. *American Journal of Orthopsychiatry, 80*(2), 237.

Rodriguez, E., Lasch, K. E., Chandra, P., & Lee, J. (2001). Family violence, employment status, welfare benefits, and alcohol drinking in the United States: what is the relation? *Journal of Epidemiology & Community Health, 55*(3), 172–178.

Rubinstein, R. I., & Parmelee, P. A. (1992). *Attachment to place and the representation of the life course by the elderly. Place attachment*. Springer139–163.

Sapat, A., & Esnard, A. M. (2012). Displacement and disaster recovery: Transnational governance and socio-legal issues following the 2010 Haiti earthquake. *Risk Hazards & Crisis in Public Policy, 3*(1), 1–24.

Shklovski, I., Palen, L., & Sutton, J. (2008). *Finding community through information and communication technology in disaster response. Paper presented at the proceedings of the 2008 ACM conference on computer supported cooperative work*.

Silove, D., Steel, Z., & Psychol, M. (2006). Understanding community psychosocial needs after disasters: Implications for mental health services. *Journal of Postgraduate Medicine, 52*(2), 121.

Simon, T., Goldberg, A., & Adini, B. (2015). Socializing in emergencies—A review of the use of social media in emergency situations. *International Journal of Information Management, 35*(5), 609–619.

Singh, J. P., Dwivedi, Y. K., Rana, N. P., Kumar, A., & Kapoor, K. K. (2017). Event classification and location prediction from tweets during disasters. *Annals of Operations Research*, 1–21.

Smith, S. K., & McCarty, C. (1996). Demographic effects of natural disasters: A case study of Hurricane Andrew. *Demography, 33*(2), 265–275.

Steinglass, P., & Gerrity, E. (1990). Natural disasters and post-traumatic stress disorder short-term versus long-term recovery in two disaster-affected communities. *Journal of Applied Social Psychology, 20*(21), 1746–1765.

Stieglitz, S., Mirbabaie, M., Ross, B., & Neuberger, C. (2018). Social media analytics—Challenges in topic discovery, data collection, and data preparation. *International Journal of Information Management, 39*, 156–168.

Sullivan, S. C. (2006). The work-faith connection for low-income mothers: A research note. *Sociology of Religion, 67*(1), 99–108.

Sutton, J., Spiro, E. S., Johnson, B., Fitzhugh, S., Gibson, B., & Butts, C. T. (2014). Warning tweets: Serial transmission of messages during the warning phase of a disaster event. *Information, Communication & Society, 17*(6), 765–787.

Taylor, M., Wells, G., Howell, G., & Raphael, B. (2012). The role of social media as psychological first aid as a support to community resilience building. *Australian Journal of Emergency Management, 27*(1), 20.

Tierney, K., & Oliver-Smith, A. (2012). Social dimensions of disaster recovery. *International Journal of Mass Emergencies & Disasters, 30*(2).

Toya, H., & Skidmore, M. (2014). Do natural disasters enhance societal trust? *Kyklos, 67*(2), 255–279.

Tumasjan, A., Sprenger, T. O., Sandner, P. G., & Welpe, I. M. (2010). Predicting elections with twitter: What 140 characters reveal about political sentiment. *ICWSM, 10*(1), 178–185.

Twitter (2016). *Twitter terms of service.* Retrieved fromsupport.twitter.com.

Verma, S., Vieweg, S., Corvey, W. J., Palen, L., Martin, J. H., Palmer, M., ... Anderson, K. M. (2011). *Situational awareness. Paper presented at the ICWSM*.

Ward, M. D., & Gleditsch, K. S. (2008). *Spatial regression models*, *Vol. 155*. Sage.

Wold, G. H. (2006). Disaster recovery planning process. *Disaster Recovery Journal, 5*(1), 10–15.

Wright, K. M., Ursano, R. J., Bartone, P. T., & Ingraham, L. H. (1990). The shared experience of catastrophe: An expanded classification of the disaster community. *The American Journal of Orthopsychiatry, 60*(1), 35–42.

Yang, T.-C. (2005). Modifiable areal unit problem. *GIS Resource Document, 5*, 65.

Yates, D., & Paquette, S. (2010). Emergency knowledge management and social media technologies: A case study of the 2010 Haitian earthquake. *Paper presented at the proceedings of the 73rd ASIS&T annual meeting on navigating streams in an information ecosystem, Vol. 47*.

Zhao, L., Chen, F., Dai, J., Hua, T., Lu, C.-T., & Ramakrishnan, N. (2014). Unsupervised spatial event detection in targeted domains with applications to civil unrest modeling. *PLoS One, 9*(10), e110206.