A Literature Review on Big Data and Social Media Usage in Disaster Management

Türkay DERELİ[1], Nazmiye ÇELİK[2] ve Cihan ÇETİNKAYA[3]

Abstract

Most of the disaster management activities are naturally related to traditional operation research and management science applications. But recently, big data information technology and social media in particular has become an integral part of disaster management. Relevant information taken from social media and the intelligent web has increased the situational awareness of decision makers. Disaster management decisions have important impacts on; the safety of disaster victims, environment, economic systems, organizations etc. Reliable, timely, consistent, sufficient and qualified information is critical in the phases of disaster management. In this study, a literature review is conducted considering big data and social media in the light of disaster management and specifically disaster relief.

Keywords: Big Data, Disaster Management, Disaster Relief, Literature Review, Social Media.

Afet Yönetiminde Büyük Veri ve Sosyal Medya Kullanımı Üzerine Literatür Taraması

Özet

Afet yönetimi ile ilgili faaliyetlerin çoğu asıl olarak geleneksel yöneylem araştırması ve yönetim bilimi uygulamalarına dayanmaktadır. Ancak son zamanlarda, büyük veri bilgi teknolojisi ve bunun özelinde sosyal medya kullanımı afet yönetiminin ayrılmaz bir parçası olmuştur. Sosyal medyadan ve akıllı ağlardan alınan ilgili bilgiler karar vericilerin durumsal farkındalığını arttırmaktadır. Afet yönetimi kararları afetzedelerin güvenliği, çevre, ekonomi sistemi, organizasyonlar vb. üzerinde önemli etkilere sahiptir. Güvenilir, zamanında, istikrarlı, yeterli ve nitelikli bilgi afet yönetiminin aşamalarında kritik rol oynamaktadır. Bu çalışmada, afet yönetimi ve özellikle afet yardımı alanında büyük veri ve sosyal medya dikkate alınarak bir literatür çalışması gösterilmiştir.

Anahtar Kelimeler: Büyük Veri, Afet Yönetimi, Afet Yardımı, Literatür Taraması, Sosyal Medya.

1. INTRODUCTION

[1] Office of the President, Iskenderun Technical Uni. & Prof. Dr., Gaziantep Uni. Dept. of Industrial Engineering. Gaziantep
[2] Res. Assist, Gaziantep University, Department of Industrial Engineering, Gaziantep
İlgili yazar / Corresponding author: nazmiye@gantep.edu.tr
[3] Assoc. Prof. Dr., Adana Science and Technology University, Department of Management Information Systems

According to the International Federation of Red Cross and Red Crescent Societies (IFRC), a disaster is "a sudden, calamitous event that seriously disrupts the functioning of a community or society and causes human, material and economic or environmental losses that exceed the community's or society's ability to cope using its own resources" (IFRC, 2018). Disasters potentially cause devastating economic losses, large affected communities and significant environmental damages (Galindo and Batta, 2013). Thus; a sensitive emergency management should be applied to minimize these effects. Information technology plays a significant role for minimizing disaster effects. Recently big data information technology has become an integral part of disaster management. Big data facilitates process automation, decision making, and allows enhanced insight. It is high-volume, high-velocity and/or high-variety information values that demand cost-effective, innovative forms of information processing (Su, 2018). In this paper, a literature review is done and some specific search syntaxes are used to determine studied subjects in literature. There are important literature studies integrated big data and disaster management together but proposed study also considers social media specifically that is part of unstructured big data. The role of social media for disaster management is important and which studies have been done in this area is the significant question to determine future gaps. So, effects of social media on disaster relief and disaster management are shown in this paper. The remainder of the paper as follows: section 2 outlines the social networks statistics, section 3 describes research methodology and literature review. Section 4 is the conclusion section.

2.SOCIAL MEDIA DATA SOURCES

Social network data is one of the big data sources generated from the social media platforms such as YouTube, Facebook, Twitter, LinkedIn, and Flickr. Social media analytics is being utilized by organizations for gaining insight into big data.

Penetration of leading social networks in Turkey during 3rd and 4th quarter 2017 is shown in Figure 1.
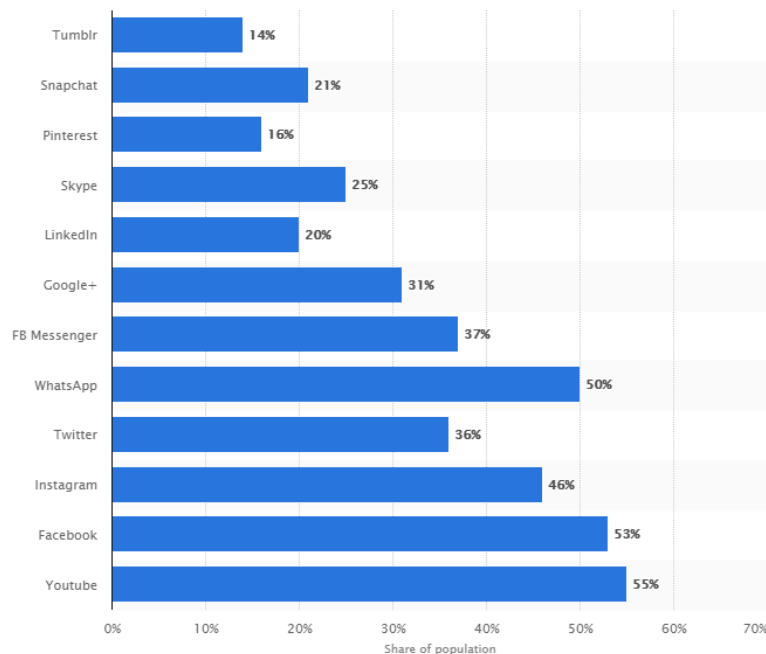

Figure 1. Social networks user counts in Turkey (Statista, 2018)

As shown in the figure, the most used social network is YouTube with a 55 percent penetration rate. Twitter has 36 percent penetration rate that is significant rate for a micro blogging.

Internet, social media, mobile use statistics in Turkey is shown in Figure 2.
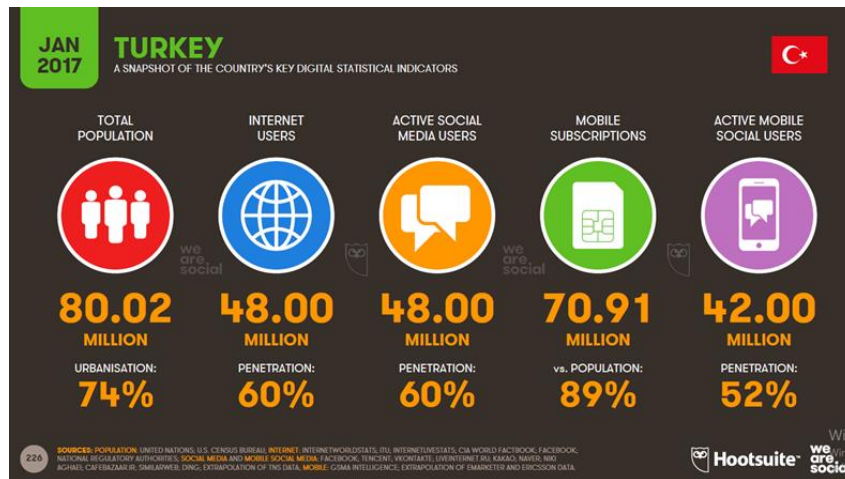

Figure 2. Statistics of internet, social media, and mobile subscriptions (2017 Digital Yearbook)

It is clearly seen that using of internet is quite common in Turkey.

Twitter is one of the most popular websites which is used for micro blogging. The user can write thoughts limited 140 words and other people can follow the blogs. Countries based on number of twitter users are shown in figure 3.
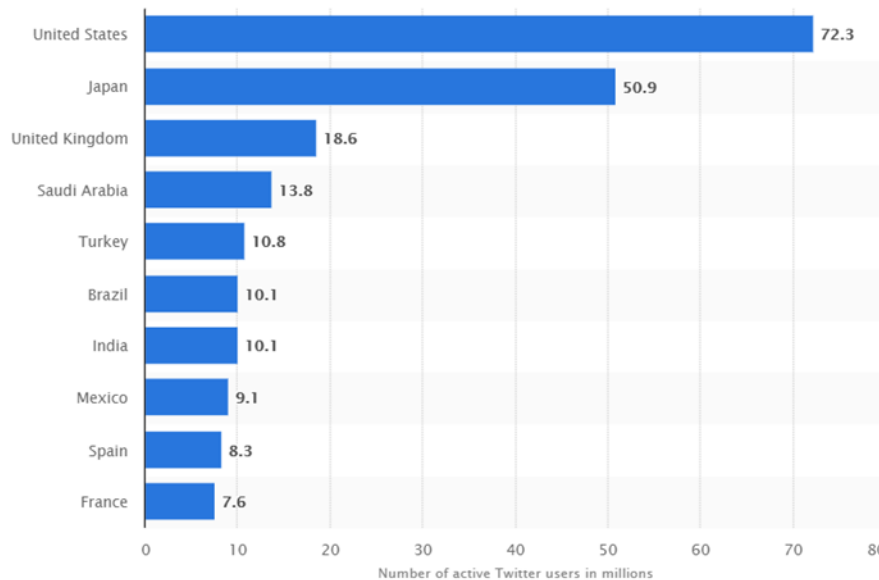

Figure 3. Leading countries based on number of Twitter users as of April 2018 (in millions) (Statista, 2018)

During this time interval, twitter had 10.8 million monthly active users in Turkey.

3. RESEARCH METHODOLOGY AND LITERATURE REVIEW
The literature review has been divided into parts as follows:

3.1 Searching based on "Big Data" String
The sources of this study are the Scopus and Web of Science databases because of being largest databases for academic papers. The first search syntax on Scopus is shown in Table 1.

Table 1. First search strings on Scopus

Database	Search strings
Scopus (May 09,2018)	(TITLE-ABS-KEY ("big data") OR TITLE-ABS-KEY ("large-scale data") OR TITLE-ABS-KEY ("predictive analy") AND TITLE-ABS-KEY ("disaster relief") OR TITLE-ABS-KEY ("humanitarian supply chain") OR TITLE-ABS-KEY ("aid distribution") OR TITLE-ABS-KEY ("disaster help"))

First of all, two main titles are considered. The first title is "big data" that includes 'big data', 'large-scale data', and 'predictive analy' words. The latter is "humanitarian supply chain" that includes 'disaster relief', 'humanitarian supply chain', 'aid distribution' and 'disaster help' words. Search is resulted in 18 papers; 8 of them are conference papers, 5 of them are articles, 2 of them are conference reviews, one of them is article in press and 1 of them is a book chapter. "Predictive analy" is used because of including analysis, analyses, and analytic keywords (Gupta vd., 2017).

The second syntax on Web of Science is shown in Table 2.

Table 2. Search strings on Web of Science

Database	Search strings
Web of Science(May 09,2018)	TS=("big data" OR "large-scale data" OR "predictive analy") AND TS=("disaster relief" OR "humanitarian supply chain" OR "aid distribution" OR "disaster help")

Same strings are used in Web of Science as in the Scopus database. 9 results are handled. 5 of them are articles and 4 of them are conference proceedings.

Only "disaster" search string is used for the second main title to search for a general point of view. The third syntax on Web of Science is shown in Table 3.

Table 3. Comprehensive search on Web of Science

Database	Search strings
Web of Science (March 19,2018)	(TS=("Big data" OR "large-scale data") AND TS="disaster")

Search is resulted in 167 papers. 92 of them are proceeding papers, 69 of them are articles, 5 of them are reviews, 1 of them is a book review, and 1 of them is an editorial material. Flow diagram of first phase search is seen in Figure 4.

Some of important studies in the light of above mentioned strings are as follows:
Monaghan & Lycett (2013) discussed the big data phenomenon with its characteristics of volume, velocity, variety, and veracity in humanitarian supply networks and emphasized the distinction between humanitarian aid and development aid. Humanitarian aid efforts and impact of delivering humanitarian relief was assessed in the light of big data techniques and technologies.
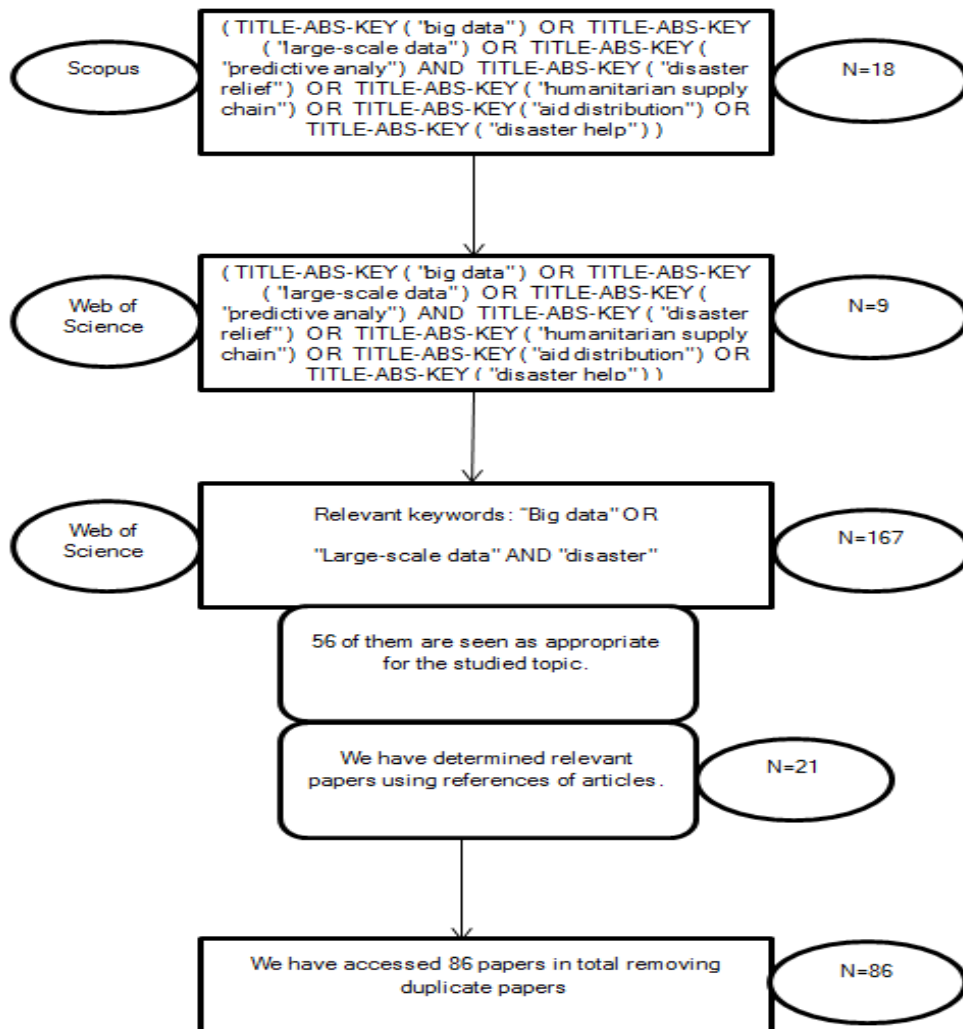
Figure 4. Flow diagram for the first search phase

Papadopulos, et al.(2016) proposed a theoretical framework to indicate the role of big data in disaster resilience for supply chains. They used unstructured big data included tweets, Google+, YouTube, news, Facebook, WordPress and also structured data involved in disaster relief activities after Nepal Earthquake in 2015. They examined the role of swift trust and quality information in supply chain networks unlike other studies. 36,422 tweets were collected for analysis and also 205 responses collected via disaster relief workers. They paid attention to the importance of public-private partnerships, swift trust, information quality, critical infrastructure resilience, community resilience and resource resilience using context of Nepal.

Cherichi & Larodec (2016) proposed an event detection system based on data obtained from twitter. The aim of this system is determine new events, recognize temporal indicators and classify significant events using big data values in the case of social events. Semantic analysis was done into two clusters: positive class and negative class.

Mulder , et al. (2016) proposed the processes of "big data making" in their papers through crowdsourcing open data platforms, within the framework of two crises: 2010 earthquake in Haiti and the 2015 earthquake in Nepal. The study included humanitarian response. Aim of the study was to discover what obstacles stood in the way of "big data making" approach.

Gupta, Altay, & Luo (2017) proposed a systematic literature review study using Scopus database. They examined big data in humanitarian supply chain management using 28 journal papers after eliminating unnecessary studies. Classification was done according to the various theories such as the enablers for big data in humanitarian supply chain management (volume, variety, velocity, veracity organizational mindfulness), the concerns identified in humanitarian supply chain management (humanitarian logistics, remote sensing, information security, social media). The future research directions were suggested in above mentioned fields.

3.2 Searching based on "Social Media" String

In another stage, a new search has been done. Social media is a part of unstructured big data. The studied area is narrowed and focused on social media. Different strings are used on the biggest database Scopus for providing new results. So, based on these strings a new methodology search is conducted. The fourth syntax on Scopus is shown in Table 4.

Table 4. Search strings on Scopus based on social media

Database	Search strings
Scopus (May 09,2018)	(TITLE-ABS-KEY ("social media") AND TITLE-ABS-KEY ("humanitarian supply chain") OR TITLE-ABS-KEY ("disaster relief") OR TITLE-ABS-KEY ("aid distribution") OR TITLE-ABS-KEY ("emergency management") AND NOT TITLE-ABS-KEY ("predictive analy") AND NOT TITLE-ABS-KEY ("big data") AND NOT TITLE-ABS-KEY ("large-scale data"))

'Social media', 'humanitarian supply chain', 'disaster relief', 'aid distribution', and 'emergency management' string words are used for analysis removing 'big data', 'large-scale data', and 'predictive analy' string words. Search resulted in 306 papers; 137 of them are conference papers, 122 of them are articles, 16 of them are book chapters, 9 of them are articles in press and 9 of them reviews. Subject area of these papers is shown in Figure 5.
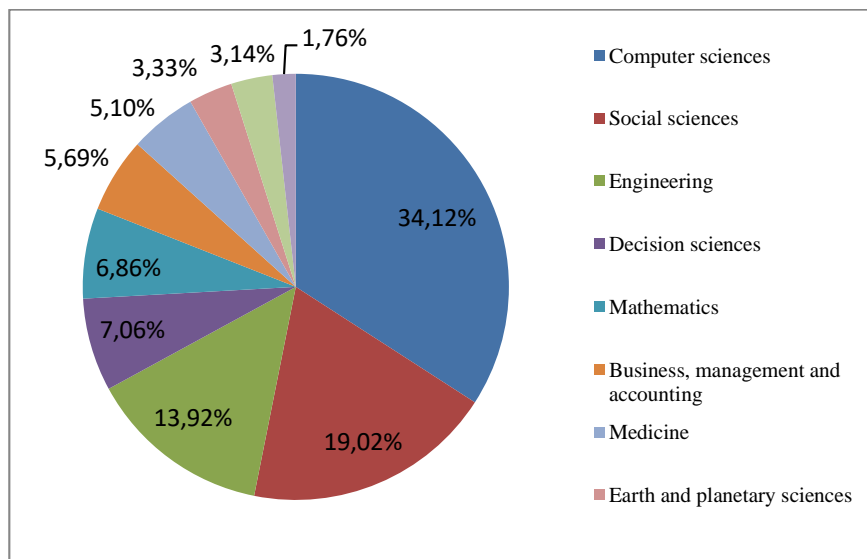

Figure 5. Subject area of 306 papers on Scopus based on social media

As seen in the Figure 5, proportion of the studies based on social media and humanitarian fields conducted within computer sciences subject area are greater than other areas. Social sciences and

engineering fields respectively follow computer science area. 223 papers can be accessed among 306 papers. Among accessed studies, 159 papers include disaster studies based on social media. 80 papers are determined among these 159 papers which used twitter platform for their studies. Word cloud of studies based on twitter using their abstracts is seen in Figure 6. At first glance, 'social', 'twitter', and 'response' words are attracted the attention. In Figure 7, word frequencies of word cloud is seen.

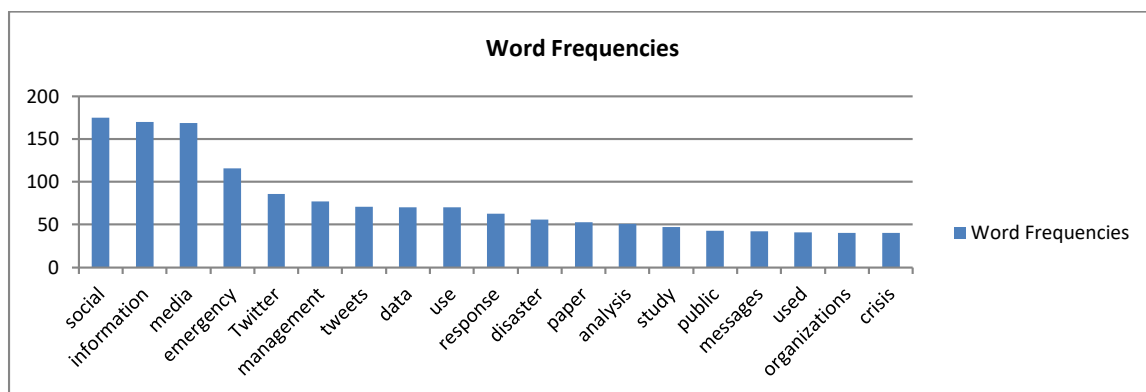*Figure* 6. Word cloud of studies based on twitter

Figure 7. Word frequencies of word cloud

As can be seen in figure 7, social and information words are the most frequently used ones in studies. When graphic is evaluated, response, crisis, and analysis words are striking among first twenty words.

Some of important studies related social media context are as follows:

Caragea, et al. (2011) proposed a classification study for the Haiti Earthquake. They developed an information technology infrastructure, called "Enhanced Messaging for the Emergency Response Sector" (EMERSE). EMERSE includes a twitter crawler component, an iPhone application, machine translation and automatic message classification. The data set contained 3,598 text messages and manually labeled into 10 categories: medical emergency, food shortage, water sanitation, collapsed structured, hospital/clinic services, people trapped, water shortage, shelter needed, food distribution, and person news.

Landwehr, et al. (2016) proposed strength and weakness of twitter to determine the features needed in a twitter system to assist disaster phases: planning, warning and response. They used a specific implementation, called TWRsms within a local context, Padang Indonesia in order to

support early warning and planning tsunamis. The study resulted around some goals such as: scalable to multiple types of social media and data volume, data survivability, interoperable with existing and proposed systems, rapid processing for material useful etc.

Wang & Ye (2018) presented a social media analytic using four dimensions of social media (space, time, content, and network) within natural disasters subject. They reviewed existing studies and evaluated these study analyzes on four dimensions, summarized extensive techniques and suggested research pros and cons in mining social media data.

Taking into consideration the years 2017 and 2018, some of the disaster related studies based on twitter platform are shown in the Table 5.

Table 5. Some of twitter related papers published between 2017-2018 years on Scopus

Author, Year	Content	Method	Disaster type
(Lai, 2016)	The argument for a systemic understanding of technology use for event-triggered collective action based on citizen response organizations. Twitter was used consistently for diverse purposes across the phases of the disaster response, whereas Facebook usage among organizations first diverged and then converged two months after the hurricane.	Multi-method and longitudinal case study (technological affordances)	Hurricane Sandy
(Alshareef & Grigoras, 2016)	Proposed a mobile cloud service that allows users to provide on-the-ground information regarding such an event, as well as early notification to people who are in the vicinity of the location of an emergency situation. The system matches users' requests to a set of pre-defined labels that will help rescuers to understand the situation more clearly.	Machine learning, classification, and text mining- Naïve Bayes protocol, bag-of-words approach.	Flood, fire and personal health
(Wukich, Siciliano, Enia, & Boylan, 2016)	Examines the structure and antecedents of two networks facilitated by the platform Twitter platform operating in the policy domain of emergency management. First includes national-level government agencies responsible for disaster response and recovery operations; the other includes nongovernmental organizations in the form of Red Cross and Red Crescent national societies.	Logistic regression quadratic assignment procedure.	General
(Ancheta, Sy, Maceda, Oco, & Roxas, 2017)	Topic modeling to automate the process of analyzing big data generated from Twitter. There are four themes for topic modeling in the study: communication and information dissemination, humor, resilience and hope, and relief.	Machine Learning for Language Toolkit (MALLET)	Typhoon
(Spielhofer, Markham, Greenlaw, & Hahne, 2017)	Making Twitter data more usable to emergency services and exploring what information can be extracted from the tweets. The work includes an exploratory analysis based on three	Naïve Bayes Classifier	Flood

	areas: noise reduction, location detection, consolidation.		
(Acerbo & Rossi, 2017)	A machine learning approach for creating a classifier able to distinguish between informative and not informative messages, and to understand common patterns inside these two classes, obtained with real data extracted from Twitter during past emergency events.	Leave One Out approach, bagging, boosting, neural networks, SVMs	Fire, floods, earthquake
(Basu, et al., 2017)	Identifying tweets that inform about needs and availabilities of resources, termed as need-tweets and availability-tweets respectively (for post disaster relief operations)	Information Retrieval (IR) techniques, Indri IR system(language modeling), Word2vec (Word embedding based model)	Earthquake
(Li, Wang, Gao, & Shi, 2017)	An automatic analytics for reasoning human emotional responses to extreme events based on event-related textual contents released by social and public media. First, based on the words co-occurrences in event-related reports or tweets or ReliefWeb, a clustering-based analysis is used to remove irrelevant contents from the data stream and identify different event episodes during the period of an entire event.	LIWC (Linguistic Inquiry and Word Count) as semantic analysis tool, word matrix and modularity measurement Q.	Earthquake and Swine influenza
(Wukich, C; Khemka, A, 2017)	Examines how and to what extent national-level Red Cross and Red Crescent organizations (i.e. national societies) use the social media platform Twitter.	Descriptive statistics, logistic regression, and binary correlation analysis	e.g., flooding in Bosnia and Serbia, the Gikomba market bombings in Kenya, and the ongoing humanitarian crisis in and around Syria
(Soni & Pal, 2017)	Creation of gold standard data for automatic retrieval of helpful tweets based on real-time information during disasters and can be utilized to aid rescue, relief and rehabilitation efforts.	Machine learning-SVM (Support vector machine), text categorization.	Earthquake
(Laylavi & Rajabifard, 2017)	Novel method for detecting event-specific and informative tweets that is likely to be beneficial for emergency response. Each sample tweet is given an event-relatedness score which indicates how related a tweet is to the event.	Machine learning-ground truth	Storm
(Zou, Lam, Cai, & Qiang, 2018)	Deriving a set of common indexes from Twitter data so that they can be used for emergency management and resilience analysis; examining whether there are significant geographical and social disparities in disaster-related Twitter use; and testing whether Twitter data can improve post disaster damage estimation.	Sentiment analysis, stepwise linear regression	Hurricane Sandy

(Onorati, Diaz, & Carrion, 2018)	Proposed a visual analytics solution that filters and visualizes relevant information extracted from Twitter. The tool offers multiple visualizations to provide emergency operators with different points of view for exploring the data based on geolocation, topic identification, and topic search. An exploratory study in which 20 practitioners answered questions about the integration of social networks in the emergency management process.	Visual analytic tool(treemap, word cloud, bubble chart, animated map, and infowindow)	Paris attacks

The table is prepared by considering the author, year, study, and content, type of method and disaster type. As seen in the table, machine learning techniques are very important for the disaster management based on social media. Furthermore, earthquakes, hurricanes and floods are the most studied disaster types.

Besides the articles, Open Access Theses and Dissertations (OATD, 2018) database has used for finding open access graduate theses and dissertations published around the world. Search syntax is shown Table 6.

Table 6: Search strings for thesis and dissertations

Database	Search strings
Open Access Theses and Dissertations (OATD) (July 04,2018)	"social media", "disaster"

Search is resulted in 68 matches using English language; 55 of them are related studied field. Approximately 50 studies have been conducted between years of 2014 and 2018. 61 papers can be accessed among 68 papers.

4. CONCLUSION

In this study, a literature review is conducted considering big data and social media in the light of disaster management and specifically disaster relief. So, with this literature review, the importance of big data and social media in disaster management and disaster relief fields can be identified. Such a literature studies allow defining opportunities and future gaps in the related area. There are few studies about disaster management considering big data when literature is analyzed. But recently, social media into big data has been used frequently for disaster management studies. It is determined that most studies are done with twitter analysis. In addition, the response phase of disaster management has been studied more frequently. Additionally, proportion of studies based on social media and humanitarian fields conducted within computer sciences subject area are greater than other areas. It is also seen that machine learning algorithms have an important place in these studies. Furthermore; hurricane, flood and earthquake have been studied more as disaster type. Victims can express their problems through social media and information can be provided about their situation. Which aids victims require are important after disaster and social media can facilitate the crisis situation. How many people needs assistance, how to reach these people, and quantity of aids are the critical questions that can be found the answers through social media. In addition, social media is important in terms of raising awareness of people before disasters. It is also important for psychological support to

victims after the disaster. As it can be seen, the impact of social media on disaster management has not been overlooked, and the literature has begun to devote a large area to this topic.

REFERENCES

Acerbo, F., & Rossi, C. (2017). Filtering Informative Tweets during Emergencies:A Machine Learning Approach. *Proceedings of the First CoNEXT Workshop on ICT Tools for Emergency Networks and DisastEr Relief*, (pp. 1-6). Incheon, Republic of Korea.

Alshareef, H., & Grigoras, D. (2016). Using social media and the mobile cloud to enhance emergency and risk management. . *15 th International Symposium on Parallel and Distributed Computing*, (pp. 92-99). Fuzhou, China.

Ancheta, J., Sy, C., Maceda, L., Oco, N., & Roxas, R. (2017). Computer-assisted thematic analysis of Typhoon. *Proc. of the 2017 IEEE Region 10 Conference (TENCON)*, (pp. 723-726). Malaysia.

Basu, M., Ghosh, K., Das, S., Dey, R., Bandyopadhyay, S., & Ghosh, S. (2017). Identifying Post-Disaster Resource Needs and Availabilities from Microblogs. *IEEE/ACM International Conference on Advances in Social Networks Analysis and Mining*, (pp. 427-430). Sydney, Australia.

Caragea, C., McNeese, N., Jaiswal, A., Tarylor, G., & Mitra, P. (2011). Classifying Text Messages for the Haiti Earthquake. *8th International ISCRAM Conference*, (pp. 1-10). Lisbon,Portugal.

Cherichi, S., & Larodec, R. (2016). Using Big Data Values to Enhance Social Event Detection Pattern. *Computer Systems and Applications (AICCSA), 2016 IEEE/ACS 13th International Conference of.* Morocco.

Galindo, G., & Batta, R. (2013). Review of recent developments in OR/MS research in disaster operations management. *European Journal of Operational Research*, 201-211.

Gupta, S., Altay, N., & Luo, Z. (2017). Big Data in Humanitarian Supply Chain Management: a Review and Further Research Directions. *Ann Oper Res*.

IFRC. (2018, May). Retrieved from International Federation of Red Cross and Red Crescent Societies: http://www.ifrc.org/en/what-we-do/disaster-management/about-disasters/what-is-a-disaster/

Lai, C. (2016). A study of emergent organizing and technological affordances after a natural disaster. *Organizing and technological affordances*, 507-523.

Landwehr, P., Wei, W., Kovalchuck, M., & Carley, K. (2016). Using Tweets to Support Disaster Planning, Warning and Response. *Safety Science*, 33-47.

Laylavi, F., & Rajabifard, A. K. (2017). Event relatedness assessment of Twitter messages for emergency response. *Information Processing and Management*, 266-280.

Li, X., Wang, Z., Gao, C., & Shi, L. (2017). Reasoning human emotional responses from large-scale social and public media. *Applied Mathematics and Computation*, 182-193.

Monaghan, A., & Lycett, M. (2013). Big Data and Humanitarian Supply Networks: Can Big Data ive Voice to the Voiceless?, (pp. 432-437).

Mulder, F., Ferguson, J., Groenewegen, P., Boersma, K., & Wolbers, J. (2016). Questioning Big Data: Crowdsourcing crisis data towards an inclusive humanitarian response. *Big Data & Society*, 1-13.

OATD. (2018). Retrieved from https://oatd.org/

Onorati, T., Diaz, P., & Carrion, B. (2018). From social networks to emergency operation centers: A semantic visualization approach. *Future Generation Computer Systems.*

Papadopulos, T., Gunasekaran, A., Dubey, R., Altay, N., Childe, S. J., & Fosso-Wamba, S. (2016). The Role of Big Data in Explaining Disaster Resillience in Supply Chains for Sustainability. *Journal of Clnear Production*, 1-11.

Soni, R., & Pal, S. (2017). Microblog Retrieval for Disaster Relief: How To Create Ground Truths? *Proceedings of the First International Workshop on Exploitation of Social Media for Emergency Relief and Preparedness* , (pp. 42-51). Aberdeen, UK.

Spielhofer, T., Markham, D., Greenlaw, R., & Hahne, A. (2017). Data mining Twitter during the UK floods: Investigating the potential use of social media in emergency management, 2016 3rd International Conference on. *Information and Communication Technologies for Disaster Management (ICT-DM).* Vienna, Austria: IEEE.

Su, X. (2018, May). *Introduction to Big Data.* Retrieved from https://www.ntnu.no/iie/fag/big/lessons/lesson2.pdf

Wang, Z., & Ye, X. (2018). Social media analytics for natural disaster management. . *International Journal of Geographical Information Science*, 49-72.

Wukich, C., Siciliano, M., Enia, J., & Boylan, B. (2016). The formation of transnational knowledge networks on social media. *International Public Management Journal*, 381–408.

Wukich, C; Khemka, A. (2017). Social media adoption, message content, and reach:an examination of Red Cross and Red Crescent national societies. *Int. J. Emergency Management*, 89-116.

Zou, L., Lam, N., Cai, H., & Qiang, Y. (2018). Mining Twitter Data for Improved Understanding of Disaster Resilience. *Annals of the American Association of Geographers*, 1-20.