

Assessing post-disaster recovery using sentiment analysis: The case of L'Aquila, Italy

Earthquake Spectra

1–28

© The Author(s) 2021



Article reuse guidelines:

sagepub.com/journals-permissions

DOI: 10.1177/87552930211036486

journals.sagepub.com/home/eqs

Diana Contreras, Sean Wilkinson, Nipun Balan, and Philip James

Abstract

Memorial days of disasters represent an opportunity to evaluate the progress of recovery. This article uses sentiment analysis (SA) to assess post-disaster recovery on the 10th anniversary of L'Aquila's earthquake using Twitter data. We have analyzed 4349 tweets from 4 to 10 April 2019 with the hashtag: #L'Aquila that we have obtained from a third-party vendor. The polarity is first defined using a supervised classification based on experts' rules on post-disaster reconstruction and Grammarly tones. Then, this polarity is compared with the outcome of an unsupervised classification based on the pre-trained SA machine learning algorithm developed by *MonkeyLearn*. We have found a significant negative assessment of the post-disaster recovery process in L'Aquila. About 33.1% of the tweets had a negative polarity, followed by 29.3% tweets with a neutral polarity, 28.7% with positive polarity, and 8.9% unrelated to the anniversary. Further analysis of the tweets confirms that after 10 years, the reconstruction is still ongoing and that criticism of the recovery reported in the literature is also found in the tweets. Based on our analysis, the critical day to collect most of the data is the anniversary's exact day. Tweets from citizens and/or news agencies, which are more likely to express the reality experienced, are therefore more useful in understanding recovery than tweets from government officials and/or governmental institutions. From the total 4349 tweets, we can state that 2488 (57%) were correctly classified by the pre-trained SA machine learning algorithm developed by *MonkeyLearn*, while 1861 (43%) were misclassified. It means an overall accuracy (ACC) of 57% and a misclassification rate of 43% by the algorithm. We argue that our results have the potential to serve as a benchmark that can be used to compare other post-disaster recovery processes using the same Twitter-based SA on their anniversaries.

Newcastle University, Newcastle upon Tyne, UK

Corresponding author:

Diana Contreras, School of Engineering, Newcastle University, Research Office 2.24, 2nd Floor Drummond Building, Newcastle upon Tyne NE1 7RU, UK.

Email: diana.contreras-mojica@newcastle.ac.uk

Keywords

Post-disaster recovery, earthquakes, Twitter, sentiment analysis, machine learning, L'Aquila

Date received: 28 September 2020; accepted: 7 July 2021

Introduction

The memorial days of disasters represent a window of opportunity not only to remind us of the human and material losses (Rossetto et al., 2014) but also to evaluate the progress of the post-disaster recovery process. There often seems to be little appetite for governments and/or scientists to monitor recovery in the long term (Shibuya and Tanaka, 2019). Therefore, recovery is poorly understood (Rossetto et al., 2014; Smith and Wenger, 2007; Yan et al., 2020). There is an argument for employing alternative sources of data and methods such as social media (SM) (Wilkinson et al., 2018) and sentiment analysis (SA) based on machine learning to improve our understanding of the recovery process. These methods are useful to either make post-disaster recovery assessments or complement more traditional means taking advantage of SM timeliness (Shibuya and Tanaka, 2019). Social media data are suitable to identify the satisfaction of the inhabitants of the case study area with the recovery process by analyzing posts (Kumar et al., 2019), linked media, user reactions, and the relationship between users (Sánchez-Rada and Iglesias, 2019).

In this research, we use the case of the L'Aquila earthquake of 6 April 2009 to investigate the application of SA (Mäntylä et al., 2018) or opinion mining (OM) (Medhat et al., 2014; Munir, 2019) as a method for the assessment of post-disaster recovery processes using text data obtained from posts on Twitter using the hashtag #L'Aquila that was included in tweets around the 10th anniversary of the earthquake in L'Aquila (specifically from 4 to 10 April 2019). The earthquake happened in L'Aquila on 6 April 2009. The main research question covered in this study is whether SA is a valid method for assessing post-disaster recovery processes. Few studies have explored people's sentiments and perspectives over the post-disaster recovery process (Yan et al., 2020). In previous relevant studies, interviews, focus groups, photovoice, and participatory mapping have been the qualitative approaches mostly adopted to assess post-disaster recovery's non-physical elements (Schumann, 2018; Yan et al., 2020). After 4 and 8 years of the Kobe earthquake, the government of Kobe undertook a comprehensive recovery assessment using the Citizen Happiness Index. This index covered the 16-point Plan of Action and consisted of 45 individual indexes. Twelve workshops were conducted in Kobe to establish (1) what life recovery means for earthquake victims and (2) what are the factors that citizens consider useful to promote recovery. The results identified seven elements in descending order: (1) housing; (2) social ties; (3) community rebuilding; (4) physical and mental health; (5) preparedness; (6) economy, livelihood, and economic and financial situations; and (7) relationship to the government (Honjo, 2011). The problem with this kind of approach is that it heavily relies on the researcher's contacts, with the field limited to a small group of stakeholders (Yan et al., 2020). In contrast, SM is a big data source (Schroeder, 2014), which overcomes the aforementioned limitations, richness in content, volume, velocity, and variety (Gandomi and Haider, 2015; Yan et al., 2020). Social Media delivers sentiments, observations, perspectives, and firsthand information about user experience. Although biased or misleading data could also be a problem, according to Linus's Law, "given enough eyeballs, all bugs are shallow" (Goodchild and Li, 2012; Yan et al., 2027, 2020). However, recent studies still focus on the quantitative assessment of post-disaster

recovery's physical dimension (Yan et al., 2020). The date 6 April 2019 was the commemoration of the 10 years after the earthquake in L'Aquila. In this case, no survey among its inhabitants was planned to assess the recovery process's success. Therefore, we selected this city as a case study area to apply the SA for post-disaster recovery assessment. The SM platform selected for this research is Twitter, given the ease to manually identify text data related to the anniversary based on the hashtags included in the tweets (Cervone et al., 2016; Jamali et al., 2019; Mejri et al., 2017; Mendoza et al., 2019; Neppalli et al., 2017; Wu and Cui, 2018; Yuan and Liu, 2018). Other authors have worked previously with Facebook (Mejri et al., 2017), Instagram, or Flickr (Cervone et al., 2016; Yan et al., 2017). The general objective is to demonstrate that SA is a valid method for assessing post-disaster recovery after earthquakes. Here, we extract features from tweets to assess the post-disaster recovery process's progress following the principle of the citizen as a sensor (Cervone and Hultquist 2018; Laituri and Kodrich, 2008; Shibuya and Tanaka, 2019) and citizen science (Fallou et al., 2020). The other objectives are to (1) ascertain when should data be collected to get the best information, (2) identify which tweets are likely to contain the most useful information, (3) find whether tweets truly reflect the reality, and (4) test the accuracy (ACC) of the pre-trained SA machine learning algorithm for unsupervised classification of text data developed by MonkeyLearn.

This article is divided into eight sections. The introductory section explains this research's novelty, based on the approach taken; elucidates the main objective; and describes the article's structure. The second section corresponds to the literature review. The third section describes the state of the recovery process in the case study area. This section was constructed based on peer-reviewed papers, related articles in newspapers, blogs, and videos published on the 10th anniversary of L'Aquila's earthquake. The fourth section contains a detailed description of the methodology applied. The fifth section presents the SA's application results, and the sixth section discusses the results. The seventh section addresses the conclusions, and the eighth section includes a set of recommendations based on the results and conclusions of applying the selected methodology.

Literature review

SA has been mainly used for customer reviews of products and places (restaurants and hotels); stock markets (Hagenau et al., 2013; Yu et al., 2013); news articles (Xu et al., 2012); products', brands', and services' online feedback (MonkeyLearn, 2020b); or political debates (Maks and Vossen, 2012). In this research, we apply SA for the post-disaster recovery assessment of L'Aquila, classifying the tweets posted by citizens according to their polarity (positive, negative, neutral, or not related) on the occasion of the 10th anniversary of the earthquake. In this research, we conceive post-disaster recovery processes as a service provided by governments in areas affected by earthquakes and their citizens as customers of this service. Therefore, they are the most relevant people to rate their quality.

SA is a natural language processing (NLP) method to analyze automatically (Hausmann et al., 2020) within text data through computational treatment, sentiments, emotions, opinions, attitudes, and subjectivity about a specific topic or toward an entity (Medhat et al., 2014; Zucco et al., 2020). NLP is a branch of artificial intelligence (AI) that enables machines to understand human language (Roldós, 2021). To analyze the text's emotional load, it is essential to understand its meaning (Gurman and Ellenberger, 2015; Ragini et al., 2018). SA identifies the sentiments that are contained in the text and classifies their polarity into positive, negative (Ragini et al., 2018), neutral, or not related to the

specific topic. This classification can be performed at three primary classification levels in SA: document level, sentence level, and aspect level (individual words) (MonkeyLearn, 2020b). There are two techniques for SA: machine learning and lexicon-based. The machine learning approach can be supervised or unsupervised. Supervised learning uses different classifiers such as decision tree, linear, rule-based and probabilistic (Medhat et al., 2014). Each classifier uses different algorithms, the linear classifier uses a support vector machine or neural network, and the probabilistic classifier uses Naïve Bayes, Bayesian network, and maximum entropy. The lexicon-based approach can be dictionary-based or corpus-based, with corpus-based being either statistical or semantic. The feature extraction methods point toward keywords that express an opinion or a sentiment. The keywords extraction methods can be lexicon-based and statistical such as a bag of words (BOW) (Zucco et al., 2020) or word clouds (Roldós, 2020). The word clouds are more frequently applied because they are fully automatic and simplify the feature selection process based on frequency counts (Medhat et al., 2014; Zucco et al., 2020). There are different statistical approaches to produce word clouds: word frequency, word collocations (e.g. bigrams, trigrams, and so forth), and co-occurrences; term frequency-inverse document frequency (TF-IDF; and rapid automatic keyword extraction (RAKE) (MonkeyLearn, 2021). In this research, we used the AI-powered word cloud generator developed by *MonkeyLearn* (Wolff, 2020), which visualizes term frequency and relevance. The relevance score is an algorithm that combines key words frequency with other factors such as how descriptive and how long a word is (Wolff, 2020a).

Recent work on emergencies and disasters includes using Twitter during the emergency phase to share information about needs. These needs are usually water, food, shelter, medical emergency, and electricity (Ragini et al., 2018). Neppalli et al. (2017) identified the divergence of sentiments expressed during Hurricane Sandy and displayed how Twitter users' sentiments change geographically. The authors demonstrated how the user's sentiment changed according to their locations and the disaster's distance. They also found that the polarity of sentiments expressed in the tweet affects the retweetability of the tweet. The extraction of sentiments during a disaster contributes to a vital situational awareness of the disaster zone dynamics. For example, Wu and Cui (2018) used SA to measure each tweet's emotion or mood and classified it as positive, negative, or neutral. Then, they used these data to quantify specific features (Zucco et al., 2020). The original contribution of our research lies in the fact that we use Twitter data for assessing recovery. In contrast, previously, it has been mostly used during the emergency phase for Post-Disaster Needs Assessment (PDNA).

Case study area

On 6 April 2009, an earthquake with a magnitude of 6.3. M_w and a hypocentral depth of 10 km struck the Italian city of L'Aquila, with a population of 72,800 inhabitants at the earthquake. L'Aquila is the capital of the namesake province and the Abruzzo region's administrative capital in central Italy (Contreras et al., 2014, 2018). Its location is shown in Figure 1a to c.

The recovery process has been severely criticized for numerous reasons. For example, immediately after the earthquake, the standard procedure of establishing temporary restricted areas (primarily for safety) was conducted. While these areas were meant to be temporary, some buildings are still cordoned off after 10 years (Imperiale and Vanclay,

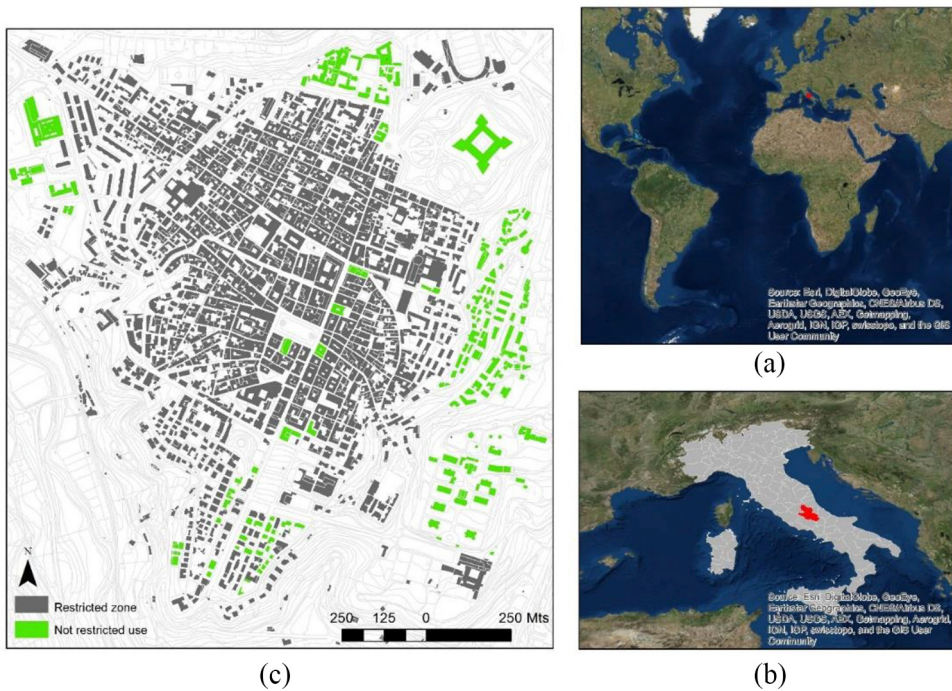


Figure 1. L'Aquila location: (a, b) location of the case study area: L'Aquila, Italy and (c) restricted zone in L'Aquila after the earthquake on 6 April 2009.

Source: Contreras (2016).

2019). Other criticisms relate to excluding the population from recovery decision-making (Fois and Forino, 2014; Özerdem and Rufini, 2013). Another critical aspect is constructing 19 new settlements that are far from the city center (4300 earthquake-proof housing and 1200 temporary housing modules) (Alexander, 2013; Contreras et al., 2013; Guerrieri, 2019). The high cost of the apartments in the new settlements (over €700 million) (Alexander, 2010, 2013; Guerrieri, 2019), their construction quality (three balconies collapsing in 2014) (Fantasia, 2020), and problems with water infiltration (Alexander, 2013; Guerrieri, 2019) were other issues. Additional criticisms relate to the lack of urban facilities around new settlements (Alexander, 2013; Contreras et al., 2017; Rossi, 2019) and the mismanagement of financial resources (Alexander, 2010, 2013; Fois and Forino, 2014; Özerdem and Rufini, 2013). These factors have contributed to the delay in the reconstruction of the historical city center, which 10 years later is still ongoing (Imperiale and Vanclay, 2019; Özerdem and Rufini, 2013). Consequences of this failure to reconstruct L'Aquila is the economic stagnation, the reduction in the sources of employment (Alexander, 2010, 2013), and the resulting dissatisfaction of the inhabitants of L'Aquila with the recovery process, as well as the gradual and very significant depopulation of the city as shown in Figure 2.

For this research, we contemplate four post-disaster phases: emergency, early recovery, recovery, and development (Contreras, 2016). Considering the time since the original earthquake, it is not unreasonable to expect that L'Aquila must have by now achieved the recovery and development phase; however, the city is still considered the largest construction site in Europe, given that significant work still continues 10 years after the

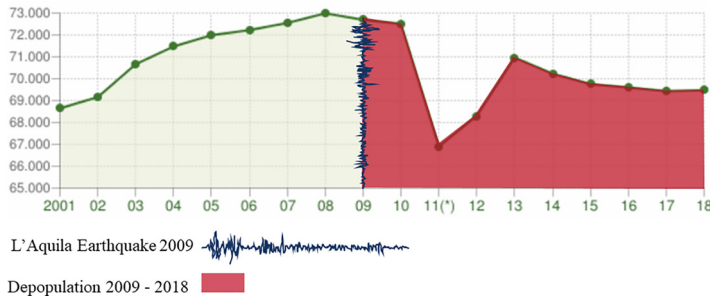


Figure 2. Gradual depopulation of L'Aquila after the earthquake in 2009 until 2018.

Source: Adapted from <https://www.tuttitalia.it/abruzzo/98-l-aquila/statistiche/popolazione-andamento-demografico/>

earthquake, mainly in the historic city center. One specific example is the main street in the city center named Corso Federico II, which was struggling to start again with only a few shopkeepers deciding to reopen in beautiful but empty seventeenth-century buildings (TGCOM24, 2019). Currently, L'Aquila is still a peripheral city with 19 settlements around without urban facilities, like in 2009, but now in some cases with large shopping centers built in record time that can only be reached by car (Contreras et al., 2017; Rossi, 2019). No schools have been rebuilt (Carboni, 2019; Nicola, 2019), and 3600 students and teachers are still conducting lessons in the temporary modules (MUSP by its acronym in Italian) (Fonzi, 2019; Nicola, 2019; Rossi, 2019; TGCOM24, 2019). As in 2009, after the earthquake, the school buildings remain abandoned (i.e. unoccupied and not demolished). Of 59 structures, 29 are unusable. After the earthquake, between €42 (Fonzi, 2019) and €44 million were allocated between 2011 and 2017 by the decrees of the Commissioner of reconstruction and by resolutions of the Inter-ministerial Committee for Economic Programming (CIPE by its acronym in Italian) (Fonzi, 2019) to the reconstruction of schools in the historic center. However, in 2019 only two private schools had been reconstructed in masonry: the Christian Doctrine and Micarelli Sisters (Nicola, 2019; TGCOM24, 2019). For public schools, the reconstruction has also been very long. Of 59 schools, 31 have been damaged. Thirty are to be rebuilt, with others earmarked to be demolished and rebuilt or relocated (Nicola, 2019). One school is still under construction, having started in 2017 (Fonzi, 2019). Guerrieri (2019) argues that parliamentary systems and bureaucracy have slowed down the reconstruction and, therefore, block the city's rebirth. With its 20,000 students, the ancient university was reopened but with the addition of students from the Gran Sasso Science Institute (GSSI). These actions were oriented to convert L'Aquila to be the new 'city of knowledge' (Carboni, 2019).

Based on the water consumption measured by Gran Sasso Acqua (the company in charge of the water management in the capital and other cities in the province), it was estimated that by 2019 the historic city center was inhabited by approximately 4000 people (3015 residents and 894 non-residents). This number of residents is less than half of those that were living there before the earthquake (AbruzzoWeb, 2019). Of 1200 shops opened in the historic center before the earthquake, 10 years later, only between 50 (Fantasia, 2020) and 86 (Carboni, 2019) stores have reopened; therefore, the city is still named "ghost city" (Contreras et al., 2014; INGENIO, 2019). The allocation of this nickname is justified with the presence of rubble from damaged buildings, curtains with dust in the windows and furniture ruined by the sun and rain, and surrounded by bulldozers, trucks, and workers. House prices have gone from €3000/€5000 per m² to a little more than a thousand;

hence, nobody sells properties to avoid monetary losses. The new city meeting point, “the Aquilone shopping center” is out of the historical city center (Fantasia, 2020). While private reconstruction is estimated to finish in 2023, with two-thirds of houses already rebuilt (Carboni, 2019), the reconstruction of public buildings is only 50% complete (TGCOM24, 2019). The Special Reconstruction Office of the Municipality (USRA by its acronym in Italian), created at the end of 2012, received 30,000 requests for financial support for private reconstruction, which is estimated to end in 2023/2024. Almost €18 billion were allocated, with €5.5 billion already spent by 2019 in private construction loans (INGENIO, 2019).

Much of the progress in the reconstruction is the result of work from many associations. These institutions have been motivated by feelings of solidarity with the affected population (Guerrieri, 2019). It was decided to rename the Municipal Council of Fossacesia, a city located in the province of Chieti and the Abruzzo region, to “L’Aquila square,” a large area in the city center on occasion of the 10th anniversary of the earthquake. This decision was made to pay tribute to L’Aquila and its citizens, with whom the city of Fossacesia forged a bond of solidarity, providing them with human, psychological, and logistical support in the post-disaster phase (Editorial, 2019). The master Nicola Piovani composed the “Symphony of the seasons” to tell with music the path followed by the city after the earthquake. The concert was held on the Basilica of Santa Maria di Collemaggio (Fattitaliani.it, 2019), a symbol of the city, already reconstructed and opened to the public in 2017. The church Santa Maria del Suffragio, another city symbol, has also already been restored (Fantasia, 2020). The reaction of a full creative force that has led citizens so self-determined through social reconstruction processes from below is the only positive achievement (Rossi, 2019), besides the prize that awarded the city with the Spirit of Rugby (Editorial, 2019b). The review above shows that there has been a very patchy recovery from the impacts of the earthquake, and there is still a great deal of work to do to return the town to its pre-earthquake status. As there is currently no official review of the efficacy of the earthquake recovery process, it is difficult to formulate a better plan and recovery protocols that could be implemented with better effect in the future. This situation is not unique to the city of L’Aquila. So we now perform an SA using Twitter data on the occasion of the 10th anniversary of the earthquake to provide a benchmark of opinion on the L’Aquila recovery that could be used to compare with other post-disaster recovery processes.

Methodology

The methodology is comprised of seven steps: hashtags identification, hashtags selection, data collection, extraction, processing, classification of polarity, and keywords extraction per polarity. First, we identify the hashtags related to the memorial of the 10th anniversary of the L’Aquila earthquake. Monitoring social media in the days before and after the 10th anniversary, we identified eight hashtags: #L’Aquila, #Laquila10annidopo, #LAquilaGrandiSperanze, #PortamiDoveSeiNata, #ionodomentico, #terremoto del #6aprile, #6aprile2019, and #forzalaquila (Contreras et al., 2020). We manually searched Twitter data using all of them separately.

Second, we select the most frequent hashtag among a set of hashtags related to the anniversary: #L’Aquila. We noticed that #L’Aquila was the most common hashtag in all the tweets related to the 10th anniversary of the earthquake. At the same time, we observed that Twitter users most of the time included more than one of those hashtags or even all of

them in their post, and the hashtag #L'Aquila was in most of them, but not necessarily the others. For this reason, we have assumed that most of the tweets with this specific hashtag posted during the analysis period would relate to the post-disaster recovery.

Third, we purchase tweet data from a third-party vendor named "TweetBinder" (TB) (TweetBinder, 2019) with the selected hashtag. This third-party vendor is a social media monitoring tool. It allows customers to generate reports and track hashtags on Twitter and Instagram. In this research, we only used Twitter data. Fourth, we obtain the data in a report in excel prepared by the third-party vendor. Individuals and organizations (e.g. National Association of Italian Dentists (ANDI), scouts, the A.S. Roma Football club, and so forth) post tweets included in the data set. Fifth, we process the data, which implies eliminating twitter handles, emoticons, and hyperlinks and translating the tweets written in Italian, German, French, Japanese, Spanish, Dutch, Portuguese, Greek, Corsican, Polish Norwegian into English. We eliminate words with already known information such as the phenomenon that caused the disaster, the anniversary day, the time when the earthquake happened, the resulting number of deaths and injured people, and the amount of time passed. This cleaning step is taken to eliminate noise and extract data that provide new information. Sixth, we classified the tweets according to their polarity. To perform this task, we first use a supervised classification and then compared it with an unsupervised. The supervised approach relies on a rule-based system and the tones detected by Grammarly (2009), while the unsupervised on a pre-trained SA machine learning algorithm developed by *MonkeyLearn* (MonkeyLearn, 2020).

Tweets are used because they are microblogs that can be utilized to extract features (keywords) at sentence level (Medhat et al., 2014) to determine their polarity (positive, negative, neutral, or not related). Combining the experience of monitoring post-disaster recovery processes from the first two authors of this article, the specific experience in L'Aquila's case of the first author, and the fourth author's expertise in urban topics, we set the rules to classify the polarity of the tweets. We considered the tweets that relate to life returning to normality as having a positive polarity. Tweets that do not provide information about the recovery are considered neutral. We classified tweets that highlight failures in the recovery process with a negative polarity. In developing our polarity rule set, we acknowledge that not everyone will agree with our classification, nor will all tweets associated with a specific rule fall into a unique category (i.e. always be positive). However, we argue that the vast majority of our rules would not be controversial to most people, and the vast majority of tweets associated with each rule will fall into the category we have chosen for it. The complete and specific rule set to define the polarity of tweets related to the 10th anniversary of the earthquake in L'Aquila can be read in Table 1.

Grammarly is an online writing assistant with a tone detector feature (Lardinois, 2019). This feature can detect different tones covering a broad range of emotions from more than 120 characters. Although the tone detector from *Grammarly* was developed to inform users of the tone of their emails and documents (Lardinois, 2019), not for SA, we decided to use this feature to see if it was appropriate in this application. We considered that tones detected by *Grammarly* could be related to a specific polarity and therefore support the expert judgment. The tones detected by Grammarly and classified in polarity by authors are listed in Table 2.

We use the no-code AI platform: *MonkeyLearn* to understand the human language contained in the tweets. *MonkeyLearn* is a user-friendly AI platform that supports users in getting started with NLP using pre-trained models or building customized ones that fit their

Table 1. Polarity classification rule set

Polarity	Rules
Positive	<p>The hope of the final reconstruction of the city. Callings for reconstructing for the future. Learning the lessons from the event. Calls to not forget what happened. Stories of survivors and rescue teams. Honors to victims. Solidarity messages. Acknowledgments to governing authorities. The role and the importance of the University of L'Aquila in the existence of the city with its students and projects. Commitment of governmental institutions such as the National Association of Public Assistance (ANPAS by its acronym in Italian) with the city announcements of the construction of infrastructure projects. Contributions of the private sector to the reconstruction, such as Sanofi Italy to the reconstruction of Scoppito. Comments of buildings already reconstructed. Announcements of opening of new business. Job announcements. School registrations. Comments of people moving to L'Aquila after the earthquake (Graziani, 2019). Anniversaries of institutions located in L'Aquila. Places to visit in L'Aquila and the surroundings, such as the Gran Sasso and Monti della Laga National Park. Science developments and cultural activities taking place in L'Aquila. Tributes to the city on radio. Social and artistic initiatives. Promotion of products of the region Promotion of sports events related to rugby (L'Aquila had one of the most famous and successful clubs in Italy) and sport victories. Praises to the value of the renaissance architecture and monuments in the city. Memories about how beautiful L'Aquila used to be. Good memories about foreign communities there, considered as an expression of topophilia (Pacione, 2009). Visit of prominent personalities. Charity works. Welcoming of immigrants. Newborn babies posts.</p>
Neutral	<p>The seismic information of the event such as magnitude, aftershocks, and geological changes caused by the earthquake. Invitation to people to attend the torchlight procession and the commemoration ceremony at the main square. Calls to watch TV and radio programs about the anniversary. Calls to remember the students who died due to the collapse of the student hostel in the city. Remembrances of the hailstorm in the night after the earthquake (Castellani, 2019). Memories of Tommy, the dog that rescued three people after the earthquake.</p>
Negative	<p>Expressions of inability to forget the impact of the earthquake. Remembrances of the red zone.</p>

(continued)

Table 1. Continued

Polarity	Rules
	<p>Mentions of the existence of cordoned houses, rubble, and barriers still 10 years after the earthquake.</p> <p>Comments about depopulation in the city center despite the reconstruction efforts.</p> <p>Complains about the delay in the reopening of schools.</p> <p>Criticism toward the amount of money designated for the reconstruction of churches.</p> <p>Comparison of the time required to finish the reconstruction of private and public buildings.</p> <p>Comparisons of the reconstruction of L'Aquila with the reconstruction of other Italian cities, also affected by earthquakes such Amatrice or other countries also affected by earthquakes such as Japan.</p> <p>Complains about the mismanagement of the financial resources and the emergency.</p> <p>Claims for an effective prevention and DRR policy.</p> <p>Complains related to the lack of information and lack of leadership by the government.</p> <p>Mourning the promising players from the L'Aquila rugby team who died due to the earthquake.</p> <p>Reports about the presence of asbestos among the debris.</p> <p>Complains about the slowness to remove debris.</p> <p>Complains of new shorings.</p> <p>Expressions of desire to leave the city.</p> <p>Incidents taking place during the commemoration ceremonies</p>

ANPAS: National Association of Public Assistance; DRR: disaster risk reduction.

Table 2. Polarity classification of tones detected by Grammarly

Positive	Neutral	Negative
Admiring	Confident	Angry
Appreciative	Curious	Anxious
Friendly	Frank	Cautionary
Joyful	Formal	Confused
Optimistic	Informative	Disapproving
	Informal	Dissatisfied
	Neutral	Sad
	Objective	Skeptical
	Compliant	Worried
		Accusatory
		Disheartening

needs (Roldós, 2021). *MonkeyLearn* is a software as a service (SaaS), known as subscribe-ware or rentware software with an online sentiment analyzer built using machine learning (Stecanella, 2020). In this research, we used the *MonkeyLearn* pre-trained SA model. We undertook the *unsupervised* classification using the pre-trained SA machine learning algorithm developed by *MonkeyLearn*. SA is modeled as a classification problem, whereby a classifier is fed a text and indicates a polarity category such as positive, negative, or neutral. *MonkeyLearn* has developed a model pre-trained to associate a particular input (i.e. text data) to the corresponding output (tag). The feature extractor transfers the text input into

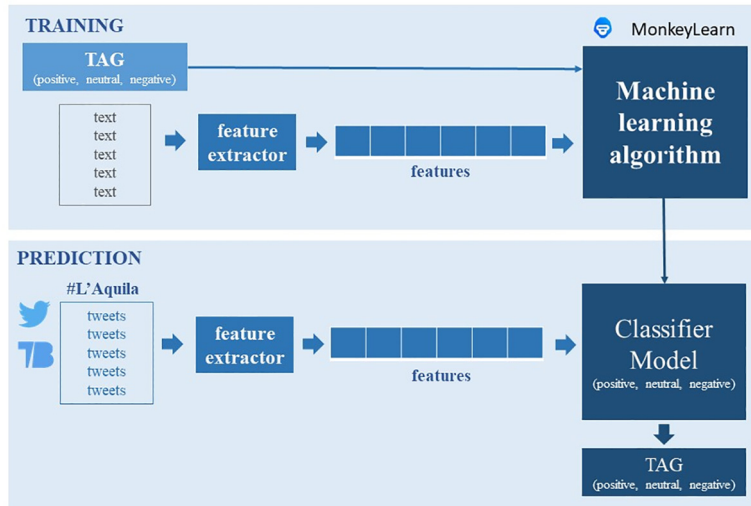


Figure 3. Machine learning classifier developed by MonkeyLearn adapted to the case study. Source: Adapted from MonkeyLearn (2020).

a feature vector. Pairs of feature vectors and tags (e.g. positive, negative, or neutral) are fed into the machine learning algorithm to generate a model. In the prediction process, the feature extractor is used to transform unseen text inputs into feature vectors. These feature vectors are fed into the model, which generates predicted tags such as the polarity categories (i.e. positive, negative, or neutral) (MonkeyLearn, 2020). The schematic representation of using the machine learning classifier developed by *MonkeyLearn* adapted to the case study is depicted in Figure 3. The tag's allocation includes a percentage of confidence based on the polarity allocated to each tweet based on the trained model. Examples of classification for each polarity are depicted in Figure 4a to c.

The first author processed the text data contained in the tweets and made a preliminary classification of them considering the rule set defined by the authors, the tones allocated by *Grammarly*, and the polarity allocated by the pre-trained SA model developed by *MonkeyLearn*. Internal discussion between the first two authors solved disagreements for the rule definition and the final classification of tweets.

The seventh and last step is the keywords extraction, for which we use the AI-powered free word cloud generator also developed by *MonkeyLearn* (Wolff, 2020). We selected this tool because it visualizes not only term frequency but also relevance. This word cloud generator integrates N-gram statistics, therefore recognizing word collocations that other word cloud tools focused solely on frequency do not identify (Roldós, 2020). The graphic description of the methodology applied is depicted in Figure 5.

Results

We purchased a report in .xls format of 4349 tweets from a third-party vendor's data, including retweets with the hashtag #L'Aquila from 4 to 10 April 2019. Looking at the numbers of tweets made each day (Figure 6), we can see that the number of tweets grows steadily from 4 April, with 242 tweets, to 1289 tweets on 5 April and peaked at 1794 tweets

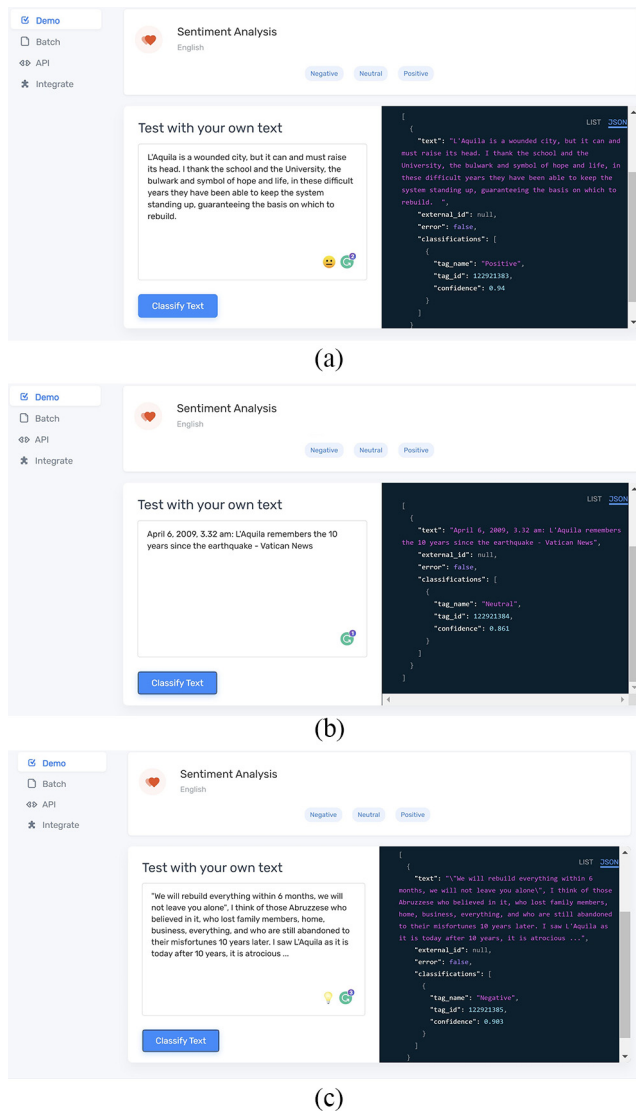


Figure 4. *MonkeyLearn* classifier model interface: (a) tag: Positive, (b) tag: Neutral, and (c) tag: Negative. Source: Adapted from *MonkeyLearn*.

on 6 April (the anniversary day). There is a dramatic downturn with 448 tweets the next day and only 265 tweets on 8 April, 215 on 9 April, and finally 96 on 10 April (the last day of the observation period).

Analyzing the tweets, we found that 33.1% (1440) of the tweets have a negative polarity, followed by 29.3% (1275) tweets with a neutral polarity, then 28.7% (1246) with positive polarity, and 8.9% (388) that were not related to the anniversary. Each segment of the pie in Figure 7 represents the proportion of the polarity described before. On 4 April, the predominant polarity was neutral (see Figure 8a), followed by positive, negative, and not related. On 5 April, the prevalent polarity was negative (see Figure 8b), followed by neutral, positive, and not related. On 6 April, the negative polarity increased (see Figure 8c)

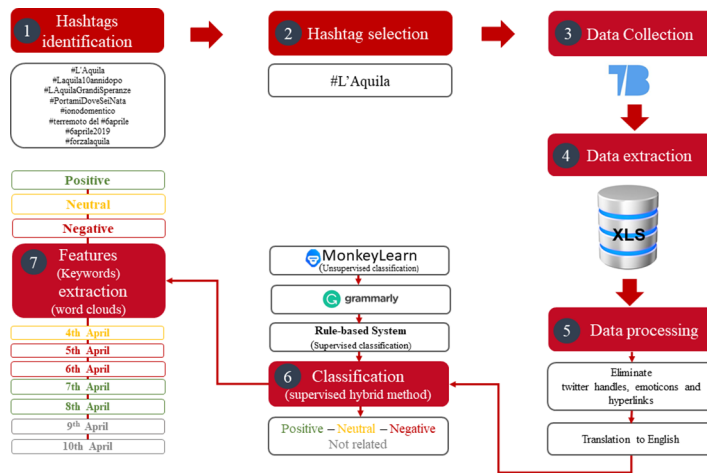


Figure 5. Methodology for the post-disaster recovery assessment using sentiment analysis.

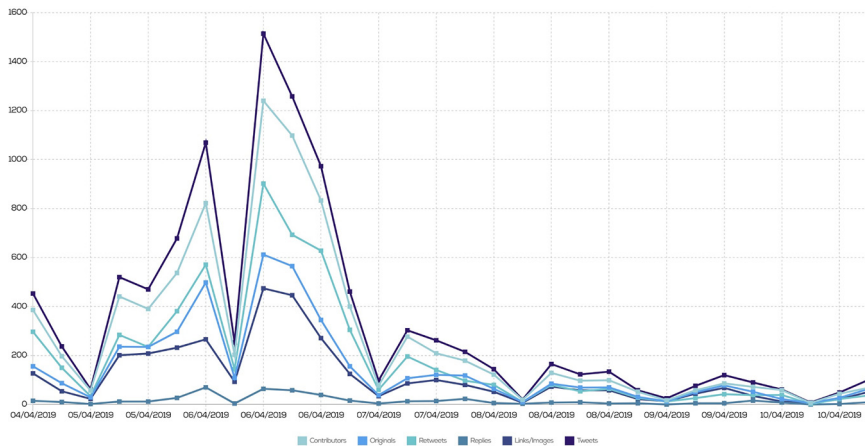


Figure 6. Twitter activity from 4 to 10 April 2019 with the hashtag: #L'Aquila. Source: TweetBinder.

Table 3. The trend distribution of the tweets' polarity with the hashtag: #L'Aquila during the observation period

Date	Polarity				
	Positive	Neutral	Negative	Not related	Total
4 April 2019	64	98	60	20	242
5 April 2019	360	420	452	57	1289
6 April 2019	498	547	692	57	1794
7 April 2019	134	112	126	76	448
8 April 2019	93	54	65	53	265
9 April 2019	69	28	32	86	215
10 April 2019	28	16	13	39	96
Total	1246	1275	1440	388	4349

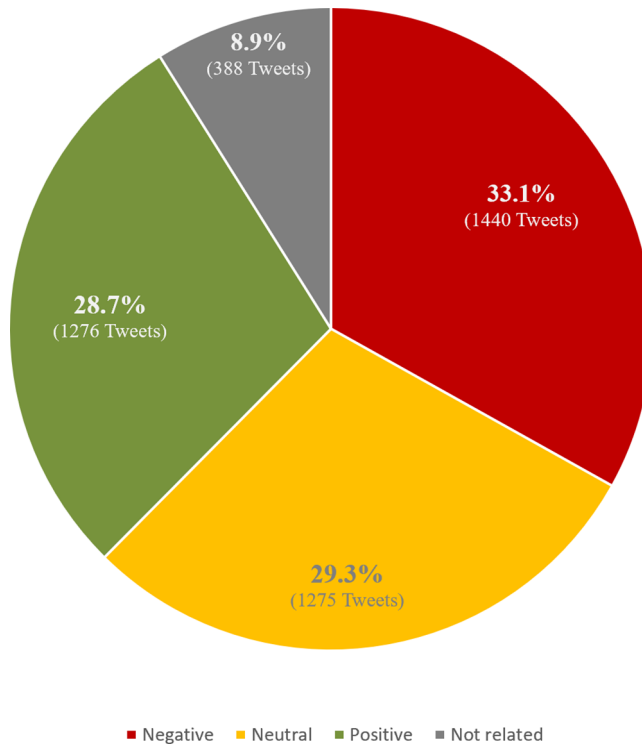
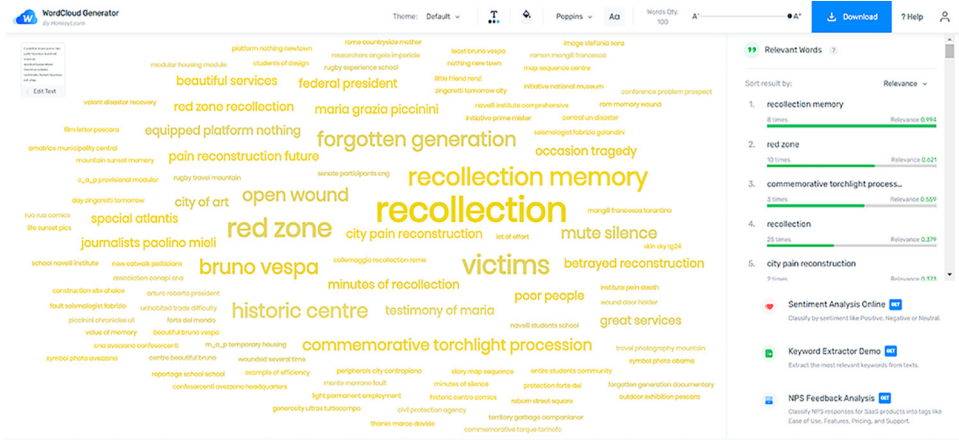


Figure 7. Polarity of the tweets with the hashtag: #L'Aquila.

compared with the previous day, followed by the neutral polarity, positive, and not related. On 7 April, the positive polarity was, for the first time, the primary polarity (see Figure 8d), followed by the negative polarity, then neutral and not related. On 8 April, the positive polarity was still the most prevalent (see Figure 8e), followed by negative, neutral, and not related. Still, this time the difference in the number of tweets with neutral polarity (54) and not related tweets (53) is minimal. On 9 April, the number of not related tweets exceeds, for the first time, the tweets with other polarities (see Figure 8f), with the second representative polarity being positive, followed by negative and then neutral. On 10 April, the last day of the observation period, the polarities were the same as the previous day (see Figure 8g). The varying polarity of the tweets during the observation period and their trends are detailed in Table 3 and plotted in Figure 9.

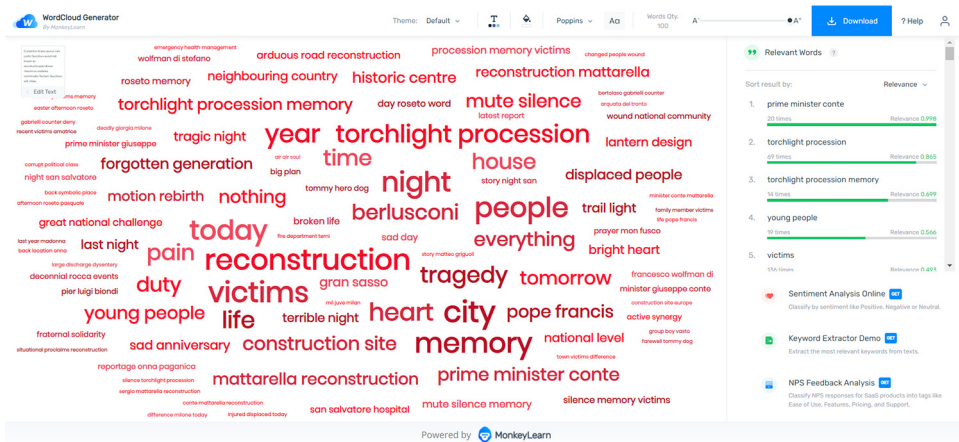
After the polarity analysis, we then extracted keywords using word clouds or tag words to identify the positive, neutral, and negative aspects of the post-disaster recovery. In the current research, we used the AI-powered word cloud generator developed by *MonkeyLearn* (Wolff, 2020), which visualizes term frequency and relevance. The relevance score is an algorithm that combines keywords frequency with other factors such as how descriptive and how long a word is (Wolff, 2020). The keywords extracted from tweets with negative polarity are depicted in Figure 10a. Considering those with the higher frequency and relevance, we can interpret that people in L'Aquila feels still pain as displaced victims and a forgotten generation. They still have a wounded historic city center initially named "red zone" and later "ghost town," "second Pompei" and "construction site" given the still presence of rubble on it after 10 years of the earthquake. According to Twitter



(a)

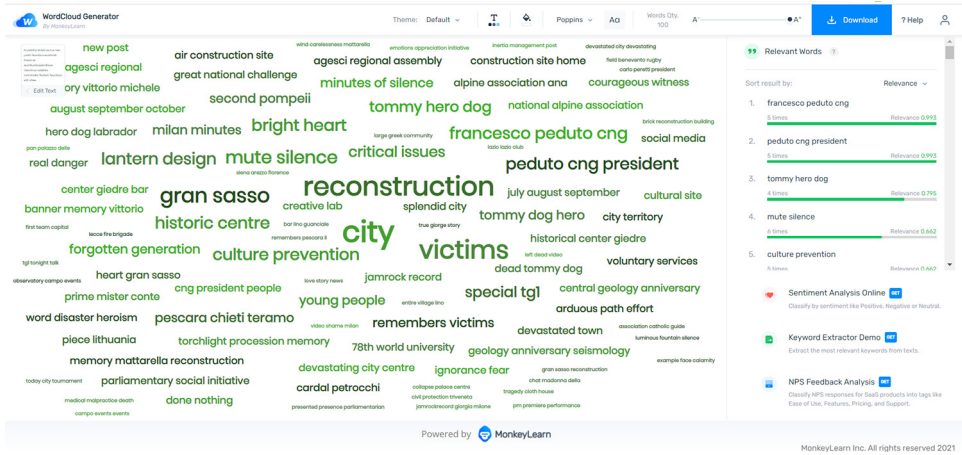


(b)

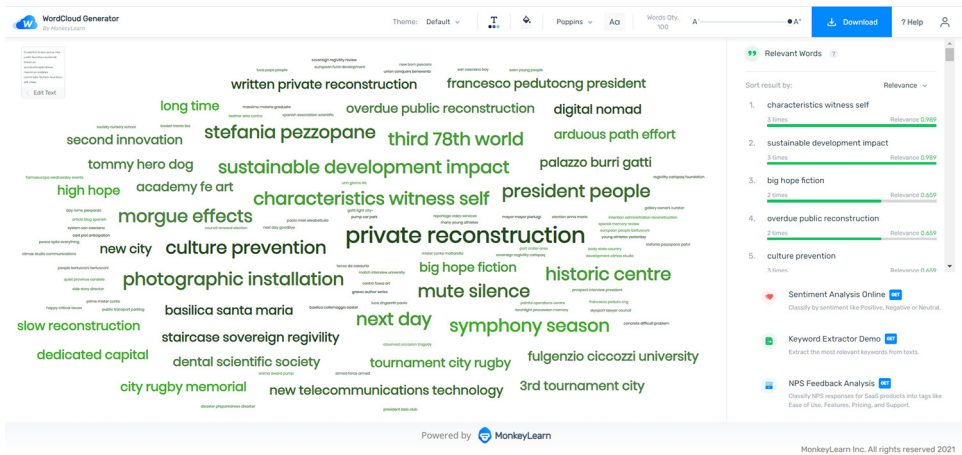


(c)

Figure 8. Continued



(d)



(e)



(f)

Figure 8. Continued



Figure 10. Keywords extraction from tweets per polarity using word clouds: (a) negative, (b) neutral, and (c) positive.

Tommy, the hero dog. Tommy was the first search and rescue (SAR) dog of the Provincial fire Brigade Command of Lecce, who rescued three people from the rubble the day of the earthquake in L'Aquila and died 2 days before the anniversary due to its age

(13 years old). This rescue dog also participated in the SAR activities in Amatrice and received the national award “Dogs with stars” for the lives saved from the earthquakes’ rubble that hit Central Italy, awarded by the then President of the Republic Giorgio Napolitano. The keywords extracted from tweets with positive polarity are presented in Figure 10c. These frequent and relevant keywords reflect the statements to the press from Alessandra Priante. She considered that L’Aquila’s touristic potential had not yet been tapped and planned to lay the city’s tourism industry’s foundation. Ms. Priante was the Head of International Relations and Protocol–Tourism Policies from the Ministry of Agricultural, Food, and Forestry (MIPAAF) at the time of the 10th anniversary and currently is the Director of the Regional Department for Europe at World Tourism Organization (UNWTO). This fact may explain why the historic city center, the Basilica Santa Maria de Collemagio, the Church of Saint Mary of Suffrage, and the Gran Sasso and Monti della Laga National Park also appear among the keywords extracted.

Discussion

Monitoring SM and then searching manually for tweets with the hashtags related to the event before 4 April, we found some tweets about the 10th anniversary of the earthquake in L’Aquila starting on 2 April. After some literature review, we realized that third-party vendors (Ragini et al., 2018) could collect Twitter data more efficiently (more data in less time). After checking several vendors, the authors opted for TweetBinder, given their fair conditions to sell the data. We only processed the data collected by the third-party vendor from 4 to 10 April.

The third-party vendor sells data collected in windows of 7 days, 1 month, and historical data. As we purchased the data on 10 April, we considered we had enough data from a period of 7 days after this date. Tweets were posted in 13 languages listed in the “Methodology” section. About 86.7% of tweets were written in Italian. We cannot assume that the tweets written in Italian are posted only by the population living or born in L’Aquila. To confirm this fact, we would need to check each Twitter user’s location that included the hashtag #L’Aquila in the selected data collection period. This information is not always available or does not necessarily indicate the current location of the user. We could have selected as a sample only the tweets in Italian, but then we would have missed the external view of the progress of the recovery process in L’Aquila, which, according to Shibuya and Tanaka (2019), has a different public sentiment to the local population during the recovery phase. The global population concentrates on images of people suffering, while the local population focuses on business details of the damage and response. The results of these authors demonstrated that when the local population in the area affected by the Great East Japan Earthquake and tsunami in 2011 expressed more complex emotions, which means a combination of positive and negative words, there were more socio-economic recovery activities ongoing. In contrast, when non-local people expressed fewer emotions, there were more socio-economic recovery activities. It seems that while local people do their best to be strong and return to their daily routines, people outside the disaster area continue to worry as part of their collective gaze on the suffering of others (Bica et al. 2017). Tweets from non-local people reflect their sentiments and perspectives rather than the real situation in the affected area (Yan et al., 2020). The recovery effort is criticized in a detailed manner in the case study area section. This section was written based on press articles included in links of tweets on which SA was applied. According to the method applied, most of these tweets had negative polarity (33.1%). This result explains the high level of criticism of this section.

Figure 6 shows that the activity on Twitter relating to the 10th anniversary of the earthquake in L'Aquila rose steadily before 6 April 2019 (which is the day of the anniversary) and peaked on the anniversary day, falling off dramatically after the event. Although we cannot prove by analyzing only one event, we argue that this Twitter activity, in terms of the total number of tweets and duration of increased Twitter activity relating to the anniversary of an earthquake, is likely to be correlated with the magnitude of the impact (expressed in among other things: casualties, injured people, buildings damaged, and economic losses) as well as the progress and the effectiveness of the recovery. This statement is in agreement with Yuan and Liu (2018). These authors found a strong positive correlation between the number of damage-related tweets and the degree of damage for Hurricane Matthew's case. For the same event, Kryvasheyev et al. (2015) demonstrated that the per-capita Twitter activity strongly correlates with the hurricane's per-capita economic damage. Neppalli et al.'s (2017) research shows how sentiments change according to their location and distance from the disaster and how the divergence of sentiments affects the tweet retweet ability. Mendoza et al. (2019) considered SM as a valuable source of spatial information for quickly estimating earthquake damages. Simultaneously, Eyre et al. (2020) developed a methodology to estimate the post-disaster recovery of small business areas similar to L'Aquila's historical city center, indirectly counting the online posting on SM.

We propose to apply our method to other similar earthquakes and earthquake anniversaries to determine whether this is indeed the case. Regarding polarity, we observe a neutral polarity at the beginning of the observation period, which turns negative as the day of the 10th anniversary is approaching to become positive for the two following days and not related for the last 2 days of the observation period. Altogether, the majority of the tweets have a negative polarity. This polarity is followed relatively closely by tweets with neutral polarity and then tweets with positive polarity. In the last place are tweets that are not related to the earthquake's anniversary and therefore are not useful for the post-disaster recovery assessment. Yan et al. (2020) also found in their research studies variations of people's sentiment and perspective regarding the post-disaster recovery process of Lombok and Bali (Indonesia) through time.

For L'Aquila's case, positive tweets did not relate to concrete achievements concerning the city's recovery but instead mainly referred to the hope for completing the reconstruction process. Only those tweets about the emotion expressed by the President of the consortium when returning the house keys to inhabitants of the city center (RepubblicaTv, 2019) and the increase in the enrolments at the engineering faculties in the University of L'Aquila (which is consistent with L'Aquila being a university town) could be considered tangible facts related to post-disaster recovery. The tweets with negative polarity correlate well with sentiments expressed in the literature reviewed at the start of this article reporting buildings still cordoned off, exclusion of population for decision-making, "temporary housing" far from the city center, the high cost of those housing units considering their construction quality, and the lack of urban facilities around. Other aspects reported on tweets were the delay in reconstructing, having the slow repopulation of this zone as a consequence, and the city's depopulation giving the economic stagnation and reducing employment sources. They relate to delays in the reconstruction of the historical city center, politics, bureaucracy, corruption, lack of information, presence of toxic elements among the rubble and the slowness to remove them, new shoring, and the desire to leave the city, and a few include xenophobic messages. Comparisons of the recovery process in L'Aquila with the recovery process of Portofino (2019) after the coastal storm, Norcia and Amatrice (2016), Molise and Apulia (2002), Irpina (1980), and Friuli (1976) after the

earthquakes, and even with the destruction of Pompeii (AD 79) are quite frequent on the tweets. One gap that the population identifies during the recovery/redevelopment phase in L'Aquila is that private buildings' reconstruction is more advanced than public ones. There is a delay in the reopening of schools and other urban facilities in the historic city center. In contrast, the reconstruction of churches is advanced. Some tweets compare the different funding sources for reconstructing specific buildings such as churches and their reconstruction speed with public buildings such as schools. Another aspect is that citizens identify the continued existence of rubble and barriers in the street as an open wound due to the earthquake. The topic of improvement in construction practices is not considered in any of the tweets collected, but there are numerous calls to learn the earthquake lessons. Several tweets about the reconstruction cost had different amounts and this could be interpreted as a lack of accurate official information, which is essential for assessing a post-disaster recovery process. One of the difficulties with the method is, for example, that positive information can be detected as neutral because the algorithm does not know the context (MonkeyLearn, 2020) of the post-disaster recovery process. To do this automatically would require careful training of the algorithm. After a careful manual revision of the links that they contained, some tweets that were initially classified as not related to the event were classified as positives. They announced an artistic performance that commemorates the earthquake's 10th anniversary. The machine learning algorithm from *MonkeyLearn* identifies tweets referring to the TV show to commemorate the 10th anniversary as positive, but in the rule set defined by the authors, a neutral polarity was allocated to these tweets as they do not tell us anything about the success of the recovery. Most of the tweets not related to the earthquake's anniversary, but still using the #L'Aquila hashtag, were weather forecasts. In some cases, it is necessary to check the links in the tweets to establish their polarity. Some of the links that we classified as "not related" were referring to other issues such as football games or racist comments, and political issues but still included related hashtags to the anniversary, such as #forzalaquila.

We detected a problem in the machine learning algorithm developed by *MonkeyLearn* to identify expressions that comprise more than three words such as "Basilica Santa Maria di Collemaggio," which the algorithm identified; however, it broke up into three different expressions and allocated different relevance: "maria di collemaggio," "santa maria di," and "basilica santa maria." It seems that the algorithm does not recognize more than 3-grams. The machine learning algorithms from the *Grammarly* tone detector accurately recognize messages with a neutral tone. Nevertheless, it still needs training and supervision. The algorithm does not recognize sarcasm in the tweets, for example, "Then they must hire them indefinitely. Evidently, Landini thinks that in L'Aquila the reconstruction will never end. excellent perspective." In this case, the messages that the algorithm recognizes with a friendly, joyful, and forceful tone have an opposite message. Stern expressions in local languages such as *sticazzi dei vivi* in Italian are not translated by tools such as *Google Translate*. Then, for this case, a native speaker is needed because those words contain a strong negative polarity, essential for the post-disaster recovery assessment using SA. Some messages have negative polarity, but as they do not use negative words, the algorithm recognizes them as neutral, and sometimes this also happens with positive tweets; this is another reason to use supervised classification. In addition, we realized that the tones detected by *Grammarly* are not exclusive to one polarity, and even more, those tones classified as neutral such as "confident" can be detected as well on tweets with positive or negative polarity. *Grammarly* added new tones such as "forceful" and "direct" during this research, which was not considered initially, and it was detected in tweets with positive and negative polarity. Another aspect to consider is that between one and three tones can be detected in tweets, which we did not expect at the beginning of this research.

Another aspect to consider for further research is to replace the emoticons by their meanings instead of eliminating them and considering them as tex-data for the classification.

One of the limitations of this research is that we are capturing the sentiments, emotions, opinions, and attitudes only from Twitter users. Still, we consider that given the widespread use of this SM platform, which is also connected to others such as Instagram, we collect a representative amount of data to apply SA for L'Aquila's post-disaster recovery assessment.

Conclusions

In this research, we have used SM to obtain a sense of L'Aquila's recovery process by interrogating Twitter data with the hashtag #L'Aquila around the time of the 10th anniversary of the initial earthquake. The other option is to select a more general hashtag that contains the words earthquake and/or anniversary to collect data. For objective 1, we have found that to mine data about the anniversary of an earthquake, it is advisable to start looking for tweets that contain hashtags with the name of the affected area and the time passed after the disaster at least 2 days before the anniversary and 3 days after. This action increases the probability that the data will be related to the event of interest. However, manual analysis did show that some tweets relating to the 10th anniversary of the earthquake in L'Aquila occurred as early as 2 April. We also found that there are more related tweets in the days before the anniversary of the earthquake than after, with the critical day to collect most of the data being on the day of the anniversary. We still need to validate this finding for other events by monitoring the Twitter activity related to anniversaries of other earthquakes (preferably also after 10 years) to compare and elucidate the factors that determine the amount of activity expressed in a number of tweets related.

For objective 2, in terms of usefulness, tweets from the citizen and news agencies are more suitable to assess post-disaster recovery since they express the reality they experience and the observed gaps (i.e. they are analogous to customers in more traditional SA). At the same time, official sources are more likely to focus on positive achievements, for example, the Mayor of L'Aquila highlights that it is the province with the highest number of graduate students in Italy and that they have returned the life to the suburbs. Some individuals were highly critical of the recovery. In the observation period, one citizen tweeted, "A decade of corruption, political instrumentalization and money burned in propaganda. Eighteen thousand million and six prime ministers later, L'Aquila still does not recover from the earthquake that left 309 dead."

We believe using tools such as *Grammarly* and *MonkeyLearn* is suitable for detecting polarity in tweets and *MonkeyLearn* to visualize keywords in the word clouds; however, some improvements (mentioned in the "Discussion" section) must be implemented. Word clouds confirm our method's feasibility to identify the positive or negative assessment reasons by revealing achievement and gaps in the recovery process. We argue that the words extracted according to frequency can be used to construct a lexicon (Ragini et al., 2018) about post-disaster recovery to be applied to any case study. The expressions extracted are more case-specific. Tweets regarding the memorial of disasters are not necessarily trending topics like they are for the actual disaster (i.e. #Laquilaearthquake is likely to be trending while #Laquila10annidopo is not). Some people use a hashtag to get more comments and retweets on their posts on Twitter, without the content being necessarily related to the hashtag. Some tweets contain only links; in this case, it is worthwhile to check their content and establish their polarity so that they can be used in the SA. It is also necessary to identify

acronyms in the tweets that can be incorrectly translated into English. Supervised classification allows to detect ironic tweets and harsh colloquial expressions. Finally, it is necessary to compare the content in the tweets with the evidence. For example, those tweets with a negative polarity that state that “nothing has been done in 10 years” do not recognize some progress in the recovery process of L’Aquila, such as that reported by Fois and Forino (2014) and Contreras et al. (2014) and Contreras et al. (2018).

For objective 3, related to the reflection of reality from tweets, we found a lot more negative assessments of L’Aquila’s post-disaster recovery process in the tweets, which agrees with most of the literature: for example, relating to, after 10 years the reconstruction is still ongoing, with even the presence of the rubble in the street; cordoned-off areas have delayed the reactivation of the city center, which includes the lack of schools (Nicola, 2019; TGC0M24, 2019); and priority for reconstruction allocated to churches. It is understandable the annoyance of the population with the differences between the reconstruction of churches and schools. Churches, independent of their religious connotation and symbolism in the city, are urban facilities used only by catholic worshippers, which are not essential for the city’s development and repopulation, while schools are. In addition, the speed in the reconstruction is confirmed by our SM analysis (i.e. speed of reconstruction is related to the different sources of funding, with the reconstruction of buildings financed by private sources and donations being more efficient than those which are reconstructed using public resources). Negative tweets criticize the commemoration ceremonies and the mismanagement of the financial resources for the reconstruction. Neutral tweets invite people to remember the victims and their families, acknowledge the work done by the SAR teams, attend the commemoration ceremonies, watch the TV programs about the anniversary, or provide information about the recovery process. Tweets with positive polarity call for hope and, not forget, solidarity, reconstruction, encouragement, memories, acknowledgment of contributions, job and service offers, social initiatives, sense of belonging, survivors’ stories entrepreneurship, but highlight very few achievements in the recovery process. Negative tweets outweigh positive tweets, and the negative tweets are often quite specific in the failings of the recovery, while the positive tweets nearly always referred to the general sentiments and hardly ever to concrete achievements. Whether this means that L’Aquila has had a relatively poor recovery or is due to bias in the motivations for people to report these events on SM is an open question, but monitoring other 10th anniversaries of disasters may shed light on this. On average, researchers agree that SA models need to have at least 50% ACC to be considered effective, while an ACC around 65% is considered good (Maksimava, 2020). For objective 4, from the total 4349 tweets, we can state that 2488 (57%) were correctly classified, while 1861 (43%) were misclassified. It means an overall ACC of 57% and a misclassification rate of 43% by the algorithm, which we argue is acceptable.

The feature extraction allows us to identify the most influential people in the recovery in the case study area during the decennial. Several Italian politicians are mentioned in the tweets, such as Silvio Berlusconi, Matteo Renzi, Guido Bertolaso, Sergio Mattarella, Massimo Cialente Roberto Salvini, among others. Silvio Berlusconi appears mentioned in the tweets with all the polarities. Bertolaso and Cialente are included in tweets with positive polarity, but curiously they and the other politicians are contained in the tweets with negative polarity. Renowned figures express their support to the recovery process, such as Pope Francis, actors Lino Guanciale, Luca Zingaretti and Dario Acocella. The Italian singers sang together a song to commemorate the anniversary. The activism of the prominent

people mentioned before gives visibility to the recovery process and attracts investment and tourism to boost the recovery. The Italian radio and television (RAI) and the TV program, “Propaganda Alive” are highly mentioned in the tweets given the documentary and the special show prepared to commemorate the anniversary. It is also noticeable the mention on the tweets of the presence of the Italian trade unions with their General Secretaries in the commemoration of the earthquake, that is, Maurizio Landini, Secretary-General of Italian General Confederation of Labour (CGIL), Anna Maria Furlan Secretary-General of the Italian Confederation of Workers’ Trade Unions (CISL), and Carmelo Barbagallo from the Italian Labour Union (UIL).

The result of both the supervised (rule-based labeling) and unsupervised classification (*MonkeyLearn*) shows mainly a negative polarity and we, therefore, infer a negative assessment of the post-disaster recovery process. This finding is also documented by the sources linked to the tweets analyzed in this article and the observations of the first author, who monitored the recovery process of L’Aquila through fieldwork between 2010 and 2016 (Contreras et al., 2014, 2016, 2018). Therefore, we consider that we managed to demonstrate through the case of L’Aquila that SA is a valid method for assessing post-disaster recovery after earthquakes. However, we agree that we need to improve the method and test it in other cases. This exercise has allowed us to demonstrate that SA is a feasible tool to evaluate the post-disaster recovery process by identifying achievements and gaps detected by citizens as sensors (Cervone and Hultquist 2018; Laituri and Kodrich, 2008; Shibuya and Tanaka, 2019) and citizen science (Fallou et al., 2020). Automatic methods of interrogating the data are largely successful, although there are some areas where these methods struggle. According to the tweets, we argue that much of L’Aquila is still in the recovery phase without reaching a developmental one. The fact that most tweets were written in Italian indicates that they are more likely to express the feelings of inhabitants of the city or at least people connected to the case, therefore potentially giving more confidence to the information provided in the tweets. We can conclude that anniversaries reawaken the memory in citizens and motivate them to comment on their own experiences and compare the success or failure of other post-disaster recovery processes near the affected area.

Recommendations

We recommend starting the search for data related to these events at least 4 days before the anniversary. According to L’Aquila’s case, we suggest searching for related tweets only 2 days after the anniversary. After this day, the number of not related tweets, even containing the selected hashtag, starts overcoming the number of related tweets.

While we have extracted a great deal of information about this particular event, we are not yet in a position to be able to rate how this recovery compares with other events based on our data. Future work will be to apply SA using the same hybrid approach (supervised and unsupervised classification) to other recovery processes to validate it and verify if it is consistent with what is observed by more traditional means. We suggest an unbiased way of doing this to compare the results of the SA analysis to more quantified information such as depopulation or increases in people experiencing negative impacts such as mental health issues and comparing both at the time of the disaster and during the anniversary.

As authors, we would like to call to citizens that during the anniversaries of earthquakes or any other disaster, please share on SM your opinions, feelings, and attitudes toward the post-disaster recovery process to facilitate its assessment improvement of these kinds of processes.

Data and resources

Contreras, D., Wilkinson, S., Balan, N., & James, P. (2021). Polarity supervised classification of Twitter data related to the 10th anniversary of the L'Aquila 2009 earthquake. Retrieved from: https://data.ncl.ac.uk/articles/dataset/Polarity_supervised_classification_of_Twitter_data_related_to_the_10th_anniversary_of_the_L_Aquila_2009_earthquake/14579196
<https://doi.org/10.25405/data.ncl.14579196>

Declaration of conflicting interests

The author(s) declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

Funding

The author(s) disclosed receipt of the following financial support for the research, authorship, and/or publication of this article: The authors would like to express their thanks to the Engineering and Physical Sciences Research Council (EPSRC) (Grant No.: EP/P025641/1). We appreciate the support of Dr. Agostino Bruno from Newcastle University to find the statistical information about the gradual depopulation of L'Aquila and Dr. Antonio Torrisi and Dr. Giuseppe Forino for their support with the Italian language. We also thank Mr. Javier Hervás Ciudad for his support with the images edition.

References

- AbruzzoWeb (2019) L'Aquila: Ten years less than 4 thousand residents in the historical city center. Home/chronicle/earthquake and reconstruction. Available at: <https://www.abruzzoweb.it/contenuti/notizia/685518-302/> (accessed 29 May 2020).
- Alexander D (2010) The L'Aquila earthquake of 6 April 2009 and Italian government policy on disaster response. *Journal of Natural Resources Policy Research* 2(4): 325–342.
- Alexander D (2013) An evaluation of medium-term recovery processes after the 6 April 2009 earthquake in L'Aquila, Central Italy. *Environmental Hazards-Human and Policy Dimensions* 12(1): 60–73.
- Bica M, Palen L and Bopp C (2017) Visual representations of disaster. In: *Proceedings of the 2017 ACM conference on computer supported cooperative work and social computing*, Portland, OR, 25 February–1 March, pp. 1262–1276. New York: Association for Computing Machinery.
- Carboni G (2019) *Tv7*. Padova: Tv7.
- Castellani G (2019) And then came the hail ... L'Aquila 2009/2019. *Agora Magazine*. http://www.agoramagazine.it/index.php?option=com_k2&view=item&id=31815:e-dopo-venne-la-grandine-l-aquila-2009-2019&Itemid=485
- Cervone G and Hultquist C (2018) Citizen as indispensable sensors during disasters. Population-Environment Research Network Cyberseminar, People and Pixels Revisited http://geoinf.psu.edu/publications/2018_PERN_CitizenScience_Cervone.pdf (accessed on July 26, 2021).
- Cervone G, Sava E, Huang Q, Schnebele E, Harrison J and Waters N (2016) Using Twitter for tasking remote-sensing data collection and damage assessment: 2013 Boulder flood case study. *International Journal of Remote Sensing* 37(1): 100–124.
- Contreras D (2016) Fuzzy boundaries between post-disaster phases: The case of L'Aquila, Italy. *International Journal of Disaster Risk Science* 7(3): 277–292.
- Contreras D, Blaschke T and Hodgson ME (2017) Lack of spatial resilience in a recovery process: Case L'Aquila, Italy. *Technological Forecasting and Social Change* 121: 76–88.
- Contreras D, Blaschke T, Kienberger S and Zeil P (2013) Spatial connectivity as a recovery process indicator: The L'Aquila earthquake. *Technological Forecasting and Social Change* 80(9): 1782–1803.

- Contreras D, Blaschke T, Kienberger S and Zeil P (2014) Myths and realities about the recovery of L'Aquila after the earthquake. *International Journal of Disaster Risk Reduction* 8(0): 125–142.
- Contreras D, Blaschke T, Tiede D and Gilge M (2016) Monitoring recovery after earthquakes through the integration of remote sensing, GIS, and ground observations: The case of L'Aquila (Italy). *Cartography and Geographic Information Science* 43(2): 115–133.
- Contreras D, Forino G and Blaschke T (2018) Measuring the progress of a recovery process after an earthquake: The case of L'Aquila, Italy. *International Journal of Disaster Risk Reduction* 28: 450–464.
- Contreras D, Wilkinson S, Balan N, Phengsuwan J and James P (2020) Assessing post-disaster recovery using sentiment analysis. In: *The case of L'Aquila, Haiti, Chile and Canterbury 17th world conference on earthquake engineering (17WCEE)*, Sendai, Japan, 13–18 September.
- Editorial (2019a) Fossacesia, a square named after the city of L'Aquila. Available at: <https://www.abruzzonews.eu/fossacesia-un-piazzale-intitolato-alla-citta-di-laquila-555779.html> (accessed 10 June 2020).
- Editorial (2019b) The Red Zone: On the Nine the special on L'Aquila, 10 years later. Available at: <https://www.oblo.it/la-zona-rossa-sul-nove-laquila-10-anni-dopo/> (accessed 23 May 2020).
- Eyre R, De Luca F and Simini F (2020) Social media usage reveals recovery of small businesses after natural hazard events. *Nature Communications* 11(1): 1629.
- Fallou L, Bossu R, Landès M, Roch J, Roussel F, Steed R and Julien-Laferrière S (2020) Citizen seismology without seismologists? Lessons learned from Mayotte leading to improved collaboration. *Frontiers in Communication* 5: 49.
- Fantasia G (2020) L'Aquila bella me: Ten years later what is missing is not “the” city where we born, but “that” city. Available at: https://www.huffingtonpost.it/2019/04/05/laquila-bella-medicci-anni-dopo-cio-che-manca-non-e-la-citta-dove-siamo-nati-ma-quella-citta_a_23706896/?ncid=tweetlnkithpmsg00000001 (accessed 4 May 2020).
- Fattitaliani.it (2019) L'Aquila ten years later, Piovani symphonic work for the tenth anniversary of the earthquake. Available at: <https://www.fattitaliani.it/2019/04/laquila-dieci-anni-dopo-opera-sinfonica.html?spref=tw> (accessed 24 May 2020).
- Fois F and Forino G (2014) The self-built ecovillage in L'Aquila, Italy: Community resilience as a grassroots response to environmental shock. *Disasters* 38(4): 719–739.
- Fonzi M (2019) *In L'Aquila, Girls and Boys Ask to Return to Their Schools*. Roma: Internazionale.
- Gandomi A and Haider M (2015) Beyond the hype: Big data concepts, methods, and analytics. *International Journal of Information Management* 35(2): 137–144.
- Goodchild M and Li L (2012) Assuring the quality of volunteered geographic information. *Spatial Statistics* 1: 110–120.
- Grammarly (2009) *Grammarly: Free Online Writing Assistant*. San Francisco, CA: United States.
- Graziani MR (2019) L'Aquila, who are the citizens who arrived after the earthquake: The story of Michela. Available at: <https://www.dire.it/04-04-2019/317608-laquila-10-anni-dal-terremoto-cittadini/> (accessed 23 May 2020).
- Guerrieri A (2019) Earthquake: Shipyards opened 10 years later—L'Aquila only half reborn. Available at: <https://www.avvenire.it/attualita/pagine/cantieri-aperti-10-anni-dopo-laquila-rinata-soltanto-a-met> (accessed 23 May 2020).
- Gurman TA and Ellenberger N (2015) Reaching the global community during disasters: Findings from a content analysis of the organizational use of Twitter after the 2010 Haiti earthquake. *Journal of Health Communication* 20(6): 687–696.
- Hagenau M, Liebmann M and Neumann D (2013) Automated news reading: Stock price prediction based on financial news using context-capturing features. *Decision Support Systems* 55(3): 685–697.
- Hausmann A, Toivonen T, Fink C, Heikinheimo V, Kulkarni R, Tenkanen H and Di Minin E (2020) Understanding sentiment of national park visitors from social media data. *People and Nature* 2: 750–760.
- Honjo Y (2011) Implementation of the Kobe City recovery plan. *Japan Social Innovation Journal* 1(1): 1–11.

- Imperiale AJ and Vanclay F (2019) Command-and-control, emergency powers, and the failure to observe United Nations disaster management principles following the 2009 L'Aquila earthquake. *International Journal of Disaster Risk Reduction* 36: 101099.
- INGENIO (2019) *L'Aquila 10 anni dopo: Il Reportage*. San Francisco, CA: INGENIO.
- Jamali M, Nejat A, Ghosh S, Jin F and Cao G (2019) Social media data and post-disaster recovery. *International Journal of Information Management* 44: 25–37.
- Kryvasheyeu Y, Chen H, Moro E, Van Hentenryck P and Cebrian M (2015) Performance of social network sensors during Hurricane Sandy. *PLoS ONE* 10(2): e0117288.
- Kumar S, Yadava M and Roy PP (2019) Fusion of EEG response and sentiment analysis of products review to predict customer satisfaction. *Information Fusion* 52: 41–52.
- Laituri M and Kodrich K (2008) On line disaster response community: People as sensors of high magnitude disasters using Internet GIS. *Sensors* 8(5): 3037–3055.
- Lardinois F (2019) Grammarly gets a tone detector to keep you out of email trouble. *Join Extra Crunch*, 24 September. Available at: <https://techcrunch.com/2019/09/24/grammarly-gets-a-tone-detector-to-keep-you-out-of-email-trouble/2021> (accessed on July 26, 2021).
- Maks I and Vossen P (2012) A lexicon model for deep sentiment analysis and opinion mining applications. *Decision Support Systems* 53: 680–688.
- Maksimava M (2020) Sentiment analysis: What is it and how does it work? *Awario*, 28 September. <https://awario.com/blog/sentiment-analysis/> (accessed on July 26, 2021).
- Mäntylä MV, Graziotin D and Kuutila M (2018) The evolution of sentiment analysis: A review of research topics, venues, and top cited papers. *Computer Science Review* 27: 16–32.
- Medhat W, Hassan A and Korashy H (2014) Sentiment analysis algorithms and applications: A survey. *Ain Shams Engineering Journal* 5(4): 1093–1113.
- Mejri O, Menoni S, Matias K and Aminoltaheri N (2017) Crisis information to support spatial planning in post disaster recovery. *International Journal of Disaster Risk Reduction* 22: 46–61.
- Mendoza M, Poblete B and Valderrama I (2019) Nowcasting earthquake damages with Twitter. *EPJ Data Science* 8(1): 3.
- MonkeyLearn (2020a) Sentiment analysis: Guides. <https://monkeylearn.com/sentiment-analysis/> (accessed on July 26, 2021).
- MonkeyLearn (2020b) Topic analysis. *MonkeyLearn*. <https://monkeylearn.com/topic-analysis/> (accessed on July 26, 2021).
- MonkeyLearn (2021) keyword extraction. *MonkeyLearn*. <https://monkeylearn.com/keyword-extraction/> <https://towardsdatascience.com/basic-binary-sentiment-analysis-using-nltk-c94ba17ae386> (accessed on July 26, 2021).
- Munir S (2019) Basic sentiment analysis using NLTK. *Towards Data Science*, 16 March.
- Neppalli VK, Caragea C, Squicciarini A, Tapia A and Stehle S (2017) Sentiment analysis during Hurricane Sandy in emergency response. *International Journal of Disaster Risk Reduction* 21: 213–222.
- Nicola AD (2019) L'Aquila, the shame of the classrooms still unusable between scaffolding and rubble. Available at: https://www.ilmessaggero.it/italia/l_aquila_terremoto_scuole_inagibili-4412907.html (accessed 11 September 2020).
- Özerdem A and Rufini G (2013) L'Aquila's reconstruction challenges: Has Italy learned from its previous earthquake disasters? *Disasters* 37(1): 119–143.
- Pacione M (2009) *Urban Geography: A Global Perspective*. London: Routledge.
- Ragini JR, Anand PMR and Bhaskar V (2018) Big data analytics for disaster response and recovery through sentiment analysis. *International Journal of Information Management* 42: 13–24.
- RepubblicaTv (2019) L'Aquila 10 anni dopo il terremoto: Roberta, presidente di consorzio: “Restituire le chiavi di casa è l'emozione più bella.” L'Aquila 10 anni dopo il terremoto—Roberta, presidente di consorzio: “Restituire le chiavi di casa è l'emozione più bella.” @RepubblicaTv. <https://twitter.com/repubblicatv/status/1113938561294712832> (accessed on July 26, 2021).
- Roldós I (2020) Best word visualization tools to create your own word cloud. *MonkeyLearn Blog*. <https://monkeylearn.com/blog/word-visualization/> (accessed on July 26, 2021).
- Roldós I (2021) NLP, AI, and machine learning: What's the difference? *MonkeyLearn*. <https://monkeylearn.com/blog/nlp-ai/> (accessed 21 June 2021).

- Rossetto T, D'Ayala D, Gori F, Persio R, Han J, Novelli V, Wilkinson SM, Alexander D, Hill M, Stephens S, Kontoe S, Elia G, Verrucci E, Vicini A, Shelley W and Foulser-Piggott R (2014) The value of multiple earthquake missions: The EEFIT L'Aquila Earthquake experience. *Bulletin of Earthquake Engineering* 12(1): 277–305.
- Rossi E (2019) L'Aquila, ten years after the earthquake, is a completely peripheral city. Available at: <https://contropiano.org/altro/2019/04/04/laquila-dieci-anni-dopo-il-sisma-e-una-citta-del-tutto-periferica-0114120> (accessed 23 May 2020).
- Sánchez-Rada JF and Iglesias CA (2019) Social context in sentiment analysis: Formal definition, overview of current trends and framework for comparison. *Information Fusion* 52: 344–356.
- Schroeder R (2014) Big Data and the brave new world of social media research. *Big Data & Society* 1(2): 563194.
- Schumann RL (2018) Ground truthing spatial disaster recovery metrics with participatory mapping in post-Katrina Mississippi. *Applied Geography* 99: 63–76.
- Shibuya Y and Tanaka H (2019) Using social media to detect socio-economic disaster recovery. *IEEE Intelligent Systems* 34(3): 29–37.
- Smith GP and Wenger D (2007) Sustainable disaster recovery: Operationalizing an existing agenda. In: Roringuez H, Donner W and Trainor JE (eds) *Handbook of Disaster Research: Handbooks of Sociology and Social Research*. New York: Springer, pp. 234–257.
- Stecanella B (2020) Tutorial on sentiment analysis with Python. *MonkeyLearn*. <https://monkeylearn.com/blog/sentiment-analysis-with-python/> (accessed on July 26, 2021).
- TGCOM24 (2019) L'Aquila ten years after the 2019 earthquake: Still barriers and rubble. *Chronicle*, 5 April. Available at: https://www.tgcom24.mediaset.it/cronaca/foto/l-aquila-a-dieci-anni-dal-terremoto-del-2009-ancora-transenne-e-macerie_3097137-2019.shtml?fbclid=IwAR2im6LI0ajLJHIMUx5dybTksI2aTKoIRzyBrjVpcIpMDxAY8oVkB0re2Q (accessed 29 May 2020).
- TweetBinder (2019) Tweet Binder. Available at: <https://www.tweetbinder.com/> (accessed 16 April 2019).
- Wilkinson S, Stone H, D'Ayala D, Verrucci E, James P, Rossetto T, So E and Ellul C (2018) How can new technologies help us with earthquake reconnaissance. In: *Proceedings of the 11th national conference in earthquake engineering, Earthquake Engineering Research Institute*, Los Angeles, CA, 25–28 June.
- Wolff R (2020a) Free AI-powered word cloud tool: MonkeyLearn. *MonkeyLearn*. <https://www.youtube.com/watch?v=QiPwnsw371> (accessed on July 26, 2021).
- Wolff R (2020b) Visualize sentiments in a word cloud: Free tools & tutorial. *MonkeyLearn Blog*. <https://monkeylearn.com/blog/word-cloud-sentiment-analysis/> (accessed on July 26, 2021).
- Wu D and Cui Y (2018) Disaster early warning and damage assessment analysis using social media data and geo-location information. *Decision Support Systems* 111: 48–59.
- Xu T, Peng Q and Cheng Y (2012) Identifying the semantic orientation of terms using S-HAL for sentiment analysis. *Knowledge-Based Systems* 35: 279–289.
- Yan Y, Chen J and Wang Z (2020) Mining public sentiments and perspectives from geotagged social media data for appraising the post-earthquake recovery of tourism destinations. *Applied Geography* 123: 102306.
- Yan Y, Eckle M, Kuo C-L, Herfort B, Fan H and Zipf A (2017) Monitoring and assessing post-disaster tourism recovery using geotagged social media data. *ISPRS International Journal of Geo-information* 6(5): 144.
- Yu L-C, Wu J-L, Chang P-C and Chu H-S (2013) Using a contextual entropy model to expand emotion words and their intensity for the sentiment classification of stock market news. *Knowledge-Based Systems* 41: 89–97.
- Yuan F and Liu R (2018) Feasibility study of using crowdsourcing to identify critical affected areas for rapid damage assessment: Hurricane Matthew case study. *International Journal of Disaster Risk Reduction* 28: 758–767.
- Zucco C, Calabrese B, Agapito G, Guzzi PH and Cannataro M (2020) Sentiment analysis for mining texts and social networks data: Methods and tools. *Wires Data Mining and Knowledge Discovery* 10(1): e1333.