ASSESSING POST-DISASTER RECOVERY USING SENTIMENT AND TOPIC ANALYSIS ON SOCIAL MEDIA POSTS: L'AQUILA, ITALY

D. Contreras[1], G. Forino[2], C. Pezzica[1], F. Liberatore[1], S. Wilkinson[3]

[1] Cardiff University, Cardiff, United Kingdom, contrerasmojicad@cardiff.ac.uk

[2] University of Salford, Manchester, United Kingdom

[3] Newcastle University, Newcastle upon Tyne, United Kingdom

Abstract *In this conference paper, we assess the progress in post-disaster recovery by analysing 4349 tweets posted between 4th and 10th April 2019 that we collected around the 10th anniversary of the earthquake in L'Aquila. Text data collected from social media is unstructured; therefore, we need to use natural language processing techniques such as topic and sentiment analysis to extract meaningful information to assess recovery. Sentiment Analysis (SA), or opinion mining, classifies people's opinions, expressed in written text, into a specific polarity, i.e. positive, negative or neutral. Topic Analysis (TA) classifies data into recurrent themes or topics users address in their posts. These analyses can be done at the tweet or sentence level, and the classification into polarities and topics can be done manually or automated. We analysed a sample of 10% of tweets covering 24 hours daily. The SA and TA were performed at both levels, and the classification was done manually. The SA at the tweet level indicates that most of the posts were classified into neutral polarity, followed closely by positive and negative. Similar results were obtained from the SA analysis at the sentence level, only with variation in percentages of the sentence classified into each polarity. The TA at the tweet and sentence levels indicate that the most frequently addressed topics by users at both levels were commemoration actions, restoration, reconstruction, governance, and distress. The SA per topics at the tweet level indicates that the topics with neutral polarity are critical infrastructure, commemoration actions, and seismic information. Topics with positive polarity are cultural heritage, early recovery, emergency response, lifelines, preparedness, restoration, solidarity messages/actions and urban facilities, which are considered successful aspects of the recovery process in this methodology. Topics with negative polarity are building damages, construction practices, depopulation and displacement, distress, governance, injuries and casualties, intensity, prevention, and reconstruction, considered the failures of the process. We also conducted a two-tailed Pearson correlation analysis between polarity and topics of tweets for each day, which confirmed, in most cases, the results of the SA for each topic at the tweet level. According to the methodology applied, we can conclude that the perception of the recovery of L'Aquila by the 10th anniversary is mainly neutral.*

1. Introduction

The use of social media data extracted from platforms such as Twitter/X and Facebook in the field of disaster management (Radianti *et al.* 2016)) has increased (Xiao *et al.* 2015); however, its research potential has not yet been fully explored (Ogie *et al.* 2022). 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 (Contreras *et al.* 2021).

On April 6th, 2019, the first author collected tweets about the 10th anniversary of the L’Aquila earthquake (a trending topic on Twitter at that time). Those tweets reflected user perception relating to the recovery process of L’Aquila after the earthquake, and here we test if those impressions could be used to evaluate the progress of the recovery process in this city and any other area affected by a major earthquake (Contreras *et al.* 2020).

Psychological and physical rehabilitation processes are part of post-disaster recovery. Perceptions and people's sentiments regarding post-disaster recovery have been addressed by a few studies. Mixed methods combining sentiment analysis (SA) and topic modelling were used by Yang et al. (2020). These authors analysed tweets posted by nonlocal Twitter users after the 2018 earthquakes about frustration with the housing reconstruction, living conditions and post-disaster tourism recovery in Lombok and Bali, Indonesia.

In 2024, it will already be 15 years since the 2009 L’Aquila earthquake, which means a new opportunity to assess the recovery process having the progress reported on the tweets for the 10th anniversary as a base line, combining SA and TA for the post-disaster recovery assessment. This is the main contribution of this conference paper considering that empirical studies regarding the use of social media to formulate or evaluate post-disaster recovery strategies have been undertaken by very few authors so far (Yang et al. 2020).

2. Methods

To assess the progress in the recovery, the first and last author collected 4,349 tweets with the hashtags #L’Aquila, #Laquila, #laquila, #LAquila10annidopo, #terremoto, #6aprile, #PortamiDoveSeiNata, among others posted between April 4th and 10th in 2019. Text data collected from social media platforms like Twitter is unstructured; therefore, we need to use natural language processing (NLP) techniques (Roldós 2021), such as sentiment and topic analysis, to extract meaningful information to assess recovery. Sentiment analysis, or opinion mining (Shibuya 2020), classifies people's opinions, expressed in written text, into a specific polarity, i.e. positive, negative or neutral. Topic Analysis (TA) classifies data into recurrent themes or topics users address in their posts (MonkeyLearn 2022). SA and TA can be done at document, sentence or aspect level (Liu 2015), and the classification into polarities and topics can be done manually or automated. However, it is necessary to count with curated datasets labelled by human annotators to train language models (Wolf *et al.* 2020), which are further finetuned for automated SA and TA (Antypas *et al.* 2022).

In this conference paper, we analysed a sample of 10% (436) of tweets posted during the mentioned period, covering 24 hours. The SA and TA were performed at tweet and sentence level, and the classification was done manually jointly by authors who are Italian experts in human geography, architecture and urbanism, data and knowledge engineering, and the first and last author with expertise in earthquake reconnaissance and post-disaster recovery after earthquakes. Besides their expertise, all authors are familiar with the recovery process of L’Aquila, ensuring their knowledge of the case study and their competence in the classification of the text data for the SA and TA. The flow chart with the methodology is presented in Figure 1.

2.1. Sentiment analysis

Rules for classifying tweets into a specific polarity were agreed upon among the authors, and the SA was applied at the document or tweet level and the sentence level. These rules are listed in Table 1. The authors analysed changes in polarity for each day in the sample during the observation period.

2.2. Topic analysis

Rules for classifying tweets into a specific topic were agreed upon among the authors and applied the TA (also at the tweet and sentence level). These rules are listed in Table 2. The authors analysed changes in topics addressed each day in the sample during the observation period.

2.3. Polarity for each topic

The results of the SA and TA are combined to determine the polarities for each topic to identify the successes and failures of the recovery process based on the highest polarity for each topic, e.g. if the highest polarity in the topic ‘reconstruction’ is negative, then we will assume that reconstruction has been one of the failures in the recovery process.

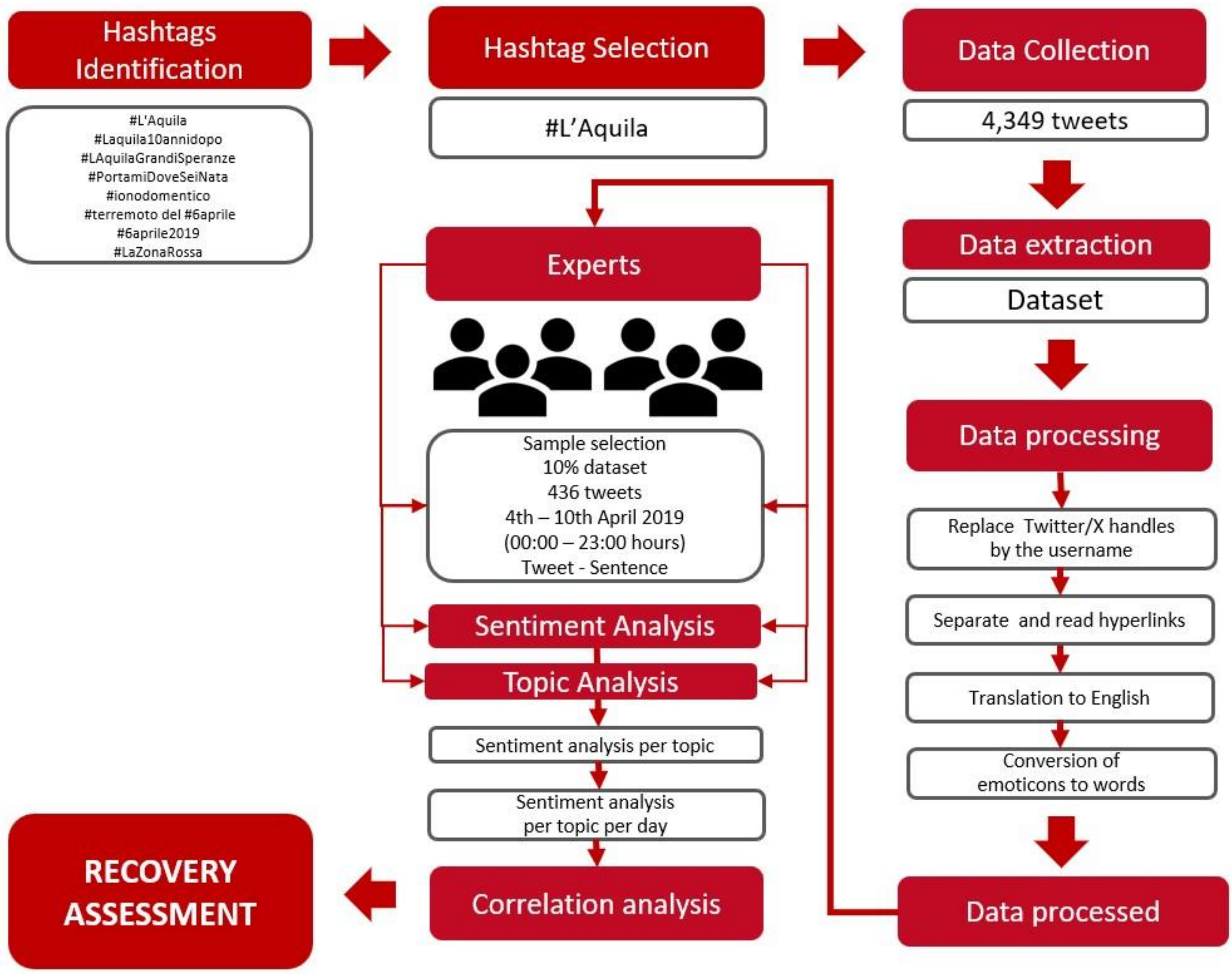


Figure 1. Methodology.

Table 1. Classification rules set for sentiment analysis. Adapted from Contreras et al. (2021)

Polarity	Rules
Positive	• Announcements of opening of new business and job opportunities. • Calls to not forget what happened. • Comments of people moving to L'Aquila after the earthquake. • Mentions of buildings already reconstructed. • Learning the lessons from the event. • Praises to the value of the renaissance architecture and monuments in the city. • Promotion of products of the region and sports events e.g. rugby and football. • Solidarity messages. • Stories of survivors and rescue teams.
Negative	• Comments about depopulation in the city centre despite the reconstruction efforts. • Complains about the delay in the reconstruction. • Complains about the lack of urban facilities in the city centre. • Complains about the mismanagement of the financial resources for the reconstruction. • Expressions of inability to forget the impact of the earthquake. • Mentions of the existence of cordoned houses, rubble, and barriers. • References to victims.
Neutral	• Mention of commemoration ceremonies or actions to honour the victims. • Magnitude, aftershocks, and geological changes caused by the earthquake.

Table 2. Classification rules set for topic analysis. Adapted from Contreras et al. (2023)

Topic	Rules
Building damages	Report of additional damages in buildings or buildings still damaged.
Construction practices	Housing quality after the earthquake.
Commemoration actions	Interviews with survivors, torchlight processions, minutes of silence, etc.
Critical infrastructure (CI)	Facilities needed to respond to the emergency, e.g. Health care posts.
Cultural heritage	Physical artifacts and intangible attributes inherited from the past.
Depopulation & Displacement	Reduction in the number of inhabitants after the earthquake.
Distress	Expressions of sorrow and pain about the earthquake.
Early recovery	Actions taken to return to normality e.g. cleaning debris.
Emergency response	Actions to save lives, e.g. search and rescue (SAR) activities.
Funding	Sources and management of money allocated for the recovery.
Geotechnical effects	E.g. landslides, rockfalls and cracks on the soil.
Governance	Role assumed by the government during the post-disaster phase.
Hate	Expressed prejudice against protected characteristics.
Injuries & casualties	Mentions of casualties or injured population due to the earthquake.
Intensity	Severity of ground shaking without mentioning a magnitude.
Lifelines	E.g. Water, electricity, gas, communication and roads.
Preparedness	Anticipated actions to respond to an emergency.
Prevention	Actions to avoid potential adverse impacts of hazards.
Reconstruction	Rebuilding of houses, infrastructure and/or monuments.
Restoration	Restoring sustainable living conditions in the socio-economic dimension.
Seismic information	Date, magnitude, epicentre, and depth of the earthquake.
Solidarity messages/actions	Encouraging messages to survivals.
Urban Facilities	Facilities different to CI, e.g. Schools, temples, post offices, etc.
Unrelated	Topics not related to the anniversary of the earthquake.

2.4. Correlation analysis

In addition to classifying tweets into a polarity and a topic, we have also conducted a two-tailed Pearson correlation analysis of polarity for each topic for each day. We argue that this analysis will confirm the relationship between polarities and topics at tweet level, previously identified in the analysis of SA for each topic.

3. Results

3.1. Sentiment analysis

We present the results of the SA in Figure 2 at a) tweet level and b) sentence level for the entire set of tweets. From the figure, it can be observed that there is not a substantial difference in the results of the SA at the tweet and sentence level; in both cases, the highest polarity identified is neutral, followed by positive, negative and unrelated tweets, as observed in Figure 2. The results of the SA at the tweet level for each day are presented in Figure 3. The plot of the polarity analysis on each day indicates that, on most days, the highest polarity is positive, except for the exact day of the anniversary when the highest polarity is neutral, followed closely by negative polarity, and the last day of the observation period when again the most frequent polarity is neutral.

3.2. Topic analysis

We present the results of the topic analysis at the tweet level in Figure 4. On the one hand, the most frequent topics addressed at the tweet level in the sample were commemoration actions, followed by solidarity messages/actions, restoration, reconstruction, unrelated tweets, governance, distress, and others. On the other hand, the most frequent topics addressed at the sentence level in the sample were seismic information, commemoration actions, restoration, injuries and casualties, reconstruction, governance, and distress, followed by others, as indicated in Figure 5. In both cases, topics such as commemoration actions, restoration, reconstruction, governance, and distress are included in the most frequent topics addressed by users on tweets about the 10th anniversary of the L'Aquila earthquake. Geotechnical effects and hate are not the main topics in any tweet and are the less frequently addressed topics by the TA at the sentence level. The topic of seismic information is the most frequent at the sentence level but one of the less mentioned at the tweet level.

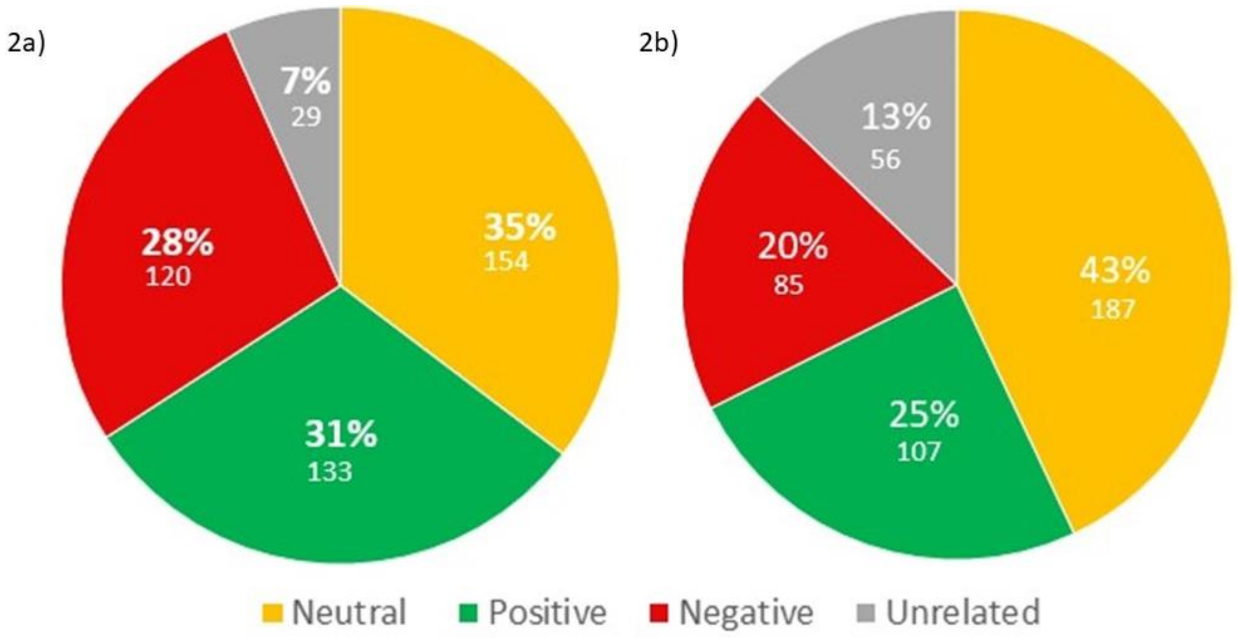
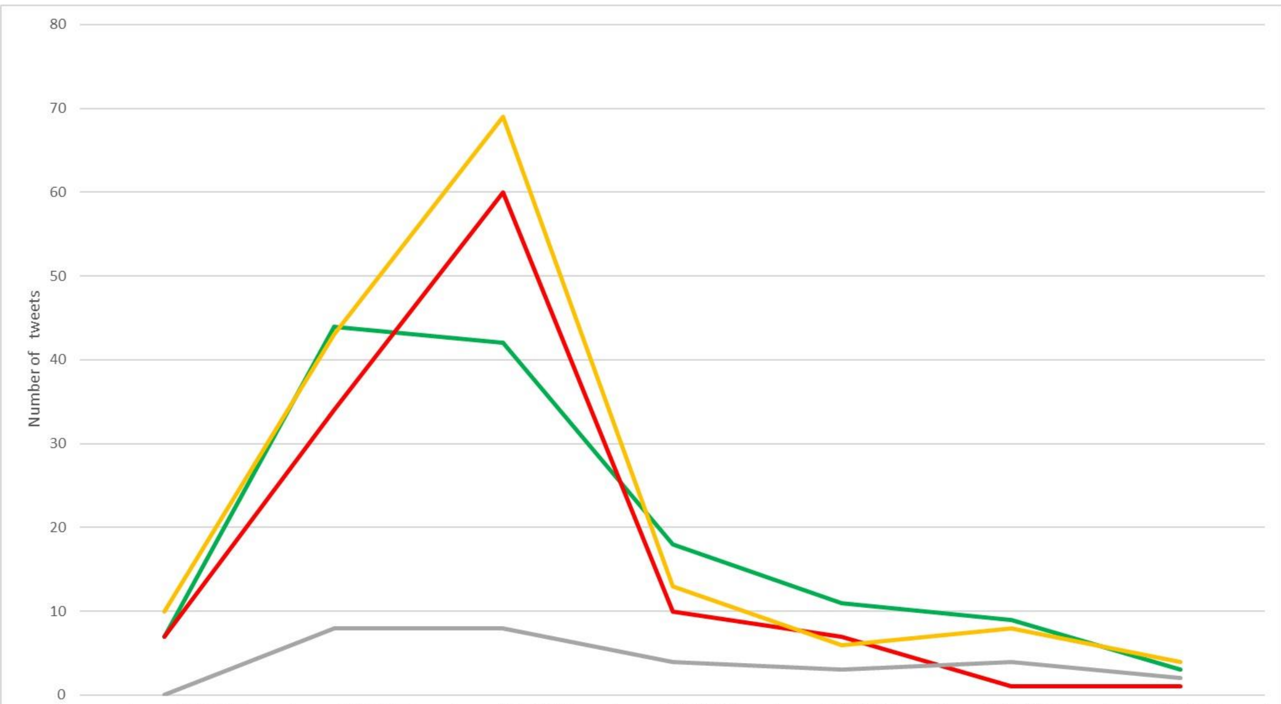


Figure 2. Sentiment analysis result at a) tweet level and b) sentence level.



	04/04/2019	05/04/2019	06/04/2019	07/04/2019	08/04/2019	09/04/2019	10/04/2019
Positive	7	44	42	18	11	9	3
Negative	7	34	60	10	7	1	1
Neutral	10	43	69	13	6	8	4
Unrelated	0	8	8	4	3	4	2

Figure 3. Sentiment analysis result at tweet level for each day.

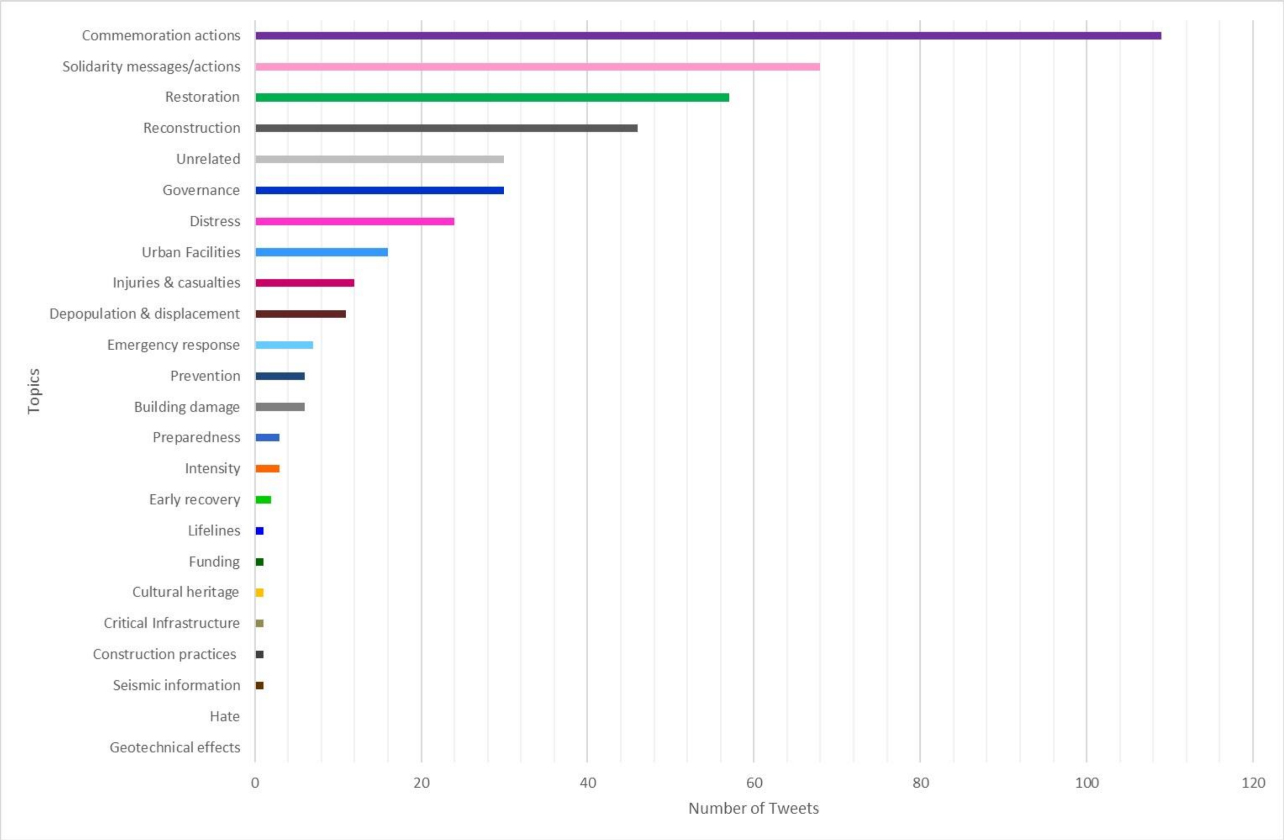


Figure 4. Topic analysis result at tweet level.

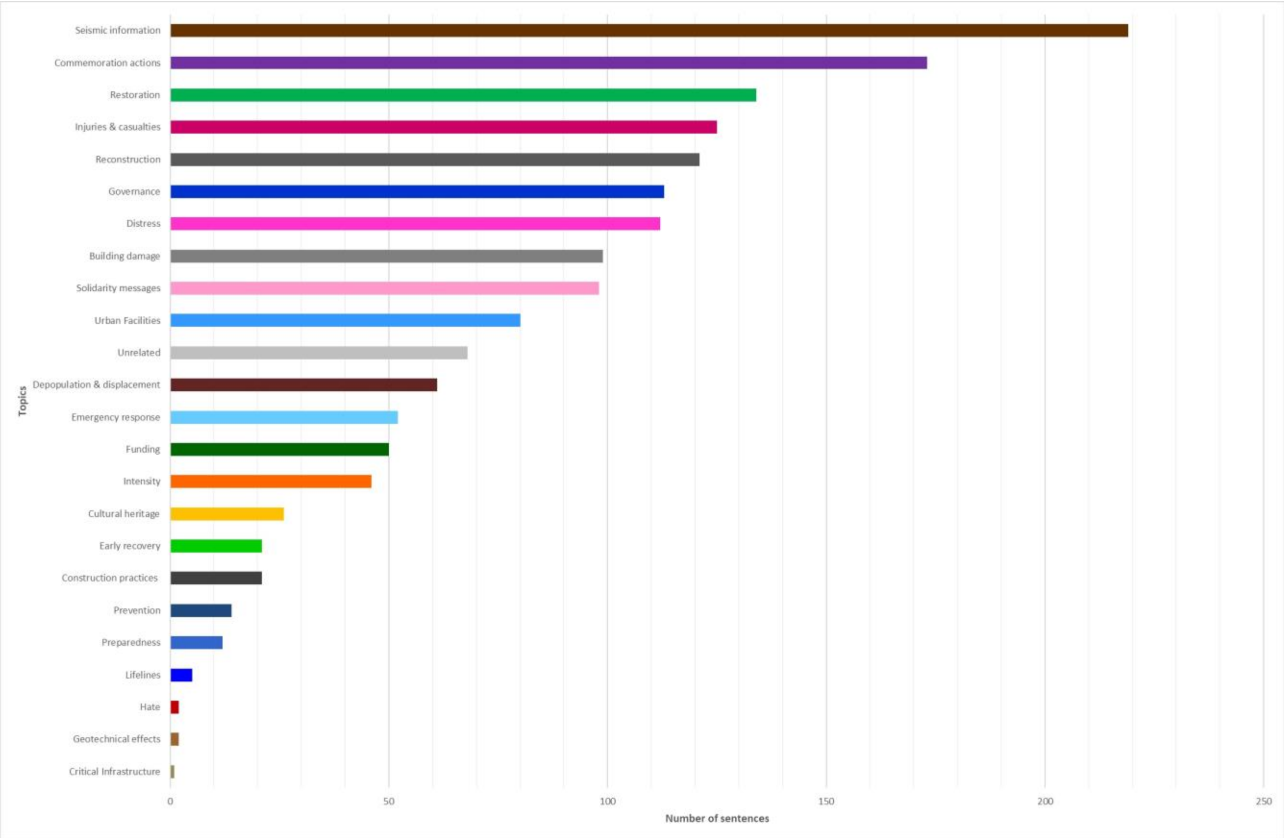


Figure 5. Topic analysis result at sentence level.

Since restoration is the most frequently addressed topic at the tweet and sentence levels, we produced a cloud word to highlight the key aspects identified by Twitter users as contributing to the restoration of sustainable living conditions in the socio-economic dimension. Those words are highlighted in Figure 6. The font size and the colour tone represent the frequency of the phrase mentioned in the tweets related to the anniversary of the earthquake.

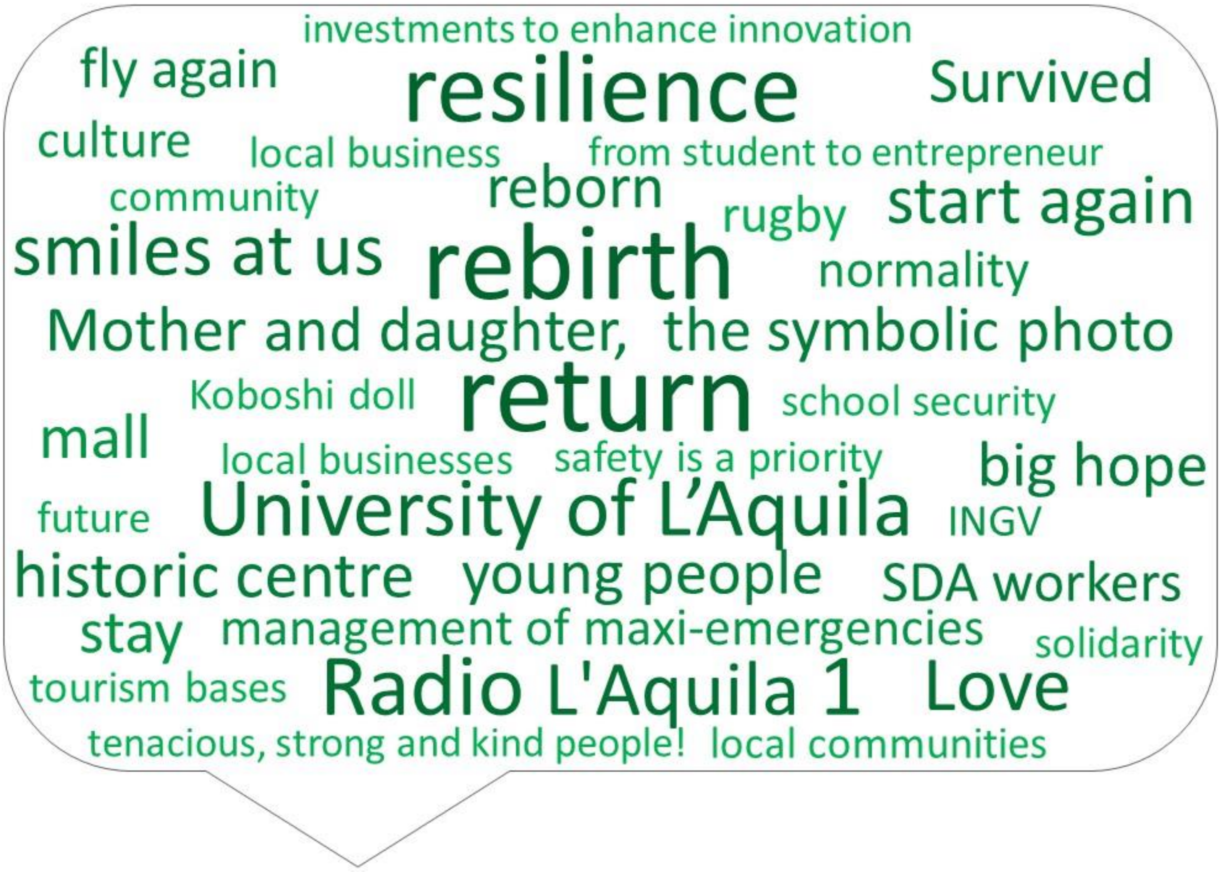


Figure 6. Key words extraction from tweets addressing the topic of restoration at tweet and sentence level.

Results of the TA for each day indicate that commemoration actions were the most frequent topic at the tweet level during most days of the observation period, including the exact day of the 10th anniversary, except for the fourth day, 7th April 2019, when the most frequent topic of tweets changed to restoration. On 8th April 2019, there was a tie between commemoration actions and reconstruction as the most frequent topics; the same happened on April 9th with a tie between commemoration actions and restoration.

We present the TA results at tweet level for each day in Figure 7 (and for clarity in Table 3). Restoration is the second most frequent topic at the tweet level during the first two days of the observation period. This changes for the date of the anniversary when the second most frequent topic is solidarity messages/actions and continues for April 7th with a tie with commemoration actions. On April 8th, several topics appeared as the second most frequent: restoration, solidarity messages/actions, governance, and unrelated tweets. On April 9th, there was a tie again as the second most important topics, reconstruction and unrelated, and the same happened on April 10th with the topics of urban facilities and unrelated tweets.

3.3. Sentiment analysis for each topic

The SA for each day at the tweet level indicates that the topics with neutral polarity as the highest polarity are critical infrastructure (CI), commemoration actions, and seismic information. The same analysis at the tweet level indicates that the topics with positive polarity as the highest polarity are cultural heritage, early recovery, emergency response, lifelines, preparedness, restoration, solidarity messages/actions and urban facilities, considered successful aspects of the recovery process in this methodology. The SA for each day at the tweet level shows that the topics with negative polarity as the highest polarity are building damages, construction practices, depopulation and displacement, distress, governance, injuries and casualties, intensity, prevention and reconstruction, considered the failures of the process. The numbers of tweets classified into a specific polarity for each topic are listed in Table 4.

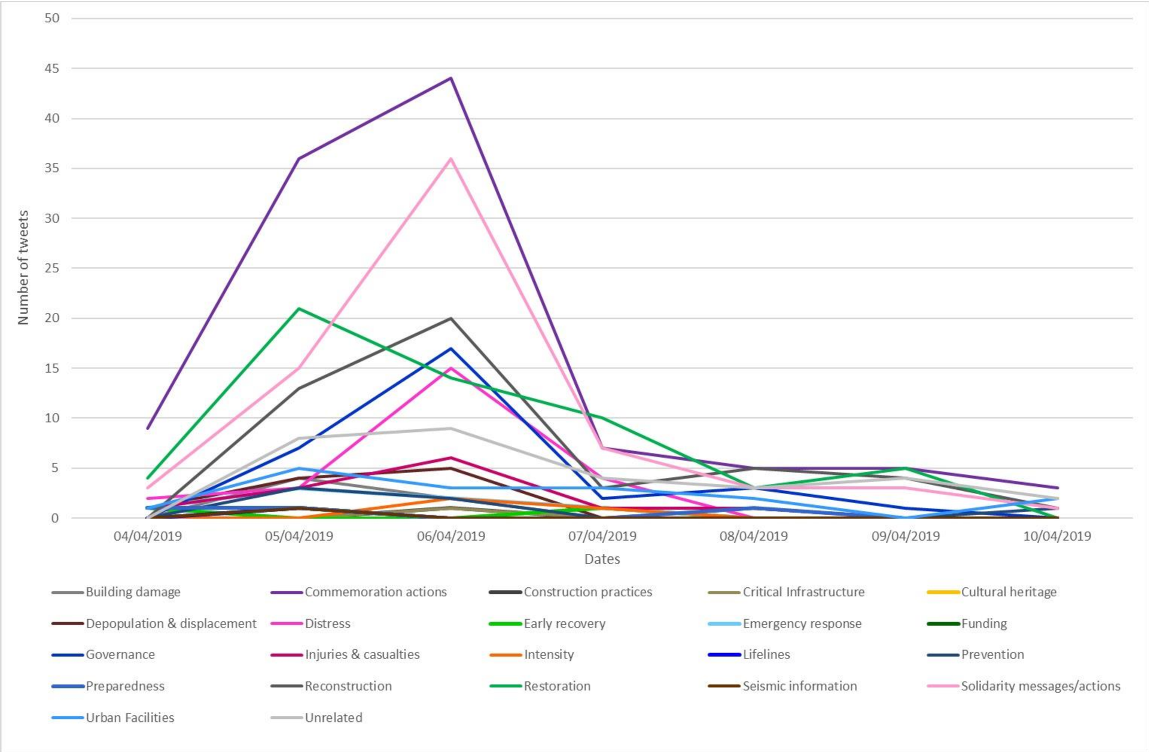


Figure 7. Topic analysis result at tweet level for each day.

Table 3. Topics addressed for each day between April 4th and 10th 2019.

Topic	**4th**	**5th**	**6th**	**7th**	**8th**	**9th**	**10th**	**Total**
Building damages	0	4	2	0	0	0	0	6
Commemoration actions	9	36	44	7	5	5	3	109
Construction practices	0	0	1	0	0	0	0	1
Critical Infrastructure	0	0	1	0	0	0	0	1
Cultural heritage	0	0	0	1	0	0	0	1
Depopulation & displacement	1	4	5	0	1	0	0	11
Distress	2	3	15	4	0	0	0	24
Early recovery	1	0	0	1	0	0	0	2
Emergency response	1	3	2	1	0	0	0	7
Funding	0	1	0	0	0	0	0	1
Governance	0	7	17	2	3	1	0	30
Injuries & casualties	1	3	6	1	1	0	0	12
Intensity	0	0	2	1	0	0	0	3
Lifelines	0	1	0	0	0	0	0	1
Prevention	0	3	2	0	0	0	1	6
Preparedness	1	1	0	0	1	0	0	3
Reconstruction	0	13	20	3	5	4	1	46

Topic	4th	5th	6th	7th	8th	9th	10th	Total
Restoration	4	21	14	10	3	5	0	57
Seismic information	0	1	0	0	0	0	0	1
Solidarity messages/actions	3	15	36	7	3	3	1	68
Urban Facilities	1	5	3	3	2	0	2	16
Unrelated	0	8	9	4	3	4	2	30
Total	24	129	179	45	27	22	10	436

Table 4. Sentiment analysis for each topic at tweet level.

Topics	Positive	Negative	Neutral	Unrelated	Total
Building damages	0	5	1	0	6
Critical infrastructure	0	0	1	0	1
Commemoration actions	19	10	80	0	109
Construction practices	0	1	0	0	1
Cultural heritage	1	0	0	0	1
Depopulation-Displacement	3	8	0	0	11
Distress	1	18	5	0	24
Early recovery	2	0	0	0	2
Emergency response	3	2	2	0	7
Funding	1	0	0	0	1
Governance	3	21	6	0	30
Injuries-casualties	0	6	6	0	12
Intensity	0	2	1	0	3
Lifelines	1	0	0	0	1
Preparedness	2	0	1	0	3
Prevention	1	3	2	0	6
Reconstruction	8	27	11	0	46
Restoration	42	10	5	0	57
Seismic information	0	0	1	0	1
Solidarity messages/actions	34	4	30	0	68
Urban facilities	12	3	1	0	16
Unrelated	0	0	0	30	30
				Total	436

3.4. Correlation analysis

We present the two-tailed Pearson correlation analysis between polarities and topics in Table 5. The correlation analysis between polarities and topics at tweet level for each day indicates that the number of tweets classified into the topic of building damage is highly correlated with tweets classified into the topic of positive polarity (.908**) and with tweets with neutral polarity (.768*) to a lesser degree. The number of tweets classified into the topic of commemoration action is highly correlated with the number of tweets classified into all the polarities, but mainly with the neutral polarity (.982*). The number of tweets classified into construction practices and CI are correlated with the number of tweets with neutral (.843*) but mainly with the number of tweets classified into the negative polarity (.860*) in both cases. The number of tweets classified into the topic of depopulation and displacement is highly correlated with the number of tweets classified into negative (.958**), neutral (.955**) and positive polarity (.920**). The number of tweets classified into the topic of distress is highly correlated with the number of tweets classified into negative (.914**) and neutral (.899**) polarities. The number of tweets classified into the topic of emergency response actions is highly correlated with the number of tweets with positive (.912**) polarity and, to a lesser degree, with tweets classified into neutral (.802*) and negative polarities (.781*). The number of tweets classified into the topic of governance is highly correlated with the number of tweets with negative (.976**) and neutral (.964**) polarity and, to a lesser degree, with the number of tweets classified into positive polarity (.848*). The number of tweets classified into the topic of injuries and casualties is highly correlated with the number of tweets for each day classified into all the polarities; however, this number is mainly correlated with the number of tweets classified into the negative polarity (.994**) followed by those classified into the neutral (979**) and positive polarity (.876**). The number of tweets classified into the topic of intensity of the earthquake are correlated with the number of tweets classified into the negative polarity (.757*). The number of tweets classified into the topic of prevention is correlated with the number of tweets classified into positive (.844*) polarity, and neutral (.778*) polarity but to a lesser degree. The number of tweets classified into the topic of reconstruction is highly correlated with the number of tweets classified into all the polarities: neutral (.964**) and negative (.962**), and positive polarity (.919**). The number of tweets classified into the topic of restoration is highly correlated with the number of tweets classified into positive polarity (.949**) and the number of tweets classified into the neutral (.784*) and negative (.755*) polarity to a lesser degree. The number of tweets classified into the topic of solidarity messages and actions is highly correlated with the number of tweets classified into negative (.983**) and neutral (.978**) polarities and also correlated with tweets classified into positive (.851*) polarity to a lesser degree. The number of tweets classified into the topic of urban facilities is correlated with the number of tweets classified into positive (.790*) polarity. The number of tweets for each day classified into cultural heritage, early recovery, lifelines, preparedness and seismic information is not correlated with the number of tweets classified into any polarity. The results of the correlation analysis are presented in Table 5.

4. Discussion

It is necessary to remember that the results presented in this conference paper are based on a dataset sample and, therefore, must be considered preliminary results. While the sample covered all the hours on daily in the observation period, it still does not represent the entire dataset. We argue that similar results in SA mean that, in this case, further analysis at the sentence level is not necessary, but we do not assert that this will generally be true. The increase in the negative polarity of tweets on the exact day of the anniversary of the earthquake can be explained by the high number of tweets classified into the topics of reconstruction, governance and distress posted on April 6th, 2019. Understandably, being the anniversary of the earthquake, tweets classified into the topic of commemoration actions appear as the majority in the TA at the tweet and sentence level. However, we decided to elaborate on the word cloud using the text data classified into the topic of restoration, because we consider it more relevant for the recovery assessment. We found the SA for each topic at the tweet level useful to visualise the topics on which the recovery process has been successful and the topics that had failed and, therefore, have delayed the recovery process. We expected to confirm the result of this analysis with the correlation analysis between polarity and topics addressed for each day during the observation period. On the one hand, the high correlation between tweets classified into the topics of construction practices, depopulation and displacement, distress, governance, injuries and casualties, intensity and the number of tweets classified into negative polarity was expected. Equally, it was expected between the number of tweets classified into the topic of emergency response, prevention, restoration and urban facilities and tweets classified into positive polarity, according to the results presented in Table 4. On the other hand, we also found inconsistencies in the correlation analysis: the number of tweets classified into the topic of building damage resulted highly correlated with the number of tweets classified into positive polarity, which was not expected. The number of tweets classified into reconstruction was slightly more correlated with the number of tweets classified into neutral than with tweets with negative polarity. The number of tweets classified into solidarity messages and actions was highly correlated with the number of tweets classified into negative polarity and barely correlated with the number of tweets classified into positive polarity, as was expected.

5. Conclusion

According to the methodology applied, we can conclude that the perception of the recovery of L'Aquila after the earthquake by the 10th anniversary is mainly neutral. Twitter users recognize the quick and effective emergency response after the earthquake as well as the efforts done by the government during the early recovery, but are frustrated with the slow reconstruction, the poor quality of the temporary houses and the fact that there are still buildings damages after a decade, which produce stress among the population and have encouraged the depopulation of the city. The same users consider that lessons have been learned regarding construction practices, emergency response and preparedness. It is demonstrated in the word cloud the role of the University of L'Aquila in the recovery of the city and it is expected that this will be the cornerstone of the recovery process with the return of the population to inhabit the city, but for that, it is necessary to reconstruct all the urban facilities necessary to provide services and sources of employment for its residents.

Table 5. Two tailed Pearson correlation analysis between polarities and topics.

		Positive	Negative	Neutral	Building damages	Commemoration actions	Construction practices	Critical infrastructure	Cultural heritage	Depopulation & displacement	Distress	Early recovery	Emergency response	Funding	Governance	Injuries & casualties	Intensity	Lifelines	Prevention	Preparedness	Reconstruction	Restoration	Seismic information	Solidarity messages & actions	Urban Facilities
Positive	Pearson Correlation	1	.912**	.924**	.908**	.955**	0.595	0.595	-0.030	.920**	0.715	-0.268	.912**	0.648	.848*	.876**	0.558	0.648	.844*	0.084	.919**	.949**	0.648	.851*	.790*
	Sig. (2-tailed)		0.004	0.003	0.005	0.001	0.158	0.158	0.949	0.003	0.071	0.561	0.004	0.116	0.016	0.010	0.193	0.116	0.017	0.858	0.003	0.001	0.116	0.015	0.034
Negative	Pearson Correlation	.912**	1	.994**	0.738	.971**	.860*	.860*	-0.143	.958**	.914**	-0.269	.781*	0.338	.976**	.994**	.757*	0.338	0.750	-0.049	.962**	.755*	0.338	.983**	0.609
	Sig. (2-tailed)	0.004		0.000	0.058	0.000	0.013	0.013	0.759	0.001	0.004	0.560	0.038	0.458	0.000	0.000	0.049	0.458	0.052	0.918	0.001	0.050	0.458	0.000	0.147
Neutral	Pearson Correlation	.924**	.994**	1	.768*	.982**	.843*	.843*	-0.158	.955**	.899**	-0.287	.802*	0.378	.964**	.979**	0.734	0.378	.778*	-0.083	.964**	.784*	0.378	.978**	0.595
	Sig. (2-tailed)	0.003	0.000		0.044	0.000	0.017	0.017	0.735	0.001	0.006	0.533	0.030	0.403	0.000	0.000	0.061	0.403	0.040	0.859	0.000	0.037	0.403	0.000	0.158
Building damages	Pearson Correlation	.908**	0.738	.768*	1	.866*	0.320	0.320	-0.240	.848*	0.424	-0.372	.917**	.881**	0.630	0.679	0.192	.881**	.946**	0.283	.765*	.912**	.881**	0.625	.811*
	Sig. (2-tailed)	0.005	0.058	0.044		0.012	0.484	0.484	0.604	0.016	0.343	0.411	0.004	0.009	0.130	0.093	0.680	0.009	0.001	0.538	0.045	0.004	0.009	0.134	0.027
Commemoration actions	Pearson Correlation	.955**	.971**	.982**	.866*	1	0.740	0.740	-0.223	.982**	.809*	-0.305	.877**	0.531	.913**	.948**	0.603	0.531	.854*	0.060	.947**	.845*	0.531	.924**	0.668
	Sig. (2-tailed)	0.001	0.000	0.000	0.012		0.057	0.057	0.631	0.000	0.027	0.506	0.010	0.220	0.004	0.001	0.151	0.220	0.014	0.898	0.001	0.017	0.220	0.003	0.101
Construction practices	Pearson Correlation	0.595	.860*	.843*	0.320	0.740	1	1.000**	-0.167	0.730	.954**	-0.258	0.382	-0.167	.919**	.884**	.881**	-0.167	0.415	-0.354	.814*	0.352	-0.167	.927**	0.196
	Sig. (2-tailed)	0.158	0.013	0.017	0.484	0.057		0.000	0.721	0.062	0.001	0.576	0.398	0.721	0.003	0.008	0.009	0.721	0.355	0.437	0.026	0.439	0.721	0.003	0.673
Critical infrastructure	Pearson Correlation	0.595	.860*	.843*	0.320	0.740	1.000**	1	-0.167	0.730	.954**	-0.258	0.382	-0.167	.919**	.884**	.881**	-0.167	0.415	-0.354	.814*	0.352	-0.167	.927**	0.196
	Sig. (2-tailed)	0.158	0.013	0.017	0.484	0.057	0.000		0.721	0.062	0.001	0.576	0.398	0.721	0.003	0.008	0.009	0.721	0.355	0.437	0.026	0.439	0.721	0.003	0.673
Cultural heritage	Pearson Correlation	-0.030	-0.143	-0.158	-0.240	-0.223	-0.167	-0.167	1	-0.335	0.047	0.645	0.000	-0.167	-0.165	-0.147	0.320	-0.167	-0.311	-0.354	-0.216	0.112	-0.167	-0.096	0.196
	Sig. (2-tailed)	0.949	0.759	0.735	0.604	0.631	0.721	0.721		0.463	0.920	0.117	1.000	0.721	0.723	0.753	0.484	0.721	0.497	0.437	0.641	0.812	0.721	0.838	0.673
Depopulation & displacement	Pearson Correlation	.920**	.958**	.955**	.848*	.982**	0.730	0.730	-0.335	1	.772*	-0.354	.837*	0.517	.908**	.947**	0.541	0.517	.833*	0.194	.937**	.784*	0.517	.896**	0.645
	Sig. (2-tailed)	0.003	0.001	0.001	0.016	0.000	0.062	0.062	0.463		0.042	0.437	0.019	0.234	0.005	0.001	0.210	0.234	0.020	0.677	0.002	0.037	0.234	0.006	0.117
Distress	Pearson Correlation	0.715	.914**	.899**	0.424	.809*	.954**	.954**	0.047	.772*	1	-0.055	0.567	-0.035	.925**	.930**	.939**	-0.035	0.473	-0.308	.828*	0.533	-0.035	.962**	0.372
	Sig. (2-tailed)	0.071	0.004	0.006	0.343	0.027	0.001	0.001	0.920	0.042		0.907	0.185	0.940	0.003	0.002	0.002	0.940	0.284	0.501	0.022	0.218	0.940	0.001	0.411
Early recovery	Pearson Correlation	-0.268	-0.269	-0.287	-0.372	-0.305	-0.258	-0.258	0.645	-0.354	-0.055	1	0.000	-0.258	-0.368	-0.228	0.062	-0.258	-0.482	0.091	-0.476	-0.106	-0.258	-0.258	-0.122
	Sig. (2-tailed)	0.561	0.560	0.533	0.411	0.506	0.576	0.576	0.117	0.437	0.907		1.000	0.576	0.417	0.623	0.895	0.576	0.273	0.846	0.280	0.820	0.576	0.577	0.795
Emergency response	Pearson Correlation	.912**	.781*	.802*	.917**	.877**	0.382	0.382	0.000	.837*	0.567	0.000	1	.764*	0.639	0.743	0.367	.764*	.832*	0.270	0.714	.944**	.764*	0.681	.810*
	Sig. (2-tailed)	0.004	0.038	0.030	0.004	0.010	0.398	0.398	1.000	0.019	0.185	1.000		0.046	0.123	0.056	0.418	0.046	0.020	0.558	0.071	0.001	0.046	0.092	0.027
Funding	Pearson Correlation	0.648	0.338	0.378	.881**	0.531	-0.167	-0.167	-0.167	0.517	-0.035	-0.258	.764*	1	0.196	0.265	-0.240	1.000**	.778*	0.471	0.390	.773*	1.000**	0.186	0.746
	Sig. (2-tailed)	0.116	0.458	0.403	0.009	0.220	0.721	0.721	0.721	0.234	0.940	0.576	0.046		0.673	0.565	0.604	0.000	0.040	0.286	0.388	0.042	0.000	0.689	0.054
Governance	Pearson Correlation	.848*	.976**	.964**	0.630	.913**	.919**	.919**	-0.165	.908**	.925**	-0.368	0.639	0.196	1	.978**	.803*	0.196	0.658	-0.146	.972**	0.651	0.196	.987**	0.501
	Sig. (2-tailed)	0.016	0.000	0.000	0.130	0.004	0.003	0.003	0.723	0.005	0.003	0.417	0.123	0.673		0.000	0.030	0.673	0.108	0.755	0.000	0.114	0.673	0.000	0.252
Injuries & casualties	Pearson Correlation	.876**	.994**	.979**	0.679	.948**	.884**	.884**	-0.147	.947**	.930**	-0.228	0.743	0.265	.978**	1	.778*	0.265	0.687	-0.021	.944**	0.704	0.265	.982**	0.563
	Sig. (2-tailed)	0.010	0.000	0.000	0.093	0.001	0.008	0.008	0.753	0.001	0.002	0.623	0.056	0.565	0.000		0.039	0.565	0.088	0.965	0.001	0.077	0.565	0.000	0.189
Intensity	Pearson Correlation	0.558	.757*	0.734	0.192	0.603	.881**	.881**	0.320	0.541	.939**	0.062	0.367	-0.240	.803*	.778*	1	-0.240	0.249	-0.510	0.678	0.392	-0.240	.845*	0.283
	Sig. (2-tailed)	0.193	0.049	0.061	0.680	0.151	0.009	0.009	0.484	0.210	0.002	0.895	0.418	0.604	0.030	0.039		0.604	0.590	0.243	0.094	0.385	0.604	0.017	0.538
Lifelines	Pearson Correlation	0.648	0.338	0.378	.881**	0.531	-0.167	-0.167	-0.167	0.517	-0.035	-0.258	.764*	1.000**	0.196	0.265	-0.240	1	.778*	0.471	0.390	.773*	1.000**	0.186	0.746
	Sig. (2-tailed)	0.116	0.458	0.403	0.009	0.220	0.721	0.721	0.721	0.234	0.940	0.576	0.046	0.000	0.673	0.565	0.604		0.040	0.286	0.388	0.042	0.000	0.689	0.054
Prevention	Pearson Correlation	.844*	0.750	.778*	.946**	.854*	0.415	0.415	-0.311	.833*	0.473	-0.482	.832*	.778*	0.658	0.687	0.249	.778*	1	0.110	.765*	.788*	.778*	0.655	.794*
	Sig. (2-tailed)	0.017	0.052	0.040	0.001	0.014	0.355	0.355	0.497	0.020	0.284	0.273	0.020	0.040	0.108	0.088	0.590	0.040		0.814	0.045	0.035	0.040	0.110	0.033
Preparedness	Pearson Correlation	0.084	-0.049	-0.083	0.283	0.060	-0.354	-0.354	-0.354	0.194	-0.308	0.091	0.270	0.471	-0.146	-0.021	-0.510	0.471	0.110	1	-0.073	0.152	0.471	-0.203	0.222
	Sig. (2-tailed)	0.858	0.918	0.859	0.538	0.898	0.437	0.437	0.437	0.677	0.501	0.846	0.558	0.286	0.755	0.965	0.243	0.286	0.814		0.876	0.745	0.286	0.662	0.632
Reconstruction	Pearson Correlation	.919**	.962**	.964**	.765*	.947**	.814*	.814*	-0.216	.937**	.828*	-0.476	0.714	0.390	.972**	.944**	0.678	0.390	.765*	-0.073	1	.760*	0.390	.948**	0.584
	Sig. (2-tailed)	0.003	0.001	0.000	0.045	0.001	0.026	0.026	0.641	0.002	0.022	0.280	0.071	0.388	0.000	0.001	0.094	0.388	0.045	0.876		0.047	0.388	0.001	0.169
Restoration	Pearson Correlation	.949**	.755*	.784*	.912**	.845*	0.352	0.352	0.112	.784*	0.533	-0.106	.944**	.773*	0.651	0.704	0.392	.773*	.788*	0.152	.760*	1	.773*	0.675	.804*
	Sig. (2-tailed)	0.001	0.050	0.037	0.004	0.017	0.439	0.439	0.812	0.037	0.218	0.820	0.001	0.042	0.114	0.077	0.385	0.042	0.035	0.745	0.047		0.042	0.096	0.029
Seismic information	Pearson Correlation	0.648	0.338	0.378	.881**	0.531	-0.167	-0.167	-0.167	0.517	-0.035	-0.258	.764*	1.000**	0.196	0.265	-0.240	1.000**	.778*	0.471	0.390	.773*	1	0.186	0.746
	Sig. (2-tailed)	0.116	0.458	0.403	0.009	0.220	0.721	0.721	0.721	0.234	0.940	0.576	0.046	0.000	0.673	0.565	0.604	0.000	0.040	0.286	0.388	0.042		0.689	0.054
Solidarity messages & actions	Pearson Correlation	.851*	.983**	.978**	0.625	.924**	.927**	.927**	-0.096	.896**	.962**	-0.258	0.681	0.186	.987**	.982**	.845*	0.186	0.655	-0.203	.948**	0.675	0.186	1	0.495
	Sig. (2-tailed)	0.015	0.000	0.000	0.134	0.003	0.003	0.003	0.838	0.006	0.001	0.577	0.092	0.689	0.000	0.000	0.017	0.689	0.110	0.662	0.001	0.096	0.689		0.258
Urban facilities	Pearson Correlation	.790*	0.609	0.595	.811*	0.668	0.196	0.196	0.196	0.645	0.372	-0.122	.810*	0.746	0.501	0.563	0.283	0.746	.794*	0.222	0.584	.804*	0.746	0.495	1
	Sig. (2-tailed)	0.034	0.147	0.158	0.027	0.101	0.673	0.673	0.673	0.117	0.411	0.795	0.027	0.054	0.252	0.189	0.538	0.054	0.033	0.632	0.169	0.029	0.054	0.258	

**. Correlation is significant at the 0.01 level (2-tailed)

*. Correlation is significant at the 0.05 level (2-tailed).

6. References

Antypas, D., Asahi, U., Camacho-Collados, J., Silva, V., Neves, L. and Barbieri, F. (2022) 'Twitter Topic Classification', in *29th International Conference on Computational Linguistics* 3386-3400.

Contreras, D., Shrestha, R., Dhungel, R., Shrestha, S., Paudel, K., Sharma, S., Adhikari, R.K., Whitworth, M., Giardina, G., Wilkinson, S., Jaquin, P., Macchiarulo, V., Burton, P., Pattison, J., Foroughnia, F., Dönmez, K., Chandran, K. and Milillo, P. (2023) 'Assessment of the Recovery Process using Sentiment and Topic analysis: the case study of the 2015 Gorkha Earthquake', in *SECED 2023*, Cambridge, United Kingdom, 14th September 2023, SECED, 43.

Contreras, D., Wilkinson, S., Balan, N. and James, P. (2021) 'Assessing post-disaster recovery using sentiment analysis: The case of L'Aquila, Italy', *Earthquake Spectra*, 38(1), 81-108, available: http://dx.doi.org/10.1177/87552930211036486.

Contreras, D., Wilkinson, S., Balan, N., Phengsuwan, T. and James, P. (2020) 'Assessing Post-Disaster Recovery Using Sentiment Analysis. The case of L'Aquila, Haiti and Chile', in *17th World Conference on Earthquake Engineering* Sendai, Japan.

Liu, B. (2015) 'Introduction ' in *Sentiment Analysis: Mining Opinions, Sentiments, and Emotions*, Cambridge: Cambridge University Press, 1-15.

MonkeyLearn (2022) 'Topic Analysis: The Ultimate Guide', *MonkeyLearn*, 2022(29th September 2022).

Ogie, R., Moore, A., Wickramasuriya, R., Amirghasemi, M., James, S. and Dilworth, T. (2022) 'Twitter data from the 2019–20 Australian bushfires reveals participatory and temporal variations in social media use for disaster recovery', *Scientific Reports*, 12(1), 16914, available: http://dx.doi.org/10.1038/s41598-022-21265-6.

Radianti, J., Hiltz, S.R. and Labaka, L. (2016) 'An Overview of Public Concerns During the Recovery Period after a Major Earthquake: Nepal Twitter Analysis', in *2016 49th Hawaii International Conference on System Sciences (HICSS)*, 5-8 Jan. 2016, 136-145, available: http://dx.doi.org/10.1109/HICSS.2016.25.

Roldós, I. (2021) 'What Is Data Analysis? Examples & Why It Matters', *MonkeyLearn*, 2021.

Rossetto, T., D'Ayala, D., Gori, F., Persio, R., Han, J., Novelli, V., Wilkinson, S.M., 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, available: http://dx.doi.org/10.1007/s10518-014-9588-y.

Shibuya, Y. (2020) *Social Media Communication Data for Recovery Detecting Socio-Economic Activities Following a Disaster*, 1st ed. 2020.. ed., Singapore : Springer Singapore : Imprint: Springer.

Wolf, T., Debut, L., Sanh, V., Chaumond, J., Delangue, C., Moi, A., Cistac, P., Rault, T., Louf, R., Funtowicz, M., Davison, J., Shleifer, S., Platen, P.v., Ma, C., Jernite, Y., Plu, J., Xu, C., Scao, T.L., Gugger, S., Drame, M., Lhoest, Q. and Rush, A. (2020) 'Transformers: State-of-the-Art Natural Language Processing', in *2020 Conference on Empirical Methods in Natural Language Processing: System Demonstrations*, Association for Computational Linguistics, 38–45, available: http://dx.doi.org/10.18653/v1/2020.emnlp-demos.6.

Xiao, Y., Huang, Q. and Wu, K. (2015) 'Understanding social media data for disaster management', *Natural Hazards*, 79(3), 1663-1679, available: http://dx.doi.org/10.1007/s11069-015-1918-0.