



Assessing emergency response and early recovery using sentiment analysis. the case of Zagreb, Croatia

Diana Contreras¹, Sean Wilkinson², Laure Fallou³, Matthieu Landès⁴,
Ivan Tomljenovich⁵, Rémy Bossu⁶, Nipun Balan⁷, Philip James⁸

¹ *Research Associate, Newcastle University, diana.contreras-mojica@newcastle.ac.uk*

² *Reader, Newcastle University, sean.wilkinson@newcastle.ac.uk*

³ *Sociologist, Euro-Mediterranean seismological centre (EMSC), fallou@emsc-csem.org*

⁴ *Seismologist, Euro-Mediterranean seismological centre (EMSC), matthieu.landes@emsc-csem.org*

⁵ *Head of Remote Sensing and GIS laboratory, Oikon Ltd.-Institut of Applied Ecology, itomljenovic@oikon.hr*

⁶ *Secretary General, Euro-Mediterranean seismological centre (EMSC), remy.bossu@cea.fr*

⁷ *Postgraduate research student, Newcastle University, n.b.thekkummal1@newcastle.ac.uk*

⁸ *Professor, Newcastle University, philip.james@newcastle.ac.uk*

Abstract

The 2020 Zagreb earthquake occurred on Sunday 22 March 2020. This earthquake was the first that happened during the lockdown imposed by governments to stop spreading the COVID-19. This fact makes the event interesting as a multi-hazard phenomenon. The lockdown made it not possible to deploy an earthquake reconnaissance mission. Therefore, it was necessary to undertake a remote mission supported by the monitoring and analysing social media (SM) platforms, such as Twitter and Instagram. This paper presents details and analysis of this monitoring and how it may help understand the impacts of an earthquake. In our work, we first identified the hashtags related to the event. Through the LastQuake app, we obtained the intensity reports from affected people and comments and pictures useful for damage assessment. The team obtained 59,246 tweets posted between the 20th March and the 30th April 2020 and 31,911 comments from LastQuake app users written on the day of the earthquake. Images from posts and comments were used for remote assessment of damage in buildings. Sentiment analysis (SA) was applied to tweets and comments related to the event to assess emergency management during the relief phase after the earthquake. Our work shows that only a limited number of pictures collected through social media were suitable for damage assessment of individual buildings. However, they were still useful as a proxy estimation of damages in some areas of Zagreb and surroundings. We also found SA supported by machine learning a valuable method to assess and identify critical aspects of the emergency and early recovery post-disaster phases. Applying SA we identified the most affected areas, the damages in the non-structural elements in hospitals, the support of collaborative networks for the evacuation of patients and the role of Ministers in the early recovery.

Key words: Post-disaster recovery, earthquakes, Twitter, sentiment analysis, social media, citizen science, LastQuake app, COVID-19

1 Introduction

The 2020 Zagreb earthquake occurred on Sunday 22 March 2020. One fatality and 26 injuries were directly caused by the earthquake, while two people died during clean-up operations afterwards. This earthquake was the first significant earthquake event to occur during a government lockdown to stop the spread of the COVID-19. This fact makes this event interesting as a multi-hazard phenomenon. It allows studying events such as people not knowing if they should stay at home, evacuate, and be exposed to the virus and the low temperatures outside. The lockdown made it impossible to deploy an earthquake reconnaissance mission by the Earthquake Engineering Field Investigation Team (EEFIT). Therefore, it was necessary to undertake a remote mission supported by monitoring and analysis of social media (SM) platforms, such as Twitter and Instagram. Shortly after the event, we obtained georeferenced image and text data provided by the Euro-Mediterranean Seismological Centre (EMSC) captured through the LastQuake app developed by them. LastQuake app is a crowdsourcing-based earthquake information app that allows eyewitnesses to share information about earthquakes they felt, combined with seismic data [1-3].

The main aim of using SM is extracting useful information from images and text data. The object is to support data-driven decisions in earthquake reconnaissance, post-disaster needs assessment (PDNA), improve preparedness, and assess the emergency response and the post-disaster recovery processes. To analyse text data, we used sentiment analysis (SA), a natural language processing (NLP) technique applied to determine data polarity (positive, negative, neutral, or unrelated), as well as feelings and emotions, urgency, and intentions (interested v. not interested) [6]. In this paper, we have applied sentiment analysis (SA) [7], also known as opinion mining (OM) [8, 9], as a method to extract information from the text data contained in the post on Instagram, tweets and comments from LastQuakeapp users. Sentiment Analysis has been mainly used before for customer reviews of products and places (restaurants and hotels), stock markets [10, 11], new articles [12], products, brands, services online feedback [6] or political debates [13].

2 Methodology

We identified the hashtags related to the event monitoring Twitter the day of the earthquake. Through the LastQuake app, we obtained the intensity reports from affected people and comments and pictures useful for damage assessment. The team obtained 59,246 tweets with the hashtags: #croatia, #croatiaearthquake, #CroatiaQuake, #earthquakecroatia, #earthquakeinzagreb, #hrvatskapotres, #LoveZagreb, #potres, #potreszagreb, #potresZg, #PrayforZagreb, #Zagreb, #zagrebearthquake, #zagreb-potres, #zagrebquake, #zagrebstrong, #zagrebstaystrong posted between the 20th March and 30th April 2020. The 33.4 % (19840) were original tweets, 66.5 % (39,406)

were retweets, 1,369 were replies and 15,415 (26 %) included links and pictures. The activity on Twitter is plotted in Figure 1.

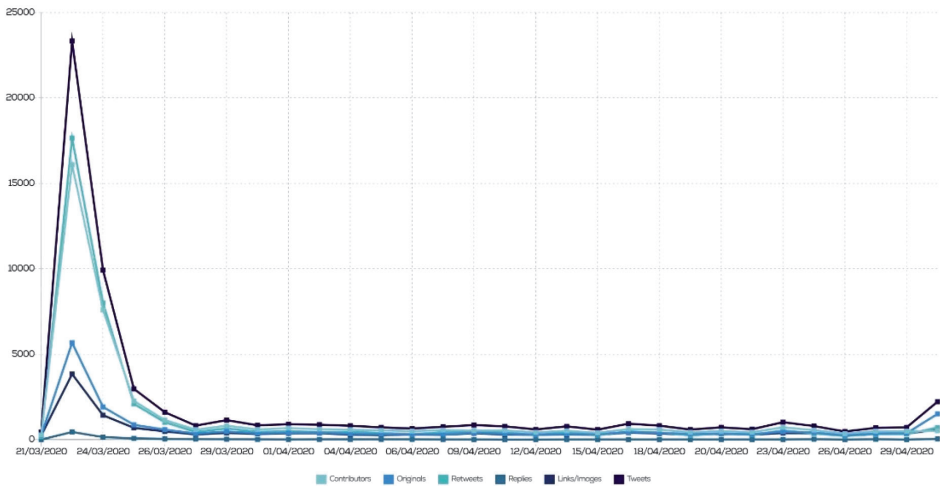


Figure 1. Twitter activity between the 20 March and 30 April 2020, 2020 related to the earthquake in Zagreb. Source: TweetBinder

We took a sample of 19 tweets with pictures showing damages to buildings and debris on the streets similar to the pictures uploaded by LastQuake app users. These pictures needed to be georeferenced because image data obtained through Twitter and Instagram are no longer geotagged due to privacy issues. It is important to note that we are not interested in the users' location, but rather in the location of the damage depicted in the photograph. To georeference the pictures obtained from Twitter, we first located images of landmarks on the Cathedral or the Croatian Society of Fine Arts images. Second, we looked at the streets' names (if this was displayed in the photo) or contacted the photographers to know the pictures' coordinates. Third, we validated the coordinates identified using the street view function from *google maps* to produce a map with the spatial distribution of damage in Zagreb and surroundings. We have assumed that debris on the streets is an indication of damages in a building nearby. Images from tweets and posts on Instagram were used for remote assessment of damage in buildings [4]. A sample of tweets with pictures are plotted on Figure 2.

LastQuake app obtained 31,911 intensity reports with comments from which it has been possible so far to translate 31,403 (98 %). The citizens included in their comments, 361 pictures. After data processing, 314 (87 %) pictures were selected for damage assessment. 45 LastQuake app users (0.1 %) included pictures with comments on their intensity reports; some examples of comments and pictures are plotted in Figure 3. After observing the location of the damage, we process the text data included in the posts on Twitter and Instagram and comments on LastQuake app to perform the SA analysis identifying keywords, their polarity, and the topics addressed on the text data.

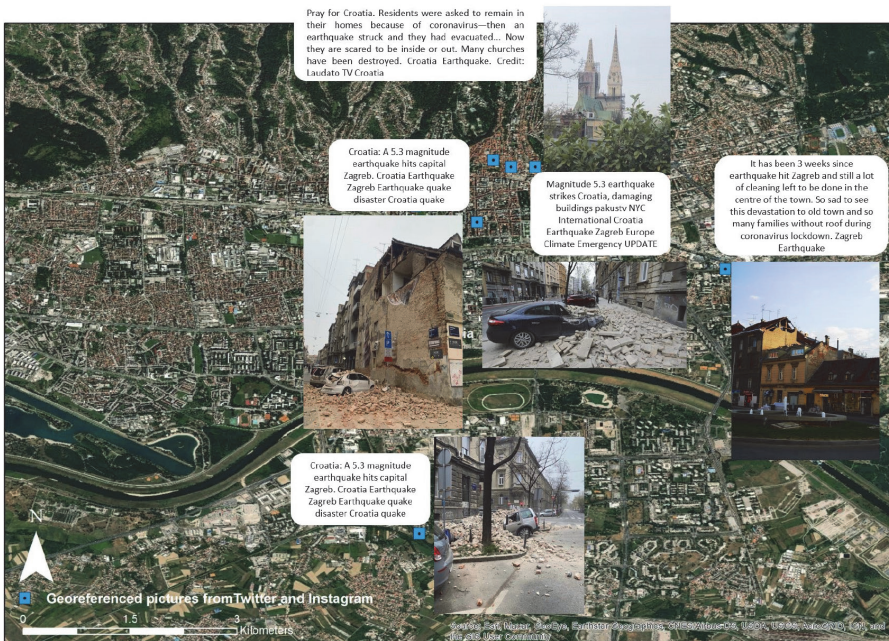


Figure 2. Sample of georeferenced pictures uploaded by Twitter users the day of the 22nd March 2020 earthquake



Figure 3. Examples of pictures and comments done by LastQuake app users after the Earthquake in Zagreb

To extract information from the text data contained in the post on Instagram, tweets and comments from LastQuakeapp users, we use a bag of words (BOW)[9] or word clouds [10]. This method allows us to identify the keywords in the text. We used the AI-powered free word cloud generator developed by *MonkeyLearn* [11], which visualises term frequency and relevance.

The categories of polarity are namely, positive, negative, neutral and unrelated. In case of an emergency due to an earthquake, most of the text data will have a negative polarity because it will contain words related to damage, fear, and anxiety. However, there will also be data that include words related to the event, such as magnitude, intensity, or the epicentre location, that will be considered neutral. Other data will contain solidarity messages, announcements of support with humanitarian aid, or descriptions of help. These are considered to be positive as they demonstrate instances of success [6].

Topic analysis, also called topic detection, topic modelling or topic extraction, is another NLP technique to automatically extract meaning from text by identifying recurrent themes or topics. This technique uses machine learning to organise and understand large text-datasets [7], such as those coming from Twitter, Instagram and LasQuake app, by assigning 'tags' or categories to each text's topic or theme. These categories can then be used to understand what impacts were most experienced and potentially, what resources should be deployed to help in the response and recovery, and where best to allocate those resources for maximum effect. Based on the study of dataset related to the recent earthquakes in Albania (2019) [8] Croatia [4] and Turkey [6] we chose 12 categories: building damage, tsunami effects, geotechnical effects, lifelines affected, injuries and casualties, emergency response, solidarity messages, preparedness, seismic information, intensity, early recovery and unrelated. In Zagreb's case, we ignored tsunami effects and unrelated as they are not relevant to this event.

Our analysis employed a hybrid approach to classification, where both automatic (using *MonkeyLearn*) and rule-based methods were used to make the classifications. The authors defined the rulesets used in this classification based on their experience in disaster management and post-disaster recovery. The methodology is presented in Figure 4.

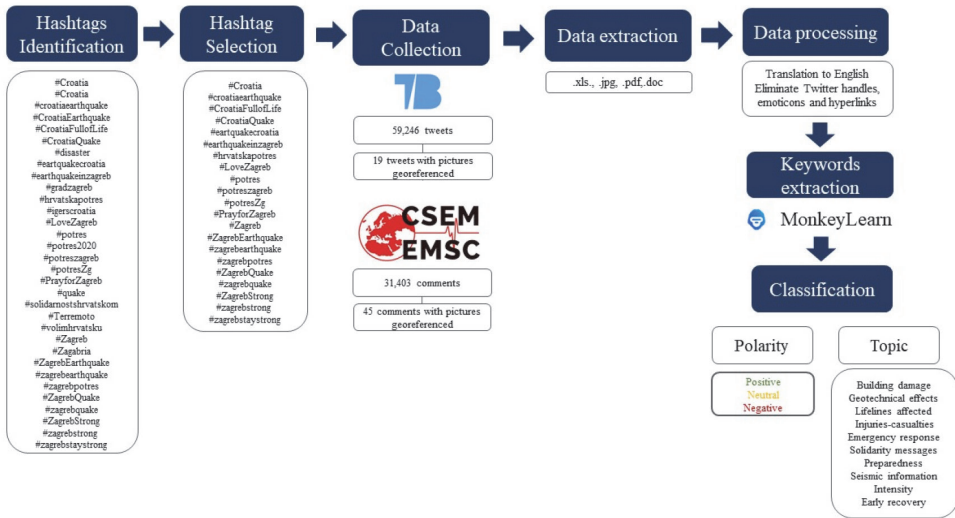


Figure 4. Methodology

4 Results

Pictures uploaded by Twitter and Instagram users are mainly taken from the outside of buildings, except for some interiors of hospitals where damage in non-structural elements can be seen and images of evacuations actions. In contrast, pictures uploaded by LastQuake app users include coordinates and come from inside and outside the buildings. LastQuake app dataset contains pictures of individual damaged buildings from different angles, which are particularly useful for damage assessment. Also, LastQuake app users identified geotechnical failures outside of the city of Zagreb, which were not highlighted in any post, neither on Twitter nor on Instagram.

The most frequent and relevant terms identified in the text data makes reference to the 'University Hospital Centre Zagreb', also known as 'Rebro', followed by 'earthquake', 'horrible experience', 'Zagreb', 'Croatia' and 'maternity hospital petrova'. The most frequent and relevant terms are depicted in Figure 5. Polarity classification listed in Table 1 shows that while Twitter and Instagram users who upload pictures of damages include posts with all polarities. Tweets or posts on Instagram that include pictures with damages do not necessarily include a text with a negative polarity; they sometimes include a solidarity message. In contrast, comments from LastQuake app users who also upload pictures for this case study have mainly a negative and neutral polarity, never positive. In the case of text with neutral polarity, or when text data does not include enough information to identify a topic, it was necessary to look at the pictures to identify their polarity. The spatial distribution of the polarity of the text data is mapped on Figure 6. Topic classification indicates that Twitter and Instagram users address mainly four topics: building damage, emergency response, seismic information and solidarity message.

es. In contrast, LastQuake app users focused on building damages and seismic intensity. The list of topics addressed by Twitter, Instagram and LastQuake app users is listed on Table 2 and mapped on Figure 7.

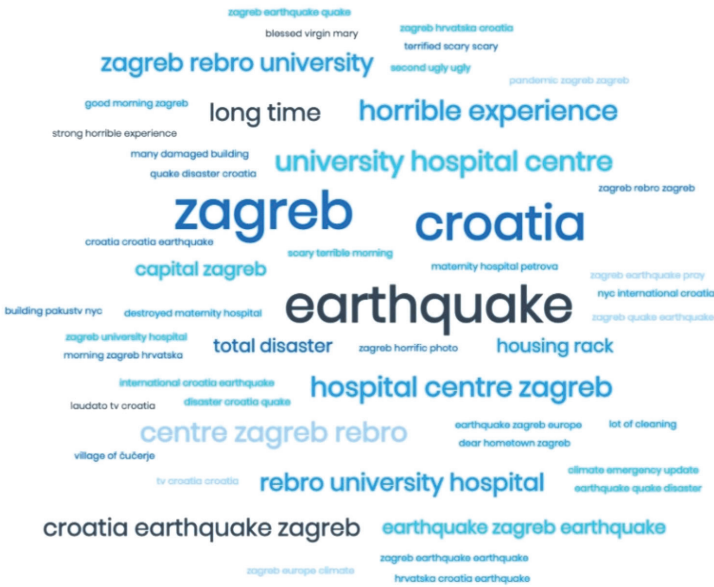


Figure 5. Polarity tweets and LastQuake app comments.

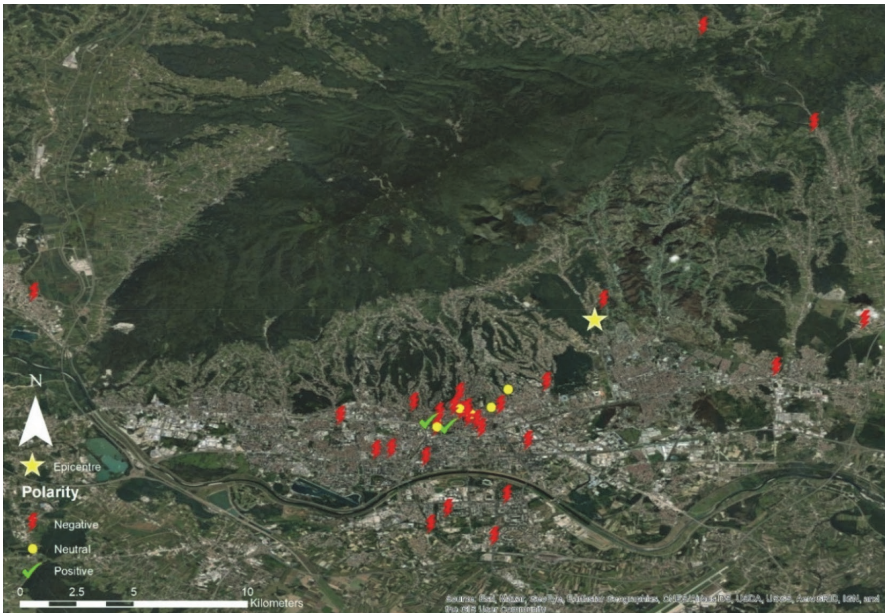


Figure 6. Polarity tweets and LastQuake app comments.

Table 1. Polarity post on Instagram, tweets and comments from LastQuakeapp users.

Polarity	Twitter & Instagram	LastQuake app	Total	Percentage
Negative	12	33	45	70
Neutral	5	12	17	27
Positive	2	0	2	3
Total	19	45	64	100

Table 2. Topics addressed on posts on Instagram, tweets and comments from LastQuakeapp users

Topics	Twitter & Instagram	LastQuake app	Total	Percentage
Building damage	5	23	28	44
Emergency response	5	0	5	8
Seismic information	6	0	6	9
Solidarity	3	0	3	5
Intensity	0	22	22	34
Total	19	45	64	100

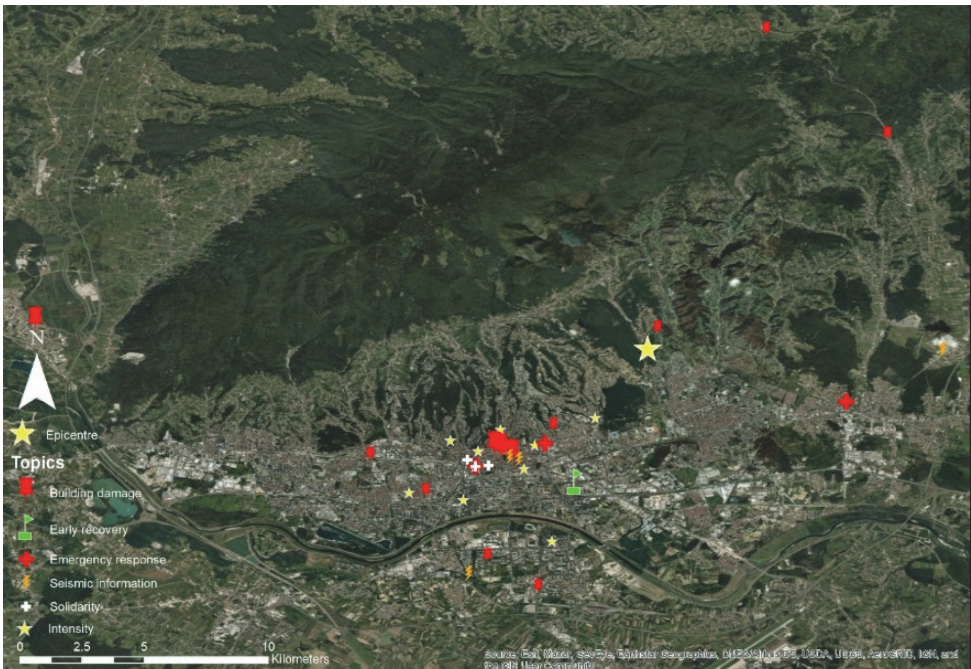


Figure 7. Topics addressed in tweets and LastQuake app comments

5 Conclusions

The users of the different social media platforms have different profiles. Therefore, text data, polarities, and topics addressed concerning the same event can vary according to the platform. However, the subject of images were similar on the three platforms (except for the evacuation images uploaded only by Twitter users and geotechnical failures identified only by LastQuake app users).

The focus on building damages, geotechnical failures and seismic intensity of pictures and comments from LastQuake app users can be explained by the fact that this app was specifically developed to share information about felt earthquakes and seismic data. In contrast, Twitter and Instagram were created to share general information. It is also possible to extract the image and text data after earthquakes from these platforms. Still, it would be more useful for situational awareness, than for damage assessment, which is more the case of image and text data shared by LastQuake app users. The general purpose of platforms such as Twitter and Instagram can also explain the variety of polarities and topics addressed in their posts. LastQuake app users are recurrent in seismic zones. Therefore it would be possible to train them to post pictures useful for emergency management and damage assessment and later transfer this training to users of other platforms such as Twitter, Instagram and Facebook.

The range of intensities reported by LastQuake app users, who upload pictures, include IV (light), V (moderate), VI (strong), VII (very strong) and VIII (severe) in the Modified Mercalli Intensity(MMI) scales. As far as the intensity reported is higher, the number of pictures and comments uploaded by users increases until intensity VII, then it plummeted. This fact allows us to conclude that higher intensities encourage users to upload pictures and to make comments. LastQuake users reporting intensities from IV to VII comments about building damage and intensity, but users who report intensity of VIII focus their comments only on intensity. The percentage of confidence with which *MonkeyLearn* classifies the text data rise also when the intensity reported is higher because users tend to include fewer words, even using only one when the intensity reported is VIII (severe), making the classification more precise. In conclusion, we believe that SM harvested from a specialised app such as LastQuake app has the potential to quickly inform disaster manager and responder about the impact of earthquakes and make the corresponding assessment. At the same time, ordinary SM provides general information as to what has occurred on the ground.

Acknowledgements

The authors would like to express their thanks to the Engineering and Physical Sciences Research Council (EPSRC) [Grant No.: EP/P025641/1] and Mr Javier Hervás Ciudad to support the process of databases.

References

- [1] Bossu, R., et al. (2018): LastQuake: From rapid information to global seismic risk reduction. *International Journal of Disaster Risk Reduction*, 28, 32-42, doi: <https://doi.org/10.1016/j.ijdr.2018.02.024>
- [2] Bossu, R., et al. (2020): Rapid Public Information and Situational Awareness After the November 26, 2019, Albania Earthquake: Lessons Learned From the LastQuake System. *Frontiers in Earth Science*, 8 (235), 1-15, doi: <https://doi.org/10.3389/feart.2020.00235>
- [3] Fallou, L., et al. (2020): Citizen Seismology Without Seismologists? Lessons Learned From Mayotte Leading to Improved Collaboration. *Frontiers in Communication*, 5 (49), 1-17, doi: <https://doi.org/10.3389/fcomm.2020.00049>
- [4] So, E., et al. (2020): The Zagreb Earthquake of 22 March 2020. EEFIT report, London, UK.
- [5] Contreras, D., et al. (2020): Assessing Post-disaster Recovery Using Sentiment Analysis. The case of L'Aquila, Haiti, Chile and Canterbury, 17th World Conference on Earthquake Engineering (17WCEE), Sendai, Japan, 12 pages.
- [6] Aktas, Y.D., et al. (2021). The Aegean earthquake and tsunami of 30 October 2020 (on revision). EEFIT report, London, UK.
- [7] MonkeyLearn: Topic Analysis. <https://monkeylearn.com/topic-analysis/>, (accessed 8th March 2021).
- [8] Andonov, A., et al. (2020): The Mw6.4 Albania Earthquake on the 26th November 2019. EEFIT report, London, UK.
- [9] Zucco, C., et al. (2020): Sentiment analysis for mining texts and social networks data: Methods and tools. *WIRES Data Mining and Knowledge Discovery*, 10(1), 1-32, doi: <https://doi.org/10.1002/widm.1333>
- [10] Roldós, I.: Best Word Visualization Tools to Create Your Own Word Cloud, in MonkeyLearn Blog. <https://monkeylearn.com/blog/word-visualization/>, (accessed 8th March 2021).
- [11] Wolff, R., Visualize Sentiments in A Word Cloud - Free Tools & Tutorial. <https://monkeylearn.com/blog/word-cloud-sentiment-analysis/>, (accessed 8th March 2021).