The Integration of Social Media Data in Emergency Management: an Innovative Decision Support System

Valentina Di Talia, Giacomo Antonioni*

University of Bologna, CIRI Fonti Rinnovabili, Ambiente, Mare ed Energia, Via Sant'Alberto 163, 48123, Ravenna
giacomo.antonioni3@unibo.it

The upsurge of social media platforms has opened to the prospect of integrating the information provided by citizens through these channels into the traditional emergency management process. This paper presents the Civil Protection Emergency System model designed for the Italo-Croatian decision support system developed in the Interreg project E-CITIJENS. Seismic, flood and forest fire are the risk typologies addressed. The model specifies the key steps that allow the system, a semantically enriched web-enabled platform, to identify and analyse significant social media posts that can provide Civil Protection authorities with additional real-time data regarding potential or ongoing emergencies in a designated geographical area. The approach chosen is to retrieve from social media the posts containing specific terms used by citizens during emergencies (i.e. an initial project terminology was developed) and to classify them according to their relative importance compared to the other posts selected, identifying those to be evaluated first by the Civil Protection staff. This is achieved by calculating the total score of each post as the sum of the scores attributed to the initial terminology keywords therein contained (i.e. three severity scales were defined to rank the terms according to their potential hazard level). This novel Civil Protection Emergency System model has been applied to a set of simulated emergency events with the aim of testing the algorithm and to verify the effectiveness of the platform in order to assess if it could provide helpful additional information to Civil Protection, improving its overall emergency coping capacity.

1. Introduction

Natural and man-made disasters are becoming more and more extreme and complex in Europe as all over the globe (Pourebrahim et al., 2019), making a priority to establish appropriate feedback mechanisms to minimise as much as possible their frequency of occurrence and their impacts. It is well known that emergency management is a complex system that allows the correct allocation of resources and more rapid and effective restoration of pre-intervention conditions. This continuous process requires a constant adaptation to changing circumstances and social requirements to prevent risks and reduce vulnerability in the risk-prone areas, improving their resilience. According to the Sendai Framework for Disaster Risk Reduction 2015-2030 (UNDRR), it is indispensable to set the conditions to boost the utilisation of social media as an additional source of data to support the emergency management process at all administrative levels and in all its phases.

Social media platforms' potential to help the authorities better detect hazardous events, estimate the damages and locate aid seekers has gained rising attention in the last decades (Bhuvana and Aram, 2019). Since the beginning of the millennium, in fact, the upsurge of social media platforms such as Facebook, Twitter and Instagram has wholly changed communication during emergencies (Imran et al., 2015). While top-down and citizen-to-citizen communication is already well-established, much needs to be done to enable Civil protection authorities and, more in general, emergency management agencies to effectively retrieve and analyse information on evolving crises posted by citizens on social media (Zhang et al., 2019). This is related mainly to the multiple issues that must be addressed during this process. Among them, the most significant are semantic interoperability (Sheth, 1999), data geolocation (Burton et al., 2012) and data quality verification (Lazreg et al., 2019). Concerning the latter, it is particularly significant to mention the approach proposed by Ludwig et al. (2015). They define five data quality attributes that could be extracted from existing metadata to be specified by the user, providing a guide to address the issue of filtering relevant information from social media messages.

Implementing innovative Decision Support Systems is the key to allowing the authorities to easily retrieve and process the information provided by citizens through social media to improve their overall capacity to cope with emergencies (Hellmund et al., 2019). In this work, the cross-border Civil Protection Emergency Decision Support System (EDSS) designed in the context of the Interreg Project E-CITIJENS will be presented, concentrating primarily on the Civil Protection Emergency System (CPES) model that was at the basis of its development. This system has been devised specifically for Italy and Croatia and addresses forest fire, seismic and flood risk. Its main novelty is the inclusion of real-time situational awareness data from citizens, seen as "active sensors", in the emergency management process to improve the overall resilience of the territories at stake.

The main aim of this study is to verify the overall effectiveness of the CPES model implemented in the EDSS platform using social media messages produced during the platform deployment tests.

2. Methodology

The idea beyond the new EDSS developed during the project E-CITIJENS is to allow the retrieval and processing of relevant messages posted on social media by citizens to have additional information to complement institutional data before and during crises. The system is a semantically enriched web-enabled platform that processes and presents to Civil Protection data from these different sources to allow a prompt detection and assessment of hazardous events.

Recognising that citizens commonly share information regarding crises and dangerous situations, this platform has to search in social media for messages containing early warning and situational awareness information, and process them to assess the potential hazard level related to their content, addressing a selected geographical area. The goal is the individuation of the social media posts that need to be presented with higher priority to the Civil Protection staff through the EDSS.

The CPES model defines the steps required for this process. In particular, it specifies the approaches to retrieve relevant messages regarding the three risk typologies at interest (which will be analysed simultaneously by the system) from social media and to rank these according to their potential hazard level.

A common terminology and a project hashtag have been defined to address the message retrieval issue. Concerning common terminology, the objective has been to individuate a set of terms frequently used by citizens during crises and ask the system to search for them in social media. Evidently, the platform must activate APIrest connectors or directly search for information "capture" solutions on selected social media platforms such as Facebook and Twitter.

The posts containing these terms are the ones to be processed by the EDSS. The definition of this terminology required the individuation of the main domains (and sub-domains) to address to ensure semantic interoperability and, therefore, to allow an effective message processing.After a survey within the partnership, and considering the necessity of simplicity, the domains and subdomains individuated are risks (forest fires, floods, seismic, and common), operations (early warning, situational awareness, relief operations, common), and languages (Italian, Croatian, English). Evidently, these terms have to contemplate all the risk typologies at stake and need to be translated into Italian, Croatian, and English. Moreover, the emergency management process phases have to be taken into account. It is also to note that a term can refer to more than one risk or phase; therefore, a sub-domain 'common' has been introduced in both categories. After this preliminary analysis, an initial set of key terms has been identified through a survey within the partnership; evidently, this terminology should be continuously updated and enhanced based on the experience gained using the platform.

'#EDSS' has been chosen as the project hashtag in order to allow the citizens to "speak" directly with the system and to facilitate and strengthen the retrieval of potentially useful social media messages (i.e., the messages containing '#EDSS' have to be given higher priority). It is noteworthy that the effectiveness of this hashtag is strictly related to the Awareness Campaign planned within the project.

Three severity scales, one per risk typology, will be calculated (for simplicity and considering the application, they are identical) to allow the ranking of the messages selected by the system. The idea is to attribute a score to each common terminology keyword to quantify its corresponding hazard level (the higher the score, the higher the potential hazard) per risk typology. The defined scores and their description are reported in Table 1.

The initial score attribution derives from expert elicitation. It is to note that score attribution can strongly influence the results obtainable with the EDSS, therefore as for the common terminology, it needsto be refined continuously in order to improve the overall performance of the platform. To aid this process, an Artificial Intelligence (AI) component based on Natural Language Solutions could be implemented in the future.

Table 1: Scores to be assigned to each key term per risk typology

Score	Description
0	No relevance; the keyword has no connection to the risk typology addressed.
0.25	Low relevance; the keyword cannot be regarded as an indicator of a potential hazard concerning the risk typology addressed though it can be crucial to contextualise the post
0.50	Medium relevance; the keyword can indicate a potential hazard regarding the risk typology addressed.
0.75	High relevance; the keyword explicitly indicates a potential hazard regarding the risk typology addressed.
1	The term unquestionably implies a hazard concerning the risk typology addressed.

The system needs to process the messages selected in the retrieval phase to individuate the primary risk they address, quantify their potential hazard level and identify the ones with higher priority. For this last step, it is first necessary to calculate the total score S_i per risk typology of each post as defined by Eq(1)

$$S_i = \sum_{j=1}^{n} s_{i,j} \tag{1}$$

In the equation, i embodies the risk typology, n is the number of key terms in the post, and $s_{i,j}$ is the keyword score for key term j and risk typology i. This calculation allows the straightforward identification of the principal risk typology addressed by the post (it matches the maximum value obtained) and of the posts related to a higher hazard level (S_i indicates the relative importance of each post compared to the other posts selected). At this stage, the system has to give higher priority at an equal total score to the posts that include the project hashtag '#EDSS' and, secondarily, any multimedia. Concerning the latter, evidently, photos or videos regarding the crisis could supply Civil Protection with valuable additional information.

To conclude, an overview of the main steps that allow the system to retrieve and process social media messages relevant to civil protection is presented. After the pre-selection of the geographical area of interest by the user, the EDSS searches in social media for the posts containing one or more common terminology words. According to the set of scores attributed to the keywords per risk typology (which represent the potential hazard level related), the system calculates for each post its total score S_i and ranks the messages according to the values obtained. The messages with higher scores (which correspond to a higher hazard level) are then presented geolocated to the user prioritising at the equal total score the ones addressed directly to the EDSS or containing any multimedia. The user can then validate this information with institutional data and activate eventually the necessary procedures to face the emergency.

3. Results

The model has been applied to verify its comprehensive ability to effectively select relevant social media messages and correctly classify them according to the potential hazard level related to their content. The data used for this application consists of 296 messages in the Italian language collected in the pilot tests deployed during the project (November 2021) by Veneto Region, Civil Protection, Security and Local Police Department. These pilot tests were deployed online on a websitein order not to create false alarms in real social networks . Three simulated events of different severity (low, medium, high) for each risk typology were defined (altogether 9 scenarios). The idea has been to present to the participants the detailed description of one of these scenarios, asking them to formulate a realisticmessage as they lived that particular situation and used social media.

The following sections will report the presentation and discussion of the results obtained by applying the proposed model using the just described data.

Considering that the analysed messages were written answering to specific simulated events, the number of posts containing a specific number of keywords (presented in Figure 1) can help assess, in the first instance, the completeness of the common terminology defined and its actual capacity to select relevant messages. It can be noticed from the graph that most of the messages (≈80%) enclose at least one of the key terms identified in the common terminology. Moreover, the majority of the posts contain one to four keywords, which can be assumed a satisfactory result considering the application. To properly read these values and to understand why 20% of the messages do not contain any keyword, it is essential to acknowledge that, since the EDSS will be asked to process a notable amount of variegated data in real situations, all the messages were included (even fake ones) in this first part of the analysis. It is moreover to underline that the texts produced by the pilot participants contained also empty fields, misspelt words and expressions in other languages that the system could not detect. It can be concluded that the initial terminology defined comprises a robust foundation for

retrieving social media messages relevant to Civil Protection. As already discussed, it should be continuously updated to improve the overall system's performance.
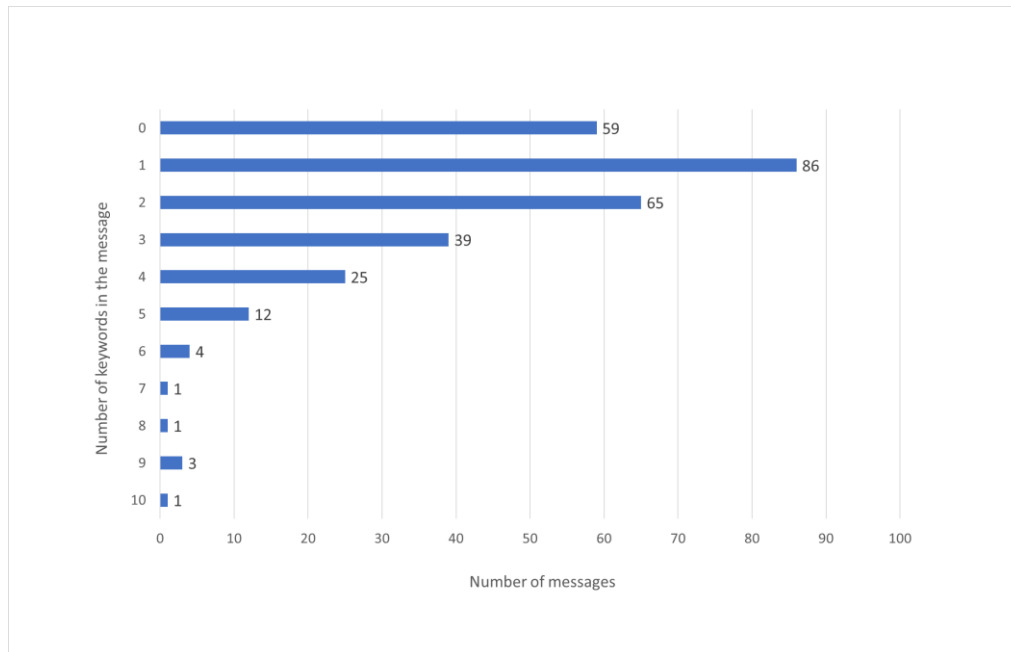

Figure 1: Analysed messages containing a specific number of keywords.

The capability of the CPES model to individuate the primary risk typology related to a message can be assessed by calculating the average values of the total scores obtained for all the messages written in response to the scenarios related to the same risk typology regardless of severity. The values obtained (with their relative uncertainties) are presented in Table 2 (the abbreviations 'FF', 'F' and 'S'correspond to forest fire, flood, and seismic risks, respectively).

Table 2: Average total scores per risk typology obtained for the messages written in response to the three main scenario groups.

	$\overline{S_{FF}}$	$\overline{S_F}$	$\overline{S_S}$
FF scenarios	2.09 ± 1.50	0.73± 0.74	0.70± 0.74
F scenarios	0.88 ± 1.00	1.45± 0.96	0.76± 0.82
S scenarios	0.73± 1.09	0.69± 0.93	2.11± 1.20

From the values presented, it is evident that, on average and despite uncertainties, the model can pinpoint the risk typology related to a message in the case study analysed. In fact, the higher values obtained per scenario group (on the diagonal of the table) match the correspondent risk typology, and the values related to the other risks are significantly lower. Even if this application is not required for most emergency situations, this test allows assessing the general correctness of the developed model.

The last issue to verify concerns the ability of the model to detect the level of severity associated with a message. The average values of the scores obtained per each scenario proposed were calculated per risk typology for this intent. The results are presented in the graph reported in Figure 2. It can be seen from the graph that the model can estimate, on average, the level of severity of a post but still with some limitations. It is clear that the best performance is obtainable for seismic risk, even though the differences between medium and high severity scenarios could not be detected by the EDSS. About forest fire risk, this difficulty regards mainly low and medium-high severity scenarios. Finally, the average values obtained for flood risk are very similar for all severity levels making nearly impossible the desired differentiation. It is evident that common terminology (and related score attribution) ought to be further implemented to improve the EDSS performance, primarily for flood riskbut also for forest fire risk, as it can be seen from the uncertainties bars. Despite the required refinements, it can be said that the model can assess the level of hazard related to a post and, consequently, can perform an appropriate ranking of the post analysed to identify the most relevant ones.
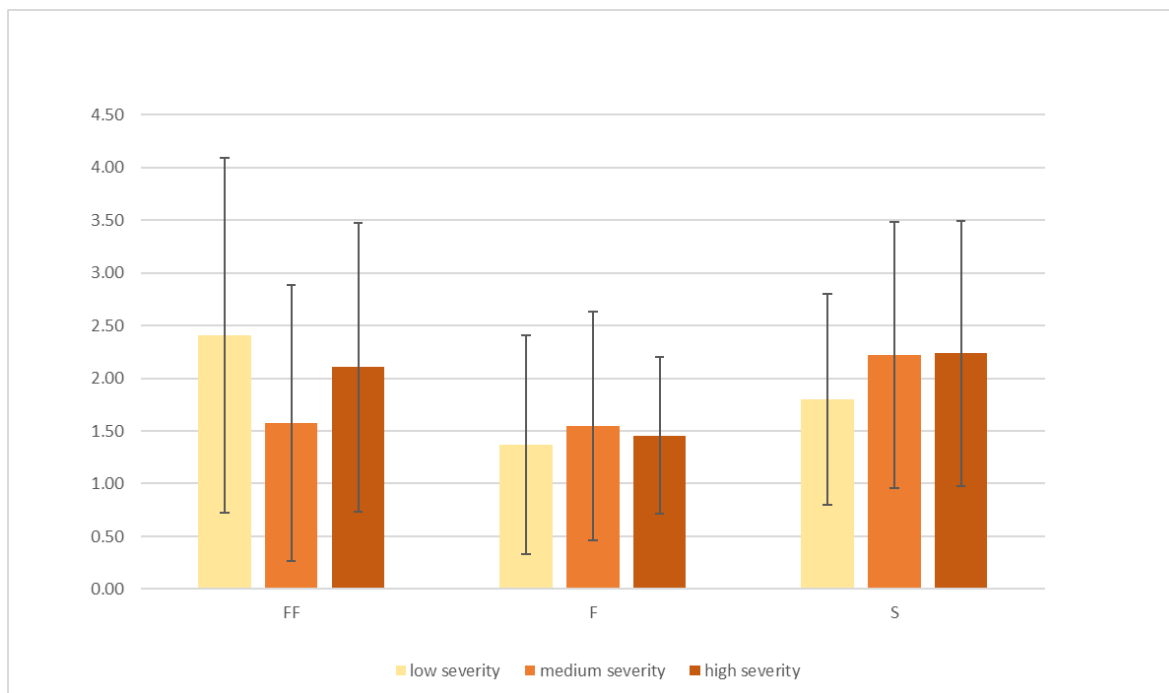
Figure 2: Average score values obtained for the risk typology and severity level addressed by each scenario.

4. Conclusions

Social media platforms have the potential to provide real-time valuable and actionable data to Civil Protection authorities to better cope with emergencies in all the stages of the crisis, improving the overall emergency management process. In this paper, the new EDSS developed within the Interreg Project E-CITIJENS has been presented, focusing primarily on the CPES model at the base of its development.

The EDSS aims to provide support to decision-makers in the analysis of the emergency situation, integrating institutional data with the information provided by citizens through social media. While institutional data are continuously collected and processed by emergency management agencies and are readily available, the retrieval and processing of social media data are not straightforward. The CPES model defines the steps required to allow the system to obtain in output a selection of social media messages with higher priority, therefore containing information relevant to Civil Protection that can enhance the overall capacity of decision-makers to cope with emergencies.

The model has been tested on a sample of 296 messages collected during the pilot tests deployed within the project to verify its overall correctness. From the results obtained, it is clear that the model can effectively select the post containing potentially useful information and is able to identify, on average, the primary risk typology addressed by a post as its level of severity. The selection issue is strictly correlated to the completeness of the common terminology. Considering that at least one keyword was found in 237 messages out of 296 (in the analysis were included all the simulations, including the ones containing no words or unrecognisable terms), the terminology defined can be considered a solid basis to allow a proper message retrieval.

Regarding the message processing, the results obtained can be assumed as good. The model is in fact capable of identifying the primary risk typology addressed in a message despite uncertainties . Calculating the average values of all the scores obtained per risk typology for each of the three main scenarios groups (forest fire, flood, seismic), the values related to the risk typology effectively addressed by the message are at least 50% higher than the ones concerning the other risks. Moreover, the CPES approach allows the identification, on average, of the level of severity correlated to a post. This aspect is fundamental to perform a correct raking of the selected message (i.e. to identify the messages that have to be presented with higher priority to the Civil Protection staff for evaluation). The average values of the scores calculated for each risk scenario per risk typology show that even though the results give an indication of the actual potential hazard related to the posts, some refinements are required. In particular, the model features (common terminology and related scores) associated with flood and forest fire risks need to be improved to obtain a better performance.

To conclude, the model presented in this paper is a fundamental tool to allow decision-makers to consider also real-time data collected through social media provided by citizens during crises. This additional information can help detect and assess the damage during an emergency, enhancing the overall emergency management process and the resilience of the affected communities. Considering further improvements and applications, is it to note that the CPES initial skeleton can be expanded regarding other risk typologies and languages. In particular, the risk of major accidents (as defined by the Seveso III Directive) could be included in the system defining a specific common terminology and the related set of scores.

Acknowledgements

References

Burton S.H., Tanner K.W., Giraud-Carrier C.G., West J.H., Barnes M.D., 2012, "Right Time, Right Place" health communication on Twitter: Value and accuracy of location Information, Journal of Medical Internet Research, 14(6):e156.

Bhuvana N., Arul Aram I., 2019, Facebook and Whatsapp as disaster management tools during the Chennai (India) floods of 2015, International Journal of Disaster Risk Reduction, 39, 101135, ISSN 2212-4209.

Hellmund T., Schenk M., Hertweck P., Moßgraber J., 2019, Employing Geospatial Semantics and Semantic Web Technologies in Natural Disaster Management, Semantics Conference, Karlsruhe (Germany).

Imran M., Castillo C., Diaz F., Vieweg S., 2018, Processing Social Media Messages in Mass Emergency: A Survey, Companion of The Web Conference 2018, 507-551.

Lazreg M. B., Goodwin M., Granmo O., 2019, An Iterative Information Retrieval Approach from Social Media in Crisis Situations, International Conference on Information and Communication Technologies for Disaster Management (ICT-DM), 1-8.

Pourebrahim N., Sultana S., Edwards J., Gochanour A., 2019, Understanding communication dynamics on Twitter during natural disasters: A case study of Hurricane Sandy, International Journal of Disaster Risk Reduction, 37:101176.

Sheth A.,1999, Changing Focus on Interoperability in Information Systems: From System, Syntax, Structure to Semantics. In: Goodchild M., Egenhofer M., Fegeas R., Kottman C. (Eds.) Interoperating Geographic Information Systems. The Springer International Series in Engineering and Computer Science, vol 495, Boston, MA.

UNDRR United Nations Office for Disaster Risk Reduction , Sendai Framework for Disaster Risk Reduction 2015-2030, <https://www.undrr.org/publication/sendai-framework-disaster-risk-reduction-2015-2030> accessed 25.10.2021.

Zhang C., Fan C., Yao W., Hu X., Mostafavi A, (2019), Social Media for intelligent public information and warning in disasters: An interdisciplinary review, International Journal of Information Management, 49, 190-207.