

Emotional analysis of safeness and risk perception of London and Rome railway stations during the COVID-19 pandemic

Fabio Garzia
*Safety & Security Engineering
Group - DICMA*
SAPIENZA – University of Rome
Rome, Italy
&
Wessex Institute of Technology
Southampton, UK
&
European Academy of Sciences
and Arts
Salzburg, Austria
fabio.garzia@uniroma1.it

Francesco Borghini
*Safety & Security Engineering
Group - DICMA*
SAPIENZA – University of Rome
Rome, Italy

Mara Lombardi
*Safety & Security Engineering
Group - DICMA*
SAPIENZA – University of Rome
Rome, Italy

Luca Canfora
*Safety & Security Engineering
Group - DICMA*
SAPIENZA – University of Rome
Rome, Italy

Soodamani Ramalingam
*School of Engineering and
Computer Sciences*
University of Hertfordshire
Hatfield, UK

Abstract- The evaluation of perceived safeness and risk by persons is particularly valuable for safety and security handling. Each individual is influenced by others' opinions on safety and security and the Internet embodies the place where these opinions are mostly pursued, acquired, and estimated. From this point of view, social networks play a considerable influential role. For this reason, Opinion Mining and Sentiment Analysis have found noteworthy uses in a variety of situations and one of the most interesting is embodied by public safety and security.

The goal of this work is to examine the perception of safety and risk within the railway stations of London (UK) and Rome (Italy) during COVID-19 pandemic. In particular, the railway stations of London Victoria and Waterloo and the railway stations of Rome Termini and Tiburtina were considered, from March 23 to July 9, 2020, highlighting the emotional components in three distinct pandemic phases of the considered period in the two countries, by means of the semantic analysis of the textual contents present in Twitter.

Keywords—Perceived risk assessment for security & safety, Open-Source Intelligent Techniques for security & safety, OSINT, opinion mining for security & safety, sentiment analysis for security & safety, COVID-19 pandemic, railway stations security & safety.

I. INTRODUCTION

Sense of safety and risk perception evaluation are very important for security and safety administration. In this context, every person is influenced by the opinion of other persons and the Internet embodies the place where these opinions are mostly pursued, acquired, and estimated. From this point of view, social networks are characterized as influencing media. In fact, about 92% of buyers assert they have more confidence in social media evaluations than in any other sort of advertising. For this reason, Opinion Mining and Sentiment Analysis [1-13] have found noteworthy uses in a variety of

situations and one of the most interesting is embodied by public safety and security.

Security and safety managers can therefore exploit people's perception to detect any unforeseen weakness in the perceived risk of a site that can vary from time to time and take appropriate measures. This requires making continuous judgements about risk perception [1 - 13] for effectively decreasing the level of risk.

The aim of this work is to examine the perception of safeness and risk within the railway stations of London (UK) and Rome (Italy) during COVID-19 pandemic. In particular, the railway stations of London Victoria and Waterloo and the railway stations of Rome Termini and Tiburtina were considered, from March 23 to July 9, 2020, highlighting the emotional components in three distinct pandemic phases of the considered period in the two countries, by means of the semantic analysis of the textual contents present in Twitter. The assemblage of opinions to be utilized for this objective requires exploring various open sources (OSINT - Open-Source INTelligence) and consequently the treating of massive volumes of data (Big Data) in digital form from which to obtain information and knowledge. In our case, Twitter was used as a source and a suitable analysis of the keywords included in the various tweets was done. The keywords have been chosen pondering that the estimation of perceived risk is deeply connected to the psychological aspects, via the emotional responses produced by a certain location.

II. LONDON AND ROME RAILWAY STATIONS

Victoria station in London (UK) is in the City of Westminster, one of the districts of London, and is a central railway terminus of the metropolis. The station

is mainly used for connections between the city of London and the south-east area of Great Britain, offering fast connections to seaside resorts and seaports. In addition, the Gatwick Express departs from here, allowing direct connection to London - Gatwick airport. As one of the main railway stations in London, it guarantees the transit of approximately 205,000 passengers per day for a total number of approximately 74 million passengers per year.

Waterloo station in London (UK) is one of the main stations in London and the first in the United Kingdom in terms of passenger flow. The station is in the neighbourhood of Lambeth, and in the Waterloo area. It has an annual passenger flow that exceeds 94 million, including passengers who do not use medium and long-distance rail transport, but local transport. The station has 20 tracks.

Termini station in Rome (Italy) is characterized by a total area of 225,000 square meters with a passenger flow of approximately 480,000 per day for a total of over 150 million passengers per year and a daily transit of over 850 trains. It is the first railway station in Italy for the number of passengers carried. The railway station consists of 32 tracks, also representing the largest railway station in Italy. It is characterized by having both long-distance services, which include high-speed trains, and regional ones. There is also the Leonardo Express service which connects directly to the Rome - Fiumicino airport.

Tiburtina station is the second railway station in Rome both by volume and by type of traffic with about 140,000 daily transits and 51 million users a year and a daily transit of about 500 trains.

III. DESCRIPTION OF THE UTILIZED METHODOLOGY

The semantic knowledge contained in the data can be attained utilizing the Sentiment Analysis, also well-known as Opinion Mining. This estimation captures the personal beliefs of the users on an object of interest by the evaluation of the polarity of the feeling (which can be positive, negative, or neutral) utilized to illustrate the experience with the latter. Normally, the Sentiment Analysis is performed according to distinct phases showed in Fig. 1.

The selected data provider is embodied by Twitter that is the source where to obtain texts. It is one of the most used social networks in the world since it has one of the greatest percentages of public profiles with respect to others (Facebook® profiles and comments are normally not public). Furthermore, the text is the central component (Instagram® contains mostly photos) whose concision, 280 characters / tweet, lets assess each tweet in a more easy, fast, and effective manner.

There is much software available for social network assessment, but NodeXL [14] is often the most common as it does not require the knowledge of any programming language. NodeXL is a free plug-in for Excel® and the basic edition allows to import data and process via a suitable metrics. Due to the limitations of Twitter API, it is conceivable to obtain up to 2000 tweets or at least tweets no older than 7-8 days in the free of charge version, but these restrictions expand utilizing

the pro version. As this work demanded an enormous number of tweets to be collected and handled, the pro version was utilized. In the particular case, the metric ‘Words and word pairs’ has been employed. This particular metric computes the words and the pairs of words which are enclosed in the tweets. Basically, this metric achieves the Sentiment Analysis via text scanning, detecting the words which belongs to the two ‘Lists of Sentiment Words’ with the related number of occurrences.

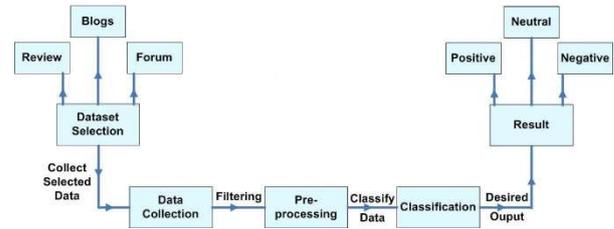


Fig. 1. General opinion mining steps.

The study focused on the examination of the tweets (both in English and Italian) within the railway stations of London (UK) and Rome (Italy) during COVID-19 pandemic, from March 23 to July 9, 2020, with the objective of getting the emotions and opinions of individuals to estimate their perception of safeness and risk in the considered places.

During the above studied periods, three distinct phases have been detected and examined both for London and Rome. They are shown in Tab. I and II respectively.

TABLE I. PHASES AND DESCRIPTIONS FOR LONDON (UK).

Phase	Period	Description
1	From the begin of pandemic to 11 th May 2020.	Total lockdown where the entire country stopped.
2	From 12 th May 2020 to 1 st June 2020.	Gradual release of restrictive measures and permission to do outdoor activities without any limitation.
3	From 2 nd June 2020 till the end of data collecting period.	Greater permission to inter-regional and international travels.

TABLE II. PHASES AND DESCRIPTIONS FOR ROME (ITALY).

Phase	Period	Description
1	From the begin of pandemic to 4 th May 2020.	Total lockdown where the entire country stopped.
2	From 5 th May 2020 to 3 rd June 2020.	Gradual release of restrictive measures.
3	From 4 th June 2020 till the end of data collecting period.	Greater permission to inter-regional and international travels and gradual increase of the activities authorized to reopen.

The primary English keywords used as a search term for London is represented by ‘London’ while the secondary keywords are represented by ‘Victoria (OR) Waterloo (OR) station (OR) train’. The primary Italian keywords used as a search term for Rome is represented by ‘Roma’ while the secondary keywords are represented by ‘Termini (OR) Tiburtina (OR) Stazione

(OR) FS' (where FS is the name of the main railway corporate in Italy).

After that the data were collected, it was necessary to filter the pertinent tweets, thus reducing their initial number, as shown in Tab. III.

TABLE III. NUMBER OF TWEETS BEFORE AND AFTER FILTERING.

City	Phase	Before filtering	After filtering
London	1	12495	5834
	2	4208	2531
	3	6490	3308
Rome	1	4817	3319
	2	3272	1797
	3	2955	1928

Later, suitable lists in English and Italian are produced. They are associated to positive words (list#1), negative words (list#2) and words connected to Covid-19 (list#3). All the other words not fitting the above lists are included in the non-categorized words section, as shown in Tabs. IV and V.

To ensure a meaningful outcome of the analysis, the word list, that clusters the categorization in one of the three groups, is produced step by step, trusting on the default English and Italian settings of the software. It is well known that dictionaries for the English and Italian language fitting for this type of assessment are already obtainable but, as some words are specific to the context, a novel enhanced lexicon of sentiments was formed.

Then, to obtain the emotional responses of persons, every word (positive or negative) was correlated with one of the 8 primary emotions aroused by it. As it is not conceivable to contemplate all the words of all the lists, only the first 12 more recurring words for each of the three periods relative to London and Rome (6 groups totally) were selected. To achieve this objective, suitable surveys were designed and filled by two randomly collected statistical reference sample (1 composed by English native speakers and 1 composed by Italian native speakers) consisting of 41 people, uniformly split between male and female, aged between 18 and 70. People were asked first the positive or negative valency of every word of the survey and then the correlated emotion stimulated by it, selecting between the 8 essential ones. After the outcomes of the surveys were collected and estimated, it was conceivable to get the prevailing emotion associated to every word.

IV. RESULTS

The managing of the data involved produced a considerable quantity of information associated to the studied collecting period. They are illustrated in Tab. IV and V.

As it is possible to see from Tables IV and V, the classes of un-categorized words are numerous. This is due to the insertion in these lists of all the keywords considering that, in a conventional conversation, the words that communicate a sure polarity, whether positive or negative, represent a subgroup. Regarding London, it is possible to see that the number of words of list#1 is always lesser than the number of words of list#2

in the 3 studied phases, indicating a persistent negative predominance. Regarding Rome, it is possible to see that number of words of list#1 is lesser than the number of words of list#2 in phase 1 and 3 while becomes a bit greater in phase 2, showing a different perception with respect to London situation.

TABLE IV. PARTITION OF WORDS USED FOR THE ANALYSIS OF THE 3 PHASES OF LONDON.

Words	Count Phase 1	Count Phase 2	Count Phase 3
Words in Sentiment List#1: Positive	4954	1311	2864
Words in Sentiment List#2: Negative	7863	5780	4270
Words in Sentiment List#3: Covid-19	1970	2140	2790
Non-categorized Words	125625	64787	72766
Total Words	140412	74018	82690

TABLE V. PARTITION OF WORDS USED FOR THE ANALYSIS OF THE 3 PHASES OF ROME.

Words	Count Phase 1	Count Phase 2	Count Phase 3
Words in Sentiment List#1: Positive	3540	4498	1457
Words in Sentiment List#2: Negative	5921	4420	2637
Words in Sentiment List#3: Covid-19	1089	825	611
Non-categorized Words	40157	45585	40842
Total Words	50707	55328	45547

Fig.s 2 and 3 shows the distribution of the number of tweets across the considered timeline, split into the 3 phases of London and Rome respectively.

It should be remarked the distinct behaviour of the timeline graphics for London and Rome due for the distinct developments of general situation in both UK and Italy in the 3 different periods, as well as the usage of Twitter.

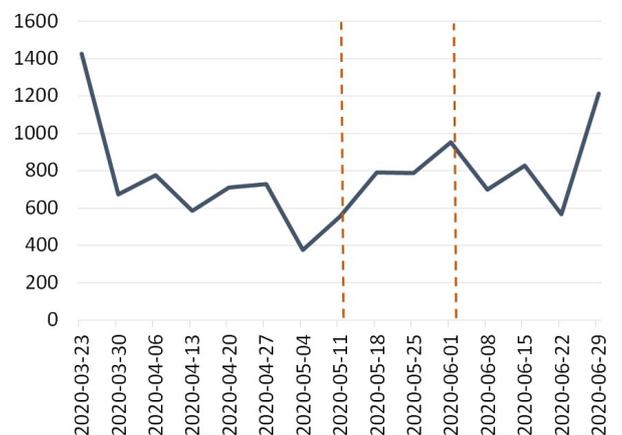


Fig. 2. Considered timeline of tweets for the 3 individuated phases of London.

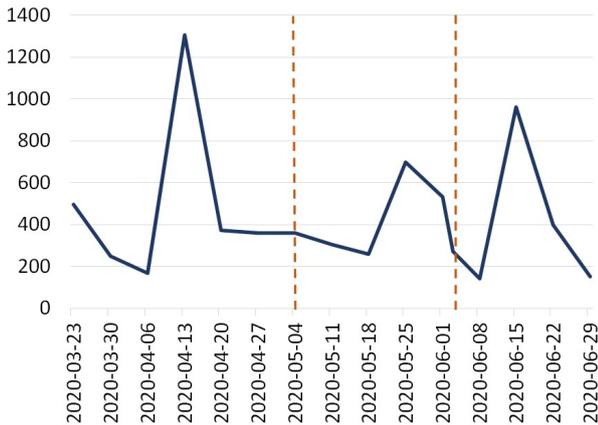


Fig. 3. Considered timeline of tweets for the 3 individualized phases of Rome.

Regarding phase 1 for London, the sharing of every primary emotion is shown in Fig.4, while the ranking of the most cited words is shown in Fig.5.

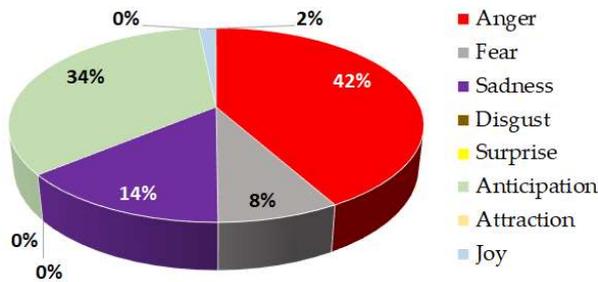


Fig. 4. Percentage distribution of every primary emotion for phase 1 of London.

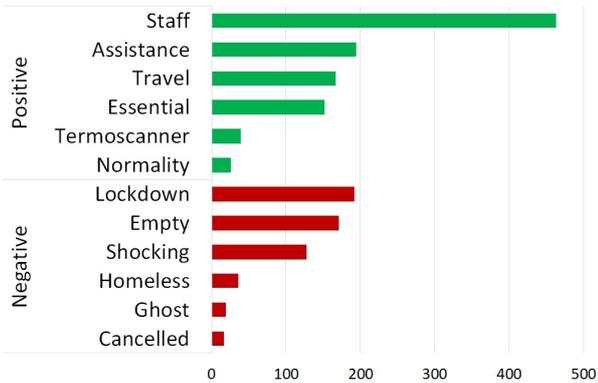


Fig. 5. Ranking of the most cited English words for phase 1 of London.

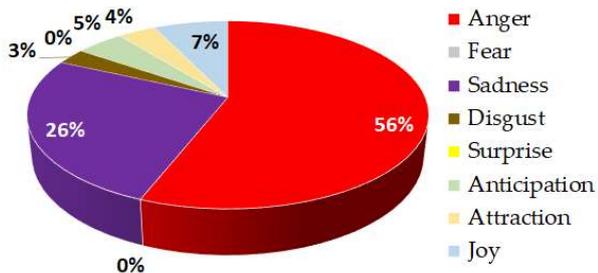


Fig. 6. Percentage distribution of every primary emotion for phase 2 of London.

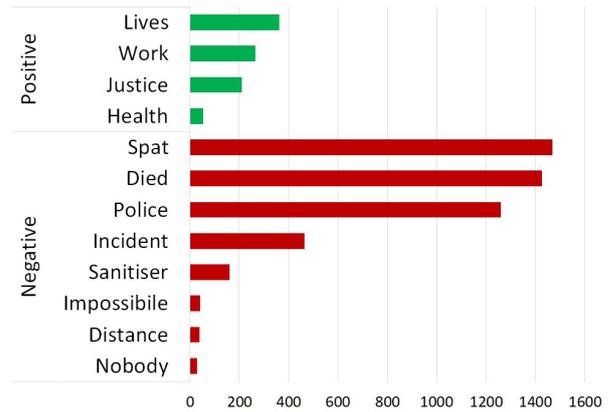


Fig. 7. Ranking of the most cited English words for phase 2 of London.

As it is possible to see from Fig.6, the dominant emotion occurring from the experience of train passengers and users of London railways during Phase 2 of the Coronavirus pandemic is 'Anger' (56%), followed by 'Sadness' (26%), 'Joy' (7%), 'Anticipation' (5%), 'Attraction' (4%) and 'Disgust' (3%). The emotions of 'Fear', and 'Surprise' are not present. Even in this phase it is possible to affirm that there is a dominance of negative emotions.

Regarding phase 3 for London, the sharing of every primary emotion is shown in Fig.8, while the ranking of the most cited words is shown in Fig.9.

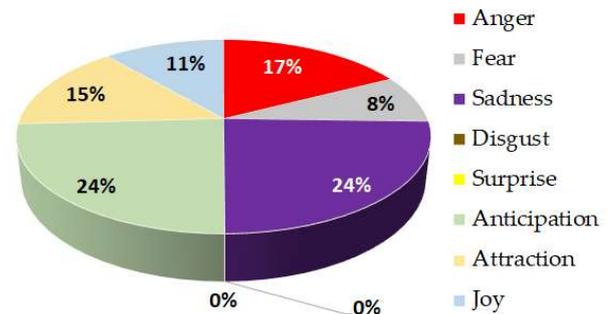


Fig. 8. Percentage distribution of every primary emotion for phase 3 of London.

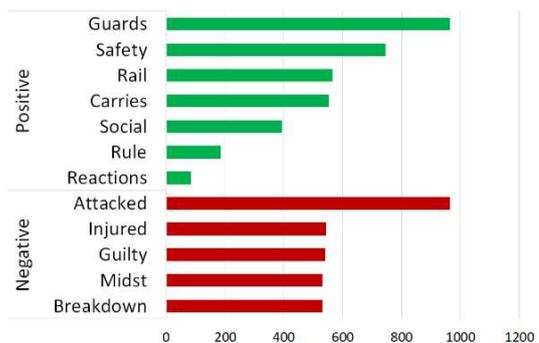


Fig. 9. Ranking of the most cited English words for phase 3 of London.

As it is possible to see from Fig.8, the dominant emotions occurring from the experience of train passengers and users of London railway stations during Phase 3 of the Coronavirus pandemic are 'Sadness' (24%) and 'Anticipation' (24%), followed by 'Anger' (17%), 'Attraction' (15%), 'Joy' (11%) and 'Fear' (8%).

The emotions of 'Disgust' and 'Surprise' are not present. It is therefore possible to affirm that this phase is characterized by a certain balance between positive and negative emotions.

Regarding phase 1 for Rome, the sharing of every primary emotion is shown in Fig.10, while the ranking of the most cited words is shown in Fig.11.

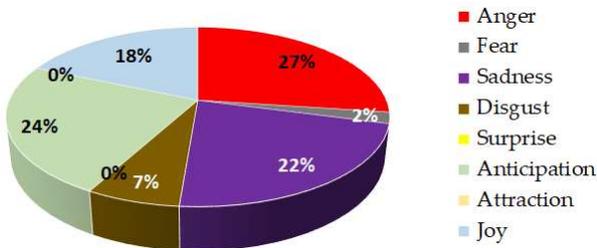


Fig. 10. Percentage distribution of every primary emotion for phase 1 of Rome.

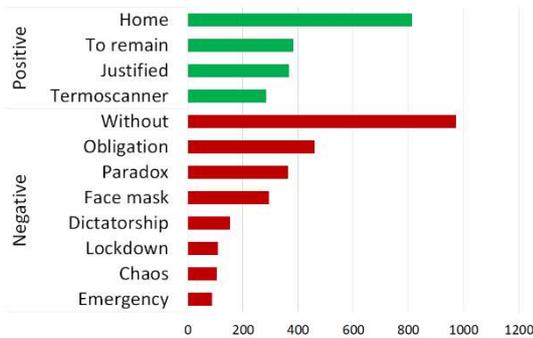


Fig. 11. Ranking of the most cited Italian words (translated in English) for phase 1 of Rome.

As it is possible to see from Fig.10, the dominant emotion occurring from the experience of train passengers and users of Rome railway stations during Phase 1 of the Coronavirus pandemic is 'Anger' (27%) followed by 'Anticipation' (24%), 'Sadness' (22%), 'Joy' (18%), 'Disgust' (7%), 'Fear' (2%). The emotions of 'Attraction' and 'Surprise' are not present. It is therefore possible to affirm that this phase is characterized by a light dominance of negative emotions.

Regarding phase 2 for Rome, the sharing of every primary emotion is shown in Fig.12, while the ranking of the most cited words is shown in Fig.13.

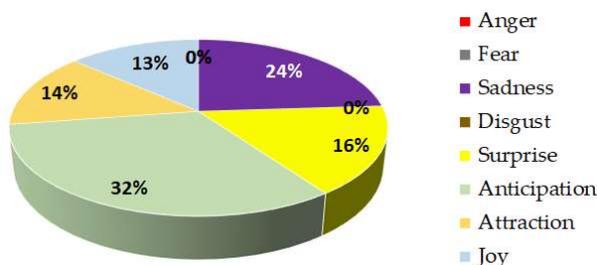


Fig. 12. Percentage distribution of every primary emotion for phase 2 of Rome.

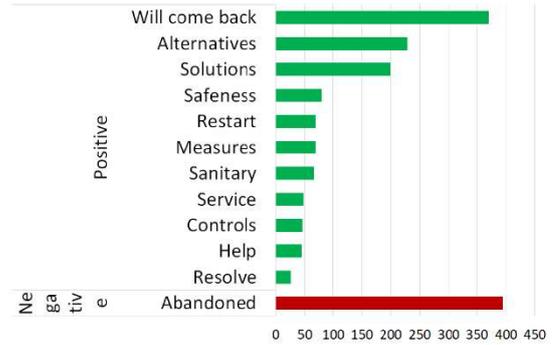


Fig. 13. Ranking of the most cited Italian words (translated in English) for phase 2 of Rome.

As it is possible to see from Fig.12, the dominant emotion occurring from the experience of train passengers and users of Rome railway stations during Phase 2 is 'Anticipation' (32%) followed by 'Sadness' (24%), 'Surprise' (16%), 'Attraction' (14%), 'Joy' (13%). The emotions of 'Anger', 'Disgust' and 'Fear' are not present. It is therefore possible to affirm that this phase is characterized by a light dominance of positive emotions.

Regarding phase 3 for Rome, the sharing of every primary emotion is shown in Fig.14, while the ranking of the most cited words is shown in Fig.15.

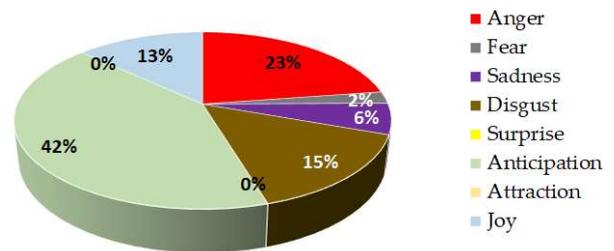


Fig. 14. Percentage distribution of every primary emotion for phase 3 of Rome.

As it is possible to see from Fig.14, the dominant emotion occurring from the experience of train passengers and users of Rome railway stations during Phase 3 is 'Anticipation' (42%), followed by 'Anger' (23%), 'Disgust' (15%), 'Joy' (13%), 'Sadness' (6%) and 'Fear' (2%). The emotions of 'Attraction' and 'Surprise' are not present. It is therefore possible to affirm that this phase is characterized by a dominance of positive emotions.

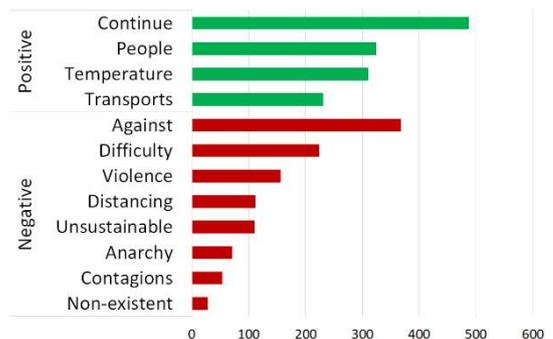


Fig. 15. Ranking of the most cited Italian words (translated in English) for phase 3 of Rome.

Regarding the London case study, the number of tweets, with a decreasing trend in the initial phase and then alternating recovery towards the end of Phase 3, indicates a progressive loss of self-esteem and trust in institutions accompanied by a negative Sentiment that fails to change, sign of the uncertainty of events. To understand the true motive of such behavior it is necessary to go into the deep dynamic analysis of emotions from which it is clear that the 'Anger', dominant in the phantasmatic impact with the Coronavirus, actually refers to the conflict in which its inseparable antithetical companion is involved, the 'Fear', removed in the unconscious because it was never presumptuously accepted. The same 'Anger' that dominates Phase 1 and Phase 2 and that in Phase 3 yields its primacy to 'Sadness' and 'Expectation' which precedes it and then accompany it with the 'Resignation' that derives from their secondary overlap.

Regarding Rome, from the psychodynamic point of view, it is possible to affirm that:

- in Phase 1 the events evoke from the subconscious memory in descending order 'Anger' (A), 'Expectation' (E), 'Joy' (J), 'Disgust' (D), 'Fear' (F), and this effect in addition to arousing complex secondary emotions in various percentages like 'Vendetta' (A + E), 'Sadism' (A + J) or 'Hope' (E + J), it refers to the respective antithetical emotions totally not expressed: 'Surprise' and 'Attraction', the maximum unforgivable conflictual provocation.
- in Phase 2 the events evoke from the subconscious memory in descending order 'Expectation' (E), 'Sadness' (SA), 'Surprise' (S), 'Attraction' (AT), 'Joy' (J), 'Disgust' (D), 'Fear' (F), and such effect in addition to arousing in various percentages complex secondary emotions such as 'Resignation' (E + SA), 'Omnipotence' (E + S), 'Disappointment' (SA + S), 'Nostalgia' (SA + AT), 'Amazement' (AT + S), refers to the only totally unexpressed antithetical emotion: 'Anger', the maximum unacceptable conflictual provocation.
- in Phase 3 the events evoke from the subconscious memory in descending order 'Expectation' (E), 'Anger' (A), 'Disgust' (D), 'Joy' (J), 'Sadness' (SA), 'Fear' (F), and this effect in addition to arousing in varied percentage of complex secondary emotions such as 'Vendetta' (E + A), 'Suffering' (E + D), 'Contempt' (A + D), etc., refers to the respective antithetical emotions totally not expressed: 'Surprise' and 'Attraction' (as in Phase 1), the maximum unforgivable conflict provocation.

V. CONCLUSIONS

In the present work, a method for the evaluation of perception of safeness and risk within the railway stations of London (UK) and Rome (Italy) during COVID-19 pandemic has been presented, giving value to the emotional components, via the semantic analysis of the textual contents present in Twitter. It represents a valuable method as it allows to have constantly opinions

about safeness and risk perception in ordinary and critical situations, such as the considered one.

REFERENCES

- [1] F. Borghini, F. Garzia, G., Borghini, A. Borghini, *The Psychology of Security, Emergency and Risk*, WIT Press (UK and Boston - USA), 2016.
- [2] W. Medhat, A. Hassan, A. H. Korashy, H., "Sentiment analysis algorithms and applications: A survey", *Ain Shams Engineering Journal*, 1093-1113, 2014.
- [3] S. Ramalingam, "Metadata Extraction and Classification of YouTube Videos Using Sentiment Analysis" *Proc. of IEEE International Carnahan Conference on Security Technology*, Orlando (USA), 2016.
- [4] E. Cambria, B. Schuller, Y. Xia, C. Havasi, "New Avenues in Opinion Mining and Sentiment Analysis", *IEEE Intelligent Systems*, Vol. 28, No.2, pp. 15-21, 2013.
- [5] F. Garzia, R. Cusani, F. Borghini, B. Saltini, M. Lombardi, S. Ramalingam "Perceived risk assessment through Open-Source Intelligent Techniques for Opinion Mining and Sentiment Analysis: the case study of the Papal Basilica and Sacred Convent of Saint Francis in Assisi, Italy", *Proc. of IEEE International Carnahan Conference on Security Technologies*, Montreal (Canada), pp. 11 - 15, 2018.
- [6] F. Garzia, F. Borghini, A. Bruni, M. Lombardi, P. Mighetto, S. Ramalingam, S. B. Russo, "Emotional Reactions to the Perception of Risk in the Pompeii Archaeological Park", *International Journal of Safety & Security Engineering, IETA (International Information and Engineering Technology Association)*, Vol.10, No.1, pp. 11 - 16, 2020.
- [7] S. M. Park, D. K. Baik, Y. G. Kim, Y.-G. (2016). Sentiment User Profile Analysis based on Forgetting Curve in Mobile Environments. *Proceedings IEEE 15th International Conference on Cognitive Informatics & Cognitive Computing*, pp. 207-211, 2016.
- [8] S. Anastasia, I. Budi, "Twitter Sentiment Analysis of Online Transportation Service Providers" *Proc. of International Conference on Advanced Computer Science and Information Systems*, pp. 359-365, 2016.
- [9] S. Akter, M. Aziz, "Sentiment Analysis On Facebook Group Using Lexicon Based Approach" *Proc. of 3rd International Conference on Electrical Engineering and Information Communication Technology*, pp. 1-4, 2016.
- [10] Shaikh T., Deshpande D., "A review on Opinion Mining and Sentiment Analysis", *International Journal of Computer Applications (0975-8887)*, 2016
- [11] E. Mayor, L. M. Bietti, "Twitter, time and emotions", *R. Soc. Open Sci.* 8: 201900, 2021. <https://doi.org/10.1098/rsos.201900>.
- [12] B. Liu, *Sentiment Analysis: Mining Opinions, Sentiments, and Emotions*, Cambridge University Press, 2015.
- [13] M. Vijarania, *Sentiment Analysis: A Combined Approach*, Lambert Academic Publishing, 2020.
- [14] <https://www.smrfoundation.org/nodexl/>