

## **The perceptions of migration during the pandemic: what Twitter data tell us?**

Sara Miccoli and Elena Ambrosetti

Department of Methods and Models for Economics, Territory and Finance, Sapienza University of Rome

### *Abstract*

The paper focuses on the perceptions on the immigration phenomenon in Italy expressed through social media. In particular, the paper aims to analyze sentiments and emotions about migration in Italy; analyzing the sentiment about this phenomenon in relation to the COVID-19 emergency. The case study of Italy is particularly interesting for several reasons: Italy has been severely affected by the COVID-19, it is one of the largest recipients of immigrants in the European Union and, Italy is among the few countries that have implemented an amnesty for irregular migrant workers during the pandemic. The study is conducted by applying a text mining and sentiment analysis to the tweets which contain hashtags and keywords related to the migration and to the COVID-19 pandemic. Preliminary results show that tweets related to migration express a sense of emergency and also invasion and that the migration issue is often addressed on social media focusing on particular aspects of the migratory phenomenon. Both negative and positive sentiments are present in the tweets, confirming also a certain polarization in the public discourse about migration.

**Keywords:** migration, COVID-19, twitter, sentiment analysis

### *1. Introduction*

The paper focuses on the perceptions on the immigration phenomenon in Italy expressed through the social media<sup>1</sup>. In particular, the paper has the following objectives: analyzing the perceptions, sentiments and emotions expressed on the immigration flows in Italy, immigrants and the migratory phenomenon in general; analyzing the perceptions, sentiments and emotions about this topic during and in relation to the COVID-19 emergency and to the amnesty for irregular migrant workers implemented to respond to the lack of seasonal workers during the pandemic.

This work is conducted by applying text mining and sentiment analysis to the tweets which present hashtags and keywords related to the migratory phenomenon and to the same issue in connection with the COVID-19 pandemic. Twitter presents some characteristics that make it particularly appropriate for the study of perceptions and ideas on certain phenomena. In a few minutes, via a computer or smartphone, users can share their thoughts on certain issues (by using a maximum of 280 characters), appreciations of other users' tweets, comments on the principal daily news. By analyzing the content of the tweets, it is possible to understand

---

<sup>1</sup> The paper is developed within the project PERCEPTIONS, funded by the European Commission H2020 Research & Innovation Action under Grant Agreement No 833870 and the project A multidisciplinary approach to the representation of migration in Italy (2013-2020) funded by Sapienza University of Rome

negative or positive sentiments towards certain topics, such as migration phenomenon in relation to COVID-19 emergency.

The study focuses on an increasingly important topic for receiving countries, such as Italy. International migration is an important driver of demographic change and a relevant source of uncertainty for governance and population projections. Understanding the perceptions towards migration and immigrants is an essential step to developing specific actions to counteract some narratives and shed light about migration phenomenon among the public opinion. At the same time, COVID-19 pandemic has probably altered the perceptions and sentiment of uncertainty and insecurity often connected with migratory phenomenon. Analyzing the way in which this new emergency relates with perceptions about migration is essential to understand another consequence of the health uncertainty caused by the pandemic.

## *2. Motivation of the study*

The pandemic of the coronavirus disease (COVID-19) is the greatest challenge that Italy and the rest of the world have faced since World War II. International migration trends have changed since the beginning of the year 2020 because most of the countries of the world closed their borders due the pandemic. Despite the COVID-19 emergency and the decrease in migration flows, the topic of migration continued to be addressed and discussed in political and public debate. Over the years in Italy the debate around international migration has been constantly accompanied by the dichotomy between the “economic acceptance of immigration” and the “political rhetoric of growing hostility and apparent closure” (Ambrosini 2013: 176). Ambrosetti & Papparuso (2018) argued that until 2011, recognition of the economic role of immigrants provided political and social legitimacy for their presence in Italy. However, since then the perception of immigration as a security and public order problem has prevailed over economic acceptance. Then, as a consequence of the so-called refugee crisis the debate on migration in Italy in the last decade became increasingly complex: the governments and the public opinion constantly asked for more cooperation and solidarity from the EU.

The public debate on migration in Italy is shaped by the same traits of major Western hosting countries. As stressed by Colombo (2018: 163), “Politicians and journalists have commonly employed terms like ‘massive invasion’ to describe the phenomenon”. Migration is perceived as a threat, an invasion and most often migrants are defined as clandestine, irregular, illegal and undocumented (e.g. Colombo 2018; Terwal 1996). Despite the heterogeneity of approaches, previous research in the Italian context keeps on detecting robust stereotyped narratives of immigrants appearing as threats for host society or as victims of exclusion/injustice (Binotto et. al. 2016). Such a consistent body of evidence emerges from content and language analysis (Calvanese 2011; Cerase 2013; Stolfi & Brancato 2019) discourse analysis (Tudisca et al. 2019; Lucchesi 2019), and visual sociology (Pogliano 2011, 2015).

Given the growing importance of social media in shaping the debate about several social phenomenon and considering the recent Italian context, analyzing how the migration phenomenon is addressed in social media is very important. The case study of Italy is particularly interesting for several reasons: firstly, because Italy has been severely affected by the COVID-19 paying a large death toll and experiencing one of the stricter and

Draft: please do not circulate or quote without permission of the authors

longer lockdown within the Western world, second because Italy is one of the largest recipients of immigrants in the European Union (EU) as one of the main port of access to the EU, third, Italy is among the few countries that have implemented a regularization program for irregular migrant workers during the pandemic.

### 3. *Research design and methodology*

The aim of this work is to analyze the perceptions expressed through the social media Twitter on the migration phenomenon in Italy and, in particular, to study the perceptions about this topic in connection with the COVID-19 emergency and the amnesty.

Words, sentences and thoughts expressed in the tweets are analyzed by applying text mining and sentiment analysis. In general, this kind of analysis permits to study communication processes in various areas, such as the commercial sector, political sphere, traditional mass media and social media. Through this analysis it is possible to discover perceptions, opinions and ideas about certain products, politicians or social phenomena. Nowadays, the sentiment analysis is often used to study the communication that occurs within social media. Among the social media, Twitter has specific characteristics which make it particularly appropriate for this analysis. Indeed, thanks to an essential layout both for smartphones and pc, users can easily and quickly express and communicate opinions and ideas, show appreciation for other tweets, share ideas of other users thanks to the re-tweeting instrument. Because of the maximum length of 280 characters, users need to be very concise in expressing their thoughts and opinions. The need to write short texts prompt users to use specific terms, while the simple layout and shortness of the messages pushes them to post messages frequently and regularly. In addition to the use of specific words within the messages, another relevant feature of Twitter is the use of the hashtag (#) followed by keywords. These hashtags make it possible to easily define the topic of the tweets. Tweets were collected by accessing Application Programming Interfaces (APIs) through the Twitter Academic Research product that allows to access Twitter's real-time and historical public data. We collected tweets from February 01, 2020 to April 30, 2021. In order to find tweets covering the topics migration and COVID-19 we looked for tweets containing various Italian words connected to migration phenomenon and the hashtag COVID-19. Therefore, the search query included the Italian words for *migrations*, *migrants*, *immigration*, *illegals*, *Mediterranean*, *refugees*, *amnesty* and the hashtags *#covid19*, *#coronavirus*. The word *illegals* was included since illegal migration is a significant component of the total immigrant population in Italy. Our keywords also included the word *amnesty*, since an amnesty was applied in spring 2020 to compensate for the pandemic-related lack of seasonal workers. After cleaning and checking the dataset, we obtained a dataset of 25,000 tweets.

We applied text mining and sentiment analysis techniques to analyze the words used in these tweets and to understand what kind of sentiment these tweets express. Specifically, we analyzed the most frequent words within tweets and the associations between some of our keywords and other words (we used “TM” R package). We then conducted a sentiment analysis to understand what sentiments and emotions the tweets expressed (we used “syuzhet” R package). Sentiments can be classified as positive, negative or neutral. The “syuzhet” package has four sentiment dictionaries and four methods to detect sentiments. These four methods are *syuzhet*,

Draft: please do not circulate or quote without permission of the authors

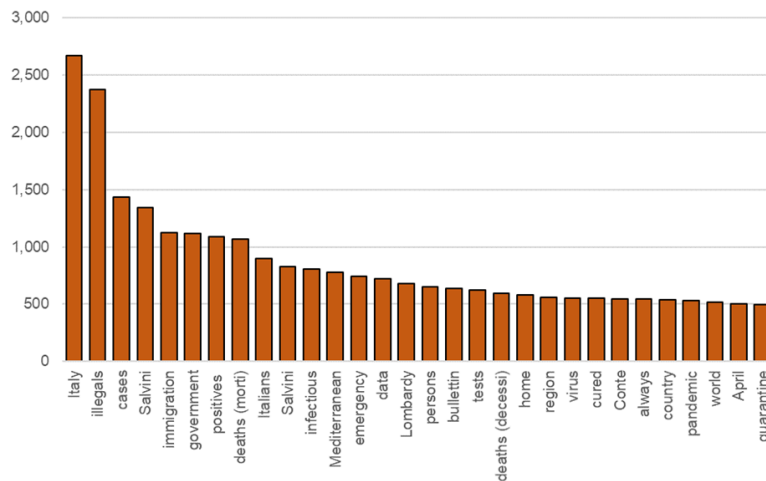
*bing*, *afinn* and *nrc* and they use different scales and give different results. The *syuzhet* lexicon was developed by the Nebraska Literary Lab under the direction of Matthew L. Jocker and the scores are equal or above +1 for words associated to the positive sentiment and equal or below -1 for negative words. The *bing* lexicon was developed by Mingqing Hu and Bing Liu as the Opinion Lexicon and the scores are equal or above +1 for words associated to the positive sentiment and equal or below -1 for negative words. The *afinn* algorithm was developed by Finn Arup Nielsen as the Afinn Word Database, and extended by analyzing Twitter posts, and it contains slang and words used on the web. Scores are equal or above +5 for positive words and equal or below -5 for the negative ones (Naldi 2019; Widyaningrum et al. 2019). All these methods assign a score (from negative to positive) to the words which compose tweets and the score for every tweet is the algebraic sum of the word's scores. The *nrc dictionary* was developed by Mohammad, Saif M. and Turney, Peter D. as the NRC Emotion Lexicon and it is different from the others because it does not only focus on polarity considering positivity and negativity, but it also assigns words to 8 different types of sentiments, categorized as emotions (Naldi 2019; Widyaningrum et al. 2019). The 8 emotions are anger, fear, sadness, disgust, anticipation, trust, surprise, and joy. For each tweet the method counts the number of words associated with each category (Naldi 2019).

This package and these methods, like others used in sentiment analysis, have limitations, such as the inability to detect sarcasm or analyze context. Also, *syuzhet* and *bing* dictionaries do not always consider negativity in a proper way. Despite this, however, the use of the four dictionaries allows to have an informative overview of the sentiments expressed in the tweets; furthermore, the possibility to investigate also the typology of the sentiments, that is the emotions, certainly enriches the analysis.

#### 4. Preliminary findings and conclusions

The three most frequent words in the analyzed tweets are *coronavirus*, *covid* and *migrants*, three of our keywords. They appear more than 3000 times and they are followed by words which appear less than 3000 times. Many of these words refer to the COVID-19 emergency (Table 1). References to political and governmental figures often appear in these tweets. We found the name of Salvini, the former interior minister and head of the main Italian sovereigntist party, and Conte, head of the Italian government during the first phases of the emergency related to the pandemic. Words related to arrivals by sea also appear frequently in our dataset, although they do not are among the 30 most frequent words. The associations between some migration-related keywords and other words show that the words related to migration are connected with words that express the sense of an uncontrolled and risky phenomenon. For example, the term *immigration* is associated with the term *facilitation*, the term *regularization* is even associated with the word *drug trafficking*. This suggests that immigration is perceived as uncontrolled, dangerous, and a problem.

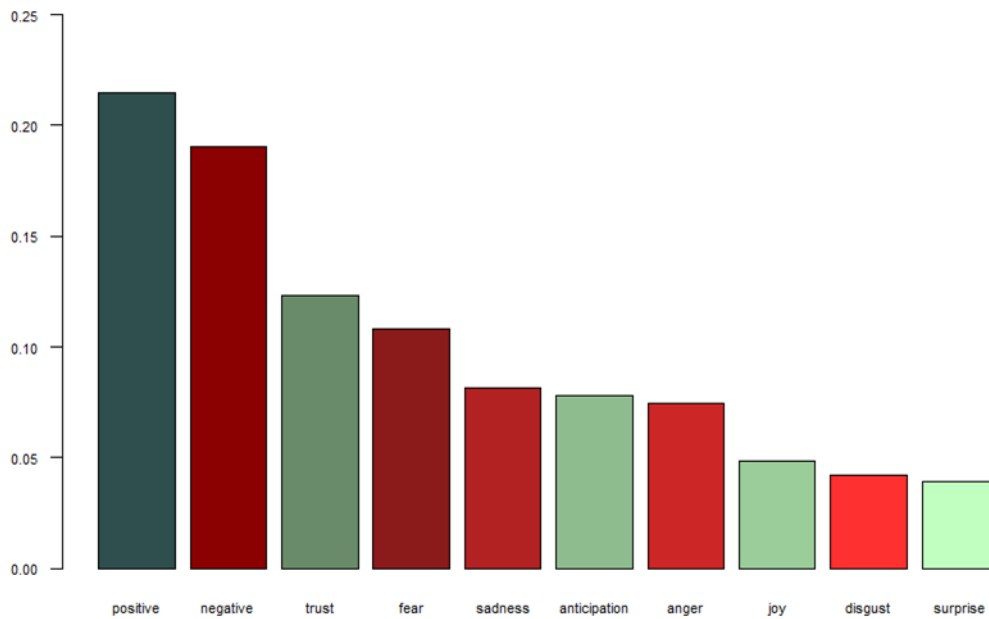
Table 1. 30 most frequent words (<3000 counts). 25,000 tweets. 01/02/2020-30/04/2021



Source: elaboration on TWITTER data

As far as sentiment analysis is concerned, looking at the mean scores of all tweets for the first three methods, the values are very close to 0. This result shows that, on average, the tweets seem to express neutrality. However, by looking at the minimum and maximum scores, it seems that this average neutrality hides also peaks of positivity and negativity. By checking sentiments and emotions of all 25,000 tweets, we found out that that positive and negative words are equivalent, with a slight preponderance of positive ones (Figure 1). This demonstrate what we found out with the other methods, which detected a neutrality in average behind which there are peaks of negativity and positivity. Regarding the emotions, the first one is trust, followed by the emotion fear. Words associated with this emotion account each more than 10% of all words. The words associated with the positive emotion of surprise are the least frequent.

Figure 1. Sentiments and emotions of the words. 25,000 tweets. 01/02/2020-30/04/2021



Source: elaboration on TWITTER data

These first results show, on the one hand, that the COVID pandemic has somehow masked the focus on migration, at least in a certain period. Indeed, many of the most frequent words are strictly related to the COVID-19 emergency. The analysis of the association between some keywords and other words, shows that migration-related keywords are associated with negatively connotated words that transmit the idea of migration as a phenomenon difficult to manage, uncontrollable, forerunner of problems and risks, and strongly associated with illegality, thus providing a negative view of the phenomenon.

Sentiment analysis, exploited with four different methods, gave the same results regarding the sentiments. In average, there is a neutrality, but this average neutrality hides peaks of negativity and positivity, sentiments that are both present. In our database, there is almost the same number of words associated with positive and negative sentiments.

Our work confirms that the public debate regarding migration on social media is strongly linked to certain aspects of the migration phenomenon in Italy, thus strongly polarized (de Rosa et al. 2021). As emerged in other studies and as observed in the current reality, migration seems to be perceived as an emergency and a problem to be solved and migrants are perceived either as a threat or as victims. The COVID-19 emergency exacerbated the existing negative sentiments and perceptions associated with the migratory phenomenon. The future of the pandemic is uncertain, and the consequences of the restrictive measures will continue for a long time. In some sectors, the consequences could even worsen when some direct government aids will no longer be granted. Feelings of injustice, fear, social tensions may worsen the negative feelings related to the migration phenomenon. At the same time, in a world where movements are limited for health reasons, migration flows could further decrease.

Draft: please do not circulate or quote without permission of the authors

Undoubtedly, if it is true that social media is somehow a reflection of what is written and said in other media and what is actually conveyed in different sectors of society, at the same time it is also true that the discourse on social media contributes to focus the attention of politicians and society on certain issues. For this reason, knowing, through the analysis of terms and sentiments, how social audiences deal with migration issues is of particular interest, within the scientific community and in political institutions, in order to be able to direct studies, research and debates on a complex phenomenon.

### References

- Ambrosetti, E., & Paparusso, A. (2018). Migrants or refugees? The evolving governance of migration flows in Italy during the “refugee crisis”. *Revue europeenne des migrations internationales*, 34(1), 151-171.
- Ambrosini, M. (2013). Immigration in Italy: between economic acceptance and political rejection. *Journal of International Migration and Integration*, 14(1), 175–194.
- Binotto, M., Bruno, M., & Lai, V. (2016). Tracciare confini: l'immigrazione nei media italiani. *Tracciare confini*, 1-284.
- Bonifazi, C., Strozza S. (2020). La regolarizzazione al tempo de coronavirus. *Neodemos*. <https://www.neodemos.info/2020/05/29/la-regolarizzazione-al-tempo-del-coronavirus/>
- Calvanese, E. (2011). *Media e immigrazione tra stereotipi e pregiudizi. La rappresentazione dello straniero nel racconto giornalistico: La rappresentazione dello straniero nel racconto giornalistico*. FrancoAngeli.
- Cerese, A. (2013). Colpevoli per elezione: gli immigrati nella lente della cronaca nera. *Comunicazione Punto doc*, 69-88.
- Colombo, M. (2018). The representation of the “European refugee crisis” in Italy: Domopolitics, securitization, and humanitarian communication in political and media discourses. *Journal of immigrant & refugee studies*, 16(1-2), 161-178.
- de Rosa, A. S., Bocci, E., Bonito, M., & Salvati, M. (2021). Twitter as social media arena for polarised social representations about the (im) migration: The controversial discourse in the Italian and international political frame. *Migration Studies*. <https://doi.org/10.1093/migration/mnab001>
- Lucchesi, D. (2019). La “crisi migratoria” e la crisi della sfera pubblica. *ECHO*, (1), 179-194.
- Naldi, M. (2019). A review of sentiment computation methods with R packages. *arXiv preprint arXiv:1901.08319*.
- Stolfi, M., & Brancato, G. (2019). Una issue al servizio dei talk show: la rappresentazione del fenomeno migratorio nei format televisivi d'approfondimento in Italia. *Una issue al servizio dei talk show: la rappresentazione del fenomeno migratorio nei format televisivi d'approfondimento in Italia*, 211-228.
- Pogliano, A. (2011). Nominare mostrando: l'influenza della visibilità sul racconto giornalistico dell'immigrazione. *Nominare mostrando: l'influenza della visibilità sul racconto giornalistico dell'immigrazione*, 231-251.
- Terwal, J. (1996). The social representation of immigrants: the Pantanella issue in the pages of La Repubblica. *Journal of Ethnic and Migration Studies*, 22(1), 39-66.

Draft: please do not circulate or quote without permission of the authors

Tudisca, V., & Valente, A. (2019). The Gender Issue in Migrant Representation: a Case Study on Italian School Textbooks and Online Newspapers. *of the Journal Scuola Democratica*, 169.

Vieira, C. M., Franco, O. H., Restrepo, C. G., & Abel, T. (2020). COVID-19: The forgotten priorities of the pandemic. *Maturitas*.

Widyaningrum, P., Ruldeviyani, Y., & Dharayani, R. (2019). Sentiment Analysis to Assess the Community's Enthusiasm towards the Development Chatbot Using an Appraisal Theory. *Procedia Computer Science*, 161, 723-730.

Zincone, G. (2006). The making of policies: immigration and immigrants in Italy. *Journal of Ethnic and Migration Studies*, 32(3), 347–375.

DRAFT