Twitter Opinion Mining related to Refugee Crisis: A comparison between the 2015 Syrian refugee crisis and the 2022 Ukrainian refugee crisis

Amina Matt
amina.matt@epfl.ch
EPFL

Yurui Zhu
yurui.zhu@epfl.ch
EPFL

Margaux L'Eplattenier
margaux.leplattenier@unil.ch
University of Lausanne

Abstract

In the past ten years, major political conflicts have led millions of people to flee from their home, seeking asylum in neighboring countries. The most important refugee crisis is the Syrian 2015 refugee crisis, which has raised many political and ethical concerns in European and "Western" countries. Recently, the Russian invasion of Ukraine has generated another refugee crisis, and these refugees from Ukraine seem to be receiving a much warmer welcome in Western countries. We show that media coverage is also biased toward those refugee crises, exhibiting double standards; Syrian refugees tend to be presented as "others", while Ukrainian refugees tend to be presented as "like us". Based on previous research (Öztürk et al., 2018 [**?**], Pope et al., 2018 [**?**], Wallace, 2018 [**?**]), a comparative analysis was performed on twitter reactions about the two refugee crises. With tweets extracted from a Kaggle [**?**] dataset containing tweets related to the Ukrainian-Russian conflict, we then conduct frequency analysis, sentiment analysis, and topic modelling. We show that although the topics covered are similar to those highlighted in previous research about the Syrian refugee crisis, some indicate a difference in treatment. The result of our sentiment analysis also records more "positive" tweets about Ukrainian refugees than Syrian refugees.

Keywords: refugee crisis, Ukraine, Syria, sentiment analysis, topic modelling, computational social media, twitter

1 Introduction

1.1 Political Context

In 2015, 1.3 million people came to the European continent to request asylum, making it the biggest movement of refugees in a single year since World War II. Those refugees were mainly Syrian civilians fleeing the civil war which began in response to the Arab Spring protests of March 2011. Most of these refugees were internally displaced within Syria or had fled to Turkey or Lebanon. Meanwhile, in 2014 Russia invaded and subsequently annexed Crimea, a strategic location situated in Ukrainian territory. This event can be seen as the prelude of the 2022 Russian invasion of Ukraine, following president Zelensky's intention of joining NATO. The timeline of the main events of the two conflicts are in appendix A.1.

1.2 Double Standards in Media

Since the beginning of the Russo-Ukrainian war, journalists have been making controversial statements concerning the Ukrainian refugees. To cite a few: Daniel Hannan, the Daily Telegraph, reported: "Ukraine is a European country. Its people watch Netflix and have Instagram accounts... War is no longer something visited upon impoverished and remote population."; Charlie D'Agata, from CBS News said, when talking about Ukraine: "This isn't a place, with all due respect, like Iraq or Afghanistan ... This is a relatively civilised, relatively European - I have to choose those words carefully too - city, where you wouldn't expect that or hope that is going to happen."; or Philippe Corbé, for BFM TV in France: "We're not talking here about Syrians fleeing the bombing of the Syrian regime backed by Putin. We're talking about Europeans leaving in cars that look like ours to save their lives."

These few examples – and there are many more like this – highlight a systematic double standard. When talking about Syrians, one highlights the differences: "This isn't a place like…", they come from a "remote" and "impoverished country", etc. On the other hand, the discourse about Ukrainians refugees accentuates the similarities: "cars that look like ours", "relatively civilized, relatively European", "its people watch Netflix and have Instagram accounts", *like we do.* This bias is also suggested by previous research. In her article about the Syrian crisis representation in media, Savannah Day has shown that there is a tendency of massification and passivation in the discourse about Syrian refugees; leading to a form of dehumanization, and what she calls a phenomenon of *othering* [**?**]. This tendency doesn't seem to apply so much to the discourse about Ukrainian refugees. While Syrians are presented as remote, poor, and generally culturally different, Ukrainians are the white Europeans, with "middle class" attributes *who could be our neighbours.*

1.3 Research goal: Comparative Analysis on twitter

This double standard in media has constituted the motivation for our study. Indeed, in her article Day has pointed out that history of media coverage regarding global humanitarian crises shows that with various tools and processes, media can shape public opinion and policy in whichever direction it desires [**?**]. Therefore, we want to know if this difference

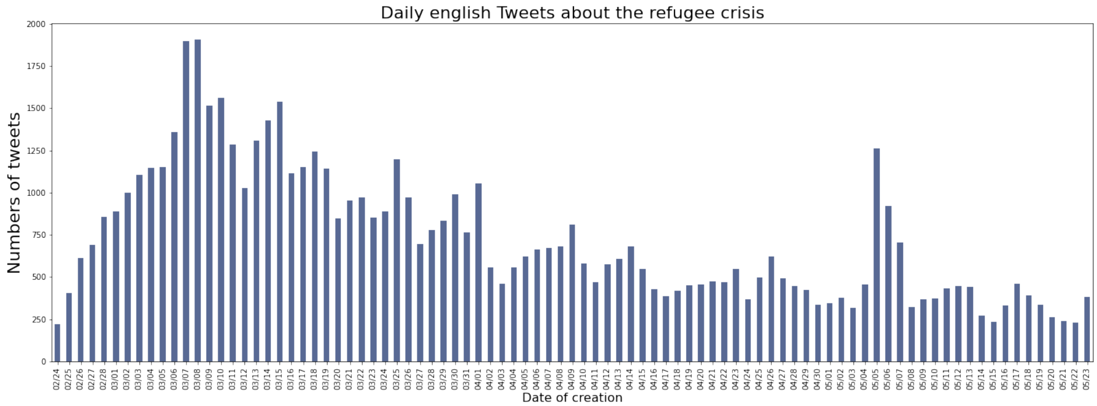


Figure 1. Daily English tweets distribution about the refugee crisis

in media coverage has an impact on the sentiments displayed in tweets talking about the topic.

1.4 Roles and Responsibility of Team

During the project, Margaux L'Eplattenier is responsible for the qualitative research on the Media coverage and the literature review for the previous work, establishing the project's direction, and generating some preliminary conclusions. Amina Matt is responsible for the initial dataset filtering and pre-processing, preparing the data for subsequent analysis and words frequency analysis. Yurui Zhu is responsible for constructing the machine learning pipeline, executing the pipeline to obtain results, aggregating the results, generating plots for analysis, and comparing the results to previous work. As for the report, Margaux L'Eplattenier handles the Introduction and Conclusion sections, as well as the interpretation of the results in the Result section. Amina Matt is responsible for the Data and Limitations sections, a portion of the results interpretation and Comparison section. Yurui Zhu is in charge of the Methods section and part of the Result and Comparison section.

2 Data

The data are extracted from a dataset on the topic of the Ukraine-Russia conflict shared by BwandoWando under the creative commons license, with a non-commercial attribution and share alike status[**?**]. The dataset is publicly available on Kaggle with the top usability rate according to Kaggle's administrators, ensuring the traceability of its contents. Early June 2022, the dataset size is 9GB and contains 37.43M tweets. The documentation available with the dataset mentions twitter.com as unique source. The data collection is computed with Azure ML and Anaconda notebooks. The code runs every 15 minutes to extract tweets based on 3 processes related to the Ukraine-Russia conflict. The 2 initial keywords lists and the third condition are the followings:

1. "#SlavaUkraini OR #Russia OR #RussiaUkraineWar OR #Putin OR #RussiaUkraine OR #RussianWar OR #ww3 OR #moscow OR #RussianConflict"
2. "#ukraineunderattack OR #Ukriane OR #Ukraine OR #RussianUkrainianWar OR #UkraineRussia OR #UkraineConflict OR #UkraineWar OR #Kharkiv OR #StopPutinNow"
3. Geolocation UKRAINE country.

From each process 2200 tweets are retrieved. Note that for the 2 processes based on hashtags filtering, hashtags are updated on the second, third and fourth run with the 15 most common hashtags from the previous sets[**?**]. Each hour, the updating is reset to the initial set of keywords. The replacement of hashtags allows to follow trends in hashtags.

The tweets used for our project are a selection of the Kaggle dataset.

We selected tweets from February 23 (starting date of Kaggle dataset), to May 23 included. Three months of data represents 35'120'500 tweets. Only the tweets with English as language were kept, which reduced the set to 23'742'066 tweets. As a large percentage (67.6%) of the tweets is in English, similar analysis would have been difficult to the lack of data. The second criteria for filtering focus on contents and select tweets related to the refugee crisis surrounding the conflict. Using similar filtering methods in [**? ?**], tweets were filtered so that they contain one of the followings keywords.

1. refugee
2. migrant
3. asylum seeker

The contents filtering resulted in 5'812'409 tweets. After the removal of retweets and deduplication, the final set contains 64'589 unique tweets.

2.1 Descriptive statistics

The distribution of tweets over time is shown in Figure 1. Meanwhile the extraction processes retrieve a constant amount of tweets per day, the publication creation time as well as the filtering processes modify the time distribution.

The tweets location were also retrieved when possible, although 27% of locations were unknown. The known locations and their frequency are presented in Figure 2. As locations aren't standardized we processed the metadata, including translation in the city/country language as well as grouping of cities into country. Cities with high percentage by themselves are kept to highlight the importance.

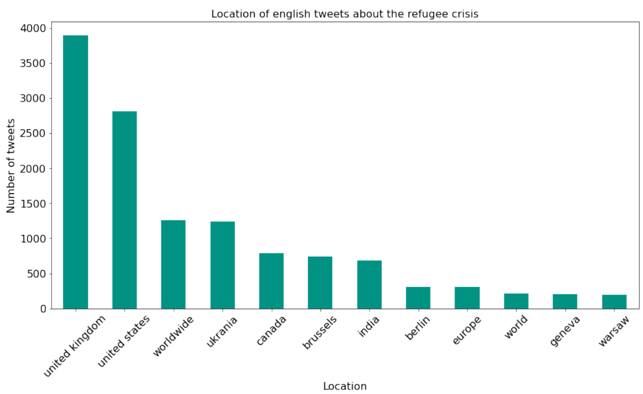


Figure 2. Location distribution for English tweets about the refugee crisis

3 Methods

In this project, we first conducted a qualitative analysis of the news and social media coverage of the Syrian refugee crisis and the Ukrainian refugee crisis, as well as a descriptive analysis to better comprehend our data. Furthermore, our analysis is comprised of two primary phases: sentiment analysis and topic modelling.

Before conducting the analysis, the text of tweets was tokenized, and contents that we deemed irrelevant to sentiment analysis were removed. We eliminated stopwords, punctuation, and special characters such as *# and @*. We also replaced in-text hyperlinks with the word *URL* and user names with *User*. Additionally, emojis were excluded from the analysis. Emojis have become integral to people's expression online and contain specific emotions. However, most sentiment analysis models cannot interpret emojis directly. And short names for emojis can not always be accurate because people generally have their own interpretations regarding using these emojis. Thus, directly translating emojis into short names as text may affect the model's accuracy. Deleting emojis will result in the loss of some information, but we believe can improve the sentiment model's accuracy accordingly.

We chose the state-of-art *twitter-roberta-base-sentiment* model [?] to get the sentiment result. This model is a pre-trained roBERTa model that was finetuned for sentiment analysis using the TweetEval benchmark and trained on 58 million tweets. For every tweet, the model will return three scores (negative, neutral and positive), representing both probability to be certain sentiment and the degree of this sentiment. We randomly selected tweets with different sentiment score to verify the if the model generalize well in our case. Based on the model result, we could conduct further analysis to understand the moods state relative to the refugee crisis and how they changed. During the analysis, the scores are treated as the degree of each sentiment, and the sentiment indexes, the positive scores minus the negative scores, are calculated to represent the sentiment of tweets. By doing so, each tweet is considered to have both positive and negative sentiments. However, the degree of the two sentiments varies, and the sentiment index will fluctuate more between 1 and -1 depending on the degree of the two sentiments. A very positive tweet will have a positive score close to 1 and a negative score close to 0, hence have a sentiment index around 1. A highly negative tweet will have the opposite, whereas a more neutral tweet will have a more balanced negative and positive score, resulting in a sentiment index around 0. It must be acknowledged that this disregards the impact of neutral scores, as the sentiment index is only derived from positive and negative scores. However, such an index can reflect more nuanced shifts in sentiment. The number of tweets expressing a particular sentiment on a given day is not directly proportional to the level of sentiment on that day. Simpson's paradox is an excellent illustration of the disparity between quantity and quality. As a result, we choose to represent the change in sentiment using the sentiment index, i.e., the probability of different emotions.

Topic extraction from tweets is done using BERTopic [?], which employs c-TF-IDF and transformers to create dense clusters that are easy to understand while maintaining essential words in the topic descriptions. The training of the BERTopic model can be viewed as unsupervised clustering, in which the topic number is not determined in advance; only the minimum topic size is required. In particular, we use the *all-mpnet-base-v2* sentence-transformers model to calculate the word embedding, which provides the highest quality among all existing sentence-transformers models. The minimum topic size is set to 100. The model parameters are detailed in Appendix A.2. Combining the topic extractions result with the sentiment index, we aim to infer how the mood is related to political events.

4 Result[1]

4.1 Words Frequency

The results, as shown in Figure 3 highlight the selection criteria as well as other features. Words related to the conflict

[1]The code is available in https://github.com/RuiaRui/project_CSM.
The interactive version of the figure in this section can be accessed through https://ruiarui.github.io/project_CSM/

itself such as *Ukraine/ukrainian, war/ukrainewar, Russia/Russian, Poland* occur frequently. This matches with on of the core themes (conflict) from Wallace [?].

Another set of frequent words is related to the core themes of Family and Services, as *child, home, need* but even more frequent are terms related to forced migration as *border, fleeing, migrant.* From the word frequency analysis, the existence of some sort of solidarity can be deduced with frequent unigrams as *help, support, standwithukraine.*

A subset of tweets containing at least one of the following keywords : *refugeeswelcome, care, shelter, host, community,* was subjected to a unigrams analysis in an attempt to highlight features relative to an recognition of Ukrainian as "like us" individuals. The results didn't show any unigrams that will indicates such phenomenon.

Together these results shows that meanwhile general topics can be obtained with word analysis (see section 5). The lack of context doesn't allow quantifying of sentiment associated with it. In addition, understanding the possible racist bias is impossible without more details. The sentiment analysis and topic modeling methods are used to pursue these goals.

Note that the acronym *URL* is one of the most frequent words and shows that a large proportion of tweets include a redirection towards another website or platform. However, in this study we didn't investigate further the URL contents.

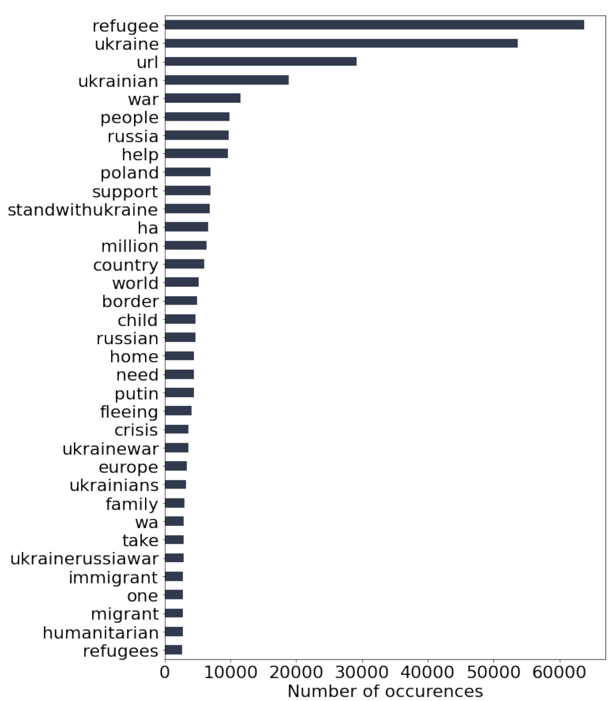


Figure 3. Most frequent unigrams in the English tweets.

4.2 Sentiment Analysis

According to the sentiment model's results, 44.27% of the statements were neutral, 38.97% were negative, and 16.74% were positive. Examining these results in greater detail reveals that the neutral sentiments are primarily news reports describing other countries' new policies or relevant actions. In contrast, the negative sentiments are people's sadness about the war and sympathy for refugees. The positive sentiments are people welcoming refugees in their communities or sharing their experiences helping refugees.

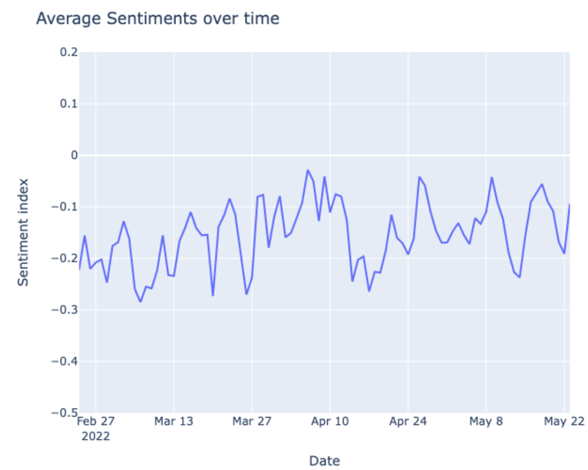


Figure 4. Average Sentiment Index Over Time

Figure 4 shows how average sentiment index changes over time. Over the 4 months period the sentiment index oscillates but stays below zero. The negative score in the tweets overcomes the positive score more often that the opposite. For the topic at hand, an armed conflict, an average negative index as shown can be expected. It indicates that topics with negative sentiments such as the war itself and sympathy for the refugees are either more present or more intense than the expression of positive sentiments.

With respect to the set research questions it is necessary to decipher of what importance of specific topics is in the average sentiment index presented above. Figure 5 shows sentiment indices for the 30 most frequent hashtags.

The aggregation by hashtags displays a few topics with a positive sentiment index (opposite to the average trend). The #standwithukraine has a positive average sentiment index as well as #refugeeswelcome. The average positive scores illustrates that the use of both hashtags is associated with a solidarity movement towards the refugees even if the other topics counterbalance the effect in the average analysis. The #Moldova has the highest positive index, and it is also linked to the welcoming of the refugees as Moldova has welcomed 450,000 refugees, which is the largest number of refugee welcomed in proportion to the size of the country [?]. Together, the three hashtags related to the welcoming of refugees and

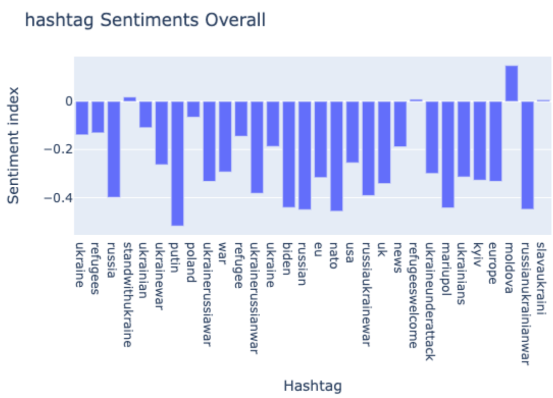


Figure 5. Average Sentiment Index By Hashtags

associated with positive sentiment index reflects a positive moods toward the welcoming in refugee. Meanwhile, this trend note that the #refugee/#refugees themselves have a sentiment below zero, which requires to investigate further the topics of the tweets.

4.3 Topic Modeling

The BERTopic model ultimately returned 27 topics, including one 'outline' topic (marked -1), which contained all tweets not categorized into any topic. The specific keywords and counts for the 27 topics are shown in the appendix A.2. The largest topic (topic 0) is related to UK and EU policies and government reactions. And topic 1, the second most important topic, is a condemnation of the different attitudes of society towards refugees between Europe and Africa.

The topic modeling reveals that racism at hand in the refugee crisis is a significant feature of our dataset, as it is specific enough and present enough to be a topic itself. While this is an explicit exhibition of the phenomenon investigated here, other topics are implicitly linked to a positive welcoming of refugees that are "like us". Topic 3 is related to donation and has a very high positive sentiment index, which indicates that people are encouraging, and enthusiast towards donating money or supplies to refugees. This feature isn't prominent in the tweets from the Syria refugee crisis.

Topic 8 is related to pets, mainly cats and dogs, and also shows a positive sentiment index. The attitude towards pets has individual and cultural influences. The cultural influences may include historical and religious beliefs. The presence of this topic point out to the believed shared practices of having cats and dogs as pets. This supposedly common cultural reference might be a feature that participates to the consideration Ukrainian as people with "middle class" attributes - such as pets - and also participates to make them look as individuals, and goes against the "massification" phenomena that Day pointed out for Syrian refugees. Topic 9, which is relative to religion and church with a positive sentiment index, might also fall into the category of shared practices, or represent an incentive to "pray for...", which still shows the integration of Ukrainians in the group of *people we care for*. Both topics are then proxies for racist discourse in which some features create similarities and other features differences (e.g. churches are similar and mosque different). Some limitations with respect to these hypothesis are discussed later on.

We also conducted the dynamic topic modeling to see how the topics evolve through time, and the change of the top 10 topics are revealed in Figure 7. The evolution of sentiments over time is necessary to confirm that the trends analyzed represents a constant trend over time. Figure 7 demonstrates that while there are oscillation the topics have distinct trends staying in the positive or respectively negative sides.

Combining the topics with the sentiment result, we can see the sentiment index for each topics (shown in figure 7).

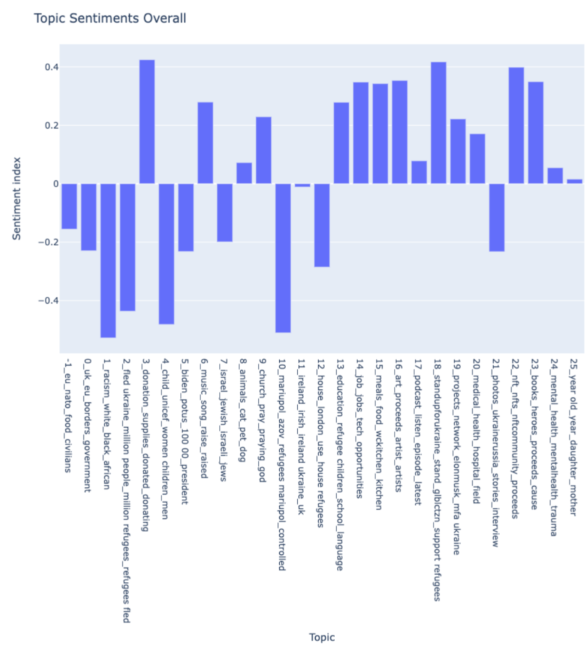


Figure 7. Average Sentiment Index For Different Topics

The topics with the most negative sentiment indexes are topic 1, 2, 4 and 10. Topic 1, as we already mentioned is concerned about racism, and topic 10 is about Mariupol, a city besieged and largely destroyed by the Russian invasion. These topics were indeed obvious candidates for tweets with negative sentiments. Topic 2, with words like "millions", "fled", is concerned about the massive number of refugees fleeing Ukraine; while topic 4 with "child", "woman", "men" is about refugees as people. It is interesting to note that even if those topics refer on the one hand to refugees as a mass and on the other hand to refugees as individuals – a distinction that we discussed above – both topics have a negative index. However, we must note that the negative index is a broad

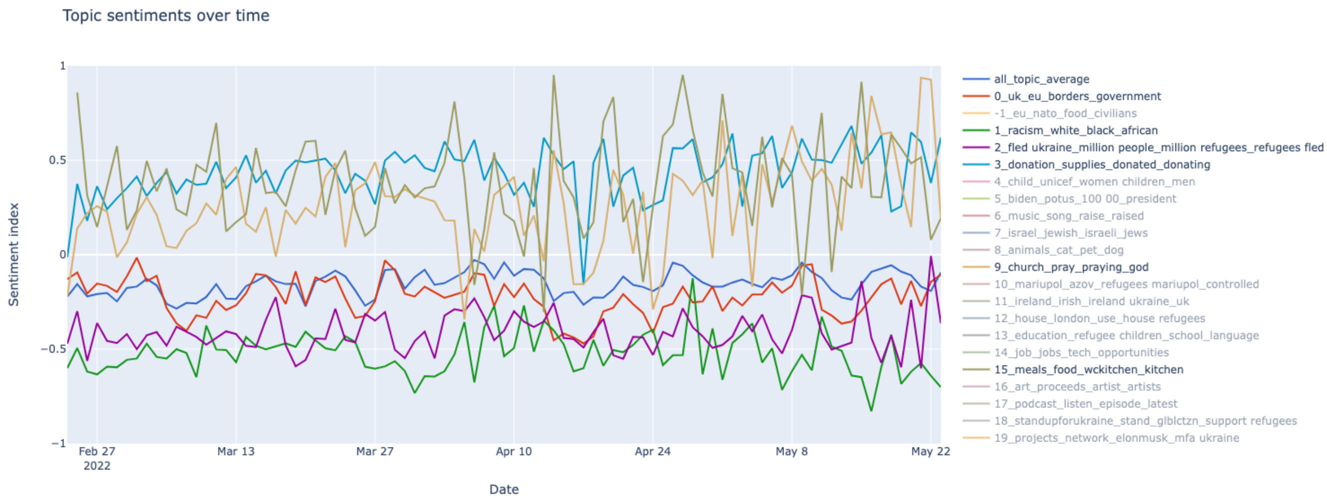


Figure 8. Topic change over time.
Only top 4 representative topics are shown in this figure, the full interactive version can be seen in https://ruiarui.github.io/project_CSM/

metric, and it cannot therefore distinguish between negative tweets about the sadness of the situation and angry tweets about the fact that refugees are coming to Europe. If making that distinction were possible, it would be interesting for future work to investigate the difference between topics 2 and 4 for that matter.

On the other hand, topics with the highest positive sentiment index are topics concerned with donations and support, like topics 3, 9 and 18, or services in general, like topics 13, 14, 15 and 20. These results are coherent and therefore support the accuracy of our sentiment analysis.

Moreover, the timelines for each topic are plot in figure 8. We can see that some topics fluctuate more than others. Generally the topics with positive sentiment indexes seem to fluctuate more. That could be due to the fact that all the tweets are related to the war, a highly negative event, and therefore even tweets about the positive aspects of the tragedy can easily "turn" to be negative.

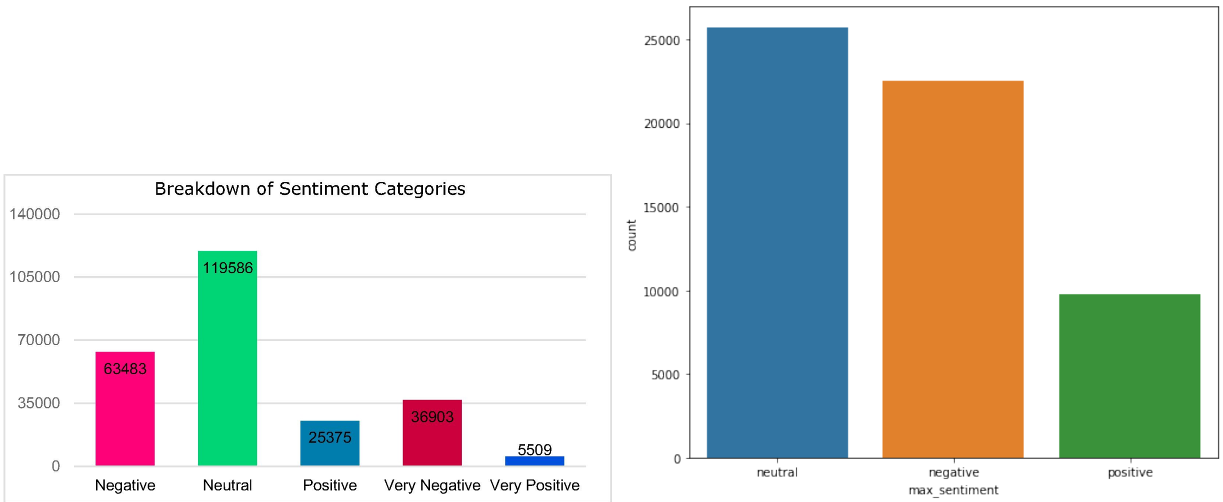
5 Comparison

For comparison we use both the results presented above and results aggregated in a similar manner as Öztürk[**?**].

The overall sentiment comparison, Figure 9, illustrates that in both cases there are more negative sentiments than positive. The reasons for such overall results are linked to the topic as a whole, which in both cases is a deadly conflict, as mentioned above. However, the proportion of positive tweets has increased in the Ukrainian refugee crisis, with more people showing welcome, gratitude or attempts to help.

The sentiment evolution, Figure 10, displays the variation of tweets counts (similar to the time distribution) as well as the variation inside of the positive, negative and neutral tweets. Compared to the Syrian refugee crisis, the number of positive tweets for the Ukrainian refugee crisis increased throughout the timeline, rather than being concentrated at a few points in time. The occurrence of its peak also coincides more closely with the overall change. This suggests that English tweets contain an overall more positive sentiment index than the Syrian refugee crisis.

A direct comparison can be made between the most frequent words in both cases at hand. For the Syrian Refugee Crisis, in English tweets, the frequent words include : *Obama, Trump, Putin, regime, illegal, president, conflict,crisis, policy, immigrant, government.* In this case, the tweets seemed to focus more on the politics and legal side of the refugee crisis and slightly less on the details about the war in Syria. In the case of the Ukrainian Refugee Crisis the frequent words become *people, help, poland, support, standwithukraine, border, child, home, need, putin.* There is in this case more words related to the solidarity or humanitarian category. In the comparison study [**?**] tweets have been classify into three categories: Politics, War and Humanitarian. The results showed, that in the case of English language, 47% of tweets were in the Politics category, 34% were clustered into War category and 19% were considered as Humanitarian. We manually used a similar method to classify topic, after removing outlier (topic -1), the percentage become: 65.35% about politics, 27.46% about Humanity and 7.17% about War in Ukraine crisis. While the politics is always the most frequent topic, even more in the case of the Ukraine-Russia conflict the War and Humanitarian categories have different proportions and Humanitarian topics have a higher proportion in the Ukrainian crisis tweets.



(a) Overall sentiment for the Syrian Refugee Crisis [?]

(b) Overall sentiment for the Ukraine Refugee Crisis

Figure 9. Overall Sentiment Result Comparison between Syrian Refugee Crisis[?] and Ukraine Refugee Crisis

The difference in topics highlights the different perspective on these conflicts. The rise of humanitarian concern is, as shown earlier, explained by the surge of topic were similarity and compassion are shared. For example the topic of donation, with positive sentiment, is part of the humanitarian category for the dataset of this study but is absent from the terms selected for the humanitarian category in the Syrian Refugee Crisis. In a similar way, *support* or *stand with* keywords are also absent. Such differences reveal that while, as presented in this study, English tweets related to the refugee crisis in the Ukraine-Russia conflict focus a lot on solidarity towards similar individuals, these aspects are missing in English tweets related to the Syrian Refugee Crisis.

6 Limitations and Future Work

In the dataset, initial filtering with a limited set of three English keywords was used to select tweets of interest in the larger dataset about the Ukraine-Russia conflict. To enlarge the keywords set, a loop mechanism using the most frequent words of the selected tweets could have been tested.

The URLs have been systematically replaced with an acronym. Due to the prominent *donation* topic further exploration of the URLs could allow to quantify more precisely the characteristic of this topic or another if many links revealed other topics.

For further historical considerations the full available period should be investigated as sentiments might change after the dates investigated here.

For locations, if a deeper geographical analysis is run, the cleaning process should include a check-speller as well as more translations than only the original language and a list of possible acronyms.

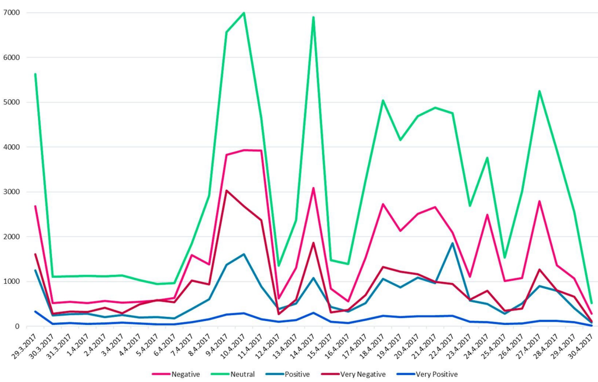
Moreover, the viewpoint taken in this study in an "English speaking" viewpoint. However, as 27% of locations are unknown and left locations include United Kingdom, USA and India, it is difficult to define who is included in the "us". First it would be an abuse to define this as a Western viewpoint uniquely. Secondly, if a breakdown by country was included the definition of Western must be investigated. A similar breakdown should be done for the comparison data.

In addition, different proxies have been used to highlight the racism that lies behind the tweets sentiments. For example, we used pets and religion topics as indicators or racism. However, in order to define more precisely these proxies, precise definitions of the racist discourse features should be defined. Indeed, in our case the analysis mentions *believed shared practices* and *supposedly* common cultural references. We can't assert, without verifying the statistics on the topic, if these practices are shared or not, do Syrian really have less dogs as pets as Ukrainian, or it is a belief. By defining the characteristics of the racist or biased discourse we could link them directly to our topics.

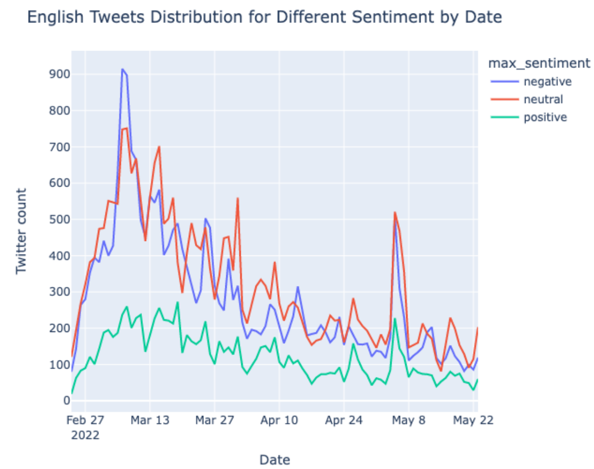
Finally, the exclusive selection of English tweets limits the interpretations of the situation. For example, Germany plays a more important role in welcoming refugees as it is closer to Ukraine than United Kingdom or USA, but is absent from our analysis. If we wanted to define a "Western" viewpoint additional languages to cover neighboring countries should be included.

7 Conclusion

The study of English tweets related the ongoing conflict shows a increase in the ratio of "positive" tweets with respect to the Syrian Refugee Crisis. The results support the hypothesis that there is a difference both in media treatment as well



(a) Tweets count for three sentiment catalog change over time for for the Syrian Refugee Crisis [?]



(b) Tweets count for three sentiment catalog change over time for for the Ukraine Refugee Crisis

Figure 10. Tweets count for three sentiment catalog change over time between Syrian Refugee Crisis [?] and Ukraine Refugee Crisis

as in opinions displayed on Twitter. The word-frequency analysis highlights different most common words pointing towards solidarity actions absent from the Syrian tweets analysis. The topic analysis further support this argument by showing several topics linked to solidarity. The topic analysis gives enough insights to reveal some proxies for feeling close to people that are white and European.

A Appendix

A.1 Timeline of Political Events

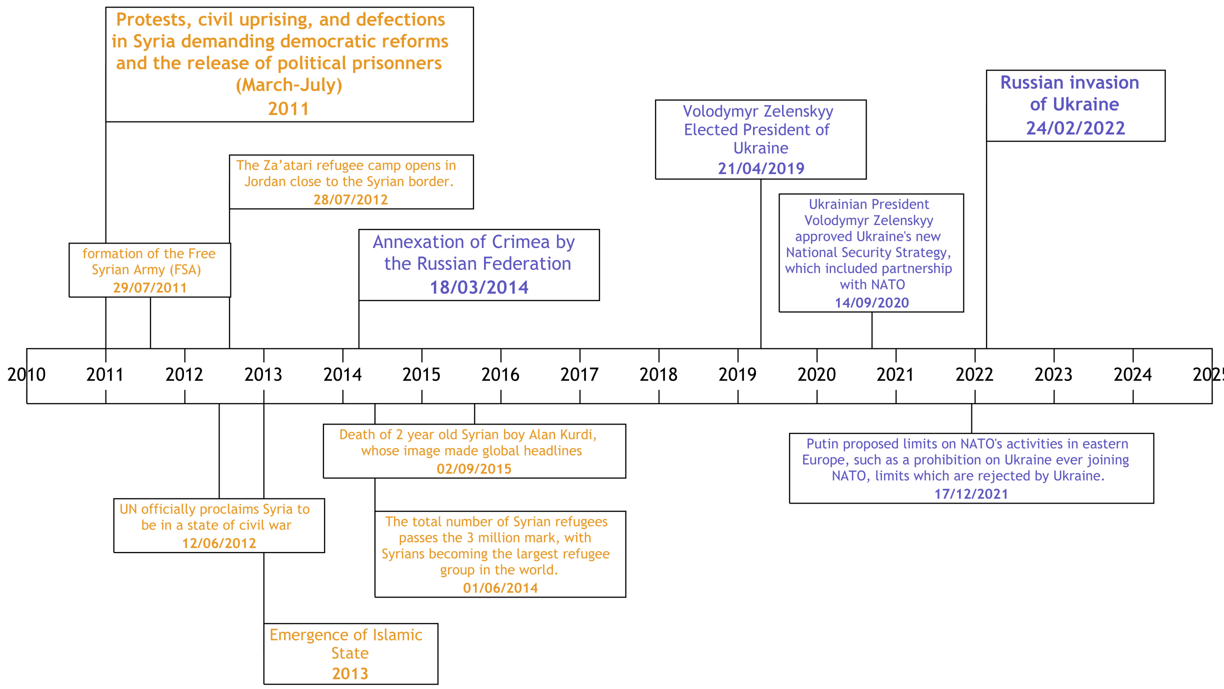


Figure 11. Timeline of events related to the Syrian conflict and the Ukrainian war

A.2 Topic Modeling - BERTopic

We customize the word embedding model, *HDBSCAN* model and the *CountVectorizer* to the BERTopic model. And the implementation details are as followed:

```
hdbscan_model = HDBSCAN(min_cluster_size=200,
                        metric='euclidean',
                        cluster_selection_method='eom',
                        prediction_data=True,
                        min_samples=5)

vectorizer_model = CountVectorizer(ngram_range=(1, 2),
                                   stop_words="english",
                                   min_df=10,
                                   max_df=0.95)

topic_model = BERTopic(verbose=True,
                       embedding_model='all-mpnet-base-v2',
                       vectorizer_model=vectorizer_model,
                       hdbscan_model=hdbscan_model,
                       calculate_probabilities=False,
                       min_topic_size=100,
                       nr_topics="auto")
```

Figure 12. BERTopic Implementation

And the 27 topics with their names and counts are shown in table 1.

Topic	Count	Name
0	24052	0_uk_eu_borders_government
-1	20924	-1_eu_nato_food_civilians
1	2536	1_racism_white_black_african
2	2495	2_fled ukraine_million people_million refugees_refugees fled
3	1602	3_donation_supplies_donated_donating
4	1450	4_child_unicef_women children_men
5	1366	5_biden_potus_100 00_president
6	835	6_music_song_raise_raised
7	824	7_israel_jewish_israeli_jews
8	813	8_animals_cat_pet_dog
9	712	9_church_pray_praying_god
10	695	10_mariupol_azov_refugees mariupol_controlled
11	607	11_ireland_irish_ireland ukraine_uk
12	558	12_house_london_use_house refugees
13	536	13_education_refugee children_school_language
14	487	14_job_jobs_tech_opportunities
15	476	15_meals_food_wckitchen_kitchen
16	466	16_art_proceeds_artist_artists
17	445	17_podcast_listen_episode_latest
18	441	18_standupforukraine_stand_glblctzn_support refugees
19	437	19_projects_network_elonmusk_mfa ukraine
20	420	20_medical_health_hospital_field
21	409	21_photos_ukrainerussia_stories_interview
22	349	22_nft_nfts_nftcommunity_proceeds
23	230	23_books_heroes_proceeds_cause
24	223	24_mental_health_mentalhealth_trauma
25	201	25_year old_year_daughter_mother

Table 1. Full Topic Extraction Result