A Longitudinal Study of Italian and French Reddit Conversations Around the Russian Invasion of Ukraine

Francesco Corso
Politecnico di Milano
Milan, Italy
francesco.corso@polimi.it

Giuseppe Russo
Zurich ETH
Zurich, Switzerland
russog@ethz.ch.

Francesco Pierri
Politecnico di Milano
Milan, Italy
francesco.pierri@polimi.it

ABSTRACT

Global events like wars and pandemics can intensify online discussions, fostering information sharing and connection among individuals. However, the divisive nature of such events may lead to polarization within online communities, shaping the dynamics of online interactions. Our study delves into the conversations within the largest Italian and French Reddit communities, specifically examining how the Russian invasion of Ukraine affected online interactions. We use a dataset with over 3 million posts (i.e., comments and submissions) to (1) describe the patterns of moderation activity and (2) characterize war-related discussions in the subreddits. We found changes in moderators' behavior, who became more active during the first month of the war. Moreover, we identified a connection between the daily sentiment of comments and the prevalence of war-related discussions. These discussions were not only more negative and toxic compared to non-war-related ones but also did not involve a specific demographic group. Our research reveals that there is no tendency for users with similar characteristics to interact more. Overall, our study reveals how the war in Ukraine had a negative influence on daily conversations in the analyzed communities. This sheds light on how users responded to this significant event, providing insights into the dynamics of online discussions during events of global relevance.

CCS CONCEPTS

• **Information systems → Social networks**.

KEYWORDS

Reddit, sentiment analysis, Russian invasion of Ukraine, online moderation

1 INTRODUCTION

On February 24th, 2022, the Russian army invaded Ukraine, after years of tension and conflict that began with the annexation of Crimea in 2014. Since then, the war in Ukraine has caused over 10,000 casualties [1] and forced over six million Ukrainians to abandon their country. This action drew widespread condemnation from the international community, as many countries viewed it as an infringement on Ukraine's sovereignty and territorial integrity [2]. Concurrently with the invasion, Russia became actively involved in spreading propaganda and misinformation regarding the conflict, aiming to manipulate public sentiment and erode support for Ukraine [10]. Such online information disorders and coordinated harm on social platforms are particularly significant during times of crisis when access to accurate and reliable information is essential [8, 22].

Due to the gravity of the conflict and the potential impact on regional stability, European countries, such as Italy and France, collectively pledged diplomatic and military support to Ukraine. This united front aimed not only to condemn the breach of Ukraine's sovereignty but also to deter an extension of the conflict to other European countries. This decision sparked a heated debate about the stance European countries should take regarding the Ukrainian war, quickly spreading to social media platforms like Twitter and Reddit, where most users took sides on either supporting military intervention or advocating for more diplomatic solutions [25]. Therefore, examining the various stances taken by users in online debates can enhance our understanding of public opinion on this controversial issue.

Present Work. We carry out a longitudinal study of conversations around the war taking place on Reddit, which is a discussion platform that allows users to engage in discussions and share content [3] in various communities known as "subreddits." Specifically, we analyze the subreddits r/italy and r/france as these two countries are amongst those states that more strongly support the Ukrainian defense against the Russian invasion[4]. In both countries, the debate around the support of Ukraine against Russia was heated, reflecting the broader European sentiment towards the conflict. Our study analyzes submissions and comments in these subreddits to understand how the war affected users' behavior in these two subreddits. We articulate our contributions into three research questions.

---

[1] https://news.un.org/en/story/2023/11/1143852 - Accessed on 29/11/2023
[2] https://www.consilium.europa.eu/en/policies/eu-response-ukraine-invasion/#invasion - Accessed on 6/12/2023
[3] https://thebrandhopper.com/2023/03/16/the-rise-of-reddit-how-the-platform-became-cultural-phenomenon/ - Accessed on 23/08/2023
[4] https://www.aljazeera.com/news/2023/2/16/mapping-where-every-country-stands-on-the-russia-ukraine-war - Accessed on 6/12/2023

- **RQ1**: How does moderators' activity change after the start of the invasion?
- **RQ2**: How does the invasion of Ukraine affect interactions between Reddit users?
- **RQ3**: Do we find evidence of homophily among Reddit users discussing the invasion?

To answer these questions, we collected 1.1M comments and 2.2M comments from `r/italy` and `r/france`, respectively. First (**RQ1**), to understand moderators' activity after the start of the war, we investigate the prevalence of moderated comments during 2022. We find a positive correlation in `r/italy` between the number of comments discussing the war and those removed, with a peak of over 10% of moderated comments on the same day of the invasion, while in `r/france` this result does not apply, since there is not a significant correlation.

Second (**RQ2**), we investigate the interactions between users by analyzing the text of comments posted in each subreddit. We employ state-of-the-art multilingual BERT-based models to infer the sentiment and toxicity of comments, finding that those discussing the war are significantly more negative and more toxic than others. Moreover, we find an overall negative sentiment with a significant negative peak value at the beginning of the invasion.

Third (**RQ3**), drawing upon results from the previous research question, we build and analyze the network of interactions between users. We assign an average sentiment score to each user based on their comments. Additionally, we consider three social scores[30] (age, gender, partisanship) based on users' pre-invasion activities in other subreddits. We investigate relationships between users and their neighbors, finding a small correlation between user's sentiment and the average sentiment of their neighbors. Interestingly, we do not find evidence of homophilic interactions among specific demographic groups particularly engaged in discussions around the war.

Implications. Overall, our work provides new evidence about how the Russian invasion of Ukraine shaped the online debate in two European countries strongly involved in the support of the Ukrainian cause. We observed that users in these Reddit communities generally express a negative attitude towards the war. Notably, discussions about these events often exhibit a more toxic behavior. This suggests that a significant portion of users align their sentiments with the prevailing stance of their country regarding the Russian invasion of Ukraine. We also find that, in accordance with existing literature, Reddit interactions around the war did not lead to the formation of echo chambers.

2 RELATED WORK

There are a number of existing contributions that analyze the interplay between online conversations and the ongoing Russian invasion of Ukraine.

In March 2022, the Observatory on Social Media at Indiana University published a series of white papers [19, 20] in which they investigate the presence of suspicious activity on Twitter during the first weeks of the invasion, showing distinct spikes in the creation of new accounts on Twitter. Furthermore, the analysis reveals the existence of multiple coordinated clusters of accounts, encompassing a wide spectrum of behaviors including spam, promotional content,

and even hate speech. The Social Media Lab at Toronto Metropolitan University deployed an online platform, called Russia-Ukraine ConflictMisinfo Research Portal [29], aimed to be a resource for researchers interested in studying the magnitude and the nature of online misinformation related to the conflict. They offer two main resources: the Russia-Ukraine ConflictMisinfo Dashboard, which is a tool where it's possible to track misinformation articles and fact-checking statements; the Russia-Ukraine ConflictMisinfo Geo-Map, a map where debunked claims about the conflict are shown with their specific geographical location.

Hanley et al. [12] explored news coverage of the war in three different media ecosystems, Western, Russian and Chinese, by employing topic modeling and differential sentiment analysis. Their work finds important differences between the Western and the Chinese/Russian ecosystems, with the last two having a tighter reciprocal influence. In another work [13], the same authors apply sentence-level topic analysis (MPNet model) to articles published by ten pro-Russian propaganda outlets, to understand the prominence of Russian state media narratives on selected subreddits. They show that in r/Russia, almost 40% of comments correspond to pro-Russian narratives, while in r/politics it is around 9%.

Another line of research [1, 3, 10, 28] study the diffusion of Russian propaganda on social media and the role of social bots in the campaigns of information warfare.

Pierri et al. [23] investigate the dynamics of account creation and suspension on Twitter during the Russian invasion of Ukraine and the 2022 French Presidential elections. The same authors carry out in [24] a longitudinal study of the diffusion of misinformation about the Russian invasion of Ukraine which originated from the low-credibility sources posted on Facebook and Twitter and from Russian state outlets, during the first months of the conflict.

Other research focuses on studying users' reactions to these events, such as the work by Caprolu et al. [4], in which the authors analyze over 5 million tweets collected during the month before and after the beginning of the war highlighting abnormal patterns in users' sentiment. Guerra and Karakus [11] study how to measure fear and hope in war-related discussion on Reddit in 6 different European countries. They observed how hope spiked positively in case of important victories of the Ukrainian resistance and non-military related events, while it spiked negatively in the correspondence of relevant strategic losses. Lastly, we report several datasets published [9, 15, 21, 26, 32] over the last years which allow researchers and practitioners to study online conversations around the war on a variety of platforms such as Reddit, Twitter, Weibo, VKontakte, and Telegram.

3 METHODOLOGY

3.1 Data Collection

To conduct our analysis, we collected all activity in the two subreddits during the period 01/02/2022 to 31/12/2022, leveraging the Pushshift dataset [2]. The resulting data contains over 3M comments, in the local languages, and 80k submissions across the two subreddits. In Figures 1 and 2 we show the temporal trends of activity of `r/france` and `r/italy` respectively. We see different volumes plotted in these figures, due to the difference in the user base of the communities (`r/france`: 600k, `r/italy` : 370k as of
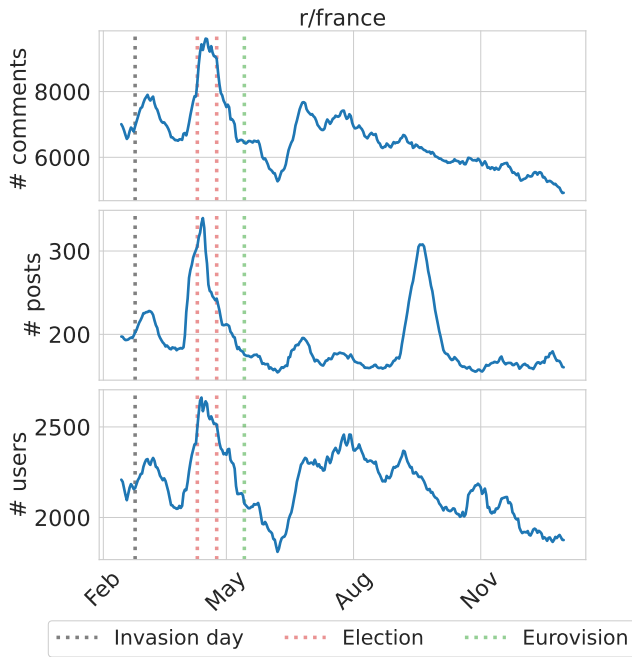
Figure 1: Time series showing the 7-day moving average of the daily number of comments, submissions and active users in `r/france`.
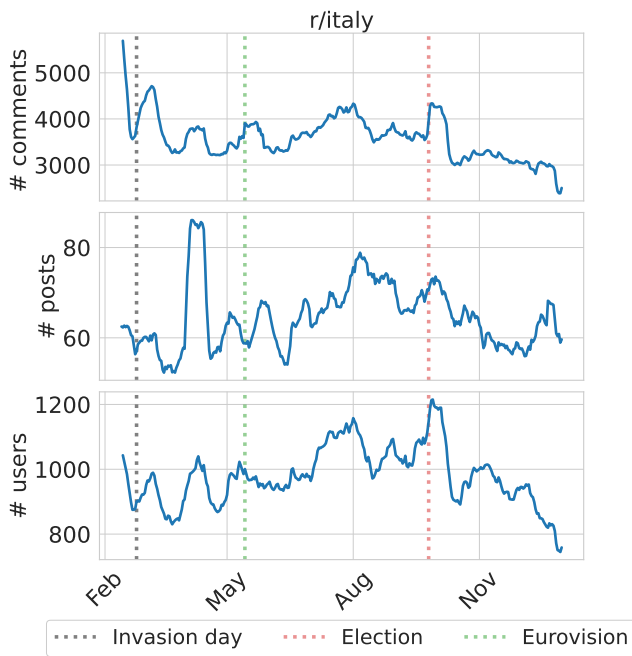

Figure 2: Time series showing the 7-day moving average of the daily number of comments, submissions and active users in `r/italy`.

March 2022). The number of daily comments for `r/france` ranges between 5k and 9k, while in `r/italy` it is between 3k and 5k. Similarly for daily submissions and active users, with `r/france` having an average of around 200 daily submissions and over 2k daily users, while `r/italy` shows an average of just over 60 daily submissions and 1k daily users.

In `r/france` we can observe the presence of a peak of daily comments, submissions and active users around the election period; this behavior is also found, with minor intensity, in `r/italy`. This pattern is present in both communities around the beginning of the invasion, even though is smaller compared to the one for the elections. To identify war-related comments, we applied a TF-IDF-based snowball sampling approach with keyword matching and manual inspection. This offers a valuable alternative to traditional probability sampling methods and it is widely employed in the literature [6, 7, 27]. We inspected a random sample,one for each subreddit, of 5.000 comments posted during the first month of the war that matched any of the following seed keywords. In Italian: 'Russia', 'Ucraina', 'Zelensky', 'Putin','guerra'; in French: 'Russie','Ukraine','Putin','Zelensky','guerre'. This way we could see what other meaningful words were used with known war-related keywords.

We then applied the previously defined sampling technique to obtain a final set of 20 relevant keywords, manually improved with case-sensitive alternatives to a total of 40, to be used to filter the body of all the comments in our dataset.

We applied the same methodology both to the title and the body of the submissions, to identify those that were related to the Russian invasion of Ukraine and extract all the comments of these submissions. This process yielded over 300k war-related comments and 6.9k submissions between the two subreddits.

3.2 Tracking moderators activity

Reddit moderators are volunteer users who are tasked with the enactment of the subreddit rules and, more in general, to keep the community in order.So, moderators can remove comments that they deem to be against the rules of the community [5]. In our dataset, removed comments are registered as a record with the body field equal to [removed] and the author field equal to [deleted], which implies that the content of removed comments cannot be inspected. Since we cannot discern if a removed comment is war-related, we consider as war-related those that are submitted under a war-related post. We identified a total of 26k removed comments in `r/italy` and 40k removed comments in `r/france`.

3.3 Sentiment and toxicity

To assign a score of sentiment to Reddit comments, we employed the last available version, as of November 2023, of a multilingual BERT-based model by Lik Xun Yuan [16] for both Italian and French languages. To analyze toxicity, we relied on Detoxify [14], a BERT-based model that proved to be performing in both languages. We leveraged the `Transformers` library by Huggingface [31] to deploy both models and to split the text of comments in our dataset into phrases we relied on the NLTK python library, obtaining a total of

---

[5]https://support.reddithelp.com/hc/en-us/articles/204533859-What-s-a-moderator - Accessed on 6/12/2023

over 5.5M sentences. For this analysis, we did not consider deleted and removed comments, since their textual content is not available. We relied on the previously mentioned models and tools because of their state-of-the-art performance, popularity in the research community, and wide availability.

The sentiment analysis model outputs two numbers which indicate the probabilities for the text being negative and positive. To assign a continuous score, we consider the positive probability minus the negative probability, obtaining values in the range (-1, +1). We notice that we are not interested in the absolute value of the sentiment, but rather comparing the sentiment of online conversations across different periods and topics. Toxicity scores are instead provided by Detoxify on a continuous scale between 0 and 1. If a comment is composed of multiple sentences, we take the average of sentiment and toxicity values.

As a reference for the sentiment and toxicity of online conversations, we consider three other important events that happened during the period of our analysis: Eurovision 2022 (10-14 May), the 2022 Italian political election (25 September), and the 2022 French political Election (10-24 April).

3.4 Users' sentiment and socio-demographic scores

Reddit is fundamentally different from social media platforms such as X (formerly Twitter), in which submissions naturally lead to interactions between users. To build a network of interactions between Reddit users, we follow previous literature [5]: there is an interaction between User A and User B only if A has commented under a comment of User B. The resulting network is a directed graph where users are represented by nodes, and edges represent the action of commenting under another user's comment. We enrich our network by adding several node attributes: Sentiment, Age, Gender, and Partisanship.

Sentiment is computed as the average of sentiment scores of all comments posted by the user during the period of analysis. Age, Gender and Partisanship are based on the scores by Waller and Anderson [30], and are obtained by analyzing the pre-invasion activity of users following an approach similar to Monti et al. [17]. Specifically, for user i we define the social score X_i as:

$$X_i = \frac{\sum_j^n v_{ij} X_j}{V_i} \tag{1}$$

which corresponds to the weighted average of the score X_j of all subreddits j in which user i has posted a comment during the period November 2021 to January 2022, weighted by the number of comments v_{ij} posted by user i, normalized by the total number of comments V_i posted by user i during the same period.

Scores can be interpreted as the tendency of a user to visit subreddits that are usually populated by users with a given particular leaning. For example, r/teenagers has a younger population than r/RedditForGrownups. Another example is r/liberal versus r/conservative: the former will have a more left-leaning population than the latter. Scores are indicated as polarities: For Age, a negative value means a young-leaning score while a positive indicates an old-leaning score. For Gender, a negative value indicates a male-leaning score while a positive indicates a female-leaning score. For political leaning, a negative value indicates more of a liberal-leaning score, while a positive value indicates a more conservative-leaning score.

4 RESULTS

4.1 Conversations around the war

We first analyze the prevalence of Reddit comments related to the war. We observe similar patterns in the two communities: around the beginning of the invasion (grey dotted line) there is a peak of war-related comments, reaching over 20% and 40% of the daily comments posted in r/france and r/italy, respectively. These values reduce to an average of 10% for r/france and around 5% for r/italy for the following months.

4.2 Patterns of content moderation

To answer our **(RQ1)**, we analyze the activity of moderators in both subreddits during the period of analysis. We first observe a similar daily prevalence of removed comments in the two communities, as shown in the middle panel of Figures 3 and 4, with values in the range 2-4%. However, we report different temporal patterns: In r/france we find a peak of removed comments in April, a month after the invasion, in correspondence with the 2022 Election. In r/italy, there is a peak of over 10% daily removed comments (not shown in the Figure, which provides a 7-day moving average) in correspondence of a peak of war-related comments, i.e., the invasion day. Consequently, we do not find a significant Pearson correlation between the prevalence of war-related comments and the daily percentage of removed comments in r/france, while we find a positive significant Pearson correlation (R=0.28, p<0.001) in r/Italy.

4.3 Sentiment and toxicity of war-related conversations

To answer our **(RQ2)**, we first compute the daily sentiment of conversations in the two subreddits by averaging the sentiment of all comments, for each day. The bottom panels of Figures 3 and 4 show that there is a significant negative correlation with the daily prevalence of war-related comments in both r/france and r/italy, with Pearson correlation coefficients respectively R=-0.50 and R=-0.38 (P < 0.001). In particular, we notice drops in the daily sentiment in correspondence with peaks of war-related comments. We also see the same behavior in r/france during the Election held in April. Lastly, we observe an increase in the daily sentiment towards the end of the year in both subreddits, likely due to conversations around Christmas and New Year's Eve.

To investigate whether different topics of conversation exhibit a different sentiment, we compare the daily values for four different cases, as described in the methods: war-related comments, election-related comments, comments about the Eurovision, and all other comments. We observe a general pattern which is common to the two subreddits: war-related comments are significantly more negative than all other comments. Besides, we observe that comments related to the Eurovision festival are much more positive than other conversations, and those related to the elections are more positive than war-related ones but less than Eurovision. All distributions
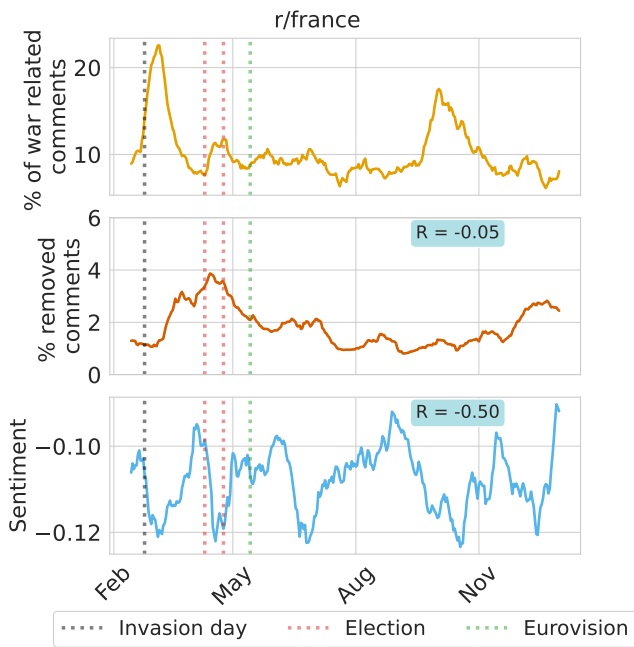
Figure 3: Daily proportion of war-related comments (Top) and removed comments (Middle), and daily average sentiment (Bottom) for `r/france`. We show a 7-day moving average. We report the Pearson correlation (computed on the daily observations and not the moving average) between each time series and the proportion of war-related comments in the legend of each panel.
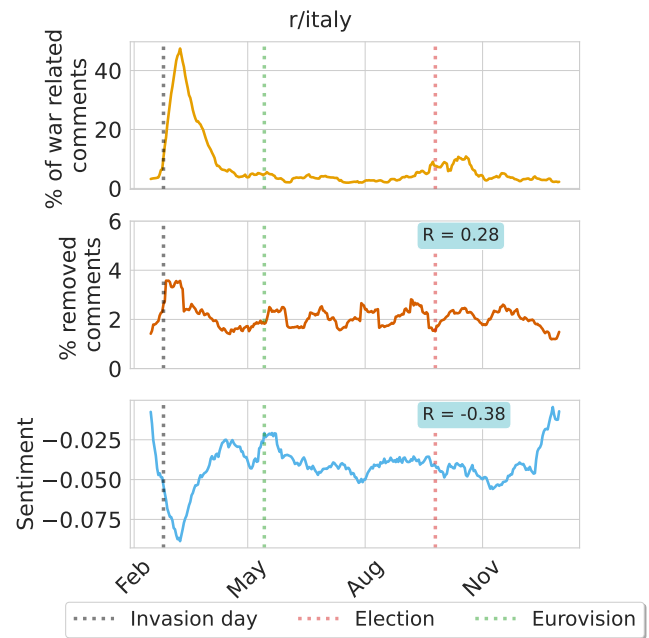

Figure 4: Daily proportion of war-related comments (Top) and removed comments (Middle), and daily average sentiment (Bottom) for `r/italy`. We show a 7-day moving average. We report the Pearson correlation (computed on the daily observations and not the moving average) between each time series and the proportion of war-related comments in the legend of each panel.

are statistically different according to a two-way Mann-Whitney test (P < 0.001).

We replicate this analysis for the toxicity, finding no correlation and no interesting patterns for the daily average toxicity in the two subreddits (we omit the figure for the sake of brevity). Interestingly, we find discrepancies across topics indicating that toxicity is correlated with negative sentiment, as shown in Figure 6. Indeed, war-related comments are significantly more toxic than any other conversation in both subreddits. Comments related to Eurovision are the least toxic ones, with those about the elections being slightly less toxic than the average comment (cf. Other), although with a high variance in both subreddits which indicates the presence of outlier comments with an extreme toxicity. All distributions are significantly different according to a two-way Mann-Whitney test (P < 0.001), except for the case Eurovision versus Elections in `r/france` (P = 0.36).

4.4 Homophily in war-related conversations

To answer our **(RQ3)**, we first analyze the network of interactions between users in the two subreddits, as described in the methodology.

Following Cinelli et al. [5], we test the presence of homophily between users according to the four user attributes: `Sentiment`, `Age`, `Gender`, `Partisanship`. Specifically, we study the joint distribution of users' attributes and the average value of their neighbors, as

derived from the network of interactions. As shown in Figures 7 and 8, we do observe a weak significant correlation for what concerns the sentiment of users; this correlation is still significant for the socio-demographic scores but the coefficient is almost 0 (except for `Gender` in `r/Italy`, where P = 0.78). These results extend previous findings on the absence of political homophily and echo chambers in Reddit [5], as we show the same pattern for `Age` and `Gender`.

We further investigate whether users more active on the topic of war exhibit different attributes. In Figures 9 and 10 we show the distribution of all attributes for three different groups of users: those that never commented on the war, those with at least one and less than 100 war-related comments, and users with more than 100 war-related comments. We find a common pattern in the two subreddits: the `Age` median shifts from negative values to more positive ones, while the `Gender` median behaves in the opposite way, shifting from a more positive (female-leaning) to a more male-leaning one. `Partisanship` shifts from predominantly conservative-leaning scores to more central ones in both subreddits. Overall we can observe that users more engaged in war-related discussions show to be more male, older and center-oriented than the users who do not comment about the war. Despite being significant (Kruskal-Wallis test, P<0.05), however, the differences in the socio-demographic scores are very small (cf. Table 1).

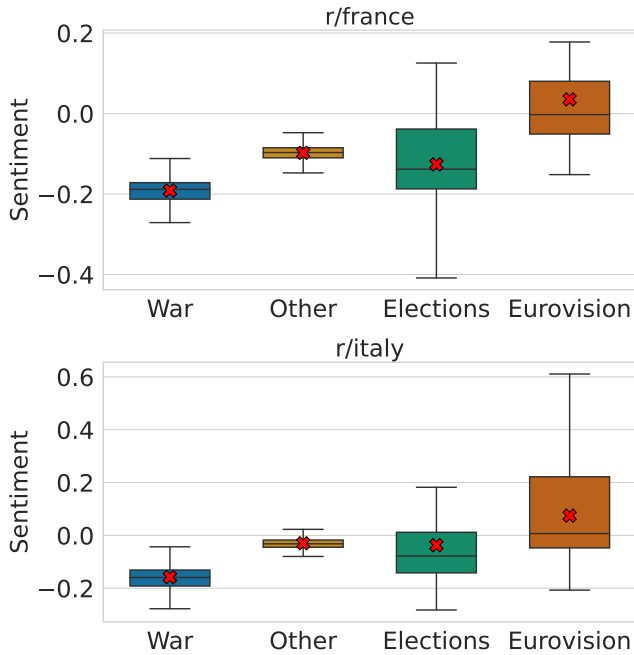Figure 5: Distributions of the daily average sentiment in `r/france` (Top) and `r/italy` (Bottom) computed for non-overlapping subsets of comments corresponding to specific topics. Outliers are not shown, and the red cross indicates the mean value of the distribution. Medians for `r/france` are: War=-0.19, Other=-0.1, Elections=-0.14, Eurovision=-0.001. Medians for `r/italy` are: War=-0.16, Other=-0.03, Elections=-0.08, Eurovision=0.006.
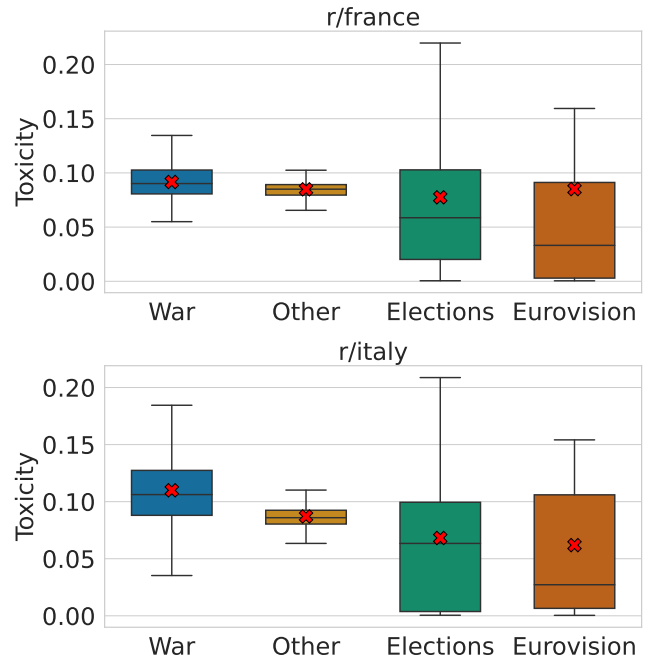

Figure 6: Distributions of the daily toxicity in `r/france` (Top) and `r/italy` (Bottom) computed for non-overlapping subsets of comments corresponding to specific topics. Outliers are not shown, and the red cross indicates the mean value of the distribution. Medians for `r/france` are: War=0.091, Other=0.8, Elections=0.05, Eurovision=0.033. Medians for `r/italy` are: War=0.11, Other=0.085, Elections=0.063, Eurovision=0.027.

		Activity	Sentiment	Age	Gender	Partisan
	0	-0.02	-0.016	0.02	0.019
r/france	[1,100)	-0.11	-0.002	0.01	0.017
	100+	-0.16	0.015	0.006	0.012
	0	0.004	-0.02	0.03	0.021
r/italy	[1,100)	-0.07	-0.007	0.01	0.019
	100+	-0.09	0.006	0.004	0.013

Table 1: Median values of the distribution of figures 9,10 divided by subreddit and class of war related activity. All the medians between for each score in each subreddit are statistically different according to the Kruskal-Wallis test (p<0.05)

5 DISCUSSION

5.1 Contributions

We conducted a longitudinal study, tracking eleven months of activity in the two most representative subreddits of Italian and French communities from the beginning of the Russian invasion of Ukraine to the end of 2022. We identified over 200k comments and 10k submissions that discussed the invasion of Ukraine in both subreddits, highlighting how the war changed the moderation behavior in

`r/italy`, while in `r/france` this effect was more visible during the Election.

We found a significant correlation between the prevalence of the war topic and the negativity of daily sentiment in comments shared by Reddit users in both communities; this difference was even more evident when compared to other significant events that happened during the time frame of our study (Eurovision 2022 and 2022 French and Italian Political Elections). We observed the same results when analyzing the toxicity of online conversations.

Finally, we found no evidence of political homophily on Reddit, as previously seen in the literature, extending this result to other socio-demographic scores and sentiment.

Our results prove that there are astonishing similarities between two representative Reddit communities of two of the most important European countries. Both highlighted a strong negativity towards the war, which is a shared sentiment with the alignment of the EU and respective countries. [2]

We also showed how the debate in these communities about this topic brought together a wide range of users with different demographic profiles. We also showed some differences, for example in the moderation activity, shedding light on how the moderators of these communities perceive the importance of discussed topics in the community they administrate.
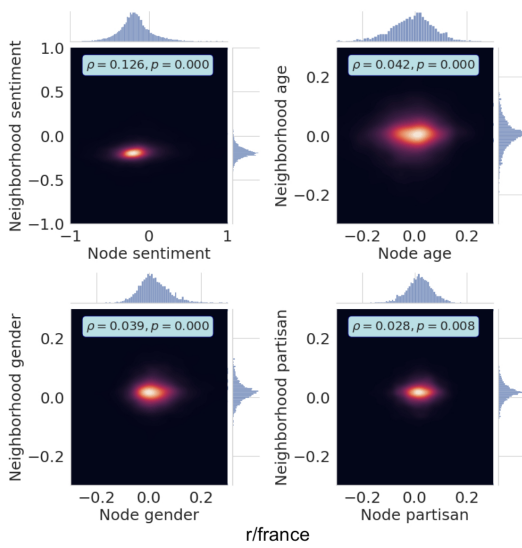
Figure 7: Joint distribution of the score of users (nodes) and the average score of their neighbors in `r/france`. `Sentiment` values show a weak significant correlation (Pearson R=0.126 P<0.001).
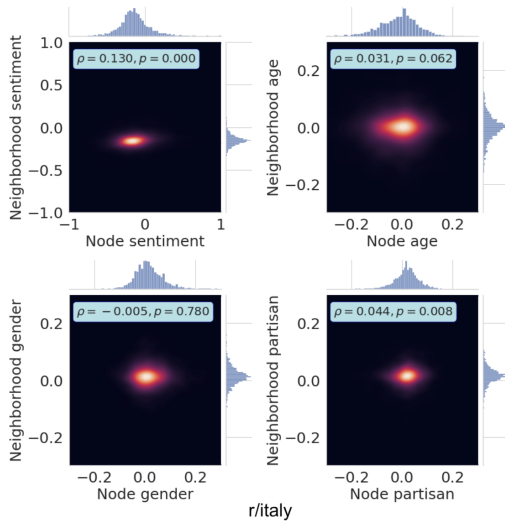

Figure 8: Joint distribution of the score of users (nodes) and the average score of their neighbors in `r/italy`. `Sentiment` values show a weak significant correlation (Pearson R=0.130,p<.001).

5.2 Limitations

There are a number of limitations to our study. First of all, even though we filtered out comments produced by known Reddit bots, such as u/RemindMeBot, u/savevideobot, we could not check for comments generated by inauthentic accounts.
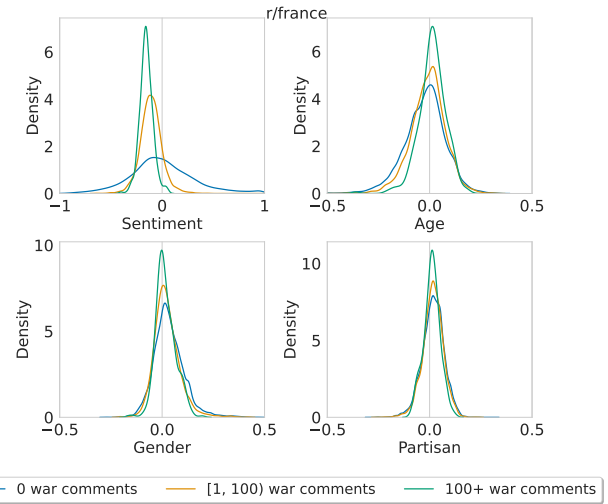

Figure 9: Distributions of `Sentiment`, `Age`, `Gender` and Partisanship scores for users of `r/france` for the three classes of war-related activity.
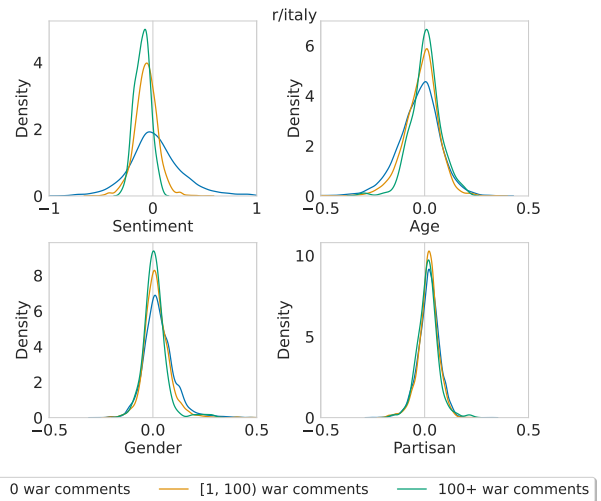

Figure 10: Distributions of `Sentiment`, `Age`, `Gender` and `Partisanship` scores for users of `r/italy` for the three classes of war-related activity.

Moreover, we used a keyword-matching-based method to identify comments related to the Russian invasion of Ukraine which tended to also signal which might entail false positives (i.e., generic comments about war). A manual validation on a random sample of 200 comments results in a small percentage (<10%).

We analyzed sentiment and toxicity with black-box deep learning models that might not be completely accurate [18]. We sampled 200 random sentences (100 for each subreddit) and manually validated the classifiers, finding a small percentage of inaccurate results (<10%).

Finally, only 56% of the users in our study were also active in the three months before February 2022 (necessary to obtain social scores) with the remaining 44% joining the subreddits in the following months. We did not notice any abnormal pattern around the way these new users joined the communities.

Furthermore, the scores we used for describing the political leaning of the users are referring more to a US-oriented political system than the Italian or French system.

5.3 Future Work

There are several opportunities for future work related to our research and this topic in general. It would be relevant to conduct the same analysis on other European subreddits, maybe of the countries geographically closer to Ukraine, to highlight eventual differences or points of similarity with the ones we analyzed in this study. Our results could also be compared with X (formerly Twitter) or Facebook, which are fundamentally different platforms and could highlight structural differences that could improve or worsen the discussion on these controversial topics.

5.4 Ethical Considerations

In this study, we gave utmost attention to the principles of privacy, transparency, and fairness. Data privacy concerns are addressed through rigorous anonymization and aggregation procedures, without any attempt of backtracking to the real user. We employed transparent methodologies, detailing data collection, preprocessing steps, and modeling choices, enabling reproducibility. The study is committed to ensuring fairness in the analysis and of results. We acknowledge the unavailability of the Reddit API, so we publish in an online repository the IDs of the comments that we employed in our dataset along with the Python notebooks we used to produce the results. The dataset can be reconstructed following the instructions provided in the description of the repository. [6]

REFERENCES

[1] Maxim Alyukov, Maria Kunilovskaya, and Andrei Semenov. 2023. Wartime Media Monitor (WarMM-2022): A Study of Information Manipulation on Russian Social Media during the Russia-Ukraine War. In *Proceedings of the 7th Joint SIGHUM Workshop on Computational Linguistics for Cultural Heritage, Social Sciences, Humanities and Literature.* 152–161.

[2] Jason Baumgartner, Savvas Zannettou, Brian Keegan, Megan Squire, and Jeremy Blackburn. 2020. The Pushshift Reddit Dataset. *International Conference on Web and Social Media* (2020). https://doi.org/10.5281/zenodo.3608135

[3] Kenneth J Boyte. 2017. An analysis of the social-media technology, tactics, and narratives used to control perception in the propaganda war over Ukraine. *Journal of Information Warfare* 16, 1 (2017), 88–111.

[4] Maurantonio Caprolu, Alireza Sadighian, and Roberto Di Pietro. 2023. Characterizing the 2022-Russo-Ukrainian Conflict Through the Lenses of Aspect-Based Sentiment Analysis: Dataset, Methodology, and Key Findings. In *2023 32nd International Conference on Computer Communications and Networks (ICCCN)*. IEEE, 1–10.

[5] Matteo Cinelli, Gianmarco De Francisci Morales, Alessandro Galeazzi, Walter Quattrociocchi, and Michele Starnini. 2021. The echo chamber effect on social media. *Proceedings of the National Academy of Sciences of the United States of America* (2021). https://doi.org/10.1073/pnas.2023301118

[6] Michael D Conover, Bruno Gonçalves, Alessandro Flammini, and Filippo Menczer. 2012. Partisan asymmetries in online political activity. *EPJ Data science* 1, 1 (2012), 1–19.

[7] Matthew R DeVerna, Francesco Pierri, Bao Tran Truong, John Bollenbacher, David Axelrod, Niklas Loynes, Christopher Torres-Lugo, Kai-Cheng Yang, Filippo Menczer, and John Bryden. 2021. CoVaxxy: A collection of English-language

[8] Twitter posts about COVID-19 vaccines. In *Proceedings of the International AAAI Conference on Web and Social Media*, Vol. 15. 992–999.

[8] Marco Di Giovanni, Francesco Pierri, Christopher Torres-Lugo, and Marco Brambilla. 2022. VaccinEU: COVID-19 vaccine conversations on Twitter in French, German and Italian. In *Proceedings of the International AAAI Conference on Web and Social Media*, Vol. 16. 1236–1244.

[9] Yi R Fung and Heng Ji. 2022. A weibo dataset for the 2022 russo-ukrainian crisis. *arXiv preprint arXiv:2203.05967* (2022).

[10] Dominique Geissler, Dominik Bär, Nicolas Pröllochs, and Stefan Feuerriegel. 2023. Russian propaganda on social media during the 2022 invasion of Ukraine. *EPJ Data Science* 12, 1 (2023), 35.

[11] Alessio Guerra and Oktay Karakus. 2023. Sentiment analysis for measuring hope and fear from Reddit posts during the 2022 Russo-Ukrainian conflict. *Frontiers in Artificial Intelligence* (2023). https://doi.org/10.3389/frai.2023.1163577

[12] Hans WA Hanley, Deepak Kumar, and Zakir Durumeric. 2023. " A Special Operation": A Quantitative Approach to Dissecting and Comparing Different Media Ecosystems' Coverage of the Russo-Ukrainian War. In *Proceedings of the International AAAI Conference on Web and Social Media*, Vol. 17. 339–350.

[13] Hans WA Hanley, Deepak Kumar, and Zakir Durumeric. 2023. Happenstance: Utilizing Semantic Search to Track Russian State Media Narratives about the Russo-Ukrainian War On Reddit. In *Proceedings of the international AAAI conference on web and social media*, Vol. 17. 327–338.

[14] Laura Hanu and Unitary team. 2020. Detoxify. Github. https://github.com/unitaryai/detoxify.

[15] Ehsan-Ul Haq, Gareth Tyson, Lik-Hang Lee, Tristan Braud, and Pan Hui. 2022. Twitter dataset for 2022 russo-ukrainian crisis. *arXiv preprint arXiv:2203.02955* (2022).

[16] Lik Xun Yuan. 2023. distilbert-base-multilingual-cased-sentiments-student (Revision 2e33845). https://doi.org/10.57967/hf/1422

[17] Corrado Monti, Jacopo D'Ignazi, Michele Starnini, and Gianmarco De Francisci Morales. 2023. Evidence of Demographic rather than Ideological Segregation in News Discussion on Reddit. *The Web Conference* (2023). https://doi.org/10.1145/3543507.3583468

[18] Gianluca Nogara, Francesco Pierri, Stefano Cresci, Luca Luceri, Petter Törnberg, and Silvia Giordano. 2023. Toxic Bias: Perspective API misreads German as more toxic. *arXiv preprint arXiv:2312.12651* (2023).

[19] OSM. 2022. *Analysis of Twitter accounts created around the invasion of Ukraine.* https://osome.iu.edu/research/white-papers/Ukraine_OSoMe_White_Paper_May_2022.pdf

[20] OSM. 2022. *Suspicious Twitter Activity around the Russian Invasion of Ukraine.* https://osome.iu.edu/research/white-papers/Ukraine_OSoMe_White_Paper_March_2022.pdf

[21] Chan Young Park, Julia Mendelsohn, Anjalie Field, and Yulia Tsvetkov. 2022. Challenges and Opportunities in Information Manipulation Detection: An Examination of Wartime Russian Media. *Conference on Empirical Methods in Natural Language Processing* (2022). https://doi.org/null

[22] Francesco Pierri. 2020. The diffusion of mainstream and disinformation news on Twitter: the case of Italy and France. In *Companion proceedings of the web conference 2020*. 617–622.

[23] Francesco Pierri, Luca Luceri, Emily Chen, and Emilio Ferrara. 2023. How does Twitter account moderation work? Dynamics of account creation and suspension on Twitter during major geopolitical events. *EPJ Data Science* 12, 1 (2023), 43.

[24] Francesco Pierri, Luca Luceri, Nikhil Jindal, and Emilio Ferrara. 2023. Propaganda and Misinformation on Facebook and Twitter during the Russian Invasion of Ukraine. In *Proceedings of the 15th ACM Web Science Conference 2023*. 65–74.

[25] Filippos Proedrou. 2010. Ukraine's foreign policy: accounting for Ukraine's indeterminate stance between Russia and the West. *Southeast European and Black Sea Studies* 10, 4 (2010), 443–456.

[26] Giuseppe Russo, Niklas Stoehr, and Manoel Horta Ribeiro. 2023. Acti at evalita 2023: Overview of the conspiracy theory identification task. *arXiv preprint arXiv:2307.06954* (2023).

[27] Giuseppe Russo, Luca Verginer, Manoel Horta Ribeiro, and Giona Casiraghi. 2023. Spillover of antisocial behavior from fringe platforms: The unintended consequences of community banning. In *Proceedings of the International AAAI Conference on Web and Social Media*, Vol. 17. 742–753.

[28] Fei Shen, Erkun Zhang, Hongzhong Zhang, Wujiong Ren, Quanxin Jia, and Yuan He. 2023. Examining the differences between human and bot social media accounts: A case study of the Russia-Ukraine War. *First Monday* (2023).

[29] Social Media Lab Toronto Metropolitan University. 2022. *Russia-Ukraine ConflictMisinfo Research Portal.* https://conflictmisinfo.org

[30] Isaac Waller and Ashton Anderson. 2021. Quantifying social organization and political polarization in online platforms. *Nature* (2021). https://doi.org/10.1038/s41586-021-04167-x

[31] Thomas Wolf, Lysandre Debut, Victor Sanh, Julien Chaumond, Clement Delangue, Anthony Moi, Pierric Cistac, Tim Rault, Remi Louf, Morgan Funtowicz, Joe Davison, Sam Shleifer, Patrick von Platen, Clara Ma, Yacine Jernite, Julien Plu, Canwen Xu, Teven Le Scao, Sylvain Gugger, Mariama Drame, Quentin Lhoest, and Alexander Rush. 2020. Transformers: State-of-the-Art Natural Language

---

[6]https://github.com/orsoFra/LS_FRIT_UKR

Processing. In *Proceedings of the 2020 Conference on Empirical Methods in Natural Language Processing: System Demonstrations*, Qun Liu and David Schlangen (Eds.). Association for Computational Linguistics, Online, 38–45. https://doi.org/10.18653/v1/2020.emnlp-demos.6

[32] Yiming Zhu, E. Haq, Lik-Hang Lee, Gareth Tyson, and Pan Hui. 2022. A Reddit Dataset for the Russo-Ukrainian Conflict in 2022. *preprint arXiv:2206.05107* (2022). https://doi.org/10.48550/arxiv.2206.05107