

Analyzing Sentiment of Reddit Posts for the Russia-Ukraine War

Armin Krivičić¹, Sanda Martinčić-Ipšić^{1,2}

¹Faculty of Informatics and Digital Technologies, ²Center for Artificial Intelligence and Cybersecurity,
University of Rijeka,
Radmile Matejčić 2, 51000 Rijeka, Croatia;
Email: armin.krivicic@student.uniri.hr, smarti@uniri.hr

Abstract—Currently, we are witnessing a major global crisis caused by the ongoing war in Ukraine. In this paper, we deepen the insights into people’s opinions about the Russia-Ukraine war by analyzing the posts on the Reddit subreddit r/UkrainianConflict. The extracted posts are preprocessed and the sentiment polarity (neutral, positive or negative) is determined using the Valence Aware Dictionary and sEntiment Reasoner (VADER). We analyze the frequency of the most representative keywords in two longer periods in May and October 2022. The keyword Zelensky appeared more in positive contexts than in negative ones. Keywords Ukraine, Russia, and especially Putin appeared more frequently in negative contexts. Finally, the trend of Google searches on the war is compared with the number of daily subreddit posts.

Keywords—Russia-Ukraine war, Reddit, r/UkrainianConflict NLP, tokenization, stop words, sentiment analysis, NLTK

I. INTRODUCTION

Social media plays an important role in communication during times of crisis, such as the COVID-19 pandemic [1], [2], wars, climate change [3], economic [4] or energy crises, etc. Social media can serve as a key communication platform [5] and is potentially a source of valuable information [6]. Over the past two decades, social media have amplified the spread of information, as well as misinformation and disinformation, which can lead to an infodemic as a negative side effect [7]. Understanding how people post comments in online communication can shed light on the fundamental mechanisms by which collective thinking emerges in a social group [8].

In this light, natural language processing (NLP) [9] offers methods that contribute to better understanding and monitoring the content of crisis communication on social media. Namely, automatic keyword extraction [10], [11], topic modelling [12], [13], named entity recognition [14], text classification [15], sentiment analysis [16], fake news detection, text summarization [17], etc. are all NLP tasks employed for the social media content analysis.

Reddit is a popular social media website with about one million daily active users (as of 2015) [18]. Reddit has many subreddits – communities or user groups that focus communication on a specific topic.

The goal of the conducted research is to discover the polarity of people’s sentiments regarding the war between Russia and Ukraine and how this polarity changes over

time. Hence, this paper analyses the content of communication, related to the topic of the war crisis in Ukraine on the subreddit r/UkrainianConflict. We collect data from this subreddit and employ NLP techniques for sentiment analysis.

The analysis is performed on posts from the subreddit r/UkrainianConflict during the period from January 1 to October 30, 2022. The extracted posts are preprocessed and the sentiment polarity (neutral, positive or negative) is determined using the Valence Aware Dictionary and sEntiment Reasoner (VADER) dictionary [19]. Finally, the trend of Google searches on the war is compared to the number of daily subreddit posts.

Section II contains a short overview of the sentiment research based on Reddit. In Section III, we describe the collected dataset. In Section IV, we present and discuss the results of sentiment analysis and statistics of representative keywords over different time periods along with Google Trends. In the last section, we give some concluding remarks and plans for future work.

II. RELATED WORK

Reddit has evolved in one of the most prominent social platforms on the Internet over the past decade, with 52 million daily active users, over 430 million active monthly users [20] and over 100,000 active topical communities called subreddits. Reddit is ranked as the 20th-most-visited website in the world and the 5th in the US [20], [21]. The authors of [22] performed a systematic analysis of 727 manuscripts that analyzed Reddit data. They show that the scope of Reddit research broadly spans many scientific fields, from computer science, medicine and health, natural sciences to humanities and social sciences, indicating that Reddit has gained important status as a source of information for the research and has an impact on the community [18].

Sentiment analysis (or opinion mining) is a field of research that analyzes people’s opinions, sentiments, evaluations, appraisals, attitudes, and emotions toward entities such as products, services, organizations, individuals, issues, events, topics, and their attributes [23]. Reddit has been the source of the sentiment analysis of user communities in a wide spectrum of domains spanning from COVID-19 vaccination [1], [2], video games stock

prices [4], advancement of robotic technologies [24], the education of gifted children [25], public perception of electric vehicles [3], and so on.

Melton et al. in [1] studied COVID-19 vaccine-related discussions on 13 subreddits performing sentiment analysis and topic modelling. Polarity analysis suggested that Reddit communities expressed more positive sentiment than negative sentiment, and this remained consistent over time.

In [2], Yan et al. examined location-based subreddits to identify city-level variations in sentiment toward vaccine-related topics. The random forest regression model was trained to assign sentiment scores for joy, sadness, fear, and anger to comments. The authors confirm that Reddit comments can be used to better understand concerns and sentiments around COVID-19 vaccines at the local level.

The authors in [4] investigated the impact of sentiment in discussions on the r/WallStreetBets subreddit on the price dynamics of the American online retailer GameStop in 2021. They extend the VADER [19] sentiment analysis package to perform textual analysis of 10.8 million comments and confirm that sentiment has an impact on stock prices.

Savela et al. in [24] investigated sentiment in social media discussions about robotic technologies. Comments were processed using VADER and LIWC (Linguistic Inquiry and Word Count) [26] dictionaries, and logistic regression models were used to infer polarity. The results indicate that the terms "robot" and "artificial intelligence" were used less frequently in a positive context.

Ruan and Lv in [3] studied the public perception of electric vehicles from the lenses of ecology and energy to mitigate climate change and global warming. They utilized VADER dictionary for sentiment inference and concluded that issues that impact the environment (e.g., climate change) have lower sentiment scores.

The study in [27] detected emotions, specifically "hate" and "fear" in Reddit posts during the period from May to July 2022. The authors used the NRC Emotion Lexicon [28] and topic modelling and concluded that the "spikes in hope/fear, both positives and negatives, are present after important battles, but also some non-military events, such as Eurovision and football games".

With this paper, we aim to fill the gap in the analysis of the Reddit discussion on the Russia-Ukraine war in May and October 2022.

III. METHODOLOGY AND EXPERIMENTS

A. Dataset

According to Zhu et al. in [29], r/UkrainianConflict is one of the twelve most prominent subreddits that focuses on the Russia-Ukraine war. It has one of the highest numbers of active users, posts and comments. Posts from the said subreddit were collected using the PSAW (Python Pushshift.io API Wrapper) library [30] implemented in the Python 3 language. We defined a function that accepts

three parameters: start date, end date and the name of the subreddit from which we want to retrieve posts. Along with the text, the date of the post was also collected in Python *datetime* format. Our function then stores the collected data in a *pandas DataFrame* format [31] and also extracts it to a comma-separated values (CSV) file. The CSV file was then separated into two parts, one for May 2022 and one for October 2022, to allow for a deeper analysis of these two time periods. These two time periods hopefully provide a stable reflection of the war-related communication, since they are not the beginning, which is characterized by excessive posting, reflecting the initial impulses of reaction (please see the frequency for February in Fig. 1). From January to October 2022, we collected a total of 101,977 posts.

B. Dataset preprocessing

In order to apply NLP techniques, the text must first be converted to lowercase. Then the text must be tokenized, that is, it must be broken down into words. Tokenization is performed using the NLTK (Natural Language Toolkit) tokenizer from the NLTK library [32]. The NLTK tokenizer uses regular expressions to determine the words in each sentence (*tokens*) and stores them in lists for each post. These lists are then stored in the new column of our *DataFrame*.

C. Dataset analysis

For each day of the period January 2022 – October 2022 we counted the posts and created a key-value dictionary. The keys are dates, while the values represent the number of posts for the corresponding date. Fig. 1 shows the created dictionary as a graph, with the dates on the x-axis and the number of daily posts on the y-axis. A peak in the number of posts can be seen in the graph, which correlates with the start of the Russian invasion of Ukraine. The invasion began on February 24 [33], while the peak was reached three days later.

One of the most informative analyses is based on the frequency of words in users' posts. To measure frequency,

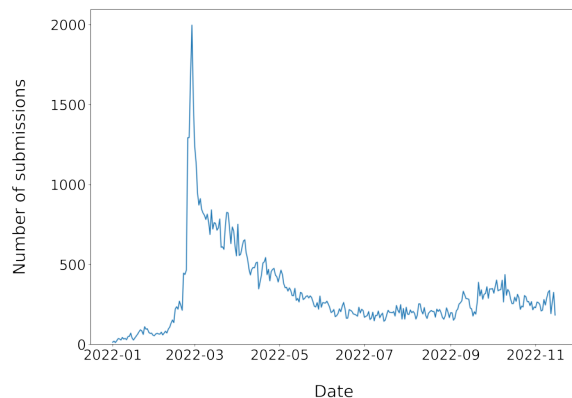


Fig. 1: Number of posts per day

we need to further preprocess the texts by removing stop words. Stop words are words that do not carry any meaning, e.g., *the, a, an, is, are, in, on*, but have a syntactical function in a text. To this end, we use a list of English stop words from the NLTK library. We create a function for comparison of the tokens in the text with a list of stop words and remove the stop words from further analysis. Then we calculate the frequency of each word in the dataset.

The top ten most frequent words (in descending order) are: **ukraine** (39,252), **russian** (32,016), **russia** (22,614), **ukrainian** (14,876), **war** (11,812), **putin** (9,470), **forces** (6,063), **military** (5,506), **says** (4,489) and **russians** (4,178). Fig. 2 shows the histogram of the top ten most frequent words.

IV. RESULTS

A. Sentiment analysis

In order to detect the polarity of Reddit posts, we use the NLTK Sentiment Intensity Analyzer [32]. The analyzer is implemented with the VADER lexicon. The VADER lexicon is a parsimonious rule-based model with over 7500 words labelled with sentiment intensity, where -4 denotes extremely negative sentiment and +4 denotes extremely positive sentiment [19]. Labelling was performed using Amazon Mechanical Turk (AMT). These words were obtained from Twitter and also include emojis and abbreviations. For this reason, the VADER lexicon is well suited for sentiment analysis in social networks.

Next, the polarity scores for each of the posts in the dataset are calculated, resulting in the *positive, negative, neutral* and *compound* scores. Positive, negative, and neutral indicate the proportion of the text that falls into each category, while the compound score is the sum of

the valence scores of the individual words in the lexicon. The score is then normalized to a range of -1 to 1. The following formula is used to calculate the compound score:

$$\frac{s}{\sqrt{s^2 + \alpha}}$$

where s is the sum of the valence scores while α is a constant set to 15.

As proposed by the authors of VADER lexicon [19], we retain the thresholds for the compound score as follows:

- Negative sentiment: ≤ -0.05 ,
- Neutral sentiment: > -0.05 and < 0.05 ,
- Positive sentiment: ≥ 0.05 .

The initial hypothesis that the majority of posts would show negative sentiment proved to be correct, as can be seen in Fig. 3. Negative sentiment predominated with 47.2%, followed by 31.3% neutral posts and 21.5% posts with positive sentiment.

The average negative sentiment score is -0.5177, the average positive sentiment score is 0.4012, and the average neutral sentiment score is 0.0001. The average overall sentiment score is -0.1582, again confirming the hypothesis about the majority of negative posts.

It is important to examine how sentiment has changed over time. For this purpose, we need to calculate the mean sentiment for each day. If we plot these daily sentiment means, we will get a very impractical chart since there are over 300 values. Instead, we can use the rolling window calculation with a window size of 30, which is implemented in the *pandas* library. The rolling window function creates a *rolling mean* of the sentiment values, and the mean value is updated each day. The resulting graph is smoothed and plotted in Fig. 4. The most interesting part is from the end of February, when we see

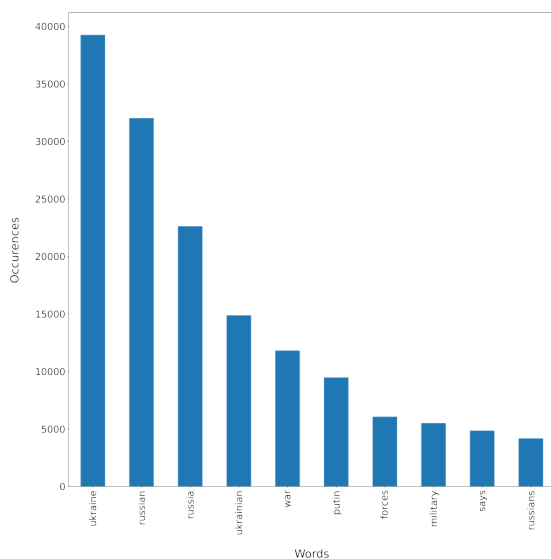


Fig. 2: Frequencies of most common words

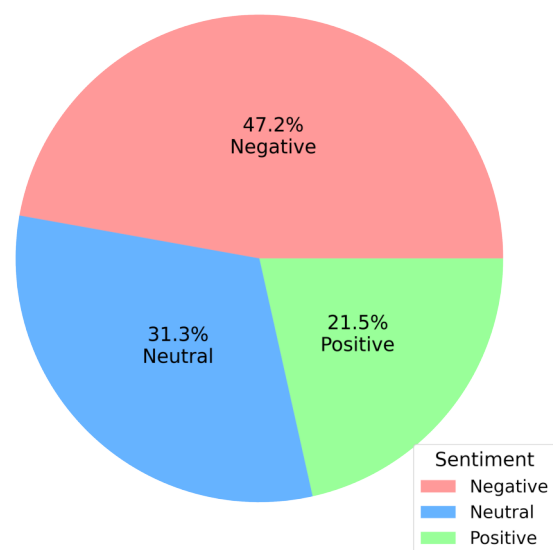


Fig. 3: Distribution of positive, negative and neutral sentiment

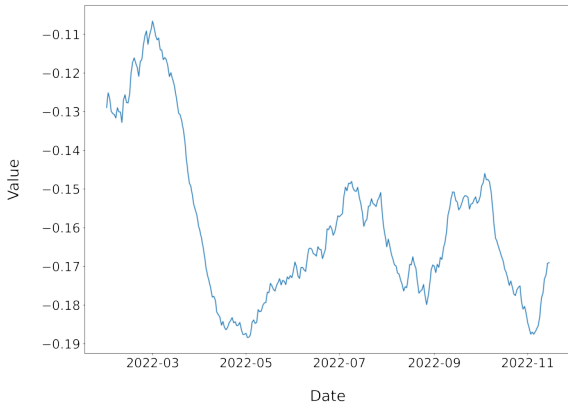


Fig. 4: Sentiment through the time with a window size of 30 days

a steep decline in the sentiment. This decline correlates strongly with the start of the Russian invasion of Ukraine.

B. Representative keywords comparison

Representative keywords used in this paper are **Zelensky**, **Putin**, **Ukraine**, and **Russia**. We divided the dataset into two parts: a positive and a negative part. In order to find out which of these representative keywords are more likely to occur in the positive context and which in the negative context, we need to calculate their percentage. By percentage we mean the ratio between the number of occurrences of the keyword (token) and the total number of tokens. The reason for this normalization is that there are many more negative posts, and the absolute numbers would be meaningless. The keyword **Zelensky** accounts for 0.25% of all tokens in positive posts. In negative posts, this percentage is 0.15%. The keyword **Putin** occurs in 0.49% of positive posts and 0.61% of negative posts. The keyword **Russia** appears in 2.63% of positive posts and 3.83% of negative posts. Finally, the keyword **Ukraine** appears almost equally in both sentiments, 2.26% in the positive and 2.33% in the negative (Tab. 1).

In addition, two longer periods are compared in terms of the most representative words. The first period is May 2022 (April 15 to June 15), and the second is October 2022 (September 15 to November 15) (Tab. 2).

As for the sentiment in these two periods, the average negative value for the month of May is -0.5268, the average positive value is 0.3965, and the average neutral value is 0. The overall average sentiment value is -0.1739,

TABLE I: Share of the most representative keywords in positive and negative sentiment

Keyword	% in positive	% in negative
Zelensky	0.25	0.15
Putin	0.49	0.61
Russia	2.63	3.83
Ukraine	2.26	2.33

TABLE II: Share of the most representative keywords in May and October 2022

Keyword	% in May	% in October
Zelensky	0.0029	0.0009
Putin	0.0132	0.0113
Russia	0.234	0.2226
Ukraine	0.0218	0.0179

slightly more negative than the overall value for 2022, which is -0.1582.

For the month of October, the average negative sentiment value is -0.525, the average positive sentiment value is 0.3938, and the average neutral sentiment value is 0.0001. The average overall sentiment value is -0.1687, which is again slightly more negative than the overall sentiment value for the year 2022. The reason for these negative-leaning sentiment values in May and October is that for the year 2022, the part up to February 24, when sentiment had higher total values around zero (i.e. was more positive in general), was also included.

C. Subreddit posts vs. Google Trends

Comparison of subreddit posts with Google Trends is performed for the same two time periods (May and October 2022). Google Trends measures the popularity of a particular term in Google searches [34]. A term can be a single word such as Ukraine or Russia or a phrase such as the Russia-Ukraine war. Google normalizes the number of searches to a range from 0 to 100, where 100 represents the day when the term was searched for the most. Additionally, we can specify a range where we want to see the popularity of the term, a region in the world, etc. The obtained data can be easily exported to a CSV file with two columns (popularity and date).

In order to compare the number of daily posts from our dataset and the popularity on Google Search, we need to normalize the values since the daily number of posts is in the range of a few hundred per day, while the popularity is at most 100. In Fig. 5 and Fig. 6, we see two graphs for May and October 2022, respectively. These graphs show the comparison, where the orange line represents the trend of Google Search for the term Russia-Ukraine war and the blue line represents the number of daily posts. In May 2022, users were more engaged on Reddit until May 9, after which activity on Reddit and Google was roughly equal. May 9 is known in Europe as Victory Day – a holiday commemorating the Soviet victory over Nazi Germany in 1945, and in Russia it is a national holiday. It was expected that the Russian president would announce an escalation of military action but that did not happen. Nevertheless, the activity on Google was very high on that day. In October 2022, activity remained the same, except for a slightly higher activity on Google in the middle of the month. The jump in activity on September 21 correlates with the day the Russian president announced a partial mobilization of military reservists [35]. The highest activity was on October 10, when Russia

launched a massive missile strike across the entire territory of Ukraine, including the capital, Kyiv [36]. The activity on Reddit follows the activity on Google, suggesting that Reddit is a legitimate source for studying war-related communication through sentiment quantification.

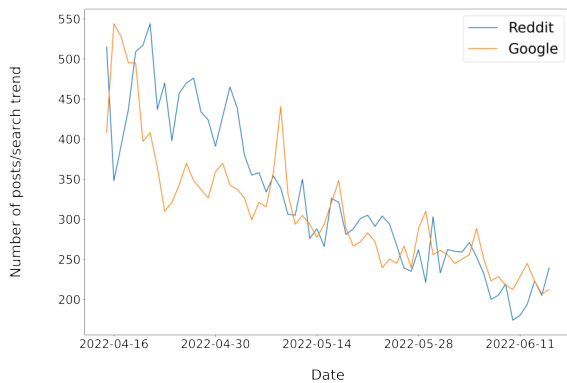


Fig. 5: Number of subreddit posts vs. Google Search trend (term Russia-Ukraine war) for May 2022

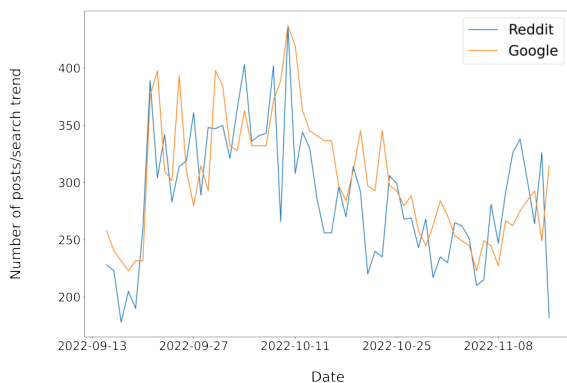


Fig. 6: Number of subreddit posts vs. Google Search trend (term Russia-Ukraine war) for October 2022

V. CONCLUSION

In this paper, we deepen the insights into people’s opinions about the Russia-Ukraine war by analyzing the posts on the Reddit subreddit *r/UkrainianConflict*. The extracted posts are preprocessed and the sentiment polarity (neutral, positive or negative) is determined using VADER. We analyze the frequency of the most representative keywords in two longer periods in May and October 2022 and compared them with Google Trends.

The subreddit *r/UkrainianConflict* was created on 20 February 2014 [29] during Russia’s occupation of the Ukrainian peninsula of Crimea [37]. A month later, Russia annexed Crimea, and to this day there are ongoing tensions between the two countries. Activity in the subreddit was low until late February 2022, which strongly correlates with the escalation of tensions and the date the Russian

president ordered the invasion of Ukraine. Sentiment remained negative throughout the observed period (January 1 to October 30, 2022), dropping sharply after the invasion. Thereafter, it remained relatively stable in the range of -0.15 and -0.2.

The keyword Zelensky appeared much more frequently in positive posts than in negative ones. Keywords such as Ukraine, Russia, and especially Putin appeared more frequently in negative posts. Moreover, the keywords Zelensky and Ukraine appeared more frequently in May than in October. For keywords like Putin and Russia, the occurrence is almost the same in both months.

This research has several limitations. The first is a shorter period of the observation narrowed to May and October of 2022. In the future, we plan to expand this preliminary analysis to the whole period since February 2022. The second one is methodological. We narrowed the detection of the sentiment to the VADER dictionary. The used method showed adequate performance for the task, but for the more granular analysis of sentiments, we plan to train a multi-class classification model.

ACKNOWLEDGMENT

This work has been fully supported by the University of Rijeka under the project number uniri-drustv-18-20.

REFERENCES

- [1] C. A. Melton, O. A. Olusanya, N. Ammar, and A. Shaban-Nejad, “Public sentiment analysis and topic modeling regarding COVID-19 vaccines on the Reddit social media platform: A call to action for strengthening vaccine confidence,” *Journal of Infection and Public Health*, vol. 14, no. 10, pp. 1505–1512, 2021, special Issue on COVID-19 – Vaccine, Variants and New Waves. [Online]. Available: <https://www.sciencedirect.com/science/article/pii/S1876034121002288>
- [2] C. Yan, M. Law, S. Nguyen, J. Cheung, and J. Kong, “Comparing Public Sentiment Toward COVID-19 Vaccines across Canadian Cities: Analysis of Comments on Reddit,” *Journal of Medical Internet Research*, vol. 23, 08 2021.
- [3] T. Ruan and Q. Lv, “Public perception of electric vehicles on Reddit over the past decade,” *Communications in Transportation Research*, vol. 2, p. 100070, 2022. [Online]. Available: <https://www.sciencedirect.com/science/article/pii/S2772424722000208>
- [4] S. C. Long, B. Lucey, Y. Xie, and L. Yarovaya, ““I just like the stock”: The role of Reddit sentiment in the GameStop share rally,” *Financial Review*, vol. 58, no. 1, pp. 19–37, 2023. [Online]. Available: <https://onlinelibrary.wiley.com/doi/abs/10.1111/fire.12328>
- [5] D. C. Glik, “Risk communication for public health emergencies,” *Annual Review of Public Health*, vol. 28, no. 1, pp. 33–54, 2007, pMID: 17222081. [Online]. Available: <https://doi.org/10.1146/annurev.publhealth.28.021406.144123>
- [6] C. Cuello-Garcia, G. Pérez-Gaxiola, and L. Amelsvoort, “Social media can have an impact on how we manage and investigate the COVID-19 pandemic,” *Journal of Clinical Epidemiology*, vol. 127, 06 2020.
- [7] G. Eysenbach, “Infodemiology: The epidemiology of (mis) information,” *The American journal of medicine*, vol. 113, pp. 763–5, 01 2003.
- [8] A. N. Medvedev, R. Lambiotte, and J. Delvenne, “The anatomy of Reddit: An overview of academic research,” *CoRR*, vol. abs/1810.10881, 2018. [Online]. Available: <http://arxiv.org/abs/1810.10881>
- [9] C. Manning, H. Schütze, and L. Lee, *Foundations of Statistical Natural Language Processing*. Cambridge, Massachusetts: The MIT Press, 08 1999. [Online]. Available: <http://nlp.stanford.edu/fsnlp/>

- [10] S. Beliga, A. Meštrović, and S. Martinčić-Ipšić, "An Overview of Graph-Based Keyword Extraction Methods and Approaches," *Journal of Information and Organizational Sciences*, vol. 39, no. 1, pp. 1–20, Jul. 2015. [Online]. Available: <http://jios.foi.hr/index.php/jios/article/view/938>
- [11] S. Beliga, A. Meštrović, and S. Martinčić-Ipšić, "Selectivity-Based Keyword Extraction Method," *International Journal on Semantic Web and Information Systems (IJSWIS)*, vol. 12, no. 3, pp. 1–26, 2016.
- [12] P. K. Bogović, A. Meštrović, and S. Martinčić-Ipšić, "Topic Modeling for Tracking COVID-19 Communication on Twitter," in *Information and Software Technologies*, A. Lopata, D. Gudonienė, and R. Butkienė, Eds. Cham: Springer International Publishing, 2022, pp. 248–258.
- [13] P. K. Bogović, A. Meštrović, S. Beliga, and S. Martinčić-Ipšić, "Topic Modelling of Croatian News During COVID-19 Pandemic," in *2021 44th International Convention on Information, Communication and Electronic Technology (MIPRO)*, 2021, pp. 1044–1051.
- [14] S. Beliga, S. Martinčić-Ipšić, M. Matešić, I. Petrijevcin Vuk-sanović, and A. Meštrović, "Infoveillance of the Croatian Online Media During the COVID-19 Pandemic: One-Year Longitudinal Study Using Natural Language Processing," *JMIR Public Health and Surveillance*, vol. 7, no. 12, p. 15, 2021.
- [15] S. Martinčić-Ipšić, T. Miličić, and L. Todorovski, "The Influence of Feature Representation of Text on the Performance of Document Classification," *Applied Sciences*, vol. 9, no. 4, pp. 743–770, 02 2019.
- [16] K. Babić, M. Petrović, S. Beliga, S. Martinčić-Ipšić, M. Matešić, and A. Meštrović, "Characterisation of COVID-19-Related Tweets in the Croatian Language: Framework Based on the Cro-CoV-cseBERT Model," *Applied Sciences*, vol. 11, no. 21, p. 10442, Nov 2021. [Online]. Available: <http://dx.doi.org/10.3390/app112110442>
- [17] D. Aljević, L. Todorovski, and S. Martinčić-Ipšić, "Extractive Text Summarization Based on Selectivity Ranking," in *2021 International Conference on INnovations in Intelligent SysTems and Applications (INISTA)*, 2021, pp. 1–6.
- [18] K. E. Anderson, "Ask me anything: what is Reddit?" *Library Hi Tech News*, vol. 32, no. 5, pp. 8–11, 2015.
- [19] C. Hutto and E. Gilbert, "VADER: A Parsimonious Rule-Based Model for Sentiment Analysis of Social Media Text," *Proceedings of the Eighth International AAAI Conference on Weblogs and Social Media*, vol. 8, no. 1, pp. 216–225, 2014.
- [20] P. Bump. (2022, Oct) 24 Reddit Stats and Facts to Know in 2022. [Online]. Available: <https://blog.hubspot.com/marketing/reddit-stats>
- [21] Semrush. (2022, 03) Reddit Website Traffic, Ranking, Analytics. [Online]. Available: <https://www.semrush.com/website/reddit.com/overview/>
- [22] N. Proferes, N. Jones, S. Gilbert, C. Fiesler, and M. Zimmer, "Studying Reddit: A Systematic Overview of Disciplines, Approaches, Methods, and Ethics," *Social Media + Society*, vol. 7, p. 205630512110190, 04 2021.
- [23] B. Liu, "Sentiment Analysis and Subjectivity," in *Handbook of Natural Language Processing*, 2010.
- [24] N. Savela, D. Garcia, M. Pellert, and A. Oksanen, "Emotional talk about robotic technologies on Reddit: Sentiment analysis of life domains, motives, and temporal themes," *New Media & Society*, p. 146144482110672, 12 2021.
- [25] J. Hodges, M. Simonsen, and J. Ottwein, "Gifted Education on Reddit: A Social Media Sentiment Analysis," *Gifted Child Quarterly*, vol. 66, p. 001698622210764, 02 2022.
- [26] J. Pennebaker, M. Francis, and R. Booth, "Linguistic inquiry and word count (LIWC)," *Mahway: Lawrence Erlbaum Associates*, 01 1999.
- [27] A. Guerra and O. Karakuş, "Sentiment Analysis for Measuring Hope and Fear from Reddit Posts During the 2022 Russo-Ukrainian Conflict," 2023. [Online]. Available: <https://arxiv.org/abs/2301.08347>
- [28] S. M. Mohammad and P. D. Turney, "NRC Emotion Lexicon," *National Research Council, Canada*, vol. 2, p. 234, 2013.
- [29] Y. Zhu, E.-u. Haq, L.-H. Lee, G. Tyson, and P. Hui, "A Reddit Dataset for the Russo-Ukrainian Conflict in 2022," 2022. [Online]. Available: <https://arxiv.org/abs/2206.05107>
- [30] D. Marx, "Python Pushshift.io API Wrapper (for comment/submission search)," 2018. [Online]. Available: <https://github.com/dmarx/psaw>
- [31] The pandas development team, "pandas-dev/pandas: Pandas," Feb. 2020. [Online]. Available: <https://doi.org/10.5281/zenodo.3509134>
- [32] S. Bird, E. Klein, and E. Loper, *Natural language processing with Python: analyzing text with the natural language toolkit*. " O'Reilly Media, Inc.", 2009.
- [33] The Kyiv Independent news desk. (2022) Putin Declares War on Ukraine. [Online]. Available: <https://kyivindependent.com/national/putin-declares-war-on-ukraine>
- [34] Google. Google Trends. [Online]. Available: <https://trends.google.com/trends/>
- [35] S. McCarthy. (2022, 09) Russia announces immediate 'partial mobilization' of citizens, escalating its invasion of Ukraine. [Online]. Available: <https://edition.cnn.com/2022/09/21/europe/ukraine-russian-referendums-intl-hnk/index.html>
- [36] (2022, 10) Timeline of missile strikes on Ukraine on October 10. [Online]. Available: <https://www.ukrainianworldcongress.org/timeline-of-missile-strikes-on-ukraine-on-october-10/>
- [37] (2014) Russia's invasion timeline. [Online]. Available: <https://war.ukraine.ua/russia-s-invasion-timeline/>