

МАТЕМАТИКА И МАТЕМАТИЧЕСКО ОБРАЗОВАНИЕ, 2024  
MATHEMATICS AND EDUCATION IN MATHEMATICS, 2024  
*Proceedings of the Fifty-Third Spring Conference  
of the Union of Bulgarian Mathematicians  
Borovets, April 1–5, 2024*

SENTIMENT ANALYSIS AND CHANGE: A CASE STUDY  
OF EVALUATING USER PERCEPTIONS DURING  
INTERNATIONAL COMPANIES' REBRANDING\*

Anastasiia Medvedenko<sup>1</sup>, Ekaterina Gruzova<sup>2</sup>, Svetlana Gaivoronskaya<sup>3</sup>,  
Snezhana Besolova<sup>4</sup>, Milena Dobрева<sup>5</sup>, Krassimira Ivanova<sup>6</sup>

<sup>1,2,3,4</sup>ITMO University, St. Petersburg, Russia

e-mails: <sup>1</sup>asya.medvedenko@gmail.com, <sup>2</sup>gruzovae@gmail.com, <sup>3</sup>nordensvetra@gmail.com,

<sup>4</sup>snejan.besolova@yandex.ru

<sup>5,6</sup>Institute of Mathematics and Informatics, Bulgarian Academy of Sciences,  
Sofia, Bulgaria

e-mails: <sup>5</sup>milena.dobрева@gmail.com, <sup>6</sup>kivanova@math.bas.bg

The article presents a sentiment analysis of the process of rebranding international companies in Russia following their withdrawal from the local market after 24.02.2022. The comparative analysis of social media content related to two pairs of companies (McDonald's / Vkusno I Tochka; Zara / Maag) showed different patterns of users' sentiments about the change. This is a pioneering attempt to explore to what extent social media analysis methods and tools can shed light on rebranding processes. This research demonstrates the potential for conducting such types of studies, but also to highlight the still existing imperfections in terms of sentence-level opinion assessment, especially in non-English texts.

**Keywords:** Data Mining, Sentiment Analysis, Twitter, RapidMiner, Phantom-Buster, Python

ИЗВЛИЧАНЕ НА МНЕНИЯ И ЧУВСТВА ПО ВРЕМЕ НА  
ПРОМЕНИ: ПРОУЧВАНЕ НА ПОТРЕБИТЕЛСКИТЕ  
ВЪЗПРИЯТИЯ ПРИ РЕБРАНДИРАНЕТО НА  
МЕЖДУНАРОДНИ КОМПАНИИ

Анастасия Медведенко<sup>1</sup>, Екатерина Грузова<sup>2</sup>, Светлана Гайворонская<sup>3</sup>,  
Снежана Бесолова<sup>4</sup>, Милена Добрева<sup>5</sup>, Красимира Иванова<sup>6</sup>

<sup>1,2,3,4</sup>ИТМО университет, Санкт Петербург, Русия

e-mails: <sup>1</sup>asya.medvedenko@gmail.com, <sup>2</sup>gruzovae@gmail.com, <sup>3</sup>nordensvetra@gmail.com,

<sup>4</sup>snejan.besolova@yandex.ru

---

\* <https://doi.org/10.55630/mem.2024.53.085-091>

The research was inspired by the course *Cultural Data Mining and Visualization* within the master's program *Digital Methods in Humanities/Digital Humanities* at the DH-Center of ITMO University.

2020 Mathematics Subject Classification: 68T50.

<sup>5,6</sup>Институт по математика и информатика, Българска академия на науките,  
София, България  
e-mails: <sup>5</sup>milena.dobрева@gmail.com, <sup>6</sup>kivanova@math.bas.bg

Статията представя приложение на методите за проучване на мнения и чувства на потребителите, приложени към процеса на ребрандиране на международни компании в Русия след оттеглянето им след събитията на 24 февруари 2022. Направеният примерен сравнителен анализ на две двойки компании (McDonald's/Вкусно и точка; Zara/Маг) показва различна реакция на потребителите относно промяната. Това е пионерски опит да се проучи до каква степен методите и инструментите за анализ на социалните медии могат да хвърлят светлина върху процесите на ребрандиране. Проучването цели да демонстрира потенциала на съществуващите методи и средства за анализ на съдържание от социални медии при процеси на ребрандиране, но също така да открие все още съществуващите несъвършенства по отношение на оценката на мненията на ниво изречение, особено в текстове, които не са на английски език.

**Ключови думи:** извличане на данни, анализ на мнения и чувства, Twitter, RapidMiner, PhantomBuster, Python

**1. Introduction.** In recent years, the development of Natural Language Processing (NLP) has reached sufficient maturity to move from the research stage of technology development to widespread use in practice. One commonly applied NLP technique is sentiment analysis, also known as opinion mining. It is used to determine and extract meaningful insights from sentiment data. Essentially, sentiment analysis involves using computational methods and algorithms to analyse and quantify subjective information within textual data, allowing businesses, organisations, and researchers to gain insight into public opinion, customer feedback, or overall sentiment around specific topics, products, brands, or events. This information can be valuable for making data-driven decisions, monitoring public perception, and understanding consumer sentiment in various contexts, such as marketing, customer service, and social media monitoring.

Our first objective is to investigate the possibilities of applying sentiment analysis to assess people's attitude change during a rebranding process by exploring opinions expressed on the social media platform Twitter. For this purpose, we examine the users' perception of two rebranded companies in Russia following the events of 24.02.2022.

There are several techniques applied to sentiment analysis. Lexicon-based methods use the sentiment polarity of the words to calculate the sentiment score of the text on the base of the presence of these words in the text. Machine and deep learning approaches involve training models on labelled datasets to classify text into different sentiment categories. Aspect-based sentiment analysis focuses on identifying the sentiment associated with specific aspects or features mentioned in the text, especially in product reviews or social media posts. Old-fashioned rule-based techniques involve defining a set of rules or patterns to identify sentiment expressions in the text. Semantic analysis aims to understand the meaning and context of text by analysing the relationships between words and phrases [20]. However, research in this area still mainly focuses on building systems and resources using the English language. There is a growing set of research on non-English texts, but it continues to lag in its development into tools for widespread use by non-specialists [3, 7].

Our second objective is to study the level the readiness of the tools that can be used to conduct sentiment analysis, especially on non-English informal texts.

The paper is organised as follows: Section 2 introduces the problem statement and hypotheses; Section 3 describes the methodology; Section 4 explains how the data was collected and preprocessed; Section 5 is dedicated to the Python-based analysis and the manual postprocessing of the data; Section 6 presents the results; the final section explores future research directions.

**2. Background.** Due to the events of 24.02.2022 and the following turmoil, according to the World Bank Group, 150 multinational companies have announced complete withdrawals, while a further 250-plus have suspended operations in the Russian Federation. The withdrawal has been the subject of research for more than a year [5]. Following the withdrawal and suspension of businesses, certain companies announced selling their assets to franchisee companies or local entrepreneurs, which had their businesses reopened and rebranded. For example, Zara stores in Russia have been transitioned to the Daher group and reopened under the name Maag; meanwhile, McDonald’s was sold to a Russian franchisee Alexander Govor who rebranded the business as Vkusno I Tochka.

The withdrawal and the rebranding of companies are widely discussed on social media. Moreover, several studies have been conducted on users’ responses to the events [4, 19]. Due to the acuteness of the issue, Twitter released an official statement where the board of directors announced the stiffening of the platform’s policies to protect its audience [9]. Twitter (currently X) was a growing social media platform as it allowed people to share and express their opinions and post messages with little censorship, mostly related to suspending accounts of questionable content [16]. How Elon Musk’s management has affected the platform has yet to be researched, but as of May 2022, when the investigation and analysis were primarily conducted, there were no quantitative restrictions on providing content. Despite being officially blocked in Russia, Twitter remained relatively popular, with over 700,000 officially registered Russian users [8].

Our study focuses on the information behaviour aspects related to this process, and more specifically on the change of attitudes to the rebranded entities compared to their ‘maternal companies’. Our initial hypothesis is that the attitude has significantly worsened with the number of tweets rising. In addition, we aim to verify whether modern Python-based data analysis models can assess attitudes to a specific object of discussion despite the general tone of discussion being different (i.e., the subject complains about being unable to obtain a certain item which they like – in this case, the attitude to the object is positive while the tone of the statement is negative).

**3. Methodology.** In this study, we apply various techniques and methods of social media data analysis. The data collection was done via low-code tools such as RapidMiner, a data mining desktop application, and PhantomBuster, a cloud-based data extraction software. This allowed us to access a greater source of posts in comparison to other methods. The next stage in the data processing is preprocessing and data cleansing. A quantitative analysis step follows, using an algorithm written in Python and based on its Hugging Face library (RuBERT-RuSentiTweet model). Then, the postprocessed data is manually verified and validated to ensure the correctness of the Python code and alignment with our research goals. Finally, the comparative method is applied by considering the differences and fluctuations in attitudes to the original brands and their

rebranded counterparts. The analysis is illustrated by visualization of the data in Tableau Public online.

**4. Collecting and Preprocessing Data.** We decided to limit the timeframes applied to the study for each pair of companies; this would allow comparing the rebranded entities with their original counterparts (Maag/Zara; Vkusno I Tochka/McDonald’s). We chose the year 2019 (01.01.2019–31.12.2019) for both Zara and McDonald’s. Among the major reasons justifying the choice are several factors: first, that was a pre-pandemic year when physical mobility to enter public spaces was not restricted; and second, that was a period before the withdrawal of companies. While Vkusno I Tochka was reopened in June 2022, the first Maag shop in Russia was established only by the end of April 2023. To provide an objective analysis, we decided to collect tweets related to either entity published by Russian-speaking users between 27.04.2023 and 27.05.2023.

As the data collection tool, we initially experimented with RapidMiner [15], a free tool from the Technical University of Dortmund, which provides a convenient environment for data collection and knowledge extraction. Although generally fast and efficient, RapidMiner was unable to collect tweets published later than one week due to Twitter’s official policy regarding access to archived posts by external tools [18]. An additional data collection method was selected to address the limitations of RapidMiner. PhantomBuster was chosen as an alternative to RapidMiner. The cloud software was initially designed for marketing purposes but is applicable in other research works and helps to scrape the data from different social media, including Twitter in a formatted way [13]. It limits the number of attempts at collecting data which requires each participant of the group to get involved in data mining. The collected data is kept in a temporary storage and can be uploaded as a CSV or JSON document.

We performed a manual preprocessing and cleansing of the data based on the following criteria:

Tweets typed in the Cyrillic alphabet but not in the Russian language were excluded;

Tweets in Russian but by a person who is not a Russian Federation resident were excluded. Immigrants of Russian citizenship (including those who left the country after 24.02.2022) are to be the exception to the rule.

Context advertisements of items unrelated to the study’s topic (i.e., homonyms) were excluded.

Verification of the users was conducted by investigating their accounts. After the preprocessing, the number of tweets decreased to 542. 119 of them were related to Maag (April-May 2023), 129 to Vkusno I Tochka (April-May 2023), and 294 were divided between Zara and McDonald’s (97 and 197, respectively). We have shared the dataset providing a link to it in the given paper [11].

**5. Python-algorithm Application and Postprocessing of the Dataset.** The major data-processing algorithm applied to the prepared dataset is a RuBERT-RuSenti-Tweet model from the Hugging Face library. The model was initially developed to provide Tone and Sentiment Analysis for posts written in Russian and extracted from Twitter [6]. The method’s choice was justified by the fact that similar libraries and models developed for the analysis of English tweets could not provide adequate results since both languages are grammatically, semantically, and syntactically different. We needed to download and install certain libraries and models, such as Transformers, Pandas, and Hugging Face

(RuBERT-RuSentiTweet model). Before conducting the sentiment analysis, it was vital to establish a new data frame, indicate classes, define additional functions, and assign the right columns to the classifier.

The RuBERT-RuSentiTweet model and Russian Toxicity Classifier were applied to the dataset. The Python-based algorithm provided the following assessment categories: 'label' (positive/ neutral/negative), which identified the general perception and attitude of a tweet, 'score' (floating number), 'label\_toxic' (toxic/neutral), which classified each tweet either as toxic or neutral, and 'score\_toxic' (floating number). After summarising the results of Python processing, we decided to focus on the 'label' and 'score' since the expressed general attitude of a tweet author was more related to the subject of research. The code file is accessible online [10].

Since one of our major concerns was the ability of the RuBERT-RuSentiTweet model to assess attitude toward the specific object of discussion despite the general tone of a post being different, postprocessing was a vital procedure to perform before proceeding to the visualisation and the analysis stages. We made an independent evaluation of each of the tweets, which we compared with the results of the algorithm. It was found that 259 out of 542 tweets were incorrectly classified ( $\sim 47.8\%$ ), a significantly high number caused by the inability of the algorithm to extract a subject of assessment, or a named entity in our case. For instance, in one of the Vkusno I Tochka segments, a user expressed their appreciation of a cafe: 'Самый вкусный на ВДНХ в МСК, недалеко от центральной аллеи. Вкусно и точка.' (translation: 'The most delicious at VDNH in MSK, the one not far from the central alley. Vkusno I Tochka.'). However, the model recognized it as a 'negative' statement. In some instances, the algorithm was unable to identify a sarcastic notion, marking it as 'positive': 'Врум! Врум! Это бздит твой ЖКТ после Вкусно и Точка' (translation: 'Vroom! Vroom! This is your GI tract rumbling after Vkusno I Tochka').

**6. Analysis of the Results.** Analysis was conducted relying on the visualisation of the data arranged in Tableau Public online service represented in charts available in the public domain [12].

**McDonald's (2019) and Vkusno I Tochka.** The analysis of the two companies which have larger shares of tweets both in 2019 and 2023, McDonald's and Vkusno I Tochka showed the following. The biggest share of tweets related to McDonald's are neutral in sentiment (44%) prevailing over both positive (33%) and negative (23%) ones separately. Similarly, for Vkusno I Tochka, its neutral score (68%) prevails over both positive (13%) and negative (19%). We can conclude that most people viewed the companies in a rather neutral way both in 2019 and 2023. However, the newer company, Vkusno I Tochka attracts more apathy and indifference. This needs to be explored further, possibly applying some qualitative research. In addition, the counterpart's shares of positive and negative assessments differ. In the case of McDonald's, positive opinions are prevalent; as for Vkusno I Tochka, the negative evaluation is higher.

**Zara (2019) and Maag.** Neutral sentiments towards Zara (39%) prevails over both positive (30%) and negative (31%). The shares of negative and positive evaluations are nearly equal. In the case of Maag, its neutral score also prevails over both positive (8%) and negative (29%) summarised (63% vs. 37%). It should be mentioned as well that the share of negative tweets is considerably larger in comparison to positive tweets.

The analysis of Zara and Maag shows different patterns. We might dispute whether Zara was predominantly seen in a neutral light or not, whereas the prevalence of a neutral attitude towards Maag is indisputable. The general trends evidence an increase in both neutral and negative estimates accompanied by a decrease in positive sentiments. Overall, the data evidence suggests that Maag's rebranding disengaged certain customers.

**7. Conclusions and Further Research.** The results of this study showed that the computer science methods and tools applied for sentiment analysis at that moment need to be improved. Currently, they lack the ability to identify named entities in text (i.e. tweets) written in Russian and evaluate the attitude to them. Before developing a new sentiment-analysis model (or, to be more specific, a Python library based on a new model), we need to advance the existing model since it is capable of providing a general assessment of a statement: among the options we are considering at the moment is the implementation of a loop which identifies and separates named entities, recognises and analyses words and expressions which characterise the entity, and provides the result as a set of attributes where each recognised entity is assessed.

The hypothesis of our research related to the information behaviour of users following companies' rebranding was that the attitude to the rebranded companies has worsened in comparison to the original ones. The above analysis shows that this statement is partly true: we registered both growth of neutral responses and a decrease in positive attitudes in cases of both pairs of companies. The differences across the pairs suggest that a wider study involving more brands would refine the information behaviour patterns. Identifying the reasons behind these different behaviours is an exciting domain for larger exploration. In addition, we could explore other environments with changes in brands.

In addition, there are issues, such as changes in the Twitter/X policy regarding the number of tweets allowed per day, which might compromise the study's credibility in the future unless such limitations are properly handled. Finally, a notable domain of growing exploration is the responsibility of social media platforms in maintaining accounts and dealing with content within the context of the Code of Practice in Disinformation of the European Commission.

**Acknowledgements.** We thank Karina Agafonova (ITMO) for her invaluable assistance in the study.

## REFERENCES

- [1] R. CHURIKOV. Russian Clothing Brands Plug Gap Left by Western Rivals, Reuters, June 2023, <https://www.reuters.com/business/retail-consumer/russian-clothing-brands-plug-gap-left-by-western-rivals-2023-05-31/>.
- [2] S. EVENETT, N. PISANI. Less than Nine Percent of Western Firms Have Divested from Russia, Social Science Research Network, December 2022, <https://dx.doi.org/10.2139/ssrn.4322502>.
- [3] F. FERDIANA, R. DJATMIKO, M. FARIS. A Review of Sentiment Analysis for Non-English Language. 2019 Int. Conf. of Artificial Intelligence and Information Technology, Yogyakarta, Indonesia, pp. 448-451, 2019. <https://doi.org/10.1109/ICAIIIT.2019.8834552>.
- [4] M. B GARCIA, A. CUNANAN. Public Sentiment and Emotion Analyses of Twitter Data on the 2022 Russian Invasion of Ukraine, 9th Int. Conf. on Information Technology, Com-

- puter, and Electrical Engineering, Semarang, Indonesia, 2022, <https://doi.org/10.1109/ICITACEE55701.2022.9924136>
- [5] J. D.GUÉNETTE et al., Implications of the War in Ukraine for the Global Economy, World Bank Group, EFI Policy Note 3, 6-11 April 2022. <https://thedocs.worldbank.org/en/doc/5d903e848db1d1b83e0ec8f744e55570-0350012021/related/Implications-of-the-War-in-Ukraine-for-the-Global-Economy.pdf>.
  - [6] Hugging Face library, RuBERT-RuSentiTweet, <https://huggingface.co/sismetanin/rubert-rusentitweet>.
  - [7] M. KAITY, V. BALAKRISHNAN. Sentiment Lexicons and non-English Languages: a Survey. *Knowledge and Information Systems*, **62** (2020), 4445–4480, <https://doi.org/10.1007/s10115-020-01497-6>.
  - [8] S. KEMP. Digital 2023: the Russian Federation, DataReportal, February 2023, <https://datareportal.com/reports/digital-2023-russian-federation#:~:text=Twitter%20users%20in%20the%20Russian,total%20population%20at%20the%20time>.
  - [9] S. MCSWEENEY. Our Ongoing Approach to the War in Ukraine, Twitter Blog, March 2022, [https://blog.twitter.com/en\\_us/topics/company/2022/our-ongoing-approach-to-the-war-in-ukraine](https://blog.twitter.com/en_us/topics/company/2022/our-ongoing-approach-to-the-war-in-ukraine).
  - [10] A. MEDVEDENKO, E. GRUZOVA, S., ROGOVTSOVA, S. BESOLOVA. Project code, <https://colab.research.google.com/drive/1I7eixZp3EGqjvXF1x6sbwXVtae8fcYW6?usp=sharing>.
  - [11] A. MEDVEDENKO, E. GRUZOVA, S., ROGOVTSOVA, S. BESOLOVA. The collected dataset, <https://docs.google.com/spreadsheets/d/1VJ-K53pWmoZhhszfh10drVqut4iGKCvVxILlftaf10Q/edit?usp=sharing>.
  - [12] A. MEDVEDENKO, E. GRUZOVA, S., ROGOVTSOVA, S. BESOLOVA. The visualization charts of the dataset, [https://public.tableau.com/app/profile/snezhana.besolova/viz/tone\\_analysis/2023](https://public.tableau.com/app/profile/snezhana.besolova/viz/tone_analysis/2023).
  - [13] PhantomBuster, a cloud software data collecting tool, <https://phantombuster.com/>.
  - [14] O. POPOVA. Russia's McDonald's Heir Growing Quickly, in for the Long Haul, Reuters, June 2023, <https://www.reuters.com/business/retail-consumer/russias-mcdonalds-heir-growing-quickly-long-haul-2023-06-05/>.
  - [15] RapidMiner, enterprise-ready data science platform, <https://rapidminer.com/>.
  - [16] N. S. REYNA et al., Twitter: More than Tweets for Undergraduate Student Researchers. *Journal of Microbiology & Biology Education*, **23**, no. 1 (2022), e00326-21, <https://doi.org/10.1128/jmbe.00326-21>.
  - [17] J. A. SONNENFELD et al. The Russian Business Retreat – How the Ratings Measured Up One Year Later, YSM and KSE, 2023. <https://kse.ua/wp-content/uploads/2023/02/paper-on-leave-Russia-project.pdf>.
  - [18] Twitter Helper Center, <https://help.twitter.com/en/using-twitter/missing-tweets>.
  - [19] P. VYAS et al. RUemo – The Classification Framework for Russia-Ukraine War-Related Societal Emotions on Twitter through Machine Learning. *Algorithms*, **16**, no. 2 (2023), 69, <https://doi.org/10.3390/a16020069>,
  - [20] M. WANKHADE, A. C. S. RAO, C. KULKARNI. A Survey on Sentiment Analysis Methods, Applications, and Challenges. *Artificial Intelligence Review*, **55** (2022), 5731-5780, <https://doi.org/10.1007/s10462-022-10144-1>.