Understanding Public Perception of Waste Management in Western Balkans using Tweets and News Reports Analysis

Ivana Gjorshoska
Macedonian Academy of Sciences and Arts
Skopje, Republic of North Macedonia
ivana.gjorshoska@students.finki.ukim.mk

Angela Madjar
Macedonian Academy of Sciences and Arts
Skopje, Republic of North Macedonia
angela.madzhar@students.finki.ukim.mk

Aleksandra Dedinec
Faculty of Computer Science and Engineering
Ss. Cyril and Methodius University
Skopje, Republic of North Macedonia
aleksandra.kanevche@finki.ukim.mk

Jana Prodanova
Macedonian Academy of Sciences and Arts
Skopje, Republic of North Macedonia
jprodanova@manu.edu.mk

Abstract—**The rapid urbanization and economic growth greatly contribute to the increasing amount of produced solid waste worldwide. As part of their pro-European policies, many countries from the Western Balkans have started to accept and apply the EU Waste Management Law in their legal systems as well as in practice in order to create adequate waste management schemes. The aim of this paper is to analyze the public response regarding waste management regulations and practices using Twitter and News reports data analysis within a period of four months, between November 1st, 2021 and February 28th, 2022. Sentiment analysis and Dynamic Topic Modeling were used to assess the public attitude in both news reports and Twitter responses, in addition to bigrams similarity using soft cosine measure to determine the semantic correlation between the two sources. The results suggest that there is a predominantly negative attitudes towards waste practices in both media outlets. Additionally, there is a strong semantic correlation between the discussed topics in news reports and on social media, leading to the conclusion that news outlets play a major role in promoting pro-environmental behavior in the general public in Western Balkans.**

Keywords—**waste management, Western Balkans, social media, twitter analysis, news reports, sentiment analysis, dynamic topic modeling, semantic similarity**

I. INTRODUCTION

With the acceleration of urban development, the amount of generated waste in the world increased dramatically and this trend is only said to continue in the following years. According to the World Bank, rapid urbanization, population growth and economic development will increase annual waste generation by 70 percent over the next 30 years, from 2.01 billion tonnes being generated globally in 2016 to 3.40 billion tonnes in 2050 [1]. Around 37 percent of the global waste is disposed in some type of landfill and 33 percent is mismanaged through open dumping or burning, whereas adequate waste disposal is almost exclusively done in high-income countries. Inadequate waste management in urban areas negatively impacts the environment and the climate, public health and ultimately the economy. Environmental pollution of waste dumping affects health through short-term effects such as congenital anomalies, asthma and respiratory infection , while long-term exposure to waste may cause chronic respiratory and cardiovascular diseases, cancer and even brain, nerves, liver or kidneys diseases [2], [3]. To reduce the negative impacts of waste, different counties have

adopted various actions. Since 2000, Japan began implementing the Containers and Packaging Recycling Law [4]. The United States enforced the Resource Protection and Recycling Act as well as different cities establishing their own regulations [5].

In the European Union, 5.2 tonnes of waste were generated per EU inhabitant in 2018, 38.5 percent of which was landfilled while 37.9 percent was recycled [6]. Due to greater economic development and EU legislation, many of the higher recycling and collection rates happen in Western Europe. However, many countries, especially countries from Western Balkans aim to meet the EU Waste Management Law as part of their pro-European policies. The governments of Serbia and Montenegro reported 338 kg and 548 kg of municipal solid waste per inhabitant in 2019 respectively, while North Macedonia reported consistent increase over the last 10 years - from 351 kg in 2010 to 452 kg per inhabitant in 2020 [7]–[9]. Additionally, the amount of waste generated annually per capita is expected to increase linearly until it reaches the corresponding waste projections of EU28 in 2035 [11]. Latest reports in the aforementioned countries also suggest high collection rates, between 70 percent and 90 percent of collected waste, however only up to 3 percent of the collected waste is being recycled and the rest is either landfilled, in both regulated and unregulated landfills, or openly burned [10]. The current waste treatment and increasing trend of waste generation plays a major role in greenhouse gas emissions, where up to 80 percent of emissions caused by the waste management sector come from solid waste disposal in landfills [10]. The main struggle for waste management in Western Balkans can be attributed to outdated infrastructure for collection and transport, illegal dumpsites and the lack of facilities to sort, clean and aggregate recyclable materials [7]–[9]. Lack of adequate policies, inefficient allocation of external resources and lack of awareness among people and companies have been identified as barriers for transitioning to green economy [12].

There is significant research done in the field of public perception for waste management and other environment-related topics based on interactions on social media platforms. Some of the exemplary research includes analyzing attitude towards municipal waste policy as well as promotion of

campaigns to encourage consumers' awareness [13], [14]. However, there is a lack of research about public perception towards waste management in the Western Balkan region. The aim of this paper is to assess the public perception about waste management via Twitter posts, as well as to research the possible correlation between tweets related to waste and news reports in the Western Balkans. The paper is organized as follows. Section II describes the data collection process, storing and preprocessing. Section III presents and explains the methods used for sentiment analysis, topic modeling and similarity assessment between tweets and news reports. Section IV presents and discusses the results obtained from the analysis and finally, Section V provides conclusions of this study and presents possibilities for future work.

II. DATA

For the purpose of the study, we collected tweets related to waste using Twitter's stream API with an authenticated application that connects to a public stream consisting of a sample of tweets that are posted on Twitter [15]. Although many studies for sentiment analysis rely on using hashtags to filter tweets, the filtering of relevant tweets in our case was done using keywords instead of hashtags. This decision was done due to a lack of hashtags relevant to this topic used in the Western Balkan region. The tweets were filtered based on keywords related to waste in Macedonian and Serbo-Croatian i.e. "отпад" (waste), "otpad" (waste), "ѓубре" (garbage) and "рециклира" (recycle). Waste-related tweets were collected on a weekly basis over a period of 4 months starting November 1st, 2021 until February 28th, 2022.

After the initial collection of the tweets, they were saved as an Excel spreadsheet for further data cleansing which included checking and filtering the appropriate tweets about waste-related topics. This secondary filtering step was necessary for this study due to language ambiguity. Namely, words such as "trash" and "garbage" tend to be used as derogatory terms for people, objects or situations and this usage of the aforementioned keywords used for filtering can be seen in the Western Balkan region as well. The collected tweets were translated into English using an online document translation tool and parts of the tweets such as "RT" standing for retweet and usernames were removed from the tweets in order to focus on the tweets' content [16].

The decision to analyze news reports was done due to the fact that journalism on Western Balkans plays a major role in enhancing societal transitions [17]. In order to obtain news reports data, a web scraping method was used for data collection. We used the Beautiful Soup python library to scrape news reports from the websites time.mk and time.rs, which are cluster-based news websites that collect and aggregate news from 120 distinct sources [18]–[20]. Similarly as the tweets, the news reports were filtered using the keywords "otpad" (waste) and "отпад" (waste) and were collected weekly over the same time period as the Twitter data. During this process, we collected both the news title and the news preview to get a better picture about the contents of the news reports for the respective week. The news reports datasets were stored and translated using the same techniques as the waste-related tweets.

III. METHODOLOGY

A. Sentiment Analysis

Sentiment analysis is the practice of using algorithms to classify various samples of related text into positive and negative categories. This method is widely applied in analyzing customer reviews and responses, however it has also been used in online social media platforms for analyzing online opinion. Sentiment analysis within microblogging has verified Twitter as a valid online indicator of political sentiment reflecting the offline political landscape [21]. It has also been used for analyzing tweets for social events and health care tweets [22], [23].

There are several models and tools that can be used for sentiment analysis of text, however none of the tools available right now can do sentiment analysis on Macedonian or Serbo-Croatian text, resulting in the previously explained English translation of the collected data in Section II. For the purpose of sentiment analysis in our research, we used a pre-built Sentiment Intensity Analyzer [24], [25]. The analyzer is built using VADER, a human-centered sentiment lexicon that combines lexical features from well-established sentiment word-banks and features that are common in microblogging, including acronyms, slang and emojis. Every candidate feature of this lexicon has been assessed using a Wisdom-of-the-Crowd approach as well as using identified heuristics humans use to assess sentiment intensity of text, which lead to creating a human-validated sentiment intensity on corpora within four domains including social media text, movie reviews, technical product reviews and opinion news articles. The Sentiment Intensity Analyzer from the NLTK library uses this lexicon to assess the sentiment intensity on a piece of text by verifying whether any of the words in the content are available in the lexicon and producing measurements that assess the sentiment intensity in the text from the identified words [24], [25]. Using this tool, the tweets and news reports for each week were classified in three groups, i.e. tweets and news reports containing positive, negative or neutral sentiment.

B. Dynamic Topic Modeling

Topic modeling is a method for unsupervised classification of documents which finds natural groups of items within a collection and is used for discovering hidden semantic structures in a body of text. Within our analysis, we used a BERTopic model that leverages BERT embeddings and c-TF-IDF to create dense clusters allowing for easily interpretable topics [26]. BERTopic uses document embedding representation with pre-trained Sentence-BERT framework which allows converting sentences and paragraphs into vector representations using the well-known BERT language model. These sentence embeddings are then clustered using Hierarchical DBSCAN, an extention of the DBSCAN algorithm that is able to find clusters of varying density, and Uniform manifold approximation and projection (UPAM) for vector dimensionality reduction. Lastly, topic representations are extracted from the clusters using class-based TF-IDF:

$$\mathrm{W}_{t,c} = tf_{t,c} \cdot log(1 + \frac{A}{tf_t}) \qquad (1)$$

where term frequency models the frequency of the term t in a class c and the inverse document frequency is replaced by

an inverse class frequency to measure how much information a term provides to a class calculated as the logarithm of the average number of words per class A divided by the frequency of term t across all classes.

For the purpose of our analysis, fine-tuning on a pre-trained BERTopic model was done using minimum topic range of 50 and monograms, bigrams and trigrams which were derived from the translated tweets and news reports during the fine-tuning process. Additionally, because the BERTopic model is used to extract topics from large documents, a count vectorizer model was used to ensure relevant topic extraction [27]. The topic over time modeling was done using the fine-tuned BERTopic model with weekly timestamps and global and evolution tuning.

C. Bigrams similarity

In order to evaluate the similarity between the tweets and news reports, we used Soft Cosine Similarity as a method of assessing semantic similarity between two documents. This type of measure considers two vectors as pairs of features and considers their similarity in a vector space model [28]. Given two N-dimension vectors a and b, the soft cosine similarity is calculated as follows:

$$\text{soft\_cosine}_1(a, b) = \frac{\sum_{i,j}^{N} s_{ij} a_i b_j}{\sqrt{\sum_{i,j}^{N} s_{ij} a_i a_j} \sqrt{\sum_{i,j}^{N} s_{ij} b_i b_j}} \quad (2)$$

where $s_{ij} = similarity(feature_i, feature_j)$.

To measure the soft cosine similarity between the most common terms found in tweets and news reports for each week, the top 10 most common bigrams for tweets and news reports were determined using a count vectorizer model [27]. All bigrams taken from tweets and news reports were respectively sorted in ascending order and a similarity matrix was calculated using GloVe vector representations [29]. Soft cosine similarity was calculated on the respective vector representations of the bigrams for each week.

IV. RESULTS

Table I and Figure 1 show statistics for the collected data over the 4 months of data collection period. Sentiment analysis of the data from both sources shows predominantly negative sentiment over positive and neutral content. Additionally, the percentage of texts classified as negative is much higher in the news reports compared to the percentage of negative sentiment in tweets. In Figure 1, a stable trend of news reports over time can be detected, however the trend of collected tweets shows a few bigger peaks. Namely, the first peak around December 6th, 2021 is connected to the ban of single-use plastic bags and additional rise of the cost of biodegradable plastic bags from 2 Macedonian Denars to 15 Macedonian Denars [30]. On the other hand, the second peak of tweets around the end of January 2022, happened after the news that over 300.000 expired COVID vaccines would be landfilled by the end of February due to lack of interest for COVID vaccination. Other smaller peaks include topics connected to generated plastic waste during the COVID-19 pandemic (between November 15th and November 29th), total generated municipal waste (between the end of December 2021 and the beginning of January 2022), as well as electronic waste and landfills regulations (around January 16th, 2022).

TABLE I. Number of Collected Tweets/News Reports

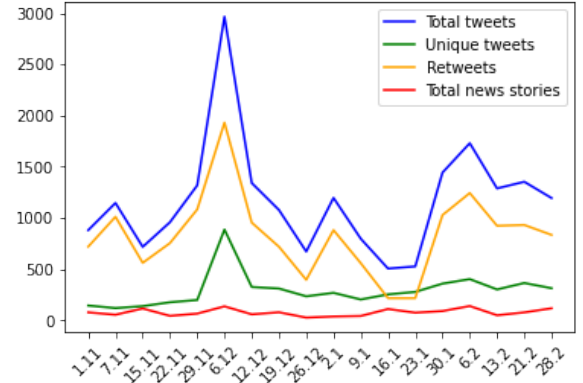| Source	Total	Unique	Positive(%)	Negative(%)	Neutral(%)
Twitter	21124	5287	22.55	68.10	9.35
News	1405	1137	15.23	81.35	3.42

Fig. 1. Number of Tweets and News Reports over time

The sentiment analysis of tweets and news reports for each week follows similar trends as the overall collected data, as shown in Figure 2 and Figure 3. The number of tweets with negative sentiment is significantly bigger than the number of tweets with positive and neutral sentiment and the news report trends adhere to these trend patterns as well. The dynamic topic modeling analysis also shows similarities within the tweets and the news reports over the 4 month period, as displayed in Figure 4. Figure 4a shows the trends of the 10 most common topics identified in tweets. Some of these topics include: people inadequately throwing waste, plastic bags, nuclear waste as a result of nuclear power plants, medical waste and its management by the municipalities, burning wood and other materials for the purpose of heating, disposal of electronic waste such as lithium batteries, composting bio waste and vehicle recycling. Over the period of 4 months, all of the identified topics have similar frequency rates with the exception of the topic about plastic bags, that is connected to the North Macedonia's ban for single-use plastic bags [30]. Figure 4b shows the trends of the most common topics identified in news reports. There are fewer identified topic within the news reports data compared to the twitter data. The topics include: waste landfills and waste management in Skopje, the "plastic waste pandemic", medical waste and its management by the municipalities, plastics bags for packaging and space debris. The trends for the previously mentioned topics show bigger peaks during the beginning of December and toward the end of January, corresponding to the two aforementioned news events about the plastic bags ban and vaccines disposal.

Lastly, Figure 5 shows the semantic similarity between the most common bigrams found in tweets and news reports over time. The figure shows trends for bigram similarity for all collected data during the week, as well as similarity between bigrams of tweets and news reports classified as positive or negative accordingly. The obtained results show that there is a strong semantic similarity between the most common topics over time, since the threshold for strong soft cosine similarity is 0.5 [28]. The biggest values for cosine similarity for the data while not taking into consideration the sentiment is 0.86, 0.81 and 0.79, accordingly corresponding

(a) Sentiment trends in tweets
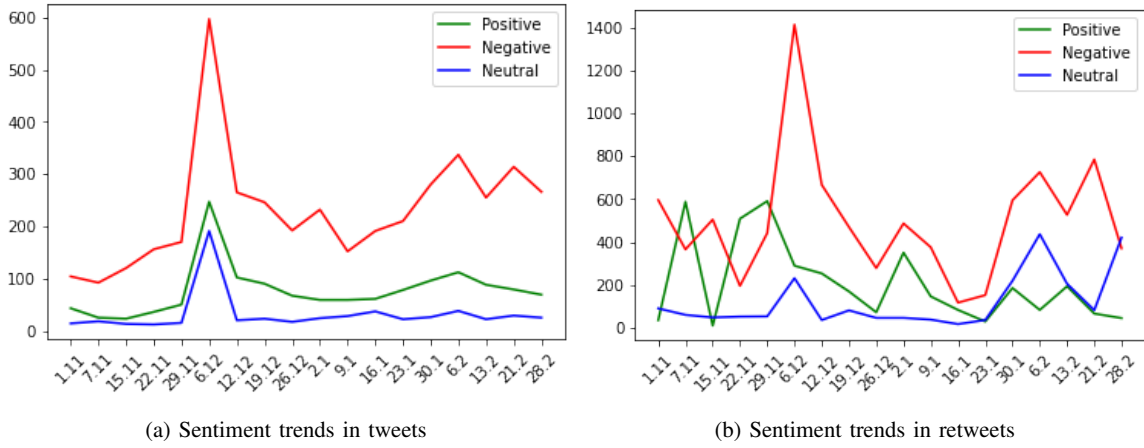

(b) Sentiment trends in retweets
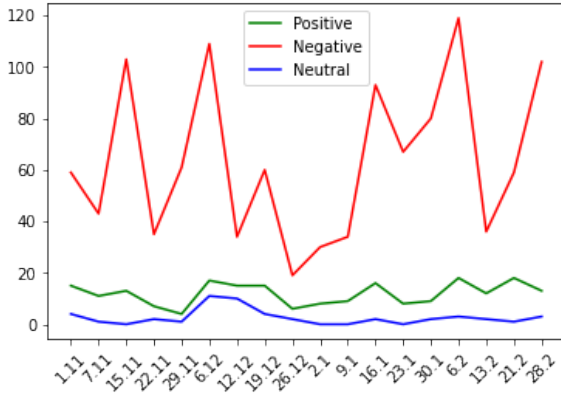
Fig. 2. Sentiment analysis

Fig. 3. Sentiment Analysis - News Reports

to the periods of end of February 2022, beginning of February 2022 and beginning of December 2021. The last two events are in correspondence with the previously mentioned news about plastic bags ban and vaccine disposal, whereas the first event with highest similarity value is connected to the war conflict between Russia and Ukraine, specifically the Russian occupation of Chernobyl and possible risks of major nuclear disaster and radioactive contamination across Europe [32]. Similar scores can be observed for cosine similarity of topics with negative sentiments, where the corresponding cosine similarity scores for the three major news events are 0.84, 0.83 and 0.78.

V. DISCUSSION AND CONCLUSION

The results obtained from all three analytic methods suggest strong correlation between the news reports and waste-related tweets which means that journalism has a major impact on the overall public response connected to waste management policies. Three major events were identified to have incited larger public response during the time frame between November 1st, 2021 and February 28th, 2022. The first event was the ban of single-use plastic bags in North Macedonia that began to be enforced as of December 1st, 2021 [30]. The main goal of these legal changes is to promote a new habit of using reusable bags made of eco-friendly materials. Despite the good intentions of the legal change, the results from the sentiment analysis show negative sentiment across both data sources. The main reason for this result is due to people's scepticism towards the governing bodies

for waste management. Namely, the vast majority of Twitter users during this period commented on the ineffectiveness of the ban due to the common practice of disposing mixed household waste in single-use plastic bags. Additionally, topics such as low recycling and waste selection rates as well as active burning of toxic waste were discussed during this event.

Another event that posed as a major initiator of public response was the great-scale COVID vaccines disposal due to expired dates in North Macedonia. In this instance, the majority of the Twitter users showed concerns about the management of medical waste, not only for the event of vaccine disposal, but also for the increased amount of medical waste generated over the last 2 years as a result of the COVID-19 pandemic [31]. The Twitter response about this issue is centered around the distrust of the waste managing facilities in the ability to adequately manage medical waste, since poor management of this type of waste can greatly threaten human and environmental health. The last identified provoking event was the armed conflict happening in Ukraine. The Russian invasion of Ukraine began on February 24th, 2022 and shortly after, the towns of Chernobyl and Prypyat were under Russian control. The major concern about nuclear disaster happened a few days after the initial invasion following the reports that the radiation level in Chernobyl drastically increased due to failure to maintain the nuclear reactor facility [32]. The Twitter response in the Western Balkans follow these concerns, mainly because of the geographical proximity of the affected region.

Over the last few years, the governments in the Western Balkan nations have made significant efforts in plotting a nation-wide plan to tackle the growing waste management issue. Last year, the North Macedonian Ministry of Environment and Physical Planning released a long-term strategy on climate action, which includes reports about waste generation trends as well as future plans for waste management [11]. Among the plans for building more efficient waste management systems, the North Macedonian government also has the intention of further incorporating the topics of climate change and sustainable development in the overall national education system, in addition to enforcing measures to raise climate awareness among the general public. The ban of single-use plastic is just one of these measures whose main intend is to limit the amount of municipal plastic waste and to raise the public awareness about waste management. Reducing

Global Topic Representation
0_bags_bag_garbage_plastic
1_serbia_waste_people_serbs
2_nuclear_nuclear waste_waste_power
3_waste_waste waste_just_turn
4_waste_plastic_container_containers
5_skopje_macedonia_garbage_hygiene
6_medical_medical waste_waste_masks
7_garbage_people_city_neighbor
8_throw_garbage_throw garbage_trash
9_wood_heating_burn_waste

(a) Topics over time - Tweets

Global Topic Representation
-1_waste_skopje_city_city skopje
1_plastic_plastic waste_waste_pandemic
2_medical_medical waste_pandemic_waste
3_bags_packaging_plastic_plastic bags
4_space_satellite_space debris_debris
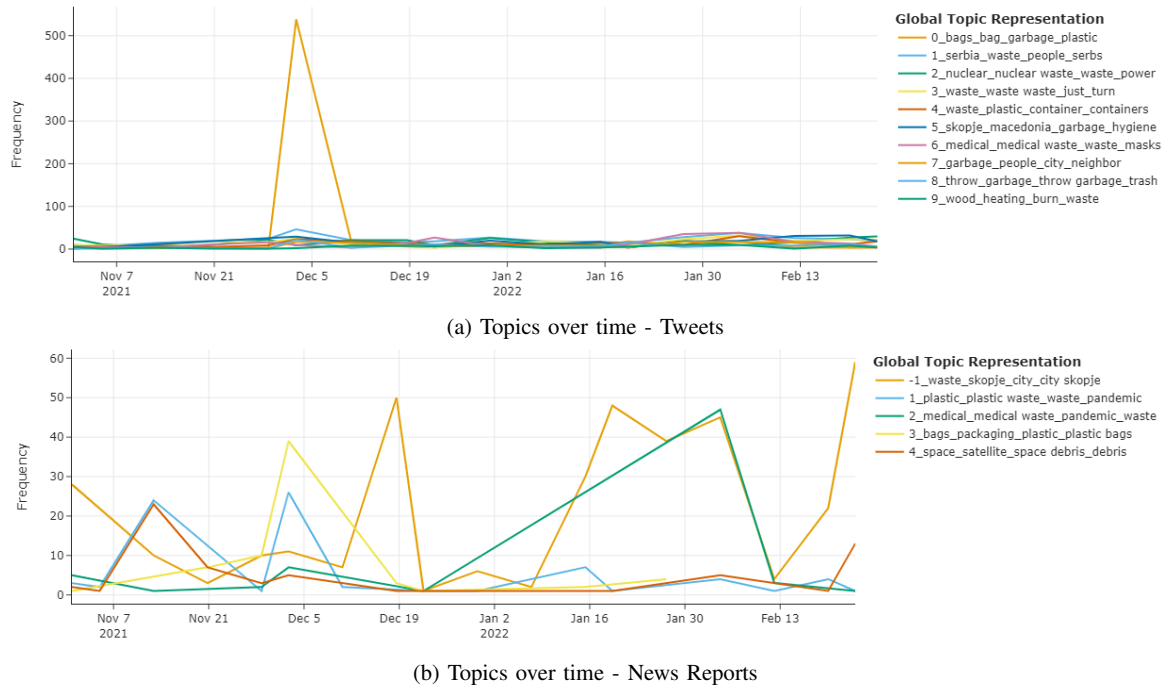
(b) Topics over time - News Reports
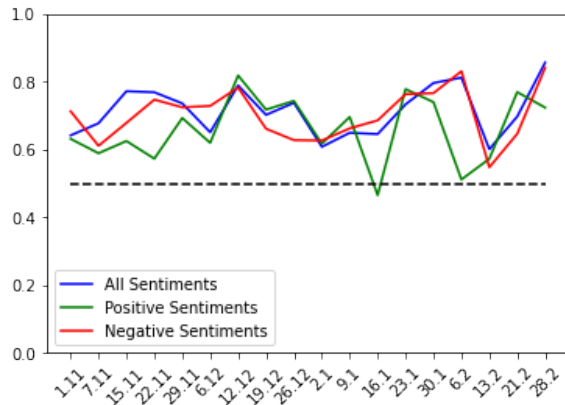
Fig. 4. Sentiment analysis - Tweets

Fig. 5. Soft Cosine Similarity between Tweets and News Reports

the amount of waste through waste selection and recycling are one of the main goals of the laws regulating waste treatment, therefore it would be convenient to promote the positive aspect of these measures initially by the news media outlets. Since there is a strong positive correlation between the news reports and the social media response concerning waste management, the positive news reports may play a major role in promoting pro-environmental attitude on social media, and consequently increase pro-environmental behavior in the general public. Additionally, analysis of social media attitudes regarding waste management could be used for assessing and improving the existing waste regulations as well as developing new approaches to promote pro-environmental behaviour.

All things considered, this paper presented methods for data collection and analysis of social media posts and news reports in order to evaluate public response related to topics about waste management. The results show that news reports greatly influence the public response on Twitter. Moreover, events associated with political and health-related topics are the main reason for the increased responses and discussions on social media in the Western Balkan region. Some possibilities for future work include analyzing public response of waste management schemes and events for a specific area of waste management, conducting a public survey to assess the level of awareness about waste management in Western Balkans, as well as collecting and analyzing data about public attitudes for waste-related topics over a longer period of time.

REFERENCES

[1] D. Hoornweg and P. Bhada-Tata, "What a Waste : A Global Review of Solid Waste Management," Worldbank.org, 2012, doi: http://documents.worldbank.org/curated/en/2012/03/16537275/waste-global-review-solid-waste-management.

[2] M. Kah, L. Levy and C. Brown (2012) Potential for Effects of Land Contamination on Human Health. 2. The Case of Waste Disposal Sites, Journal of Toxicology and Environmental Health, Part B, 15:7, 441-467, DOI: 10.1080/10937404.2012.736855.

[3] D. Porta, S. Milani, A. I. Lazzarino, C. A. Perucci, and F. Forastiere. "Systematic review of epidemiological studies on health effects associated with management of solid waste." Environmental health 8, no. 1 (2009): 1-14.

[4] M. Fujii, T. Fujita, X. Chen, S. Ohnishi, and N. Yamaguchi. "Smart recycling of organic solid wastes in an environmentally sustainable society." Resources, Conservation and Recycling 63 (2012): 1-8.

[5] G. E. Louis, "A historical context of municipal solid waste management in the United States." Waste management & research 22, no. 4 (2004): 306-322.

[6] "Waste statistics - Statistics Explained," Europa.eu, 2022. https://ec.europa.eu/eurostat/statistics-explained/index.php?title=Waste_statistics#Total_waste_generation[Accessed: 15-Mar-2022].

[7] Vonschoenberg.info. Waste Management in Montenegro, 2022. [online] Available at: https://vonschoenberg.info/avscfiles/uploads/2021/06/CMS_Montenegro_Fact_Sheet_final.pdf [Accessed: 15-Mar-2022].

[8] Vonschoenberg.info. Waste management in the Republic of Serbia, 2022. [online] Available at: https://vonschoenberg.info/avscfiles/uploads/2021/06/CMS_Serbia_Fact_Sheet_final.pdf [Accessed: 15-Mar-2022].

[9] Vonschoenberg.info. Waste management in North Macedonia, 2022. [online] Available at: https://vonschoenberg.info/avscfiles/uploads/2021/06/CMS_North_Macedonia_Fact_Sheet_final.pdf [Accessed: 15-Mar-2022].

[10] Р. Македонија, "Четврт национален план за климатски промени извештај за националниот инвентар на стакленички гасови" [Online]. Available: https://api.klimatskipromeni.mk/data/rest/file/download/3260224e5a2745c8597aa095f5f7f30f56c5e11b9b2085c5b68258a87c6f93d.pdf [Accessed: 15-Mar-2022].

[11] "Long-term strategy on climate action and action plan" [Online]. Available: https://unfccc.int/sites/default/files/resource/MKD_LTS_Nov2021.pdf. [Accessed: 15-Mar-2022].

[12] A. Licastro, and S. S. Bruno. "Drivers and barriers to a green economy. A review of selected balkan countries." Cleaner Engineering and Technology 4 (2021): 100228.

[13] Z. Wu, Y. Zhang, Q. Chen, and H. Wang. "Attitude of Chinese public towards municipal solid waste sorting policy: A text mining study." Science of the Total Environment 756 (2021): 142674.

[14] L.M. Heidbreder, M. Lange, and G. Reese. "# PlasticFree-July–Analyzing a Worldwide Campaign to Reduce Single-use Plastic Consumption with Twitter." Environmental Communication 15, no. 7 (2021): 937-953.

[15] J. Roesslein, Tweepy: Twitter for Python! URL: https://github.com/tweepy/tweepy. (2020).

[16] Online Doc Translator, 2021. [Online]. Available: https://www.onlinedoctranslator.com/en/. [Accessed: 1-Nov-2021].

[17] K. Andresen, A. Hoxha, and J. Godole. "New roles for media in the Western Balkans: A study of transitional journalism." Journalism Studies 18, no. 5 (2017): 614-628.

[18] L. Richardson. Beautiful soup documentation. April (2007).

[19] "Вести," Time.mk, 2021. [Online]. Available: https://time.mk/st/vesti. [Accessed: 1-Nov-2021].

[20] "Vesti," Time.rs, 2021. [Online]. Available: https://time.rs/. [Accessed: 1-Nov-2021].

[21] A. Tumasjan, T. Sprenger, P. Sandner, and I. Welpe. "Predicting elections with twitter: What 140 characters reveal about political sentiment." In Proceedings of the International AAAI Conference on Web and Social Media, vol. 4, no. 1, pp. 178-185. 2010.

[22] X. Zhou, X. Tao, J. Yong, and Z. Yang. "Sentiment analysis on tweets for social events." In Proceedings of the 2013 IEEE 17th international conference on computer supported cooperative work in design (CSCWD), pp. 557-562. IEEE, 2013.

[23] S. Gohil, S. Vuik, and A. Darzi. "Sentiment analysis of health care tweets: review of the methods used." JMIR public health and surveillance 4, no. 2 (2018): e5789.

[24] S. Bird, E. Klein, E. Loper. Natural language processing with Python: analyzing text with the natural language toolkit. "; Reilly Media, Inc." (2009).

[25] C. J. Hutto, & E. E. Gilbert. VADER: A Parsimonious Rule-based Model for Sentiment Analysis of Social Media Text , (2014). Eighth International Conference on Weblogs and Social Media (ICWSM-14). Ann Arbor, MI, June 2014.

[26] M. Grootendorst. BERTopic: Leveraging BERT and c-TF-IDF to create easily interpretable topics (Zenodo v0.9.4). Zenodo v0.9.4. (2020) doi:10.5281/zenodo.4381785

[27] F. Pedregosa, G. Varoquaux, A. Gramfort, V. Michel, B. Thirion, O. Grisel, M. Blondel et al. "Scikit-learn: Machine learning in Python." the Journal of machine Learning research 12 (2011): 2825-2830.

[28] G. Sidorov, A. Gelbukh, H. Gómez-Adorno, and D. Pinto. "Soft similarity and soft cosine measure: Similarity of features in vector space model." Computación y Sistemas 18, no. 3 (2014): 491-504.

[29] J. Pennington, R. Socher, and C. D. Manning. "Glove: Global vectors for word representation." In Proceedings of the 2014 conference on empirical methods in natural language processing (EMNLP), pp. 1532-1543. 2014.

[30] N. Georgievski, "Single-use plastic bags banned in North Macedonia, biodegradable ones sell at €0.25," Meta.mk, Dec-2021. [Online]. Available: https://meta.mk/en/single-use-plastic-bags-banned-in-north-macedonia-biodegradable-ones-sell-at-e0-25/ [Accessed: 23-Mar-2022].

[31] World, "Tonnes of COVID-19 health care waste expose urgent need to improve waste management systems," Who.int, Feb-2022. [Online]. Available: https://www.who.int/news/item/01-02-2022-tonnes-of-covid-19-health-care-waste-expose-urgent-need-to-improve-waste-management-systems#:~:text=Tens%20of%20thousands%20of%20tonnes,to%20a%20new%20WHO%20report [Accessed: 23-Mar-2022].

[32] P. Polityuk, F. Crellin, "Ukraine reports higher Chernobyl radiation after Russians capture plant," Reuters, 25-Feb-2022. [Online]. Available: https://www.reuters.com/world/europe/ukraine-nuclear-agency-reports-higher-chernobyl-radiation-levels-due-heavy-2022-02-25/. [Accessed: 23-Mar-2022]