TweEvent: A dataset of Twitter messages about events in the Ukraine conflict

Samyo Rode-Hasinger
Technical University of Munich
samyo.rode@tum.de

Matthias Häberle
Technical University of Munich
matthias.haeberle@tum.de

Daniel Racek
Ludwig-Maximilians-Universität München
daniel.racek@stat.uni-muenchen.de

Anna Kruspe
Technische Hochschule Nürnberg
anna.kruspe@th-nuernberg.de

Xiao Xiang Zhu
Technical University of Munich
xiaoxiang.zhu@tum.de

ABSTRACT

Information about incidents within a conflict, e.g., shelling of an area of interest, is scattered amongst different data or media sources. For example, the ACLED dataset continuously documents local incidents recorded within the context of a specific conflict such as Russia's war in Ukraine. However, these blocks of information might be incomplete. Therefore, it is useful to collect data from several sources to enrich the information pool of a certain incident. In this paper, we present a dataset of social media messages covering the same war events as those collected in the ACLED dataset. The information is extracted from automatically geocoded Twitter text data using state-of-the-art natural language processing methods based on large pre-trained language models (LMs). Our method can be applied to various textual data sources. Both the data as well as the approach can serve to help human analysts obtain a broader understanding of conflict events.

Keywords

Conflict, Ukraine, Dataset, Social Media, NLP

INTRODUCTION

Obtaining information about modern conflicts in the age of disinformation and fake news is increasingly difficult. Press outlets often have limited opportunities to report from war zones and therefore have to rely on official information from governmental sources. Additionally, information about conflict zones from non-governmental sources is not only interesting for press reporting, but also for research and for decision makers. An example of such an information source is the Armed Conflict Location & Event Data Project (ACLED, Raleigh et al. 2010)[1] which documents conflict incident data of conflict regions across the globe. In the resulting dataset, events are recorded using various sources and enriched with additional metadata such as location and conflict type (see section ACLED).

Since social media data is also globally available, this source could yield additional on-site information (Zhu et al. 2023). Unfortunately, geo-located social media data with precise coordinates is scarce (Kruspe, Häberle, et al. 2021). Therefore, we propose an algorithm to identify location information in multilingual Twitter text and match those tweets with ACLED incidents. To this end, we use state-of-the-art natural language processing (NLP) methods.

[1] https://acleddata.com/about-acled/

For the identification of location mentions, we employ the WikiNEuRal Named Entity Recognition (NER) model (Tedeschi et al. 2021) on ACLED event locations and Twitter text messages. To align the recognized geographic entities of both sources, the tagged locations are embedded via a Sentence-BERT language model (Reimers and Gurevych 2019), and the Cosine similarity of the feature vectors is calculated. Next, we embed the ACLED event descriptions and tweets' text using the same embedding model, and again compute similarities between both sources. Our approach is able to match event and tweet text and can add details to a specific event. One challenge here is the multilinguality of Twitter and event data. Twitter is itself highly polyglot (Mocanu et al. 2013), and Ukraine as the area of interest poses further difficulties for NLP methods due to the discrepancy between Russian and Ukrainian spelling of words in the Cyrillic script. Therefore, we employ highly multilingual variants of the selected language models (Reimers and Gurevych 2020). We publicly provide the resulting "TweEvent" dataset which contains conflict incident IDs derived from ACLED and corresponding tweet IDs for Twitter messages with high semantic similarity.

This algorithm and dataset can serve as a first step to identify additional event information hidden in social media data such as tweets. Twitter threads, associated replies, or linked images and videos could help to clarify and correctly code the respective events reported in datasets such as ACLED and further help to improve their quality, as misinformation continues to be an issue during the coding process (Miller et al. 2022). If social media discourse only is of interest, our dataset can also be used without the corresponding ACLED data as it assembles sets of Twitter messages related to the same events.

RELATED WORK

Several sources collect conflict-related text data, but often exclude social media data. One of these, the Uppsala Conflict Data Program Georeferenced Event Dataset (UCDP GED) (Sundberg and Melander 2013) reports on events on armed violence, similar to ACLED. Neither this dataset nor the monthly candidate dataset release by the same team (Hegre, Croicu, et al. 2020) contain social media sources. Sacco and Bossio 2015 argue that social media has become a key source of information and can complement traditional media when covering conflicts. Dowd et al. 2020 show that tweets capture events of political violence during the Kenyan elections in 2017. In their work, they conduct a manual matching strategy to assign events captured by tweets to ACLED events and find a clear overlap between the two. Steinert-Threlkeld et al. 2022 classify Twitter images to better understand protest dynamics and violence. They specifically highlight the valuable contribution social media can make in understanding sub-national conflict.

When considering social media sources only, several crisis-focused datasets have been released, such as Crisisbench (Alam et al. 2021). The authors compiled various datasets for crisis-related research and classification, including CrisisLex (Olteanu et al. 2014), CrisisNLP (Imran, Mitra, et al. 2016) and ISCRAM2013 (Imran, Elbassuoni, et al. 2013). The dataset mainly contains social media data from natural disaster events such as hurricanes and earthquakes, which was first collected via keyword/hashtag filtering and then labeled by human assessors. Due to the relatively isolated nature of these events, no additional geographic filtering or similarity matching was necessary. Kruspe, Kersten, et al. 2020 provide an overview over manual and automatic methods to detect relevant social media messages in crisis situations. In some cases, detecting previously unseen developments may be of particular interest (Kruspe 2020; Kruspe 2019).

With regard to the Russia-Ukraine war, Chen and Ferrara 2022 and Haq et al. 2022 have published worldwide Twitter datasets covering the public discourse directly after the outbreak, but do not consider geographic information and only cover the first few weeks. Park et al. 2022 examine fake news and disinformation campaigns by Russian media outlets on Twitter and VKontakte immediately before and during the war using their newly released dataset. Fung and Ji 2022 have released a Weibo dataset covering discourse on the war through a keyword search.

DATASETS

In this section, we introduce the two datasets used in this study. First, we present the collection and filtering processes for our own Twitter dataset. In the subsequent paragraph, we present the existing ACLED dataset in more detail.

Twitter Dataset

Our Twitter data collection has been running since 2018 using the free 1% stream of the Twitter API and covering the whole world. We restrict the sample to tweets containing geoinformation using the Filter API (see e.g. Pfeffer et al. 2022 for a description of the different Twitter APIs). Each tweet object consists of several metadata attributes such as username, creation time, geoinformation and the actual tweet text. We exclude any messages not originating from Ukraine by keeping only tweets with the `country_code` attribute set to "UA" (Ukraine). The resulting Ukrainian subsample includes 7.9M tweets from the timeframe of January 23rd, 2018, to October 12th, 2022.

ACLED

The Armed Conflict Location and Event Dataset (ACLED), officially introduced in 2010 (Raleigh et al. 2010), covers events in conflicts across the globe from 1997 onwards, and has been highly relevant for conflict research ever since (Hegre, Metternich, et al. 2017, Donnay et al. 2019). It reports events of political violence and protests with information on date, location, involved actors, fatalities, and types of violence for each event. The dataset is collected by an experienced team of researchers drawing on information acquired from newspaper articles, government and NGO reports, and partner organizations' social media. After a careful reviewing process, new events are added to the dataset on a weekly basis (i.e. coding).

At present, social media data is not part of ACLED on a large scale due to its varying quality[2]. Instead, accounts are vetted in advance and social media posts are only drawn from these few pre-selected accounts. In this work, we seek to augment this process by introducing an algorithm which is able to automatically detect social media posts that are related to a specific event. The algorithm could simplify the process for humans in the loop to find and draw on all relevant social media posts during the coding process.

We match social media data to ACLED data from Ukraine for the timeframe from January 1st, 2018, to November 11th, 2022, with a total number of 80,365 events (downloaded on November 21st, 2022). In the following sections, we introduce our matching algorithm and the resulting dataset in more detail.

DATASET CREATION METHOD

The aim of the proposed algorithm is to add related information collected from social media platforms, e.g., Twitter, to a conflict dataset such as ACLED. The algorithm is capable of detecting i) matching conflict locations and ii) a high agreement between texts acquired from conflict datasets and social media text messages. A dataset can then be constructed out of the matched information.

Figure 1 shows the dataset creation pipeline. First, we feed the ACLED event dataset and the described Twitter dataset into the system. We preprocess the tweets' text by replacing URLs with the 'HTTPURL' token, e-mail addresses with the 'EMAIL' token, and user mentions with the '@USER' token. We also delete the '#' sign from hashtags and convert emojis into their corresponding string shortcodes.[3] The successive main steps of the framework involve Named Entity Recognition (NER) of text data, temporal filtering and location matching, and sentence similarity event matching.

Named Entity Recognition NER is the task of identifying semantic elements occurring in a given text and assign them to pre-defined categories such as 'person', 'company', or 'location'. In this work, we use the multilingual WikiNEuRal model (Tedeschi et al. 2021) which supports 9 languages.[4] The NER model is capable of detecting named entities in languages other than English, e.g., Russian. Since we are interested in location mentions in social media posts, the tweets' preprocessed text fields are scanned for location entities (with LOC tags).[5] Tweet texts without LOC entities are discarded. We also apply the NER model to the 'location' field in the ACLED crisis events dataset. The latter is needed to obtain location entities consistent with those found in the social media posts.

Temporal Filtering and Location Matching In order to limit the amount of possible events in the ACLED dataset for the following matching step, we first apply a temporal filtering method by specifying a time window of ±3 days. We assume that, in a war situation, many messages with repetitive content are disseminated over a long period of time, as events such as shelling, bombardments, air raid warnings, etc. are constantly repeated. Therefore, it seems reasonable to consider tweets in a certain timeframe around a given event, but exclude messages with similar content which may have occurred at a different time within an ongoing conflict. We then use a pre-trained multilingual Sentence-BERT (SBERT) model (Reimers and Gurevych 2019; Reimers and Gurevych 2020) to generate a vector representation $\mathbf{v}_L \in \mathbb{R}^d$ of each LOC entity string found in the tweets' text and in the event location information present in ALCED. Specifically, we employ the `paraphrase-multilingual-mpnet-base-v2` model

[2]ACLED (2019). Armed Conflict Location & Event Data Project (ACLED) Codebook, 2019.
[3]https://pypi.org/project/emoji/
[4]https://github.com/Babelscape/wikineural
[5]The NER model is case-sensitive, which is why we do not apply lowercasing to the input text.

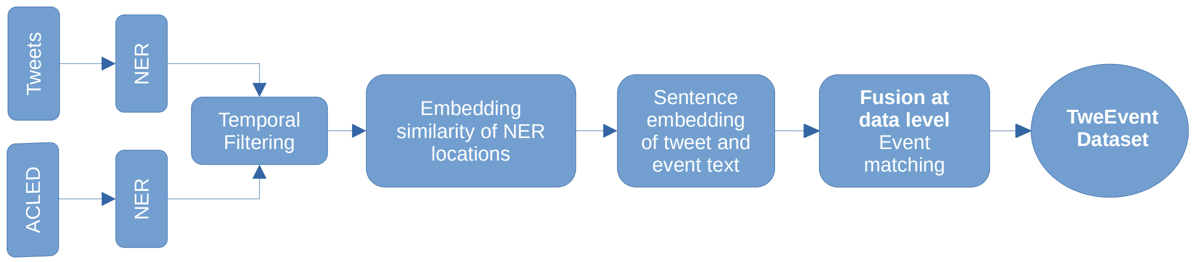

Figure 1. Workflow of our event matching algorithm.

(with $d = 768$) obtained from the SentenceTransformers library[6] in order to get semantically meaningful location embeddings for all given languages. Next, we compute the Cosine similarity between each embedded LOC entity in the tweets and the embedded ACLED event locations, and retain all tweets with similarity ≥ 0.7 per ACLED event as candidates for further processing.

Event Matching After obtaining a subset of tweets with matching timeframe and location for each event, we embed the full texts of both data sources, i.e. the ACLED 'notes' field and the tweets' texts. Again, we utilize the above-mentioned multilingual SBERT model as a text encoder. Each text is represented by a real-valued vector $\mathbf{v}_T \in \mathbb{R}^{768}$. Next, the Cosine similarity between the embedding vectors is calculated to estimate the relatedness of tweet-event pairs. Alam et al. 2021 stress the importance of the right choice of a Cosine similarity threshold for near-duplicate filtering in the context of social media data. We follow their approach and select tweets for a given event description with Cosine similarity ≥ 0.7 in order to include sufficient textual information that might be relevant.

Dataset Construction As a formalized expression of the matching step, let $e_i \in \mathcal{E} = \{e_1, e_2, ..., e_M\}$ be an embedded event in the ACLED conflict dataset of size M, and $t_j \in \mathcal{T} = \{t_1, t_2, ..., t_N\}$ be an embedded post in the social media dataset of size N. We construct our dataset $\mathcal{D} = \{(e_1, (t_1, t_2, ..., t_N)), ...(e_M, (t_1, t_2, ...t_N)\}$ such that $\cos(e_i, t_j) \geq \theta$, where θ is the Cosine similarity threshold parameter. Intuitively, each conflict event e_i is matched with 0 or more tweets t_j according to their semantic similarity (which is well represented in the embedding space).

Our dataset is publicly available under https://doi.org/10.14459/2023mp1703244.

# matched tweets total	20,491
# matched tweets unique	6,739
# matched events	7,500
Max. # tweets per event	136
Avg. # tweets per event	2.73

Table 1. Statistics of final TweEvent dataset with Cosine similarity ≥ 0.7.

DATASET STATISTICS AND ANALYSIS

In this section, we present some statistics and analyses of the achieved matching results, and discuss uncovered challenges. We obtain a final dataset with 7,500 unique events and 20,491 matching tweets in total. Table 1 shows the basic statistics of the dataset compiled with our proposed method and the Cosine similarity threshold parameter $\theta \geq 0.7$ for event-tweet pairs.

In Figure 2, we display the language distribution of the top 10 tweet languages in the final dataset according to Twitter's `lang` attribute. We observe that English is the most common language with a share of 37.94%, followed by German with 24.59%. Russian and Ukrainian are equally distributed with 13.04%. All other languages have a share of less than 3% each, including 'Undefined' with roughly 1%. The tag 'Undefined' is given to text messages where Twitter is not able to detect a specific language, mostly due to code-switching or too little textual information.

Figure 3 shows the distribution of the top 10 sub-event types according to ACLED within TweEvent dataset. Most tweets refer to the ACLED category 'Shelling/artillery/missile attack' with more than 50%. The rather vague event

[6] https://www.sbert.net/

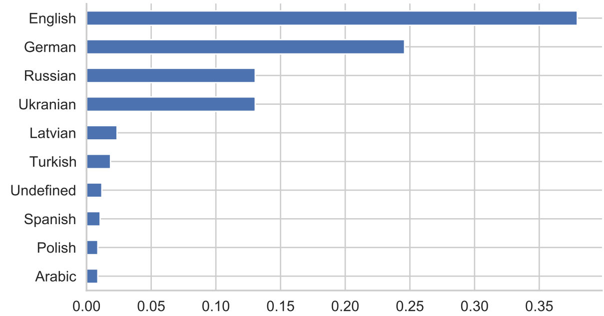

Figure 2. Distribution (ratio) of top 10 tweet languages in TweEvent with $\theta \geq 0.7$.

description 'Armed clash' is the second most frequent sub-event type with about 25%. 'Peaceful protest' makes up almost 7% and mainly refers to protesters in Ukraine. 'Air/drone strike' and 'Disrupted weapons use' have a share of 5.6% and 3.5%, respectively.

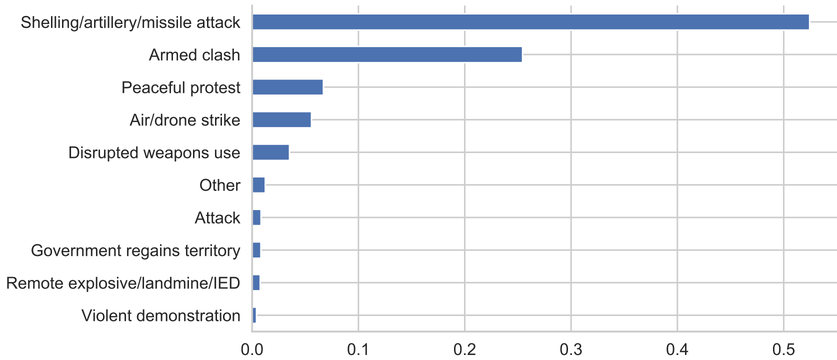

Figure 3. Distribution (ratio) of top 10 ACLED sub-event types in TweEvent with $\theta \geq 0.7$.

Regarding the locations mentioned in ACLED, we find a total number of 809 unique place names in the final dataset. Kharkiv, Mykolaiv, and Kyiv are the most frequent place mentions with 238, 183, and 164 in absolute numbers, respectively (see Fig. 4). It should be noted that some larger cities are divided into districts ("raions") in ACLED. For example, Kyiv, the capital of Ukraine, is mentioned 401 times in total with district information such as Shevchenkivskyi or Pecherskyi appended. Similarly, many ACLED events with matching Twitter messages took place in the city of Donetsk, located in eastern Ukraine. We find a total of 802 mentions here with Donetsk Airport (132) and Donetsk (129) being the most frequent ones. Further work may serve to align these related or nearby geographic entities to each other.

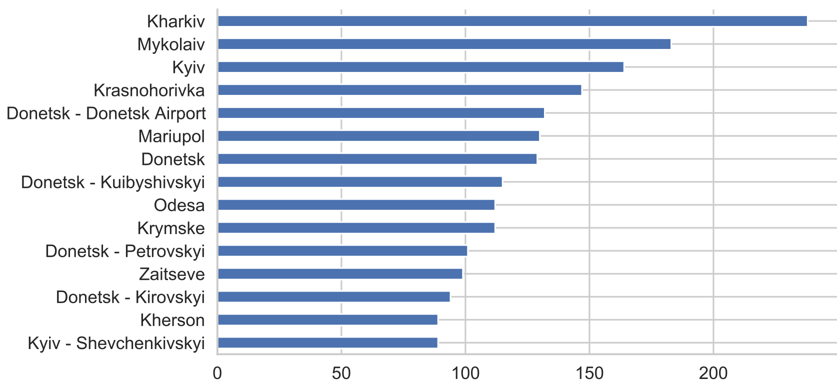

Figure 4. Number of 15 most frequent ACLED locations in TweEvent with $\theta \geq 0.7$.

Looking at the information sources for the events within ACLED, we identify various media outlets and governmental press services. The Ministry of Defence of Ukraine is the source of information with the most Twitter messages corresponding to its reports (about 20%). The source with the second-most aligned Twitter messages is DPR Armed

Forces Press Service with roughly 13%. The Donetsk People's Republic (DPR) was declared by pro-Russian separatists in the Donbas region in 2014 and recognized as sovereign state by Russia in 2022. 24 Channel, a Ukrainian news channel, is the third most frequent source of information with about 12%. Other sources include the Institute for the Study of War[7], an American research organization, with almost 3%, and the press center of the Ukrainian Joint Forces Operations headquarters (JFO HQ press centre) with over 3%. The JFO area denotes Donetsk and Luhansk regions in eastern Ukraine occupied by Russian forces. Further analysis of this data may reveal information about the trustworthiness of these sources (e.g., incorrect information about an event will likely not be corroborated by affected social media users).

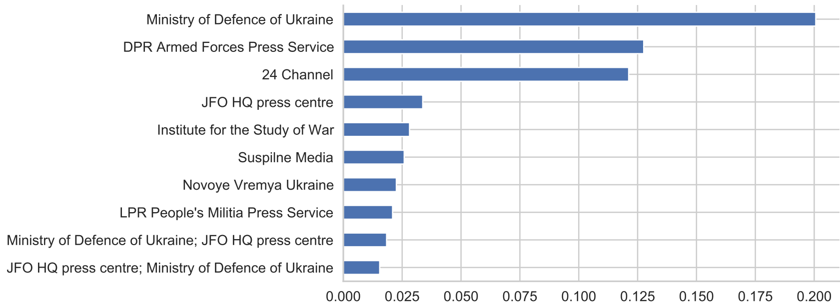

Figure 5. Distribution (ratio) of top 10 ACLED sources in TweEvent with $\theta \geq 0.7$.

Text Similarity Analysis

Table 2 shows ten English-language text examples selected via manual inspection. The first five rows contain examples where the matching algorithm yielded good results, whereas examples 6–10 show examples of possibly failed matches between events and tweets. These amounts are not representative of the whole dataset, but merely serve to illustrate strengths and weaknesses of our approach. We believe the matching capabilities of the algorithm are very promising. In example 1, the Twitter text provides a summary of the ACLED event. This match shows the capabilities of the applied sentence embedding model. Even a very short text can appropriately be represented by the embedding vector so that a high similarity score is achieved. In example 2, the ACLED text documents a protest in favour of a Ukrainian filmmaker imprisoned in Russia. The tweet text mentions that specific incident with fewer details, but with two links attached to the tweet. Subsequent analysis could crawl the linked websites to exploit additional materials such as photographs or news paper text, which could increase the information density of the event in question. Example 3 is about an attack on British journalists. The ACLED text briefly describes this incident while the tweet text offers additional insights: it mentions that the team of journalists identified itself as press. Not only can the algorithm match the correct event, but also add information valuable for documentation or for consecutive research. Example 4 is about a bombing with several civilian victims. The tweet mentions "northeastern city" as geo-spatial information as well as the possible composition of the victim group, once again adding details to the ACLED record. In example 5, the ACLED event and Twitter text both describe a rocket attack in Synelnykove Raion. The tweet text provides additional information about victims, which could be important for later analysis of the incident.

The negative examples 6–10 uncover challenges of the proposed algorithm. The two text paragraphs of example 6 have high similarity because the used vocabulary is similar, but a human reader is able to spot discrepancies of the content. The ACLED text documents a land mine explosion, while the tweet text mentions a bomb attack on a Ukrainian military vehicle. Here, the analysis of the attached web URL could reveal additional information to clarify if ACLED text and tweet refer to the same event. The texts of example 7 demonstrate the same challenge, even though region and attack vector are the same. However, the tweet provides additional information which may or may not relate to the same incident. The ACLED text and the tweet text of example 8 differ significantly in length and content. The ACLED text discusses detailed information about the incident, whereas the tweet text mentions shelling, but states the wrong city. Example 9 presents a similar challenge: the ACLED text refers to a drone strike, while the tweet text mentions that a single drone was repelled. Only the words drone and and the location are correct. Here, the size of the two text sequence might be the issue. Because the sentence embedding cannot embed much distinctive linguistic information, the matching words are "biasing" the outcome of the matching. The same can be seen in example 10, where only a few words coincide. These texts may be the results of two different events.

The "good" examples provide some evidence of the functionality and potential of our proposed algorithm. We could detect additional details about a certain incident or obtain summaries of it, as well as gain insights into personal

[7] https://understandingwar.org/

#	ACLED ID	Similarity	ACLED Text	Tweet
1	UKR5443	0.756	On 30 May 2018, in Kiev, Ukrainian authorities detained a Ukrainian citizen, Borys Herman, accused of being recruited by Russia's secret services to organize a murder plot against self-exiled Russian reporter and Kremlin critic Arkady Babchenko. Ukraine's Security Service (SBU) says it thwarted the planned killing by working together with Babchenko to fake his death. Ukrainian authorities also said a total of up to 30 people in Ukraine had been targeted for killing as part of the alleged Russian plot. Both SBU officials and Babchenko have defended their decision to fake Babchenko's death - rejecting criticism from reporters and journalism advocates who warned that it has undermined the credibility of law enforcement agencies and independent media organizations.	Security service of Ukraine: "we detained organizer of Babchenko's murder few hours ago in Kyiv"
2	UKR5185	0.814	On 24 May 2018, about 25 people protested in front of the Russian consulate general in Kharkiv, Ukraine, in support of the Ukrainian filmmaker Oleh Sentsov imprisoned in Russia. [size=about 25]	Rally in support of Oleg Sentsov and other Ukrainian political prisoners in Russia near Russian consulate in Kharkiv HTTPURL HTTPURL via @USER Ukraine
3	UKR52865	0.770	Russian military forces opened fire on British journalists of Sky News near Kyiv. One of the journalists was wounded.	The Sky News team are clearly & loudly identifying themselves as journalists but the rounds from the Russians keep coming. A 'professional ambush' of Russian forces of journalists on the ground in Ukraine.
4	UKR53066	0.765	On 8 March 2022, Russian air force dropped a bomb in Sumy, killing at least 22 people, including three children and four soldiers, and injuring around 20 people. Around 20 houses were destroyed.	Russian air strikes kill 22 people in Sumy overnight on March 8. Head of Sumy regional state administration Dmytro Zhyvytskyy said that three children were among those killed in the northeastern city. Russia war Ukraine
5	UKR69840	0.752	On 24 August 2022, Russian forces launched 3 rockets to Synelnykove district (coded to Synelnykove, Dnipropetrovsk region), destroying a private house and an infrastructure facility. A child was killed, a woman and another child were rescued under the rubble.	Recently Russians bombed Synelnykove Raion in Dnipropetrovsk Oblast. 11 years old girl died, many private houses are damaged . You can hear explosions on the video RussiaIsATerroristState RussianWarCrimesInUkraine GenocideOfUkrainians ArmUkraineNow
6	UKR33976	0.752	As reported on 4 March 2020, four Ukrainian soldiers were wounded in an explosion of a landmine that they were placing in the area of Krymske, Luhansk.	One Ukrainian soldier killed, 3 wounded in road side bomb explosion that targeted Ukrainian army BMP vehicle near Krymske HTTPURL Ukraine
7	UKR75504	0.767	On 3 October 2022, Russian forces launched a rocket attack at 2 villages in Zaporizhia district. There were no casualties.	@AFP This morning russians launched another rocket attack on the outskirts of Zaporizhzhia region. 16 rockets fired, 4 hit a convoy of cars with civilians who were going to pick up their relatives from the temporarily occupied territory. 25 dead, 50 wounded. RussiaIsATerroristState
8	UKR66629	0.768	On 3 August 2022, Russian forces shelled Krasnopillia community (coded to Krasnopillia), Sumy region, with 122 mm artillery and 120 mm mortars, the missiles hit near the railway station, a grain storage facility and town center. Casualties unknown.	russiaisateroriststate The russian army hit Kharkiv with cluster shells
9	UKR76077	0.777	On 24 September 2022, Russian forces conducted a drone strike using Iranian Shahed 136 drones on Odesa, Odesa region as a result of which one civilian was killed.	Ukrainian air defense shot down a Russian drone over Odesa
10	UKR64611	0.753	On 22 July 2022, Ukrainian forces shelled Dolomytne, Donetsk with 120mm mortars and 155mm artillery. Russian forces assaulted and were repulsed from Ukrainian positions. Casualties unknown.	During the day, the Armed Forces of Ukraine destroyed 4 ammunition depots and 3 Russian bases Destroyed warehouses and bases were located in Kadievka, Donetsk, Makiivka, and Horlivka. StopRussianAggression Ukraine Ukrainian UkraineWillWin UkrainianArmy GloryToUkraine

Table 2. Analysis of matched text pairs: Entries 1–5 denote good, 6–10 bad matching examples. ACLED ID and ACLED Text columns are cited from the ACLED dataset (Raleigh et al. 2010).

situations and perspectives. This information might enrich event datasets such as ACLED with additional insights, which might be valuable for processing conflicts. The negative examples, on the other hand, expose weaknesses of our approach: we may see superfluous matches whenever similar vocabulary is used in the tweet and the ACLED event text, but the semantic content or context does not fit or is u nclear. Those weaknesses mark spots to concentrate research. In some cases, a stricter matching of locations could be helpful; attack vectors could also be emphasized as a matching criterion. Furthermore, attached web URLs can be scraped and embedded. In many cases, the matching is somewhat subjective or fuzzy, also leading into future research opportunities (see section Future work).

It is important to note that the proposed method is currently a work in progress and presents some challenges due to the absence of reliable ground truth data. The determination of ground truth heavily relies on the specific use case, and in certain cases, it may not be readily discernible, thereby complicating the accurate evaluation of the method's performance. In addition, the amount of matched results depends on the similarity hyperparameter, which determines the threshold for data matching and how many social media posts will be included in the final dataset. Setting the threshold to a higher value results in less data points, while reducing the value leads to more, but sometimes fuzzy data. However, the latter may also reveal relevant information that was not previously observed. In our experiments, we used the given threshold parameter as it yielded promising results upon random manual inspection. Nevertheless, thorough investigation and assessment for each specific use case are necessary to fully understand the method's capabilities and limitations, particularly in terms of relevance and accuracy of the data acquired.

CONCLUSION

In this ongoing research, we present a dataset based on the data level fusion of ACLED event data and geo-referenced Twitter text messages, which is available under https://doi.org/10.14459/2023mp1703244. We use state-of-the-art natural language processing methods such as WikiNEuRal (Tedeschi et al. 2021) for named entity recognition and SBERT (Reimers and Gurevych 2019; Reimers and Gurevych 2020) for semantic textual similarity. These methods are utilized for the identification of locations in user posts (tweets) and for the matching of such locations by comparing their linguistic features via Cosine similarity. The same approach is used to compare the text of events documented in the ACLED dataset with tweets that were posted within a timeframe of three days around the ACLED event. The proposed approach is useful to support human-in-the-loop processes by adding possibly supplementary information and finally provide an invaluable resource for researchers and practitioners in the field of crisis management.

The results demonstrate that the proposed algorithm is not only capable of matching locations, but also identifying high text similarity scores of associated ACLED texts and tweets. In our study, we find that ACLED event descriptions and Twitter messages cannot only be matched, but we can also detect additional information about events, e.g. geospatial information or extra details of a specific situation (see section Text Similarity Analysis).

As in many applied machine learning research tasks within real-life scenarios, obtaining ground truth data is challenging. In this scenario, the correctness of matches is hard to determine for human assessors due to the limited textual information as well as the fuzziness of the question itself (i.e., what is an event?). As such, even when obtaining high Cosine similarity scores, an absolute conclusion whether a tweet is related to an event or not remains challenging and needs further assessment. However, the achieved results are promising and demonstrate that the proposed methodology is effective in building first steps towards a pipeline of mechanisms to match locations within social media data automatically and enrich conflict data with additional information from social media sources.

FUTURE WORK

For the Ukraine conflict, many tweets or data sources such as VIINA (Zhukov 2022) contain data in the Cyrillic alphabet. Therefore, a larger focus should be placed on more multilingual approaches to adapt to a multilingual world. Based on our approach, future work should conduct detailed analyses of multilingual text pairs in order to enhance the cross-language matching of event descriptions. In addition, future tasks could encompass a human-driven evaluation of the location matching approach and the exploration of different similarity levels. This includes, for example, the investigation of optimal hyperparameter thresholds and postprocessing pipelines for NER tags found in the texts. Also, some guidelines for a reasonable location matching could be valuable for the community. For example, is *Donetsk* and *Donetsk train station* a location match, even though we are dealing with different levels of geographic granularity?

In general, the findings from this research provide valuable insight into how people respond to crisis situations and how one could leverage this beneficial information. Potential applications of these findings could include the development of more targeted interventions for those in need, as well as improved strategies for managing and responding to crises at both an individual and organizational level. Additionally, the results may be used to inform policy decisions on how best to support individuals during times of distress or disruption so that they can access necessary resources in a timely manner. Furthermore, in light of increasing disinformation campaigns (Keller et al. 2020) and the inexorable spread of misinformation on social media platforms (Rode-Hasinger et al. 2022; Park et al. 2022), our dataset could encourage researchers to apply and investigate automatic detection algorithms to mitigate the threat of fake news in the context of crisis situations.

REFERENCES

Alam, F., Sajjad, H., Imran, M., and Ofli, F. (May 2021). "CrisisBench: Benchmarking Crisis-related Social Media Datasets for Humanitarian Information Processing". In: *Proceedings of the International AAAI Conference on Web and Social Media* 15.1, pp. 923–932.

Chen, E. and Ferrara, E. (2022). "Tweets in time of conflict: A public dataset tracking the twitter discourse on the war between ukraine and russia". In: *arXiv preprint arXiv:2203.07488*.

Donnay, K., Dunford, E. T., McGrath, E. C., Backer, D., and Cunningham, D. E. (2019). "Integrating conflict event data". In: *Journal of Conflict Resolution* 63.5, pp. 1337–1364.

Dowd, C., Justino, P., Kishi, R., and Marchais, G. (2020). "Comparing 'new'and 'old'media for violence monitoring and crisis response: evidence from Kenya". In: *Research & Politics* 7.3.

Fung, Y. R. and Ji, H. (2022). "A Weibo Dataset for the 2022 Russo-Ukrainian Crisis". In: *arXiv preprint arXiv:2203.05967*.

Haq, E.-U., Tyson, G., Lee, L.-H., Braud, T., and Hui, P. (2022). "Twitter dataset for 2022 russo-ukrainian crisis". In: *arXiv preprint arXiv:2203.02955*.

Hegre, H., Croicu, M., Eck, K., and Högbladh, S. (2020). "Introducing the UCDP Candidate Events Dataset". In: *Research & Politics* 7.3.

Hegre, H., Metternich, N. W., Nygård, H. M., and Wucherpfennig, J. (2017). *Introduction: Forecasting in peace research*.

Imran, M., Elbassuoni, S., Castillo, C., Diaz, F., and Meier, P. (2013). "Extracting information nuggets from disaster-Related messages in social media". In: *International Conference on Information Systems for Crisis Response and Management*.

Imran, M., Mitra, P., and Castillo, C. (May 2016). "Twitter as a Lifeline: Human-annotated Twitter Corpora for NLP of Crisis-related Messages". In: *Proceedings of the Tenth International Conference on Language Resources and Evaluation (LREC'16)*. Portorož, Slovenia: European Language Resources Association (ELRA), pp. 1638–1643.

Keller, F. B., Schoch, D., Stier, S., and Yang, J. (2020). "Political Astroturfing on Twitter: How to Coordinate a Disinformation Campaign". In: *Political Communication* 37.2, pp. 256–280. eprint: https://doi.org/10.1080/10584609.2019.1661888.

Kruspe, A. (Dec. 2019). "Few-shot tweet detection in emerging disaster events". In: *AI+HADR Workshop @ NeurIPS*.

Kruspe, A. (May 2020). "Detecting novelty in social media messages during emerging crisis events". In: *International Conference on Information Systems for Crisis Response and Management (ISCRAM)*.

Kruspe, A., Häberle, M., Hoffmann, E. J., Rode-Hasinger, S., Abdulahhad, K., and Zhu, X. X. (Nov. 2021). "Changes in Twitter geolocations: Insights and suggestions for future usage". In: *Proceedings of the Seventh Workshop on Noisy User-generated Text (W-NUT 2021)*. Online: Association for Computational Linguistics, pp. 212–221.

Kruspe, A., Kersten, J., and Klan, F. (2020). "Review article: Detection of actionable tweets in crisis events". In: *Natural Hazards and Earth System Sciences (NHESS)* Special Issue: Groundbreaking technologies, big data, and innovation for disaster risk modelling and reduction.

Miller, E., Kishi, R., Raleigh, C., and Dowd, C. (2022). "An agenda for addressing bias in conflict data". In: *Scientific Data* 9.1, pp. 1–6.

Mocanu, D., Baronchelli, A., Perra, N., Gonçalves, B., Zhang, Q., and Vespignani, A. (Apr. 2013). "The Twitter of Babel: Mapping World Languages through Microblogging Platforms". In: *PLOS ONE* 8.4, pp. 1–9.

Olteanu, A., Castillo, C., Diaz, F., and Vieweg, S. (2014). "CrisisLex: A Lexicon for Collecting and Filtering Microblogged Communications in Crises". In: *Proceedings of the International AAAI Conference on Web and Social Media*.

Park, C. Y., Mendelsohn, J., Field, A., and Tsvetkov, Y. (2022). "VoynaSlov: A Data Set of Russian Social Media Activity during the 2022 Ukraine-Russia War". In: *arXiv preprint arXiv:2205.12382*.

Pfeffer, J., Mooseder, A., Hammer, L., Stritzel, O., and Garcia, D. (2022). "This Sample seems to be good enough! Assessing Coverage and Temporal Reliability of Twitter's Academic API". In: *arXiv preprint arXiv:2204.02290*.

Raleigh, C., Linke, A., Hegre, H., and Karlsen, J. (2010). "Introducing ACLED: an armed conflict location and event dataset: special data feature". In: *Journal of peace research* 47.5, pp. 651–660.

Reimers, N. and Gurevych, I. (Nov. 2019). "Sentence-BERT: Sentence Embeddings using Siamese BERT-Networks". In: *Proceedings of the 2019 Conference on Empirical Methods in Natural Language Processing*. Association for Computational Linguistics.

Reimers, N. and Gurevych, I. (Nov. 2020). "Making Monolingual Sentence Embeddings Multilingual using Knowledge Distillation". In: *Proceedings of the 2020 Conference on Empirical Methods in Natural Language Processing*. Association for Computational Linguistics.

Rode-Hasinger, S., Kruspe, A., and Zhu, X. X. (Oct. 2022). "True or False? Detecting False Information on Social Media Using Graph Neural Networks". In: *Proceedings of the Eighth Workshop on Noisy User-generated Text (W-NUT 2022)*. Gyeongju, Republic of Korea: Association for Computational Linguistics, pp. 222–229.

Sacco, V. and Bossio, D. (2015). "Using social media in the news reportage of War & Conflict: Opportunities and Challenges". In: *The journal of media innovations* 2.1, pp. 59–76.

Steinert-Threlkeld, Z. C., Chan, A. M., and Joo, J. (2022). "How state and protester violence affect protest dynamics". In: *The Journal of Politics* 84.2, pp. 798–813.

Sundberg, R. and Melander, E. (2013). "Introducing the UCDP georeferenced event dataset". In: *Journal of Peace Research* 50.4, pp. 523–532.

Tedeschi, S., Maiorca, V., Campolungo, N., Cecconi, F., and Navigli, R. (Nov. 2021). "WikiNEuRal: Combined Neural and Knowledge-based Silver Data Creation for Multilingual NER". In: *Findings of the Association for Computational Linguistics: EMNLP 2021*. Punta Cana, Dominican Republic: Association for Computational Linguistics, pp. 2521–2533.

Zhu, X. X., Wang, Y., Kochupillai, M., Werner, M., Häberle, M., Hoffmann, E. J., Taubenböck, H., Tuia, D., Levering, A., Jacobs, N., et al. (2023). "Geo-Information Harvesting from Social Media Data". In: *IEEE Geoscience and Remote Sensing Magazine*.

Zhukov, Y. (2022). "VIINA: Violent Incident Information from News Articles on the 2022 Russian Invasion of Ukraine". In: Ann Arbor: University of Michigan, Center for Political Studies.