

Telegram Discourse as a Proxy for War-Related Event Detection in Israel

Yosef Meltser 205468283

Abstract—Social media platforms have become powerful open-source intelligence (OSINT) tools for real-time conflict monitoring. In this paper, we analyze two major Telegram channels—Abu Ali Express (Hebrew) and Al Jazeera Arabic—to investigate whether message activity and content trends can serve as proxies for war-related events in Israel. We present a multilingual analysis pipeline that scrapes messages, translates non-English text using Meta AI’s NLLB and Helsinki-NLP models, extracts topics with KeyBERT, identifies country co-mentions via NER, classifies sentiment using a Hebrew BERT-based model, and generates geospatial visualizations. Our results show that message surges align with key conflict events, while discourse patterns differ sharply across channels: Hebrew content emphasizes military alerts and tactical updates, whereas Arabic content foregrounds humanitarian themes. We further integrate rocket alert data from Israel’s Home Front Command to map the physical distribution of missile attacks. This cross-lingual, multi-source analysis highlights Telegram’s potential as a real-time sensor for both event detection and narrative framing in wartime.

I. INTRODUCTION

Major conflict events are often accompanied by bursts of online discourse as information is rapidly disseminated and discussed. In the context of Israel and its surrounding region, platforms like *Telegram* host popular channels that broadcast real-time updates during wars, military operations, and crises. Monitoring such channels offers a potential early-warning signal for emerging events, complementing traditional news and intelligence sources.

This paper investigates “Abu Ali Express,” a prominent Hebrew Telegram news channel, as a proxy sensor for war-related events in Israel. We hypothesize that spikes in message activity on this channel correspond to significant events (e.g., outbreaks of violence or major operations), and that analyzing the content of these messages can reveal the narrative and sentiment around the countries involved. To test this, we collected the channel’s messages over several years and applied natural language processing (NLP) techniques to extract temporal and semantic patterns. By correlating message volume trends with known events and examining country co-mentions and sentiment, we assess how well Telegram discourse reflects real-world conflict dynamics.

Our contributions are fourfold. First, we present a data collection and analysis pipeline for Telegram messages (Three groups), leveraging state-of-the-art multilingual NLP models to handle Hebrew and Arabic text. Second, we provide an empirical analysis of the temporal patterns in Telegram messaging around four major conflict events in Israel (from 2021 to 2023), demonstrating clear surges that align with

those events. Third, we conduct topic modeling using KeyBERT and clustering methods to extract and visualize high-level semantic themes from translated messages across both the Hebrew-language *Abu Ali Express* and Arabic-language *Al Jazeera* channels. This comparative topic analysis highlights divergent narratives—military-focused discourse versus humanitarian framing—revealing how different communities portray conflict through social media. Fourth, we enrich the analysis by integrating geospatial and temporal data on rocket alerts (“Red Alerts”) extracted from the official Telegram channel of Israel’s Home Front Command. This allows us to directly correlate public discourse patterns with the physical distribution and intensity of missile attacks across Israeli cities during the same periods.

II. RELATED WORK

Prior research has extensively explored platforms like Twitter for identifying emergencies, coordinating responses, and understanding public sentiment. For instance, Imran *et al.* [1] presented a comprehensive survey of computational methods for processing social media messages in mass emergencies, addressing key tasks such as filtering, classification, and summarization to support situational awareness. Kruspe *et al.* [2] provided an in-depth review of techniques for detecting actionable tweets during crisis events. Their work emphasized the evolving definitions of message relevance and informativeness, advocating for use-case-specific criteria to distinguish actionable content from noise in high-volume social streams. In the context of newer platforms, Taney *et al.* [3] demonstrated that Telegram messages can be used to detect and geolocate battle events during the Russo-Ukrainian war. Their system leveraged NLP techniques to extract armed conflict events.

Cross-lingual and multilingual NLP methods are also increasingly critical in crisis response. Chowdhury *et al.* [4] proposed a multilingual tweet classifier using BERT and Manifold Mixup, showing strong zero-shot transfer capabilities to unseen languages and disasters. Meta AI’s “No Language Left Behind” (NLLB) initiative [5] represents a milestone in massively multilingual machine translation. The NLLB-200 model supports translation across over 200 languages. This enables high-quality translation of social media messages from conflict zones in underrepresented languages. Despite these advances, relatively few studies have explored non-English Telegram discourse using multilingual pipelines for conflict monitoring. Our work addresses this gap by integrating

Hebrew and Arabic-language Telegram channels into a unified framework.

III. METHODS

A. Data Collection

We collected Telegram messages from three major public channels: *Abu Ali Express* (Hebrew), *Al Jazeera Arabic* (@AjaNews), and the official channel of Israel’s Home Front Command (@PikudHaOref_all). Each channel serves a distinct role in conflict-related communication, offering complementary perspectives on war-related events.

For **Abu Ali Express**, we used the Telethon Python library to programmatically scrape over 100,000 messages posted between 2017 and 2025. Each message included a timestamp, message ID, sender ID (if available), and message text. This channel primarily broadcasts updates and commentary during active military operations in Israel.

For **Al Jazeera**, we followed a similar scraping procedure focused on messages posted from 2019 onward. We filtered the Arabic-language messages using Israel-related keywords (e.g., "Tel Aviv" or "Israel") and retained only those relevant to the regional conflict context.

We also collected all public messages from the **Pikud HaOref** (Home Front Command) Telegram channel between 2019 and 2025. This channel issues real-time alerts during emergencies, especially rocket attacks. We filtered messages for relevant civil defense alerts using keyword matching (e.g., "Red Alert," "Rocket Fire," "Sirens Activated"), and extracted location names using regex-based parsing.

B. Time Series Analysis

To explore how Telegram discourse reflects real-world conflict dynamics, we analyzed the daily message volume from the *Abu Ali Express* Telegram channel between 2017 and 2025. Our hypothesis was that significant security events in Israel or Gaza would correlate with surges in message activity.

We aggregated the number of messages per day and computed the mean (μ) and standard deviation (σ) across the entire period. Any day with a message count exceeding $\mu + 2\sigma$ was flagged as anomalous. These anomalies were interpreted as potential indicators of real-world events.

To contextualize these spikes, we queried ChatGPT with the prompt: "What major events occurred on [DATE] in Israel or Gaza?" This investigation revealed clear alignment between message spikes and key military operations, including *Guardian of the Walls* (May 2021), *Breaking Dawn* (August 2022), *Shield and Arrow* (May 2023), and the *October 7 Attack* (October 2023).

These findings support the use of Telegram message volume as a real-time proxy for detecting war-related events. This temporal analysis forms the first of three core components in our broader framework, which also includes semantic topic extraction and sentiment-based geovisualization. For multilingual processing, all non-English messages were translated using pretrained machine translation models prior to downstream analysis.

C. Topic Modeling

To extract key semantic themes, we applied a topic modeling pipeline to both Hebrew (*Abu Ali Express*) and Arabic (*Al Jazeera*) messages. The process consisted of the following steps:

- 1) **Translation:** Messages were translated into English using either Meta’s NLLB-200 model (Hebrew) or the Helsinki-NLP opus-mt-ar-en model (Arabic).
- 2) **Keyword Extraction:** We used KeyBERT with the all-MiniLM-L6-v2 SentenceTransformer model to extract top keywords from each translated message.
- 3) **Clustering:** The most frequent keywords were embedded and clustered into 30 groups using KMeans, capturing recurring semantic themes.
- 4) **Dimensionality Reduction:** UMAP was applied to visualize clusters in 2D space, highlighting the distribution of topics and narrative focus in each channel.

D. Sentiment Analysis and Geovisualization

To capture emotional tone and geopolitical framing:

- **NER for Country Mentions:** We applied named entity recognition (NER) using spaCy to extract country names from the translated messages.
- **Sentiment Classification:** The original Hebrew messages were analyzed for sentiment using the HeBERT model, classifying each as Positive, Neutral, or Negative.
- **Co-Mention Mapping:** We constructed a co-mention network of countries mentioned in the same message, annotated with average sentiment values.
- **Geovisualization:** Using Folium, we visualized sentiment-weighted country relationships on a map, using colored edges to represent tone and line thickness for co-mention frequency.

E. Spatiotemporal Mapping of Red Alerts

In addition to analyzing online discourse, we extracted official alert data from the Israeli Home Front Command’s Telegram channel (@PikudHaOref_all), which provides real-time notifications of rocket attacks and civil defense updates.

Using the Telethon Python library, we scraped all messages from 2019 onward, filtering them for alert-related keywords such as "Red Alert," "Rocket Fire," and "Sirens Activated." For each relevant message, we extracted the names of affected cities or regions using heuristic line parsing and regular expressions.

After extracting geographic mentions, we grouped the alert data by month and location. We then used the Nominatim geocoder to assign each location its respective coordinates. The resulting dataset was used to create monthly geospatial visualizations, with alert frequency represented by marker size and opacity. This spatial layer enables direct comparison between discourse intensity and actual impact on civilians across time and regions.

IV. RESULTS AND ANALYSIS

A. Temporal Pattern of Telegram Activity

Telegram message activity shows clear surges corresponding to major conflict events in Israel and Gaza over the last few years. Figure 1 below plots the daily message volume in the channel over time, with annotations marking four notable events: the May 2021 *Operation Guardian of the Walls* conflict, the August 2022 *Operation Breaking Dawn*, the May 2023 *Operation Shield and Arrow*, and the surprise October 7, 2023 *Hamas attacks* on Israel. These events are known flashpoints that involved significant military action and public attention. As shown in the figure, each event aligns with a pronounced spike in daily messages. For instance, during Operation Guardian of the Walls (May 2021), the channel’s activity spiked dramatically as rocket barrages and airstrikes unfolded, reflecting intense public interest and rapid information spread. A similar surge is evident in early October 2023, when the Hamas attacks (and the ensuing war) generated a record volume of messages. This temporal correlation suggests that the Telegram discourse is highly sensitive to real-world conflict events, effectively acting as a sensor: abrupt increases in message volume flag that a major incident or escalation is occurring.

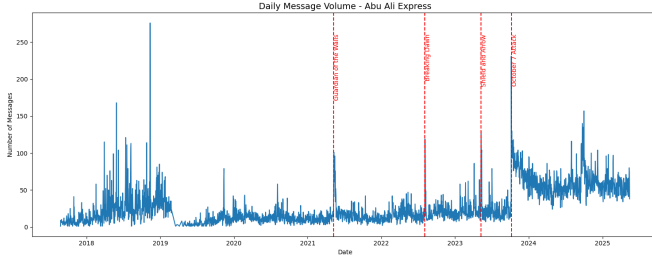


Fig. 1. Daily message volume over time for the Telegram channel (2017–2025), with four major conflict events annotated. Each labeled peak corresponds to a known war-related event in Israel/Gaza, during which message posting surged dramatically.

These temporal peaks align with historically documented military operations, supporting the hypothesis that Telegram channels can act as early indicators of conflict escalation.

B. Country Co-Mention Network and Sentiment Analysis

While the time-series analysis ties specific dates to events, we also examined the content of messages to understand the international context of the discourse. We focused on messages that mentioned at least one country, extracting pairs of countries that were mentioned together in the same message. By aggregating these co-mentions over the entire dataset, we constructed a network of countries, where an edge between two countries indicates they were co-mentioned in some Telegram messages. We then analyzed the prevalence and sentiment of each country pair. Table 1 summarizes the most frequently co-mentioned country pairs in the Telegram data and the majority sentiment of the messages in which they appear. The pairs are sorted by frequency of co-mentions (number of messages mentioning both countries).

TABLE I
TOP CO-MENTIONED COUNTRY PAIRS AND SENTIMENT

Country Pair	Mentions	Sentiment
Russia–Ukraine	134	Negative
Israel–Turkey	28	Negative
Israel–USA	6	Positive
Israel–Japan	3	Positive
France–Morocco	4	Neutral

From Table 1, we observe that the most frequent pair by far is Russia–Ukraine, which reflects how extensively the channel covered the Russo-Ukrainian war (134 messages discussed Russia and Ukraine together). The sentiment for this pair is overwhelmingly *negative*, unsurprising given that the context is an ongoing war with heavy losses and international tensions. The next most frequent pairs involve Israel in combination with various countries: Turkey (listed as “Turquía”), Morocco (“Maroko”), Russia, Brazil, Spain, etc., all of which also skew negative in sentiment. This trend likely stems from the channel discussing geopolitical repercussions or alliances around Israel’s conflicts – for example, Israel’s strained relations or incidents involving Turkey and Morocco drew negative-toned coverage. An interesting outlier is Israel–United States, which, although mentioned less frequently (6 co-mentions), has a *positive* prevailing sentiment. This suggests that when the US is mentioned alongside Israel, the context might be U.S. support for Israel (e.g., military aid, diplomatic backing), framed positively. Another relatively positive pair is Israel–Japan (a small number of mentions but positive sentiment), perhaps alluding to cooperative events or humanitarian aid. Meanwhile, European connections like France–Morocco show up with a neutral sentiment, indicating factual or mixed-tone reporting (e.g., news about diplomatic relations or events not strongly emotional). To visualize these relationships, we created a world map of the country co-mention network with edges colored by co-mention frequency (and annotated by sentiment polarity). Figure 2 depicts this map. Each line connects a pair of countries that were mentioned together, and the line’s color intensity corresponds to how many messages mentioned that pair (with more frequent pairs drawn in brighter colors). For clarity, only the prominent pairs are shown on the map. The map highlights, for instance, the heavy link between Russia and Ukraine (thick line) signifying the high volume of discussion around that war, as well as the multiple connections between Israel and other countries. Israel appears as a hub node connected to the United States, Turkey, Morocco, Russia, and others, mirroring Israel’s central role in the region’s conflict discourse. The sentiment dimension (indicated by line color hue or annotations) shows that most of these connections are discussed negatively (red or orange lines), which is intuitive given that conflict-related mentions (war, threats, attacks) carry negative connotations. The Israel–USA link is one of the few with a positive tint (green line), aligning with the Table 1 finding of positive sentiment, likely reflecting supportive statements or positive alliances. This country network analysis provides a macro-level view of the discourse. It re-

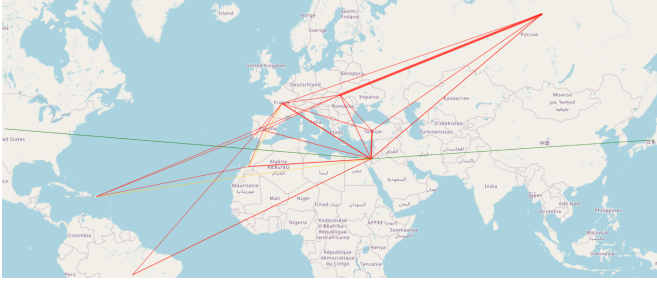


Fig. 2. Geospatial visualization of sentiment-weighted country co-mentions extracted from Telegram discourse. This map illustrates connections between country pairs that are co-mentioned within the same Telegram messages, based on multilingual analysis of conflict-related discourse. Lines are drawn between geocoded centroids of each country, with their color reflecting the dominant sentiment (red = negative, yellow = neutral, green = positive) and thickness scaled by co-mention frequency. The visualization reveals strong negative sentiment ties between countries involved in ongoing conflicts, as indicated by thick red lines.

veals which international relationships were most salient in the channel's coverage of conflicts. In particular, the prominence of Russia–Ukraine content indicates that the channel not only focuses on Israel's immediate conflict environment but also pays significant attention to major world conflicts. The strong connections between Israel and countries like Turkey and Morocco in the data could correspond to specific diplomatic crises or normalization deals making news. Overall, combining sentiment analysis with co-mention frequency gives additional context: not only do we see which pairs of countries are discussed, but we also understand the tone of that discussion, which is vital for interpreting public or media perception during war events.

C. Semantic Topic Analysis

To gain deeper insights into the thematic content of Telegram discourse, we applied a multilingual topic modeling pipeline composed of translation, semantic embedding, clustering, and dimensionality reduction. The objective was to extract and compare high-level discourse patterns from two distinct sources: *Abu Ali Express* (Hebrew) and *Al Jazeera* (Arabic). Translated messages were processed using the KeyBERT algorithm with a MiniLM-based SentenceTransformer model, allowing for the identification of the most representative keywords from each message. These keywords were then embedded into a high-dimensional semantic space, clustered using KMeans ($k = 30$), and visualized in two dimensions via UMAP.

a) Al Jazeera (Arabic):: The semantic space of Al Jazeera messages clustered around themes related to humanitarian narratives, resistance discourse, and civilian casualties. Frequent keyword clusters included concepts such as *martyrdom*, *siege*, *Gaza suffering*, and *international law*. This suggests that the Arabic-language discourse predominantly frames conflict through the lens of human rights violations, regional solidarity, and moral urgency.

b) Abu Ali Express (Hebrew):: In contrast, the topic clusters extracted from Abu Ali Express centered on tactical

and strategic discourse. Key themes involved military operations (e.g., *airstrike*, *operation shield*), threat alerts (e.g., *red alert*, *rocket siren*), and identification of hostile actors (e.g., *hezbollah responsibility*, *terrorist infiltration*). This reflects a narrative orientation focused on real-time situational awareness, military developments, and threat surveillance.

c) Comparative Insight:: The divergence in topic distributions highlights the different narrative frameworks adopted by each channel. While Al Jazeera emphasizes a humanitarian and resistance-focused framing, Abu Ali Express prioritizes operational and security-oriented discourse. This semantic contrast underscores how the same conflict events are interpreted through differing socio-political and linguistic lenses, shaping public perception in distinct ways. The UMAP visualizations in Figures 3 and 4 reinforce these findings by revealing spatially coherent topic clusters that reflect each channel's dominant discursive orientation.

Broader Semantic Topics (Clustered & Aggregated)

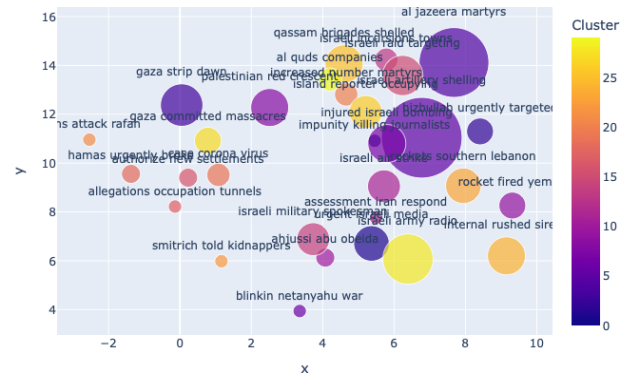


Fig. 3. Topic clusters from Al Jazeera (Arabic) channel. Topics emphasize martyrdom, humanitarian issues, and regional resistance.

Broader Semantic Topics (Clustered & Aggregated)

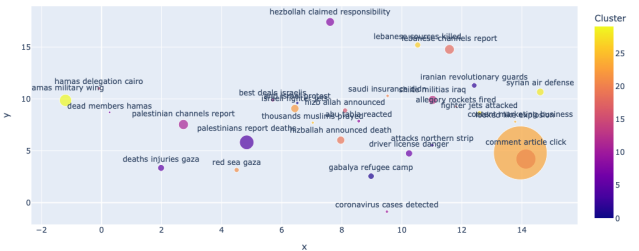


Fig. 4. Topic clusters from Abu Ali Express (Hebrew) channel. Topics emphasize military operations, alerts, and enemy identification.

D. Spatiotemporal Patterns of Red Alerts

To visualize the physical distribution of conflict intensity, we plotted monthly rocket alerts across Israeli territory based on official Home Front Command Telegram messages. Alerts were grouped by location and month, and visualized using

Folium, with circle radius and fill opacity proportional to alert frequency.

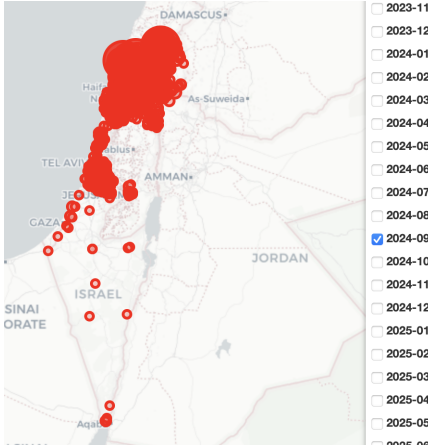


Fig. 5. Rocket alert density by city, October 2023. High concentrations appear near Gaza and central Israel during the October 7 attacks.

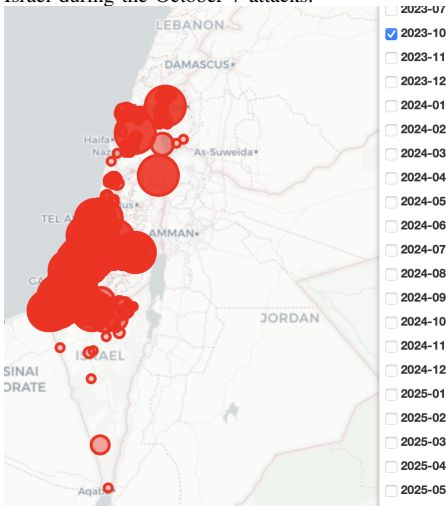


Fig. 6. Rocket alert distribution, September 2024. Alert activity shifts significantly northward, reflecting Hezbollah-related escalations.

The October 2023 map shows intense alert activity in the southern and central regions of Israel, reflecting the scale of the Hamas-led attack. In contrast, the September 2024 map shows concentrated activity in the north, consistent with reports of conflict with Hezbollah.

These geospatial patterns complement the temporal and semantic discourse analysis by offering a physical view of threat intensity. Furthermore, this data may inform future post-conflict recovery plans by identifying high-impact regions that could require psychological, infrastructural, or emergency services.

V. CONCLUSIONS AND FUTURE DIRECTIONS

This study demonstrates that Telegram discourse—exemplified by key Israeli and Arabic-language conflict news channels—can serve as an effective proxy for

detecting and interpreting war-related events. By tracking daily message volumes, we identified spikes that align closely with real-world conflicts, such as the May 2021 *Operation Guardian of the Walls* and the October 2023 Hamas attacks. These temporal peaks validate our hypothesis that message activity reflects conflict intensity and timing.

Beyond simple volume tracking, we applied a multilingual NLP pipeline incorporating machine translation, sentiment classification, and named entity recognition to uncover deeper insights. Country co-mention analysis revealed sentiment-laden international discourse, with most ties—especially involving Russia, Ukraine, and Middle Eastern actors—carrying negative tones. In contrast, some pairs like Israel–USA reflected positive framing, suggesting supportive alliances. The co-mention map visually reinforced these patterns, placing Israel at the center of a complex, emotionally charged network of regional and global relationships.

Topic modeling across both channels underscored a key narrative divergence: *Abu Ali Express* emphasizes military and tactical updates, while *Al Jazeera* frames the conflict through humanitarian and resistance lenses. This semantic contrast reflects differing political and cultural orientations in Hebrew and Arabic discourse, and highlights the value of comparative multilingual analysis.

To complement online discourse, we extracted spatiotemporal rocket alert data from Israel’s Home Front Command Telegram channel. This dataset allowed us to visualize conflict intensity by region and month, with clear geographic shifts—such as southward concentration in October 2023 and northward escalation in September 2024—mirroring real-world military dynamics.

Our contributions are fourfold: (1) a scalable multilingual pipeline for Telegram discourse analysis, (2) validation of message volume surges as indicators of conflict events, (3) semantic and sentiment-driven narrative comparison between channels, and (4) integration of physical threat data via rocket alerts.

Future Work. Several promising directions remain:

- **Multi-source Integration:** Expanding to more Telegram channels or additional platforms (e.g., X/Twitter, WhatsApp groups) could enable broader situational awareness.
- **Enhanced NLP Models:** Fine-tuned NER models for Hebrew and Arabic.
- **Image Analysis:** Many Telegram messages include visual content such as battlefield photos, propaganda, or damage reports. Incorporating image classification, object detection, or multimodal analysis could enrich interpretation of wartime discourse and help validate textual claims.

In conclusion, Telegram channels offer rich, real-time signals that reflect unfolding wartime realities. Our Israeli case study confirms that volume and content trends in public messaging can reveal both the structure and sentiment of conflict narratives. As modern warfare continues to unfold on social media, integrating multilingual discourse analytics with spatial and temporal data will be essential for future conflict monitoring and policy response.

VI. ACKNOWLEDGMENTS

I would like to acknowledge the assistance of ChatGPT 4o in refining the analysis and structuring the content of this paper, as well as the use of Google Colab for providing the computational infrastructure, including access to the A100 GPU, which facilitated the efficient processing of large-scale data.

VII. LINK TO GIT REPO

GitHub Repository

REFERENCES

- [1] M. Imran, C. Castillo, F. Diaz, and S. Vieweg, "Processing social media messages in mass emergency: A survey," *ACM Computing Surveys*, vol. 47, no. 4, pp. 1–38, 2015.
- [2] A. Kruspe, J. Kersten, and F. Klan, "Review article: Detection of actionable tweets in crisis events," *Natural Hazards and Earth System Sciences*, vol. 21, pp. 1825–1845, 2021.
- [3] H. Tanev, et al., "Detecting and Geocoding Battle Events from Social Media Messages on the Russo-Ukrainian War," in *Proceedings of the CASE Workshop*, 2023.
- [4] J. R. Chowdhury, C. Caragea, and D. Caragea, "Cross-Lingual Disaster-related Multi-label Tweet Classification with Manifold Mixup," in *Proc. of the 58th Annual Meeting of the Association for Computational Linguistics: Student Research Workshop*, pp. 292–298, 2020.
- [5] NLLB Team, et al., "No Language Left Behind: Scaling Human-Centered Machine Translation," arXiv preprint arXiv:2207.04672, 2022.