

Forced Displacement Situation Awareness with Large Language Models: Ukrainian Refugees in Poland Case Study

Brian Tomaszewski

Rochester Institute of Technology
bmtski@rit.edu

Nataliya Shakhovska

Lviv Polytechnic National University
nataliya.b.shakhovska@lpnu.ua

Harley Emery

University of Maryland
harleyem@umd.edu

Paweł Śniatała

Poznań University of Technology
pawel.sniatala@put.poznan.pl

ABSTRACT

We evaluated the utility of the large language model (LLM) ChatGPT to develop situation awareness related to the forced displacement of Ukrainian refugees into Poland. Utilizing text messages derived from the Help for Ukrainians in Poland Telegram message group, we used ChatGPT to translate messages in multiple languages into English and identify message topics and themes. Topics and themes from beginning of the war in Ukraine in 2022 were analyzed and visualized using K-means clustering and word clouds. The language identification and translation capabilities of the LLM were evaluated by two human evaluators and measured using Kappa (0.86) and BLUE scores (0.46) with the LLM performing effectively. We conclude that the ability of LLM using carefully developed language prompts for large data volume analysis with no need for manual human analysis shows promise for humanitarian analytics focused on rapidly identifying potential key trends, needs, and locations of displaced people.

Keywords

Ukrainian refugees, Poland, Large Language Model, ChatGPT.

INTRODUCTION

The unprovoked Russian invasion of Ukraine in 2022 has created one of the largest humanitarian crises in modern times. According to the UN, as of August 2022, 1,388,339 Ukrainian refugees have registered for temporary protection in Poland (United Nations High Commissioner for Refugees (UNHCR), 2022b) (Figure 1).

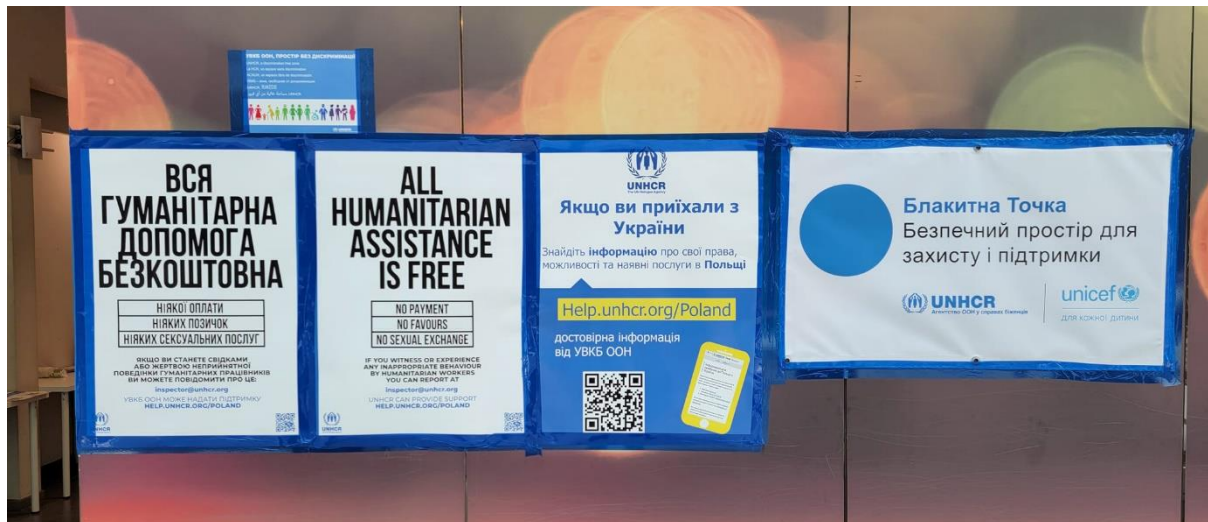


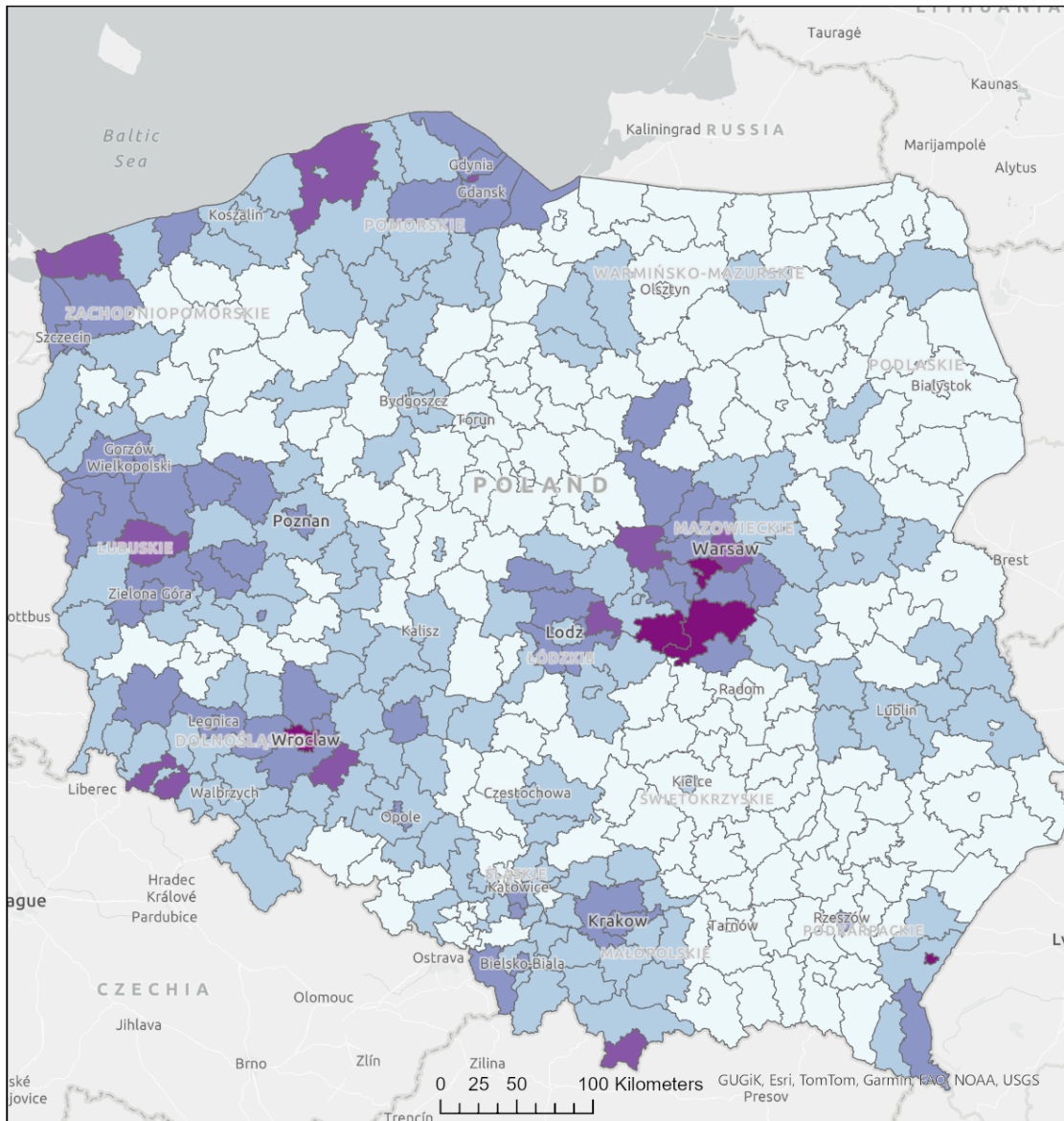
Figure 1. An example of 'Blue Point' refugee assistance center in Poznań, Poland (picture by author, 2022).

Although many Ukrainian refugees are returning home, with the war in Ukraine showing no end in sight, it is likely that many will remain in Poland for many years to come, as is typical in protracted forced displacement situations (UNHCR Regional Bureau for Europe Data Identity Management and Analysis Unit (DIMA), 2022).

International organizations that track forced displacement situations like the International Organization for Migration (IOM, UN Migration) monitor the situation of refugees once they arrive in a host country (c.f. International Organization for Migration, 2023b). In the case of the Ukrainian situation in Poland, it is important to know where and why refugees are settling in particular places in Poland and issues they are facing as there are several issues that may drive these decisions. For example, International Organization for Migration (2023a) outlines how in Warsaw, the main factors for why refugee population chose to settle in Warsaw include employment opportunities, accommodation availability and cost, and availability of health, educational, and other services.

Additionally, the general relationship between Polish society and Ukrainian refugees in Poland overtime is important to monitor. Poland received great praise from the international community for its support of Ukrainians in the first few months of conflict in 2022 in terms of social support, legal status, education and other factors (United Nations High Commissioner for Refugees (UNHCR), 2022a). However, as the war in Ukraine continues into the foreseeable future, some sources are signaling that Polish support systems have already, and will continue, to come under strain (Reidy, 2022).

A significant challenge exists for keeping track of the situations that individual displaced people may be experiencing once inside a host country. For example, official Polish government data sources on PESEL registrations (a system to track foreigners in Poland, see: (gov.pl, n.d.)) can give an indication of where refugees may be settling at the administrative level 2 in Poland (also known as a Powiat, similar to a county or district). Although this is a very useful starting point, refugees do not always reside in the location where they register for PESEL, making this data source not necessarily an accurate representation of what displaced people may be experiencing (Figure 2).



Legend

UKR PESEL Registrations as percent of Powiat Population (Equal Interval Classification)

- 0.003229 - 0.01707
- 0.01708 - 0.03092
- 0.03093 - 0.04476
- 0.04477 - 0.05860
- 0.05861 - 0.07245

Data:

October 2023 PESEL: <https://dane.gov.pl/en/dataset/2715/resource/51583/table>

2023 Powiat Population: <https://stat.gov.pl/en/topics/population/population/area-and-population-in-the-territorial-profile-in-2023,4,17.html>



Figure 2. Ukrainian PESEL registrations in Poland as of October 2022. Understanding of where Ukrainian Refugee populations in Poland are is a challenging problem due to data granularity and temporality issues. A map like this is often the best representation of refugee locations but often inaccurate as refugees frequently move after receiving a PESEL.

Besides knowing *where* placed people may be located, are the additional challenges of understanding *what* the issues are that displaced people may be experiencing. Social media has a long established precedent in ISCRAM-related research as one potential source of data that can be used for developing situation awareness about a given crisis (c.f. (Hughes & Palen, 2009; Power, Robinson, & Cameron, 2022)).

Generative AI is offering new promise for conducting social media analysis due to the lower entry barriers of accessing Large Language AI models (LLM) using human dialogue prompts. We argue that this approach presents a significant opportunity for non-technical specialists, such as humanitarian professionals, to utilize the power of LLM for social media analytics to ideally fill gaps in situation awareness knowledge about displaced people. A comparative study has been made among PaLM, BERT and GPT language models for topic categorization using AG News Dataset (Yerramreddy, Marasani, Gowtham, & Abhishek, 2023). However, the investigation for social analysis in Ukrainian language is not given. Thus, in this pilot work we investigated the question: What can LLM reveal about the Ukrainian forced displacement situation in Poland? To investigate this question, we conducted a pilot investigation to specifically investigate and evaluate the use of LLM tools to (1) translate social media messages to English and (2) classify translated social media messages by topic and theme to understand the Ukrainian forced displacement situation in Poland.

BACKGROUND

LLM like ChatGPT are perhaps one of the most significant technology paradigm shifts in recent years. The ability of the LLM to receive input via human dialogue and provide natural language responses is quickly demonstrating its ability in application domains ranging from business to education. In the humanitarian domain, AI language models can help address several challenges related to natural language processing (NLP)-centered analytics such as language translation and thematic analysis. NLP is generally unknown to humanitarians (Rocca, Tamagnone, Fekih, Contla, & Rekabsaz, 2023). Limited academic research has been conducted to date on the efficacy of LLM models for humanitarian applications. The literature that can be found on the use of AI models like ChatGPT for humanitarian applications has highlighted operational issues such as content creation for advocacy, supporting funding applications, legal document generation, and identifying support organizations (Crew et al., 2023; Roane, 2023). Tanner (2023) identified several issues identified from several humanitarian AI projects related to the broader application of artificial intelligence in humanitarian applications such as effective data strategies, the need for skilled data science professionals and ensuring proper data custody. However, the report explicitly stated that all the projects were conducted before the advent of LLM like ChatGPT.

METHODOLOGY

Figure 3 graphically outlines the workflow used to investigate what LLM can reveal about the Ukrainian forced displacement situation in Poland using social media.

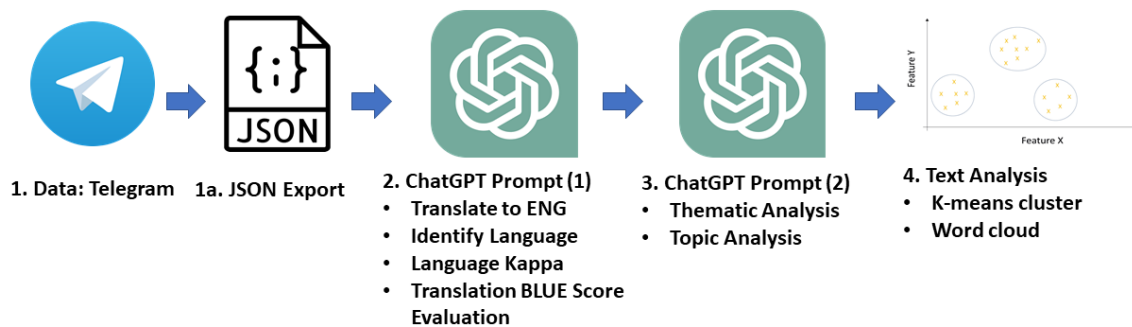


Figure 3. Ukrainian Refugee Social Media Analysis with ChatGPT work flow steps.

Step 1: Data Source

Telegram is a social media platform that is popular in Ukraine. In particular, the service has been extensively utilized by Ukrainians to discuss the ongoing war with Russia and as a tool for information-sharing among Ukrainian refugees in Poland. In the research presented here, we utilized the Telegram message community called Help for Ukrainians in Poland (UKR: Помощь украинцам в Польше) (Telegram, 2024). At the time when this research was conducted, this group had over 17,000 members with frequent daily posting (Figure 4).

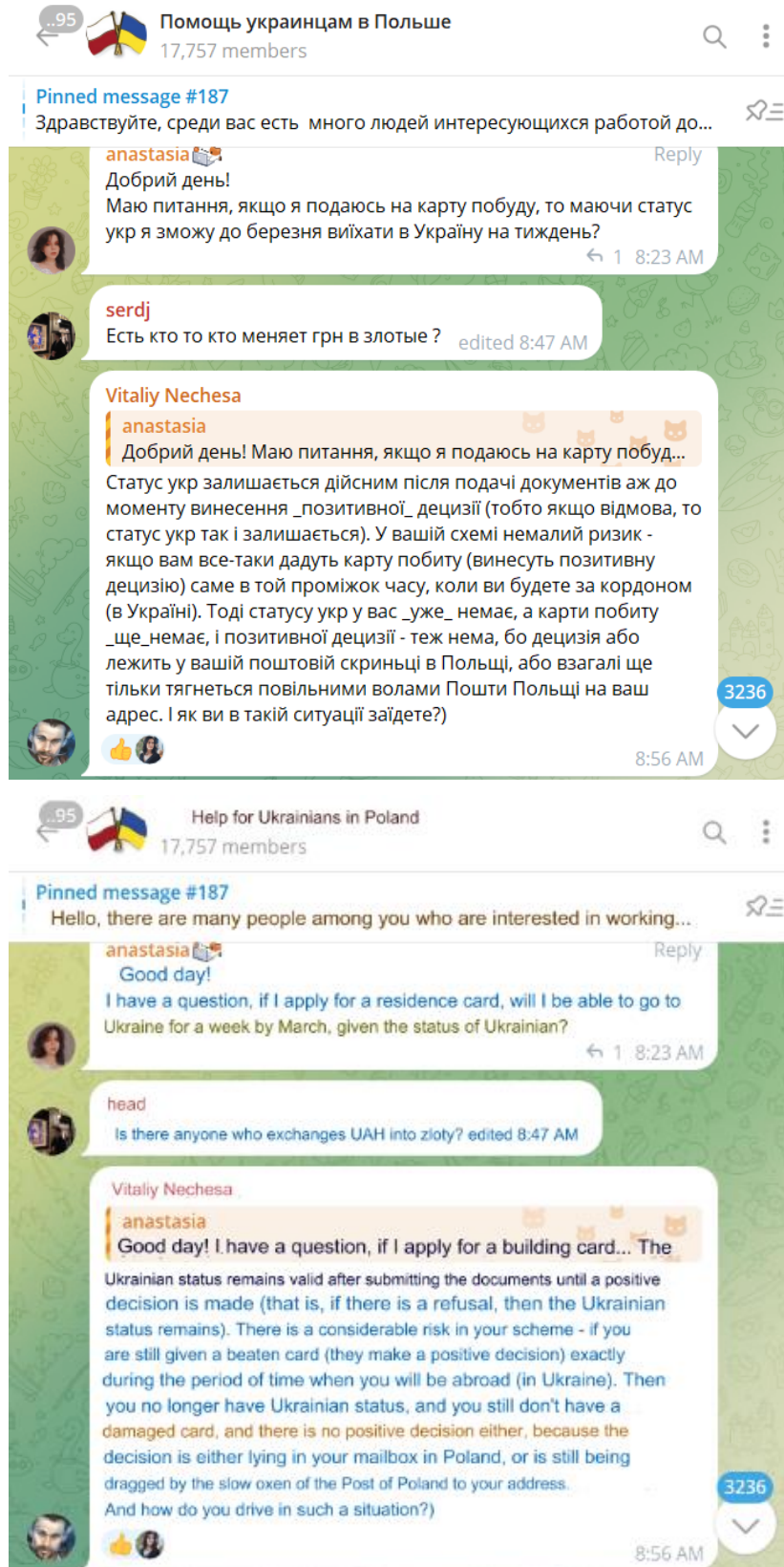


Figure 4. Screen shot of the Help for Ukrainians in Poland Telegram group. The image on the top is in the original source language, the image on the bottom is a Google image translation to illustrate the types of conversions conducted in the group.

Data from Telegram groups can be easily downloaded without requiring any special API programming like other social media platforms. Thus, we downloaded the entire collection data from this group dating back to 2022 at the beginning of the war up until January 2024. This yielded over 2.39 gb of data ranging from images to audio clips to text messages. In the work presented here, we focused only analysis of text messages that were provided in a JSON format as per step 1a seen in Figure 3.

Step 2: ChatGPT Prompt 1 – Message Translations and Language Identification

We utilized the gpt-3.5-turbo-0613 model via the OpenAI API in Python (OpenAPI, 2024). As this was an exploratory pilot project, this model was deemed sufficient vs other alternatives.

Perhaps the most important area of research with LLMs is prompt engineering, or how a human makes a natural language request of the AI model. Engineering plays a crucial role in designing AI systems that can understand and fulfill user queries accurately. By crafting prompts that are specific, well-structured, and tailored to the task at hand, engineers can guide the AI to produce results that meet user expectations. Beyond relevance, engineering also involves formatting the output of AI systems in ways that facilitate further analysis and processing. This includes organizing the information, presenting it in a structured manner, and ensuring compatibility with downstream tools and workflows. Careful prompt engineering is essential to ensure that the AI returns back relevant results as well as results that are formatted in ways that can be utilized for further analysis. By structuring prompts and results returned in JSON format, they can be easily manipulated and managed programmatically, allowing for automation and integration with other workflows. This can be particularly useful in scenarios where a large number of prompts must be generated or dynamic adjustment of prompt parameters.

In step 2 of the workflow, we were interested in (1) having the Telegram messages translated to English and (2) have the LLM identify what language it thought the message was in. Our motivation was twofold. First, English is generally the language used by the international humanitarian community, even though there are numerous humanitarian staff in Poland that speak Polish and Ukrainian. Therefore, we wanted to investigate how using LLM could enable a wider group of humanitarian professionals to access information in languages other than English. Second, people from Ukraine in many cases speak both Ukrainian and Russian, or primarily Russian. The messages provided in the Help for Ukrainians in Poland Telegram group board are posted in both a mix of Russian, Ukrainian and other languages. Therefore, we are interested to assess how well the language model was at identifying a particular language in order to build trust that results from the LLM were accurate.

The following prompts were used to receive both the English translation of an input text and the language that ChatGPT identified from its analysis:

PROMPT_PART_1 = "Translate this text to English <start> "

PROMPT_PART_2 = " <end>. Return back two things. The first is your translation to English of text that was between the <start> and <end> tags. The second is a one-word description of the language of text that was between the <start> and <end> tags. "

PROMPT_PART_3 = "Put the two items you return into a JSON structure. Your translation to English of text that was between the <start> and <end> tags placed inside a JSON tag named translation. Your one-word description of the language of text that was between the <start> and <end> tags inside a JSON tag named language. Do not return any additional text, descriptions of your process or information beyond two items and output format of the tags specified."

Telegram message texts was processed from the JSON file downloaded in Step 1A and placed inside of the start and end tags as per the aforementioned prompts. The use of <start> and <end> tags was used after several iterations of developing prompts as they were found to be effective for clearly defining what text chat GTP should be focused on for translations as opposed to output formatting instructions.

LLM models can also benefit from being given a motivation to help establish the task being asked to do. Therefore, the following motivation message was sent in along with the prompts.

MOTIVATION_MESSAGE = "You are a skilled linguist who is an expert in identifying what language a given text is and doing translations of that text to English."

Using these prompts, we processed over 20,000 messages into English dating from 2022 to 2023. All of the English translations along with the original Source message, date of the message ID of the message and identified language of the message were stored in a CSV file. This allowed for the next step in the workflow to conduct thematic and topic analysis on the English translations that we discuss in the next section.

Step 3: ChatGPT Prompt 2 – Thematic and Topic Analysis

Identifying themes and topic from text is a common natural language processing goal particularly in the application of developing situation awareness in an ongoing protracted crisis situation like Ukrainian refugees in Poland (Rocca et al., 2023). Therefore, in step 3 of our process, we investigated the ability of ChatGPT to identify topics and themes from the English language translations that were derived in step 2 and discussed previously. Our motivation here was to investigate how high level thematic and topic summaries of what was being discussed in social media could be derived for developing an initial situation assessment that could then be followed up upon for further deeper investigation and situation awareness at local levels as per the well-established ideas of the process of achieving situation awareness (Endsley, 1995).

The following prompts were used to receive to receive topic and thematic summaries of individual messages:

PROMPT_PART_1 = "Conduct a thematic analysis of this text <start> "

PROMPT_PART_2 = " <end>. Return back two things."

PROMPT_PART_3 = "The first is your general thematic assessment of text that was between the <start> and <end> tags. The second are any topics you found in the text that was between the <start> and <end> tags. "

PROMPT_PART_4 = "Put the two items you return into a JSON structure. Your general thematic assessment of text that was between the <start> and <end> tags placed inside a JSON tag named theme. Any topics you found in the text that was between the <start> and <end> tags inside a JSON tag named topic."

PROMPT_PART_5 = "Do not return any additional text, descriptions of your process or information beyond two items and output format of the tags specified."

Like step 2, English language message texts processed in Step 2 were placed inside of the start and end tags of the aforementioned prompts.

The following motivation was used to help establish the task that ChatGPT was being asked to do.

MOTIVATION_MESSAGE = "You are a skilled humanitarian analyst who is an expert in conducting thematic analysis of English language texts."

Topics and themes returned back from chatGPT as text were stored in a .csv file as an input to topics and theme text analysis in step 4 in the process that we discussed next

Step 4: Topics and Themes Text Analysis

Given the very large volume of unstructured text derived from step 3 thematic and topic analysis, we use the well-established methodology of K-means clustering to group similar topics and themes based on words in those messages to derive a summary representation of topics and themes (Kumbhar, Mhamane, Patil, Patil, & Kale, 2020). Specifically, topic and theme texts were first processed to remove stop words, or words with no significance like 'the' 'and' 'is', and converted into a TD-IDF vector, or the importance of a given word in a single Telegram message relative to the overall collection of messages. The optimal number of clusters were calculated using a silhouette score that determines how well the messages might cluster in general and the optimal number of clusters. K-means clustering, which will partition text found in the messages into similar and dissimilar clusters, was then performed using the optimal number of clusters derived from the silhouette score. The dimensionality of the clusters was then reduced using principal component analysis (PCA). PCA was performed to reduce the dimensionality of the data space (or reducing the number of words in a given cluster while preserving the most important words) to allow for graphical representation of clusters derived from K means processing. A graphical representation of the clusters derived by calculating the centroid of a given cluster based on the term with the highest TF-IDF score in that cluster. This term was then plotted to create a graphical representation of the cluster. Additionally, word cloud generations were generated as an additional source of graphical representation of the text analysis.

RESULTS

As discussed previously, a particular motivation of this work was to utilize LLM to conduct language translations so that content derived from social media could be more widely accessible to a wider range of humanitarian professionals who may not in fact speak Polish, Ukrainian, or Russian. Therefore, we were interested in the effectiveness of the LLM to (1) specifically identify what language a given message was in and (2) the accuracy of the translation from a given language into English. Understanding the accuracy and effectiveness of such language tasks is important so that the AI can be considered trustworthy.

To facilitate these evaluations, two native speakers of Ukrainian and Russian with fluency of English were

recruited to perform the same tasks that the LLM conducted in terms of (1) identifying what language a given message was in and (2) do a translation of the message into English. These two evaluators worked on a sample of 108 messages. This sample size was deemed sufficient given the high expense of conducting human translations. Messages were selected for human evaluation with the criteria that the message contained no emojis or other non-text characters that could not be directly translated into English. Comparison of the human vs LLM of language identification identified using Cohen's Kappa and accuracy of the LLM translation to English evaluated using BLEU (Bilingual Evaluation Understudy) score are discussed next.

Language Identification Kappa

Cohen Kappa is a well-established statistic to measure to assess the level of agreement between two evaluators (Viera & Garrett, 2005). In the work presented here, we were interested in comparing what language a given Telegram message was as rated by both a human evaluator and the LLM. The percentage of agreement between the human annotator and the LLM was 96.2% with a Kappa score of 0.86. These results indicate that the LLM was very effective at identifying what language a given Telegram message was across Russian, Ukrainian, Polish and other languages when compared to the same assessment done by a human.

Translation BLUE Score

BLEU is a very well established metric for automated evaluation of machine translations (Papineni, Roukos, Ward, & Zhu, 2002). English translations of a given Telegram message provided by the LLM were evaluated against two human translations of the same message using a Corpus BLUE score. A 4-gram precision was calculated with the weights 0.1, 0.1, 0.4, 0.4 to place higher emphasis on 3-gram and 4-gram precision (40% each) as these would be considered more useful for thematic and topic analysis and lower on 1-gram and 2-gram precision (10% each). The Corpus BLUE score result was: 0.46. Thus, it can be considered the LLM translations provided "High quality translations" (Google Cloud, n.d.).

K-MEANS TERM CLUSTERING AND WORD CLOUDS

Given the overall time frame of the data set that was processed, it was deemed more feasible to examine thematic and topic analysis based on particular time periods during the forced displacement of Ukrainians into Poland. Specifically, we performed K-means clustering on messages from the date range of 26 – 28 February 2022 within the overall data set processed by the LLM. This date range covers the very beginning of the Help for Ukrainians in Poland telegram group. It also coincides with the opening of the war in Ukraine including the first significant military attacks on large Ukrainian cities such as Kiev and Kharkiv causing thousands of people to begin fleeing to Poland (BBC News, 2022). Our goal was to do an initial assessment of themes that would be found in the data set. Figure 5a and 5b graphically show the results of K-means clustering around thematic analysis conducted by the LLM. Figure 6 shows a word cloud representation of thematic analysis conducted by the LLM. Figure 7 shows the results of a K-means clustering of the topic analysis conducted by the LLM. Figure 8 shows the results of a word cloud representation of the topic analysis conducted by the LLM. These figures are followed by a discussion of these results.

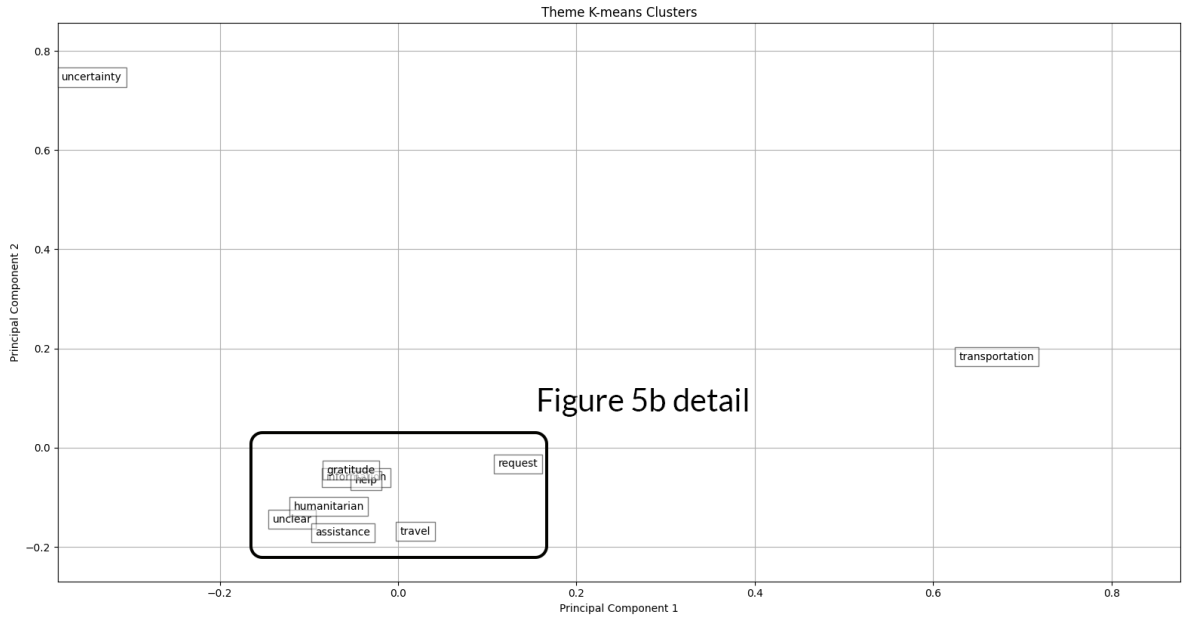


Figure 5a. Theme clusters – overview

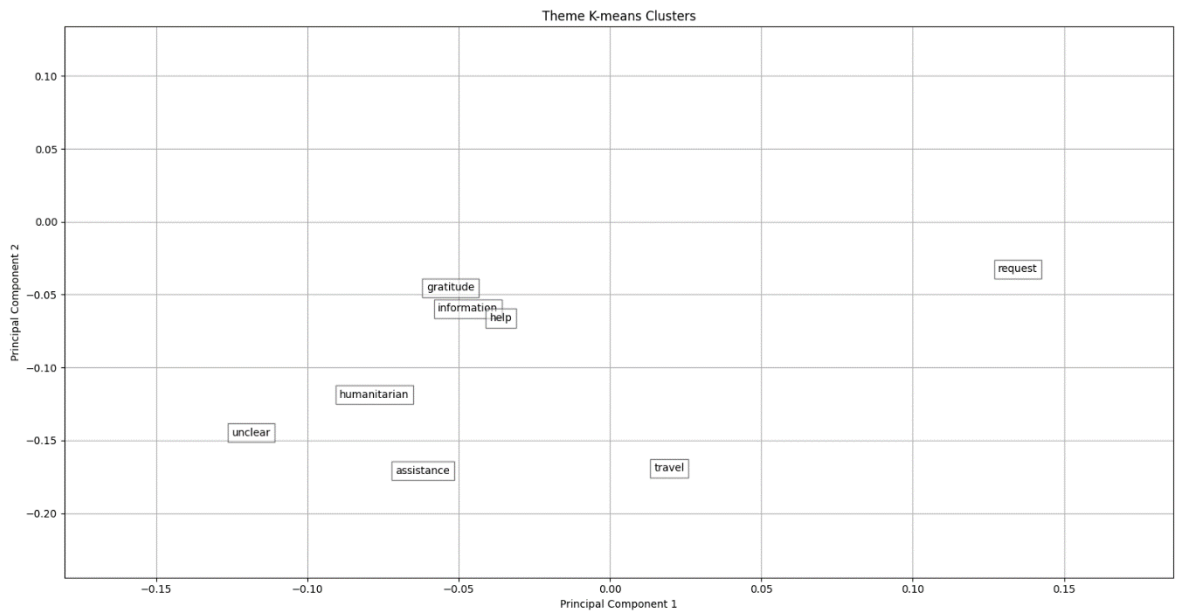


Figure 5b. Theme clusters - detail

Figures 5a and 5b. K-means Theme Clusters



Figure 6. Theme word cloud

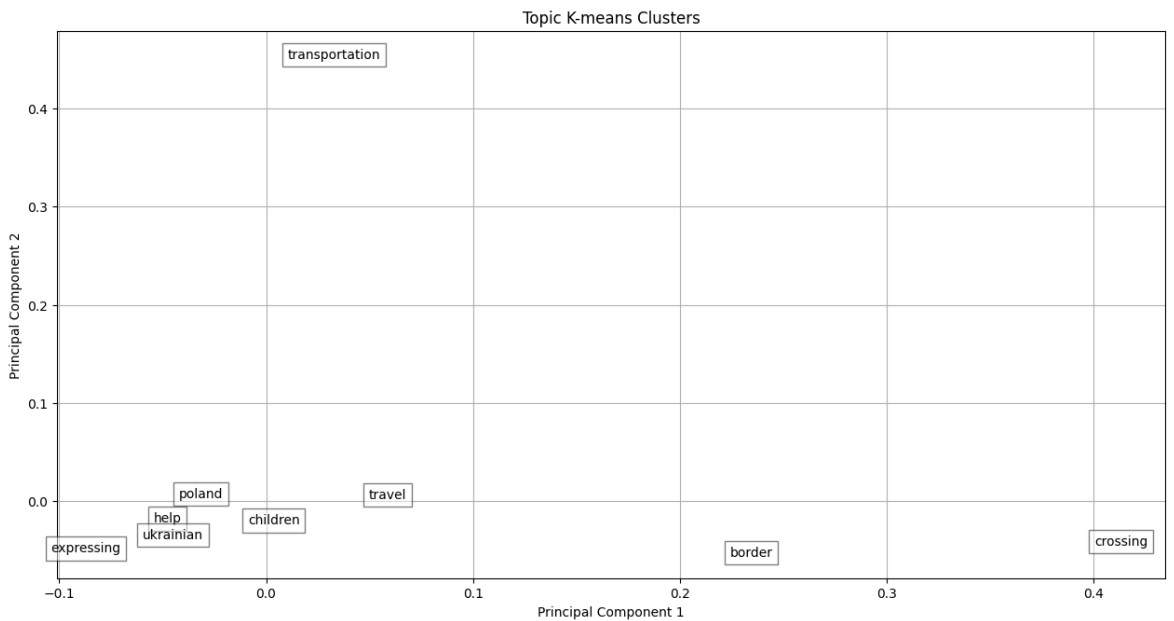


Figure 7. K-means topic clusters

bloodshed! The Russian people are not your enemies, surrender! And you will be treated well: you will receive tasty hot food and warm clothing, and if you are wounded - medical assistance."

We only worked with one message group within one source of social media (Telegram: Help for Ukrainians in Poland). There could be extensive follow-up research on combining multiple streams of social media to develop a more comprehensive picture of situation awareness related to displaced people. Additionally, Telegram has an age bias, most users are younger, more tech-savvy people. Thus, results from this analysis would likely exclude the needs and situations of many older individuals or individuals without smart phones.

We also conducted the topic and thematic analysis based on a translation to English as we acknowledge our limitation and lack of knowledge of Ukrainian and Russian. Deeper incorporation of specialists with expertise in Ukrainian and Russian languages could enable a similar analysis to be done in the source languages and compare those results with those done via English translations. Related to all these points, it is important to acknowledge that ChatGPT itself has gender, racial, and cultural bias derived from its own training data that could impact analysis results (Ray, 2023).

The current word cloud and K-means results presented here are not inherently useful for humanitarian practitioners. However, future research in automated sorting and filtering of topics, such as messages related to health care and medical needs, would allow for more detailed analysis of specific needs and potential barriers to services. Identifying methods for filtering out "non-relevant" text would also enhance the utility of the word cloud or word frequency analysis. Identifying repeated phrases (i.e. "...schedule a medical appointment..." or "...long waiting times...") could also assist in identifying more specific key themes. Thematic analysis using Latent Dirichlet Allocation (LDA) could offer an improved approach over K-means to support this research direction (Westrupp et al., 2022).

Areas of visualization follow-up research would include identification and geocoding locations within the messages to generate geospatial visualizations to understand more specifically about the location of refugees and their movements over space and time, and issues they are facing (Hübl, Cvetojevic, Hochmair, & Paulus, 2017). For example, placing geocoded messages within the context of a map like shown in Figure 1 where official government data sources could be combined with social media data sources to generate a more comprehensive overall picture of the situation of refugees in Poland (Tomaszewski, 2021). There is also extensive visualization research that could be conducted on hierarchical visualization that could be done to allow topic and thematic clusters to be investigated in terms of specific messages within those clusters for more insight and follow-up on specific messages (Elmqvist & Fekete, 2009).

Finally, the work presented here was a pilot study to evaluate the technical aspects of an LLM within the context of a humanitarian crisis situation. Metrics like Kappa and Corpus BLUE scores are easy to understand metrics that can be used to communicate with non-technical humanitarian professionals as an initial step in building trust and acceptance of LLM in the context their ability to be effective at language translations within the context of an international crisis. The ISCRAM community has long made calls for more integration between the academic research and practitioner groups within the overall community. Approaches like one presented here can help begin to cross this divide.

Longer term, it will be important that the approaches like those presented here are evaluated by working humanitarian crisis professionals engaged in situation awareness activities related to forced displacement. For example, analyzing the accuracy of LLM analysis as compared to existing ground-truthed data sources reflecting the needs and situations of Ukrainian refugees in Poland such as the Multi-Sectoral Needs Assessment (MSNA) report which is based on extensive survey data collection with Ukrainian refugee populations (United Nations High Commissioner for Refugees (UNHCR), 2024).

CONCLUSION

In this work, we evaluated the utility of a large language model (LLM) to develop situation awareness related to the forced displacement of Ukrainian refugees into Poland. Specifically, utilizing text messages derived from the Help for Ukrainians in Poland (UKR: Помощь украинцам в Польше) Telegram message group, we translated messages in a variety of languages including Russian, Ukrainian, and Polish into English using ChatGPT. The contents of the messages were then analyzed with ChatGPT to identify topics and themes within those messages. The text of topics and themes generated by ChatGPT were then analyzed using K-means clustering to identify the most important topics and themes from within the data that was collected. The language identification and translation capabilities of the LLM were evaluated by two human evaluators and measured using Kappa and Corpus BLUE scorers with the LLM performing effectively (Kappa: 0.86, Corpus BLUE: 0.46).

It can therefore be concluded that within the scope and acknowledged limitations of the work presented here that the ability of LLM using carefully developed language prompts for large data volume analysis with no need for

manual human analysis shows promise for humanitarian analytics focused on rapidly identifying potential key trends, needs, and locations of displaced people.

As a final takeaway from the work presented in this paper, we asked ChatGPT itself the following for future guidance in this line of work via this prompt:

What would you say to a human academic and practitioner audience interested in information systems for crisis response and management in terms of your ability to help with understanding the situation awareness of Ukrainian refugees who have been forcibly displaced into Poland?

And received back the following reply:

As an AI language model, I can provide valuable assistance in understanding the situation awareness of Ukrainian refugees who have been forcibly displaced into Poland within the context of information systems for crisis response and management.

Here's how I can help:

Data Analysis and Interpretation: I can process and analyze large volumes of data from various sources such as news articles, social media, official reports, and humanitarian organizations' databases to extract insights about the situation of Ukrainian refugees in Poland. This includes understanding their demographics, needs, concerns, and challenges.

Trend Analysis: By analyzing patterns and trends in the data, I can help identify emerging issues, hotspots of activity, shifts in refugee movement, and changes in the humanitarian situation over time. This information is crucial for decision-making and resource allocation by crisis response teams and policymakers.

Risk Assessment: Utilizing natural language processing (NLP) techniques, I can assess the sentiment and tone of refugee communications to identify areas of distress, potential conflicts, or security risks. This aids in proactive risk management and early intervention to prevent escalation of crises.

Resource Allocation Optimization: Through optimization algorithms and predictive modeling, I can assist in determining the most efficient allocation of resources such as food, shelter, medical supplies, and transportation based on real-time needs assessments and demand forecasts.

Communication and Coordination Support: I can facilitate communication and coordination among various stakeholders involved in the crisis response, including government agencies, non-governmental organizations (NGOs), aid workers, volunteers, and affected communities. This ensures timely exchange of critical information and enhances collaboration efforts.

Decision Support Systems: By providing synthesized reports, dashboards, and visualizations, I can support decision-makers in gaining a comprehensive overview of the situation, evaluating the effectiveness of interventions, and making informed decisions to improve crisis response strategies.

Overall, my capabilities extend to assisting academic researchers and practitioners in gaining deeper insights into the situation awareness of Ukrainian refugees in Poland, enabling them to develop more effective information systems and strategies for crisis response and management.

Thus, it is notable that ChatGPT in its current form does have an awareness of how it can support humanitarian operations and it will be interesting to see how these functionalities can be further utilized in the future.

However, LLM are not a panacea. Caution should be considered in that social media and LLM contain bias, falsehoods rumors and other potential pitfalls that can interfere with true understanding of a humanitarian situation (Acerbi & Stubbersfield, 2023; Gallegos et al., 2023). Further evaluation of LLM, social media and combination of official government data sources can show a pathway forward for better understanding of what people who are forcibly displaced are experiencing and how they can be supported.

ACKNOWLEDGMENTS

This research was supported by a grant from the Polish-U.S. Fulbright Commission.

REFERENCES

- Acerbi, A., & Stubbersfield, J. M. (2023). Large language models show human-like content biases in transmission chain experiments. *Proceedings of the National Academy of Sciences*, 120(44), e2313790120. doi:10.1073/pnas.2313790120
- BBC News. (2022). Ukraine conflict: Simple visual guide to the Russian invasion. Retrieved from <https://www.bbc.com/news/world-europe-60506298>
- Crew, R., Ruetzel, S., Barton, J., Bymes, T., Wegener, E., Rivera, V., . . . Mushagalusa, J. (2023). How are humanitarians using AI tools like Chat GPT? Retrieved from <https://www.calpnetwork.org/blog/how-are-humanitarians-using-ai-tools-like-chat-gpt>
- Elmqvist, N., & Fekete, J.-D. (2009). Hierarchical aggregation for information visualization: Overview, techniques, and design guidelines. *IEEE transactions on visualization and computer graphics*, 16(3), 439-454. doi:10.1109/TVCG.2009.84
- Endsley, M. R. (1995). Toward a theory of situation awareness in dynamic systems. *Human Factors*, 37(1), 32-64. doi:10.1518/001872095779049543
- Federal Emergency Management Agency (FEMA). (2017). Harvey Rumor Control. Retrieved from <https://www.fema.gov/blog/harvey-rumor-control>
- Gallegos, I. O., Rossi, R. A., Barrow, J., Tanjim, M. M., Kim, S., Demoncourt, F., . . . Ahmed, N. K. (2023). Bias and fairness in large language models: A survey. *arXiv preprint arXiv:2309.00770*.
- Google Cloud. (n.d.). Understanding the BLEU Score. Retrieved from <https://cloud.google.com/translate/automl/docs/evaluate#bleu>
- gov.pl. (n.d.). Obtain a PESEL Number – a service provided to foreigners. Retrieved from <https://www.gov.pl/web/gov/uzyskaj-numer-pesel-usluga-dla-cudzoziemcow-en>
- Hübl, F., Cvetojevic, S., Hochmair, H., & Paulus, G. (2017). Analyzing refugee migration patterns using geo-tagged tweets. *ISPRS International Journal of Geo-Information*, 6(10), 302. doi:10.3390/ijgi6100302
- Hughes, A. L., & Palen, L. (2009). Twitter adoption and use in mass convergence and emergency events. *International journal of emergency management*, 6(3-4), 248-260. doi:10.1504/IJEM.2009.031564
- International Organization for Migration. (2023a). Poland — FGD Report on Employment, Accommodation, and Future Plans (April 2023). Retrieved from <https://dtm.iom.int/reports/poland-fgd-report-employment-accommodation-and-future-plans-april-2023>
- International Organization for Migration. (2023b). Poland | Displacement Tracking Matrix. Retrieved from <https://dtm.iom.int/poland>
- Kumbhar, R., Mhamane, S., Patil, H., Patil, S., & Kale, S. (2020). *Text document clustering using k-means algorithm with dimension reduction techniques*. Paper presented at the 2020 5th International Conference on Communication and Electronics Systems (ICCES).
- OpenAPI. (2024). OpenAI API. Retrieved from <https://openai.com/blog/openai-api>
- Papineni, K., Roukos, S., Ward, T., & Zhu, W.-J. (2002). *Bleu: a method for automatic evaluation of machine translation*. Paper presented at the Proceedings of the 40th annual meeting of the Association for Computational Linguistics.
- Power, R., Robinson, B., & Cameron, M. (2022). *Insights from a Decade of Twitter Monitoring for Emergency Management*. Paper presented at the Proceedings of the ISCRAM Asia Pacific Conference.
- Ray, P. P. (2023). ChatGPT: A comprehensive review on background, applications, key challenges, bias, ethics, limitations and future scope. *Internet of Things and Cyber-Physical Systems*. doi:10.1016/j.iotcps.2023.04.003
- Reidy, E. (2022). Is Poland's smooth reception of Ukrainian refugees heading for trouble? Retrieved from <https://www.thenewhumanitarian.org/news-feature/2022/08/04/Poland-Ukraine-refugee-concern-grows>
- Roane, A. (2023). Four ways ChatGPT could help level the humanitarian playing field. Retrieved from <https://www.thenewhumanitarian.org/opinion/2023/03/20/ways-chatgpt-could-help-humanitarian-field>
- Rocca, R., Tamagnone, N., Fekih, S., Contla, X., & Rekabsaz, N. (2023). Natural language processing for humanitarian action: Opportunities, challenges, and the path toward humanitarian NLP. *Frontiers in big Data*, 6, 1082787. doi:10.3389/fdata.2023.1082787
- Tanner, L. (2023). *Learning From Innovations Using Artificial Intelligence*. Retrieved from <https://humanitariangrandchallenge.org/research-learnings-reports/>
- Telegram. (2024). Помощь украинцам в Польше (Help for Ukrainians in Poland). Retrieved from t.me/refugeesinPoland
- Tomaszewski, B. (2021). *Geographic Information Systems for Disaster Management (Second Edition)*. Oxfordshire and New York: Routledge.

- UNHCR Regional Bureau for Europe Data Identity Management and Analysis Unit (DIMA). (2022). *Lives on Hold: Profiles and Intentions of Refugees from Ukraine*. Retrieved from
- United Nations High Commissioner for Refugees (UNHCR). (2022a). Poland welcomes more than two million refugees from Ukraine. Retrieved from <https://www.unhcr.org/news/press/2022/3/6234811a4/poland-welcomes-million-refugees-ukraine.html>
- United Nations High Commissioner for Refugees (UNHCR). (2022b). Situation Ukraine Refugee Situation. Retrieved from <https://data.unhcr.org/en/situations/ukraine>
- United Nations High Commissioner for Refugees (UNHCR). (2024). *Multi-Sector Needs Assessment - Protection Challenges for Refugees from Ukraine in Poland (MSNA 2023)*. Retrieved from <https://data.unhcr.org/en/documents/details/106915>
- Viera, A. J., & Garrett, J. M. (2005). Understanding interobserver agreement: the kappa statistic. *Fam Med*, 37(5), 360-363.
- Westrupp, E. M., Greenwood, C. J., Fuller-Tyszkiewicz, M., Berkowitz, T. S., Hagg, L., & Youssef, G. (2022). Text mining of Reddit posts: Using latent Dirichlet allocation to identify common parenting issues. *PLoS one*, 17(2), e0262529. doi:10.1371/journal.pone.0262529
- Yerramreddy, D. R., Marasani, J., Gowtham, P. S. V., & Abhishek, S. (2023). *An Empirical Analysis of Topic Categorization Using PaLM, GPT and BERT Models*. Paper presented at the 2023 Innovations in Power and Advanced Computing Technologies (i-PACT).