

Navigating Online Narratives on Israel-Hamas War with LLMs

Clara Maathuis and Iddo Kerkhof

Open University of the Netherlands, Heerlen, The Netherlands

clara.maathuis@ou.nl

email@iddo.eu

Abstract: In recent years, international conflicts have seen a steady rise, accompanied by the increased use of technological advancements to achieve strategic and military objectives. One such conflict, the ongoing hostilities between Israel and Hamas, exemplifies the complex interplay of political, ideological, social, and technological factors shaping contemporary warfare. While extensive analysis is devoted to analysing strategic, historical, and economic dimensions, still limited attention is directed through research efforts towards the vast troves of information available on global social media platforms such as YouTube. These platforms serve as conduits for a wide range of narratives, opinions, and representations related to this war, by providing valuable insights into public discourse and sentiments. Given the capability of automatically and systematically analysing vast amount of content, this research tackles the above stated knowledge gap by developing a novel explorative LLMs (Large Language Models)-based modelling approach to efficiently extract, categorize, and interpret nuanced discussions surrounding the Israel-Palestinian war. It does that following the Data Science methodological approach and focusing on YouTube data aiming at not only contributing to enriching understanding of contemporary conflicts through social media platforms, but also to reflect on the evolving role that state-of-the-art AI technological developments play in shaping the global security landscape.

Keywords: Israel-Hamas war, Israel-Palestine War, Generative AI, Large language models, LLMs, Knowledge graphs

1. Introduction

“Artificial intelligence is a tool that can be used for good or bad. It is not good or bad in itself. It is how we use it.” (Andrew Ng)

The ongoing Israeli-Palestinian conflict, with its deep historical roots, continues to have profound and far-reaching consequences. The multifaceted impacts of this war imply both societal and strategic struggle as well as widespread chaos and hardship (Buheji & Hasan, 2024). At the same time, people deal further with a wide spectrum of psychological and mental health effects including post-traumatic stress disorder (PTSD), anxiety, and depression, directly linked to prolonged exposure to violence, destruction, and loss (Abudayya et al., 2023). In this context, building intervention strategies implies understanding the meaning and implications of such actions as well as how solutions for resilience, conflict resolution, and peace-building mechanisms could be thought and actually put in place. At the same time, diverse demographics shape perspectives across different countries, with age influencing empathy levels and contributing to the growing global empathy for Palestine amidst ongoing political and social challenges (Buheji, Mushimiyimana & Ahmed, 2024). Understanding the nature of this war is crucial for applying the correct legal principles and comprehending the humanitarian impact on the civilian population involved. In this sense, the multifaceted efforts of the global strategic audience to assist through governance, legal, and social instruments, reveal the need for a comprehensive, integrated approach that addresses the diplomatic, legal, humanitarian, and psychological dimensions of this war (Akgül-Açıkmeşe & Özel, 2024; ICC, 2024).

In this context, through social media platforms, the global narratives and public opinion are (re)shaped as these contribute to cultural communication and opinion formation within this complex geopolitical context. These digital spaces have emerged as critical arenas for information dissemination, digital activism, and the moulding of public perception (Anisa, Harjoko & Khadifa (2023). Then a nuanced understanding on the narratives embedded in social media is necessary as it plays a significant role in information propagation, public sentiment formation or crystallization, and mobilization support. This digital dimension introduces an additional layer of complexity to the conflict, where information warfare and narrative control operate concurrently with physical hostilities (Dostri, 2024). In this arena, AI (Artificial Intelligence) demonstrated important success in both building analytical, content generation, and user experience solutions as well as building mechanisms of social media manipulation and efforts of countering it (Fard & Maathuis, 2024), LLMs started to emerge as a powerful force in this domain. Taking into account the fact that LLMs have shown remarkable versatility across various applications, including language translation, text generation, and question answering. In the context of social media, these capabilities translate into enhanced content creation, more sophisticated chatbots, improved sentiment analysis, and more accurate content moderation. The social impact of LLMs in social media is multifaceted, presenting both challenges and opportunities (Ya et al., 2023: Maathuis & Chockalingam, 2024). As noted by Myers (2024), public scepticism surrounding AI implementation

in social media, particularly concerning privacy and data security, necessitates careful management and transparent communication. The potential for LLMs to perpetuate or amplify biases present in their training data is a critical concern, especially in the context of social media where such biases could have far-reaching societal impacts. Additionally, the (substantial) computational resources required for LLM operation raise questions about the environmental sustainability of their widespread adoption in social media platforms. Despite these challenges, the potential benefits of LLMs in enhancing accessibility, improving content quality, and facilitating more meaningful interactions on social media platforms are significant (Raiaan et al., 2023; Hadi, 2023). Given the potential of LLMs in the social media domain as well as the incipient stage of building LLMs-based solutions for understanding conflict narratives in various social media platforms, this research aims to explore their potential application in the context of the ongoing Israel-Palestinian war by focusing on understanding existing war narratives in YouTube. In this sense, this research contributes to a more comprehensive and objective understanding of the conflict, ultimately informing more effective conflict resolution strategies and supporting humanitarian efforts in the region. This is done by within the Design Science Research methodology and considering a Data Science research approach (Peffer et al., 2007) to build a LLMs-based solution. Accordingly, clusters of narratives are captured and knowledge graphs are built in order to reflect on the evolving discourses surrounding the Israel-Hamas conflict in an objective and transparent manner (Maathuis & Chockalingam, 2023).

The outline of this article is structured as follows. Section 2 discusses relevant studies characterizing the research perspectives adopted in this research. Section 3 presents the methodological approach considered together with implementation details. Section 4 reflects on the results obtained. And Section 5 provides concluding remarks and future research perspectives.

2. Related Research

Chen et al., (2024) propose the IsmasRed dataset containing public discourses on Reddit regarding the Israel-Hamas conflict from August to November 2023. Comprising nearly 400,000 conversations and over 8 million comments, the dataset demonstrates highly emotionally charged discussions, reflecting the complex and sensitive nature of the conflict. The research points out the intricate interplay between ideology, sentiment, and community engagement in digital spaces, examining trends in topics, controversy, and emotional and moral language over time. Moreover, the study aims to enrich understanding of online discussions about the conflict, providing valuable insights into how people perceive and engage with this contentious issue on social media platforms. Anisa, Harjoko & Khadifa (2023) analyse the role of the #Free Palestine hashtag in the current Israel-Palestinian conflict based on data collected from 06 to 25 December 2023 from X (former Twitter), Facebook, and Instagram. The study highlights the power of cultural communication through social media platforms in transforming public perceptions, influencing policy decisions, and fostering activist movements on a global scale. This is done through a movement of online activism that was further translated into real-world actions that impacted international narratives surrounding various geopolitical and social dimensions of this war. At the same time, the authors analyse the underlying sentiments of the narrative captured in these social media platforms and find that while X served as a platform for polarized political dialogue, with hashtags like #IsraelVsPalestine fostering intense debates Facebook, in contrast, has become a hub for Palestinian solidarity movements, with #FreePalestine garnering significant support and interactions. Instagram's impact is primarily visual, utilizing emotionally charged images from conflict zones to influence public opinion effectively. Focusing on conventional media, Irfan, Aqeel & Hussain (2024) stress out that biased narratives can significantly influence public perception, shape policy decisions, and impact transnational relations, potentially exacerbating tensions and obstructing peace efforts. The findings emphasize the critical role of the media in shaping global perceptions and policy formulation, highlighting the need for media literacy among viewers and the encouragement of journalistic practices that stimulate well-informed and least-biased discussions on complex geopolitical conflicts.

Knowledge graphs are complex data structures that represent information in a graphical format, consisting of interconnected entities and relationships. At their core, knowledge graphs are defined by a set of entities (E), relations (R), factual triples (T), and a function (F) that represents background knowledge. This structure allows for the organization of information in an intuitive and easily navigable manner, facilitating complex queries and data analysis. The graphical representation of knowledge enables the capture of contextual information, which is often challenging to represent in traditional data management systems, making knowledge graphs particularly valuable for handling real-world knowledge that is situational, layered, and dynamic. The primary components of a knowledge graph include entities, relations, factual triples, and background knowledge. Entities represent the main concepts or objects within the graph, while relations define the connections

between these entities. Factual triples form the basic units of information, consisting of two entities linked by a relation. Background knowledge serves as the foundation for constructing the knowledge graph, providing the rules or schema that guide the formation of factual triples and their relationships. The construction process of a knowledge graph involves three key stages: knowledge acquisition, where data is collected and initial structures are formed; knowledge refinement, which improves accuracy and completeness; and knowledge evolution, ensuring the graph remains up-to-date and relevant over time (Zhong et al., 2023). Moreover, knowledge graphs have diverse applications across multiple domains, significantly enhancing the capabilities of various systems and platforms. In search engines, they improve query understanding and result relevance, while in recommendation systems, they enable personalized suggestions by leveraging entity relationships. Question answering systems benefit from the structured information storage and retrieval capabilities of knowledge graphs, allowing for more precise responses to complex queries. In enterprise settings, knowledge graphs facilitate efficient organization and access to internal data, boosting productivity and decision-making processes. Chen et al., (2023)

Moreover, domain-specific applications in fields such as healthcare, finance, and education utilize knowledge graphs to organize specialized knowledge, enabling professionals to access relevant information swiftly and effectively. These wide-ranging applications demonstrate the versatility and power of knowledge graphs in enhancing information retrieval, analysis, and decision-making across various sectors (Zou, 2020; Abu Salih, 2021; Zhong et al., 2023). In the context of using knowledge graphs for summarization and analysis, the focus can then be done on both the textual content and the structural relationships represented in the knowledge graph, resulting in the selection of sentences that most accurately capture the main ideas of the text. The integration of knowledge graphs in this process significantly enhances the quality of the generated summaries, producing more concise and representative extracts that effectively capture the essence of the source document (Gao, 2023). This could be done in a mechanism such as the one integrating the conversational capabilities of AI models with the dynamic information retrieval strengths of QASs (Question Answering Systems). This integration presents a promising direction for developing more powerful and effective knowledge graphs chatbots, capable of combining natural language interaction with accurate, up-to-date information retrieval (Omar et al., 2023).

LLMs are seen as powerful tools in the detection of disinformation and fake news as they contain extensive world knowledge and strong reasoning abilities to detect and counter false information effectively. Nevertheless, LLMs have a dual nature as they can be both powerful tools against social media manipulation mechanisms such as disinformation and misinformation, but they can they can equally be exploited to generate deceptive content at scale, posing a critical challenge. The historical review of misinformation detection methods and current efforts to utilize LLMs for this purpose provide a comprehensive backdrop, emphasizing the need for robust intervention and attribution strategies. A major challenge identified is the increased difficulty in detecting LLM-generated misinformation, which tends to be more deceptive and harmful than human-written content (Chen & Shu, 2023). On the dual nature of LLMs and their potential fairness risks reflect Castillo-Eslava et al., (2023). Accordingly, the authors conduct a series of case studies comparing ChatGPT's responses with Wikipedia information and United Nations resolutions on controversial territories such as Crimea, the West Bank, and Transnistria. Then the paper illustrates AI's potential to consolidate power and shape public understanding of reality. The findings reveal the necessity for ongoing monitoring and analysis of AI's role in constructing legitimacy and recognizing territorial sovereignty, advocating for a critical approach to emerging technologies.

While LLMs like GPT 3.5 demonstrated the ability to identify fake news and provide multi-perspective rationales, in a study conducted by Hu et al., (2024) they unexpectedly underperform compared to simpler, fine-tuned small language models (SLMs) such as BERT. This discrepancy is attributed to LLMs' challenges in effectively selecting and integrating rationales to form conclusive judgments. However, the authors point out that a promising role for LLMs as advisors to SLMs, leveraging their capacity to generate multi-perspective instructive rationales. This suggests a potential synergy between LLMs and SLMs in enhancing fake news detection, where LLMs contribute depth and breadth of analysis, while SLMs excel in focused, task-specific decision-making. Hence, these results reflect the importance of a balanced approach in leveraging different types of language models for optimal performance in combating misinformation. In contrast, in a different context, Sprenkamp, Jones & Zavolokina (2023) analyse the capabilities of GPT-3 and GPT-4 models to identify propaganda within news articles. Tested against a complex dataset known for subtle propaganda techniques and contextual dependencies, GPT-4 demonstrated performance comparable to state-of-the-art models like RoBERTa. The research employed five variations of GPT-3 and GPT-4, applying different prompt engineering

and fine-tuning strategies, to comprehensively evaluate their propaganda detection abilities. Performance was assessed using metrics such as F1 score, Precision, and Recall, providing a robust evaluation of the models' accuracy and reliability. While the study highlights the potential of LLMs in tackling complex tasks like propaganda detection, it also underscores the challenges that remain, emphasizing the need for further research to enhance their effectiveness in real-world applications. These findings contribute valuable insights to the ongoing development of AI-driven tools for combating misinformation and propaganda in digital media. Particularly in the context of the ongoing war in Ukraine, Pavlyshenko (2023) built a fine-tuned LLaMA 2 model using efficient techniques like PEFT/LoRA to analyze a wide range of text data, including news articles, social media posts, and propaganda narratives. The fine-tuning process enables LLMs to perform complex tasks such as fact-checking, identifying manipulation tactics, and detecting fake news by examining semantic structures, frequency of specific item sets, and associative rules within texts. Additionally, LLMs analysed user behaviour on social media platforms to identify patterns indicative of disinformation campaigns Maathuis & Kerkhof (2023b), such as artificial amplification of messages.

While various studies focus on either various approaches and development of LLMs-based solutions for various analytical purposes and knowledge graphs are increasingly considered for implementation in various scientific and practical applications, based on the comprehensive analysis above conducted on the use of LLMs and knowledge graphs in conflict contexts, and especially focusing on the ongoing conflict in Gaza, it can be seen that a limited number of studies are dedicated to this topic while the complexity of the events and their corresponding narratives and underlying sentiments and emotions continue to evolve. It is then the objective of this research to tackle this knowledge gap.

3. Research Approach and Implementation

This research aims to develop and evaluate an innovative LLMs-based solution for detecting and analysing narratives related to the Israel-Hamas conflict within YouTube videos. The proposed system automates the process of video search and narrative detection, initiating from a single search term and iteratively expanding through grouping similar narratives and generating new search terms. This iterative approach facilitates the compilation of a comprehensive dataset of narratives, enabling a multifaceted analysis of the conflict. The solution extends beyond text-based narrative collection by incorporating a knowledge graph visualization, which elucidates the interconnections between identified narratives. This integrated approach combines advanced NLP (Natural Language Processing) techniques with data visualization to provide a robust tool for understanding the complex narrative landscape surrounding the Israel-Hamas conflict, offering potential applications in conflict analysis, media studies, and political science research. To this end, a Design Science Research methodological approach is considered taking a Data Science stance seeing previous successful applications in war context (Peffer et al., 2007; Maathuis, C., & Kerkhof, 2023) and is depicted in Figure 1. Accordingly, the following research question is formulated: How to detect narratives in YouTube video transcripts on the Israel-Hamas conflict using LLMs?

The data collection starts by establishing the search term which is "Israel Hamas". The term was used to retrieve YouTube videos using the DuckDuckGo Search platform. For each video, transcripts are obtained from YouTube and the videos without a transcript are discarded. Furthermore, the narratives are extracted using GPT-4 and LangChain to facilitate interaction with GPT-4. The prompt objective is defined as follows:

Your objective is to extract and summarize the narratives in this text. Extract at most 10 narratives. If there are more, extract the ones most likely to contain disinformation or hate speech.

As multiple narratives can be extracted from a single video, and the relation between them is stored, so that the reasoning of the LLM can be verified. Then the narratives are clustered with CPT-4, and for each cluster, a unified narrative is created. The prompt objective used in this step is:

Cluster these narratives in similar narratives and create a new narrative description for each cluster in max. 100 words each..

Through a series of experiments and a variation of the number of clusters, the narratives were grouped effectively and the narratives were merged to avoid duplicates in three iterations. Further, triples are extracted from each narrative to create the knowledge graph. The corresponding prompt objective used is depicted below:

Extract triples for this text that can be used in a knowledge graph. For verbs use present tense. Try to prevent complex triple parts consisting of multiple words.'

Standardize these triples by replacing synonyms and alternative phrasings in these triples with a single, consistent text throughout all triples, so that I can create a single consistent knowledge graph with them.

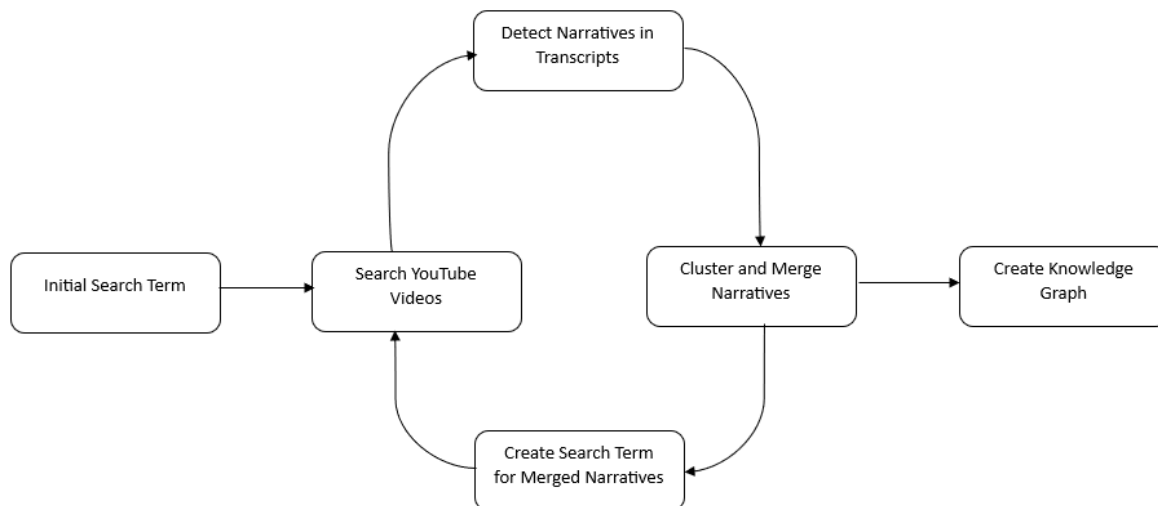


Figure 1: Research approach and implementation

4. Results

A total of 230 videos were found and 61 clusters were initially identified and then merged into 23 narratives. This is captured in Figure 2 and below a selection of three narratives are provided for exemplification and reflection.

Iteration	Narratives	Clusters
1	246	12
2	1077	29
3	882	20
Total	2205	61

Figure 2: Number of narratives and clusters per iteration

Narrative 1: Hamas and Israel pursue ceasefire negotiations, assess potential prisoner swaps, and discuss phased conflict de-escalation. Proposals include hostage releases and call for the destruction of Hamas to end hostilities, with involvement from international and regional negotiators like the US, Qatar, and Egypt. Reports show varied progress and stances on Israeli withdrawal and a halt in operations.

Narrative 6: Gaza suffers a humanitarian crisis with high casualties, infrastructure damage, and shortages of essentials, eliciting varying international responses and increased international aid efforts.

Narrative 21: Globally, security forces stay alert for conflict linked threats, and educational institutions grapple with internal divisions on the Israel-Hamas issue.

From the final set of narratives, a knowledge graph was constructed by extracting and standardizing the triplets. In this way, 164 triplets were extracted. Accordingly, a selection of triplets is further captured and a part of the knowledge graphs is illustrated in Figure 3 and 4:

(Israel, risks, war crimes prosecution),

(Israel, seeks to, dismantle Hamas),

(Israel, takes, security actions),
 (Journalistic disputes, question, authenticity),
 (Likelihood, is of, regional war expansion).

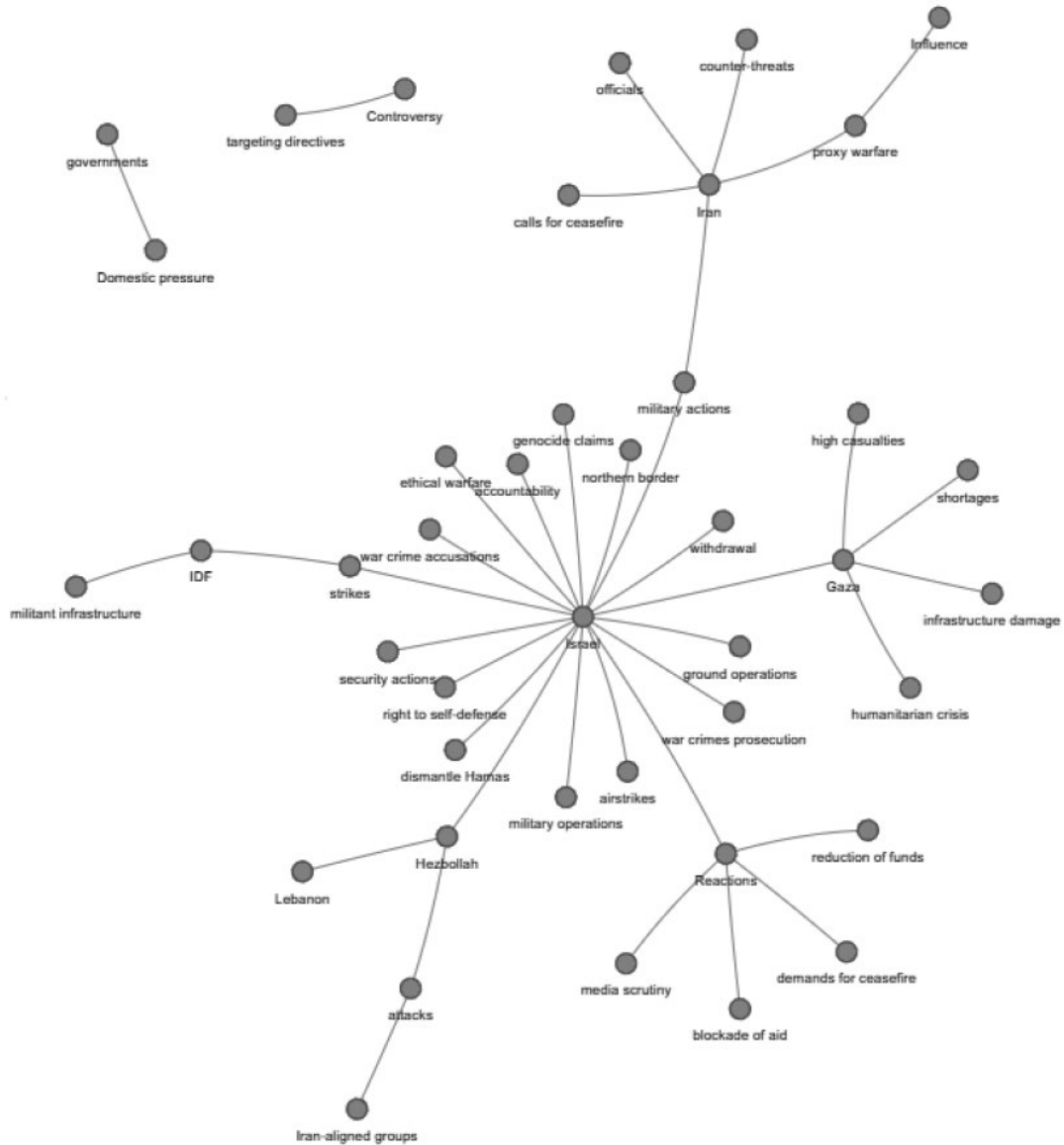


Figure 3: Part of the Knowledge Grap centred around Israel

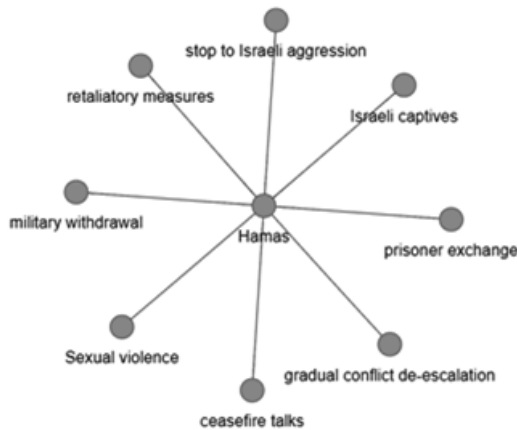


Figure 4: Part of the Knowledge Grap centred around Hamas

As Figure 3 and Figure 4 show, the ongoing narratives in the Israel-Hamas war, as captured through keywords such as the ones depicted, revolve around several critical themes. These include military actions, humanitarian crises, and high casualties, reflecting the intense and often devastating nature of the conflict. Discussions also highlight security actions, strikes, and proxy warfare, indicating the broader geopolitical implications. There are calls for ceasefires and negotiations on captives, military withdrawal, and prisoner exchanges, underscoring the complex diplomatic efforts involved. The narratives are further complicated by social media manipulation mechanisms like disinformation and misinformation since they are shaping public opinion and influencing international perspectives on the conflict.

The results obtained are analysed in a qualitative way given the approach and context considered in this research. They show that the application of LLMs for narrative analysis has potential to capture important underlying dimensions of the narratives and allows their positioning in relation to ongoing media and research trends and discourses for the ongoing Israel-Hamas war. As lessons learned, LLMs offer enhanced explainability, allowing for the tracing of clustered narratives back to their foundational elements excelling in abstractive clustering, grouping texts based on topical relationships rather than mere word similarity, surpassing traditional machine learning techniques. LLMs also demonstrate improved triple extraction and coreference resolution, and generate coherent narratives instead of hard-to-interpret bags of words common in topic modelling. Nevertheless, the subjectivity in assessing the quality and effectiveness of narrative division poses challenges in quantitative evaluation. Inconsistency is evident in the fluctuating number of clusters and narratives identified across multiple executions, highlighting the process's unpredictability. At the same time, while infrequent, formatting issues, could disrupt the parsing process, necessitating repeated attempts. Perhaps most significantly, the use of advanced LLMs like GPT-4 incurs substantial costs and slower processing speeds, limiting their feasibility for large-scale projects without significant investment.

5. Conclusions

This research proposes an innovative modelling approach to analyse the Israel-Hamas war narratives using LLMs on YouTube video transcripts. To this end, an intelligent solution is developed to iteratively expand from a single search term to detect, group, and visualize narratives through knowledge graphs. In this approach, data visualization methods are used in order to provide a method for visualizing and understanding the complex narratives in the landscape of this conflict. To this end, a Data Science approach is taken in a Design Science Research methodological approach and valuable insights are obtained for further conflict analysis, media studies, and political science research. Future research perspectives may include integrating real-time data streams for dynamic narrative tracking, incorporating multi-modal analysis by combining video and audio content with transcript data, and developing cross-platform comparisons to understand narrative propagation across different social media ecosystems. Additionally, investigating the application of this methodology to other complex geopolitical conflicts could yield valuable comparative insights. Conclusively, this research not only offers a novel approach to understanding the narrative landscape of the Israel-Hamas conflict but also paves the way for building advanced AI-based solutions in conflict analysis and media studies.

References

- Abu-Salih, B. (2021). Domain-specific knowledge graphs: A survey. *Journal of Network and Computer Applications*, 185, 103076.
- Abudayya, A., Bruaset, G. T. F., Nyhus, H. B., Aburukba, R., & Tofthagen, R. (2023). Consequences of war-related traumatic stress among Palestinian young people in the Gaza Strip: A scoping review. *Mental Health & Prevention*, 200305.
- Akgül-Açıkmeşe, S., & Özel, S. (2024). EU Policy towards the Israel-Palestine Conflict: The Limitations of Mitigation Strategies. *The International Spectator*, 59(1), 59-78.
- Anisa, N., Harjoko, I., & Khadifa, R. (2023). A Case Study of the Hashtag# Freepalestine: Cultural Communication and Public Opinion Transformation. *Ilomata International Journal of Social Science*, 4(4), 787-803.
- Buheji, M., & Hasan, A. (2024). Beyond Famine and Chaos—Case of Gaza. *International Journal of Management (IJM)*, 15(2), 2024.
- Buheji, M., Mushimiyimana, E., & Ahmed, D. (2024). Empathic engagement with Gaza: dynamics, impact, and prospects. *International Journal of Management (IJM)*, 15(1), 132-156.
- Burghardt, K., Zhang, J., & Lerman, K. (2024). IsamasRed: A Public Dataset Tracking Reddit Discussions on Israel-Hamas Conflict. In *Proceedings of the International AAAI Conference on Web and Social Media* (Vol. 18, pp. 1900-1912).
- Castillo-Eslava, F., Mougan, C., Romero-Reche, A., & Staab, S. (2023). The role of large language models in the recognition of territorial sovereignty: An analysis of the construction of legitimacy. *arXiv preprint arXiv:2304.06030*.
- Chen, C., & Shu, K. (2023). Combating misinformation in the age of llms: Opportunities and challenges. *arXiv preprint arXiv:2311.05656*.

- Chen, J., Geng, Y., Chen, Z., Pan, J. Z., He, Y., Zhang, W., ... & Chen, H. (2023). Zero-shot and few-shot learning with knowledge graphs: A comprehensive survey. *Proceedings of the IEEE*, 111(6), 653-685.
- Chen, K., He, Z., Dostri, O. (2024). Israel's Struggle with the Information Dimension and Influence Operations during the Gaza War. *MILITARY REVIEW*, 1.
- Fard, A. E., & Maathuis, C. (2021). Toward Capturing the Underlying Offensive Mechanisms of Social Manipulation: A Data Model Approach.
- Gao, Y. (2023). *A Deep Learning Approach to Extractive Text Summarization Using Knowledge Graph and Language Model* (Master's thesis, Illinois State University).
- Zhong, L., Wu, J., Li, Q., Peng, H., & Wu, X. A. (2023) comprehensive survey on automatic knowledge graph construction. arXiv 2023. *arXiv preprint arXiv:2302.05019*.
- Hadi, M. U., Qureshi, R., Shah, A., Irfan, M., Zafar, A., Shaikh, M. B., ... & Mirjalili, S. (2023). A survey on large language models: Applications, challenges, limitations, and practical usage. *Authorea Preprints*.
- Hu, B., Sheng, Q., Cao, J., Shi, Y., Li, Y., Wang, D., & Qi, P. (2024) Bad Actor, Good Advisor: Exploring the Role of Large Language Models in Fake News Detection. arXiv 2024. *arXiv preprint arXiv:2309.12247*.
- ICC (2024). PANEL OF EXPERTS IN INTERNATIONAL LAW Convened by the Prosecutor of the International Criminal Court, 20 May 2024.
- Irfan, I., Aqeel, M., & Hussain, M. S. (2024). Critical Discourse Analysis of Islamophobia Reflection in Gaza-Israel Conflict: A Case Study of BBC Reporting. *Journal of Arts and Linguistics Studies*, 2(1), 289-301.
- Maathuis, C., & Kerkhof, I. (2023). The first two months in the war in Ukraine through topic modeling and sentiment analysis. *Regional Science Policy & Practice*, 15(1), 56-75.
- Maathuis, C., & Kerkhof, I. (2023b). Social media manipulation awareness through deep learning based disinformation generation. In *International Conference on Cyber Warfare and Security* (Vol. 18, No. 1, pp. 227-236).
- Maathuis, C., & Chockalingam, S. (2023). Modelling Responsible Digital Security Behaviour for Countering Social Media Manipulation. *Indian Institute for Energy Technology*.
- Maathuis, C., & Chockalingam, S. (2024). Risk Assessment of Large Language Models Beyond Apocalyptic Visions. In *European Conference on Cyber Warfare and Security* (Vol. 23, No. 1, pp. 279-286).
- Myers, D., Mohawesh, R., Chellaboina, V. I., Sathvik, A. L., Venkatesh, P., Ho, Y. H., ... & Jararweh, Y. (2024). Foundation and large language models: fundamentals, challenges, opportunities, and social impacts. *Cluster Computing*, 27(1), 1-26.
- Omar, R., Mangukiya, O., Kalnis, P., & Mansour, E. ChatGPT versus Traditional Question Answering for Knowledge Graphs: Current Status and Future Directions Towards Knowledge Graph Chatbots. (2023) doi: 10.48550. *arXiv preprint ARXIV.2302.06466*.
- Pavlyshenko, B. M. (2023). Analysis of disinformation and fake news detection using fine-tuned large language model. *arXiv preprint arXiv:2309.04704*.
- Peffer, K., Tuunanen, T., Rothenberger, M. A., & Chatterjee, S. (2007). A design science research methodology for information systems research. *Journal of management information systems*, 24(3), 45-77.
- Pycińska, M. (2023). Israeli and Palestinian settler colonialism in new media: the case of roots. *Humanities*, 12(5), 124.
- Raiaan, M. A. K., Mukta, M. S. H., Fatema, K., Fahad, N. M., Sakib, S., Mim, M. M. J., ... & Azam, S. (2024). A review on large Language Models: Architectures, applications, taxonomies, open issues and challenges. *IEEE Access*.
- Sprenkamp, K., Jones, D. G., & Zavolokina, L. (2023). Large language models for propaganda detection. *arXiv preprint arXiv:2310.06422*.
- Ye, W., Ou, M., Li, T., Ma, X., Yanggong, Y., Wu, S., ... & Zhao, J. (2023). Assessing hidden risks of LLMs: an empirical study on robustness, consistency, and credibility. *arXiv preprint arXiv:2305.10235*.
- Zou, X. (2020, March). A survey on application of knowledge graph. In *Journal of Physics: Conference Series* (Vol. 1487, No. 1, p. 012016). IOP Publishing.