EPJ Data Science
a SpringerOpen Journal

RESEARCH   **Open Access**

Analyzing Russia's propaganda tactics on Twitter using mixed methods network analysis and natural language processing: a case study of the 2022 invasion of Ukraine

Iuliia Alieva[1]* , Ian Kloo[1] and Kathleen M. Carley[1]

*Correspondence: ialieva@cmu.edu
[1] Carnegie Mellon University, Pittsburgh, PA, 15213, USA

Abstract

This paper examines Russia's propaganda discourse on Twitter during the 2022 invasion of Ukraine. The study employs network analysis, natural language processing (NLP) techniques, and qualitative analysis to identify key communities and narratives associated with the prevalent and damaging narrative of "fascism/Nazism" in discussions related to the invasion. The paper implements a methodological pipeline to identify the main topics, and influential actors, as well as to examine the most impactful messages in spreading this disinformation narrative. Overall, this research contributes to the understanding of propaganda dissemination on social media platforms and provides insights into the narratives and communities involved in spreading disinformation during the invasion.

Keywords: Social Network Analysis; Natural Language Processing; BERTopic; Disinformation; Russia; Ukraine; Computational Propaganda

1 Introduction

The emergence of computational propaganda has given rise to a new social media phenomenon that utilizes automation and algorithms, facilitating the efficient dissemination and amplification of discourse on social media platforms. This includes the spread of disinformation and state-funded propaganda, enabling ideological control and manipulation [1]. Governments and other entities leverage computational power, Internet resources, and big data to achieve information control and manipulation objectives. The use of social media for spreading disinformation, consolidating power, exerting social control, and promoting agendas has become a recognized strategy for many states globally [2, 3]. Propaganda strategies continuously adapt to technological and media changes [4], emphasizing the need to monitor media discourse, particularly on social media platforms. Recent trends, including the rise of bots, trolls, and other manipulative efforts [5, 6], underscore the importance of identifying and analyzing these activities, as well as the narratives and communities involved in disseminating malicious information.

Springer

The 2022 Russia's invasion of Ukraine underscores the significant role of social media in modern warfare, as both sides utilize online platforms to manipulate geopolitical dynamics and shape public opinion [7]. Russia-affiliated social media accounts propagate narratives aligned with their motives, downplaying support for sanctions against Russia and undermining support for Ukraine. Conversely, Ukrainian side aims to raise global awareness of Russia's war crimes, garner Western support, emphasize their own military endeavors, and challenge prevailing perceptions of the Russian military [7, 8]. While extensive research exists on identifying malicious cyber activities, less attention has been given to investigating narratives and their role in broader conversations, particularly concerning Russia's invasion of Ukraine. This study focuses on identifying primary communities and narratives associated with the prevalent and damaging "fascism/Nazism" narrative in discussions related to Russia's invasion of Ukraine. To achieve this, we employ a mixed-methods pipeline for social media analysis, combining network science approaches, natural language processing, as well as community clustering, and qualitative analysis of the tweets and users. This comprehensive approach allows for the identification of key influencers and communities as well as examination of narratives related to this specific disinformation narrative.

2 Related works

2.1 Unveiling the tactics and impact of Russian propaganda

Russian propaganda tactics have been extensively scrutinized, especially in the context of major global events such as the 2016 US presidential elections and Brexit [2, 9, 10]. The Internet Research Agency (IRA), a Russian state-affiliated troll factory, has garnered attention for its malicious activities aimed at manipulating online opinions through divisive messages [6, 9–11]. With social media platforms and algorithms, the troll factory actively promotes strategic narratives to generate destabilization, polarization, information chaos, and distrust [6, 11]. Key characteristics of IRA trolls encompass deception, the cultivation of political discord and distrust, and the use of online troll accounts to simulate grassroots activities, commonly referred to as astroturfing. It is noteworthy that a range of account types is employed, encompassing automated bots, trolls, and sock puppets controlled by humans but presented as authentic social media users [9].

Contemporary computational propaganda employs traditional propaganda tactics, utilizing symbols, emotions, stereotypes, and existing frames to mold perceptions and manipulate cognition and behavior in pursuit of the propagandist's objectives [12, 13]. In the current media landscape, propaganda techniques have evolved to involve swift distribution across various channels while maintaining a covert presence, giving rise to the phenomenon of computational propaganda. This phenomenon harnesses computational tools like automation and algorithms to propagate and magnify discourses and opinions on social media, serving the objective of ideological control and manipulation [14]. Recent tactics employed to manipulate public opinion involve the convergence of social media platforms, autonomous bots, and big data [15]. These tools use algorithms to precisely and quickly target individuals, providing stakeholders with significant influence without fundamentally altering the nature of propaganda. Propaganda aims to sway and persuade through ideological symbols, seeking specific responses, solidifying identity, and fostering loyalty [15]. It primarily consists of persuasive communication aimed at promoting ideological objectives, shaping public opinion, and institutionalizing the loyalty of targeted groups.

The Russian propaganda apparatus includes both overt and covert participants. Overt actors openly disseminate propaganda, including state-funded media outlets such as RT and Sputnik, as well as official political entities like the Ministry of Defense, the Ministry of Foreign Affairs, and Russia's embassies. In contrast, covert actors operate through less transparent means, including low-credibility news sources referred to as "pink slime media," influencers, automated bot accounts, and deceptive human-operated trolls. These covert actors contribute to the dissemination of propaganda while employing secrecy and deception.

Russian propaganda intentionally lacks consistency and employs a strategy of deliberate confusion. It utilizes multiple explanations to cater to diverse audience preferences without offering clear guidance. The goal is to inundate readers with misleading information, creating a challenge in discerning the truth [16]. Furthermore, Russian propaganda relies on repetition to reinforce its desired narrative and promote familiarity with the message. It specifically targets groups with distinct identities, such as those with anti-West and anti-capitalist beliefs or those who mistrust government and institutions (e.g., conservative, conspiracy, and strongly left- or right-wing groups). By appealing to confirmation bias, Russian propaganda solidifies these groups' existing beliefs. Ultimately, its objective is often to erode trust and undermine the credibility of democratic institutions, sowing chaos and discord [17].

2.2 Russian propaganda in the context of war in Ukraine

Russia's invasion of Ukraine has prompted academic scrutiny of disinformation operations [7, 18–22]. Termed a "hybrid war," the invasion combines conventional warfare with unconventional disinformation tactics [7]. Tolz and Hutchings [22] describe narratives disseminated by Russian state propaganda during the invasion. Exploiting distorted historical and cultural discourses, Russian state propaganda denies ethnic diversity and portrays Ukraine as an integral part of Russia, using national imperialistic identity narratives. State-affiliated actors and opposition groups employ themes of colonization and fascism/Nazism, albeit with contrasting meanings. Russian state propaganda accuses the "collective West" of colonizing Ukraine, while opposition groups argue that Russia itself is a colonizing empire. Claims of a Nazi regime and genocide of the Russian population in Ukraine by Russian state propaganda are countered by the opposition, asserting that Russia is the perpetrator of the Ukrainian genocide [22]. While Russian state propaganda spreads distorted information to manipulate audience attitudes, it proves especially persuasive among individuals with pre-existing pro-Russian and/or anti-West sentiments, as well as conservative and alt-right groups worldwide. An example of widespread disinformation is the claim that the US has constructed military biolabs in Ukraine, allegedly developing bioweapons aimed at Russia. This disinformation campaign gained significant traction on Twitter, with dissemination across conservative, alt-right, and anti-vax communities [23].

The predominant discourse surrounding Russia's invasion of Ukraine accuses the United States of imperialism, portraying Ukraine as a victim of American aggression. This narrative, along with the existing narrative of NATO expansion, depicts Western influence as a threat and provides justification for the ongoing war. It resonates with far-left, alt-right, and conservative groups by referencing NATO expansion, US imperialism, and traditional values. Ukrainian aspirations for cultural and national sovereignty, as well as closer ties

with the West, are depicted as 'fascism/Nazism,' framing the invasion's main goal as 'denazification' and 'de-Westernization' [22, p. 13].

While the terms such as disinformation and propaganda have distinct meanings, they have their shared use of false or distorted information to manipulate audience attitudes. Propaganda, however, carries a stronger political context and relies on emotional reactions through falsification. Propaganda actors leverage cultural and historical associations to create persuasive narratives. In the context of the Russia's invasion of Ukraine, the utilization of the term "Nazism/fascism" holds significant cultural and historical connotations, justifying the invasion and reinforcing biases and anti-Ukraine narratives promoted by Russian state media since the annexation of Crimea in 2014.

Russia's state propaganda narratives display specific characteristics, incorporating the adoption of identity-related discourses, cultural and historical narratives, elicitation of emotional reactions, falsification of facts, and extensive repetition across various media channels such as social media, TV, press, and messaging apps. Twitter is viewed as a platform where Russian opposition, individuals with liberal attitudes, and foreign audiences are targeted with disinformation tropes. Analyzing these prevalent narratives, influential actors, communities, and propaganda strategies is crucial for understanding the mechanisms of state propaganda, discourse dynamics, and the consumption of disinformation and counter-disinformation efforts. This study focuses on examining the propagation and discussion of the 'fascism/Nazism' narrative specifically on Twitter, encompassing both the English and Russian segments of the platform. The following questions are presented for analysis:

RQ1: Who are the most influential and prominent actors and communities involved in the Twitter discourse about 'Nazism/fascism' in Ukraine in Russian and English Twitter discourses?

RQ2: What narratives and topics are identified in the discourse for each language?

3 Method

Python package *twarc* was used to collect tweets via an archive search with updated Twitter academic API. Twarc facilitates the retrieval of tweets using the Twitter API and simplifies the process of searching, filtering, and collecting tweets. It allows researchers and developers to efficiently gather large datasets from Twitter, which is particularly useful for conducting analyses and studies involving social media content, such as the examination of propaganda, disinformation, or other trends on the platform. Two datasets in English and Russian were collected using keywords such as 'nazi,' 'denazification,' 'Ukraine,' and others in both languages (see Python queries in Table 1).

As a result, two datasets were compiled with English and Russian tweets. The data was collected for a period of one year, starting on December 24, 2021, and ending on January 24, 2023 (see Table 2 for more information about the datasets).

For data analysis, a mixed-method pipeline was employed, which included network analysis of Twitter data to detect key actors and influencers, Leiden clustering to identify communities within the network, as well as natural language processing to understand the topics and narratives (see Fig. 1 for the data analysis process).

This methodology has already been partially implemented in previous research (e.g., [23, 24]), including the identification of influencers and qualitative network analysis. However, this study enhances it by incorporating a modified BERTopic modeling methodology

Table 1 Queries with keywords for data collection

English dataset	Russian dataset
twarc2 search '((neo-nazi OR nazi OR denazifying OR denazification OR denazify OR nazism OR fascism OR fashism OR fashist OR fascist) (ukraine OR kyiv OR kiev OR zelenskiy OR zelensky OR zelenskyy OR azov OR donbass OR donbas) OR ukronazi OR ukrofashist OR ukrofashism OR ukrofascism OR ukrofascist OR ukronazist OR ukronazism)(lang:en)' –archive –start-time 2021-12-24 –end-time 2023-01-24	twarc2 search '(неонацизм OR неонацист OR нацики OR нациками OR нацикам OR нациков OR денацификация OR денацификации OR денацификацию OR денацификацией OR денацифицировать OR денацифицировали OR денацифицируем OR нацизм OR нацизма OR нацизму OR нацизмом OR нацизме OR неонацисты OR неонацистов OR неонацистам OR неонацистами OR неонацистах OR укрофашист OR фашист OR фашисты OR фашистов OR фашистам OR фашистами OR фашистах OR фашизм OR фашизма OR фашизму OR фашизмом OR фашизме) (украина OR украины OR украине OR украину OR украиной OR киев OR зеленский OR зеленского OR зеленскому OR зеленским OR зеленском OR днр OR лнр OR донбас OR донбасс OR азов) OR (укрофашист OR укрофашиста OR укрофашисту OR укрофашистом OR укрофашисте OR укрофашисты OR укрофашизм OR укрофашистов OR укрофашистами OR укрофашистах OR укронацист OR укронациста OR укронацисту OR укронацистом OR укронацисте OR укронацисты OR укронацистов OR укронацистами OR укронацистах OR укронацистам OR укрофашизм OR укронацизм)(lang:ru)' –archive –start-time 2021-12-24 –end-time 2023-01-24

Table 2 Summary of the data used in this study

Language	English	Russian
Total Number of tweets	5,242,019	883,225
Number of retweets	3,039,867	577,635
Number of users	1,016,126	126,050
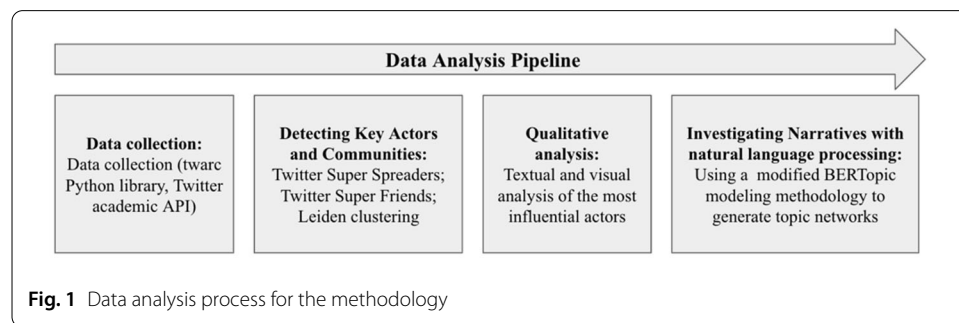
Fig. 1 Data analysis process for the methodology

to generate topic networks, aiming to provide a better understanding of the narratives and their development. The ORA software tool for network analytics [25] was used to analyze the data. ORA provides various features for Twitter data, such as identifying super spreaders (users who frequently generate and effectively spread shared content) and super friends (users who engage in frequent two-way communication, facilitating large or strong communication networks). ORA helps to identify key actors and communities for further qualitative analysis [26, 27].

Influencers are users whose tweets have a significant impact on the social network due to their follower count and network position. The narratives they disseminate can influence the opinions of other users within the network. Identifying key influencers is cru-

cial for understanding the potential harm of information operations. By conducting Twitter network analysis in ORA, it is possible to detect super spreaders, super friends, and other influential users [28]. To identify network communities participating in conversations on Twitter, we used the Leiden clustering method [23]. The Leiden clustering algorithm involves network partitioning and node movement, ensuring the formation of well-connected communities. The Leiden algorithm has been proven to be more efficient than others, such as Louvain, as it is faster and provides better partitions [29]. After identifying the communities, qualitative methods were employed to compare the content and user characteristics between the groups. We implement qualitative textual and visual analysis of the most influential agents in the network, their corresponding tweets and narratives surrounding communities of agents.

In this study, we concentrate on two distinct categories: super spreaders, identified through a combination of network analysis metrics like out-degree centrality, page rank centrality, and k-core; and super friends, determined through a combination of total degree centrality and k-core. Super spreaders comprise users who regularly generate content that is widely and effectively disseminated across the network. The super friends list comprises users engaging in frequent two-way communication, thereby contributing to the formation of extensive and robust communication networks. We run influencer analysis for the largest Leiden communities and identify the main attitudes presented among influencers in each group. Since the number of communities and actors is significantly high we choose to focus on the most influential users since their content is widely disseminated in each community. We implement manual qualitative analysis when we investigate the content of tweets and users through qualitative textual and visual analysis. We check the first hundred most influential users in each community for the five largest communities.

A modified BERTopic modeling methodology is used to generate topic networks based on the text content of tweets that show the flow of Twitter conversations between topics. BERTopic has been shown to be superior to other topic modeling strategies such as LDA for short-text documents [30]. The typical BERTopic pipeline includes using a (usually BERT-based) embedding model to create vector representations of text, a dimensionality reduction step, and a clustering step to group like-documents [31]. We used a multilingual embedding model, which allowed us to process English and Russian tweets in the same pipeline, translating the final topic labels to English for readability.

Our method enhances the typical BERTopic pipeline by using OpenAI's GPT-4 to generate human-readable topic labels instead of using common words within each topic cluster [32]. Specifically, we passed the 10 most representative documents to GPT-4 and requested a short (five-word) summary. We validated the GPT-4 results by reading documents within the topic cluster, and we found the GPT-derived labels to be accurate and more human readable when compared to the common word descriptions.

The output of our BERTopic pipeline is a topic associated with each tweet. These topics become the nodes in our graph representation. We then assign a directed edge between topics each time they are present in a reply chain. For example, a directed edge from the topic "Ukraine Conflict" to "Ukraine Nazi Soldiers" means that a user replied to a tweet about the conflict in Ukraine with a comment about Nazi soldiers. Edge weights correspond to the number of times a topic was used in reply to another topic. The result is a network showing the typical flows between topics in Twitter reply chains.

Fig. 2 A topic network for the English dataset
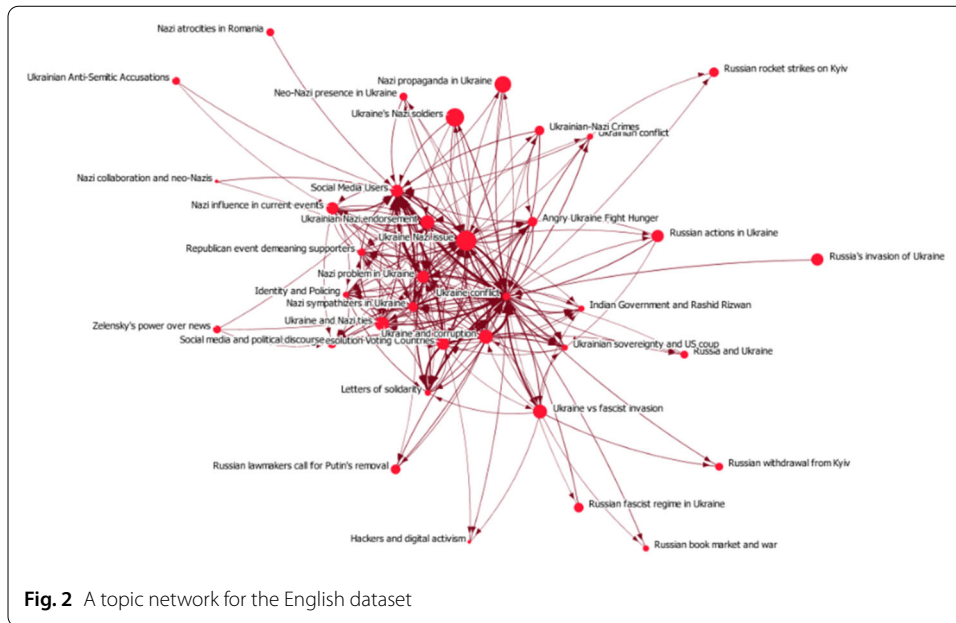
4 Results

4.1 BERTopic analysis

To identify general topics in the English and the Russian conversations, we used natural language processing approach with BERTopic modeling. For each language dataset such as English (Fig. 2) and Russian (Fig. 3), a topic network was generated with an edge meaning that a topic (target) was brought up in response to another topic (source). Node size is proportional to the number of times a topic appeared in the data and edge size is proportional to the number of edges that exist between topics. We can see that generated topics help us to understand general topics in the discussion.

These networks show similar discussions in both English and Russian tweets that include many of the expected propaganda narratives. A notable difference is the English topics tend to focus on the word "Nazi" while the Russian topics use a variation of "fascist" in narratives that seek to justify the invasion for similar reasons.

4.2 Twitter influencers and Leiden communities in the English dataset

To gain a deeper understanding of the narratives, we utilized network science analysis and conducted textual qualitative analysis of the topics. Our analysis involved identifying the main influencers in the overall conversation within each language and the largest Leiden groups. The list of super spreaders in English dataset encompasses accounts belonging to influential figures such as POTUS, Ukraine president Zelensky, Elon Musk, as well as various influencers and newsrooms reporting on the invasion. Additionally, the list includes users who propagate anti-West and anti-Ukraine narratives associated with Russia's invasion. We also observed the presence of low-credibility news accounts disseminating narratives that are anti-Ukraine, antisemitic, anti-West, anti-NATO, and promoting pro-Russia and pro-China propaganda narratives in English language (for instance, the Grayzone News and its bloggers). It is worth noting that the account of Russia's ministry of foreign affairs holds a prominent position in terms of out-degree centrality, indicating a substantial number of out-links to other users.

Fig. 3 A topic network for the Russian dataset
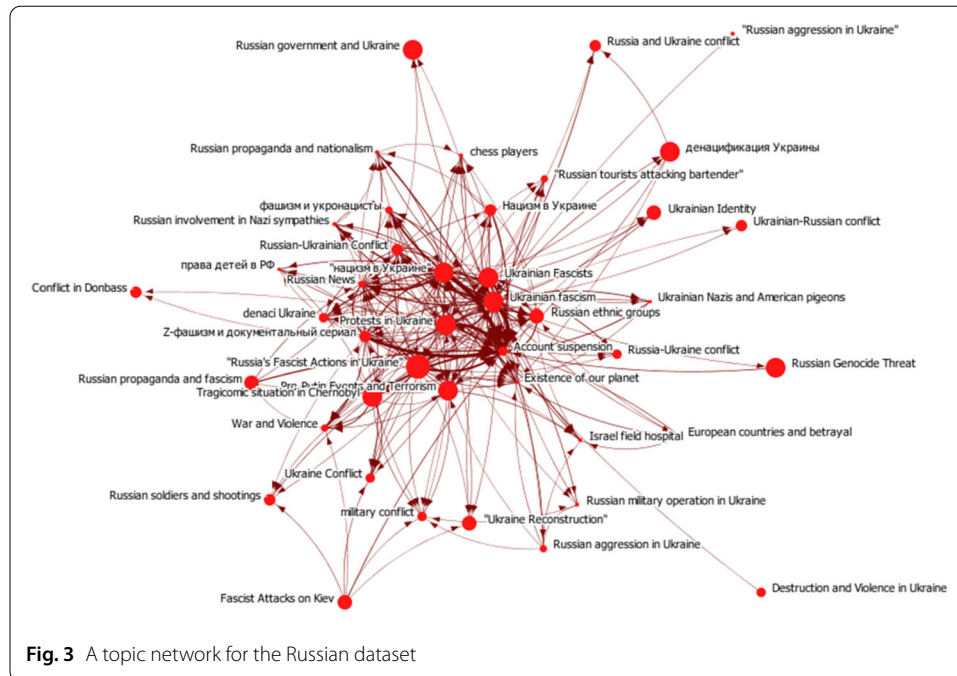
Among the super friends, certain users exploit narratives surrounding the war in Ukraine and Western support to undermine the West and exploit political polarization in the US. These users interact with others who aim to promote Russian propaganda narratives concerning neo-Nazis in Ukraine and various conspiracy theories. Generally, anti-Ukraine and pro-Russia users attempt to justify Russia's invasion of Ukraine by presenting arguments and propaganda narratives, including defending the Russian population in Ukraine, alleging discrimination, and characterizing Ukraine's government as Nazi. Identified propaganda narratives also encompass portraying Ukraine, its government, and its allies as weak, framing the Western countries as adevrsaries, referencing NATO expansion, highlighting corruption in Ukraine, asserting Western domination and hegemony, and emphasizing the perceived inability of the West to unify. Additionally, these narratives label Ukraine as a Nazi or totalitarian state, thereby undermining the Western countries that support it. Other propaganda themes involve accusing Ukraine of war provocations, criticizing the United States and other partners for overlooking internal problems while providing financial assistance to Ukraine, asserting that Crimea and other occupied territories are historically Russian, and accusations against the Ukrainian government for 'ethnic cleansing' in Eastern regions of the country with a predominantly Russian-speaking population. Many propaganda narratives are disseminated through replies in two-way communication, potentially to mimic real-life conversations and circumvent Twitter suspensions.

In the top super friends' list, there are also users who actively debunk the Nazi narratives, engaging with disinformation and presenting counterarguments. Overall, among the counter narratives, users highlight that many countries, including Russia, have a neo-Nazi problem. They also point out that Ukraine has a Jewish president, using it as an argument against the likelihood of Ukrainian government being Nazi. Additionally, users mention that Russia instigated the conflict in Eastern Ukraine with pro-Russia separatists and that there is no systemic discrimination against the Russian population in Ukraine. The main

Table 3 Statistics for Leiden clustering in English dataset (with Newman modularity 0.542)

# groups	31,900
# isolate groups	24,662
# dyad groups	4,879
# triad groups	1,420
# users in largest group	309,180

counter narrative emphasizes that Putin's regime itself behaves like Nazis and resembles Nazi Germany.

Through the analysis of influencers in Leiden groups, we identified five largest groups and investigated the top hundred influencers in each group (see more information about Leiden groups in Table 3). The largest group demonstrates pro-Ukraine attitudes and includes Ukrainian media, Ukrainian and US politicians, and other pro-Ukraine users. The second and third groups among the top influencers consist of alt-right political activists, bloggers, low-credibility websites, conspiracy theorists, and trolls. The fourth group encompasses accounts of Western politicians and accounts demonstrating support for Ukraine. Group 5 comprises accounts of Russian officials, including the Ministry of Foreign Affairs, Russia's embassies in various countries, Russian state-affiliated media (RT), and other pro-Russia accounts.

The first and fourth largest groups propagate narratives that express solidarity with Ukraine, promote support for Ukraine, and advocate for sanctions against Russia. These narratives also accuse Putin and Wagner (a Russian military group) of being Nazis, highlight Russia's war crimes, civilian casualties, and provocative actions. In contrast, the narratives spread by alt-right activists and conspiracy theorists in Groups 2 and 3 revolve around Hunter Biden's emails, alleging Biden's corrupted interests in Ukraine. They also complain about and undermine financial support to Ukraine, mention corruption in Ukraine, discuss disinformation narratives and conspiracy theories about global elites and world order, military biolabs in Ukraine, and others. Group 5 propaganda accounts primarily disseminate Russian propaganda narratives. These narratives blame the US for its participation in previous conflicts like Iraq and Syria, mention NATO bombings of Yugoslavia, highlight conspiracies about neo-Nazis and military biolabs in Ukraine, criticize NATO expansion, undermine Ukraine and its partners, accuse Ukraine of attacks on its own civilians, complain about Russophobia, and promote anti-West sentiments.

4.3 Twitter influencers and Leiden communities in the Russian dataset

In the Russian datasets, the list of top superspreaders includes satire accounts and influencers promoting pro-Ukraine content in both Russian and Ukrainian languages. The top influencer list among super friends mostly comprises pro-Ukraine accounts and Russian propaganda bloggers spreading Nazi disinformation discourse and hate speech. The posting of disinformation narratives about Nazis in Ukraine began two months before the war, based on the start date of our data collection. Additionally, many pro-Russia accounts share Telegram links to promote their Telegram channels, redirecting their audience from Twitter to alternative platforms.

Among the main pro-Russia narratives, there is an undermining of Ukraine, its politicians, and its supporters, with claims of a corrupted government and allegations that Ukraine is governed by Nazis. More typical state propaganda narratives include referring to Russia's invasion of Ukraine as a "special military operation" necessary to prevent an

Table 4 Statistics for Leiden clustering in Russian dataset (with Newman modularity 0.515)

# groups	6,457
# isolate groups	4,847
# dyad groups	1,112
# triad groups	269
# users in largest group	42,803

attack from Ukraine or to stop war and discrimination against the Russian population, aligning with Russia's official government stance. Pro-Russia state narratives also present arguments highlighting the alienation of Ukraine, suggesting that Americans and citizens of Western countries do not support military aid for Ukraine. Pro-Russia accounts also accuse Ukrainian military forces of war crimes and killing their own citizens. They depict the invasion as a "liberation" of Ukraine from "Ukrofascists" and "Ukronazis," terms commonly used by Russian state propaganda and military accounts. These narratives often assert that Russia only targets military objectives and deny responsibility for civilian casualties. Furthermore, there is a narrative that criticizes Western sanctions as unjust or ineffective, accompanied by mockery of the West and Ukraine. Frequent two-way communication is identified as narratives are promoted through replies in direct communication with other users. Quotes from Russian politicians and officials are extensively utilized to support these narratives.

Pro-Ukraine accounts actively disseminate narratives about Russia's war crimes in Ukraine, labeling Russian politicians as fascists, drawing comparisons between Russian actions and those of Nazi Germany. They also ridicule Russia, state propaganda, and state media narratives, likening Putin to Hitler and referring to Russia as a Nazi regime. These narratives advocate for sanctions against Russia, referencing previous conflicts and Russian war crimes in Chechnya and Syria, and describe the war in Ukraine as a genocide. Pro-Ukraine accounts aim to expose Russian propaganda and promote counter narratives, particularly targeting Russian-speaking audiences. They also praise the Ukrainian military forces, urging increased support and criticizing the West for not imposing sufficient sanctions and measures against Russia.

In the analysis of influencers within the Leiden communities, we identified five largest distinct groups and investigated top influencers in each group (see more information about Leiden groups in Table 4). The first group primarily consists of pro-Ukraine influencers and experts. The second group comprises Russia's liberal opposition figures and media entities that support Ukraine. The third group consists of Russia's propaganda actors and accounts, while Group 5 includes Russia's state-affiliated propaganda media and government entities. Group 4 comprises journalistic organizations covering the war.

Group 1 predominantly spreads pro-Ukraine narratives, highlighting Russia's war crimes, labeling Putin's regime as Nazi, and debunking Russian propaganda. Group 2 focuses on promoting anti-war narratives. Groups 3 and 5 primarily disseminate Russian state propaganda, blaming Ukraine for killing its own civilians, undermining Western support, and assigning blame to the West for escalating the conflict. Pro-Russia users in these groups express support for Russian troops and disseminate propaganda narratives about military biolabs, mentioning Nazi government in Ukraine, and claim that the US, West, and NATO support Nazi groups. These propaganda narratives present the invasion as an ideological war against NATO and Western hegemony. Group 4 consists of journalists covering the war in Ukraine.

5 Conclusion and discussion

This paper examines Russia's propaganda discourse on Twitter during the 2022 Russian invasion of Ukraine, focusing specifically on the narrative of "fascism/Nazism." Through a mixed-methods approach incorporating natural language processing, network analysis, community clustering, and qualitative analysis of influential users and communities the study aims to identify the prominent actors, communities, and narratives within this discourse.

The findings of this study contribute to the broader understanding of disinformation campaigns employed by governments on social media. By shedding light on the strategies, narratives, and communities associated with Russia's state propaganda discourse during the invasion of Ukraine, it enhances our knowledge of the evolving tactics used to manipulate public opinion and shape geopolitical dynamics. It also provides perspective on how counter-narratives are developed and identified in communication.

The Leiden clustering analysis reveals distinct communities with varying attitudes, ranging from pro-Ukraine to pro-Russia sentiments. Influential actors play a crucial role in shaping the narrative landscape. The study emphasizes the importance of monitoring and analyzing these actors for a nuanced comprehension of information dissemination.

Moreover, the BERTopic analysis provides insight into the general topics discussed in both English and Russian conversations. Notably, the discussions center around expected computational propaganda narratives, with variations in language use reflecting the geopolitical context.

In conclusion, the study contributes to the growing body of research on computational propaganda, emphasizing the need for continued monitoring of social media discourse for a nuanced understanding of evolving propaganda tactics. The findings underscore the importance of considering cultural and historical contexts in analyzing propaganda narratives and highlight the role of influential actors in shaping public opinion during geopolitical events.

Moving forward, it is crucial to continue research and efforts aimed at developing effective countermeasures against disinformation campaigns. This includes raising awareness among social media users about the presence and impact of computational propaganda, promoting media literacy, and improving the transparency and accountability of social media platforms. Collaboration between researchers, policymakers, and technology companies is essential in developing comprehensive strategies to combat the spread of harmful disinformation and protect the integrity of information in the digital age.

Author contributions
The contributions of each author to this research paper are as follows: 1. Iuliia Alieva: primarily focusing on research design, conducting data analysis, and paper's writing process. 2. Ian Kloo: engaging in data analysis activities; 3. Kathleen M Carley: managing the research process, providing research supervision, and offering valuable guidance and advice throughout the project's duration. All authors read and approved the final manuscript.

Declarations

Competing interests
The authors declare that there are no competing interests associated with the research presented in this paper.

References

1. Woolley S, Howard P (2017) Computational propaganda worldwide: executive summary. Computational Propaganda Project. https://ora.ox.ac.uk/objects/uuid:d6157461-aefd-48ff-a9a9-2d93222a9bfd
2. Nyst C, Monaco N (2018) How governments are deploying disinformation as part of broader digital harassment campaigns. Institute for the Future. Retrieved from https://bit.ly/2Mi8DYm
3. Weedon J, Nuland W, Stamos A (2017) Information operations and Facebook. Facebook Menlo Park
4. Wanless A, Berk M (2021) The changing nature of propaganda: coming to terms with influence in conflict. The world information war. Routledge, London, pp 63–80
5. Starbird K (2019) Disinformation's spread: bots, trolls and all of us. Nature 571(7766):449–450
6. Linvill DL, Warren PL (2020) Troll factories: manufacturing specialized disinformation on Twitter. Polit Commun 37(4):447–467
7. Smart B, Watt J, Benedetti S, Mitchell L, Roughan M (2022) # IStandWithPutin versus #IStandWithUkraine: the interaction of bots and humans in discussion of the Russia/Ukraine war. In: Social informatics: 13th international conference, SocInfo 2022, proceedings, Glasgow, UK Springer, Cham, pp 34–53
8. Geissler D, Bär D, Pröllochs N, Feuerriegel S (2022) Russian propaganda on social media during the 2022 invasion of Ukraine. ArXiv preprint. arXiv:2211.04154
9. Badawy A, Addawood A, Lerman K, Ferrara E (2019) Characterizing the 2016 Russian IRA influence campaign. Soc Netw Anal Min 9:1–11
10. Lukito J, Suk J, Zhang Y, Doroshenko L, Kim SJ, Su M-H, Xia Y, Freelon D, Wells C (2020) The wolves in sheep's clothing: how Russia's Internet research agency tweets appeared in US news as vox populi. Int J Press/Polit 25(2):196–216
11. Bastos M, Mercea D (2018) The public accountability of social platforms: lessons from a study on bots and trolls in the Brexit campaign. Philos Trans R Soc A, Math Phys Eng Sci 376(2128):20180003
12. Hemánus P (1974) Propaganda and indoctrination; a tentative concept analysis. Gazette (Leiden, Netherlands) 20(4):215–223
13. Jowett GS, O'Donnell V (2018) Propaganda & persuasion. Sage, Thousand Oaks
14. Woolley SC, Howard PN (2018) Computational propaganda: political parties, politicians, and political manipulation on social media. Oxford University Press, London
15. Hyzen A (2021) Revisiting the theoretical foundations of propaganda. Int J Commun 15(18)
16. Pena MM, Klemfuss JZ, Loftus EF, Mindthoff A (2017) The effects of exposure to differing amounts of misinformation and source credibility perception on source monitoring and memory accuracy. Psychol Conscious.: Theor, Res, Pract 4(4):337
17. Paul C, Matthews M (2019) Defending against Russian propaganda. The SAGE handbook of propaganda, 286
18. Ablazov I, Karmazina M (2021) Disinformation as a form of aggression: Ukraine and its partners amidst the Russian fake news. Politi Sci Secur Stud J 2(2):65–72
19. Leitenberg M (2020) False allegations of biological-weapons use from Putin's Russia. Nonprolif Rev 27(4–6):425–442
20. Maschmeyer L (2021) Digital disinformation: evidence from Ukraine. CSS Analyses in Security Policy, 278
21. Pomerantsev P, Gumenyuk N, Kariakina A, Borzylo I, Peklun T, Yermolenko V, Rybak V, Kobzin D, Montague M, Barbieri J, Innes M, Budu V, Dawson A (2021) Why Conspiratorial Propaganda Works and What We Can Do About It: audience Vulnerability and Resistance to Anti-Western, pro-Kremlin Disinformation in Ukraine
22. Tolz V, Hutchings S (2023) Truth with a Z: disinformation, war in Ukraine, and Russia's contradictory discourse of imperial identity. Post-Soviet Affairs, 1–19
23. Alieva I, Ng LHX, Carley KM (2022) Investigating the spread of Russian disinformation about biolabs in Ukraine on Twitter using social network analysis. In: 2022 IEEE international conference on big data (big data). IEEE, New York, pp 1770–1775
24. Alieva I, Robertson D, Carley KM (2023) Localizing COVID-19 misinformation: a case study of tracking Twitter pandemic narratives in Pennsylvania using computational network science. J Health Commun 28(1):76–85
25. Carley KMOR (2014) A toolkit for dynamic network analysis and visualization. In: Reda A, Rokne J (eds) Encyclopedia of social network analysis and mining. Springer, Berlin
26. Alieva I, Carley KM (2021) Internet trolls against Russian opposition: a case study analysis of Twitter disinformation campaigns against Alexei Navalny. In: 2021 IEEE international conference on big data (big data), pp 2461–2469
27. Alieva I, Moffitt JD, Carley KM (2022) How disinformation operations against Russian opposition leader Alexei Navalny influence the international audience on Twitter. Soc Netw Anal Min 12(1):80

28. Carley LR, Reminga J, Carley KM (2018) ORA & NetMapper. In: International conference on social computing, behavioral-cultural modeling and prediction and behavior representation in modeling and simulation, vol 3. Springer, Berlin, pp 3–7

29. Traag VA, Waltman L, Van Eck NJ (2019) From Louvain to Leiden: guaranteeing well-connected communities. Sci Rep 9(1):1–12

30. de Groot M, Aliannejadi M, Haas MR (2022) Experiments on generalizability of BERTopic on multi-domain short text. ArXiv preprint. arXiv:2212.08459

31. Grootendorst M (2022) BERTopic: neural topic modeling with a class-based TF-IDF procedure. ArXiv preprint. arXiv:2203.05794

32. OpenAI (2023) GPT-4 Technical Report. ArXiv preprint. arXiv:2303.08774

Publisher's Note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.