Article

#Azovsteel: Comparing qualitative and quantitative approaches for studying framing of the siege of Mariupol on Twitter

Michael Tschirky and Mykola Makhortykh

Institute of Communication and Media Studies, University of Bern, Switzerland

Abstract

Social media platforms play a major role in shaping how the public around the world perceives contemporary wars, including the ongoing Russian invasion of Ukraine. However, there are multiple challenges in studying how exactly these platforms represent violence and what aspects of it are made more salient by their users. One of these challenges concerns the broad range of qualitative and quantitative approaches used to study platform-based war framing and their different capabilities in tackling the large volume of available data. To address this challenge, the authors compare the performance of qualitative and quantitative approaches – i.e. qualitative content analysis and topic modelling – for studying how one of the key episodes of the Russian–Ukrainian war, the siege of Mariupol in 2022 was framed on Twitter over time. Their findings demonstrate that both approaches show the prevalence of human interest and conflict frames that aligns with earlier research on war framing in journalistic media. At the same time, they observe differences in the estimated visibility of less common frames, such as morality and responsibility frames, depending on what method is used.

Introduction

The large-scale Russian invasion in February 2022 signified a new stage in the Russian–Ukrainian war which has been ongoing since 2014. This shift is characterized not only

Corresponding author:
 Michael Tschirky, Fachhochschule Graubünden, Pulvermühlestrasse 57, Chur 7000, Switzerland.
Email: m.tschirky97@gmail.com

by the increased scale of destruction and the expansion of hostilities beyond the eastern regions of Ukraine but also by the larger amount of media attention compared with earlier stages of the war. The combination of easy access to Ukraine for foreign journalists (De Witte, 2022) and the large number of digital war-related content coming from the war zone (Bergengruen, 2022) enabled 24/7 war reporting which was further stimulated by the global audience's interest in the largest outbreak of violence in Europe since the end of the Second World War.

While legacy media, in particular TV, served as a primary source of information about the Russian invasion for many Western countries (Eddy and Fletcher, 2022), social media (e.g. Twitter and Facebook) also played a major role in informing the public about the course of the war. The Russian invasion has also been extensively discussed on both Western (Chen and Ferrara, 2022) and Russian social media platforms (Urman and Makhortykh, 2022). Similar to other recent wars (e.g. in Syria: Kayyali, 2022, or Ethiopia: Fisher, 2022), these platforms served as key constituents in the process of mediatizing the ongoing violence and defining how it is perceived worldwide.

While the importance of social media for framing the Russian invasion – i.e. by making some aspects of it more salient than others – is hardly in dispute, the question of how exactly platforms like Twitter frame the latest outburst of violence in Ukraine remains open. Studies looking at the use of social media (mainly during the earlier stages of the war) showed the multitude of aims which these platforms can serve in the context of war framing. These aims vary from reinforcing the state propaganda and legitimizing the violence (Gaufman, 2022; Makhortykh, 2018) to countering propagandist claims by spreading counter-narratives and debunking misinformation (Khaldarova and Pantti, 2020; Makhortykh and Lyebyedyev, 2015) to enhancing the war narration by professional journalists (Ojala et al., 2018; Pantti, 2019) and citizen journalists (Golovchenko et al., 2018; Makhortykh and Sydorova, 2017) to enabling peer-to-peer security advice (Schmüser et al., 2022).

One of the main challenges of addressing these complex roles of social media in the context of the Russian–Ukrainian war relates to the diverse methodological approaches used to study them. So far, the majority of studies rely on qualitative approaches ranging from discourse analysis (Gaufman, 2022) to close reading (Makhortykh, 2018) to qualitative content analysis (Schmüser et al., 2022). However, despite the usefulness of qualitative approaches, their application is limited by the relatively small samples of data to which these approaches can be feasibly applied. By contrast, quantitative approaches, such as topic modelling (Urman and Makhortykh, 2022), can be applied to large datasets, but the resulting observations often have lesser interpretative value, in particular in terms of framing analysis. Under these circumstances, there is a need to compare what insights about war framing can be generated by different approaches and how comparable they are.

To address this need, we compare the performance of two approaches – qualitative content analysis and topic modelling – for studying how one of the key episodes of the Russian invasion, the siege of Mariupol, is framed on Twitter. One of the largest cities in Ukraine before the Russian invasion, Mariupol was a key Ukrainian stronghold on the Sea of Azov coast and also the base of the Azov regiment, a major bogeyman of Russian propaganda.[1] These two factors contributed to Mariupol becoming a place of particularly brutal battles in February–May 2022, which resulted in the city being largely destroyed

before it was occupied by the Russian army. Using the two approaches noted above, we aim to compare what information about the framing of the siege of Mariupol they can provide and how this framing varied over the course of the siege.

The rest of the article is organized as follows. First, we briefly introduce the historical background of the war in Ukraine and scrutinize the course of the siege of Mariupol in 2022. Then, we discuss existing research on war framing and its challenges, in particular concerning the comparability of findings and the diverse methodological approaches used in the existing scholarship. After that, we introduce our dataset and discuss in more detail our methodological approach, including the three periods of the siege that we aim to compare. Then, we share our findings about the performance of the qualitative and quantitative approaches for analysing war frames. We conclude with the discussion of the implications of our comparison for the war framing research as well as the limitations of the current study.

Historical background

The beginning of the Russian–Ukrainian war can be traced back to the spring of 2014.[2] Following the ousting of the pro-Russian president of Ukraine, Viktor Yanukovych, in the course of the Euromaidan protests, Russia occupied and then annexed the Crimean peninsula. The crisis of legitimacy of the new Ukrainian government in the eastern and southern regions of the country together with the poor state of the Ukrainian army played an important role in the limited resistance to the annexation and encouraged the continuation of Russia's aggression.

In the spring of 2014, several Russian paramilitary groups (e.g. the Strelkov group) seized several administrative centres in Eastern Ukraine and declared the establishment of the Donetsk (DNR) and Luhansk People's Republics (LNR). The Ukrainian authorities retaliated with the use of force and conducted a series of offensives in the summer of 2014 which resulted in substantial territorial losses of the pro-Russian puppet states. These losses included Mariupol, one of the largest cities in south-eastern Ukraine and a major industrial centre, which was liberated in June 2014 by the Ukrainian paramilitary groups, including the Azov battalion.

The intervention of the Russian army in August 2014 resulted in a series of defeats of the Ukrainian forces (e.g. during the Battle of Ilovaisk) and culminated in the signing of the first Minsk agreements in September 2014. After the major outburst of violence in the winter of 2014–2015, when the Russia-backed insurgents started another offensive resulting in the capture of Debaltseve, an important transport hub between the DNR and LNR, the war transitioned to positional warfare. Except for occasional clashes, the next six years were characterized by a substantial decrease in hostilities with both the Ukrainian and the Russian side trying to make minor advances in the frontline zone to improve their strategic positions.

The nature of the war changed dramatically in February 2022, when Russia staged a large-scale invasion of Ukraine from multiple directions. While in the north and the east of Ukraine, the advancement of the Russian forces was halted, in the southern regions of Ukraine the Russian army managed a breakthrough. As a result, the Ukrainian defences were broken and the Russian army captured most of the coastal areas of the Azov sea, thus enabling the land bridge between the Russian territory and Crimea.

The siege of Mariupol turned out to be a key episode of fighting in the south. Despite Russian units already reaching Mariupol on 25 February and surrounding it by early March, it did not surrender quickly like other cities in the south of Ukraine (e.g. Kherson or Melitopol). Instead, the Ukrainian units, including the Azov regiment which turned Mariupol into its base following the liberation of the city in 2014, fought back. The fighting continued for almost three months and caused massive destruction attributed to the indiscriminate shelling of residential areas by the Russian forces. The situation was further worsened by the humanitarian crisis which, according to the Ukrainian officials (e.g. France24, 2022), was caused by Russian units preventing civilians from leaving the city and destroying the critical infrastructure.

By early April, the advance of the Russian forces resulted in the isolation of the pockets of resistance within the city. The shortage of ammunition and food caused most of the defenders to surrender by mid-April with the resistance continuing at the Azovstal plant held by the Azov regiment and Ukrainian marines and hosting numerous civilians. Despite the heavy bombardments and regular attacks, Azovstal held until 20 May, when the last Ukrainian forces there surrendered and Mariupol was fully occupied by Russia.

Literature review

Framing as a concept

Framing is one of the most used concepts in social sciences and humanities (Bryant and Miron, 2004). Originating from cognitive psychology, it looks at how social reality can be interpreted by making some aspects of it more prominent than others (e.g. Tuchman, 1978). The saliency of different aspects of specific issues is manipulated via frames which are often defined as 'principles of selection, emphasis and presentation composed of little tacit theories about what exists, what happens, and what matters' (Gitlin, 1980). The choice of frames determines how a specific issue is presented to the public and thus has a major impact on how it is understood and treated (Entman, 1993).

Originally, the concept of framing has been primarily applied to studying how specific issues are presented to the public by the mass media, in particular mainstream actors such as journalists and politicians (Lecheler and De Vreese, 2019). However, the rise of digital platforms has resulted in the expansion of the range of actors capable of promoting their interpretations of social reality to a broader audience. In addition to challenging the monopoly of the mainstream actors on framing important societal issues and interfering with their ability to use it to pursue these actors' agendas, digital platforms also enable the promotion of alternative frames (Ahmed et al., 2019).

In addition to the increasing recognition of the importance of adapting the framing paradigm to the distinct features of the digital media environment, a number of studies (e.g. Cacciatore et al., 2016; D'Angelo et al., 2019; Krippendorff, 2017) also call for the critical reconsideration of the concept of framing. The common source of criticism in this context concerns the often inconsistent definition of framing used in the scholarship which, together with its intensive use, causes overlaps between framing and other concepts in the field of communication (e.g. agenda setting or priming, see Cacciatore et al., 2016). Under these circumstances, the concept of framing arguably loses much of its

explanatory power and can impede – rather than advance – research on the media effects, which is one of the areas where it is particularly frequently used. The suggested solutions to these problems range from shifting the focus in framing research to 'the sequentiality of discursive interactions and observable behaviors' (Krippendorff, 2017: 96), thus making the operationalisation of the concept more concrete and amplifying its analytical potential, to abandoning the concept of framing in general and instead differentiating between specific forms of framing processes (e.g. equivalence and emphasis framing, see Cacciatore et al., 2016).

While these critical points are very valid, we agree with D'Angelo et al. (2019) who argue that, instead of moving beyond framing, it is important to move beyond the concept's limitations. While the use of framing within a field of media effects research has a number of shortcomings, the use of the concept has expanded far beyond this specific area and increasingly serves as a 'bridging' (Guenther et al., 2023) framework for research on a broad range of communication phenomena. Thus, framing is particularly effective for bringing together different interdisciplinary perspectives, in particular when used to study how specific phenomena (e.g. armed conflicts) are represented and not necessarily what are the effects of representation.

In the context of using framing as a prism for studying representation of societally relevant issues, there is a growing interest towards the new possibilities offered by the use of computational approaches and large volumes of textual data (Ali and Hassan, 2022; D'Angelo et al., 2019). A number of studies (e.g. Burscher et al., 2014; Khanehzar et al., 2019; Liu et al., 2019; Walter and Ophir, 2019) applied different natural language processing techniques ranging from the ensembles of supervised machine learning classifiers (Burscher et al., 2014) to the topic model networks (Walter and Ophir, 2019) to the long short-term memory neural networks and transformer models (Liu et al., 2019). While many of these techniques demonstrate solid performance, there is still limited understanding of how their performance compares with the traditional qualitative approaches for framing analysis, in particular considering the argued importance of combining big data approaches with close reading (e.g. D'Angelo et al., 2019), especially when dealing with the analysis of representation of highly nuanced phenomena, such as the ones related to mass violence.

War framing research and its challenges

One specific area where the digital turn in framing is particularly important is war framing. Similar to other types of crises, such as healthcare emergencies (Wicke and Bolognesi, 2020) and natural disasters (Houston et al., 2012), framing has a substantial impact on the course of wars by shaping how the public understands the wars' causes and what are the expectations about wars' outcomes. Such an understanding is essential for many aspects of the war, ranging from the degree to which the general public is eager to associate with the war goals or, instead, adopt an anti-war sentiment to whether more violence-oriented or, instead, de-escalating behaviour is viewed as more socially acceptable.

Until now, the majority of research on war framing focuses on journalistic media. In the case of Ukraine, which is our case study, a number of studies discuss how journalistic

media frame Russia's aggression, demonstrating a number of differences in terms of how it is represented in Ukraine and Russia (e.g. Nygren et al., 2018) as well as in other parts of the world (primarily, Europe; e.g. Fengler et al., 2020; Lichtenstein et al., 2019; Ojala et al., 2017). By contrast, there are relatively few studies looking at how the war is framed on social media platforms: the majority of them focus on regional platforms, such as VK (e.g. Gaufman, 2015; Makhortykh and Sydorova, 2017; Urman and Makhortykh, 2022), whereas the role of Western platforms such as Twitter remains less studied (for an exception, see Pantti, 2019).

In addition to the relative scarcity of research on war framing on digital platforms, in particular in the context of Ukraine, the assessment of the impact of platform-based framing on the representation of modern wars is complicated by two other challenges. The first of these challenges concerns the diverse set of methods used for studying the relationship between war framing and online platforms, in particular considering the large volume of available digital data. These methods range from visual content analysis (e.g. Makhortykh and Sydorova, 2017) to discourse analysis (Gaufman, 2015, 2022) to transformer-based computational models for text summarisation (e.g. Urman and Makhortykh, 2022).

Each of the above-mentioned approaches has its own advantages and benefits, in particular in terms of dealing with the large volumes of war-related content available on the platforms. However, the systematic investigation of what these differences mean for the analysis of war framing is currently lacking. Such an investigation is important for understanding whether insights generated by these methods can systematically differ and lead to the diverse assessment of the representations of the war. It can also facilitate the choice of the method for the researchers working on the topic of war framing, in particular considering the growing popularity of more quantitative methods in social sciences and humanities.

The second challenge relates to the comparability of conceptual constructs used for studying war framing. Over the years, multiple typologies of frames were introduced (De Vreese, 2005). The typology that is of particular interest to us relates to the generic and issue-specific frames. Issue-specific frames are particular for the context and the issue which is framed, whereas generic frames (i.e. conflict, responsibility, human interest, economic and morality, see Semetko and Valkenburg, 2000) are applicable for a broad range of contexts. Issue-specific frames are more specific but, because of this, they are difficult to compare, whereas generic frames are topic-agnostic and comparable (Lecheler and de Vreese, 2019: 4). This feature of generic frames contributed to their use for studying framing of different issues ranging from immigration (Kim and Wanta, 2018) to healthcare (Makhortykh et al., 2022).

The comparability of generic frames makes them particularly appealing for comparative research which is important considering the many questions remaining about the impact of digital platforms on war framing. However, despite the frequent use of generic news frames in other thematic areas, their application to studying the framing of wars has so far been limited. While several studies used some variation of generic news frames to study war framing (e.g. Dimitrova et al., 2005; Schwalbe, 2013), others (e.g. Makhortykh and Sydorova, 2017; Schwalbe and Dougherty, 2015) tend to rely on issue-specific frames. Under these circumstances, we argue that it is important to further examine how applicable are generic news frames for studying war framing on social media using different methodological approaches.

Data collection and analysis

To collect data, we used Academic Twitter API 2.0. On 27 June 2022, we retrieved all unique tweets (i.e. not retweets or replies) with the keyword 'mariupol' for three time periods (all in 2022): (1) 24 February–3 March; (2) 1–9 April; and (3) 13–20 May. The first period corresponds to the very beginning of the Russian invasion when Mariupol mostly remained under Ukrainian control, but Russian troops started surrounding it. During the second period, Mariupol was under siege and the Russian forces pressed into the city forcing the defenders to retreat to the Azovstal steel plant. The third period corresponds to the last days of the defence of Mariupol when Russian troops gained full control over Mariupol and the remaining Ukrainian defenders on Azovstal surrendered.

We extracted tweets in English using information about the tweets' language provided by the API. The decision to focus on English tweets was related to them being the most numerous in response to the 'mariupol' query in the Latin script as well as our reliance on language-sensitive methodology for data analysis. The complete dataset for all three periods consisted of 103,438 tweets.

To analyse collected data we used two approaches: a qualitative and a quantitative one. For the qualitative approach, we used qualitative content analysis, where two coders looked for the presence of five generic news frames using the set of questions from Semetko and Valkenburg (2000: 99–101) to identify if the frame is present. The decision was based on the number of features (i.e. positive responses to the questions) of individual frames observed per each tweet; in the case of a tie (e.g. the same number of features), the decision was made by the coder depending on which frame was perceived as the prevalent one.

The selection of frames was based on the work by Semetko and Valkenburg (2000), who identified five generic news frames: (1) conflict; (2) human interest; (3) economic; (4) morality; and (5) responsibility frame. The *conflict* frame emphasizes the conflict between certain groups or individuals. The *human interest* frame puts an emphasis on humanitarian and/or emotional angles of the framed issue and often involves personal stories related to it. The *economic* frame focuses on the economic aspects and consequences of a specific event or an issue. The *morality* frame deals with the social or ethical prescriptions and their implications for the interpretation of the framed issue. Finally, the *responsibility* frame focuses on the individual, group, or government responsible for a problem or a solution.

The qualitative content analysis was applied to three samples of tweets: one sample per each of the three periods noted above. Each sample was made of 100 tweets randomly extracted from the collection for the respective period. We used Cronbach's Alpha to measure the intercoder reliability; the aggregate value of Alpha was 0.66. Following the analysis of intercoder reliability, all disagreements between the coders were consensus-coded.

For the quantitative approach, we used topic modelling, a statistical technique used to identify sets of topics present in a textual dataset as well as the relationship between individual items in the dataset (e.g. documents) and the specific topics (Nikolenko et al., 2017). Compared with manual content analysis, topic modelling is capable of processing large volumes of textual data more effectively; however, its outputs have less interpretative value

than more qualitative techniques. Despite this limitation, a number of studies applied topic modelling to study framing on different issues ranging from the COVID pandemic (Wicke and Bolognesi, 2020) to the historical wars (Barna and Knap, 2023; Makhortykh et al., 2021) to the migration (Heidenreich et al., 2019).

To conduct topic modelling analysis, we used the latent Dirichlet allocation (LDA) technique implemented via the ldatuning R package (Moor, 2022). To identify the optimal number of topics, we used a selection of minimization/maximization metrics from earlier studies (i.e. Arun et al., 2010; Cao et al., 2009; Deveaud et al., 2014) implemented in the package. Specifically, we used these metrics to test the performance of the model with the range of topics between 2 and 40 and ended with 14 topics as the optimal option performance-wise.

Findings

Qualitative analysis of framing of the siege of Mariupol

We started our analysis by examining the presence of generic news frames in the qualitatively analysed sample of tweets coming from the three periods of observation. The results are summarized in Table 1 which demonstrates the prevalence of human interest and conflict frames that aligns with earlier observations about the framing of other wars by journalistic media (e.g. Dimitrova et al., 2005; Schwalbe and Dougherty, 2015). The responsibility and morality frames were substantially less visible (except in period 3) and the economic frame was consistently under-represented. Such a distribution of frames can be attributed both to the established practices of representing the war as well as the specific focus on the siege of Mariupol (e.g. the discussions of economic consequences of the war were likely to be more prevalent in the context of the overall Russian invasion and not necessarily in the case of Mariupol in particular).

The analysis also shows the variation in the visibility of tweets related to specific frames over the three periods we examined. During the first period, the conflict frame which emphasized the confrontation between specific groups of individuals was the most prevalent one. Its visibility can be attributed to the initial emphasis on the military aspects of the confrontation between Russia and Ukraine (e.g. 'Air strikes and MLRS attacks on #Mariupol according to local sources. Via @MarQs__'; OAICM1, 2022) which was prevalent during the first days of the invasion.

Table 1. Overview of the frames by time period and the total number / percentage of tweets.

Frame	Period 1	Period 2	Period 3	Total (%)
Responsibility	9	8	20	37 (12.3)
Human interest	31	53	32	116 (38.7)
Conflict	49	30	31	110 (36.7)
Morality	6	5	12	23 (7.7)
Economic	2	3	2	7 (2.3)
None	3	1	3	7 (2.3)

During the second period, where the heavy fighting had been going on for some time within Mariupol, human interest became the most prevalent frame. Its visibility can be attributed to the brutal nature of the urban street fights between the Russian and Ukrainian forces as well as the intensification of the humanitarian crisis and the growing awareness about war crimes committed by the Russian units both in Mariupol, but also other parts of Ukraine (e.g. the northern areas of the Kyiv region such as Bucha). An illustrative example of content articulating the human interest frame is the following tweet:

> What absolute nightmare. Each & every family in Kyiv, Kherson, Kharkiv, Mariupol is now thinking, we are ab to become the next Sarajevo, Grozny, Aleppo, what do we do? Many will stay & fight, willingly or not. Many will try to flee. Some won't be able to. Impossible choices. (theodora_dragos, 2022)

Finally, during the third period of the siege, the frames which were relatively under-represented during the first two periods became more visible. It is particularly the case of the responsibility and morality frames. The increase in their visibility can be attributed to the attempts to ensure the safety of the last Ukrainian defenders of Mariupol by empha-sizing the importance of them staying alive and not being executed as had happened with some Ukrainians who surrendered earlier (e.g. 'SAVE MARIUPOL – Sign the Petition! https://t.co/iey4v4SwPv via @ChangeorgSA'; FionaMa05008316, 2022) as well as emphasizing the responsibility of Russia and the international community for the lives of Ukrainian soldiers and civilians (e.g. 'Despite Russian assurances, aid still blocked for hard-hit Mariupol'; CarolynEvertso1, 2022).

Quantitative analysis of framing of the siege of Mariupol

Using LDA topic modelling, we identified 14 main topics related to the discussions of the siege of Mariupol on Twitter. Figure 1 shows these topics together with the five most common terms per topic. While interpreting the distribution of topics, it is important to keep in mind the unequal number of tweets per period with period 1 attracting the least of them.

The examination of topic modelling results highlights difficulties of applying this approach to studying war framing. While certain topics can be interpreted from the fram-ing point of view (e.g. 'evacu cross red besieg ukrain'), others (e.g. '"#mariupol #ukrain #russia #ukrainerussiawar #russian' or 'u 0001f1fa fe0f 0001f1e6 0001f1f7') have lim-ited interpretative value. In particular, the 'u 0001f1fa fe0f 0001f1e6 0001f1f7' topic looks rather cryptic: the five most common terms here are made of Twitter emojis which are treated as Unicode text elements by the LDA model. This specific topic aggregates shorter tweets which were mostly made of emojis.

At the same time, topic modelling does provide information which can be used to examine the framing of the siege of Mariupol on Twitter, in particular as certain topics can be connected to the five generic news frames. A number of topics (e.g. 'help defend save ukrain now', 'ukrainian stori photo hospit girl', 'evacu cross red besieg ukrain') can be interpreted as the ones related to the human interest frame, whereas others (e.g. 'ukrain russian russia forc fight', 'azov nazi ukrainian russian battalion') can be connected to the
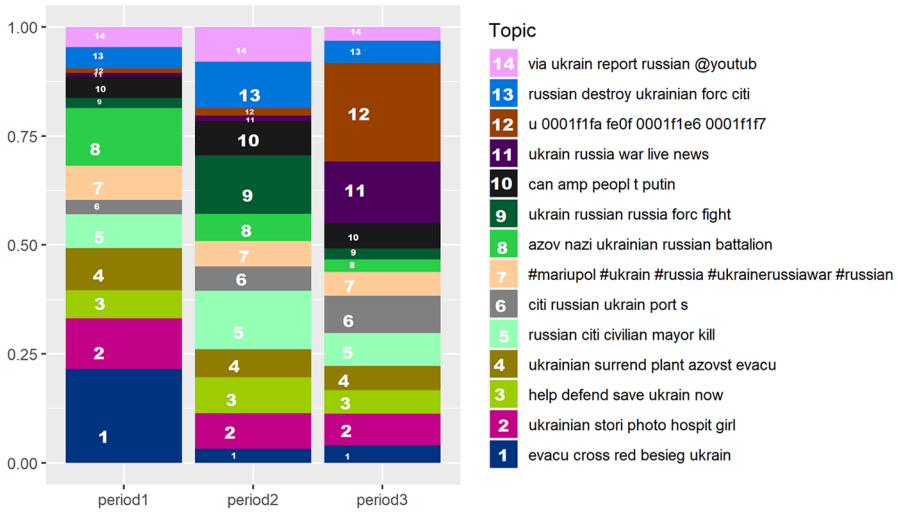
Figure 1. Outputs of LDA topic modelling per period.

conflict frame. Similarly to the findings of qualitative analysis, no topics seem to be explicitly connected to economic or morality frames. However, the responsibility frame seems to be more visible based on topic modelling data with at least one topic ('can amp peopl t putin') being potentially relevant for it.

In terms of the distribution of topics over the three examined periods, then similar to the qualitative analysis discussed in the previous section, we observed variation in what aspects of the siege of Mariupol were emphasized during each period. The topics potentially related to human interest (i.e. 'evacu cross red besieg ukrain' or 'russian citi civilian mayor kill') and conflict frames (i.e. 'azov nazi ukrainian russian battalion') were prevalent during the first period (i.e. the beginning of the siege). The prevalence of these specific aspects of the siege can be attributed to the desire to attract international attention to the humanitarian crisis threatening the city (from the Ukrainian side) as well as the propaganda campaign justifying the destruction in Mariupol by presenting it as a fight against Nazism (from the Russian side).

The second period was characterized by the human frame remaining quite visible (e.g. 'russian citi civilian mayor kill') with a different topic (allegedly related to the conflict frame; 'ukrain russian russia forc fight') also gaining more visibility. The decrease in the visibility of the evacuation-related discussions can be attributed to the growing recognition that Russian troops are not willing to enable evacuations of the civilian population to the Ukrainian territory. Interestingly, the topic potentially related to the responsibility frame (i.e. 'can amp peopl t putin') gained the largest visibility during this specific period.

Finally, during the third period, when the majority of Mariupol was occupied and Ukrainian resistance continued only at the Azovsteel factory, we observed a substantial decrease in the visibility of the topics prevalent during the earlier periods. Instead, the topic (i.e. 'u 0001f1fa fe0f 0001f1e6 0001f1f7') made of tweets composed primarily of emojis

became more prevalent. Its high visibility can be attributed to the increased emotionality of the discussion and the known tendency of social media users to rely on emojis to express strong emotional feelings (as contrasted to formulating them in the text; Kaye et al., 2016). Another topic, the visibility of which has increased, concerned the news updates (i.e. 'ukrain russia war live news') and per se was not connected to a particular frame.

Interestingly, the topic concerning the evacuation of the Ukrainian troops (i.e. 'ukrainian surrend plant azovst evacu') was not that visible during the third period which can possibly be attributed to it being more debated in the timespan between the second and the third periods of data collection. Overall, the third period turned out to be the one containing the least number of highly visible topics which could have been connected to specific generic frames. It can be attributed to it being associated with more emotional reactions to the end of the siege rather than with specific interpretations of this event through the generic frames.

Conclusions

In this article, we compared qualitative and quantitative approaches for studying the framing of the siege of Mariupol on Twitter in February–May 2022. Specifically, we looked at what insights in terms of the framing of this central episode of Russia's war in Ukraine can be generated using qualitative content analysis and topic modelling. We applied these two methods to examine the large set of tweets ($N = 103,438$) dealing with the three periods of the siege: the initial push of the Russian forces in February–March, the intense urban fighting in April and the surrender of the Ukrainian garrison in May.

The qualitative analysis of the siege's framing showed the unequal distribution of generic news frames. Similar to the observations of the framing of the earlier wars on legacy media (e.g. the US invasion of Iraq or the Second Lebanon war; Dimitrova et al., 2005; Schwalbe and Dougherty, 2015), we observed the strong prevalence of the human interest and the conflict frames. While morality and responsibility frames were less visible for the first two periods of data collection, their visibility has increased substantially for the May period associated with the surrender of the Ukrainian garrison in Mariupol. Finally, the economic frame was largely absent from the examined content and, while its presence could be expected due to the importance of Mariupol as an industrial and economic centre of the Azov region, we found little discussion of the economic implications of the siege.

Compared with the qualitative analysis, we found the outputs of quantitative analysis more difficult to interpret from the framing point of view. Some topics extracted via topic modelling were either too general to be connected to a specific frame or too difficult to interpret. Among those topics which were possible to connect to generic news frames, the largest number of topics seemed to be related to human interest and conflict frames, thus reiterating the main findings generated via qualitative analysis. At the same time, we observed a higher presence of the responsibility frame and less visibility of the morality frame (the latter, however, could be present in the topics which we associated with the human interest frame).

Together, these observations highlight two important points in relation to war framing research. The first of them concerns the comparison between the qualitative and quantitative approaches to war framing analysis. Despite the substantial differences between

the two approaches, our findings demonstrate that they largely align with each other, in particular in terms of showcasing the prevalence of two dominant types of frames (i.e. human interest and conflict). At the same time, there are also differences, in particular concerning the visibility of less prevalent frames with the qualitative approach highlighting the presence of the morality frame and the quantitative approach suggesting higher visibility of the responsibility frame.

While such a disagreement can be attributed to several reasons, we interpret it as a fine illustration of the strength and weaknesses of the two approaches. The reliance of a qualitative approach on the manual examination of a small sample of data makes it more likely that some frames might appear less common due to the random sampling not capturing enough content associated with them. For instance, in our case, the little visibility of the responsibility frame can be attributed to the relatively small random samples of tweets not including enough content with this particular frame. While the increase of the sample which is to be coded will amend this issue, the larger the sample becomes, the more difficult and time-consuming it is to code it.

By contrast, the quantitative approach is not susceptible to the issue with the sampling because it is capable of quickly processing the whole dataset. However, the problem of topic modelling is related to its limited granularity, i.e. the ability to capture fine-grained details in analysed data. It is especially an issue for complex detection tasks, such as the identification of generic news frames, in particular considering that in the case of war framing the line between some of these frames can be quite vague. In particular, the differences between the morality and human interest frame can be quite difficult to track, in particular based on a few top terms that we used to connect the outputs of topic modelling to individual frames.

The second point concerns the actual framing of the siege of Mariupol on Twitter. Despite many differences in terms of framing on social media and in journalistic media (e.g. in the case of framing conflicts; Ahmed et al., 2019), our observations about Twitter-based framing of this particular episode of Russia's war in Ukraine turned out rather similar to what we would expect based on earlier research focusing on war framing in journalistic media (i.e. the prevalence of human interest and conflict frames) with the exception of the increasing visibility of responsibility and morality frames at the end of the siege. One possible explanation of such a similarity can be attributed to the growing appropriation of the chaotic dynamic of social media-based war framing by mainstream actors (referred to as 'arrested war' by Hoskins and O'Loughlin, 2015) that leads to the growing alignment between journalistic and non-journalistic framing of armed conflicts.

Finally, it is important to note several limitations of the study. The first of them concerns our focus on comparing just two approaches – the qualitative and the quantitative one – whereas there is a larger number of approaches which can be applied to studying war framing. In particular, in the case of quantitative approaches, it is important to take into consideration more complex approaches, such as, for instance, the transformer-based ones which showed its applicability for studying the framing of the war in Ukraine on social media (e.g. Urman and Makhortykh, 2022). Another limitation concerns our exclusive focus on English tweets, whereas the comprehensive assessment of the framing of Russia's war in Ukraine on social media would require taking into consideration content in Ukrainian and Russian. Furthermore, to prevent the loss of contextual information, we opted for a

very limited preprocessing of data used for the quantitative analysis (e.g. we did not delete the emojis as shown by our resulting topic models). There is a possibility that through a more extensive preprocessing, including the deletion of emojis, our topic modelling could provide better results even if it would result in the loss of some contextual information. Finally, in the current study, we focused primarily on the representation aspect of the framing of the siege of Mariupol. Other framing aspects, however, are equally important and constitute important directions for future studies . For instance, future research can benefit from looking in detail at who were the actors involved in the process of framing the siege of Mariupol on Twitter (as well as other platforms) and investigate the possibilities and drawbacks of using qualitative and quantitative approaches for this task.

Funding

ORCID iD

Mykola Makhortykh ⃝iD https://orcid.org/0000-0001-7143-5317

Notes

1. The Azov battalion was originally a paramilitary group formed at the start of Russia's war in 2014. Among its original members, there were a number of football ultras as well as right-wing personalities, such as Andrii Biletskii. After the incorporation of the battalion in the Ukrainian National Guard and the subsequent expansion of Azov into a regiment, the majority of its extreme members left Azov. However, Russian propaganda kept framing Azov as a neo-Nazi armed group posing existential threat to Russia. For more information, see McCallum (2022).
2. For more information about the Russian–Ukrainian war prior to its recent escalation in 2022, see Kuzio (2017), Wood et al. (2015) or Zhukov (2016).

References

Ahmed S, Cho J and Jaidka K (2019) Framing social conflicts in news coverage and social media: A multicountry comparative study. *International Communication Gazette* 81(4): 346–371.

Ali M and Hassan N (2022) A survey of computational framing analysis approaches. In: *Proceedings of the 2022 Conference on Empirical Methods in Natural Language Processing*, Abu Dhabi, United Arab Emirates, 7–11 December, Association for Computational Linguistics, 9335–9348.

Arun R et al. (2010) On finding the natural number of topics with latent dirichlet allocation: Some observations. In: *Pacific-Asia Conference on Knowledge Discovery and Data Mining*, Hyderabad, India, 21–24 June. Berlin, Heidelberg: Springer, 391–402.

Barna I and Knap Á (2023) Analysis of the thematic structure and discursive framing in articles about Trianon and the Holocaust in the online Hungarian Press using LDA topic modelling. *Nationalities Papers* 51(3): 603–621.

Bergengruen V (2022) The battle for control over Ukraine's internet. *Time*. Available at: https://time.com/6222111/ukraine-internet-russia-reclaimed-territory/ (accessed 16 June 2023).

Bryant J and Miron D (2004) Theory and research in mass communication. *Journal of Communication* 54(4): 662–704.

Burscher B, et al. (2014) Teaching the computer to code frames in news: Comparing two supervised machine learning approaches to frame analysis. *Communication Methods and Measures* 8(3): 190–206.

Cacciatore MA, Scheufele DA and Iyengar S (2016) The end of framing as we know it . . . and the future of media effects. *Mass Communication and Society* 19(1): 7–23.

Cao J et al. (2009) A density-based method for adaptive LDA model selection. *Neurocomputing* 72(7/9): 1775–1781.

CarolynEvertso1 (2022) *Twitter*. Available at: https://twitter.com/CarolynEvertso1/status/1510130153493573634 (accessed 16 June 2023).

Chen E and Ferrara E (2022) Tweets in time of conflict: A public dataset tracking the twitter discourse on the war between Ukraine and Russia. *arXiv*. DOI: 10.48550/arXiv.2203.07488.

D'Angelo P et al. (2019) Beyond framing: A forum for framing researchers. *Journalism & Mass Communication Quarterly* 96(1): 12–30.

De Vreese CH (2005) News framing: Theory and typology. *Information Design Journal + Document Design* 13(1): 51–62.

De Witte M (2022) News about the conflict in Ukraine has been different – Stanford scholar and former war journalist discusses why. *Stanford*. Available at: https://news.stanford.edu/2022/03/17/reporting-war-ukraine/ (accessed 16 June 2023).

Deveaud R, SanJuan E and Bellot P (2014) Accurate and effective latent concept modeling for ad hoc information retrieval. *Document numérique* 17(1): 61–84.

Dimitrova D et al. (2005) War on the Web: The immediate news framing of Gulf War II. *International Journal of Press/Politics* 10(1): 22–44.

Eddy K and Fletcher R (2022) Perceptions of media coverage of the war in Ukraine. *Reuters Institute*. Available at: https://reutersinstitute.politics.ox.ac.uk/digital-news-report/2022/perceptions-media-coverage-war-Ukraine (accessed 16 June 2023).

Entman R (1993) Framing: Toward clarification of a fractured paradigm. *Journal of Communication* 43(4): 51–58.

Fengler S et al. (2020) The Ukraine conflict and the European media: A comparative study of newspapers in 13 European countries. *Journalism* 21(3): 399–422.

FionaMa05008316 (2022) *Twitter*. Available at: https://twitter.com/FionaMa05008316/status/1525219815014670346 (accessed 16 June 2023).

Fisher J (2022) #HandsoffEthiopia: 'Partiality', polarization and Ethiopia's Tigray conflict. *Global Responsibility to Protect* 14(1): 28–32.

France24 (2022) Civilians still trapped as Mariupol evacuation fails, says Ukraine deputy PM. *France24*. Available at: https://www.france24.com/en/europe/20220420-live-commander-in-ukraine-s-mariupol-warns-facing-last-days-if-not-hours (accessed 16 June 2023).

Gaufman E (2015) Memory, media, and securitization: Russian media framing of the Ukrainian crisis. *Journal of Soviet and Post-Soviet Politics and Society* 1(1): 141–175.

Gaufman E (2022) Damsels in distress: Fragile masculinity in digital war. *Media, War & Conflict*. DOI: 10.1177/17506352221130271.

Gitlin T (1980) *The Whole World is Watching*. Berkeley: University of California Press.

Golovchenko Y, Hartmann M and Adler-Nissen R (2018) State, media and civil society in the information warfare over Ukraine: Citizen curators of digital disinformation. *International Affairs* 94(5): 975–994.

Guenther L, Jörges S and Brüggemann M (2023) Framing as a bridging concept for climate change communication: A systematic review based on 25 years of literature. Communication Research. Epub ahead of print 20 January 2023. DOI: 10.1177/00936502221137165.

Heidenreich T et al. (2019) Media framing dynamics of the 'European refugee crisis': A comparative topic modelling approach. *Journal of Refugee Studies* 32: 172–182.

Hoskins A and O'Loughlin B (2015) Arrested war: The third phase of mediatization. *Information, Communication & Society* 18(11): 1320–1338.

Houston J, Pfefferbaum B and Rosenholtz CE (2012) Disaster news: Framing and frame changing in coverage of major US natural disasters, 2000–2010. *Journalism & Mass Communication Quarterly* 89(4): 606–623.

Kaye L, Wall H and Malone S (2016) 'Turn that frown upside-down': A contextual account of emoticon usage on different virtual platforms. *Computers in Human Behavior* 60: 463–467.

Kayyali D (2022) Digital memory, evidence, and social media. Lessons learned from Syria. *Sociologica* 16(2): 253–259.

Khaldarova I and Pantti M (2020) Fake news: The narrative battle over the Ukrainian conflict. In: Allan S et al. (eds) *The Future of Journalism: Risks, Threats and Opportunities*. London: Routledge, 228–238.

Khanehzar S, Turpin A and Mikolajczak G (2019) Modeling political framing across policy issues and contexts. In: Proceedings of the 17th Annual Workshop of the Australasian Language Technology Association, Sydney, Australia, 4–6 December, Australasian Language Technology Association, 61–66.

Kim J and Wanta W (2018) News framing of the US immigration debate during election years: Focus on generic frames. *Communication Review* 21(2): 89–115.

Krippendorff K (2017) Three concepts to retire. *Annals of the International Communication Association* 41(1): 92–99.

Kuzio T (2017) *Putin's War Against Ukraine: Revolution, Nationalism, and Crime*. CreateSpace independent publishing platform.

Lecheler S and de Vreese CH (2019) *News Framing Effects: Theory and Practice*. Abingdon: Taylor & Francis.

Lichtenstein D et al. (2019) Framing the Ukraine crisis: A comparison between talk show debates in Russian and German television. *International Communication Gazette* 81(1): 66–88.

Liu S et al. (2019) Detecting frames in news headlines and its application to analyzing news framing trends surrounding US gun violence. In: *Proceedings of the 23rd Conference on Computational Natural Language Learning (CoNLL)*, Hong Kong, China, 3–4 November. Stroudsburg, PA: Association for Computational Linguistics, 504–514.

McCallum A (2022) Much Azov about nothing: How the 'Ukrainian neo-Nazis' canard fooled the world. *Monash University*. Available at: https://lens.monash.edu/@politics-society/2022/08/19/1384992/much-azov-about-nothing-how-the-ukrainian-neo-nazis-canard-fooled-the-world (accessed 16 June 2023).

Makhortykh M (2018) #NoKievNazi: Social media, historical memory and securitization in the Ukraine crisis. In: Apryshchenko V and Strukov V (eds) *Memory and Securitization in Contemporary Europe*. London: Palgrave Macmillan, 219–247.

Makhortykh M and Lyebyedyev Y (2015) #SaveDonbassPeople: Twitter, propaganda, and conflict in Eastern Ukraine. *Communication Review* 18(4): 239–270.

Makhortykh M, Lyebyedyev Y and Kravtsov D (2021) Past is another resource: Remembering the 70th anniversary of the victory day on LiveJournal. *Nationalities Papers* 49(2): 375–388.

Makhortykh M and Sydorova (2017) Social media and visual framing of the conflict in Eastern Ukraine. *Media, War & Conflict* 10(3): 359–381.

Makhortykh M, Urman A and Ulloa R (2022) This is what a pandemic looks like: Visual framing of COVID-19 on search engines. *arXiv*. Available at: https://arxiv.org/abs/2209.11120 (accessed 16 June 2023).

Moor N (2022) Nikita-Moor/Ldatuning. *Github*. Available at: https://github.com/nikita-moor/lda-tuning (accessed 16 June 2023).

Nikolenko S, Koltcov S and Koltsova O (2017) Topic modelling for qualitative studies. *Journal of Information Science* 43(1): 88–102.

Nygren G et al. (2018) Journalism in the crossfire: Media coverage of the war in Ukraine in 2014. *Journalism Studies* 19(7): 1059–1078.

OAICM1 (2022) *Twitter*. Available at: https://twitter.com/OAICM1/status/1497962662130499588 (accessed 16 June 2023).

Ojala M, Pantti M and Kangas J (2017) Whose war, whose fault? Visual framing of the Ukraine conflict in Western European newspapers. *International Journal of Communication* 11: 474–498.

Ojala M, Pantti M and Kangas J (2018) Professional role enactment amid information warfare: War correspondents tweeting on the Ukraine conflict. *Journalism* 19(3): 297–313.

Pantti M (2019) The personalisation of conflict reporting: Visual coverage of the Ukraine crisis on Twitter. *Digital Journalism* 7(1): 124–145.

Schmüser J et al. (2022) 'Please help share!' Security and privacy advice on Twitter during the 2022 Russian invasion of Ukraine. *arXiv*. DOI: 10.48550/arXiv.2208.11581.

Schwalbe C (2013) Visually framing the invasion and occupation of Iraq in Time, Newsweek, and US News & World Report. *International Journal of Communication* 7: 239–262.

Schwalbe C and Dougherty S (2015) Visual coverage of the 2006 Lebanon War: Framing conflict in three US news magazines. *Media, War & Conflict* 8(1): 141–162.

Semetko HA and Valkenburg PM (2000) Framing European politics: A Content Analysis of Press and Television News. J*ournal of Communication* 50(2): 93–109.

theodora_dragos (2022) *Twitter*. Available at: https://twitter.com/theodora_dragos/status/1499448388248612864 (accessed 16 June 2023).

Tuchman G (1978) *Making News: A Study in the Construction of Reality*. New York, NY: Free Press.

Urman A and Makhortykh M (2022) My war is your special operation: Engagement with pro- and anti-regime framing of the war in Ukraine on Russian social media. *OSF Preprints*. DOI: 10.31219/osf.io/67snk.

Walter D and Ophir Y (2019) News frame analysis: An inductive mixed-method computational approach. *Communication Methods and Measures* 13(4): 248–266.

Wicke P and Bolognesi M (2020) Framing COVID-19: How we conceptualize and discuss the pandemic on Twitter. *PloS One* 15(9): 1–24.

Wood E et al. (2015) *Roots of Russia's War in Ukraine*. New York, NY: Columbia University Press.

Zhukov Y (2016) Trading hard hats for combat helmets: The economics of rebellion in Eastern Ukraine. *Journal of Comparative Economics* 44(1): 1–15.

Author biographies

Michael Tschirky is an independent scholars who has completed Bachelor's degree in Social Science at the University of Bern and is currently working as a scientific assistant at the Canton of Graubünden in Switzerland. His research interests are centred around the impact of digital platforms and online communities on the diverse societal phenomena.

Mykola Makhortykh is a postdoctoral researcher at the Institute of Communication and Media Studies at the University of Bern. In his research, Mykola focuses on politics- and history-centred information behaviour in online environments and how it is affected by the information retrieval systems, such as search engines and recommender systems. Recently, his research appeared in journals such as Memory Studies, International Journal of Communication, and Journal of Computational Social Science

Address: as Michael Tschirky. [email: m.tschirky97@gmail.com]