



Middlebury Institute of
International Studies at Monterey
Center on Terrorism, Extremism, and Counterterrorism

The Center on Terrorism, Extremism, and Counterterrorism

The Wagner Group's Social Footprint: A Time-Series Analysis of PMC World

MARCH 2023

Jason Blazakis, Kojin Glick, and Alice Shepard

About the Authors

Jason Blazakis is the director of the Center on Terrorism, Extremism, and Counterterrorism at the Middlebury Institute of International Studies (MIIS). He is also a professor at MIIS.

Kojin Glick is a graduate researcher with the Center on Terrorism, Extremism, and Counterterrorism at the Middlebury Institute of International Studies who focuses on textual data analysis, open-source intelligence, and the study of emerging threats.

Alice Shepard (nom de guerre) is a graduate researcher with the Center on Terrorism, Extremism, and Counterterrorism at the Middlebury Institute of International Studies who specializes in financial crimes, open-source investigations, and politics. The author has family connections in Russia and wishes to be anonymous for security reasons.

Acknowledgements

The authors gratefully acknowledge the support for this project provided by the Kathryn Wasserman Davis Collaborative in Conflict Transformation.

Introduction

Growing out of the Ukrainian Donbas War in 2014, the Russian private military company (PMC) known as the Wagner Group is inextricably entwined with the events in Ukraine. Given this history, it is unsurprising that with the escalation in Russian and Ukrainian conflict, the group has grown in its ambitions within Russian society and abroad. An understanding of the Wagner Group's public perception by Russian society is imperative to an accurate evaluation of the organization's next steps. For example, based on whether certain communities associate positively with the Wagner Group as a participant in civil society, different forward paths may be available to Wagner's founder Yevgeny Prigozhin's efforts to broker political and social influence. If Prigozhin can nurture a public presence among significant sectors of the Russian population, he can further cement his role as one of the decision makers that shape Russia's future policies.

For the purpose of understanding the changing nature of how conflict actors interact with domestic audiences, this paper offers several novel insights. By framing Prigozhin as a public figure and the Wagner Group as a broker for his influence, we were able to harness computerized statistical methods to underscore the significance of Prigozhin and the Wagner Group's new public orientation. To monitor how discourse around the Wagner Group changed, we monitored the lifespan of a channel on VK—a popular Russian social media platform—known as “PMCWorld.” We first tracked the change in sentiment from its inception in the middle of 2020 through the end of 2022. Sentiment analysis, though limited in its ability to reveal detailed insights about granular elements of the discussion, is a useful tool to aggregate how a community is feeling about issues relevant to its topic. Since the stated purpose of the “PMCWorld” community is to provide information about the Wagner Group to those interested,¹ the sentiment analysis offers more specificity in determining inflection points in the evolution of discourse surrounding the Wagner Group. We found that by projecting himself into popular culture, Prigozhin now has the tools and ambitions to exert incredible influence in Russian domestic affairs and set the stage for broader Russian foreign policy impact, like the expansion of the Russo-Ukrainian war.

Methodology for Sentiment Analysis

When deciding where to look for communities close to the Wagner group, we chose to scrape the social media platform VK due to its popularity specifically among Russian speakers. In 2020, VK enjoyed the second highest social media penetration rate with 76.4% of Russians with a VK account, below WhatsApp (80.9%) and above Instagram (63.7%).² Although scraping WhatsApp for sentiment would theoretically increase representativity, WhatsApp poses a significant challenge due to the private, end-to-end encrypted nature of activity of the app.

VK, which is structured similarly to Facebook, contains various accounts for users to follow and post on topic-specific forums. For this research initiative, we chose the community “PMC World,” a landing page for Russian speakers interested in private military company related topics. Currently, “PMC World” boasts the largest membership of any community related to the Wagner Group. Using computerized

¹ Вакансии в ЧВК Вагнер / Работа / Курсы / PMC.” VK, June 2020. <https://vk.ru/pmeworld>.

Putin, Vladimir. ““On The Historical Unity of Russians and Ukrainians.”” President of Russia, July 12, 2021.

² We Are Social, Hootsuite, and DataReportal. “Russia: Most-Used Social Media Platforms 2021.” Statista. Statista, May 19, 2022. <https://www.statista.com/statistics/867549/top-active-social-media-platforms-in-russia/>.

methods for collecting and parsing aggregate data,³ we identified over 150,000 posts from “PMCWorld”. After collecting the data, we applied a sentiment analysis model⁴ on each engagement (textual posts and comments), resulting in four additional data points: three softmax scores consisting of a neutrality score, a positivity score, and a negativity score, all of which add up to 1; and one argmax score that summarizes the sentiment, where 0 represents neutral, 1 represents positive, and 2 represents negative.⁵

Because the argmax summary of sentiment analysis lacks both granularity and continuity, we sought to construct a single summary variable, called the weighted differential, to evaluate the community’s sentiment in aggregate for any given time period. These weighted differential scores were calculated in order to distill the positive and negative softmax scores of engagements posted in a particular time frame into a single variable. To calculate the weighted differential, n_{positive} is the number of positive posts, Mdn_{positive} is the median value of the positive posts, n_{negative} is the number of negative posts, and finally, Mdn_{negative} is the median value of the negative posts. These values were aggregated by week—a time frame that was large enough to aggregate enough engagement samples, while also narrow enough to provide a significant degree of granularity across data points.

$$\textit{Weighted Differential} = \frac{((n_{\text{positive}} \times Mdn_{\text{positive}}) - (n_{\text{negative}} \times Mdn_{\text{negative}}))}{(n_{\text{positive}} + n_{\text{negative}})}$$

Finally, we removed weeks which lacked enough data to allow us to construct the weekly weighted differential. It is important to note that we removed an outlier, which was the data for the week of November 2, 2020. All posts coded positively read “Happy Holidays” in reference to the recent Unity Day, and had little to offer by way of analytical utility.

The utilization of quantitative methods to conduct analysis of sentiment is intended to clarify the following key points: First, by organizing the aggregate sentiment data over time, we seek to understand where major inflection points lie, or in other words, where aggregate public sentiment seems to change. Second, building on the research conducted in aggregate analysis, we seek to build a time-sensitive lexicon of this group, which acts as a window into the topics that are discussed between these inflection points.

Establishing a Timeline of Change in “PMCWorld”

³ Glick, Kojin. “Moonstripe/CTEC_WG_RESEARCH_RESOURCES_1.” GitHub. GitHub, December 7, 2022. https://github.com/moonstripe/ctec_wg_research_resources_1.

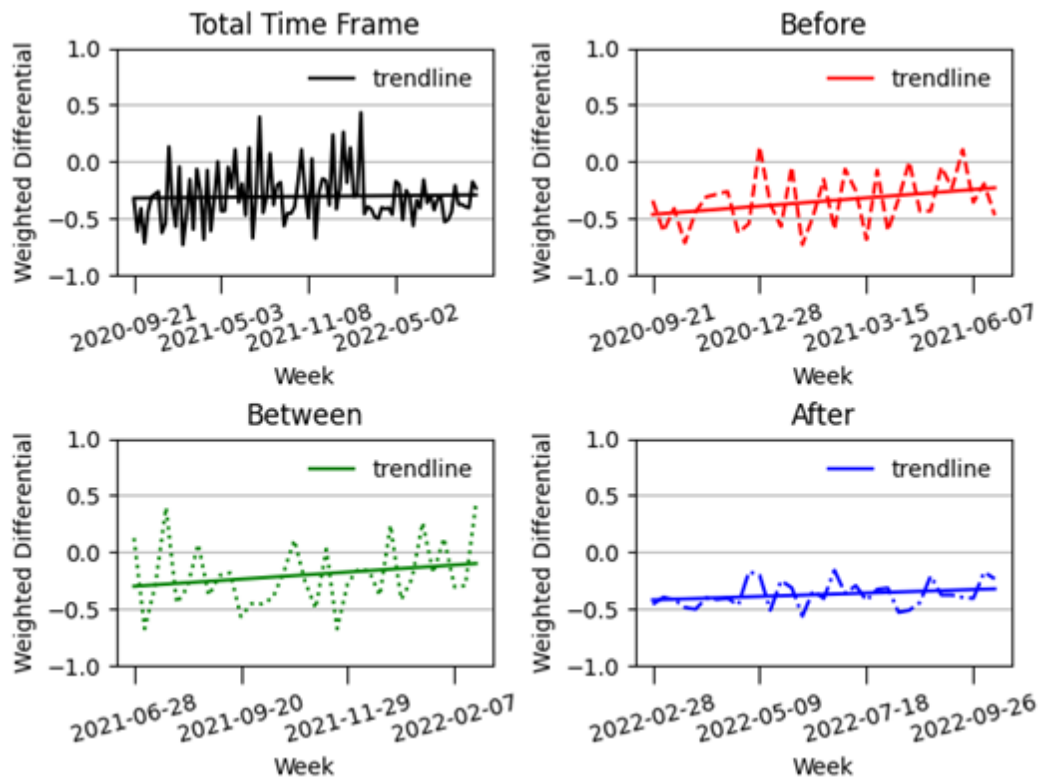
⁴ Blanchefort. “Blanchefort/Rubert-Base-Cased-Sentiment-Rusentiment.” Hugging Face. Hugging Face, January 10, 2021. <https://huggingface.co/blanchefort/rubert-base-cased-sentiment-rusentiment>.

⁵ Softmax and argmax functions are different ways of evaluating the output of a machine learning process. While discrete argmax scores are typically used in inference, we use softmax scores which accurately represent how confident the learning process is in labelling particular category (positive, neutral, or negative) in a value between 0 and 1. This allowed us to produce a single output directly into a continuous sentiment value from -1 and 1.

The chosen method of determining how the sentiment of “PMCWorld” varied is to test how the community’s reactions align with significant events in the Russian news cycle. After calculating the weighted differential, we looked at engagement sentiment across the total timeline of “PMCWorld” (Graph 1). From the graph of the weekly weighted differential over the whole lifetime of “PMCWorld”, the first of two spikes in weekly weighted differential occurs the week of Putin’s declaration that Ukraine and Russia are the same people (late June/early July 2021) and the second coincides with the outbreak of Russia’s “Special Military Operation” in Ukraine (late February 2022).

Due to the significant spikes related to these events, we executed a statistical test that would determine if the weighted sentiment of the period between these two major events were notably different from the equally sized periods before and after.

Graph 1: Summary of weighted differentials over time



Testing for a Significant Change between Periods

To determine whether the increased weekly weighted differential of the period between Putin’s announcement and the expansion of the Russia-Ukraine conflict was significantly higher than the period before and after, we chose to apply the paired-sample t-test. This was conducted to determine whether the aggregated sentiment, represented by weekly weighted differential, significantly varied before and after the period between Putin’s speech and the expansion of the Russia-Ukraine conflict.

To conduct this test, we first ensured that our sample sizes were equally divided. This was accomplished by removing weeks from the dataset which lacked enough data to construct the weekly weighted differential statistic, followed by a division of the dataset into three 33-week samples: one before Putin’s announcement, another between Putin’s announcement and the expansion of the Russia-Ukraine conflict, and finally one after the expansion of the conflict. While this satisfied the test’s requirement that the samples be of equivalent size, we also needed to satisfy the second assumption that the samples were taken from normally distributed populations. By conducting a Shapiro-Wilk test on the three sets of weekly weighted differentials and setting a threshold of 5% (see Table 1), we failed to reject the null hypothesis that the data was drawn from a normal distribution. In other words, the samples were each generated from a normal distribution, allowing us to apply the paired-sample t-test.

Table 1: Results of the Shapiro-Wilk Test

Sample	Statistics	p-value
Pre-announcement	0.97	0.44 > 0.05
Announcement through Expansion of Conflict	0.95	0.14 > 0.05
Post-Expansion of Conflict	0.95	0.17 > 0.05

H_0 : The samples are drawn from a normally distributed population.

H_A : The samples are not drawn from a normally distributed population.

Analyzing Sentiment Over Time

To determine whether the weekly weighted differentials of the engagements posted in the period between Putin’s announcement and the expansion of the conflict were significantly varied from engagements posted during other times, we applied the paired-sample t-test to all unique combinations of samples. We refer to the comparison between the period bounded by the announcement and the expansion of the conflict and the period before Putin’s announcement as the backwards comparison, and the comparison between the bounded period and the period after the expansion of the conflict as the forwards comparison.

We found that in both the forwards and backwards comparison, we rejected the null hypothesis that the paired population means are equal. As a control comparison, we compared the distribution of pre-announcement weekly weighted differentials to the post-expansion of the conflict. In this comparison, we failed to reject the null hypothesis that the paired population means are equal.

Table 2: Results of the Paired-Sample T-Tests

Comparison	Statistic	p-value
Backwards comparison *	-2.368	0.047 < 0.05
Forwards comparison **	3.040	0.005 < 0.05
Control comparison ***	0.554	0.583 > 0.05

H_0 : The paired population means are equal.

H_A : The paired population means are not equal.

*comparison of pre-announcement period and period between Putin’s announcement and the expansion of the conflict

**comparison of period between Putin’s announcement and the expansion of the conflict and post-expansion of the conflict

***comparison of pre-announcement period and post-announcement period

Analyzing relative word frequency over time

The next phase of our approach was to analyze the change in lexicon on “PMCWorld” to assess sentiment surrounding the group’s activities in these given periods. While the first quantitative intervention relied on statistical testing, this method relies on descriptive analysis of the vocabulary used by the community in aggregate. To accomplish this, we identified the top five most common words from each period, before Putin’s declaration, between Putin’s announcement and the expansion of the Russia-Ukraine conflict, and after the expansion of the conflict. Then, we calculated relative frequency as the number of instances where this word appears, divided by the sum of all instances of the five most common words. Finally, we graphed these relative frequencies for each period (Graph 2).

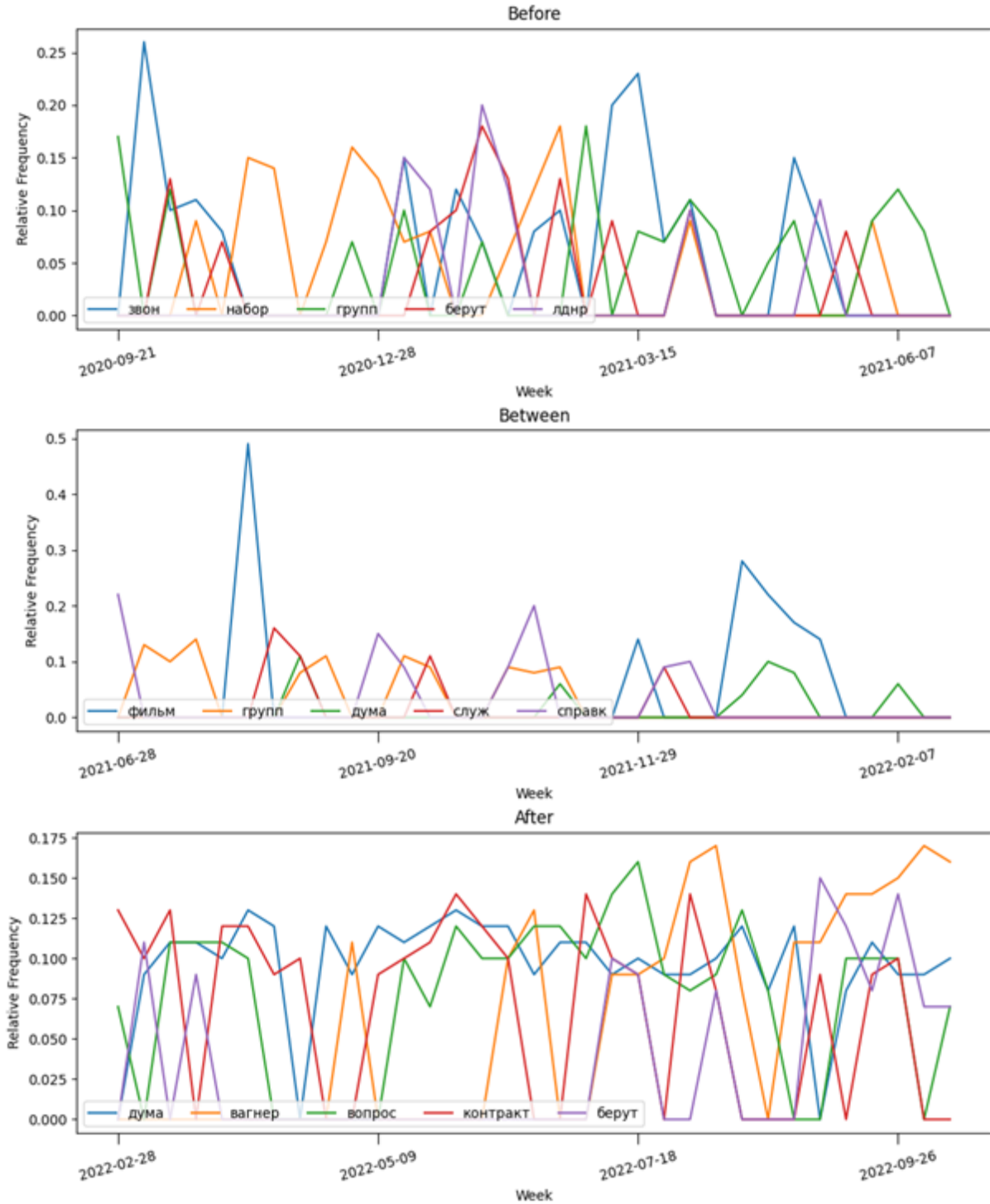
This strategy of relative frequency calculation reveals certain trends among the engagements found in the “PMCWorld” community during its lifespan. Firstly, in line with our expectations of the Wagner Group pivoting towards a more public role, significant use of the phrase “Wagner” (“Вагнер”) by the community only appeared after the start of 2022. Notably, the peaks in mentions of the group by name preceded the September 26th admission by Prigozhin that he has been leading the Wagner Group. In fact, the word “Wagner” (“Вагнер”) was the most mentioned word during the weeks of May 2nd, June 20th, the week of July 25th through the week of August 1st, and finally, the week of September 12th through the week of October 10th.

Secondly, the spread of frequencies is the widest in the periods before Putin’s announcement, where none of the most common words exceeded 25% relative frequency, and after the expansion of the conflict in Ukraine, where none of the most common words exceeded 17.5% relative frequency. During the period between these events, the word “movie” (“фильм”) exceeded 25% relative frequency, first during the

week of August 2nd and again during the week of December 27th. No other words among the most common exceeded 20% relative frequency during this period.

Graph 2: Relative frequency of the most common words over time⁶

⁶ Translated into English and in descending order of frequency, the most common words during the period labeled “Before”: ring (as in to call), kit, groups, take, LDNR (as in the Donetsk People’s Republic); “Between”: movie, groups, Duma, service, help; “After”: Duma, Wagner, question, contract, take



Discussion of the Findings

The results of the analysis revealed that aggregate sentiment changed most drastically on or around the week of June 28, 2021, and the week of February 28, 2022. These inflection points co-occurred with significant events regarding the state of Russian-Ukrainian affairs. For instance, in June 2021, Russian President Putin and U.S. President Biden met to discuss increased tensions over Ukraine, which ended

with Putin penning the essay, “On the Historical Unity of Russians and Ukrainians,” in mid-July 2021.⁷ The sentiment change in June 2021 was positive, raising the average weighted differential by 0.147 points from -0.348 to -0.201. In late-February 2022, Russia invaded Ukraine, resulting in a negative sentiment change which reduced the average weighted differential by 0.172—from -0.201 to -0.373. This data indicates a positive change in sentiment, though negative in absolute terms, during the period between Putin’s declaration of Ukrainian reunification and the expansion of the Russia-Ukraine war, which then subsided after the expansion of the war. To better understand the content of discussion during this period, a relative frequency comparison of the five most common words used in each of the three periods was conducted. Notably, the results indicated that *only during this period*, the relative frequency of the word “movie” (“фильм”), is higher than 25%.

There are two media events related to Prigozhin’s own activities that can help explain the sudden rise in mentions of the word “movie” (“фильм”) in “PMCWorld”. First, premiering in St. Petersburg on the 11th of August 2021, was “Hot Sunlight” (“СОЛНЦЕПЁК”).⁸ This drama-thriller takes place in the Luhansk Province of Ukraine in May 2014, and paints the 2014 onset of the conflict as victimizing for people with Russian connections. Second, premiering online on October 5th, 2022 was “The Best in Hell” (“ЛУЧШИЕ В АДУ”).⁹ This movie depicts brave Wagner Group soldiers engaged in a military conflict, portraying the group in a highly positive light. Both films were produced by Aurum Productions in conjunction with Paritet, which is a Russian production firm which has only produced the above two films. Most interestingly is that Yevgeniy Prigozhin owns a partial stake in Aurum, according to reporting by Google’s Threat Analysis Group.¹⁰

Conclusion

The public nature of Prigozhin and the Wagner Group’s recent activity enable researchers to observe the interaction between the Wagner Group and specific elements of Russian society. This paper offers a modern strategy of surfacing the state-of-the-art tools of influence that conflict actors like the Wagner Group can utilize in the information age. As groups like the Wagner Group take to the internet to recruit members, raise finances,¹¹ spread propaganda, and poll support, researchers and policymakers interested in their evolution can mobilize cutting-edge technologies like sentiment analysis to draw new insights and inform future actions with more nuance. As threat actors like the Wagner Group begin to adopt mainstream social media strategies to support their operation, readers are reminded of the depth of knowledge that Prigozhin, and Russian foreign policy writ large can employ to maximize the effect of propaganda in mass-media. The production of “Hot Sunlight” and “The Best in Hell” positions Prigozhin

⁷ Putin, Vladimir. “On The Historical Unity of Russians and Ukrainians.” President of Russia, July 12, 2021. <http://en.kremlin.ru/events/president/news/66181>.

⁸ “Солнцепёк (2021) - Фильм - Информация о Фильме - Российские Фильмы и ...” Kino Teatr. Accessed January 17, 2023. <https://www.kino-teatr.ru/kino/movie/ros/148053/annot/>.

⁹ “Luchshiy v Adu.” IMDb. IMDb, October 5, 2022. <https://www.imdb.com/title/tt22530256/>.

¹⁰ Tong, Tina. “Prigozhin Interests and Russian Information Operations.” Google. Google, November 10, 2022. <https://blog.google/threat-analysis-group/prigozhin-interests-and-russian-information-operations>

¹¹ Blazakis, Jason., Glick, Kojin. & Shepard, Alice. (2023). *From Military Adventurism to Venture Capitalism: The Changing Face of the Wagner Group Through Financial Analysis* [Unpublished manuscript]. Center on Terrorism, Extremism, and Counterterrorism, Middlebury Institute for International Studies at Monterey. pp. 1-2, p. 1

as a broker for nationalist sentiment and represents a significant effort to demonstrate the utility and the flexibility of the Wagner Group to tackle the most pressing issues facing Russia today.

Works Cited:

- Blanchefort. “Blanchefort/Rubert-Base-Cased-Sentiment-Rusentiment.” Hugging Face. Hugging Face, January 10, 2021. <https://huggingface.co/blanchefort/rubert-base-cased-sentiment-rusentiment>.
- Blazakis, Jason., Glick, Kojin. & Shepard, Alice. (2023). *From Military Adventurism to Venture Capitalism: The Changing Face of the Wagner Group Through Financial Analysis* [Unpublished manuscript]. Center on Terrorism, Extremism, and Counterterrorism, Middlebury Institute for International Studies at Monterey.
- Glick, Kojin. “Moonstripe/CTEC_WG_RESEARCH_RESOURCES_1.” GitHub. GitHub, December 7, 2022. https://github.com/moonstripe/ctec_wg_research_resources_1.
- “Luchshiye v Adu.” IMDb. IMDb, October 5, 2022. <https://www.imdb.com/title/tt22530256/>.
- Putin, Vladimir. ““On The Historical Unity of Russians and Ukrainians.”” President of Russia, July 12, 2021. <http://en.kremlin.ru/events/president/news/66181>.
- Tong, Tina. “Prigozhin Interests and Russian Information Operations.” Google. Google, November 10, 2022. <https://blog.google/threat-analysis-group/prigozhin-interests-and-russian-information->
- We Are Social, Hootsuite, and DataReportal. “Russia: Most-Used Social Media Platforms 2021.” Statista. Statista, May 19, 2022. <https://www.statista.com/statistics/867549/top-active-social-media-platforms-in-russia/>.
- “Вакансии в ЧВК Вагнер / Работа / Курсы / PMC.” VK, June 2020. <https://vk.ru/pmcworld>.
- С, Ассоциации. “Лето и Арбалеты | Вагнера.” YouTube. YouTube, August 11, 2022. https://www.youtube.com/watch?v=T0fSet9xC_w.
- “Солнцепёк (2021) - Фильм - Информация о Фильме - Российские Фильмы и ...” Kino Teatr. Accessed January 17, 2023. <https://www.kino-teatr.ru/kino/movie/ros/148053/annot/>.
- Петрова, Анна. “В Симферополе Снимают Драму о Событиях в Луганске 2014 Года - Фото.” РИА Новости Крым. РИА Новости Крым, May 22, 2021. <https://crimea.ria.ru/20210522/V-Simferopole-snimayut-dramu-o-sobytyakh-v-Luganske-2014-goda---foto-1119603028.html>.