Proceedings of the Nineteenth International AAAI Conference on Web and Social Media (ICWSM 2025)

A Telegram Dataset of Propaganda and its Moderation

Klim Kireev[1,2], **Yevhen Mykhno**[6], **Carmela Troncoso**[1,2], **Rebekah Overdorf**[3,4,5]
[1]EPFL
[2]Max Planck Institute for Security and Privacy (MPI-SP)
[3]Research Center Trustworthy Data Science and Security, University Alliance Ruhr
[4]Ruhr University Bochum (RUB)
[5]University of Lausanne
[6]Independent Researcher
klim.kireev@epfl.ch, carmela.troncoso@epfl.ch, rebekah.overdorf@unil.ch,

Abstract

Messaging applications like Telegram have evolved into de facto social networking platforms as they add features like broadcast channels and large groups. Yet, research on these aspects of Telegram is sparse compared to more traditional social media platforms. In this paper, we present a dataset of Telegram messages collected using the export API that returns channel histories, complemented by messages collected in real-time. This dual collection methodology allows us to label deleted messages, i.e., messages that are present in the real-time dataset but not the historical dataset. Additionally, we provide labels indicating whether messages have been sent by accounts belonging to one of two distinct propaganda networks. We provide experiments that show how this rich dataset of Telegram messages can be used to study moderation in Telegram, stances and trends on different topics, and to shed light on malicious behaviours present on Telegram. Finally, we outline other use cases where our dataset could help the research community better understand Telegram as a social network.

Introduction

Messaging applications like Telegram, Discord, and WhatsApp are widely used around the world and have greatly impacted how people communicate. Since its launch in 2013, Telegram has steadily grown its user base to 950 million monthly users as of the summer of 2024[1], making it one of the most downloaded mobile applications globally.

While the original use of messaging platforms was one-to-one communications, many have evolved into de facto social networks resembling platforms like Facebook or X in many ways. On Telegram, users can now participate in group chats with hundreds of thousands of users. These are similar in nature to Facebook groups or Reddit subreddits, where individuals can post and comment about certain topics. Users can also follow broadcast channels that share content via posts with their followers and allow for commenting and discussion on their posts. These are similar to Facebook pages, where businesses, news organizations, or individuals share content with their followers. These features have contributed to a blurring of the line between social networks and messaging apps.

[1]https://t.me/durov/337

Telegram's use also has grown in importance in a number of critical, real-world situations. Its lack of content moderation and large-scale channels for disseminating information quickly and widely, coupled with its widespread adoption has led to its use in crises including conflict situations. This includes its use by the government in the 2021 coup d'état in Myanmar (Advox 2023), its use by Hamas in the ongoing conflict in Gaza (BBC 2023), and its use by both sides of the Russo-Ukrainian War (Bergengruen 2022).

As in traditional social networks such as X (Zannettou et al. 2019), Instagram (Hosseinmardi et al. 2015), and Facebook (Scrivens and Amarasingam 2020), as Telegram becomes popular, it also attracts users that leverage the platform to spread malicious content such as toxic content (Guhl and Davey 2020), disinformation (Willaert et al. 2022), and propaganda (Kireev et al. 2024) on the platform. The lack of moderation on Telegram exacerbates this problem, as malicious content is rarely removed, and users are rarely banned from the platform as a whole (Wijermars and Lokot 2022).

Even though Telegram has increased in popularity and importance as a social network, and suffers from the same issues as traditional platforms, most studies on social networks still focus largely on X, Reddit, or Facebook. For example, in the ICWSM proceedings in 2024, 28 papers mention Twitter in their abstract and 18 mention Reddit, compared to just four that mention Telegram.

In this paper, we present a dataset of 17.3M Telegram messages collected from 13 channels. The dataset contains messages collected in real-time, as well as messages collected using the export API which allows the download of all recent messages in a channel. These two collection methods complement each other. The historical data collected via the export API contains data as far back as 36 months, however, it lacks messages that were removed by channel moderators or deleted by users. The real-time data contains these deleted messages but is collected over a shorter time frame of just two months.

Messages collected via the real-time API are labelled to indicate whether or not the message was eventually deleted (i.e., whether it appears in the historical data or not). Additionally, all messages in the dataset are labelled based on whether they are sent by accounts belonging to one of two distinct propaganda networks, one that spreads pro-Russian narratives and one that spreads pro-Ukrainian narratives, as we identified in our prior work (Kireev et al. 2024).

The dataset follows FAIR principles (Wilkinson et al. 2016), since it is published on a popular platform[2] (**F**indable), openly **a**ccessible, stored in user-friendly CSV format (**I**nterpretable), under a Creative Commons license (**R**eusable).

Related Work

Related Datasets

Although there is much more work on other social media platforms than on Telegram, there are still existing Telegram datasets that others have released for studies about social media. Additionally, others have collected and released datasets about propaganda on other social media platforms. In this section, we focus on other Telegram datasets.

Because of Telegram's popularity in Russia and other former soviet republics, many datasets about propaganda and telegram focus on Russian channels. Both Kloo, Cruickshank, and Carley (2024) and Hanley and Durumeric (2024) created datasets specifically about the Russo-Ukrainian conflict. Both sourced links for channels from external sources (links collected from Twitter and links collected from Russian media outlets respectively), allowing them to find many channels and create massive datasets. Both datasets are labelled according to the topics of the messages. Not all work in this area is focused on Eastern Europe. Venâncio et al. (2024) created a dataset of Telegram messages relating to the 2022 Brazilian Presidential Election. This paper also used Twitter to find Telegram channels.

All of the above datasets were created using the export API. Our dataset is comparatively small due to our duel collection strategy. While data from many Telegram channels can be collected with the export API, the real-time data collection does not scale. However, the data collected from the real-time API contributes an important feature to the dataset: it contains messages that were deleted by channel moderators. When studying malicious behaviours on Telegram this is of critical importance, as malicious activities are often moderated and do not exist in historical datasets.

Propaganda on Social Media

Research on propaganda has grown significantly in recent years, as its influence, especially when deployed on social media, becomes part of the public consciousness. Many of these works focus on the types of propaganda that are spread, for example public health-related propaganda (Sharma, Zhang, and Liu 2022) or propaganda around an election (Deb et al. 2019). Others focus on detecting propaganda, either by detecting the content itself (Khattar et al. 2019; Ferrara et al. 2016) or detecting the bots and accounts that spread it (Shevtsov et al. 2022a). In this paper, we present a dataset with messages labelled as being part of a propaganda network of coordinating accounts or belonging to normal users. By identifying the method of efficiently spreading propaganda messages (i.e., duplicating messages), we were able to create a rich dataset of propaganda content being actively spread in both Russia and Ukraine.

[2] https://zenodo.org/records/14661891

Dataset Collection

Telegram offers two methods for collecting data from channels. The most straightforward method, used in prior works involves directly exporting data from a channel in a single download. However, this method does not capture deleted or moderated content, which is relevant for some analyses, especially when studying malicious content that might have been removed or deleted, e.g., Elmas et al. (2021). We supplement the historical data collection with data collected in real-time, which captures potentially malicious content before it is removed. This dual collection ability is not unique to Telegram and has been exploited by others on Twitter (Chen and Ferrara 2023; Shevtsov et al. 2022b).

In the following we describe our collection method, and we summarize the characteristics of the resulting dataset in Table 1, broken down by channel.

Historical Data Collection

The "Export Chat History" API, allows to retrieve past channel content as a JSON dump. Although this collection method itself is simple, the documentation is scarce and does not state how much of the data is returned. In our collection, we noted that for less active channels the API returned the past 36 months of data. For the more active channels there appeared to be a limit on the size of the dump, and the API only returned the most recent posts and comments. To avoid losing data for the more active channels, we requested data from the API periodically to prevent gaps due to the volume of messages.

Real-time Data Collection

To ensure that we also have access to deleted or removed content, we supplement the historical data collection with data collected via the real-time data collection API. This API has a workflow similar to other social networks. We first obtain an API key by registering our application on the Telegram website using the Telegram account of one of the authors. We then created a custom Telegram client via the Telethon library[3] using the API key provided by Telegram and deployed it on a dedicated, secure server for two months (August 16 — October 16, 2023). The client's workflow is the following: 1) the client receives new events for new messages posted to any channel that the account is a part of, 2) the client checks to see if the message comes from one of the monitored channels, 3) if so, the client saves the message to a JSON string.

Channel Selection

The main challenge to collecting Telegram data is channel selection as the relevant channels must be known in advance of the data collection. This is because on Telegram the search functionality for finding content via keywords is

[3] https://docs.telethon.dev/en/stable/

Channel	Subscriptions	Content	Stance	Hist. data	Real-Time data	Earliest Message Date
Readovka	2.3M	Politics	Right-wing	2.71M	863K	25.05.2023
Topor	1.25M	Entertainment	Neutral/Mixed	1.15M	297K	07.01.2023
Nexta	1.02M	Politics	Neutral/Mixed	1.61M	824K	26.06.2023
Rtrus	809K	Politics	Right-wing	2.26M	584K	16.05.2023
KK	492K	Entertainment	Neutral/Mixed	1.36M	158K	06.10.2022
Ru2ch	479K	Mixed	Neutral/Mixed	3.25M	862K	27.04.2023
Agitprop	101K	Politics	Left-wing	720K	-	22.10.2022
Murz	97K	Politics	Right-wing	566K	-	21.10.2020
Shtefanov	78.3K	Politics	Neutral/Mixed	1.44M	281K	31.01.2023
Donrf	41.2K	Politics	Right-wing	-	90.6K	16.08.2023
SpecchatZ	26.9K	Politics	Right-wing	1.00M	359K	25.07.2023
Rudoi	26.9K	Politics	Left-wing	417K	47.3K	11.12.2021
SamaraNews	17.9K	Mixed	Neutral/Mixed	7.7K	5.0K	24.02.2022

Table 1: *Telegram dataset summary.* Characteristics of the channels that we selected including: their popularity (number of subscriptions), topic, the number of messages collected (via both export and real-time APIs) and the date of the earliest message collected via the export API.

very limited. This is in stark contrast to most social media platforms, e.g., Reddit where keyword searches can be executed via the API[4].

We selected channels according to two criteria. First, popular channels that we identified using the TGStats[5] catalogue. Second, smaller channels that we identified manually as having apparently automated propaganda activity as described in our prior work (Kireev et al. 2024).

Top Channel Lists Among the top ten channels listed in the Russian TGStats catalogue, we selected the channels in which we observed propaganda activity: *Readovka*, *Ru2ch*, *Topor*, *KK*, and *Rtrus*. We added one channel in the top Belorussian catalogue (*Nexta*). We found no Russian-language channel in the top Ukrainian channels that appeared to have the type of automated propaganda activity we described.

Manual Selection of Smaller Channels To supplement the larger, popular channels, we manually identified six less-popular channels that contained propaganda messages. Four small (*Shtefanov*, *Rudoi*, *Samaranews*, and *Donrf*) and two mid-sized (*Murz* and *Agitprop*) channels. We encountered the mid-sized channels near the end of the data collection process, so for these two channels, we only collected historical data.

Control Channel Finally, we added a control channel to the data set *SpecchatZ* in which no propaganda messages were identified.

Dataset Annotation

Channel Content

As part of the channel selection process, we aimed to have a diverse cross-section of Russian-language Telegram channels. We manually labelled each channel according to the main topic of the posts published. The majority of channels (9) are focused on politics, 2 are focused on Entertainment and 2 had varied content. We also labelled the general stance of the posts and the comments users made on these posts. Our selected channels show quite diverse stances, with five right-leaning channels, two left-leaning channels, and 6 channels with either a variety of stances or neutral stances. Table 1, columns 'Content' and 'Stance' shows these labels per channel.

Manual Labelling of Propaganda Messages

There are many different types of propaganda on social networks like bots spreading political propaganda about health crises (Robles et al. 2022), fake news during elections (Bovet and Makse 2019), or state-sponsored news with a specific narrative (Solopova et al. 2023). In this work, we focus on a specific propaganda type that we witnessed on Telegram: networks of coordinated accounts that reply to messages that contain certain *trigger* words or phrases with politically charged, pre-written replies. We describe this propaganda in more detail in Kireev et al. (2024), which aims to build a detector that can be deployed by channel owners. Below, we summarize the labelling process that we executed in our prior work.

We first observed these networks operating on multiple Russian-language networks in July 2023 and have since also been described by an anonymous activist group Vox-Harbour (2024). We observed two distinct traits of propaganda account that we use as a heuristic for labelling. 1) the accounts have random or Western-looking usernames (e.g., "Mark"), and 2) their replies are often disconnected from the message they are replying to. That is, replies might be on the same topic as the message they are replying to, but they do not fit logically into the conversation or are only tangentially related. Also, propaganda messages messages lack "bridge words" that are typically used to connect responses to the message they are replying to (e.g, "Yeah" and "So").

[4]https://www.reddit.com/dev/api/#GET_search
[5]tgstat.com

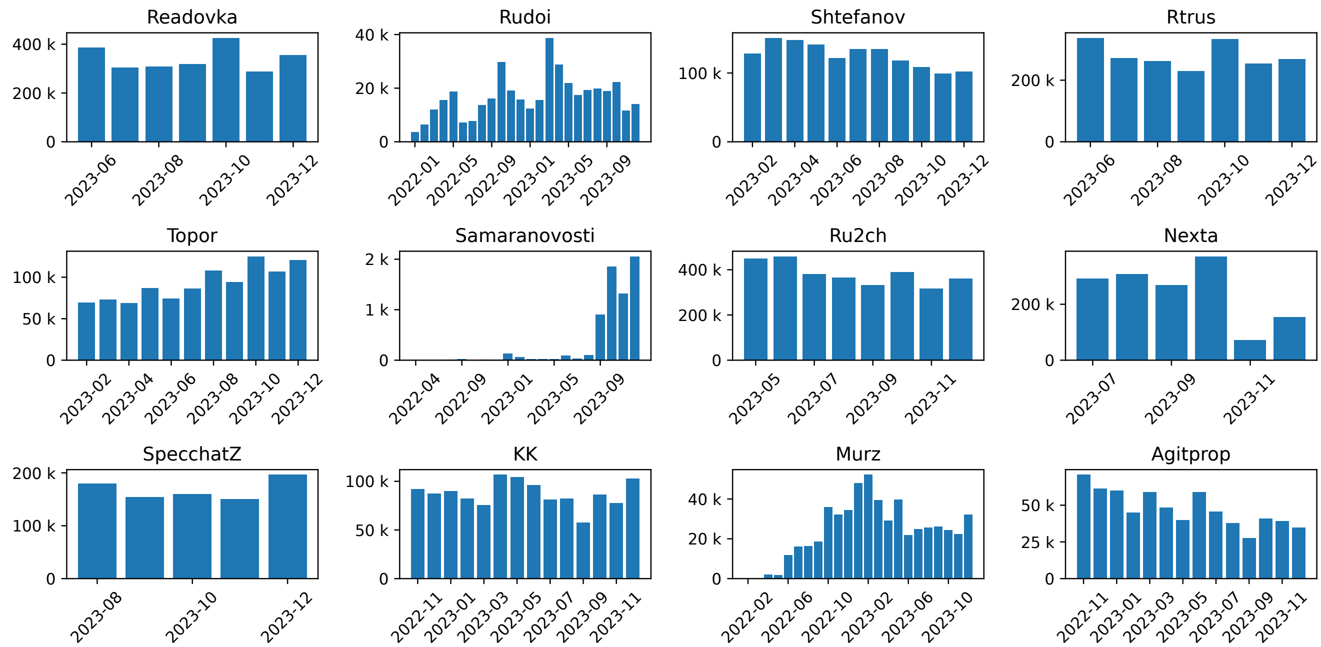

Figure 1: Channel activity over time.

Using these heuristics, two native-Russian-speaking authors independently, manually labelled all messages that we collected from the *Ridoi* channel. Each author iterated over the messages. If both heuristics were found to be true for a message, the message was marked as propaganda. The Cohen-Kappa agreement was $\sim 95.7\%$, indicating very high agreement.

Next, all messages by accounts that posted messages labelled as propaganda were also (automatically) marked as propaganda. We manually confirmed that all automatically propaganda-labelled messages were, in fact, propaganda.

Snowball Labelling

During the labelling process, the annotators noticed that some long propaganda messages appeared multiple times. But this never happened with normal user messages. These duplicate messages were posted by different accounts. We use this fact to uncover propaganda accounts in other channels of our dataset.

This snowball labelling process worked as follows: We took a list of long manually-labelled messages (at least 30 characters). Then, we searched for exact matches of these messages in other channels. If a match was found, *all* messages posted by that account in any channel were manually checked to be propaganda, and added to the seed list. We repeated this process until no messages were added to the list.

After the snowball process we ended up with 83.9K RealTime (239K in total) propaganda messages. Table 2 shows the distribution per channel as well as the prevalence of propaganda in that channel.

Channel	Propaganda Message Count	% of Propaganda
SamaraNews	270	5.4%
Readovka	37.11K	4.6%
Donrf	2.2K	2.4%
Rtrus	13.43K	2.3%
Nexta	18.13K	2.2%
Rudoi	804	1.8%
Topor	3.86K	1.3%
Shtefanov	2.25K	0.8%
KK	316	0.2%
Ru2ch	0	0%
SpecchatZ	0	0%

Table 2: *Labelled Dataset Results*. The percentages denote the ratio of propaganda messages to the total number of messages per channel. Note that the control channel (*SpecchatZ*) has no propaganda messages even after the snowballing. *Ru2ch* had no propaganda messages during the collection period, but we manually observed propaganda activity outside of the collection period.

Data Analysis

In this section, we provide a basic analysis of our dataset in terms of general characteristics, propaganda activity, and moderation. Then, we provide an example on how our dataset can be used to study opinion changes during the Russo-Ukrainian conflict.

(a) pro-Russian (b) pro-Ukrainian

Figure 2: Wordclouds for different clusters in pro-Russian and pro-Ukrainian propaganda networks

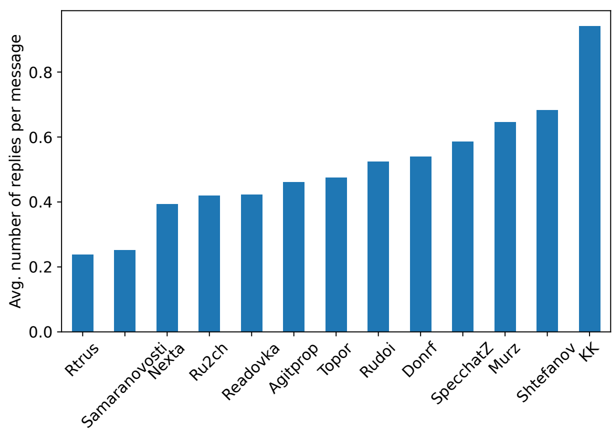

Figure 3: *Average account-to-account replies per message.* For each channel, the average number of replies per message is reported. Channel owner accounts are excluded from this evaluation, since all first-level comments are replies to their messages (channel posts).

Language	Number of Messages
Russian	13.1M
Ukrainian	207K
English	33K
Belarusian	18K

Table 3: Language distribution in the dataset.

General Characteristics

Language Distribution. The language detection problem for our dataset is complicated due to the short-message nature of Telegram communicationsWe used the language detection method (Stahl 2024) specifically tailored to detecting languages in short text. However, even this method showed a limited performance and required subsequent manual investigation. E.g., originally it spotted a significant amount of Bulgarian language in the corpus, which upon inspection appeared to be either short or new words in Russian. The results are reported in the Table 3.

Channel Activity. Figure 1 illustrates the activity over time in different channels from our dataset. We excluded the *Donrf* channel from this analysis since we do not have historical data. We see that channels popularity over time varies and no channel shows a clear consistent trend in the period of observation.

User Interaction. Telegram groups often serve as a space for users to share their opinions on various topics. Users not only comment on channel posts, but also engage in discussions with each other. Figure 3 shows that users are active in all the channels, with varying degree of inter-user interaction. The least interactive groups are Russia Today (*Rtrus*) and SamaraNews (*Samaranovosti*), which are media outlet channels where users are likely to consume content and not to talk to each other.

Propaganda Networks

Using the labelling procedure described above, we found one relatively large network of propaganda accounts containing the pro-Russian narratives. Later, we also found a significantly less active pro-Ukrainian network.

We conducted a textual analysis for both networks using a SBERT model (Reimers and Gurevych 2019) fine-tuned to the Russian language (SberDevices 2022). We first used SBERT to convert the 60K unique propaganda messages in the dataset into embeddings, and then clustered these embeddings using DBSCAN (Ester et al. 1996). We refined the cluster manually, attaching unclustered messages to specific clusters with similar topics. We then manually named the resulting $\sim$180 clusters using keyword searches (e.g. "Zelensky") to identify their topics. We follow the same procedure for the pro-Ukrainian propaganda messages, yielding 19 clusters. We present the wordcloud graph for all clusters in the Figure 2.

The pro-Russian propaganda covers a wide spectrum of topics, ranging from feminism criticism ("Feminism") and domestic issues like vape smoking legislation ("Vapes"), to international issues weakly related to Russia such as the Armenia-Azerbaijan conflict ("Armenia-Azerbaijan"), or the Israel-Hamas war ("Palestine_Israel"). The most popular topics, however, are directly related to the Russo-Ukrainian conflict ("Zelensky_hate", "Peace_negotiations", "Nazi_in_Ukraine").

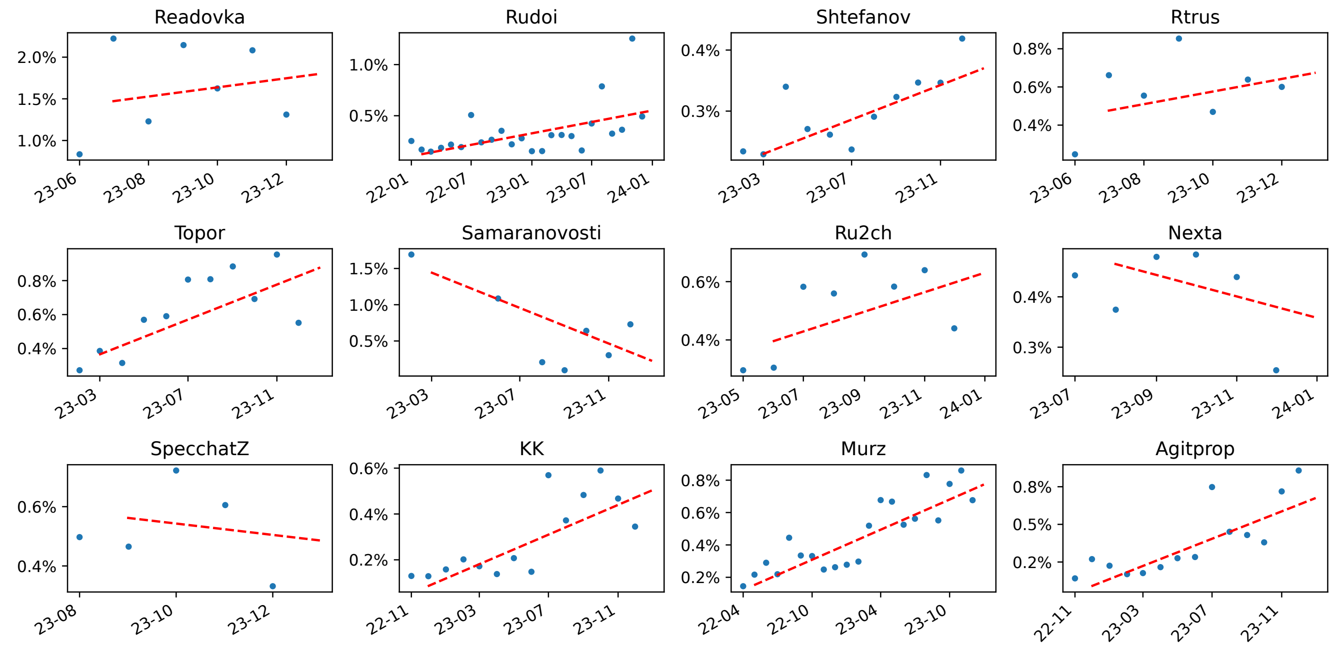

Figure 4: Immigration topic importance over time.

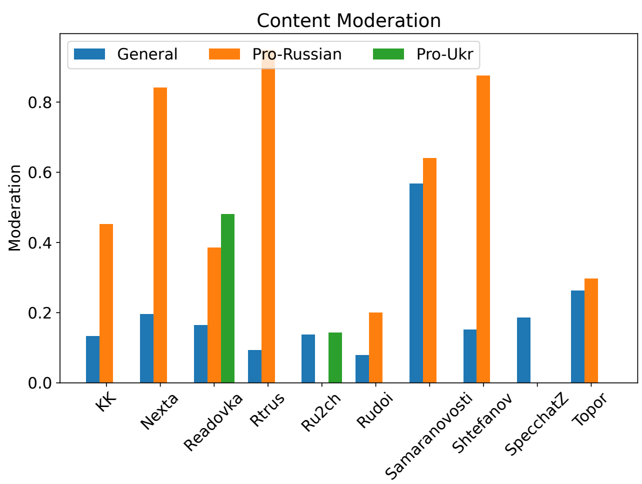

Figure 5: Content moderation across different channels.

In contrast, the pro-Ukrainian propaganda, probably due to its smaller scale, focuses mostly on internal Russian issues ("Human_right_violation", "Concentration_of_power") and war-related issues ("War_is_bad_for_all", "Nuclear_blackmail"). Interestingly, the most popular topic is composed from messages targeted at ethnic minorities ("Separatism_in_Russia"), pushing the narrative that these minorities should exit Russia and form independent states.

Moderation

One of the main advantages of the hybrid Real-time/Historical data approach is the possibility to study content moderation in Telegram, both with respect to propaganda moderation and message moderation in general (total amount of messages deleted by the moderators). In Figure 5, we report the average amount of moderation across the studied channels. All the channels perform content moderation in various degrees, deleting 10-50% of all messages depending on the channel. Regarding the propaganda moderation there is a significant discrepancy between channels: while for some channels (*Topor*, *Samaranovosti*) the moderation rate for propaganda messages is roughly the same as general moderation rate, other channels (*Rtrus*, *Nexta*, *Shtefanov*) demonstrate high selectivity for propaganda messages. The reasons for such selectivity in case of state-owned Russian media outlet (*Rtrus* - Russia Today) remain unclear.

Topic Trends

Another possible application of our dataset, is analysing how public opinion in Russian-speaking channels for certain topics changed during the conflict. To provide an example of such analysis, we measure the immigration importance dynamics along the course of the conflict. The results are reported on Figure 4. The analysis shows that in most channels, including the left-wing ones (*Rudoi*, *Agitprop*), there is an increase in messages mentioning immigration, which may indicate a rise of right-wing narrative, and nationalistic agenda in general during the conflict.

Limitations

The main limitation of the dataset is its relatively small channel selection. Due to the real-time data collection, we could not scale our dataset to the extent of (Baumgartner et al. 2020). Collecting the real-time dataset at larger scale can be a promising direction for the future work. There is a minor method limitation, since we cannot distinguish if a user deleted their own message or it was deleted by the moderators. However, we do not have any evidence supporting that users delete their messages in the public groups in any significant quantity. Another limitation to this dataset is its lack of linguistic diversity. As Table 3 demonstrates, our dataset is predominately in Russian, with a relatively small number of messages in other languages.

Finally, our dataset is labelled according to whether a message was posted by an account that is part of one of the two propaganda networks that we identified, not whether the message itself is propaganda. All of the messages that are labelled as propaganda are propaganda by our definition, however, we do not consider other types of malicious behaviours that might be present in the dataset nor do we consider other forms of propaganda, e.g., propaganda from other users or networks that do not follow the repeated messages pattern.

Discussion

In this section, we discuss other possible use cases for our dataset, as supported by the analysis we completed in this work.

Moderation Labels

One unique factor of our dataset is the dual collection methodology. That is, our dataset is made up of two distinct data collection pipelines (from the real-time API and the export API). This allowed us to label the dataset according to whether the message was deleted (generally by a channel moderator, as platform-level moderation on Telegram is very limited).

These labels can be used to study channel-level moderation on Telegram and understand what types of messages are being deleted by moderators of different types of channels. As we show in Figure 5, moderation varies across channels. This line of research is especially important given the recent arrest of the Telegram CEO in France over the lack of platform-level moderation (Sauer 2024).

These labels can also be used to detect other forms of malicious content present on Telegram. Messages deleted by channel moderators are, by definition, undesired content according to the channel owners. This may be because they are part of a propaganda network, as we explored prior (Kireev et al. 2024), but they might also have been deleted because they contain spam, hate speech, or any other undesired content for that channel. Our dual collection methodology uniquely allows for this type of analysis, and demonstrates where the standard method of collecting telegram data via exporting the channel history falls short.

Propaganda Labels

As we explored in prior work (Kireev et al. 2024), this dataset can be used to detect and remove propaganda from Telegram. We implemented a classifier to detect future propaganda based on the needs and abilities of channel owners. However, this space is open to much more analysis. The propaganda detection algorithm can be further refined, the labels can be used to train models to detect propaganda on other platforms or channels, and these labels used to detect other forms of propaganda. Additionally, the content of the propaganda messages themselves can be studied. For example, the narratives or stances of the messages in different networks can be extracted or the different strategies for spreading propaganda can be studied. As Figure 2 shows, there are stark differences between the two propaganda networks. It is worth noting that our snowballing method for detecting propaganda was strictly better than human moderation, so the moderation label alone is insufficient for studying propaganda.

Telegram Message Content

Finally, this dataset can also be used as a general purpose, Russian-language dataset that has been cleaned of many propaganda messages. For example, this test can be used for model training and fine-tuning Russian-language text models.

More specific to this particular dataset, however, would be topic or stance classification around the Russo-Ukrainian conflict or Russian political topics more generally. As we explored above and is shown in Figure 4, these data can be used to track different topics (e.g., debates over immigration). The channels are labelled according to stance, which can further be used to understand how users with different viewpoints are discussing a topic. The data were collected between 2021 and 2023, largely from popular political channels, curated by the authors to be diverse along a number of different vectors (topic, size, stance, etc.). Although this dataset is far from the largest of its kind, the hand selection of channels is advantageous and the collection time period is interesting.

Conclusion

In this work we present a novel dataset of Telegram messages labelled by whether they are part of one of two propaganda networks and whether the message has been deleted (usually by moderators). We further present some analysis of of this dataset to both explain the content of the dataset and explore potential additional uses.

Our dataset serves to provide other researchers with a tool to study Telegram as a social network, specifically in the context of the Russo-Ukrainian conflict. We provide labels for propaganda messages and moderation activity to further aid researchers in uncovering patterns of malicious behaviours.

References

Advox. 2023. How military supporters are using Telegram channels to suppress dissent in Myanmar. *Global Voices Advox*.

Baumgartner, J.; Zannettou, S.; Squire, M.; and Blackburn, J. 2020. The pushshift telegram dataset. In *Proceedings of the international AAAI conference on web and social media*, volume 14, 840–847.

BBC. 2023. Explainer: Which Hamas-linked Telegram channels are still accessible? *BBC Monitoring*.

Bergengruen, V. 2022. How Telegram Became the Digital Battlefield in the Russia-Ukraine War. *Time*.

Bovet, A.; and Makse, H. A. 2019. Influence of fake news in Twitter during the 2016 US presidential election. *Nature communications*, 10(1): 7.

Chen, E.; and Ferrara, E. 2023. Tweets in time of conflict: A public dataset tracking the twitter discourse on the war between Ukraine and Russia. In *Proceedings of the International AAAI Conference on Web and social media*, volume 17, 1006–1013.

Deb, A.; Luceri, L.; Badaway, A.; and Ferrara, E. 2019. Perils and challenges of social media and election manipulation analysis: The 2018 us midterms. In *Companion proceedings of the 2019 world wide web conference*, 237–247.

Elmas, T.; Overdorf, R.; Özkalay, A. F.; and Aberer, K. 2021. Ephemeral astroturfing attacks: The case of fake twitter trends. In *2021 IEEE European symposium on security and privacy (EuroS&P)*, 403–422. IEEE.

Ester, M.; Kriegel, H.-P.; Sander, J.; Xu, X.; et al. 1996. A density-based algorithm for discovering clusters in large spatial databases with noise. In *kdd*, volume 96, 226–231.

Ferrara, E.; Varol, O.; Menczer, F.; and Flammini, A. 2016. Detection of promoted social media campaigns. In *Proceedings of the international aaai conference on web and social media*, volume 10, 563–566.

Gebru, T.; Morgenstern, J.; Vecchione, B.; Vaughan, J. W.; Wallach, H.; Iii, H. D.; and Crawford, K. 2021. Datasheets for datasets. *Communications of the ACM*, 64(12): 86–92.

Guhl, J.; and Davey, J. 2020. A safe space to hate: White supremacist mobilisation on Telegram. *Institute for Strategic Dialogue*, 26.

Hanley, H. W.; and Durumeric, Z. 2024. Partial mobilization: Tracking multilingual information flows amongst russian media outlets and telegram. In *Proceedings of the International AAAI Conference on Web and Social Media*, volume 18, 528–541.

Hosseinmardi, H.; Mattson, S. A.; Ibn Rafiq, R.; Han, R.; Lv, Q.; and Mishra, S. 2015. Analyzing labeled cyberbullying incidents on the instagram social network. In *Social Informatics: 7th International Conference, SocInfo 2015, Beijing, China, December 9-12, 2015, Proceedings 7*, 49–66. Springer.

Khattar, D.; Goud, J. S.; Gupta, M.; and Varma, V. 2019. Mvae: Multimodal variational autoencoder for fake news detection. In *The world wide web conference*, 2915–2921.

Kireev, K.; Mykhno, Y.; Troncoso, C.; and Overdorf, R. 2024. Characterizing and Detecting Propaganda-Spreading Accounts on Telegram. *arXiv preprint arXiv:2406.08084*.

Kloo, I.; Cruickshank, I. J.; and Carley, K. M. 2024. A Cross-Platform Topic Analysis of the Nazi Narrative on Twitter and Telegram During the 2022 Russian Invasion of Ukraine. In *Proceedings of the International AAAI Conference on Web and Social Media*, volume 18, 839–850.

Reimers, N.; and Gurevych, I. 2019. Sentence-BERT: Sentence Embeddings using Siamese BERT-Networks. In *Proceedings of the 2019 Conference on Empirical Methods in Natural Language Processing*. Association for Computational Linguistics.

Robles, J.-M.; Guevara, J.-A.; Casas-Mas, B.; and Gómez, D. 2022. When negativity is the fuel. Bots and Political Polarization in the COVID-19 debate. *Comunicar*, 30(71): 63–75.

Sauer, P. 2024. Telegram CEO charged in France for 'allowing criminal activity' on messaging app. *The Guardian*.

SberDevices. 2022. BERT large model (uncased) for Sentence Embeddings in Russian language.

Scrivens, R.; and Amarasingam, A. 2020. Haters gonna "like": Exploring Canadian far-right extremism on Facebook. *Digital extremisms: Readings in violence, radicalisation and extremism in the online space*, 63–89.

Sharma, K.; Zhang, Y.; and Liu, Y. 2022. COVID-19 vaccine misinformation campaigns and social media narratives. In *Proceedings of the International AAAI Conference on Web and social media*, volume 16, 920–931.

Shevtsov, A.; Tzagkarakis, C.; Antonakaki, D.; and Ioannidis, S. 2022a. Identification of twitter bots based on an explainable machine learning framework: the US 2020 elections case study. In *Proceedings of the international AAAI conference on web and social media*, volume 16, 956–967.

Shevtsov, A.; Tzagkarakis, C.; Antonakaki, D.; Pratikakis, P.; and Ioannidis, S. 2022b. Twitter dataset on the Russo-Ukrainian war. *arXiv preprint arXiv:2204.08530*.

Solopova, V.; Popescu, O.-I.; Benzmüller, C.; and Landgraf, T. 2023. Automated multilingual detection of pro-kremlin propaganda in newspapers and telegram posts. *Datenbank-Spektrum*, 23(1): 5–14.

Stahl, P. M. 2024. Lingua: The most accurate natural language detection library for Python, suitable for short text and mixed-language text.

Venâncio, O. R.; Ferreira, C. H.; Almeida, J. M.; and da Silva, A. P. C. 2024. Unraveling User Coordination on Telegram: A Comprehensive Analysis of Political Mobilization during the 2022 Brazilian Presidential Election. In *Proceedings of the International AAAI Conference on Web and Social Media*, volume 18, 1545–1556.

Vox-Harbour. 2024. https://telegra.ph/Kremleboty-v-Telegrame-Bolshoe-issledovanie-ot-Vox-Harbor-01-26.

Wijermars, M.; and Lokot, T. 2022. Is Telegram a "harbinger of freedom"? The performance, practices, and perception of platforms as political actors in authoritarian states. *Post-Soviet Affairs*, 38(1-2): 125–145.

Wilkinson, M. D.; Dumontier, M.; Aalbersberg, I. J.; Appleton, G.; Axton, M.; Baak, A.; Blomberg, N.; Boiten, J.-W.; da Silva Santos, L. B.; Bourne, P. E.; et al. 2016. The FAIR Guiding Principles for scientific data management and stewardship. *Scientific data*, 3(1): 1–9.

Willaert, T.; Peeters, S.; Seijbel, J.; and Van Raemdonck, N. 2022. Disinformation networks: A quali-quantitative investigation of antagonistic Dutch-speaking Telegram channels. *First Monday*.

Zannettou, S.; Caulfield, T.; De Cristofaro, E.; Sirivianos, M.; Stringhini, G.; and Blackburn, J. 2019. Disinformation warfare: Understanding state-sponsored trolls on Twitter and their influence on the web. In *Companion proceedings of the 2019 world wide web conference*, 218–226.

Paper Checklist

1. For most authors...

   (a) Would answering this research question advance science without violating social contracts, such as violating privacy norms, perpetuating unfair profiling, exacerbating the socio-economic divide, or implying disrespect to societies or cultures? Yes

   (b) Do your main claims in the abstract and introduction accurately reflect the paper's contributions and scope? Yes

   (c) Do you clarify how the proposed methodological approach is appropriate for the claims made? Yes

   (d) Do you clarify what are possible artifacts in the data used, given population-specific distributions? Yes

   (e) Did you describe the limitations of your work? Yes, see Limitations

   (f) Did you discuss any potential negative societal impacts of your work? Yes, see Ethics Statement

   (g) Did you discuss any potential misuse of your work? No

   (h) Did you describe steps taken to prevent or mitigate potential negative outcomes of the research, such as data and model documentation, data anonymization, responsible release, access control, and the reproducibility of findings? Yes, see Ethics Statement

   (i) Have you read the ethics review guidelines and ensured that your paper conforms to them? Yes

2. Additionally, if you are using existing assets (e.g., code, data, models) or curating/releasing new assets, **without compromising anonymity**...

   (a) If your work uses existing assets, did you cite the creators? Yes

   (b) Did you mention the license of the assets? Yes

   (c) Did you include any new assets in the supplemental material or as a URL? Yes, as url

   (d) Did you discuss whether and how consent was obtained from people whose data you're using/curating? Yes, in the Ethics Statement

   (e) Did you discuss whether the data you are using/curating contains personally identifiable information or offensive content? No

   (f) If you are curating or releasing new datasets, did you discuss how you intend to make your datasets FAIR (see Wilkinson et al. (2016))? Yes

   (g) If you are curating or releasing new datasets, did you create a Datasheet for the Dataset (see Gebru et al. (2021))? No

Ethics Statement

All channels and groups used in the study are public now and were public at the moment of data collection and all the data used in this dataset was publicly available at the moment of data collection. The data was collected using the official Telegram API and according to Telegram's terms of service. The researchers did not use their personal accounts for this study. The published dataset contains the bare minimum of data required for the analysis described in the paper. Even though this information is public, we excluded usernames and substituted user IDs with random numbers to guard the telegram users from unnecessary attention. We do not see any major negative consequences from publishing this dataset. On the contrary, we believe that publishing this dataset will benefit society by enabling new research on propaganda in telegram.

Our institutional IRB has approved the research project and dataset publication.