2022-04-14-2022-07-06 2022-07-18 2022-08-04 2022-08-22 2023-06-12
2022-02-24-2022-07-06 2022-08-02 2022-09-02 2023-06-12
2022-02-25-2022-07-06 2022-07-19 2022-08-22 2023-03-21 2023-06-12
2022-02-24-2022-07-07 2022-11-01 2023-03-21 2023-06-12 2023-11-04
2022-03-01-2022-07-07 2022-10-26 2023-03-21 2023-06-12 2023-06-15
2022-02-22-2022-07-07 2022-11-01 2022-12-29 2023-06-12 2023-11-04
2022-02-24-2022-09-05 2022-10-31 2023-03-21 2023-06-12 2023-10-10
2022-02-22-2022-09-06 2022-11-01 2023-03-22 2023-06-12 2023-11-03
2022-02-25-2022-10-14 2022-11-03 2023-11-21 2023-12-27
2022-11-12-2022-11-19 2023-06-12 2023-11-03
2022-08-12-2022-12-12 2023-07-01 2023-08-15 2023-10-10 2023-11-03 2023-12-27
2022-01-12-2023-02-16 2023-06-12 2023-09-21 2023-11-03
2022-02-24-2022-12-22 2023-06-10 2023-06-15 2023-12-27
2022-06-18-2023-04-04 2023-06-12 2023-06-15 2023-11-04
2022-06-03-2023-06-11 2023-06-15 2023-11-03 2023-12-27
2022-11-13-2023-06-11 2023-06-15 2023-11-04
2022-05-27-2023-06-11 2023-06-15 2023-10-10 2023-11-03
2022-02-24-2023-01-24 2023-03-06 2023-06-12
2022-04-02-2022-04-03 2023-02-07 2023-06-12 2023-10-10
2022-02-16-2023-03-25
2022-02-24-2023-04-17 2023-06-12
2022-03-13-2023-05-03 2023-05-11 2023-06-12 2023-11-03
2022-02-24-2023-05-17 2023-06-08 2023-06-12 2023-12-27
2022-07-06 2023-01-06
2022-07-19 2022-08-18
2022-06-08-2023-06-10 2023-06-15 2023-10-10 2023-11-04
2023-06-09-2023-06-12 2023-06-15
2022-10-21 2023-03-21
2022-10-21 2023-03-21
2022-10-21 2023-03-21
2022-02-24-2022-07-06 2022-07-20 2022-07-28
2022-02-24-2022-07-06
2023-06-12
2022-06-26-2022-07-07 2022-11-01 2023-03-21 2023-06-12
2022-07-19 2022-07-28 2022-08-04

TELEGRAM ARCHIVE OF THE WAR

2022-02-28_
2024-02-28_

TWO-YEAR PROJECT RESULTS

CENTER FOR URBAN HISTORY

Telegram Archive of the War

An initiative to preserve and organize data from the Telegram social network regarding the Russo-Ukrainian war
Two-year results of the project
Date of publication: February 28, 2024

This report summarizes the results of the archival team's work created at the Center for Urban History as part of the initiative to document Russia's full-scale invasion of Ukraine. The document describes the context of the archive, the methodology of data collection, the experience of the archival team, the challenges faced, and the decisions made. It explains the structure of the collected materials and outlines the existing and potential ways of working with the archival data. Additionally, the report provides information on archival materials collected outside the Telegram platform as part of this archival initiative (websites, radio recordings, etc.).

Contents

1. Introduction

The Telegram Archive of the War, as an emergency archiving initiative, arose from realizing the importance of preserving historical evidence of the full-scale invasion and the risk of losing these materials, given the fragile nature of social media platforms as a historical source.

Telegram became an integral part of the public space in Ukrainian society on the eve of and particularly with the onset of the Russian full-scale invasion. In general, the active use of Telegram and social media is evidenced by sociological surveys. Telegram is the most popular source of information among Ukrainians, and its usage increased from 65.7% to 71.3% during the year of the full-scale invasion. (Opora, 2022; Opora, 2023)

The platform has emerged as a pivotal space for communicating various aspects of military reality. Telegram serves as a source of information on security threats, including official and civilian reports of shelling, as well as monitoring the movement of Russian aircraft launching missiles into Ukrainian territory. Telegram chats have become vital for mutual aid and "horizontal" communication among individuals during the war. There are numerous chats where users exchange experiences like navigating through occupiers' checkpoints and "filtering" procedures, seeking and offering transportation for evacuation from occupied territories, and sharing information regarding car queues, transport infrastructure, and traffic safety.

Various volunteer groups use the platform to coordinate cyberattacks on official Kremlin websites. These groups frequently share data leaks of Russian officials and the military. Their activities on Telegram are routinely either entirely or partially deleted. Some traces of their activities likely no longer exist beyond this Archive collection. Ukrainian authorities have also developed a chatbot on Telegram to gather data on the movements of the Russian military, suspicious individuals, and attempted sabotage. There have even been specific official requests to report certain types of military equipment, prompting users to send data when they recognize something similar in their vicinity. However, during wartime, there are restrictions on posting information about the Ukrainian military's positions and the armed forces' movements. It is also prohibited to publish details about the outcomes of Russian missile attacks before official reports are released to prevent the enemy from adjusting their strikes further.

Another integral aspect of the unfolding war on Telegram is the development of networks of channels and chats for disinformation and propaganda campaigns.

(Yuskiv et al., 2022; Ghasiya and Sasahara, 2023; Lazaruk, 2022; Maschmeyer et al., 2023; Mikhailidi, 2023; Plakhta, 2020)

Some communities undergo significant changes due to catastrophic events. For instance, the administrators and participants of a chat focusing on advertisements for buying/selling various goods in the city of Oleshky (Kherson region) strictly published posts connected to the economic themes during the first year of the full-scale invasion. However, the Kakhovka hydroelectric power plant explosion compelled the community to deviate from their established mode of communication: economic messages disappeared, and many comments related to the city's flooding emerged.

The surge in Telegram's popularity as a messaging and social media platform was propelled by its swift communication capabilities and a comparatively more lenient content moderation policy in contrast to other platforms. Consequently, users enjoy plentiful opportunities to publish content while maintaining their anonymity. The Telegram administration asserts its refusal to divulge users' personal data to third parties. However, in 2022, there were reports of the first instances of personal data being transferred to the German police, which the platform administration did not officially comment on. (Der Spiegel, 2022)

However, alongside the popularity of this platform, discussions regarding the risks of its use persist in the Ukrainian context. Despite the privacy and security assurances declared by Pavel Durov, the platform's founder, existing studies raise concerns about its commercial model and highlight its lack of transparency (Maréchal, 2018). While the link between the platform and Russian official structures remains unproven, some studies suggest that Russian intelligence services may have access to confidential user data (Loucaides, 2023).

Today, social media data is increasingly recognized as an important historical source, albeit one with inherent fragility. Social media platforms are primarily designed for rapid instant messaging or dissemination of the most recent updates rather than for long-term preservation as a historical source. This issue becomes particularly apparent and critical during war and extreme situations, where their *significance* as a source and *susceptibility* to deletion only escalate.

During wartime, posts, messages, entire chats, and channels are continually being created and deleted: channels and chats frequently cease updating, become inactive, or are deleted by administrators. Additionally, some chats may disable the export function, rendering further archiving impossible. Moreover, users can also modify or delete their comments in chats. Therefore, when archiving Telegram, we

strive to collect publicly available content, mindful of its timely preservation from deletion, considering the rapidly evolving nature of digital sources and communication practices dependent on users and platforms.

2. Archive Structure

Briefly about the structure of the Archive

- from a few chats discussing DDoS attacks and channels disseminating Russian propaganda in early March 2022, the Archive collection has expanded to 2711 items (1048 channels, 1651 chats, and 12 chatbots) with a total data volume of 27.8 terabytes (as of February 9, 2024).
- archived channels and chats are categorized into 37 thematic collections (for further details, see Collections).
- most messages in the channels/chats were archived from February 24, 2022, to June 12, 2023, marking the completion of the primary archiving stage. Nonetheless, there are instances where chats/channels were uploaded before the invasion or partially uploaded after the primary archiving stage. Consequently, the overall timeframe for messages in the Archive spans from December 2016 to December 2023.
- the initial data download occurred on February 28, 2022, with the final one on December 27, 2023. The decision to conclude the archiving process is solely based on the internal capabilities of the Archive and is not linked to any specific event (for further details, see Chronology and Geography).
- to enhance the usability of the archive, metadata for archival units and collections were gathered from channels/chats and created by archivists (for further details, see Register of Archival Units and Collection register).
- archivists endeavoured to preserve various types of files that could be manually downloaded using the Telegram export function. The Archive encompasses files such as photos, videos, video messages, audio messages, GIFs, stickers, etc (for further details, see File Formats).

Context and background

This initiative began with no prior experience or training in web or social media archiving. Our colleagues at the Center for Urban History, who work on archiving more traditional audiovisual historical sources, introduced us to the basic archiving principles. Thus, we studied the fundamental methodology and gathered advice directly while creating the archive.

The initiative emerged in late February 2022 during internal discussions at the Center on responding to a full-scale invasion from professional and civic perspectives. Previous experience documenting the Revolution of Dignity (2014) and the coronavirus pandemic (2021) prompted a similar initiative to document the Russian invasion. Alongside the collection of oral testimonies, visual materials, and personal accounts such as dreams and diaries, an initiative to archive web and social media content was launched.

With no clear understanding of what and how to archive, project coordinator Taras Nazaruk contacted the international community of web archivists for advice on archiving methodology and other similar initiatives to document the Russian war in Ukraine. Despite various ongoing projects, it became evident that Telegram channels and chats were not being archived. At that time, the fastest and most affordable method to download the data was through the desktop Telegram interface, enabling the saving of channel and chat histories onto local disks. Within the first few weeks, additional archivists joined the effort. Eventually, a permanent group of four people was formed. They continue to work on the archive.

Media researchers, journalists, and historians, some of whom originate from the frontline or temporarily occupied territories, are involved in the archive's activities. While working on the archive, some archivists acquired specific personal experiences related to the war: mobilization, residing in the temporarily occupied territories, relocating to other regions of Ukraine, fleeing abroad as refugees, or living in frontline settlements. This is one reason why, in the initial phase of the archive's establishment, data concerning it, including information about the team members, were not disclosed; this collective decision was made for the personal safety of the archivists. Clearly, as we are directly involved in the events and may undergo stress and a sense of danger, we cannot fully keep a certain level of detachment from what is being archived. On the one hand, it complicates the ability to reflect on decisions regarding the selection of materials for archiving and introduces biases when describing the collected data. Due to the absence of prior experience in archiving social media, the methodology for collecting materials may contain flaws and inconsistencies, some of which may necessitate later revisions. However, being directly involved in the events constitutes an epistemological standpoint enabling us to draw upon personal experiences and professional expertise to better monitor the information landscape, respond to fluctuations, and prioritize themes. Thus, archiving the war represents another facet of experiencing the war. In the course of their work, the archivists produced several report notes detailing their current archiving experiences, which were taken into consideration when writing this report.

Project team

Taras Nazaruk serves as the coordinator of the Archive, is the head of digital history projects at the Center for Urban History, and is pursuing a doctoral degree at FernUniversität in Hagen.

Oksana Avramenko, a historian and graphic designer, initially contributed to the Archive by curating a selection of channels and chats from the temporarily occupied Crimea. Over time, her focus shifted towards socially vulnerable groups such as internally displaced persons (IDPs), refugees, and human rights protection during the war.

Anna Rubanska, a journalist, specializes in Ukrainian refugees abroad, media, marketing, and channels and chats of military units as tools for information warfare.

Dmytro Serbulov, a historian, commenced his involvement with the Archive by assembling a collection of chats and channels from southern Ukraine. Later, he broadened his interests to encompass other topics such as religion, ecology, and animals.

Participants of the initial archiving phase

Roksoliana Holovata is a historian and the editor of the Interactive Lviv project at the Center for Urban History. She is also a PhD student at the Ukrainian Catholic University and the University of Wroclaw. She has been involved in archiving some of the Telegram channels and websites.

Oleksandr Kud is a journalist, writer, and musician. He began working with the Archive in March 2022 but later joined the Armed Forces of Ukraine. He contributed to the formation of collections of chats from the Kharkiv region and propaganda channels targeting the audience in Southern Ukraine.

Karina Lazaruk is a graphic designer and researcher. She joined the Archive in March 2022 and has focused on archiving Russian propaganda and disinformation. Her involvement in the Archive led to collaboration with scholars from UvA and INC, the results of which were cited in *Politico* and published in *Theory on Demand #44*.

Serhiy Pakhomenko is a historian and an associate professor at the Department of International Relations and Foreign Policy at Mariupol State University. During his fellowship at the Center for Urban History from May to September 2022, he archived channels and chats related to Mariupol.

Key milestones

The work on the Archive is structured into four stages: collection, processing of collected data, development of access protocols and infrastructure, and research and interpretation. The primary task of the Archive team was to collect data, systematize it, and establish conditions for effective search. Simultaneously, when creating descriptions for archival units, archivists composed descriptive summaries and, drawing from their interpretation and experience, focused on potentially significant topics and communication practices. Additionally, we aim to initiate discussions with researchers from various disciplines regarding the analysis and interpretation of the material gathered in this archive. Moving forward, we are interested in hosting collaborative events where we can present the archive's creation context and contents and discuss research directions across a range of social and humanitarian disciplines and artistic practices.

Given the transient nature of this source type, data collection has been the project's primary focus for an extended period.

The main phase of data collection lasted from February 28, 2022, to June 12, 2023. However, the archived data spans a much wider timeframe (see Chronology and Geography for more details). The conclusion of the archiving process was not linked to any specific event. Instead, the archiving deadline was determined solely by the internal capabilities of the Archive and the necessity to process the collected data. In the latter half of 2023, archivists concentrated on structuring the amassed data, finalizing descriptions and metadata, and establishing sensitivity levels and access protocols. Nevertheless, throughout 2023, archivists continued to address the current situation and selectively uploaded data for individual items on a case-by-case basis.

The scope of archiving

The scope of the archiving was not clearly defined from the outset and was shaped more intuitively based on the archivists' experience, the research focus of the Center for Urban History, and the snowball effect that occurred as the project expanded. In the initial stages, the Archive team even curated several collections with materials that did not originate from Telegram (these materials are stored separately from the main Telegram archive; see Appendix 1).

The primary task of archiving Telegram initially focused on preserving only a few dozen channels and chats that each archivist could monitor based on their individual needs and interests, as well as the threat of deletion. It soon became evident that there was a vast amount of such material, resulting in a continual increase in the number of archival units for storage.

The following criteria are taken into consideration when deciding to include a channel or chat in the archive:

- relevance to wartime events
- everyday and local practices not represented in official resources and media
- regional and social diversity
- risk of disappearance
- platform practices that have become integral to the war
- availability of publicly accessible content

This list of criteria was not established initially but rather summarizes the decisions made retrospectively within the context of emergency archiving during the initial weeks of the Russian invasion.

The collected materials encompass various wartime practices and pertain to various social groups and political perspectives.

Chronology and geography

Chronologically, most of the archival data pertains to activity on Telegram following the start of the full-scale invasion on February 24, 2022. However, some channels were archived for a specific period before or even years prior to 02/24/2022, starting from their creation time, if deemed necessary by the archivists to document their developmental trajectories. This allows us to assess how the full-scale war has influenced the worldview of Ukrainian society, the perception of certain phenomena, and audience demands based on the nature of the content before and after the invasion. In certain cases, segments of time from months preceding the aggression were archived to document the proliferation of hybrid influences of Russian propaganda or responses to the escalation of Russia's military presence near Ukraine's borders. For instance, the Center for Countering Disinformation channel commenced its activity on 11/23/2021, coinciding with the initiation of "military exercises" between Russia and Belarus. Similarly, communities of cyber volunteers started leaking data about Russian officials and the military in December 2021. As for city and regional chats, the initial archiving date was set a few days before the onset of the war to observe the shift in communication practices. Many chats were established to facilitate horizontal emergency communication during the war and have been archived from their creation date after 02/24/2022. Examples include chats for mutual assistance in evacuation and networking among internally displaced persons (IDPs).

Given the limited technical storage capacity and the necessity to prioritize high-quality processing of the uploaded data and its preparation for access, we decided

during the working meeting on 06/12/2023 to conclude the main archiving stage. In certain instances, it was decided to monitor and upload new time periods ad hoc until the end of 2023. As a result, most archival units encompass the period from February 24, 2022, to June 12, 2023. However, some archival units include messages dating as early as December 2016 and as late as December 2023.

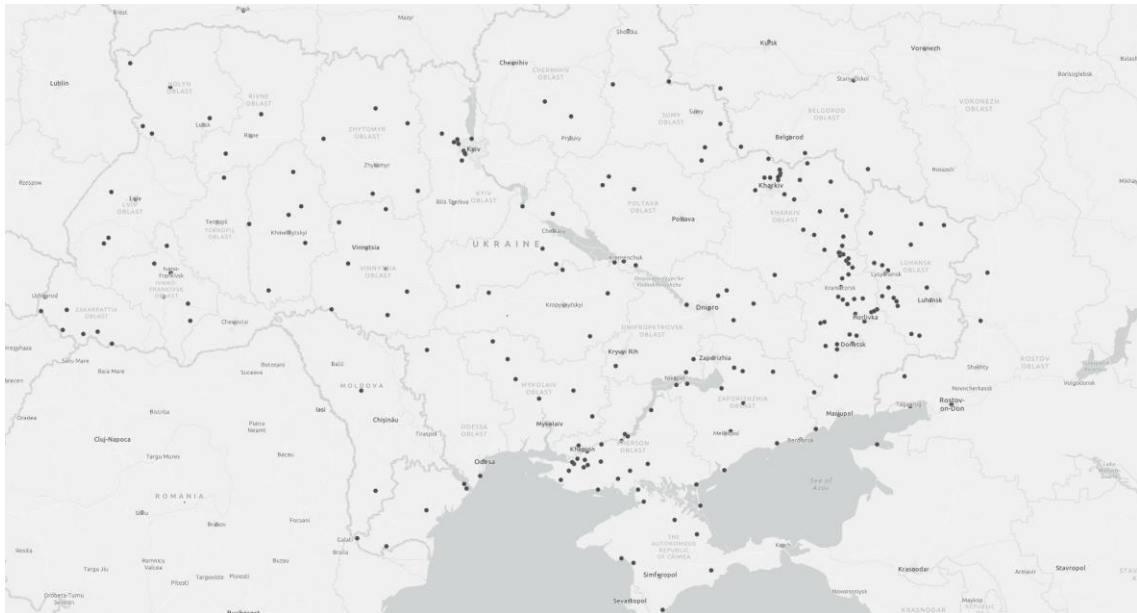The geographical localization of channels and chats is confined to the territory of Ukraine within its state borders.

Figure **1**. Map of channels and chats in Ukrainian cities, based on mentions of the city in their names or original descriptions.

In determining the geographical scope of channels and chats, archivists relied on direct references to specific localities or regions in the title or original description of the channel or chat.

However, with the expansion of the thematic focus, particularly concerning Ukrainian refugees worldwide and in Russia, the geographical factor has become less significant. Currently, the Archive contains channels and chats related to 36 countries.

Figure **2**. Map of countries with archived chats of Ukrainian refugees and migrants.

Archive terminology

Telegram has two primary tools for publicly disseminating information: channels and chats (the term "chats" is utilized in the Archive as it is more commonly recognized among users compared to the internal Telegram terminology, which employs "groups"). Each archived channel or chat constitutes an archival unit, encompassing a collection of gathered and generated metadata (for further details, see Metadata), as well as the content of the channel/chat (for further details, see Collections and File Formats). For more insights into the structure of the Telegram platform, consult our FAQ channel.

Channel

The emergence of social media has greatly changed the media landscape in many countries worldwide. Ukraine is no exception here. Since the beginning of the invasion, Telegram channels have become one of the most important sources of information for Ukrainians. This tool allows one person or group to broadcast information to a wide audience unilaterally. In Telegram channels, information is often presented concisely and consists of posts with a large number of original

photos and videos. Ordinary users often create and send this content to the channel administration. Of course, different Telegram channels have different topics and expertise. However, those that publish news (especially local news channels) often have a certain "horizontal" setup when some posts are created by the channel's administration and others are sent by readers. For instance, the Telegram channel *Trukha* (укр. "Труха") promptly disseminates photos and videos depicting the aftermath of Russian missile strikes. These materials, created by ordinary users, appear on the channel in a matter of hours (or even minutes!) after the event. Such efficiency is possible thanks to an extensive network of users who actively send photos and videos to the channel's administration. Earlier, the *Trukha* Telegram channel even used to offer a small financial reward for high-quality photo/video material. Along with user-generated content, the Telegram channel's administration publishes its own comments, links to conventional media, other social networks, etc. Conventional media cannot compete with Telegram channels in terms of the speed of information delivery and often refer to Telegram channels as the main sources of their publications. Many traditional media outlets are forced to switch to creating long-form materials, interviews, investigations, and other types of journalistic content in which promptness is not a determining factor for competitive advantage. In other words, they have to transform and develop new niches to compete with Telegram channels in the information field. One can consider channels more moderated sources compared to chats, given that a select few individuals make decisions regarding the publication of information.

Chat

Unlike channels, chats have a limited audience. A chat room can have up to 200,000 subscribers. Usually, people share information, discuss, and exchange videos, audio, and photos in chats. Chat is a horizontal communication tool that allows users to communicate on an "all to all" basis when everyone sees the message of anyone in the chat and can start a conversation with them. Unlike channels that represent the point of view of one person or a small group of people, chats contain posts from a wide range of users. Although such a large volume of messages is quite difficult to navigate, it is a priority for the Archive to upload chats as an important source of horizontal communication between people during the war.

The Archive team uses its own typology of chats: regular chats, organized chats, and inner chats. Despite the fact that these chats look almost identical in the Archive, they have different natures and properties able to affect the data structure within the chat.

Regular chat

A constant flow of information. A place where everyone communicates with each

other. Identifying any specific topic that users discuss in regular chats is often difficult.

Organized chats

As mentioned above, regular chats represent a continuous stream of chaotic messages that are difficult to navigate. Therefore, Telegram decided to fix this problem and introduced a new feature for chats with more than 100+ subscribers. Now, chat owners can create thematic subgroups within chats. There are separate thematic discussions in the left panel of the chat. For example, people discuss movies, health, and other topics. This should help organize information in chats. However, only a few chats use this feature.
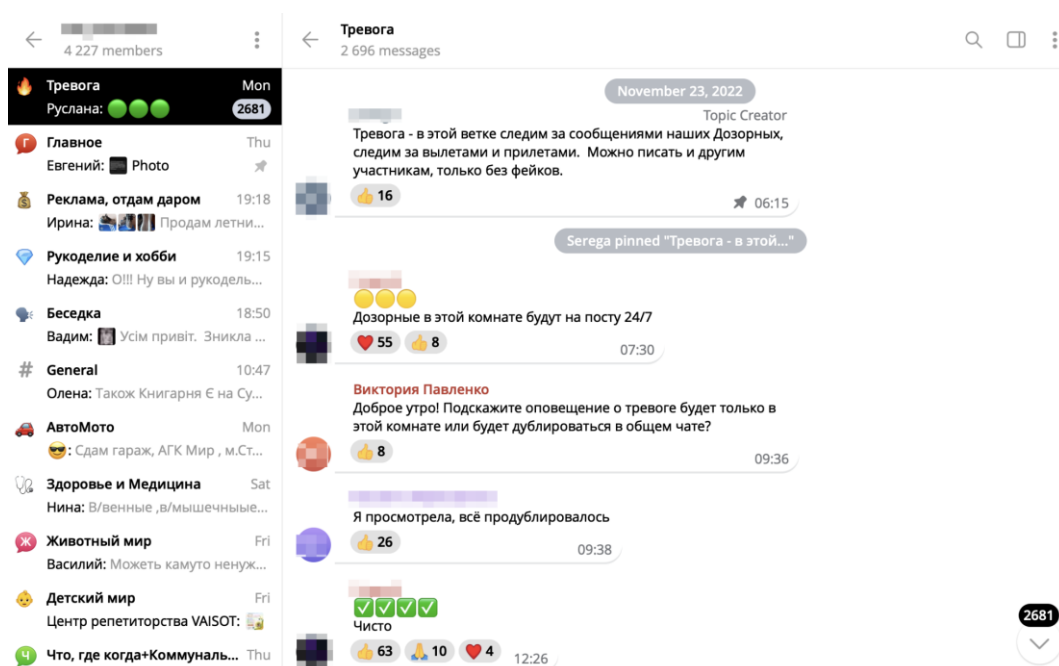
Figure **3**: Chat is divided into thematic branches (organized chat) in the Telegram interface: on the left is a list of thematic subgroups within the chat.

"Inner" chat

This one is yet another chat subtype. Archivists call this one an "inner" chat in internal communications. To enable comments under the posts, channel administrators should create a new chat or attach an existing chat to a channel. This way, a comment box will appear below each post on the channel. Users can click on this comment box and start a discussion below the post. However, if you click the button (three vertical dots) in the upper right corner of the chat and select View discussion, the chat attached to the channel opens.

Figure **4**. "View discussion": the context menu of the desktop version of Telegram to go to the "embedded" chat.

An "inner" chat is a kind of repository for comments and posts from a channel. If there is no chat, the channel has no comment function. In this internal chat, all the channel posts and comments are displayed as a continuous stream of information. From the outside, it looks like a regular chat. But, unlike a regular chat, an internal chat necessarily contains two components: a post from the channel to which it is linked and comments under these posts. "Inner" chats are the only way to download channel posts and comments. Otherwise, the comments are lost if one downloads only the channel and not its "inner" chat. Therefore, "inner" chats (if available), not just channels, have become a priority for downloading. Internal chats reflect both the point of view of channel owners and show how people perceive posts published by channel owners.
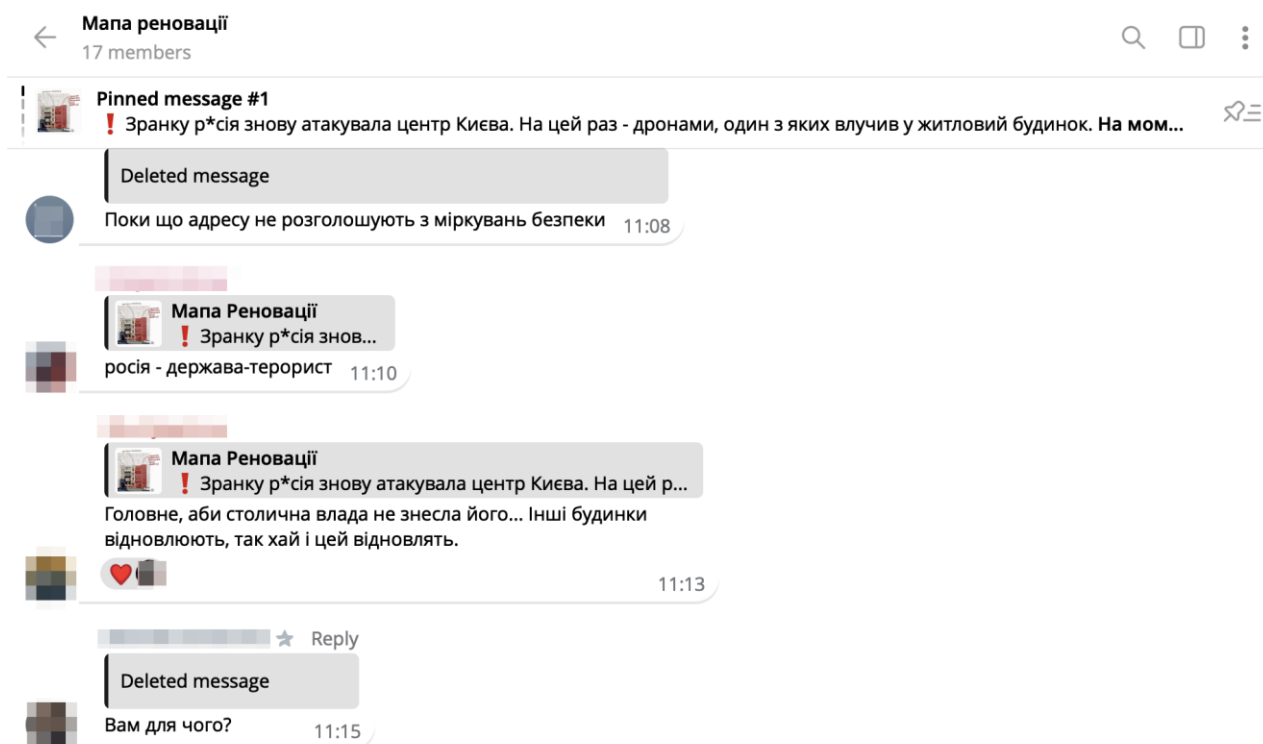
Figure **5**. The channel's subscribers comment on its posts in the "inner" chat.

Chatbots

Chatbots are the smallest part of the Archive's collection. They were used by authorities, human rights organizations, and volunteer initiatives in an attempt to structure and optimize communication with their audiences, especially during the intense communication phase of the first weeks of the invasion. Government agencies created civilian security guides with chatbots. The police collected information about missing persons and identified the dead. For the military, chatbots served as sources of crowdsourcing intelligence. For cyber volunteers, they provided samples of content for viral mailings. Human rights activists provided answers to the most frequently asked questions of legal aid recipients during the war and collected evidence of war crimes. Chatbots have become a tool for sending requests for evacuation assistance for residents of the temporarily occupied territories.
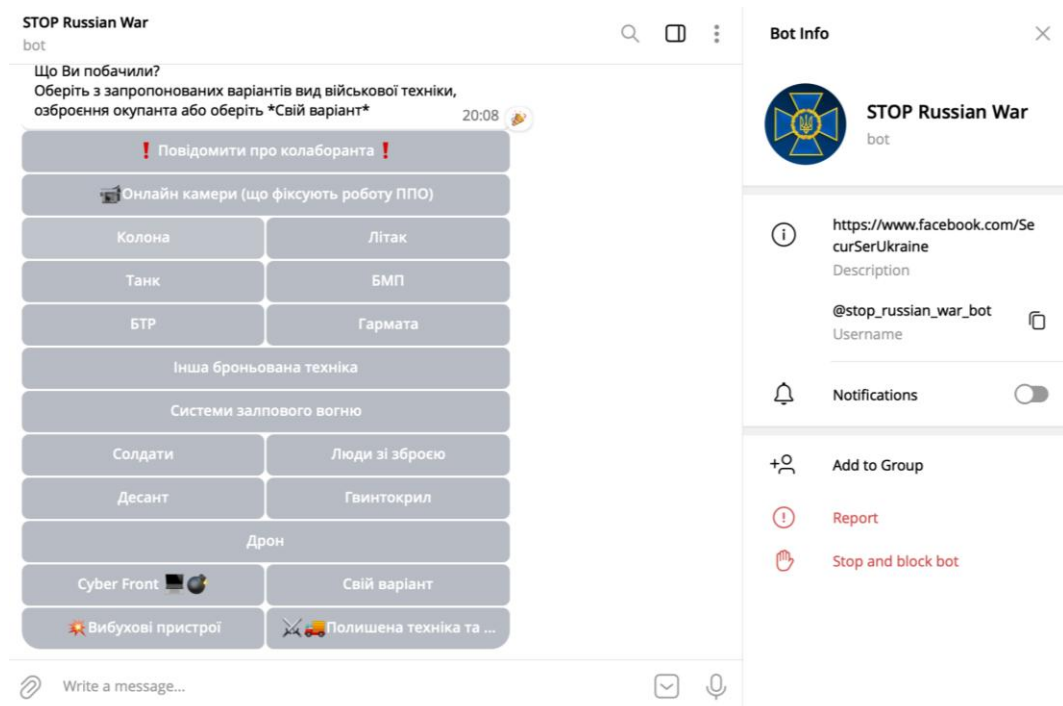
Figure **6**. The official SBU chatbot for collecting information on the location, types of Russian military equipment, and activities of collaborators.

One can see the internal chatbot's structure only when interacting with it. This way, we managed to record their internal structure, but all the information provided by the archivist in communication with the chatbot is fictitious.
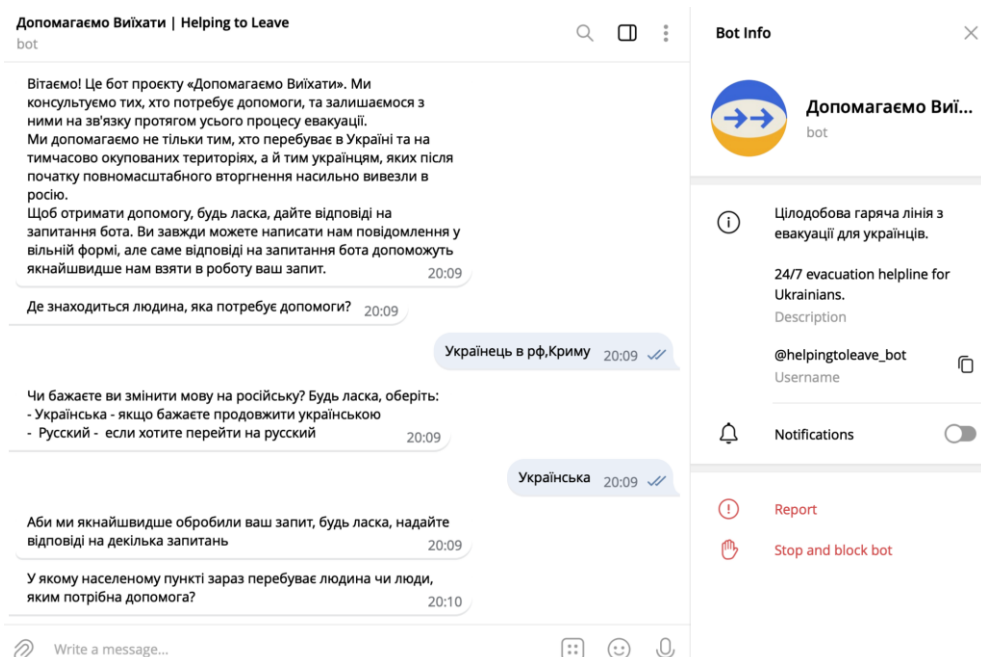

Figure **7**. An example of an archivist communicating with a chatbot.

Metadata

The collected materials of the Archive are structured with the metadata entered in the register.

Register of archival units

Each channel or chat constitutes a separate archival unit included in the register. An archival unit in the register primarily includes metadata put together with the source. It contains original title and description, link, type of unit (channel, chat, chatbot), creation date, as well as secondary metadata created by archivists: ID, affiliation with a thematic collection, number of subscribers, messaging frequency, archival description and its author, audience tags, keywords, sensitive content note, technical comments.

ID ⌄	≡ Title ⌄	⌐ URL ⌄	⊙ Ch... ⌄	⊙ Thematic Collection ⌄	⊙ Frequ...
2	Віталій Кім / Миколаївська ОДА	https://t.me/mykolaivskaODA	Channel	Urban and Local	1-10/day
5	ТРУХА⚡Харьков + Украина🇺🇦	https://t.me/truexanewsua	Channel	Urban and Local	1/few days
35	Хуевый Харьков	https://t.me/hueviykharkov	Channel	Urban and Local	1/few days
36	Харьков LIVE	https://t.me/+8Rah-lb7Q6Y2...	Channel	Urban and Local	1/few days
41	Дмитро Орлов - Енергодарсь...	https://t.me/orlovdmytroEn	Channel	Urban and Local	1-10/day
43	Запорізька обласна військова ...	https://t.me/zoda_gov_ua	Channel	Urban and Local	1-10/day
45	Хуевая Одесса	https://t.me/xydessa	Channel	Urban and Local	>10/day
46	ХД 🇺🇦 ДНЕПР	https://t.me/hyevuy_dnepr	Channel	Urban and Local	>10/day
72	СВОИ	Кривой Рог	https://t.me/svoiKR	Channel	Urban and Local	1-10/day
73	Типичный Николаев	https://t.me/tpnik	Channel	Urban and Local	1-10/day
75	Анатолій Куртєв. Секретар За...	https://t.me/kurtievofficial	Channel	Urban and Local	1-10/day
86	Экстренные Новости Белгород	https://t.me/breakingnewsBel...	Channel	Russian Bordering Regions	>10/day

Figure **8**. Metadata of the register of archival units in the table.

Primary metadata

Title. The original channel or chat name recorded at the time of the description creation. It may have changed over time, as noted by the archivist in the Technical Comment.

The original description provides a brief overview of the purpose of the channel or chat room and the topics that chat owners would like to discuss there. Quite often,

the description does not define the possible topic trajectories of the chat. In some cases, the original description is missing.
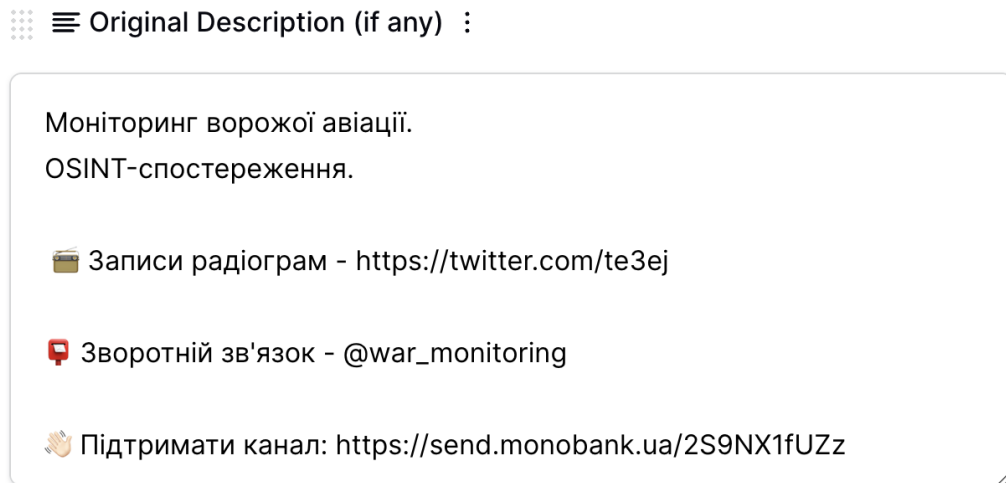

Figure **9**. The original description of one of the archived channels.

The URL of a channel or chat on the platform. Its presence or absence, as well as the name, depends on the actions of administrators. Sometimes, it is available only via an invitation link due to privacy settings. There is no link in cases where the chat is "inner" and exists only as comments to the main channel or if the chat is set to secret. The archivist can only access such a chat if they find an invitation link in another chat or if another person adds the archivist. This emphasizes the importance of assigning an ID because, in such cases, it helps identify the channel or chat in the archive.

Type (for further details, see Archive terminology).

The "Existed before the war" mark and the Creation date tag make it possible to ascertain whether a channel or chat room existed before the war or was established in response to it. The oldest item in the Archive is dated September 2015 (a news channel in the temporarily occupied Crimea), and the newest one is dated August 2023 (a chat of activists to resist mobilization in Russia). The Creation date does not always indicate the chronological framework of the stored materials; rather, it solely denotes the creation of the archival unit within Telegram.
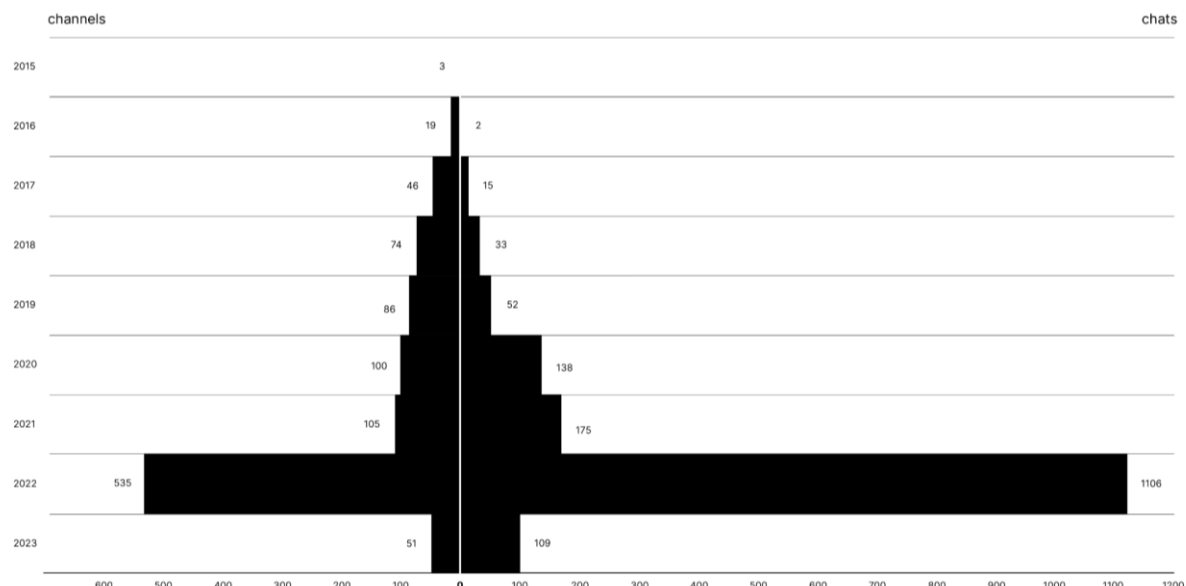
Figure **10**. Statistics on the number of archived channels and chats by the year of their creation.

Secondary metadata

ID. Each channel and chat is assigned a unique identifier as an archival unit. This identifier is created to simplify navigation and storage of archival data.

Amount of subscribers is a quantitative indicator of the audience at the time of the description creation. The Archive records quite different numbers—from a few or a few hundred people to tens and hundreds of thousands.

⊙ Channel/Chat ⌄	☰ # of followers/members(as of certain date)
Chat	16 (27.12.2022)
Channel	236 022 (03.06.2022)
Chat	1 811 (02.06.2022)
Chat	11 448 (02.06.2022)
Chat	19 326 (02.06.2022)

Figure **11**. Number of subscribers of different archival unit types.

At the same time, the actual number of participants in an "inner" chat often does not correspond to the number of active contributors in it. This is because "inner" chats are essentially repositories where a Telegram channel stores its posts and the

comments users leave under these posts. Not all users are aware of the existence of "inner" chats or, for some reason, do not want to join and use them for communication. Therefore, in some "inner" chats, we can observe many comments automatically transferred from the channel, along with a few subscribers, because users do not use the chat for communication or join it (for further details about the "inner" chat, see Archive Terminology).

Message frequency (>10/day, 1-10/day, 1/every few days, 1-10/every few days, 1/week, 1/every few weeks). This is an indicator of the contributors' activity in the chat room and the frequency of publishing new posts in the channel, determined by the archivist through a quick assessment while creating the description.

⊙ Thematic Collection ⌄	⊙ Channel/Chat ⌄	⊙ Frequency (>10/day, ...
Urban and Local	Channel	>10/day; 1-10/day
Urban and Local	Chat	>10/day
Russian Propaganda and...	Chat	>10/day
Russian Propaganda and...	Chat	>10/day
Volunteer Chats	Channel	1/few days
Volunteer Chats	Channel	1/week
Economy	Chat	1-10/day
Volunteer Chats	Chat	>10/day

Figure **12**. Frequency of notifications depending on the type and subject of the archival unit.

Most chats with active discussions or channels with prompt updates have >10 messages/day. 1-10 messages/day are the updates on the activities of human rights and religious organizations. 1/few days—small communities—personal and art blogs. 1-10/several days, 1/week—volunteer initiatives that use Telegram mainly to report on their activities. Channels often reduce or stop their updates altogether, and activity in chats decreases.

Archival description represents the archivist's endeavor to outline the prevalent topics and discussions, origin, and type of content inherent in the channel or chat room at the time of description. Fully examining all chat topics isn't always feasible due to the volume of messages. To create the inventory, the archivists reviewed each channel and chat at certain intervals to track changes in topics and frequency

of messages. Depending on the chat/channel topic, messages for the description were selected with an interval of one week or one month before the date of the archival description. Our team made efforts to avoid value judgments; however, the archival description may focus on certain emphases and topics, given the professional or civic engagement of the archivists. It reflects the professional bias of the archivist, i.e., the registry contains descriptions created from the perspective of media researchers and historians.
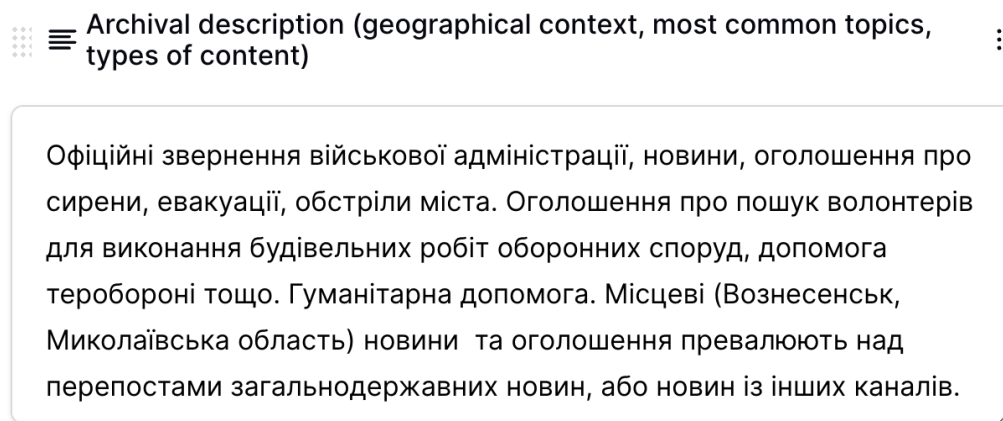
⋮⋮ ≡ Archival description (geographical context, most common topics, types of content) ⋮

> Офіційні звернення військової адміністрації, новини, оголошення про сирени, евакуації, обстріли міста. Оголошення про пошук волонтерів для виконання будівельних робіт оборонних споруд, допомога теробороні тощо. Гуманітарна допомога. Місцеві (Вознесенськ, Миколаївська область) новини та оголошення превалюють над перепостами загальнодержавних новин, або новин із інших каналів.

Figure **13**. An archival description created by an archivist for a local media outlet in the frontline areas.

Audience tag defines the geographic location and target groups of a channel or chat. The definition of target groups takes into account the purpose of the chat/channel and the topic of the chat/channel. The audience description is based on the way administrators describe their audience—whom they address and with what messages.
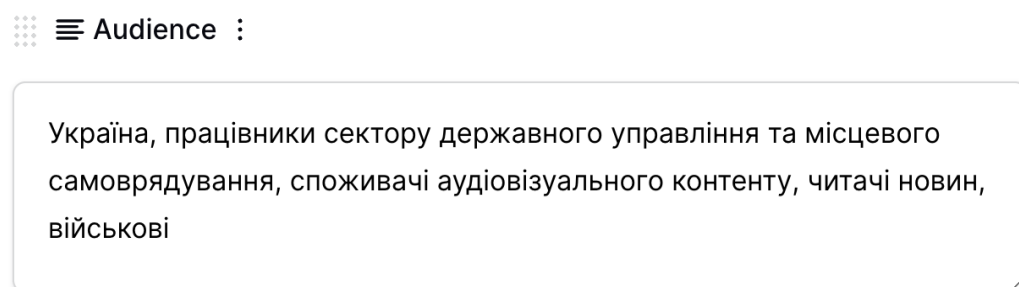
⋮⋮ ≡ Audience ⋮

> Україна, працівники сектору державного управління та місцевого самоврядування, споживачі аудіовізуального контенту, читачі новин, військові

Figure **14**. Audience tags assigned to the channel of the head of the Oblast State Administration.

The archivists identified 235 audience tags based on professional, ethno-religious, gender, social vulnerability (LGBTIQ, people with disabilities, IDPs, refugees), interests, political preferences, and socio-demographic characteristics. This set of tags is an informed assumption by archivists regarding the users to which a particular channel/chat is relevant. Together with keywords, audience tags simplify the surface navigation of archival units.

Keywords (1751) were created as generalized categories and concepts that reflect everyday life, the new military reality, Russian propaganda clichés, and authentic common vocabulary of chat room users to denote groups of people or phenomena that do not constitute a value judgment of archivists ("Khokhly," "Zhduny," "Katsaps", "Orcs," "Churki", "Ukrainian Nazis," "Ukrops," "good Russians," "bavovna" (literally "cotton," a Ukrainian word and Internet meme that originated during the Russian invasion of Ukraine, humorously used to refer to explosions). They are designed to serve as thematic clues during a user search and play a supplementary role in navigating the materials.
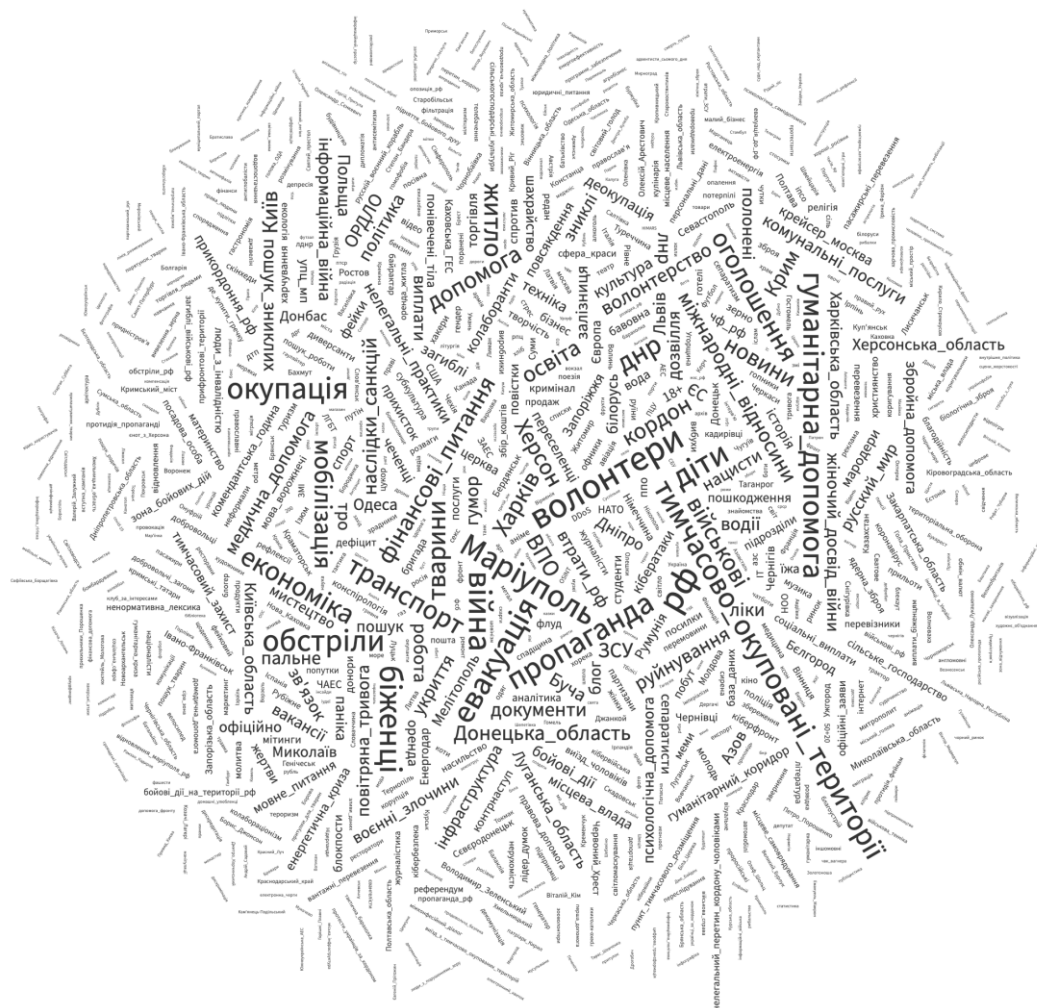
Figure **15:** Keyword cloud. One can investigate an interactive visualization here: https://perma.cc/3BLK-HUK6

Sensitive content refers to the presence of images of the dead, human remains, mutilated bodies, and sexually explicit images. The large-scale Russian invasion has led to significant changes in the Ukrainian population's perception of sensitive content: in many chats or channels, you can find photos of dead Russian soldiers,

videos of bombs dropped on the enemy, etc. The presence of violent content on social media is increasingly commonplace.

Thematic Collection	Channel/Chat	Sensitive Content
Databases and Indexes	Chat	yes
Humour	Chat	yes
Russian Propaganda and...	Channel	yes
IT Community/ Informatio...	Chat	yes
Military / Political Blogs	Channel	yes

Figure **16**. A mark in the register indicates the presence of sensitive content in the materials of a unit.

Even when discussing general resources, such as humor or meme-based platforms, images depicting deceased military personnel and sexualized images become a ubiquitous component of wartime humor, ridicule, or intimidation of the enemy. Channels and chat rooms that declare "traditional values" as one of the key ideas they fight for are characterized by a large number of homophobic images and scenes of violence. In addition, the Telegram administration's mechanism of labeling and moderating sensitive materials is practically absent or works only in some cases. As a result, archive materials often contain images of sensitive nature. At the same time, archivists try to label channels/chats where such content is published regularly.

Technical commentary contains information primarily for internal use, aiming to clarify the archival description and provide important additional details about the integrity of the archival unit. However, it also holds value for researchers.

☰ Technical commmentary
з 21.05.2022 архівується через вшитий чат (ID 1063). архівування припинено 12.06.2023
Заархівовано від 16.02.2022. Заархівовано без відео. архівування припинено 12.06.2023
Архівується з 09.08.2022
З 21.07.2022 архівується вшитий чат
архівування припинено 08.06.2023

Figure **17**: Technical comments on the termination of unit archiving.

This is where the archivists record:

- Changes to the original title, URL links

- For archived "inner" chats, information about the main channel (name, URL link, ID (if it was uploaded separately from the main channel), original description, number of subscribers, and creation date) has been added
- A comment if the data was exported without video or with a certain file size limit.
- A comment if it was appropriate to archive the data from a certain date until February 24, 2022.
- A comment in case the data export is no longer relevant. A mark about the date of termination of archiving.
- A comment if the data of a certain unit is partially (for a certain period) or completely lost and can't be restored.

Collections

The archived channels and chats are grouped into 37 thematic collections. Each collection has an introductory description explaining the principles of assigning materials to it (preference for certain topics, geographical origin) and possible research paths, marking the predominant types of content, topics of discussion, the presence of specific vocabulary, and the sporadic presence of topics from this collection in other collections.
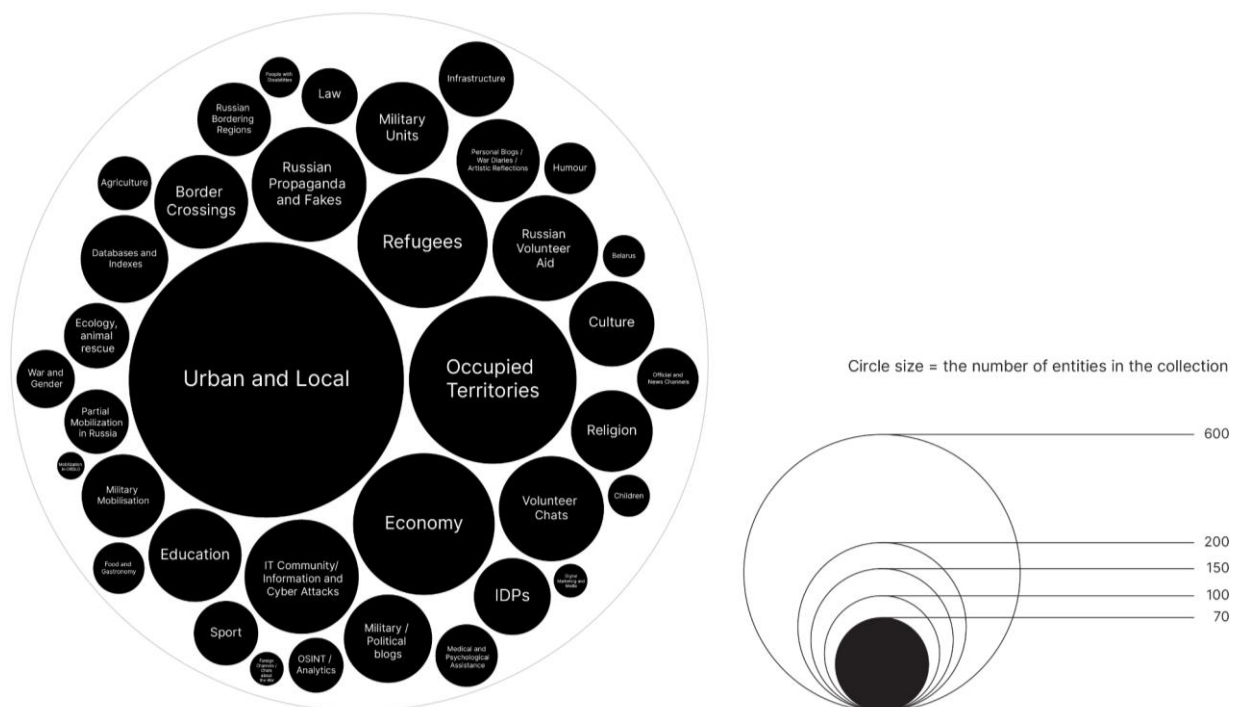
Figure **18**. Hierarchy of thematic collections by the number of archival units.

Official and News Channels (32 archival units: 22 channels, 7 chats, 3 chatbots)

This collection includes official channels of communication of state bodies used to disseminate official narratives about the military-political situation, state strategic communication, and messages on civilian security issues. In addition, mass media channels that record the coverage of the war in news updates have been preserved. In this section, the geography of sources is broad, encompassing media and news sources within a region or oblast. These can be located in the corresponding section of the collection titled "City and Regional Channels and Chats."

Foreign Channels/Chats about the War (12 archival units: 6 channels, 6 chats)

This collection includes channels and chats created to spread pro-Russian messages and narratives outside of Ukraine, as well as pro-Ukrainian channels and chats that translate news from official nation-wide sources and counter fakes. The chats in the collection allow us to trace how foreign citizens perceive the war in Ukraine and what content they discuss and share. Most of the channels and chats are dedicated to discussing Russian aggression against Ukraine, politics in Ukraine, and other news about Ukrainian-Russian relations.

Urban and Local (633 archival items: 150 channels, 483 chats)

The collection covers local news channels and city/oblast/regional chats, organized based on Ukraine's administrative division.

The "panic of the first days of the war" is noticeable in the chats. Users promptly share information about shelling and discuss the situation in their locality or district, the work of pharmacies, shops, and cash withdrawals. Depending on the proximity to the frontline, these topics gradually disappear, and users return to topics of everyday communication. For instance, participants in rear-area chat rooms engage in discussions about news and volunteer initiatives and become more active, mainly during air raids. Instead, in the occupied territories, the state of communication services, the work of utilities, the availability of essential goods, medicines, and the actions of the occupation authorities and the Russian military in the territory of the locality remain the most pressing topics for a long time.

IDPs (57 archival units: 29 channels, 28 chats)

The collection contains chats of the frontline or temporarily occupied territories residents forced to move to other regions of Ukraine. In a number of cities, modular towns have been built. IDP communities arise in and around such towns in the physical dimension, and chats play a role in digital networking environments. In addition to the main topics of discussion, such as how and where to get financial and

humanitarian aid, IDPs are returning to everyday issues, making friends, planning leisure activities, exchanging advice, and trying to organize their daily lives.

Border Crossings (72 archival units: 2 channels, 70 chats)

This collection focuses on migration and mobility during the war, including refugee experiences at the EU or Russian borders. Chats from the frontline and temporarily occupied territories reveal ongoing uncertainty regarding filtering procedures, transit, and forced mobilization.

Chat participants seek and provide evacuation vehicles and exchange information regarding traffic queues, transport infrastructure integrity, and traffic safety.

Refugees (144 archival units: 15 channels, 129 chats)

This collection brings together Ukrainians who fled the war to different countries worldwide. These are primarily countries bordering Ukraine, except for Russia and Belarus. Media reflecting the lives of Ukrainians who voluntarily moved or were forcibly displaced to Russia are included in the Russian Volunteer Groups collection. This collection includes channels created by governmental or volunteer organizations in the host countries that coordinate assistance to Ukrainians and provide useful information. The participants of the chats in this collection discuss and compare the standard of living in different EU countries and share their experience of receiving temporary protection, financial and humanitarian aid, housing, enrolling children in educational institutions, receiving medical care, and solving financial issues. Refugees create networks and plan joint leisure activities. They often discuss the conditions of a short-term visit to Ukraine. Ukrainians' attempts to earn money by offering diverse services to their compatriots are widespread, with dedicated channels and chats to promote small businesses.

Volunteer Chats (99 archival units: 38 channels, 60 chats, 1 chatbot)

The chats and channels in this collection are dedicated to volunteer activities related to solving humanitarian crisis issues and meeting the army's needs. These are chats for searching for and helping each other in evacuation, providing humanitarian aid to residents of the war zone and frontline areas, and providing shelter and assistance to IDPs. There are also channels for volunteer initiatives to coordinate and share volunteer experiences.

A separate collection comprises Russian Volunteer Aid (89 archival units: 20 channels, 69 chats), including chats dedicated to locating missing relatives and the collection of humanitarian aid by volunteers from Russia and the territories occupied since 2014 for the Russian military, residents of the occupied territories, and refugees to Russia.

Medical and Psychological Assistance (35 archival units: 13 channels, 21 chat rooms, 1 chatbot)

This collection contains chats providing psychological assistance to victims of hostilities. These include chats primarily dedicated to connecting victims with psychotherapists for further assistance in private communication; psychological mutual assistance communities; and channels with psychological self-help tips.
The collection also includes chats for volunteer doctors of various specialties. Such channels share advice on first aid and tactical combat casualty care.

Military Units(71 archival units: 46 channels, 25 chats)

This collection contains channels and chats associating with or representing Ukrainian military units and command. Some of the resources are official and operate in alignment with the stance of the Ukrainian leadership. Additionally, the collection encompasses unofficial chats of military formations, providing supplementary information on specific front areas, peculiarities of military life, lack of equipment, conflicts with the command, and blogs featuring reflections from combatants that may diverge from the official stance.

Military / Political Blogs (67 archival items: 40 channels, 27 chats)

Channels of the military, activists, volunteers, and journalists offer firsthand accounts of the war and provide original audiovisual content. They cover combat operations thematically, publish testimonies from combatants, personal reflections, and insights from the defense sector, and contain materials to boost morale and humor. Bloggers frequently interact with their audience through "inner" chats.

Military Mobilisation (55 archival items: 16 channels, 39 chats)

The collection includes chat networks created to collect and share information about public places where draft notices are served. Given the calls to resist mobilization, their simultaneous and networked creation, channels, and chats can serve as a means of hybrid warfare. This collection comprises channels and chats for people who avoid mobilization. They provide notifications about public places where notices are issued. There are also calls urging resistance against the existing mobilization order. Furthermore, there are channels operated by service providers facilitating illegal border crossings for men of military age.

The separate collection Military Mobilisation in ORDLO (6 archival units: 1 channel, 5 chats) includes chats about mobilization processes in the so-called "L/DPR" region, shedding light on its covert and compulsory nature under the de facto military dictatorship.

"Partial" Mobilisation in russia (36 archival units: 1 channel, 35 chats)

This collection contains chats dedicated to the mass migration of Russian citizens following the "partial" mobilization on September 21, 2022. Official reports indicate that 200,000 people left the country within the first week of mobilization. Most of the chats in the collection are geographically linked to specific checkpoints on the border with countries neighboring Russia, such as Kazakhstan, Georgia, and Finland.

IT Community/ Information and Cyber Attacks (107 archival items: 61 channels, 44 chats, 2 chatbots)

The channels in this collection cover the functioning and damage to digital infrastructure, information about mutual cyberattacks by Russia and Ukraine, and countermeasures to them.

The chats are used to identify targets, share instructions and technical solutions for organizing cyberattacks. These chats share data of Russian officials and classified information about government and military institutions.

The collection also contains channels of Russian hackers attacking Ukrainian digital infrastructure and online media.

OSINT/ Analytics (27 archival units: 20 channels, 7 chats)

The channels in this collection track the movement of military equipment, types of armed groups and weapons, and collect evidence of war crimes. They provide analytics on the pace, strategy, and tactics of hostilities, identify problems, and record losses of the Russian military, which is a valuable source of information for the Armed Forces of Ukraine. These publications are frequently used to debunk Russian propaganda fakes circulating in public discussions about the war.

Databases and Indexes (61 archival units: 39 channels, 21 chat rooms, 1 chatbot)

This collection contains channels with data on the losses of the Russian military. They are full of sensitive visual content and negatively colored vocabulary against the Russian army, which is a manifestation of a kind of collective dehumanization of the enemy, a public demonstration of enmity and hatred.

It contains databases of local collaborators in the temporarily occupied territories of southeastern Ukraine.

Other channels document evidence of war crimes and maintain registers of victims and missing civilians.

Infrastructure (48 archival units: 26 channels, 21 chat rooms, 1 chatbot)

Russia is constantly shelling residential buildings, critical and transportation infrastructure, and nuclear power facilities. The consequences are recorded, as well

as the prompt restoration of communication routes and telecommunications, both during hostilities and after de-occupation, as reported by the channels of state and local utility companies and Internet providers. Chats in urban communities discuss the principles of post-war reconstruction.

Personal Blogs / War Diaries / Artistic Reflections (62 archival items: 41 channels, 21 chats)

The collection includes blogs and diaries of the military, volunteers, activists, artists, residents of the occupied territories, and foreigners in Ukraine. The channels feature the author's reflections, personal stories and experiences, and artistic rethinking of the war.

People with Disabilities (13 archival units: 2 channels, 11 chats)

Here, one can find chat rooms for visually impaired people who communicate via voice messages. Such chats allow them to hear the direct speech of users and even assume their age category. Hearing-impaired persons communicate through sign language in video messages. Volunteers help children with disabilities in the frontline area. Mothers of children with disabilities make decisions regarding evacuation and discuss the healthcare systems abroad.

Belarus (14 archival units: 5 channels, 9 chats)

Channels and chat rooms in this category mainly cover opposition sentiments to Belarus' participation in the war. At the level of civil society, a guerrilla campaign was launched to sabotage the movement of Russian equipment, and volunteer military units were sent to help the Ukrainian military. Chat participants reflect on the regimes of Putin and Lukashenko and discuss the consequences of the war for the republic and ways to avoid mobilization.

Russian Propaganda and Fakes (120 archival units: 85 channels, 35 chats)

Channels of Russian state institutions and their representatives that produce main messages of war propaganda at the state level. Channels and chats of Russian media, propagandists, and military correspondents are used to spread narratives in both Russia and Ukraine. The rhetoric within these channels and chats often contains hate speech directed against Ukrainians, Ukraine, its politicians, and the Armed Forces.

Occupied Territories (238 archival units: 108 channels, 130 chats)

This collection includes channels and chats from the regions occupied since 2014 and the territories temporarily seized during the hostilities after February 24, 2022.

Their titles include the symbols *V* and *Z*, Russian flags, names of self-proclaimed republics in ROTU, and the pronoun "our" (similar to the slogan "Crimea is ours") along with the name of the occupied settlement. The discussions among participants in such chats unmistakably exhibit a pro-Russian inclination. Entire networks of regional channels have been created to spread fake news, to foster a media narrative suggesting anticipation of the arrival of the "Russian world." Fake channels and chats with pro-Russian discourse were created for certain settlements that were identified as the next military target and had not yet been occupied. Chats and channels from the temporarily occupied territories with a pro-Ukrainian tone or without pronounced political discussions are presented in the collection titled *City and Regional Channels and Chats*.

Russian Bordering Regions (49 archival units: 22 channels, 27 chats)

The collection includes specially created and original local chats used by Ukrainians to stir up panic among residents of Belgorod, Bryansk, and Kursk as part of the information war (fake messages about air raid alerts and urgent evacuation). In addition, participants in city chats from these regions share testimonies about flights of combat aircraft and discuss the strengthening of "anti-terrorist" security measures, queues at gas stations, and communication from local authorities.

Economy (167 archival units: 49 channels, 118 chats)

This collection brings together channels and chats that reflect shifts in commodity and labor markets, price dynamics, the functioning of industry, and other economic indicators during the war. In the occupied territories, there is a significant shortage of essential goods and medicines, which are in high demand among chat users. The increase in the number of displaced persons in rear regions has created a huge demand for real estate, resulting in elevated rental prices.

Special attention is paid to documenting the consequences of the fuel and energy crisis. Ukrainians are creating multiple chats to exchange information regarding fuel availability in cities. In November 2022, channels for selling electric generators and stoves appeared amid blackouts.

As part of the wartime economic reality, the collection includes channels and chats providing specific services for the military ("flea markets" for ammunition, making devices for the front). They adjust to the reality of the occupying country in the territories controlled by the Russian army (buying and selling scarce goods, currency exchange, withdrawing funds from Ukrainian cards, and commercial transportation across the front line). The anonymity provided contributes to creating chats with illegal economic practices, including assistance in avoiding mobilization and illegal border crossing by persons liable for military service.

Agriculture (23 archival units: 3 channels, 20 chats)

Recognizing the significance of the agricultural sector in Ukraine's economy, particularly in regions heavily impacted by the war, the archivists have organized the agrarian chats into a separate collection. In logistics chats, one can monitor the geography and routes of mass export (theft) of grain from the occupied territories. In farmers' chats, the discussion focuses on the price dynamics for agricultural products in the occupation and the emergence of new forced markets.

Food and Gastronomy (21 archival items: 5 channels, 16 chats)

This collection comprises chats from food blogs and channels with gastronomic news. Chat participants discuss products with a long shelf life, food storage, and ways to collect wild plants. Bloggers describe their experience of volunteering for the Ukrainian Armed Forces. A separate topic here is the wartime gastronomic business.

Digital Marketing and Media (9 archival units: 4 channels, 5 chats)

Media channels that create and disseminate messages appealing to the international community to support Ukraine represent this collection. Additionally, there are channels informing about trends and initiatives of the creative industries during the war. For example, marketing quickly captures the mood of the target audience by creating slogans on military and patriotic topics. They also adapt by transitioning to Ukrainian-language communications in regions where it's not legally mandated, but marketers see such a consumer request.

War and Gender (31 archival items: 16 channels, 15 chats)

This collection covers the gender experience of war. It includes channels of LGBTIQ persons in the Ukrainian Armed Forces. It also archives the female dimension of war represented by the military, paramedics, volunteers, activists, and mothers. Unfortunately, it also represents the victims of Russian sexual violence and human trafficking at the borders. Their stories and experiences are recorded in the channels of this collection.

Culture (59 archival units: 32 channels, 26 chats, 1 chatbot)

Cultural institutions use their channels to share information about the war with the international community. They organize charity art events, support artists and cultural workers affected by the war, and document destroyed cultural monuments. Cultural life continues where it is relatively safe and calm, as evidenced by the publications of museums, galleries, and film festivals. Research, education, and artistic pursuits continue. This collection also includes chats of music and literature lovers who discuss the war, among other topics.

Religion (55 archival units: 38 channels, 17 chats)

This collection features Christian, Jewish, and Muslim channels. From the first days of the invasion, all religious organizations have been aiding the military and the war victims, collecting humanitarian aid, evacuating people, and providing their buildings as shelter.

In church news chats, participants discuss the autocephaly of the Ukrainian Orthodox Church (Moscow Patriarchate) from Moscow, pressure exerted on the church by local residents, calls for separatism in priests' sermons, and inspections conducted by the Security Service of Ukraine.

Children (14 archival units: 8 channels, 6 chats)

In this collection, one can find channels with recommendations for parents on providing psychological support to their children during the war, educational materials, and announcements of events for children's leisure. Additionally, there are channels dedicated to supporting and helping families of internally displaced persons and refugees.

Education (69 archival units: 36 channels, 33 chats)

This collection features chats from student communities at various universities actively involved in volunteer initiatives to support the Ukrainian military. However, they mostly discuss the educational process and the new realities of war in a humorous way.

It also includes channels of educational institutions evacuated from the temporarily occupied territories to the rear regions. Another topic of discussion in the chats is the placement of refugee children in educational institutions within the country of temporary protection.

Humour (20 archival units: 16 channels, 4 chats)

The meme channels' administrators swiftly adapted their entertainment content to the wartime context. New comedic symbols appeared, with some gaining immense popularity. "Russian warship". Tractor drivers collecting Russian tanks as trophies. "Bavovna" (explosions) in Russian border towns. Humor forms a specific vocabulary for the enemy: Russian Federation—"Pigsty," "Mordor," Russians—"pigs," "orcs," "piggies." Local authorities' representatives created newsbreaks to spread memes along with their images. Users' enthusiasm for certain types of weapons provided by Western partners (Bayraktar, Javelin) is also depicted in a series of memes.

Sport (33 archival units: 8 channels, 25 chats)

This collection includes chats between amateurs and professionals of various sports, serving as a platform for communication between coaches and parents of junior athletes who are arranging their evacuation to Europe to continue training. Channels of sports federations and associations monitor sanctions against Russia at the international level.

The circumstances for training and competitions changed over time, but the desire to continue training persisted, whether in shelters, on courts, or grounds in rear regions. Daily communication has not disappeared either. Participants in the sports movement organized joint leisure activities, charity championships and tournaments, and fundraisers to support the Armed Forces of Ukraine.

Law (25 archival units: 16 channels, 7 chats, 2 chatbots)

This collection includes governmental, legal analysis, and human rights channels. They document war crimes, collect and publish stories of the de-occupied and occupied territories residents and IDPs, collect evidence of the forced displacement of Ukrainians to Russia, and help them return.

Victims of the war seek legal assistance in chats with lawyers providing free legal aid. They mostly ask questions about evacuation, crossing the border with the EU, social guarantees for IDPs and refugee status, and payments for damaged housing and property restoration. Members of the armed forces need legal assistance in resolving various disputes with their command, obtaining combatant status, registering disability, and receiving social guarantees. Relatives of killed, missing, and captured military also seek legal aid.

Ecology, Animal Rescue (38 archival items: 9 channels, 29 chats)

This collection includes channels and chats of environmental and animal rights activists. Environmental activists highlight the devastating impact of intense hostilities on the environment. Many pet owners made efforts to evacuate their pets from the war zone. However, after the Kharkiv and Kyiv regions' de-occupation and the Kakhovka Hydroelectric Power Plant explosion, numerous chats/channels appeared where animal rights activists were looking for abandoned pets' owners.

Collection register

A separate register contains all thematic collections. The register presents a detailed collection description, a list of channels and chats in each collection, an audience description, channel and chat languages, and keywords.

Collection description explains the principles of assigning materials to the collection (preference for certain topics, geographical origin) and possible research paths, marks the predominant types of content, topics of discussion, the presence of specific vocabulary, and the sporadic presence of topics from this collection in other collections.
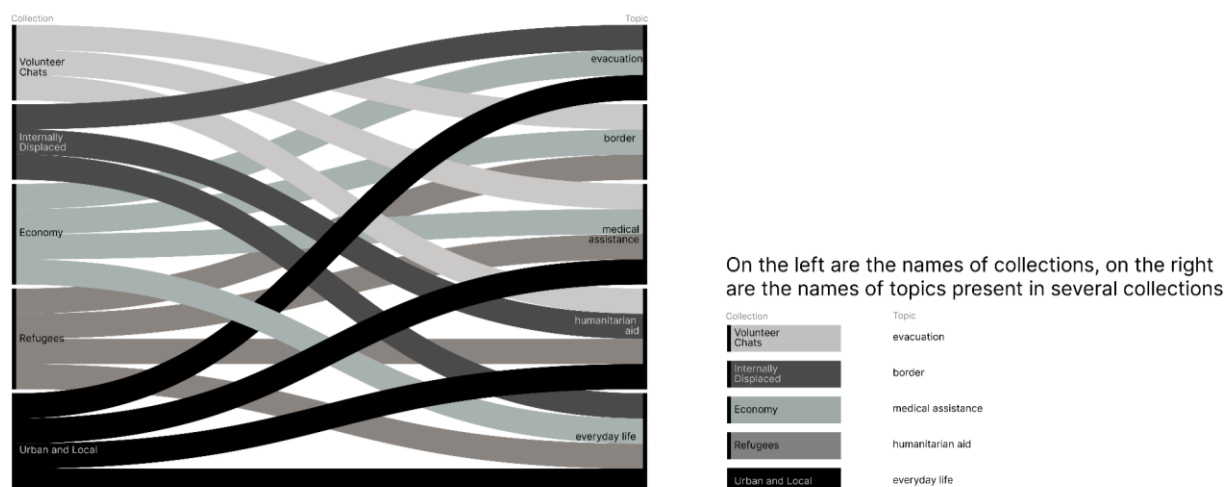
Figure **19**. An example of overlapping topics: topics related to medical, humanitarian aid, and border crossings appear in different collections at the same time.

List of channels and chats in each collection consists of the ID and name of the archival unit included in the corresponding collection.

Collection audience is a summary of information from the channel and chat register. It contains lists of audience types by geography and type (for further details, see Metadata).

Archivists determined the languages of the channels and chats in the collection based on the availability of content or communication in a particular language. For most collections with a geographical reference to the territory of Ukraine, in a bilingual society, Ukrainian and Russian were specified. For collections with cross-border interactions (Border, Volunteer groups, Refugees, "Partial" mobilization in Russia), the list of languages used is much wider and mostly corresponds to the country's official language with which the communication takes place regarding border crossing and residence. The collection "Foreign Channels/Chats about the War" contains materials in English, French, German, Spanish, Italian, and Bulgarian—based on the content of archived news channels or chats. The People with Disabilities collection contains messages in Ukrainian sign language.

The registers of archival units and collections will be available to users after registration (for further details, see Access).

From a few hacker chats and channels with Russian propaganda in early March 2022, the archive's collection has expanded to 2711 items (1048 channels, 1651 chats, and 12 chatbots) with a total data volume of 27.8 terabytes (as of February 9, 2024).
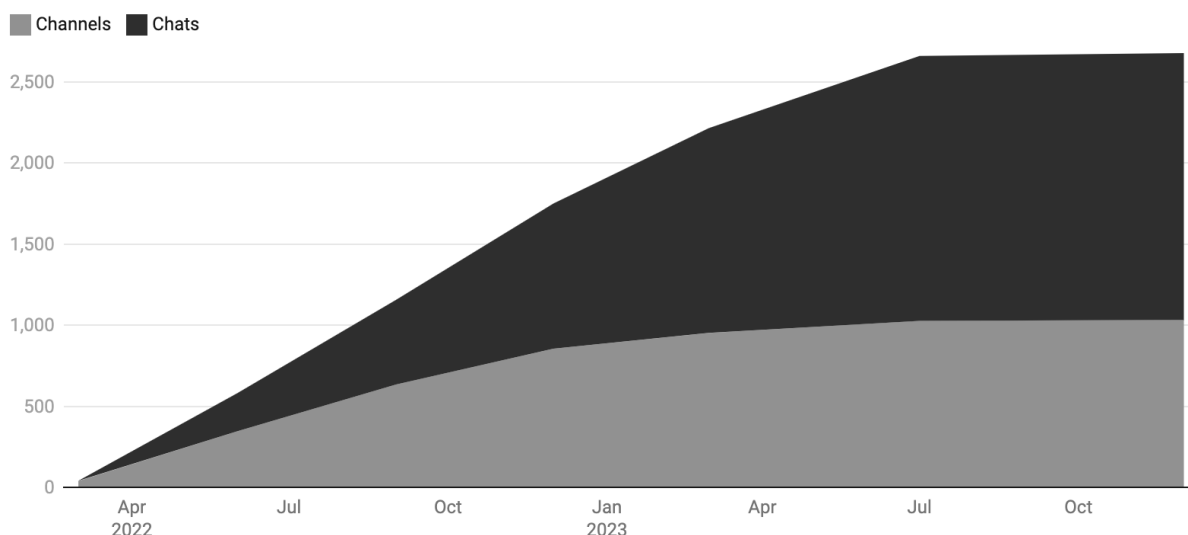
Figure **20**. The dynamics of archival materials' growth by downloaded channels and chats, interactive visualization: https://perma.cc/R22X-B7UE

Within the emergency archiving, in addition to the main archival collection of Telegram materials, archivists saved data off the platform, from various sources in a piecemeal and event-based manner. To systematize them, the archivists created a collection of data not related to Telegram and described separately (for further details, see Appendix 1).

File formats
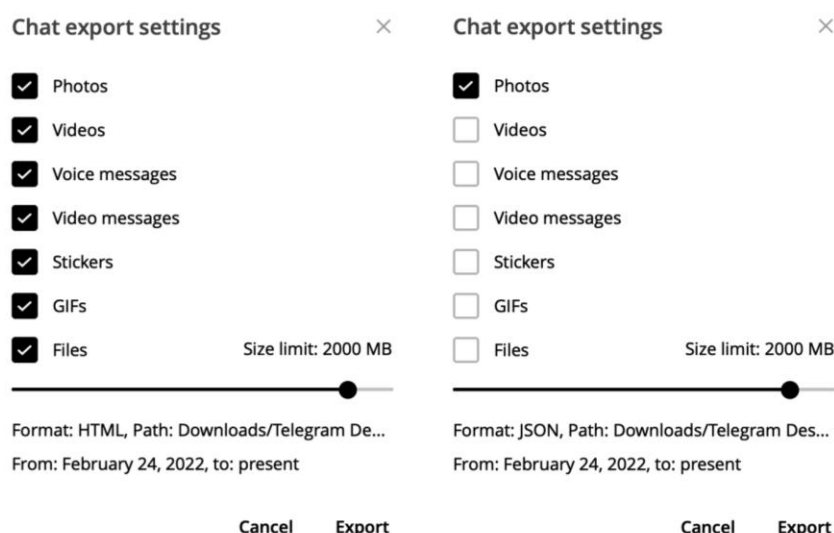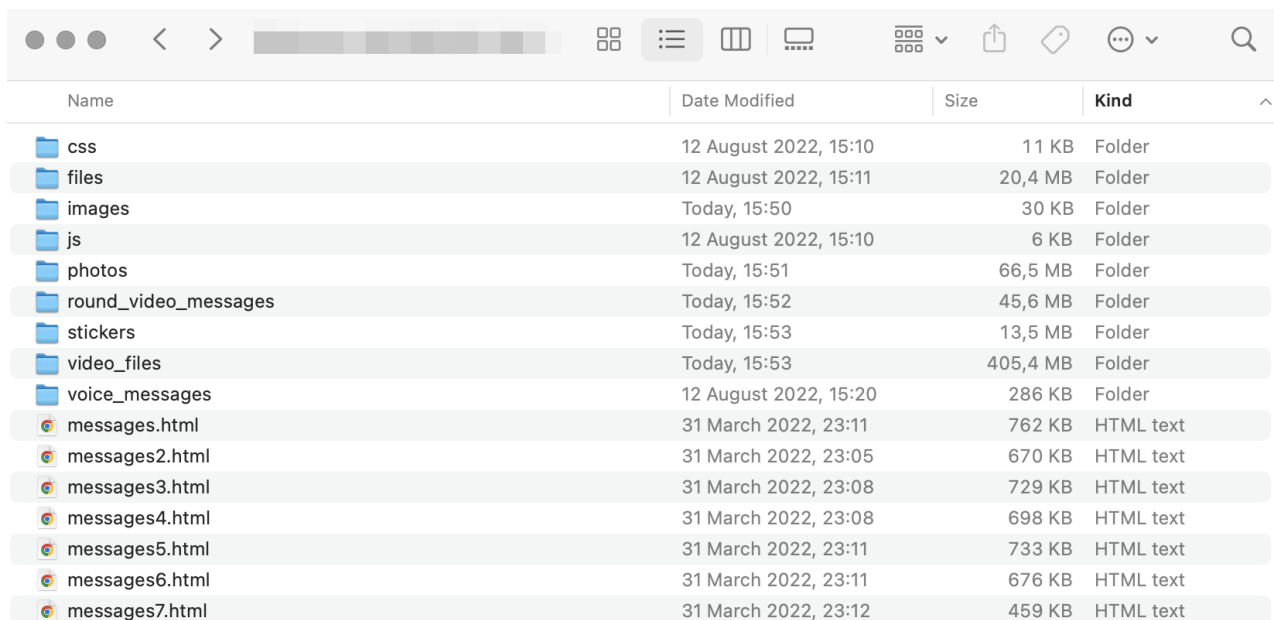
Figure **21**. Chat export settings when selecting the Export Chat History context menu. The export settings used when uploading HTML (left) and JSON (right) files to the Archive are illustrated.

The archiving function is available to all Telegram users by default in any channel and chat unless the administrator disables it. Throughout the archiving process, we uploaded data manually, as we lacked the expertise and technical solutions to swiftly organize automatic uploads. This approach limited the scope of archiving but allowed us to thoroughly study the content of the archived materials and collect a different data set than automatic uploads. This is especially true compared to archiving via web crawler, which only allows for downloading data from Telegram channels for a short period and not the entire stream of messages, as is the case with our archive. The very approach to saving entire sequences of messages from each channel or chat helps preserve primary data and makes it possible to track dynamics over an extended period.

Archived materials consist of message threads, text and voice, all images, videos, stickers, GIFs, and attachments saved sequentially at certain intervals in downloadable HTML and JSON formats.

Figure **22**. File structure of archival materials exported in HTML format.

This means that for a certain period, for each channel and chat, we have a folder with the corresponding HTML file and folders with all types of content related to this HTML file:

- messages.html—a file with the stream of messages from a certain period (for chats, there can be several dozen or hundreds of them, depending on user activity)
- js—a technical file for connecting JavaScript code
- css—a technical file with HTML file display styles
- images—a folder with technical images—icons in .png format
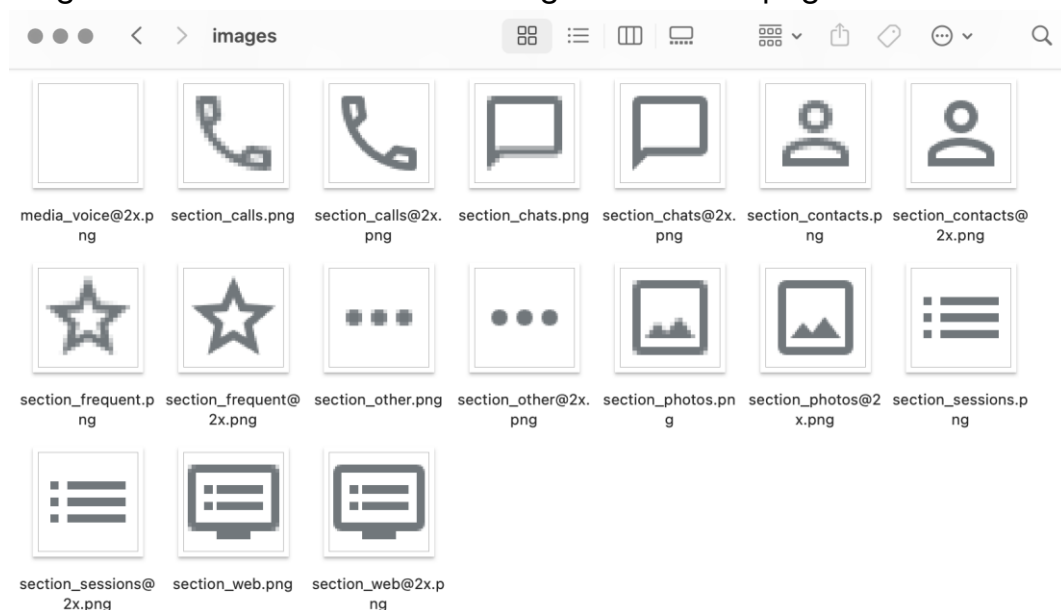
- photos—a folder with attached images, photos in .jpeg format sent in a compressed format, not as a full-size file
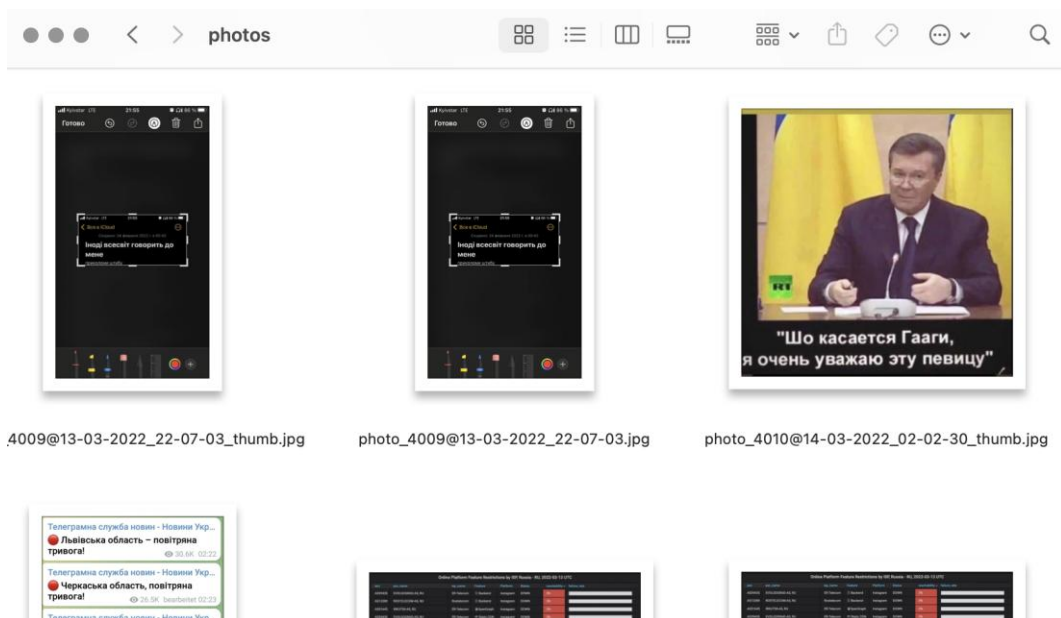
- video_files—a folder with videos sent in a compressed format, GIFs in .mp4, .MOV formats, and a preview file in .jpeg
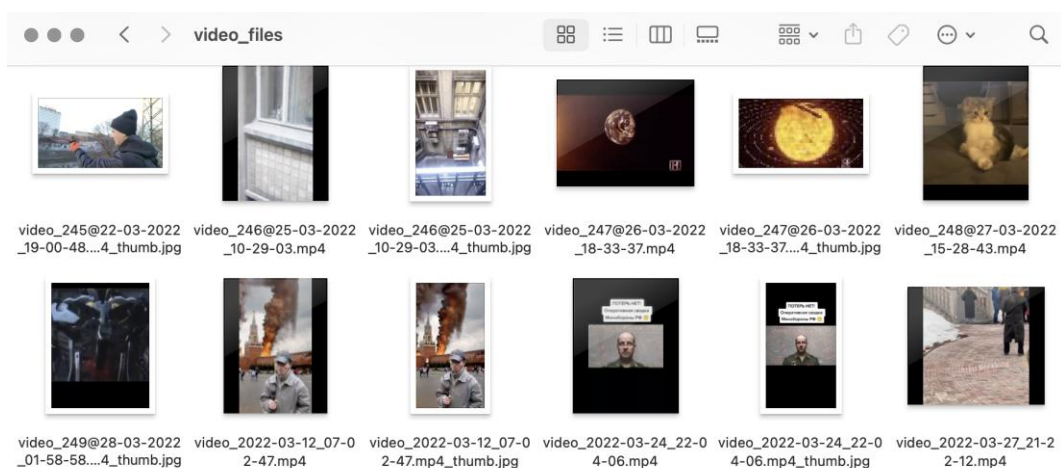
- round_video_messages—a folder with user video messages in .mp4 and a preview file in .jpeg, which have a circle format in the interface and are displayed in Voice messages
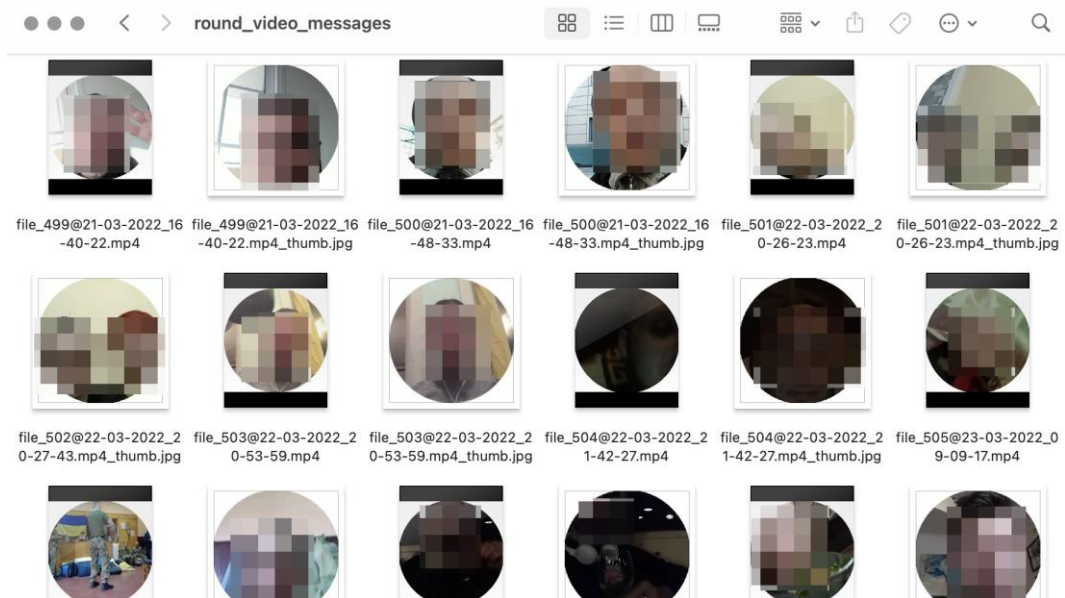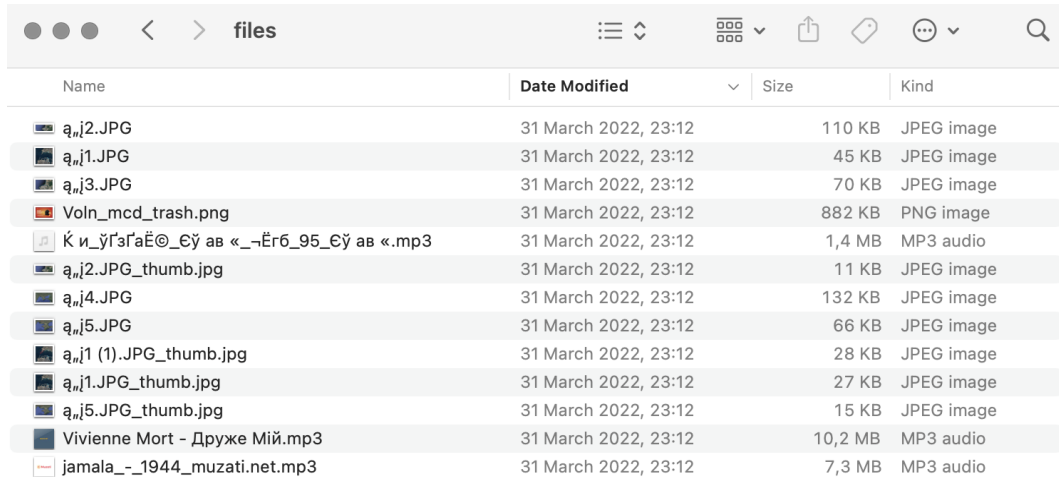
- voice_messages—a folder with voice messages in the .Ogg format, which are displayed in the interface in Voice messages

- stickers—a folder with stickers in .webp and .tgs formats and a preview file in .jpeg
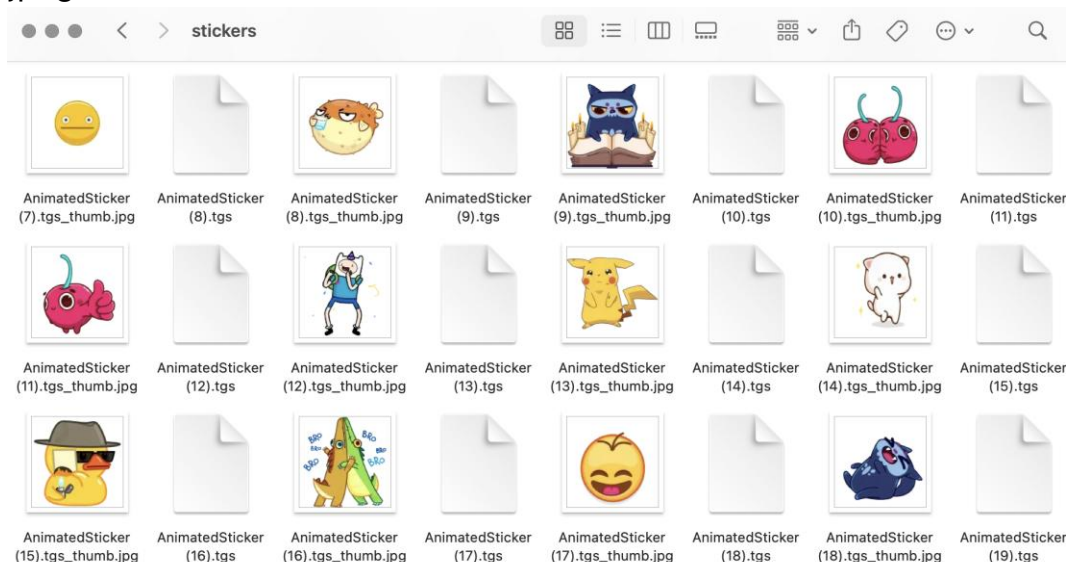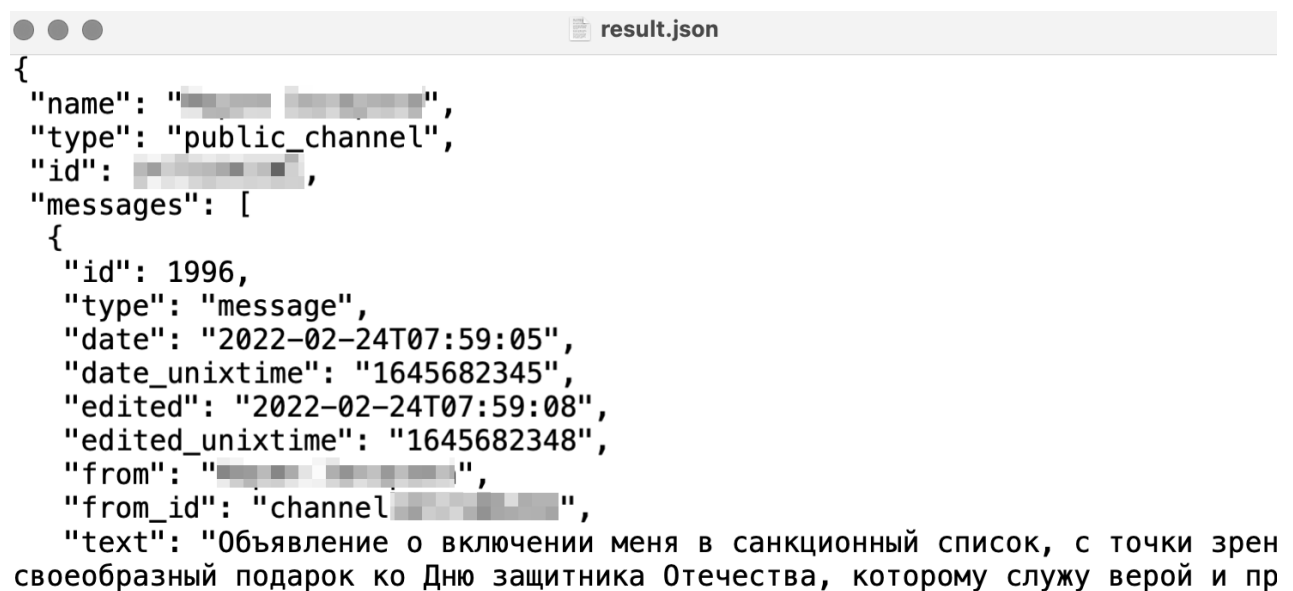
● files—a folder with attached files in various formats (.pdf, .mp3, .wav, .mp4, .mov, .gif, .docx, .fb2, .zip, etc.), including full-size images in .jpeg and .png

The decision regarding file format selection was made urgently as we needed to begin collecting data promptly. Initially, we opted to export the data solely in HTML format. This choice was motivated by HTML facilitating accessibility for both machines and, more importantly, humans. Additionally, HTML serves as the foundational format for archiving websites, enabling the conversion of materials into other formats and the extraction of structured data.

Nevertheless, the question of format selection remained pertinent. We consulted experts, examined the practices of other archival initiatives, and considered the requirements of researchers, who constitute the primary Archive audience. In January 2023, we initiated duplicating all existing archival units while re-archiving and retaining message histories for relevant periods in JSON format. This decision followed a workshop with researchers held in December 2022.

Figure **23**. Syntax and structure of an archive unit exported in JSON format.

This format is better suited for automated analysis and contains more accurate archival information, including:
- User IDs (even if they changed their Username) ("from_id": "user[...]")
- The exact time of publication (regardless of the time zone during message history export) (due to the presence of "date_unixtime")
- The time of editing the publication if the editing happened ("edited")
- Discussions in the thematic branches of organized chats ("reply_to_message_id":
- each thematic branch has its own ID at the end of the link)
- Discussions in the "embedded" chats (publications of the main channel will be marked "from_id": "channel[...]"), and comments from contributors— "reply_to_message_id":)

We managed to save 2502 units out of 2711 in JSON format.

As a result, most archival units will be available to researchers in both formats. Since downloading an archival item in HTML and JSON, even for the same period, did not occur simultaneously, this will allow us to track the dynamics of changes in messages for each archival item.

User geolocation data is not a separate type of data available for downloading and is, therefore, not separately identified in the Archive structure. However, users can publicly share their geolocation in personal correspondence or chat. Such messages with geolocation contain geographic coordinates (longitude and latitude) of a specific location. This information can pertain to the location of the user who posted the message or another location. Therefore, geolocation messages may also be contained in our Archive data.

The uploaded files are saved in .zip and .rar formats (uncompressed).

3. Access, ethical and legal issues

Archiving social media largely depends on the platforms' privacy policies and technical capabilities. Telegram's privacy policy remains relatively liberal in defining the boundaries between public and private data. The platform does not grant access to data such as phone numbers or IP addresses of users, enabling users to maintain privacy and anonymity. Additionally, there is an option to create secret (peer-to-peer) chats stored on users' devices only. One can also set up auto-delete messages. However, in most cases, the export of chat and channel content, i.e., the history of user communication, is available for download by default to anyone ("Telegram Privacy Policy," 2018). Moreover, content moderation by the platform's administration is virtually non-existent or minimal compared to other platforms.

This necessitates us to carefully define the scope of archiving, the sensitivity levels of the data, and further access to the archived materials. As a result, Archive materials have remained inaccessible for use since the archive's inception. Efforts are underway to establish an access infrastructure.

Our priority is to maintain the integrity of archival units to the greatest extent possible, as this is an integral aspect of the archive's reliability as a source. Therefore, any interference with the content of uploaded materials, including alteration or deletion of sensitive data, is prohibited. Instead, we direct our efforts toward accurately labeling different types of sensitive content and creating a secure mechanism for accessing archival materials.

We have chosen to focus solely on public channels and groups that can be freely joined. Once messages are published in such channels and chats, the data becomes public and available for archiving without violating copyright or personal data laws or the platform's privacy policy.

However, third-party access to these Archive collections is subject to several ethical restrictions, which we endeavor to consider in developing the Archive access protocol. In the future, access to archival materials will be subject to various restrictions, such as registration with personal data, institutional affiliation, anonymization and non-disclosure agreement, declaration of the purpose of data use, etc. Additionally, some archival materials will remain inaccessible to researchers indefinitely, at least until the end of martial law. We have categorized archival units according to three sensitivity levels to facilitate this. Both the principles for determining sensitivity levels and the categorization of archival units will be periodically reviewed.

Sensitivity levels

Considering security risks and ethical responsibilities regarding the future use of personal data within the archive, the Archive team is actively working to assess the sensitivity level and appropriately label the content. The Archive contains images depicting deceased and mutilated bodies, as well as scenes of violence, which necessitates a warning about the presence of graphic content that may provoke shock or evoke negative emotions across the entire archive.

The primary principle guiding the determination of sensitivity levels was the presence and nature of personal data that could infringe upon individuals' privacy rights or pose security risks.

In exceptional circumstances, we thoroughly reviewed the entire content of archived messages, verified the absence or presence of sensitive data, and assessed the source's nature. However, in most instances, the decision to categorize sensitivity levels relied on a preliminary analysis of messages within specific timeframes. Consequently, individual messages requiring further labeling may not be entirely excluded.

Due to the impossibility of verifying all archived content, sensitivity levels are primarily determined based on various structural characteristics (such as channel or chat type) and an evaluation of the source (including topic, geographical relevance, declared purpose of creation, etc.). All archived materials are classified into three categories based on the sensitivity level of the data: moderately sensitive, sensitive, and highly sensitive.

REGISTRATION, NON-DISCLOSURE AGREEMENT		UNPUBLISHED
MODERATELY SENSITIVE DATA	**SENSITIVE DATA**	**HIGHLY SENSITIVE DATA**
— public information — official sources, mass media — personal blogs or thematic channels (humour, art, analytics)	— sensitive content (18+) — personal data — propaganda — channels of occupied territories	— databases, indexes — chats of occupied territories — sensitive military information — illegal practices

Figure **24**. Sensitivity levels of archival materials.

Highly Sensitive Data (1170 archival units: 324 channels, 846 chats)

Access to these materials will be completely restricted indefinitely, a principle enforced at least for the duration of hostilities. However, the Archive will regularly reassess sensitivity levels, so the access to certain materials may change over time.

The primary criterion for determining sensitivity levels is the presence of personal data and information that could jeopardize the safety of civilians and military personnel. Materials containing such information are categorized as highly sensitive.

Most restricted access materials are chats originating from the de-/occupied territories. These chats represent invaluable sources of evidence regarding various aspects of daily life, the new military landscape, and the experiences and sentiments of the occupied population. They include chat rooms from local communities, resident blogs, chats for arranging transportation, chats for individuals of conscription age and their relatives in temporarily occupied territories of Ukraine, and chats for refugees fleeing to Russia, among others. Despite their research value and the accessibility of primary sources on Telegram, materials collected in our Archive will remain inaccessible until at least the end of hostilities. Publication within these chats poses additional risks for individuals under occupation and their families, as personal data may be used to identify and track them, potentially leading to persecution based on political beliefs.

Highly sensitive archival units also include those posing a security threat upon publication, such as channels dedicated to locating missing persons or evacuees, which deliberately disclose individuals' full information—names, dates of birth, and addresses. Additionally, channels that distribute "collaborator databases" and allegations of crimes against national security foundations, along with those systematically sharing sensitive audiovisual content and personal data about fallen soldiers, fall into this category.

Chats containing data potentially related to the course of hostilities, prisoner exchanges, or the personal safety of military personnel will not be published. These are public channels and chats related to prisoners of war and missing soldiers, their search, where their call signs, positions, and units are mentioned. These are also unofficial chats of military units with testimonies from certain areas of the frontline, discussions of military life, problems with equipment, conflicts with the command, legal advice chats for the military, and military blogs with combatants' reflections that may differ from the official position. Furthermore, pro-Russian chats disseminating coordinates of Ukrainian strategic infrastructure or information about the consequences of shelling, which could be classified under current legislation as

unauthorized dissemination of information during a state of war or emergency, will also not be published.

Highly sensitive data includes chats with content related to children—mothers' communications with photos that allow personal identification of minors; chats with information about children from the occupied territories who need adoption; content allegedly created by children (chats with a school-age audience—school chats with video and audio messages).

Highly sensitive data includes channels and chats with content likely related to illegal practices. The Archive cannot be tasked with qualifying the legality or illegality of such practices, but the collected materials may contain highly sensitive data. Mostly, these are channels and chats with discussions on mobilization avoidance. And also, all practices that are qualified as fraudulent by the users themselves (blacklists with screenshots of correspondence, contacts, personal data of "fraudsters"—carriers, landlords, sellers of goods, service providers).

Sensitive data (1190 units: 387 channels, 802 chats, 1 chatbot)

The sensitive data category includes almost all chats, provided they do not contain highly sensitive information. By volume, this one is the largest category in the archive.

Although the publication of personal data is not the main purpose of most chats, users are free to publish financial information, contact details of individuals, information about person's health or social status, as well as personal circumstances and experiences that can identify or harm a person. Telegram rules do not de facto restrict users from publishing any materials. Archivists do not have the opportunity to check all messages from archived chats and ensure that there is no personal data. Therefore, chats are defined by default as a category that may contain personal data.

The thematic range of chats classified as sensitive data is extensive. Here are some examples (for further details regarding the thematic range, see Collections):
- local chats of the rear regions or regions that were not occupied;
- professional chats where users discuss the problems of public utilities and telecommunication services;
- economic chats where users of the rear regions post their commercial ads.

Some categories of channels also fall into the category of sensitive data:
- channels containing a large amount of personal data, but do not contain information that could potentially be classified as highly sensitive data (for further details, see Highly Sensitive Data). These channels include buying/selling channels

that regularly publish personal data in ads (house addresses, phone numbers, names of sellers, etc.);
 - all channels from the occupied territories (unless they contain highly sensitive information, such as channels with lists of collaborators' personal data) and propagandists' channels. As in the case of chats, there is no way to check whether these channels contain highly sensitive information. Therefore, it was decided to place all propaganda channels in the sensitive data category. Other channels are mostly placed in the moderately sensitive data category.

Moderately sensitive data (348 units: 337 channels, 11 chatbots)

These archival units are where we assess the likelihood of encountering personal data and sensitive content that violates the right to privacy to be the lowest.

Moderately sensitive data includes channels and bots, not chats, where users can post content. Even though administrators moderate chats, various personal data may be published there. Therefore, we treat all chats as sensitive or highly sensitive data.

Firstly, moderately sensitive data includes Telegram channels of officials, public figures, and organizations. Official channels contain public information that is consistent with the position of the authorities and military command. The likelihood of violating Ukrainian law and ethical standards in such channels is low.

We can consider media channels registered under the Law of Ukraine "On Media", which imposes additional restrictions on the publication of sensitive data, moderately sensitive data. Some media outlets consider the Telegram channel a supplementary communication tool, and a significant amount of the channel content constitutes links to the main communication channels.

Author blogs and channels with analytical, cultural, socio-political, and humorous content, where there is no personal data, are classified as moderately sensitive.

Access

Third parties will be able to access the Archive (or parts of it) only after registration, outlining their intentions, and signing an agreement that obliges the researcher to anonymize personal data if it is necessary to publish sensitive content. At the same time, archival units may contain copyrighted materials. Copyright law regulates their use from the archives' collections.

Almost from its inception, the Archive has received requests for access to materials from researchers, journalists, and artists. These requests cover a wide range of topics, including research on migration, internet memes, propaganda narratives, gendered experiences of war, and the search for evidence of war crimes. However, given the above challenges, direct access to archival materials is currently not feasible. Therefore, the Archive team is developing possible access formats depending on the needs and level of sensitivity of the data: a user interface for viewing text and media messages, requests for access to raw data, and joint access workstations for archivists and researchers.

The problem of accessing and working with sensitive data was the topic of many presentations and seminars related to the archive. One of them, "Russia's War in Ukraine": Interdisciplinary Data Sprint with Telegram Archives, took place on December 16-18, 2022, and was organized as part of a collaboration between Miglė Bareikytė, Yarden Skop (SFB Media of Cooperation at the University of Siegen), and Taras Nazaruk (Center for Urban History). During the workshop, the team of archivists presented the context of the archive's creation, as well as the key methodological and ethical challenges in making archival units available for research to the interdisciplinary group of researchers. These challenges were discussed using a specially prepared test dataset. Sharing and working with the test data occurred through a prototype application developed in cooperation with TGSearch.
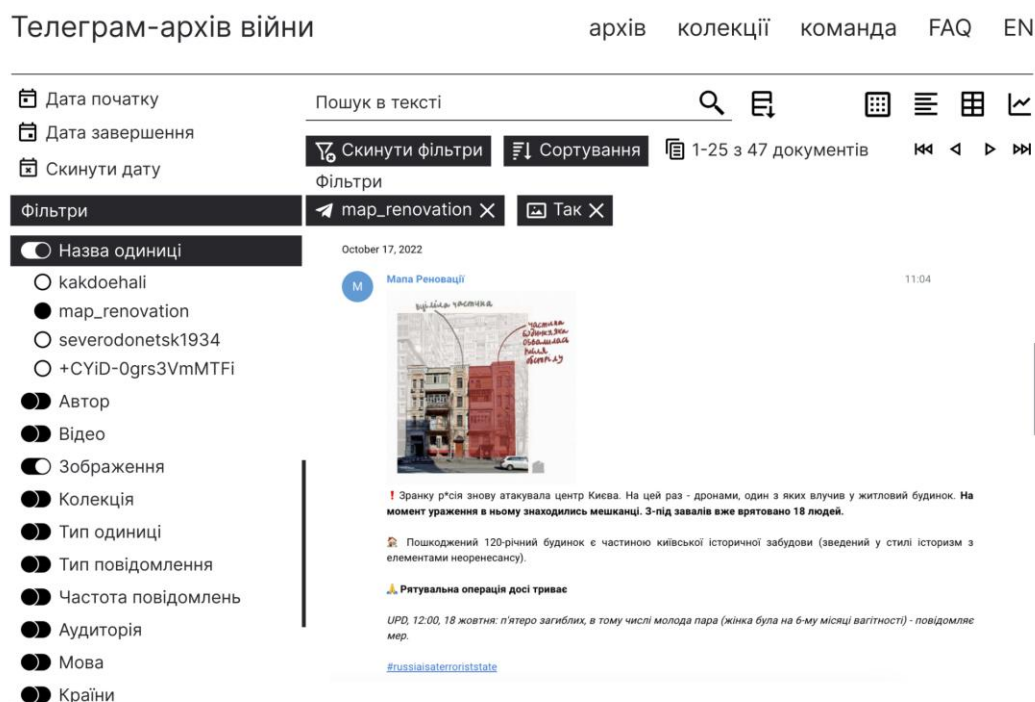

Figure **25**. Prototype of the future Archive interface.

The application allows navigation through the Archive materials with sets of filters based on metadata, text search, and time periods. Based on the results of this

project, the prototype is being finalized to resolve technical errors and improve functionality. The improved tool version will become one of the main methods of working with the accessible part of the archive.

Additionally, the Archive users will have access to the register of channels and chats, as well as the register of collections. This will enable researchers to review and work with the metadata of archival units. The mechanism of access to the raw data of the Archive has not yet been developed. Requests for such access, if any, will be considered individually.

Since November 2023, the work on creating an access format for the collective work of archivists and researchers has also been part of Taras Nazaruk's PhD research within the research group NFG 026 "Digital Stories of Violence in the 21st Century. Participatory War Publics and Autobiographical in Russia, Belarus and Ukraine" at the FernUniversität in Hagen.

4. Publications and presentations

Academic publications about the Archive / based on the Archive materials

Bareikytė M., Nazaruk T., Skop Y., Martin A., Makhortykh M. (2024) How should platforms be archived? On sustainable use practices of a Telegram Archive to study Russia's war against Ukraine (accepted, Media, Culture & Society)

Nazaruk, T. (2022). Subscribe and Follow. Telegram and Responsive Archiving the War in Ukraine. Sociologica, 16(2), 217-226. https://doi.org/10.6092/issn.1971-8853/15339

Shumylovych B., Makhanets O., Nazaruk T., Otrishchenko N., Brunow D. (2022) Preserving the now! Mediating memories and archiving experiences in Ukraine. NECSUS_European Journal of Media Studies, Jg. 11, No. 2, S. 126-149. http://dx.doi.org/10.25969/mediarep/19171.

Participation in seminars, workshops, and conferences

- Preserving the Now! Mediating memories and archiving experiences in Ukraine, NECS, 03/09/2022

- [Preserving the Now! Emergency Web Archiving as a Way to Survive and Resist the War in Ukraine,](#) University of Giessen, 05/03/2022

- [Memory Under Fire. Archiving and Data Practices during War,](#) University of Siegen, 05/23/2022

- Witnessing the Now, workshop by German Historical Institute Warsaw, 05/27/2022

- [IIPC Webinar: Web Archiving the War in Ukraine](#), 08/31/2022

- On the Telegram Archive of War at "Migration, Mobility, and Displacement in the Black Sea Region," *Annual Conference of the Center for Governance and Culture in Europe at the University of St. Gallen*, Tbilisi, 09/13/2022

- On the Telegram Archive of War at "[Documenting the War. Past and Present](#)," *European Holocaust Research Infrastructure, Center for Urban History, online*, 10/06/2022

- [Documenting War Experiences: Oral Testimonies and Social Media Archives in Ukraine](#) at the event series "Russia's War on Ukraine in Historical Perspective," *George Mason University*, 10/31/2022

- "[Russia's War in Ukraine](#)" - Interdisciplinary Data Sprint with Telegram, *The University of Siegen and Center for Urban History*, 12/16-18/2022

- Discussion program "[Memory of War](#)," Center for Contemporary Culture in Dnipro, 12/7/2022

- "Telegram Archive of the War in Ukraine. Collecting Stories by Extracting Data", "*How to Document the Present: Oral History and Memory of the War*" Winter School, *University of Insubria - Center for Research on Minorities and University of New Europe (UNE)*, 01/24/2023

- A Year of Archiving the War: From Emergency Response to a Challenge of Sustainability, symposium "*Languages of War: Interpretive Knowledge and Debates on the Perspectives of the War in Ukraine*," *Gießener Zentrum Östliches Europa Justus-Liebig-Universität Gießen,* 02/22/2023

- Telegram archive of the war. Documenting the new reality. Seminar "War: practices of documenting/archiving digital space" Institute of Archival Affairs and Documentation, 10/16/2023

- "Telegram Archive of the War: Sources to Unfold Wartime Reality," 10. MuseumDigit Conference, Hungarian National Museum, 11/22/2023

In June-July 2023, a series of online seminars and discussions on theoretical, methodological, and practical aspects of social media archiving were held, initiated by Mykola Makhortykh, Ani Baghumyan (University of Bern) in collaboration with Bohdan Shumylovych and Taras Nazaruk (Center for Urban History) and with the support from the University of Zurich. Speakers: Jeff Deutch (Mnemonic / Syrian Archive), Sara Day Thomson (University of Edinburgh), Niels Brügger (Aarhus University).

Mentions in the media

- Emergency Archiving in Ukraine. Historical Research in Times of War (ZZF Podcast, April 13, 2022)
- Ukrainische Historiker dokumentieren den Krieg (Deutschlandfunk Kultur, April 20, 2022)
- En Ukraine, "le réseau Telegram peut être plus rapide que les bombes" (Mediapart.fr, May 6, 2022)
- MayDay Magazine (May 9, 2022, друк)
- "Ми боремося за збереження пам'яті" (Детектор медіа, May 20, 2023)
- Web Archiving the War in Ukraine (International Internet Preservation Consortium Blog, July 21, 2022)
- Social Media als Faktor im Ukraine-Krieg (ORF Topos, February 19, 2023)
- Ukrainians' Google Searches Reveal a Year of Fear—and Hope (Wired, March 14, 2023)
- Ausgelöschte Erinnerung (Frankfurter Allgemeine Zietung, November 23, 2023)

5. Acknowledgements

We extend our gratitude for the direct and indirect support, valuable advice, and guidance from individuals including Oleksii Chebotarov (University of Vienna, University of St.Gallen), Miglė Bareikytė (European University Viadrina Frankfurt (Oder), Jarden Skop, Alexander Martin (University of Siegen), Mykola Makhortykh

6. Bibliography

1. Chambers S., Birkholz J., Geeraert F., Pranger J., Messens F., Lieber S., Mechant P., Michel A. & Vlassenroot E. (2021), BESOCIAL: Final report Work Package 1 an international review of social media archiving initiatives.

2. Der Spiegel, 2022
3. Ghasiya P., & Sasahara K. (2023). Messaging strategies of Ukraine and Russia on Telegram during the 2022 Russian invasion of Ukraine. First Monday, 28(8). https://doi.org/10.5210/fm.v28i8.12873
4. Lazaruk K. (2022). Bunk Debunkers: Warfakes as Russian War Participatory Propaganda. Institute of Network Cultures, April 5. https://networkcultures.org/blog/2022/04/05/bunk-debunkers-warfakes-as-russian-war-participatory-propaganda/.

5. Loucaides D, The Kremlin Has Entered the Chat, Wired, February 2, 2023. https://www.wired.com/story/the-kremlin-has-entered-the-chat/

6. Marechal N. "From Russia With Crypto: A Political History of Telegram." FOCI @ USENIX Security Symposium (2018). https://www.usenix.org/conference/foci18/presentation/marechal.

7. Maschmeyer L., Abrahams A., Pomerantsev P., Yermolenko V. (2023) Donetsk don't tell - 'hybrid war' in Ukraine and the limits of social media influence operations, Journal of Information Technology & Politics, DOI: 10.1080/19331681.2023.2211969

8. Mikhailidi K. (2023). Debunking of fake news in Telegram in 2022: comparative study (Doctoral dissertation).

9. Opora (2023) Media consumption of Ukrainians: the second year of full-scale war. Opora survey, June 10.

10. OPORA Survey. (2022). Media Consumption of Ukrainians in a Full-Scale War. https://oporaua.org/report/polit_ad/24068-mediaspozhivannia-ukrayintsiv-v-umovakh-povnomasshtabnoyi-viini-opituvannia-opori.

11. Plakhta D.. "TELEGRAM AS A TOOL FOR POLITICAL INFLUENCE AND MANIPULATION." TV AND RADIO JOURNALISM. 2020. Issue 19. P. 88-94 (2020).

12. Yuskiv B, Karpchuk N., and Pelekh O. "The Structure of Strategic Communications during the War: the Case-Study of the Telegram Channel Insider Ukraine." Politologija 107.3 (2022): 90-119.

13. "Russia's War in Ukraine" - Interdisciplinary Data Sprint with Telegram. CRC Media of Cooperation, University of Siegen & Center for Urban History, Lviv: https://www.mediacoop.uni-siegen.de/de/veranstaltungen/russias-war-in-ukraine-interdisciplinary-data-sprint-with-telegram/.

Appendix 1

Archival materials stored outside the Telegram platform

The Archive contains additional digital materials related to the topics of the main collection but collected from various resources and sources outside the platform. These sources are city forums, news websites, and local radio. These complement the main archive and are characterized by eventfulness and thematic selectivity. These materials are described in a separate internal register. Their total volume is 66.7 GB.

While creating the Archive for certain regions (particularly the temporarily occupied Crimea and Sevastopol), we noticed the popularity of other platforms outside of Telegram, such as *VK* and city forums. Therefore, to fully preserve evidence of the reception of the "SMO" ("Special military operation," an official term used by the Russian government and pro-Russian sources to denote the Russian invasion of Ukraine) and later the war among the inhabitants of the peninsula, the effects of sanctions, and military losses, we archived thematic threads on forums and, in some cases, discussions under posts on *VK* community pages.

City forums

The lack of coverage of topics valuable to the archive, such as news from the Russian Navy and the effects of economic sanctions on the lives of peninsula residents, was resolved by archiving discussions from the Sevastopol City Forum. This forum serves as a popular resource for discussions on various aspects of city life, encompassing discussions on political and military topics with a focus on the Russian Black Sea Fleet, whose bases are located in most of the city's bays. In its threads, one can trace local panic following drone attacks and the sound of explosions, attitudes towards the "special operation" and its economic consequences, as well as increasing doubts about its practicality.
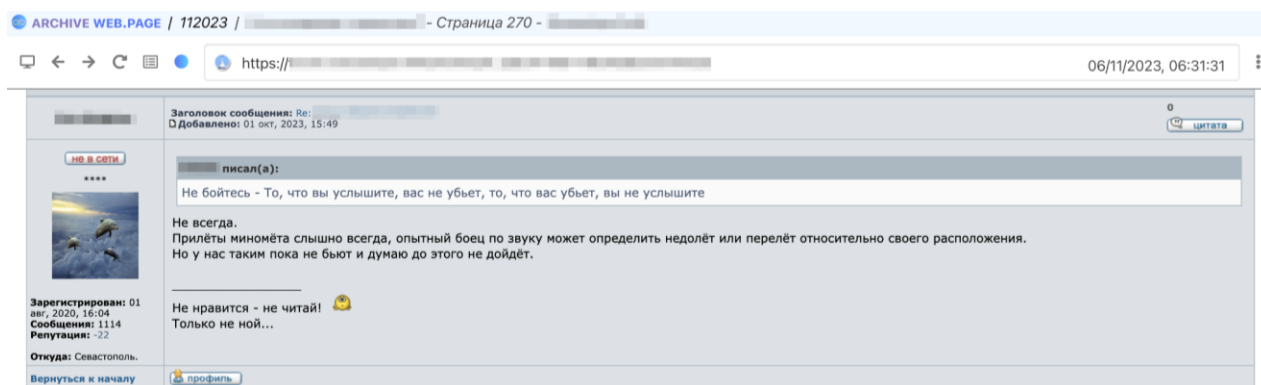
Figure **1**. Discussion in a thread on the city forum of the temporarily occupied Sevastopol, saved using the ArchiveWeb.page extension

Communication on the forum is highly moderated, and users are ranked by reputation and sanctioned for expressing unpopular opinions or those contrary to Russian propaganda. In particular, in the saved threads, there are profiles marked "Banned: indefinitely" and "Banned until [a certain date]," indicating their blocking by IP address or removal from the Forum. Some messages deleted by moderators before being saved are recorded in the Archive as screenshots (.png format and "_deleted" in the title). Sometimes, users delete the content of their messages themselves, replacing the expressed position with ".".
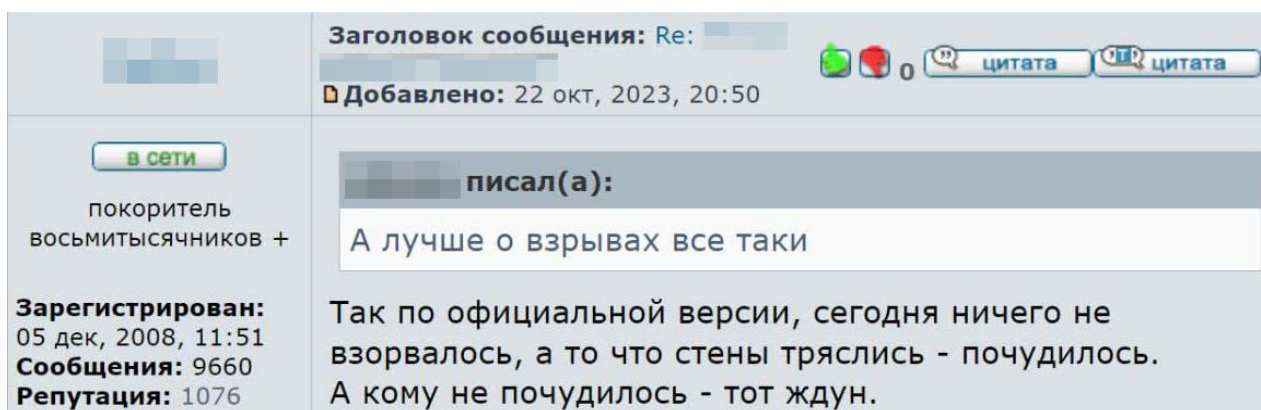

Figure **2**. Screenshot of a message deleted from the thread titled "What are these strange explosions?" on the sevastopol.info forum.

The forums can serve as an additional source for researching propaganda discourses and conspiracies common in political discussions of the pro-Russian population of the government-controlled and temporarily occupied territories. For instance, the "SMO" discussion thread of the Sevastopol Political Forum was preserved.

A signature (text automatically added to messages if specified in the user's profile), a popular means of additional self-identification among forum users, deserves special attention. During the war, it is a place to demonstrate a pro-government chauvinistic position.

Коронавирус показал, кто у нас на самом деле болен

Украинцы это не народ, а созданные из русского этнического материала мутанты, появившиеся при отсутствии каких-либо устоявшихся норм и традиций и утратившие свою национальную самоидентификацию.
__
Украинская химера — https://www.litmir.club/br/?b=136590

Когда украинец умнеет, он становится русским.

Украинец это русский, укушенный поляком.

С Россией можно дружить - она добродушная.
С Россией можно сотрудничать - она честная.
У России можно попросить - она щедрая.
Россию можно обмануть - она доверчивая.
Но с Россией нельзя воевать - она **НЕПОБЕДИМАЯ!**

Я слышал столько клеветы в Ваш адрес, что у меня нет сомнений: Вы — прекрасный человек! (с) Оскар Уальд

Настоящему русскому не стыдно , что он русский, а если стыдно, то он не русский и не с нами (с) Д.П.

Собкор, как Путин, виноват во всем (с)

Figure **3**: Signatures under the posts of Sevastopol political forum users.

Forum threads have been archived using ArchiveWeb.page in the .wacz. format and titled in the format "Thread name_resource address_save date."

Comments under the posts on VK allowed us to cover the impact of the war on the economy, tourist flows on the peninsula, personal financial security, and record losses among the Russian Black Sea Fleet.
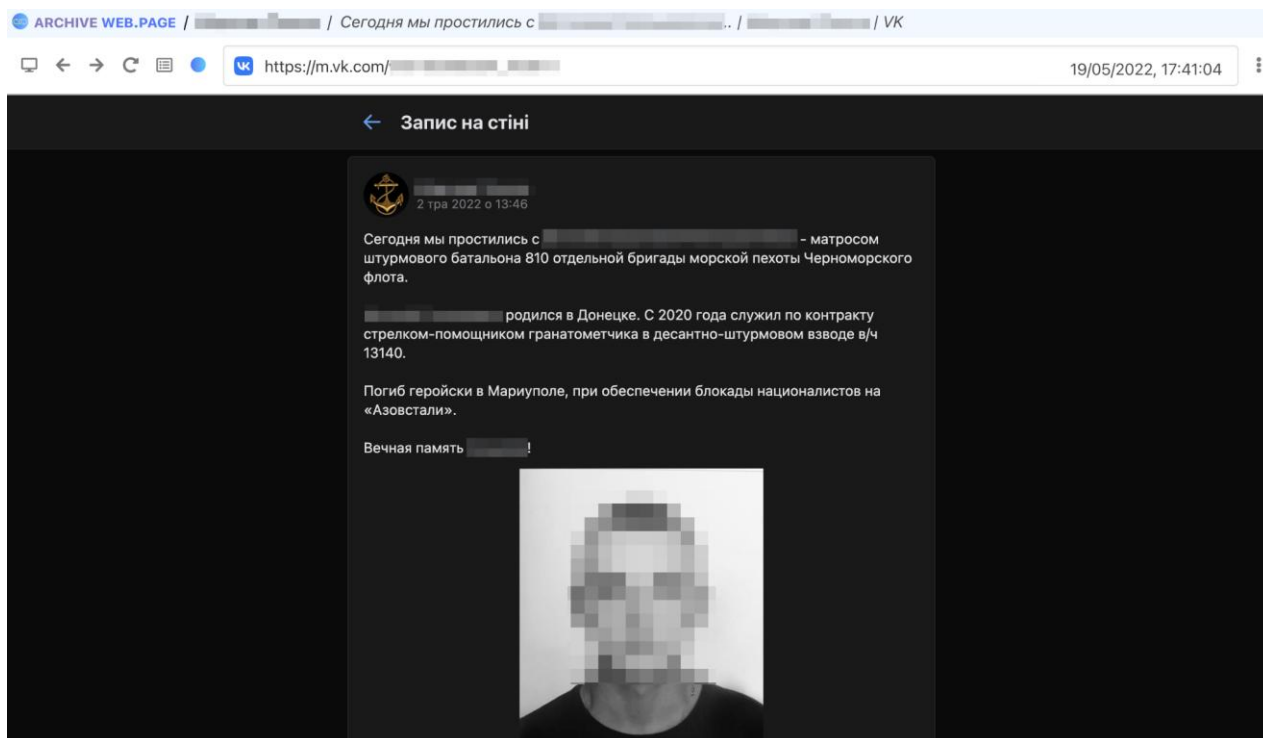
Figure **4**. A post in the VK community about a sailor of the Russian Black Sea Fleet who died during the Azovstal blockade, saved using the ArchiveWeb.page extension

The Archive has saved these materials in thematic collections of posts in the .wacz format and, partially, in cases of page-by-page preservation in the Wayback Machine.

Radio broadcasts of 105.2 station

The archive contains recordings of the radio station 105.2 broadcasts (formerly Avtoradio Melitopol), which was seized by Russian troops. 5 files in the .m4a format were archived on 03/19/2022, 03/20/2022, 03/21/2022 and 03/23/2022. On these dates, fragments of radio broadcasts were recorded with a total time of 00:49:46 (round-the-clock recording was impossible for technical reasons) and a volume of 91.8 MB. The audio recordings' metadata consists of the audio recording's creation date, file timestamp, and a brief description of the audio file content. The thematic content of the recordings: Russian propaganda, various announcements, music, "interviews," etc.

Local news websites

Additionally, the Archive-It collections include websites of news media in Ukrainian cities (Berdiansk, Chernihiv, Popasna, Hadiach) and resources dedicated to cultural heritage. They are also stored on the Archive's server in the .gz Archive format.

Figure **5**. The thematic collection of the Center for Urban History dedicated to the war is available among the Archive-It collections.

Certain publications from the Internet, deemed by the archivists to offer additional insight into the context of this war, were preserved using the Wayback Machine. Pages on third-party resources, referenced in the descriptions of archival units, have been preserved similarly.

Digital materials are organized by metadata, including ID, resource title, its original link, link to the saved page in the Wayback Machine (if available), resource city or region, archival description, keywords (borrowed from the main registry), creation date, technical comment (regarding the saving of resource pages in .wacz format).