Data for Ukraine

Report Week April 21 – 27

Introduction

This report provides an overview of the key highlights of data collected by the Data for Ukraine Project during the period of April 21-27, 2022. Below we examine three themes: patterns around what information was most prominently shared, major events in our four categories of interest, and key highlights about the quality of sources most often employed. For full information about the Data for Ukraine Project and an explanation of our methodology please see last two sections.

Most Shared Information

In our review of the most shared information for the week of April 21- 27, we find interesting differences in the nature of the information depending upon whether the language source was Ukrainian, Russian or English. Overall, the language that most closely reflects current events on the ground is Ukrainian. Not surprisingly, Russian language tweets are also useful for events in Luhansk and Donetsk Oblasts, though for the most shared ones at least, they tend to relate to events that are already at least a few days old. The most shared tweets in English tend to focus on events outside of Ukraine itself.

*Displaced Persons.* The most shared Ukrainian language information focused on the ongoing seige of Mariupol and on changes in the situation along the Polish-Ukrainian border. There, first time since the start of the war, there is busy traffic and long delays entering Ukraine as opposed to leaving. Another feature of the Ukrainian language tweets were calls for change in policy to allow ordinary citizens should be allowed to flee the war if they so choose, while only public officials should have their travel restricted.

By contrast, the most shared Russian language tweets on military activities in specific localities in Luhansk Oblast (namely, Severodonetsk and Kremenaya) and Donetsk, as well as the siege of Mariupol and ongoing blockade of the Azovstal Plant. Interestingly, most of the “news” in these tweets related to events that were about a week old. The information also included appeals by officials for citizens in Luhansk specifically to immediately evacuate.

In the English language batch of the most shared information on displacement the focus was on the broader issue of mass the mass displacement of Ukrainian children (60% had to flee their homes).

*Human Rights.* Our analysis of the most shared information on human rights abuses – finds that the Ukrainian language information most shared is focused on Kherson and the mass atrocities and war crimes taking place including the export of grain from Kherson to Crimea, the imprisonment and replacement of local councils and heads Kherson oblast, the torture of Ukrainian Orthodox Priest in Kherson and the shooting of children and elderly in the region. Russian language information is again covering an earlier period but focuses on human rights abuses in specific localities in Luhansk oblast. There were many Tweets referring to rape being used as weapon of war, about looting and Russian bombardments, often of civilian targets around Kharkiv and other eastern towns.
Humanitarian support. Again, we see a clear pattern, information shared in Russian focuses on specific localities in Luhansk and is delayed by about a week. The English language information focuses on the ongoing inability to open humanitarian corridors in Mariupol and Russia’s refusal to cooperate with international actors to do so. Tweets reflected the large number of announcements this week of international military and humanitarian assistance to Ukraine. There were also stories about the welcome given to refugees in other countries, as well as the supply shortages of the Russian army namely the fact the Russian army aid is being crowd sourced and sent by foreigners. The Ukrainian language information follows up on this theme highlighting that the Russian army is forced to conduct blood drives among civilians in Belgorod. The focus is also on crimes related to the distribution of humanitarian aid in Occupied Kherson oblast, where new Russian business ventures are accused of selling Russian products to civilians in Kherson at raised prices – whilst the majority of humanitarian aid distributed is of Ukrainian origin. This is in addition to information about Russians sabotaging evacuation efforts.

Civil Resistance. We found most interesting differences between the kind of information most shared in the three languages when it came to civilian resistance. In the Ukrainian and Russian languages, information focused on the civilian resistance to occupation and repression, mass imprisonment and execution of ordinary citizens and officials in Kherson. By contrast, the English language information most often shared focused on acts of civilian engagement and resistance outside of Ukraine. Namely, acts of arson against of Russian enlistment offices, an increase in enlistment in the Polish Territorial Defense Volunteer group, and the fleeing of Gazprom VP from Russia to Ukraine who also joined the territorial defense. This highlights a general trend in western media attention – moving away from the ongoing resistance in occupied territories such as Kherson. This also suggests that there is limited knowledge outside of Ukraine on the extent of human rights abuses occurring in Kherson, while the rest of our data suggest that the situation on the ground is dire.

Major Events

This week we identified several “major events” in all four categories.

Displaced Persons. For the category of displacement, there were major events between April 19 and 21 (Figure 1), occurring along the borders in the west and north of the country as well as in the south east of Ukraine which is in large part occupied by Russian forces (Map 1). On April 20, the major event was the result of information about the partial opening up of an evacuation corridor from Mariupol as well as reporting about increased military attacks in the Donbas and the movement of people from these areas of Ukraine.
Figure 1 Displacement of People Major Events Timeline April 15-29, 2022

Map 1 Displacement of People Location of Major Events April 15-29, 2022
Human Rights. We continue to observe the frequent occurrence of successive major events (Figure 2). These tend to occur with the highest frequency in Kharkiv, Donetsk, Luhansk, Kherson, Dnipropetrovsk, and Mykolaiv Oblasts which are all either experiencing partial occupation or increase in fighting and shelling (Map 2). We also see from the map activity that the scale of events occurring continues to be very large.

On April 24, major events of human rights abuses occurred in Kharkiv, Kherson, Donestk, Luhansk Oblasts. Forced mobilization of young people and doctors is reported Kharkiv, Kherson and Zaporizhia Oblasts.

In Kerhson, a priest was specifically targeted leading up to Easter, kidnapped and tortured. Russian military were also reported taking food and clothes from civilians, and exporting food and grain from the city and elsewhere across and the south of Ukraine. To quell ongoing resistance in Kherson, the Russian military stands accused of tricking ordinary citizens by placing Ukrainian flags on tanks and then shooting at civilians from those tanks.

In Kharkiv oblast civilian infrastructure was targeted. Civilians were targeted and killed in Chuguiv rayon. In Vovchansk also in Chuguiv rayon, a prison was set up in a factory. Russian held Ukrainian prisoners of war are being held for ransom and family members are receiving demands for cash in Euros.

Several major events occurred in Luhansk and Donetsk oblast. The encirclement and bombing of the Azovstal plant in Mariupol continued, with Russia refusing to allow safe passage or evacuation corridor. In Ocheretyn, also in Donetsk Oblast, children were targeted and killed. Civilians and civilian infrastructure were also targeted in Luhansk. In Luhansk, 7 temples and an Islamic center were targeted and destroyed. Members of the OSCE Special Monitoring Mission were captured by Russian forces in Donetsk and Luhansk.

Elsewhere on April 24, civilians killed in Pavlograd district in Dnipropetrovsk and civilian infrastructure (railways) was also specifically targeted and civilians and civilian infrastructure targeted in Sumy region, where a psychoneurological boarding school was targeted.
Figure 2 Human Rights Abuses Major Events Timeline April 15-29, 2022

Map 2 Human Rights Abuses Location of Major Events April 15-29, 2022
Humanitarian support and needs. We identified four major spikes in the data on the 17, 18, 19, and 27 (Figure 3, Map 3), though it was not clear from a study of the tweets what actual events underlay the spikes in the data. This highlights one of the challenges of working with such a large number of tweets. In cases where no one or two major events dominate the Twitter traffic, it can be hard to identify in a reasonable time what the sources of a given spike might have been.

Figure 3 Humanitarian Support/Needs Major Events Timeline April 15-29, 2022

Map 3 Humanitarian Support/Needs Location of Major Events April 15-29, 2022
Civilian Resistance. Finally, we see ongoing major events of Civil Resistance in Kherson Oblast (Map 4) and in Kropyvnytska Oblast, with major events being noted on April 19, 20, 24, and 27. We focus our discussion on the largest event that occurred on April 24. Kherson continues to be a major cite of contention and resistance against the Russian occupation. The major event that occurred on April 24 was a large auto convoy protest of citizens standing up against plans for an illegal referendum by “exiting” the city. Kherson continues to see excessive violence against peaceful civic resistance, with protesters being regularly shot at, tortured, imprisoned and killed. It is remarkable that Kherson continues to appear in our data when we know that internet access has been cut. As internet access is slowly being restored we expect to collect even more data of civilian resistance and repression in the Oblast.

Figure 3 Civilian Resistance Major Events Timeline April 15-29, 2022

![Civilian Resistance Major Events Timeline April 15-29, 2022](image)
Top sources

We again see that the top sources embedded into different language tweets is very different. In fact, we see that there is limited overlap between the sources even if most if not all the Russian and Ukrainian sources are available in both languages.

Looking at our Ukrainian language sources (outside of twitter and youtube links which come out on top in all three languages), among the top 10 the most shared reputable sources include: www.5.ua, www.pravda.com.ua, espreso.tv, tsn.ua, www.ukrinform.ua, 24tv.ua, censor.net, prm.ua, https://novynarnia.com, and https://www.europoint.com.ua (in order of most to least embedded).

Local Poltava based news source np.pl.ua was in the top 20 as was the website of Petro Poroshenko’s Party European Solidarity.
Less reputable sources included bbc-ccnn.com which includes a mix of misinformation and sensational information. Another source with more sensational information includes https://cripo.com-ua.

In the top 30 we also find https://most ks.ua a Kherson local media outlet.

In the top 50 there are multiple lower quality and tabloid style media sites. Some have examples of misinformation that could be interpreted both pro Ukrainian and Russian (examples include https://patrioty.org.ua, https://antikor.com.ua and military focused http://mil.in.ua).

The most shared Russian sources are not of the same quality as those in the top 10 in Ukrainian. These include: inforest.org, antikor.com.ua, www.dialog ua, ru.interfax.com.ua, www.ostro.org, ift.tt, novostey.com, and kanaldom.tv. In fact the major Ukrainian news media do not appear in the top 15 and we can identify 10 in the top 60 (in order of most to least embedded: zn.ua, www.pravda.com.ua, www.5.ua, ukranews.com, www.unian.net, 24tv.ua, tsn.ua, prm.ua, uatv.ua, nv.ua)

In Russian the most shared information came from https://inforest-org – which is run by Interfax a Russian state media agency. The information on this site primes tabloid information and so-called “click bait.” Ostro.org focuses on Donetsk and Luhansk oblasts but also summarizes information from other sources in Russia, Occupied LNR DNR and Crimea. It is run by the NGO Center for Research on Social Perspectives of Donbass and the Information agency "OstroV" and reports that the project is implemented with the assistance of: National Endowment for Democracy.


About the Project
Data For Ukraine (#DataForUkraine) is an international collaboration between scholars at The Kyiv School Of Economics, the MOBILISE project at the Universities of Manchester and Oxford (led by Olga Onuch), the Machine Learning for Peace project at Duke University (led by Erik Wibbles), political scientists at the University of North Carolina (led by Graeme Robertson) and the Inter-Disciplinary Lab for Computational Social Science at University of Maryland (led by Ernesto Calvo).

The aim of the project is to gather data on civilian resistance (CR), human rights abuses (HRA), displacement of people (IDP) and humanitarian support/needs (HS) across Ukraine, with the intention of providing timely information for private citizens, NGOs, INGOs and policymakers responding to the Russian invasion and war. To do so, we use hourly data from the Twitter API to report on the incidence and location of important events.
Methodology
The data are drawn from an initial collection of more than 400 Twitter accounts (and their followers) covering politicians, civil society activists, journalists and media at the national and local level all across Ukraine and including as broad a range of political positions as possible. Using this initial batch, we identify key networks/communities of accounts and identify key nodes of information. The list of accounts is regularly reviewed to maximize relevant content.

Tweets are automatically searched for more than 600 keywords in Ukrainian and Russian. Keywords were initially derived iteratively from social science theory and translated in to Ukrainian and Russian by native speakers. Through a period of consistent multi-team member verification and careful study of the living language actually used in Tweets about the phenomena of concern, we expanded the terms to include these ‘real world’ formulations. Once collated and collected, the keywords are classified into 4 categories with the goal of separately identifying civilian resistance (CR), human rights abuses (HRA), displacement of people (IDP) and humanitarian support/needs (HS). Following the collection of the tweet data, repeated human manual verification of a random sample of Tweets is used to ensure and improve the accuracy of classification of events. Our team is committed to continue and improve the verification through the data collection process as and when possible.

The results are displayed on the Data for Ukraine website (https://mlp.trinity.duke.edu/dataforukraine.php#en).

Our analysis of the data collected by our project between April 17 and 21, 2022 see that the information being share in the three different languages not only has significantly different focus but also relies on very different information. The Ukrainian language information covers the most immediate Human Rights violations and the continued civilian resistance. The Russian language information covers localities but often relies on led reputable sources and is delayed temporally. The most pressing issues highlighted in our data this week were the ongoing violent human rights abuses and violence against civilians in the east (Luhansk and Donetsk oblasts) and in the south of the country (in Kherson specifically) as well as the inability to evacuate civilians from these localities. We note that attention to these matters is much lower in our English language data – specifically the level of repression in Kherson and surprising ongoing resistance to Russian occupation even in the face of mass arrests and disappearances.