



Virtual Manipulation Brief

RUSSIA'S STRUGGLE TO CIRCUMVENT SANCTIONS AND COMMUNICATE ITS
WAR AGAINST UKRAINE

PREPARED AND PUBLISHED BY THE
**NATO STRATEGIC COMMUNICATIONS
CENTRE OF EXCELLENCE**



Executive Summary

In this first issue of the Virtual Manipulation Brief, we zoom in on the Russian language conversation, and how it changed as a result of the decision to invade Ukraine.

Sanctions hindered the Kremlin's messaging on Western platforms, while the focus on domestic audiences pulled many propagandists to Telegram, VKontakte, and RuTube. The number of Russian Telegram users has increased by two thirds, while four in five users of Facebook and Instagram have left Meta's platforms.

The increase in users understates the importance of Telegram. It acts as a reliable hub, where pro-regime voices can post (almost) without fear of platform censorship. Kremlin propagandist Margarita Simonyan's social media posting patterns reveal how communication has adapted following Russia's invasion of Ukraine. In response to sanctions, she, together with many journalists working for RT and Sputnik, moved to posting through Telegram. Telegram posts are automatically cross-posted to her Western platforms, meaning the messaging continues to flow at little or no extra effort. This method has the added benefit of evading attempts by Twitter to prevent the amplification of RT content.

Such use of automation means the Kremlin's

messaging will not disappear on Western platforms, even when domestic audiences are the primary focus. The comparison of Russian-language messaging about NATO on VKontakte and Twitter shows how this combination of push and pull factors has dramatically altered the Kremlin's reach. We estimate that the relative reach of pro-Kremlin messaging on VKontakte in March 2022 was a hundred times more than normal, compared to Twitter.

The Virtual Manipulation Brief builds on our Robotrolling reports, tracking how Russian bots and trolls manipulate the flow of information online. It expands the area of focus beyond the conversation about NATO to Russia's war against Ukraine. And it will look beyond Twitter and VKontakte to track how antagonists use other social media platforms to disinform the public.

In this issue, we start by examining how social media usage in Russia changed since February. A comparison of heavily amplified content about NATO on Twitter and VKontakte exposes the Kremlin's waning ability to control conversations on Western platforms. Finally, we show how the Kremlin propagandist Margarita Simonyan moved her communication to Telegram to evade sanctions directed at RT. ■

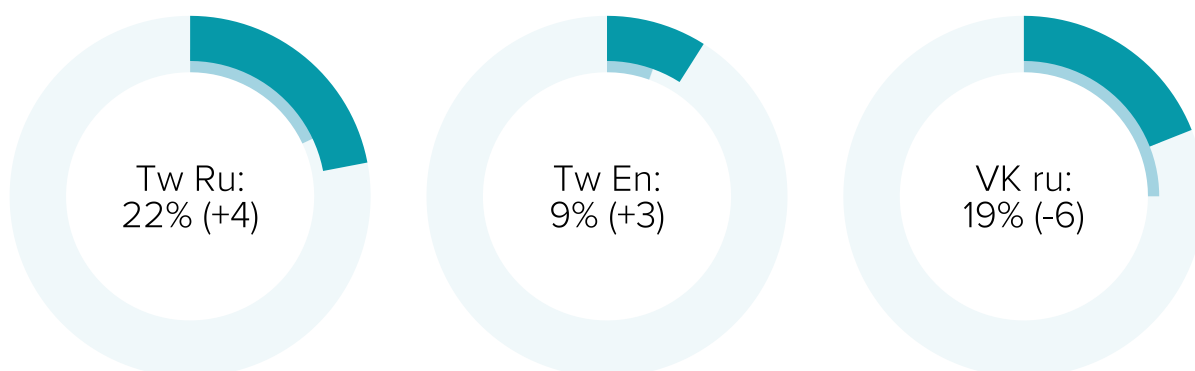


Figure 1: Percentage of automated posts about NATO in May-August 2022, by platform and compared to Feb-Apr 2022

Instagram out, Telegram in

In response to the Russian invasion of Ukraine, regulators and social media companies around the world took unprecedented steps to restrict the reach of hostile voices. Within the Russian Federation, restrictions aimed to isolate the domestic public from real information about the war. At the authorities' behest many internet providers immediately restricted access to Facebook, Instagram and Twitter.

The campaign to censor the Internet domestically within Russia was fitfully rolled out in the second week of the war. Access to Twitter, Facebook, and Instagram was first slowed, then later blocked completely. In March, VKontakte swiftly complied with media regulator Roskomnadzor's demands to geoblock access to pages that criticised the war. Additionally, Roskomnadzor blocked access to thousands of websites and, on 21 March, officially declared Meta a terrorist organisation.

At the same time, Russian Foreign Ministry Spokeswoman Maria Zakharova called on Russian content creators to move away from Western

social media platforms. Figure 2 shows a sharp growth in users of Telegram, and a significant increase in active VKontakte users. However, YouTube's continued accessibility and increase in active user numbers shows the relative failure to move Russian content producers to the domestic RuTube platform.

"Save your content, transfer it to Russian platforms. And hurry up"
- M. Zakharova, April 2022

A gradual draining away of Russian influencers, combined with slow access speeds for those using VPNs resulted in a 72% reduction in Russian users accessing Instagram by July. The reduction for Facebook was even bigger in percentage terms. While TikTok restricted the ability of Russians to upload new content and see new material created outside Russia, the number of Russians accessing TikTok remained stable in the period January–July 2022. ■

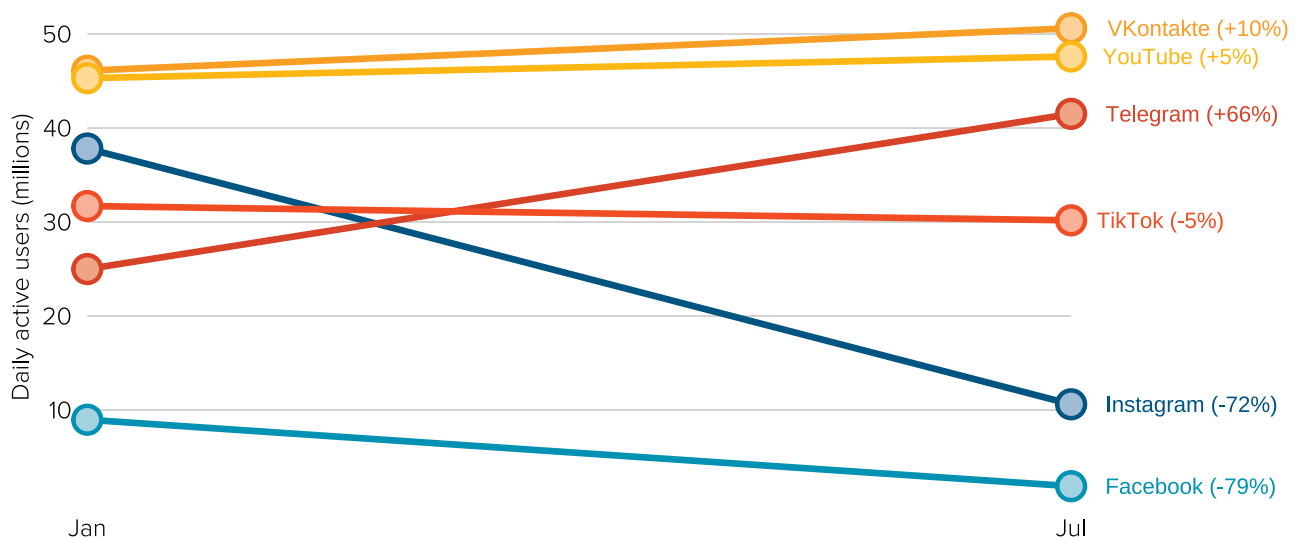


Figure 2: Changing Social Media Usage in Russia Jan-Jul 2022. Daily active users, in millions.

Twitter Targets pro-Kremlin Trolls

This section analyses the effect of Twitter policies implemented since the start of the war to reduce the spread of Russian disinformation. Measures include restricting algorithmic reach by ‘de-amplifying’ prominent propagandists (not recommending their content, removing them from search results and the ‘top’ section), and banning popular and egregious anonymous troll accounts.

Figure 3 ranks engagement statistics for the top thirty pro-Kremlin accounts mentioning NATO before 20 February 2022 (on the left) and from 1 April 2022 (on the right).

Many formerly prominent accounts have been deleted, including soulstray (Philip Maslovsky) and Vityzeva. These accounts were among the ‘more than a dozen’ accounts Twitter permanently suspended on 27 February for being linked to Russian intelligence. Others have stopped posting, for instance the notorious spacelordrock (Golos Mordora), who preempted a ban by moving to Telegram.

Another group of accounts—those in the middle of the right-hand column—remain

accessible but have lost popularity and reach. Most notably RT_russian, which has been sanctioned in Europe and demoted in popularity worldwide. The reduced reach of accounts in this group demonstrate the power of Twitter’s deamplification measures.

The most influential pro-Kremlin Twitter accounts have either been removed or have lost reach. Of the thirty accounts with reach prior to March 2022, only seven remained significant after April. Many others have seen their content de-amplified, such as the accounts of TV propagandist Vladimir Solovyov and the formerly independent news outlet Gazeta.ru. Whereas the previous ranking was dominated by troll accounts, the new ranking is made up of Russian media outlets, Russian officials, and media personalities.

This case study demonstrates how removing even a handful of influential accounts with established audiences has dramatically reduced the reach of pro-Kremlin propaganda. Fewer fake amplifiers meant fewer real users were exposed to the Kremlin’s messaging. ■

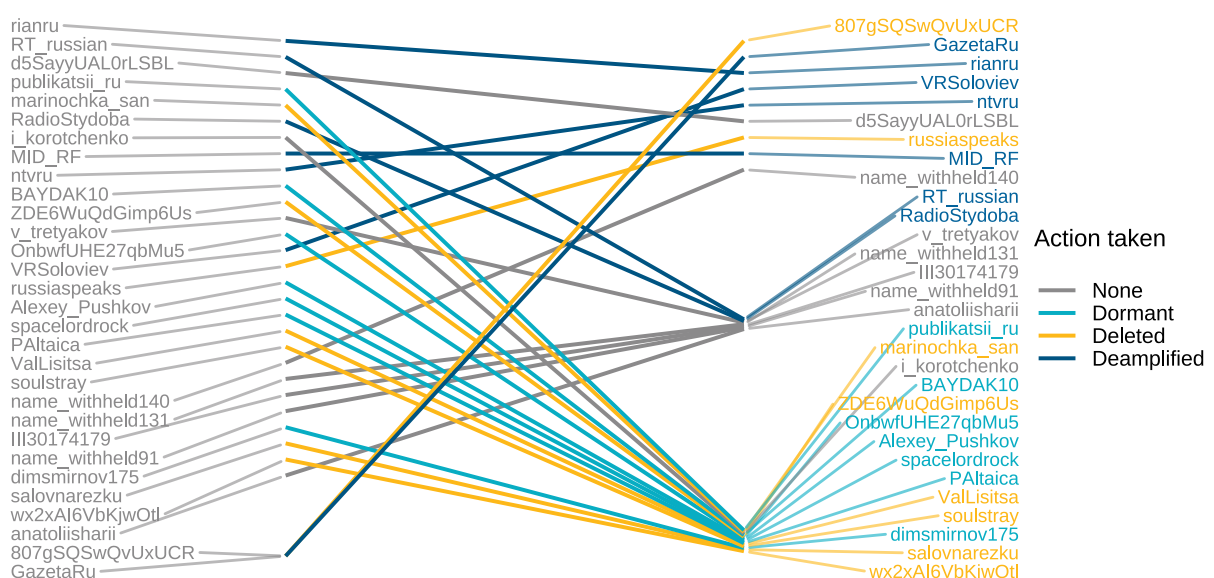


Figure 3: Pro-Kremlin Twitter accounts posting regularly about NATO before 20 February 2022 (on the left) and from 1 April 2022 (on the right). The graph shows whether the accounts were deleted, went dormant, or was deamplified.

Hostile messaging about NATO

This section analysis Russian-language tweets mentioning NATO in the period Feb-Aug 2022. Mentions of NATO and Ukraine peaked during the week of the invasion on 24 February. Since then, messaging volumes have decreased and stabilised at near pre-invasion levels.

In the first three months of the war, half of the 800 000 Russian-language posts on Twitter mentioning NATO also mentioned Ukraine. In the next three months mentions of NATO dropped to 200 000, of which only a third mentioned Ukraine.

English-language messaging about NATO in the Baltics has stabilised at around three times the pre-war volume; for Russian, the volume is modestly elevated, at a roughly 50% increase.

In the period Jun–Aug 2022, 22% of Russian-language tweets about NATO in the Baltics and Poland came from automated accounts, compared to 9% of English-language posts—an increase from 18% and 5.5% in the previous three months (see Figure 1). This increase in fake activity was driven by a drop in authentic

interest from regular social media users.

In the final edition of Robotrolling, we noted that in the three months leading up to Russia’s unprovoked invasion of Ukraine, anti-Kremlin Russian-language content was much more likely to go viral than was pro-Kremlin messaging. In the period Oct 2021–Feb 2022, hostile messaging from pro-Kremlin accounts increased steadily. At the same time, anti-Kremlin messaging spread further and faster than the narratives they responded to,

In March 2022, pro-Kremlin messaging about NATO on VKontakte had a hundred times the normal reach.

creating a novel dynamic.

Since Feb 2022, hostile pro-Kremlin messaging has decreased steadily—in absolute terms, in terms of average reach, and compared to anti-Kremlin messaging. Figure 4 shows that



Figure 4: Cumulative daily retweets for pro- and anti-Kremlin Tweets mentioning NATO.

anti-Kremlin Russian-language voices spiked in mid-May when Finland and Sweden declared they would join NATO. This spike in messaging was similar to that in response to Putin's ultimatum to NATO the previous Dec. On the other hand, the reach of pro-Kremlin accounts in May was less than half of what it was in Dec 2021.

Figure 4 shows a striking reduction in the Kremlin's control over the Russian-language debate on Twitter regarding NATO and Ukraine. In Oct and Nov 2021, 50% of retweets were of pro-Kremlin content; however, by Jan, that number had decreased to 20%. By May of 2022, that number had decreased even further to less than 7%. And through the summer of 2022 there were no successful attempts to mobilise based on anti-NATO messaging.

The reduced reach of pro-Kremlin accounts in Mar 2022 should not be attributed to NATO becoming less salient to the Kremlin. Figure 5 demonstrates that on the Russian (and Kremlin-controlled) social network VKontakte, the reach of the top pro-Kremlin posts mentioning NATO was more than 40 times higher in Mar 2022 than in Jan 2022. Simultaneously, the reach of pro-Kremlin messaging on Twitter dropped

sharply, with Mar levels less than half those of Jan. In rough terms, this places the relative reach of pro-Kremlin messaging on VKontakte in Mar 2022 at about a hundred times more than normal, compared to Twitter. While this difference shows the effectiveness of preventing misuse of Western platforms, part must also be attributed to the Kremlin increasingly focusing on domestic platforms.

Figure 5 shows the sum of daily views for posts mentioning NATO in pro-Kremlin groups on VKontakte. The timeseries of NATO mentions on VKontakte has a completely different shape to that seen on Twitter. In Oct 2021, the total view count for these posts was 0.6 million, which rose to 1.5 million in Dec and 2.2 million in Jan. Then, in mid-Feb, two weeks before the invasion, the total daily view counts increased dramatically beyond 4 million on Feb 24 alone. At the time of the Extraordinary NATO Summit, when a number of Western leaders made statements about what NATO might or might not do to support Ukraine, attention on VKontakte spiked higher even than on Feb 24. In Mar, messages from pro-Kremlin sources on VKontakte reached 50 million people, compared to 0.5 million in Oct 2021. ■

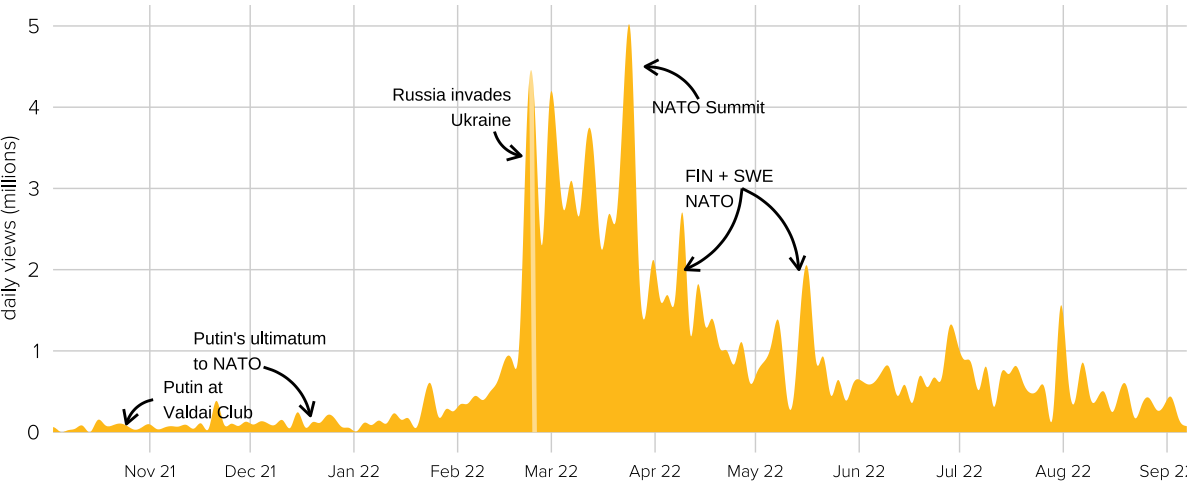


Figure 5: Cumulative daily views for pro-Kremlin posts mentioning NATO on VKontakte.

Sanctions-busting through Telegram

Telegram and YouTube are the two major platforms that avoided systematic blocking in both Europe and Russia. Supporters and opponents of the Russian war effort are both able to access and post content on the platform. Now these are the major current contested online spaces.

Telegram has emerged as especially significant in Russia. The platform experienced a massive growth in usage in Feb and Mar 2022. All channels dedicated to war-related content saw huge subscriber growth. Ramzan Kadyrov, head of the Chechen Republic, owns the largest Russian political channel on Telegram. On 23 Feb, the channel had 57 000 subscribers; in one week the number had risen to 370 000, and to 2.7 million by August.

Sanctions introduced by the EU and implemented by Twitter and Meta forced Russian state-media outlets to adapt. In early Mar 2022, Margarita Simonyan described how RT's attempts to spread propaganda to Western audiences now relied on guerilla tactics:

"We're not on Telegram on Western platforms. We exist in Russia, in Russian, but our foreign language resources are gone, on TikTok it's the same, that's over. We're now operating like partisans, so no one understands it's us. We

don't post under our own branding. We open a channel on YouTube, it gains millions of views in three days, and after three days their intelligence figures out it's ours and closes it. Then we open a new one."

We're now operating like partisans, so no one understands it's us
- M. Simonyan

Russian state media journalists help RT and Sputnik circumvent sanctions through cross-posting articles to their personal accounts. For example, Simonyan's social media accounts outperform many official channels.

In the days leading up to Russia's invasion on 24 Feb, few of Simonyan's tweets referred to RT's journalistic output. Her posts declined as the EU announced sanctions on 27 Feb and implemented them on 2–3 Mar. Propagandists may have hesitated, waiting for cues from above on how to cover the Russian war effort. Then, over the next two weeks, Simonyan's output exploded.

Sanctions immediately reduced RT's reach online. Twitter marked its posts as unavailable

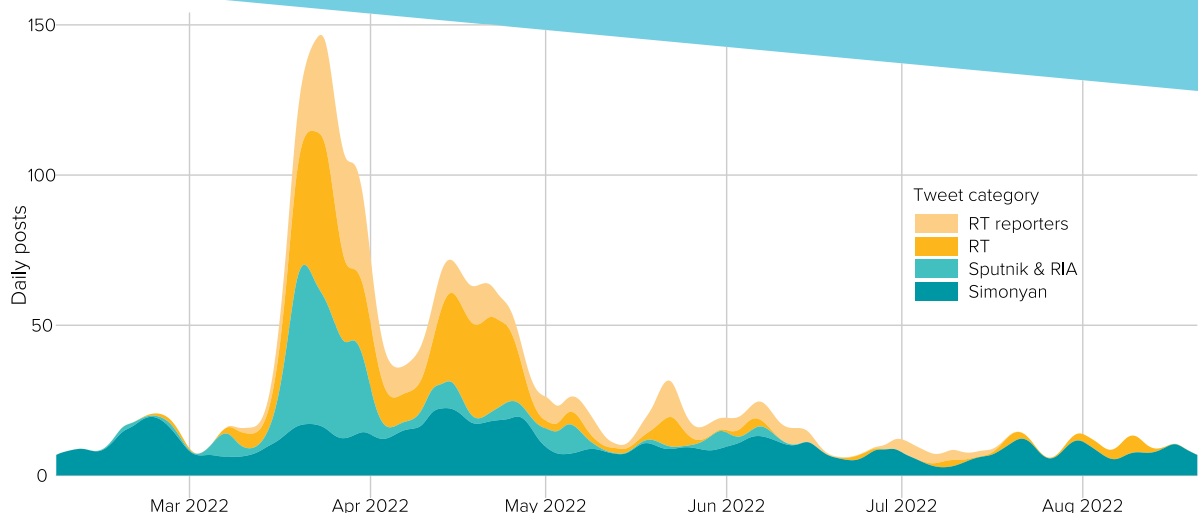



Figure 6: Categories of Tweets, posted by Margarita Simonyan



to social media users in a long list of Western states. Bots could retweet, but in Europe no one saw their posts. Instead of retweeting, Simonyan posted verbatim copies from RT. Her posts had the same text and images. The pictures often included RT's logo. Whenever Simonyan shared a message from RT on her Telegram channel, the content was automatically reposted to Twitter. Reposting via Telegram obscures the source, evades Twitter's black list, and is displayed without any disclaimer or geoblocking.

In the first days of Mar, Simonyan trialled the method for selected sources: RIA Novosti, Sputnik, and RT. In the period 15 Mar–11 Apr, 150 of Simonyan's tweets had texts and media from the sanctioned RT outlet. These messages were all visible in the EU.

The regular formatting hints that custom-made software automates simultaneous management of multiple social media accounts. The post structure is source name, line break, title text, and embedded video or image. Messages are often broken across multiple tweets, and contain the first sentences from the linked article. Although the posts are verbatim, they are not simple copy-paste. Embedded videos are hosted on Twitter, not through a link to external content.

Simonyan posts simultaneously on many Western platforms. We found examples of identical content posted to Twitter, Facebook, VKontakte, and Telegram. For example, if she posted to Telegram at 15.45, an identical post emerged on her VKontakte at 15.46, on her

Facebook at 15.47, and her Twitter at 15.50. We infer that the source of Simonyan's activity is Telegram, and that posts to other platforms are made automatically or semi-automatically. Thus the big change visualised in the graph reflects Simonyan moving most of her communications to Telegram, and automating amplification of this content to other platforms. This change was made roughly on 17 Mar and was highly active for about six weeks.

The move to Telegram coincided with a change in RT's war reporting. In March and April, most of the sources Simonyan shared on Telegram were from other RT reporters who maintained their own Telegram channels. These tactical changes could be explained by a move to personal branding of journalists to replace official state media accounts and channels.

Taken together, these tactics demonstrate how Russian state media have evaded both platform mitigation methods and the EU ban by adopting Telegram as the central node in their social media distribution mechanisms. While Twitter's measures effectively reduced the reach of pro-Kremlin voices, their ability to adapt exposes a flaw in the blacklisting approach adopted by the EU and social media companies: blocking a list of the main information sources is at best a temporary measure. The distributors of propaganda will create new accounts, domains, and channels through which they spread their content, whereas the list of sanctioned entities tends to remain static. ■

Prepared by Dr. Rolf Fredheim and Martha Stolze and published by

NATO STRATEGIC COMMUNICATIONS CENTRE OF EXCELLENCE

The NATO StratCom Centre of Excellence, based in Latvia,
is a Multinational, Cross-sector Organization which provides Comprehensive analyses,
Advice and Practical Support to the Alliance and Allied Nations.

@stratcomcoe, info@stratcomcoe.org

www.stratcomcoe.org