ROBOTROLLING

PREPARED AND PUBLISHED BY THE

**NATO STRATEGIC COMMUNICATIONS
CENTRE OF EXCELLENCE**

Executive Summary

In this edition of *Robotrolling*, we trace messaging about the build-up to Russia's invasion of Ukraine. Since the illegal annexation of Crimea, there has been an expectation that kinetic activity would be preceded by large-scale information activities.

We assess that the period from August 2021 to 20 February 2022 saw an increase in elite statements (and troop movements) unmatched by fake pro-Kremlin social media activity. This may reflect a top-down communication hierarchy, wherein lower echelons either received little guidance, or the guidance was to be silent. The volumes of automated activity were too low to offer any reliable signal of the looming invasion.

In February 2022, pro-Kremlin channels and accounts amplified the narrative that "genocide" of Russian-speakers in the Donbas justified intervention. Our analysis shows that this narrative's traction was attributable to statements by Putin personally, not online propaganda channels.

The comparative absence of pro-Kremlin activity on Twitter, combined with increased activity on VKontakte, suggests that Kremlin propagandists prioritised domestic audiences. On Twitter, the increase in Russian-language tweets about NATO was overwhelmingly driven by anti-Kremlin and pro-Ukrainian messaging.

The share of automated messages on Twitter and VK about the Baltic states, Poland and NATO dropped, amidst greatly increased overall traffic. Bots accounted for 30% of Russian-language and 12% of English-language messages on Twitter; and for 15% of messages about the region on VK. In February 2022, the number of English tweets was seven times higher than Russian tweets. ◾

The Big Picture

In this edition of *Robotrolling*, we investigate artificial social media activity prior to Russia's invasion of Ukraine on 24 February. In the months before the invasion, the expert community was divided. Military analysts said war was probable due to the large concentration of Russian forces on the Ukrainian border. Social science researchers tended to downplay the likelihood of a full-scale invasion, in part because there appeared to be an absence of messaging aimed at preparing domestic audiences for the prospect of war.

Our long-standing effort to document fake messaging about NATO's presence in the east of Europe offers raw material for a closer look. The Alliance plays a prominent role in Russian propaganda narratives. Any obvious signals would be clearly detectable through this lens. Our analysis focused on the activities of automated accounts (bots) and coordinated anonymous human accounts. In this—our eighteenth—issue, we compared the period from 1 August 2021 to 20 February 2022 with the previous reporting period, 1 February to 31 July 2021.

Overall, the messaging volume increased greatly. However, this increase was driven by English-language conversations. The number of English tweets (excluding retweets) about the NATO presence quadrupled from 11 300 during the last reporting period to 45 000. Meanwhile, the number of Russian-language tweets almost doubled from 7 200 to 12 700. On VKontakte, the total messaging volume increased by 50% from 58 000 in the previous reporting period to 86 000.

The number of English-language bots almost tripled to more than 2 000. Nonetheless, a four-fold increase in the total number of users meant that the bot-share declined from 11% to 9%. The bot-share among accounts tweeting predominantly in Russian also decreased.

We observe that the percentage of automated messaging on Twitter dropped due to greatly increased overall traffic, from 13% to 12% for English-language and from 36% to 30% for Russian-language messaging.

The change for automated activity on VKontakte is less pronounced. Although the number of bot accounts messaging about NATO, the Baltics and Poland grew by 5%, the percentage of bot users decreased from 24% to 21%. ◾

Twitter **RU** — 14%

Twitter **EN** — 9%

VK **RU** — 21%
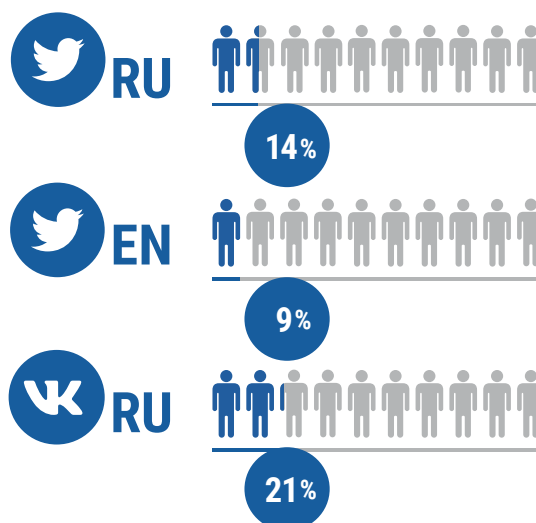
Themes: Mobilising for War

In the first months of 2014, when Russia executed a hybrid campaign culminating in the annexation of Crimea, social media channels were flooded with automatically generated content. This content served as a smoke screen, obscuring the ability of citizens to find credible information online. In 2022, the invasion could have been—and was—anticipated, based on statements by President Vladimir Putin and other elites, along with the military build-up on Ukraine's borders. However, we also assess that Twitter-monitoring in this case added little.

Since the annexation of Crimea, there has been an expectation that kinetic activity would be accompanied, even preceded, by large-scale information activities. To identify whether similar lessons can be drawn from February 2022, we investigated the volume of messaging and the role of bots, the Kremlin's ability to dominate the online conversation, the key narratives, and the Kremlin's overall communication strategy.

Messaging volumes and the role of bots.
Overall, pro-Kremlin social media activity about NATO followed the patterns observed over the previous five years of reporting. The volume of posts is high, but less than one would expect based on English-language volumes; the presence of bots is significant, but lower than usual in percentage terms.

The "threat" of NATO expansion has been constantly activated since we started monitoring in 2017. From late December, the spikes in attention largely coincided with elite statements. Putin himself spoke frequently and dramatically about the threat posed by an aggressive and expansionist NATO.

The largest spikes in bot activity about NATO and Ukraine in January 2022 related to statements by Western leaders signalling that NATO would not fight in Ukraine. Other messaging drew attention to NATO's military build-up in the Baltic states and Poland. NATO's eastern enlargement was described as illegitimate and a breach of Russia's security interests. This theme was manifest both in Russian- and English-language bot activity.

Bots promoted the narrative that the West—not Russia—was escalating tensions in Europe, inter alia pushing tweets that Russia had warned the West regarding the mistake of admitting Poland and the Baltic States into NATO. In this period, bots amplified incendiary elite statements, but the volumes were too low to in themselves signal imminent invasion, even when analysed retrospectively.

Did the Kremlin dominate the Russian-language discussion?
Retweet statistics reveal that the increase in Russian-language tweets about NATO was not driven by pro-Kremlin messaging, but by negative reactions to it. Figure 4 demonstrates that the messaging about NATO aggression failed to gain traction; instead, the vast majority of heavily retweeted messages were posted (and amplified) by Ukrainians and Russian-speakers opposed to any intervention in Ukraine. Thus, content that mattered was overwhelmingly anti-Kremlin and pro-Ukrainian.

During the period from August 2021 to 20 February 2022, 80% of the most popular tweets pushed messages critical of the Kremlin. Pro-Kremlin Russian-language tweets had a comparatively much lower reach in February 2022 than they did during November 2021, when they drew attention to migrants crossing from Belarus into the EU. Anti-Kremlin voices were even more dominant after 24 February, due to sanctions on Russian state outlets in Europe, as well as Twitter restricting the reach of Russian propaganda.
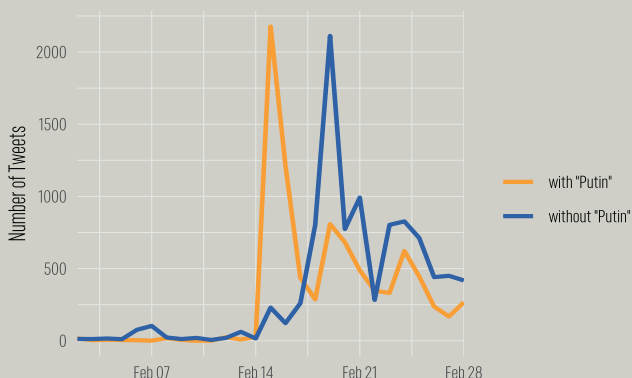

Figure 2: Russian-language Tweets mentioning "Donbas" and "genocide". Most tweets on February 15 also mentioned Putin, showing this narrative was elite-led.
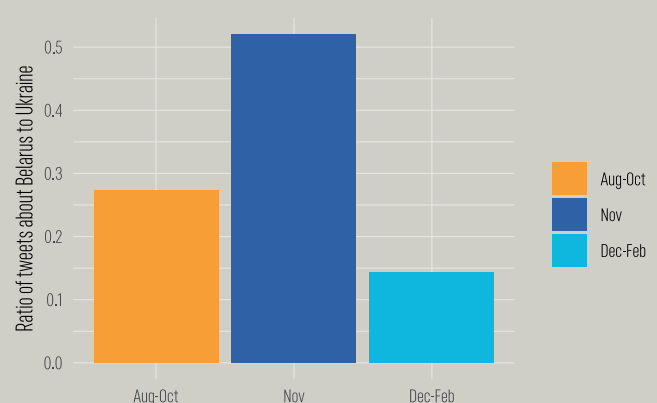

Figure 3: The ratio of tweets mentioning NATO and Belarus to tweets mentioning NATO and Ukraine.

Higher activity volumes on VKontakte than on Twitter suggest a shift in focus towards messaging **directed against domestic audiences.** In one reading, the low level of Russian-language activity on Twitter reflected a concealment strategy to deceive international audiences, whereas the domestic audience was prepared via a comparative barrage of messaging about NATO aggression.

Three factors may help explain the lack of pro-Kremlin activity on Twitter:

- The Kremlin's messaging did not ring true. It was easier to ridicule and mock it than to mobilise based on it.
- Twitter successfully reduced the reach of pro-Kremlin accounts.
- The Kremlin propagandists prioritised targeting domestic audiences through television and VKontakte, rather than Twitter's more international user base.

However, the character of the messaging both on VKontakte and Twitter was cautious. It was unusually reflective of a top-down messaging culture, whereby pro-Kremlin social media messaging took cues from Kremlin elites and state television. There are no notable examples of NATO-related content originating on social media and bubbling up to state media. We speculate that the handlers of propaganda accounts intuited (or had been warned) that now was a risky time to take initiative or show creativity.

This period contrasts with November where—from Moscow's perspective—the migrant crisis in the Baltics and Poland had obvious propaganda value. At this time, inauthentic activity on Twitter revolved around the challenge posed by illegal migration to EU and NATO solidarity. Pro-Kremlin bots on VKontakte suggested that "European fascists" had staged chaos at the Belarusian border, and that Poland and Lithuania had exploited the crisis to increase their military presence.

Key narrative: genocide in the Donbas. Some post-hoc analyses of social media data have found unusual activity in the ten days prior to the invasion, for instance, when tracking the Twitter mentions of Donbas together with the term "genocide". The notion that Russian intervention was required to protect ethnic Russians against the genocidal actions of Ukrainian Nazis, was activated in the period immediately preceding the Russian invasion (as shown by Figure 2) .

Figure 2 also demonstrates that virtually all the social media mentions of Donbas and genocide also mentioned Putin. At a 15 February press conference following talks with Chancellor Olaf Scholz, Putin described the situation in the Donbas as "genocide". The activation of the term "genocide" should be attributed to Putin, not to the social media accounts that amplified his statements and in the following days elaborated on the rhetoric.

The official Russian communication about the war was characterised by panic and improvisation, not careful planning. In the first days of the war, some military communities posted about tanks marked with the symbol Z. Within two days, it became the main mobilising idea of the invasion. Anecdotal reports claimed that Kremlin propagandists despaired at the choice, as the symbol had no legacy or prior meaning.

The coordination of the propaganda messaging has happened via Telegram. The notorious Telegram channel *War On Fakes* was created in the hours after the invasion; the group coordinating activity for a new St. Petersburg troll farm, *Cyber Front Z*, was only created on 11 March.

The campaign to censor the Russian internet was fitfully rolled out in the second week of the war, after reports of Russian failures had already been widely shared. Access to Twitter, Facebook, and Instagram was first throttled, and later blocked completely. In March, thousands of VKontakte groups were geo-blocked within the Russian Federation. The Russian media regulator Roskomnadzor blocked access to hundreds of websites. Taken together, there is little to suggest the communicators were involved in the invasion plan. ■
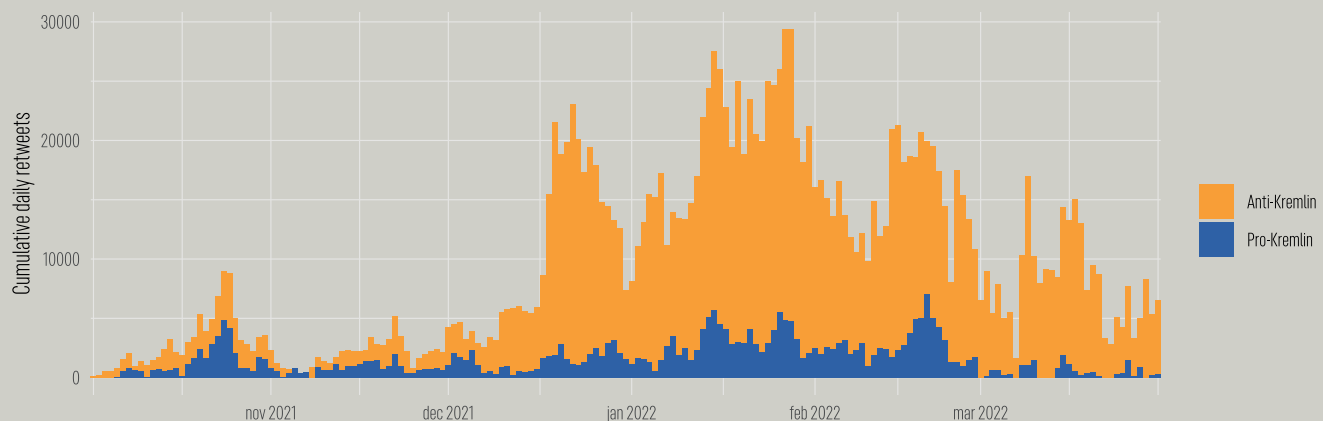

Figure 4: Daily retweets of top Russian-language tweets mentioning NATO.

Country Overview

On Twitter, Lithuania and Estonia attracted the largest number of Russian-language tweets, whereas on VKontakte, Poland and Lithuania received the bulk of attention. Poland also attracted the highest share of English-language bot activity.

Inauthentic messaging on Twitter spiked in mid-November, when Poland and the Baltic states considered triggering NATO Article 4 in response to mass migrant crossings from Belarus. The high inauthentic Russian-language messaging on 13 January, in turn, revolved around the Baltic States calling for increased NATO presence in the regioon. English-language inauthentic conversations were triggered by Germany banning the export of German-origin weapons to Ukraine in January, and by the deployment of additional US troops to Europe in February. VKontakte bot messaging was the highest on 2 February, when bots questioned Ukraine's chances of joining the EU and NATO.

Estonia

On 4 November, bots created more than 50% of Russian-language tweets mentioning NATO and Estonia, in response to Secretary General Jens Stoltenberg suggesting that Russia did not pose an imminent threat to the Baltic states. On 17 January, Russian-language bot activity on VKontakte and Twitter revolved around Estonia's readiness to accommodate 5000 soldiers of the NATO Rapid Reaction Force. The highest amount of English-language inauthentic activity took place in reaction to Germany objecting to Estonia supplying German-origin weapons to Ukraine.

Latvia

The most significant inauthentic activity revolved around the Foreign Minister Edgars Rinkēvičs warning of "incidents" between Russia and NATO troops during the upcoming "Zapad" military exercises. The majority of posts about this statement came from bot accounts. Similarly, bot messaging spiked on 29 November, when Prime Minister Krišjānis Kariņš called on the EU and NATO to signal clear consequences to Russia if it escalated tensions with Ukraine. On VKontakte, bot messaging was highest concerning the opening of a NATO military airfield in Latvia on 24 October.

Lithuania

On 28 November, around 50% of Russian- and English-language messages on Twitter mentioning NATO and Lithuania were spread by bots, prompted by Stoltenberg's and von der Leyen's joint visit to the Baltics. The share of Lithuania-specific Russian-language inauthentic activity on VKontakte was highest on 15 November, when over 60% of VKontakte activity had bot-origin, and claimed that Poland and Lithuania had taken advantage of the border crisis to raise their military readiness. The highest number of bot-origin messages on VKontakte were observed on 24 January, with accounts discussing strengthened NATO military posture.

Poland

Russian-language bot activity peaked on 14 November, when bots on VKontakte blamed "European fascists" and "deceitful Poland" for the chaos at the Belarusian border. On 5 February, 40% of Russian- and 50% of English-language tweets were of bot origin, with Russian conversations discussing the arrival of the UK military in Poland, while English-language bots mainly shared the headline "Ukraine links arms with Turkey, Poland and UK as NATO membership remains distant". ◼
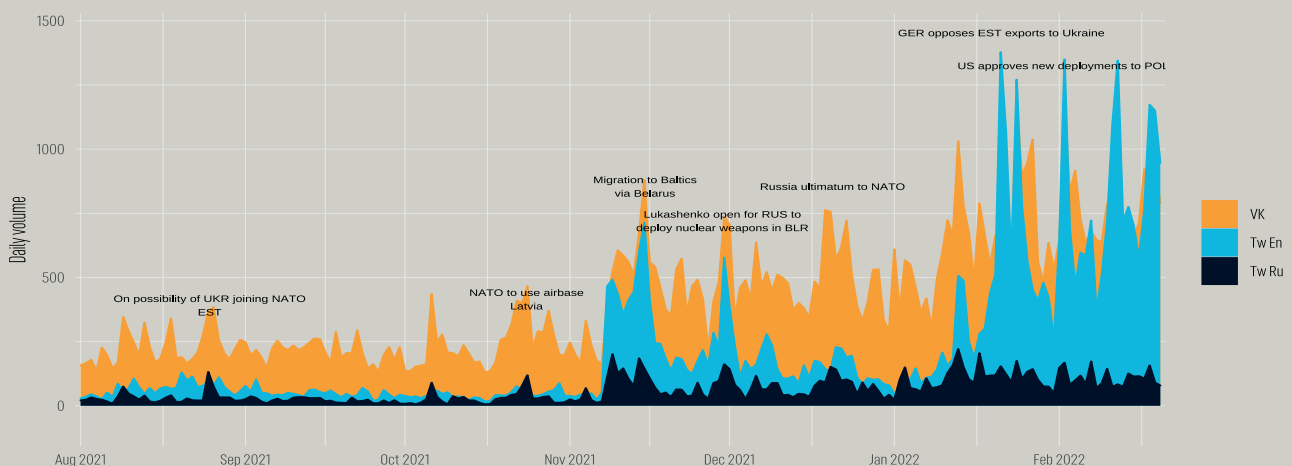

Figure 5: Timeline of VKontakte and Twitter mentions of NATO in the Baltics and Poland

Themes

Three themes of inauthentic activity stand out. While in August and September 2021, the Zapad military exercises still attracted considerable inauthentic messaging concerning NATO, the Baltic states and Poland, the attention turned to the situation of migrants at the Belarusian border in November/December. Moreover, since January 2022, the spikes in bot activity have revolved around Russia-Ukraine tensions. Bots on VKontakte increasingly amplify Russian elite statements that portray NATO as an aggressive military block that seeks conflict with Russia.

Preceding the Zapad military exercises, bot messaging comprised about 60% of Russian- and 40% of English-language tweets concerning Latvia and NATO. Some bots spread the false claim that the Russian military had clashed with NATO forces in Latvia. On 7 September, bots on VKontakte amplified the misleading message that the US had called on NATO countries to prepare for war with Russia; while others emphasised that Kaliningrad is "Russia's knife at the throat of Europe", bearing strategic importance in Russia's efforts to contain NATO in the Baltic Sea region.

Messaging about migrants at the Belarussian border peaked in November 2021, when 22% of the English- and 40% of the Russian-language tweets on this topic originated from automated accounts. Estonia and Latvia received the highest share of Russian-language bot activity; while the share of English-language bot activity was highest for Latvia and Poland. On Twitter, the share of migration-specific bot messaging reached over 60% for the Baltic states and 50% for Poland in November. The inauthentic activity primarily reacted to calls for greater EU and NATO support in response to migrant border crossings, as well demands for US leadership and retaliation for Putin's actions amidst the migrant crisis.

Inauthentic messaging on VKontakte about Belarus was more diverse, with pro-Kremlin bots mainly suggesting that "European fascists," including Poland, the US and NATO, had staged chaos at the Belarusian border, while Poland and Lithuania had taken advantage of the crisis by significantly increasing their military presence on the border with Belarus. Other bots emphasised that Latvia, Lithuania and Poland considered triggering Article 4 in reaction to the migration crisis, while some bots claimed that the Baltic states turning to NATO for help might result in a full-scale war. Fake accounts claimed the crisis at the EU's eastern border was "growing deeper and wider", and that the activity at the Polish-Belarusian border was becoming "more and more reminiscent of preparations for war".

From mid-January, Russian-language inauthentic messaging on Twitter highlighted the Baltic states asking NATO to increase its military presence on their territory; while spikes in English-language bot activity in February were mainly triggered by the US and UK deploying additional troops to the east of Europe. On VKontakte, bots increasingly amplified Russian elite statements which portrayed NATO as an aggressive, expansionist block, arguing that NATO's increased military presence at Russia's borders demands a reaction. In the weeks leading up to Russia's invasion of Ukraine, bots on VKontakte pushed the narrative that NATO's enlargement had been illegitimate and a breach of Russia's security interests. ■

Authors: Rolf Fredheim and Martha Stolze

NATO STRATEGIC COMMUNICATIONS CENTRE OF EXCELLENCE