

ISBN 978-9934-619-43-4

VIRTUAL
MANIPULATION

2025
ISSUE

1



Virtual Manipulation Brief 2025

From War and Fear
to Confusion and Uncertainty

PREPARED AND PUBLISHED BY THE

**NATO STRATEGIC COMMUNICATIONS
CENTRE OF EXCELLENCE**



ISBN: 978-9934-619-43-4

Authors: Dr. Gundars Bergmanis-Korāts, Raitis Ralfs Vecmanis, Marija Isupova

Project manager: Dr. Gundars Bergmanis-Korāts

Contributors: Guna Šnore, Pāivi Tampere

Content editing: Egil Fredheim

Design: Inga Ropša

About the Virtual Manipulation Brief

The Virtual Manipulation Brief, part of NATO StratCom COE's Robotrolling series since 2017, examines automated messaging about NATO in the Baltics and Poland, along with Russia's broader narrative on the Alliance and online campaigns targeting Ukraine.

Building on the June 2024 report, this edition focuses on the growing role of artificial intelligence (AI) in Russian foreign information manipulation (FIMI). AI-generated content and coordinated social media networks are enhancing the scale and reach of manipulation campaigns targeting NATO. Platforms like Telegram, X, and VKontakte (VK) are used to spread hostile narratives, often portraying NATO as weak or aggressive. Financial incentives from platforms like Telegram further fuel manipulation online.

This year, in contrast to the previous Virtual Manipulation Brief, we have expanded the scope to include 10 platforms: X, Facebook, Instagram, BlueSky, YouTube, Telegram, VK, OK (Odnoklassniki), Threads, and TikTok. Between June 2024 and May 2025, over 11 million posts and comments were collected across 10 key topics (around five times more data have been collected than in the previous time period, requiring a complete redesign of our long-established baseline of using two platforms in the earlier issues and three in the previous one).

The 2025 report offers insights on recent developments and strategic recommendations for addressing both ongoing and emerging threats, based on our internal expertise.

Riga, May 2025

NATO STRATCOM COE

11b Kalnciema iela

Riga LV1048, Latvia

www.stratcomcoe.org

Facebook: [stratcomcoe](https://www.facebook.com/stratcomcoe)

Twitter: [@stratcomcoe](https://twitter.com/stratcomcoe)

This publication does not represent the opinions or policies of NATO or NATO StratCom COE.

© All rights reserved by the NATO StratCom COE. Reports may not be copied, reproduced, distributed or publicly displayed without reference to the NATO StratCom COE. The views expressed here do not represent the views of NATO.



CONTENTS

Executive Summary	4
Overview	5
What Matters and Where?	5
Hostile messaging about NATO	6
Profile taxonomy	11
Baltic perspective	15
Impact of US elections on overall messaging and peace-talks	16
The Asymmetric Power of X: Exploited by Pro-Russian Actors	17
NATO's Partnerships From the "Cold War Rhetoric" Perspective	18
AI perspective	21
Conclusions	26
Recommendations	27
Appendix	28
Endnotes	29

Executive Summary

The “Virtual Manipulation Brief 2025” analyses the shifting landscape of digital influence operations targeting NATO, Ukraine, the European Union, and the United States. The current issue marks a complete redesign of our data collection and processing pipeline, expanding from only two social media platforms to ten. This significant change necessitated a fundamental restructuring of the baseline (our reference data and metrics for measuring manipulation), which will guide our future research in this area. An estimated ~7.9 % of all the interactions we tracked show statistics-based signs of coordination. Kremlin-aligned messaging bursts were roughly twice as frequent as their pro-Western counterparts, and about three times as frequent for the posts that appeared on more than one platform. This implies the broader reach and tighter synchronisation of Kremlin-orchestrated operations.

Moreover, platform importance depends on the monitored topic and language. One should not rely on the post and comment volume alone. While X had the highest number of posts, YouTube and Telegram showed greater engagement and reach in our case. This demonstrates that the platform with the most posts isn't necessarily the leader in terms of engagement or reach.

A key observation is the opportunistic nature of Russian actors, who are actively exploiting the policies of the new US administration as means to intensify their targeting of Ukraine, the EU, and NATO. In addition to Russian activities, we identified and analysed strategic narratives originating from China

that are specifically aimed at NATO, broadening the scope of concern beyond a single primary actor. China's messaging focuses on showcasing its strength and portraying the US as weak, corrupt and aggressive. At the same time Russian messaging was more emotional and defamatory.

Artificial intelligence is increasingly prominent in the digital space. Do you remember the time when early deepfakes were a primary concern? Today, AI can generate diverse content across multiple languages, create influencers from static images, etc. This capability enables the exploitation of political events and crises through the rapid creation of misleading videos, audio, images, and text. Additionally, initial practical issues of applied AI are being gradually mitigated, making the technology more usable. We estimate that approximately 15% of our current computer code for data collection and analysis was generated by AI. This percentage is expected to rise significantly in the coming years due to AI's capacity to accelerate capability building: the increased speed in developing tools and services marks the beginning of a new era where not only information dissemination accelerates, but also software development and service delivery. Moreover, the advanced reasoning abilities of new large language models, coupled with emerging AI agent frameworks, are enhancing the autonomous operation of AI systems. We expect it to be a primary focus of the next iterations of Virtual Manipulation Brief.

Overview

What Matters and Where?

Within the scope of this research we processed roughly eleven million social media posts and comments. X is the largest source (nearly half), followed by Telegram (around a quarter) and YouTube (about twelve percent, primarily longer narrative-building videos). Russian platforms VK and OK contribute seven percent, reaching audiences within Russia and its diaspora. The remaining five percent comes from Facebook, Instagram, TikTok, Bluesky, Threads, and similar services. English dominates X, Bluesky, and partly YouTube, while Russian prevails on Telegram, VK, and OK. Other platforms show a mix of European and Asian languages. Two main dissemination paths exist: X for broad reach (subject to platform rules) and Telegram/Russian sites for a loyal, harder-to-reach core audience.

But what matters where? Our ambition has drastically increased with this report and thus also the data volumes. With this iteration our aim is to observe and estimate where and how we have to look in the digital space.

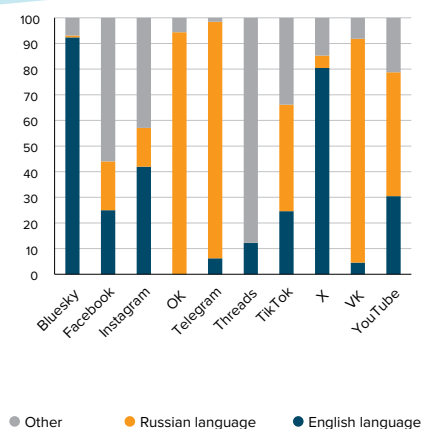
Platform use by influence actors is topic-dependent (Figure 1.A-D), with X leading discussions on broad geopolitical issues, while

The Data

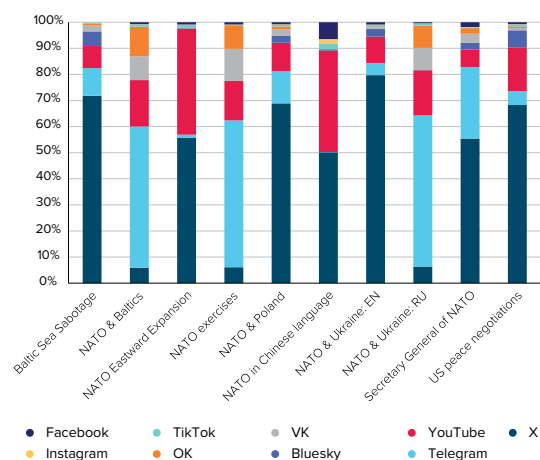
For this report, we collected & analysed:

- 11 million social media posts (~5 times more data than in 2024 issue)
- From 10 social media platforms
- Posted between June 2024 and May 2025
- 10 key topics

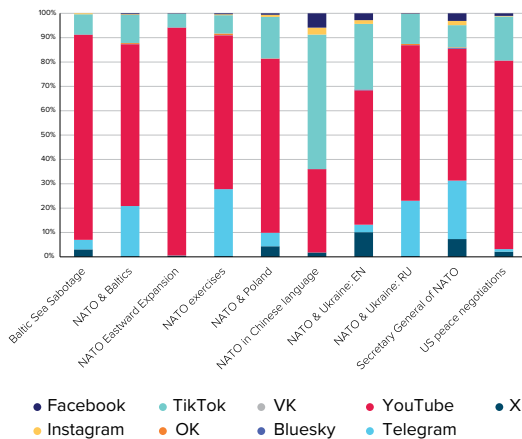
Telegram and Russian platforms VK and OK are used for specific regional and exercise-related content; although X, Telegram, and YouTube show the highest volume, YouTube, Telegram, and TikTok demonstrate the greatest engagement and reach. Due to high user engagement on certain platforms, practitioners should pay close attention to the comments section to gain deeper insights and look for potential violations and hostility.



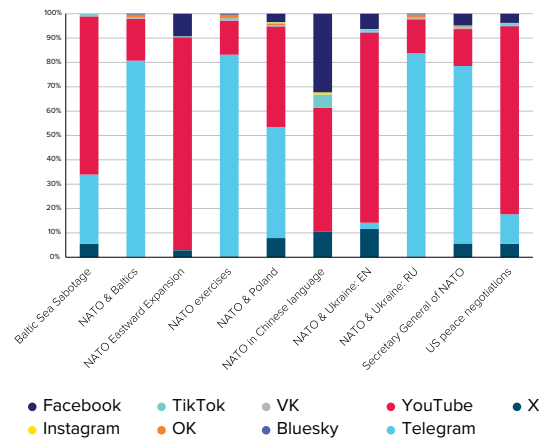
A: Relative language distribution across all platforms for the topics of interest



B: The relative volume of platforms in tracked topics



C: Relative engagement estimated across topics and social media platforms



D: Potential reach estimated across topics and social media platforms

FIGURE 1. Different metrics reveal distinct platform importance. Figures show relative distribution of languages, volume, engagement and potential reach for posts.

Hostile messaging about NATO

Let's do a deep dive into hostile messaging towards NATO. First, we analysed Kremlin-aligned posts by their post type (post, reshare, extended reshare, comments and reply-comments). As intuitively expected, the Kremlin-aligned information space heavily relies on amplification. A significant majority of high-alignment items (58.9%, totalling 74,772) are simple reposts, with an additional 3.0% (3,798) being extended reposts. X is the primary platform for these, hosting 87.8% of all reposts and 74.0% of extended reposts, while Telegram accounts for most of the remaining activity. Original content is less frequent, comprising only 18.7% (23,735 items), and is mainly found on Russian domestic platforms such as VK (42.9%), OK (14.6%), and Telegram (32.7%). Comment activity is also notable, with Telegram driving most comments (41.7%) and reply-comments (66.0%), followed by YouTube (32.8% and 12.2% respectively). Engagement data indicates a dual approach: original posts, though fewer in number, achieve greater average reach (approximately

12,100 views) and interaction rates (around 141 per post), while the much more prevalent X reposts achieve scale but garner less audience engagement.

Before examining platform-specific engagement patterns, it is essential to establish the broader communicative landscape observed between May 2024 and April 2025. During this period, overall posting volumes surged in response to salient geopolitical events, with English- and Russian-language activity rising and receding in near-parallel waves. Figure 2 shows the relationship between the total volume of posts in English and Russian languages vs. the key political events over time. Figure 3 displays a timeline comparing highly aligned Kremlin-supporting (red) and West-leaning (blue) messaging.

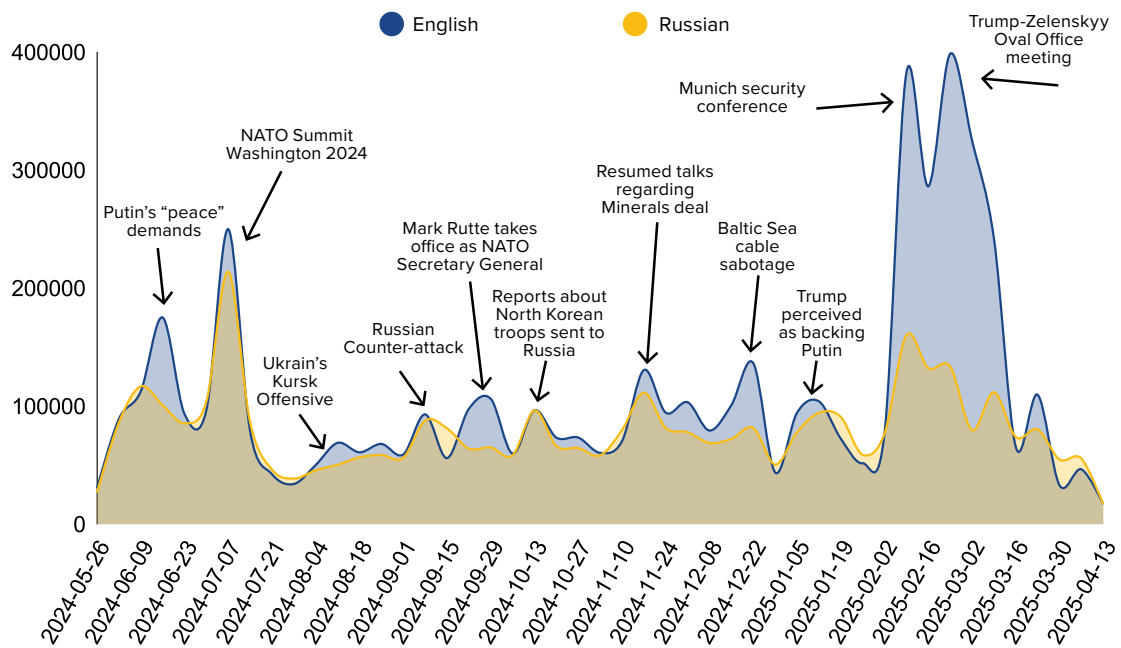


FIGURE 2. A comparison of English vs. Russian volumes across all platforms and all topics of interest

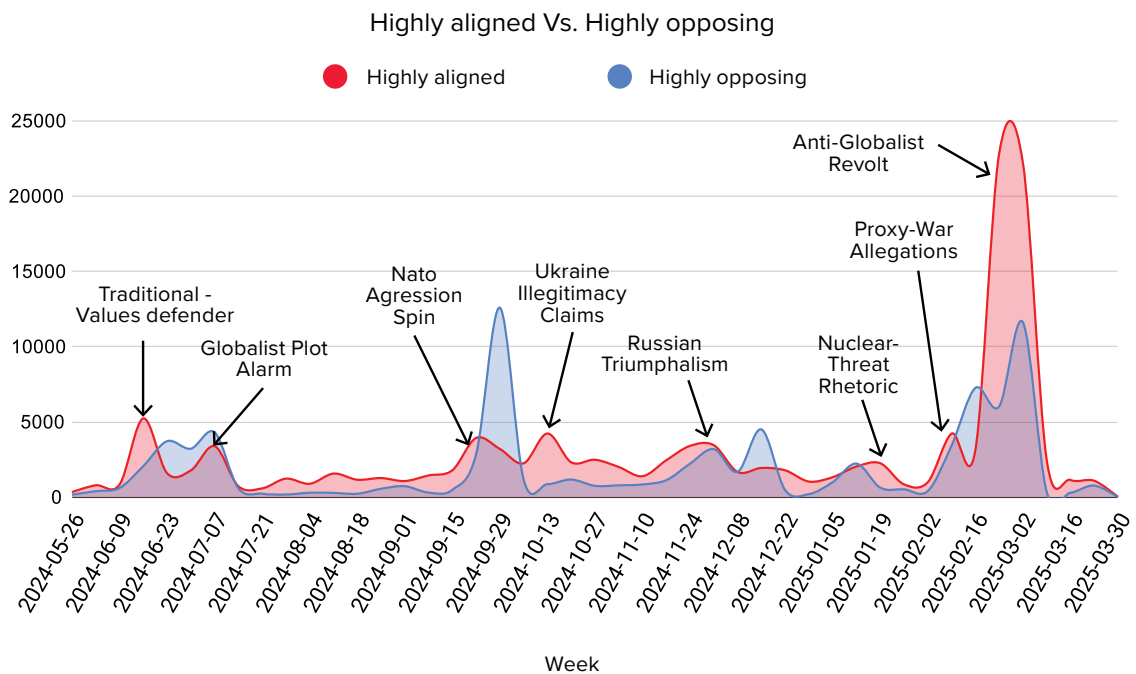


FIGURE 3. Total volumes of strongly kremlin-aligned messaging from all platforms/topics. Big pro-kremlin messaging peak coincides with Zelenskyy-Trump Oval Office meeting

The table below outlines peaks in pro-Kremlin social media messaging, each marked by intensified dissemination of aggressive, conspiratorial, and nationalist narratives:

Traditional-Values Defender June 2024	Discussions glorify Russia as a defender of traditional values and sovereignty, while portraying the US as corrupt, degenerate, and controlled by globalist interests. The West is blamed for the Ukraine conflict, accused of election fraud and warmongering. Ukraine's drone attacks are cited as provocations.
Globalist Plot Alarm July 2024	The West, NATO and Ukraine are cast as deceitful aggressors whose eastward expansion, "globalist" plots and even engineered pandemics are said to seek Russia's "strategic defeat" by stoking internal unrest and launching future F-16 strikes on Crimea; Western leaders are mocked as Russophobic puppets, Ukraine is branded a fascist proxy, and NATO an anti-Russian tool, yet the writers claim Russia will crush any assault once it addresses its own domestic ills. The rhetoric is uniformly alarmist, conspiratorial and fiercely nationalist, presenting Moscow as civilisation's last bulwark against an existential Western onslaught.
NATO Aggression Spin September 2024	Discussions with a narrative of Western weakness, deceit, and declining resolve in Ukraine, portraying NATO as aggressive and Russophobic. Zelenskyy is depicted as self-serving and desperate, more concerned about personal power than peace. The West is accused of provoking new fronts (e.g. Moldova, Baltics), while Russia is criticised for being too passive. Other posts celebrate Russian military successes and mock Ukraine's leadership, accusing the West of using Ukrainian lives for its own interests. The tone is hostile, alarmist, and triumphalist, promoting a more forceful Russian response and framing NATO as a global aggressor.
Ukraine Illegitimacy Claims October 2024	Posts in this peak advance a strongly pro-Russian, anti-Ukrainian and anti-Western narrative, portraying Ukraine and Zelenskyy as corrupt, illegitimate, and militarily desperate. The conflict is framed as a justified Russian response to NATO aggression, the 2014 "coup," and the persecution of ethnic Russians. Zelenskyy is mocked for diplomatic failures and outlandish claims, while NATO is depicted as manipulative, hesitant, and ultimately doomed. The tone is derisive, triumphalist, and conspiratorial, emphasising Russia's military strength, Ukraine's collapse, and the West's moral and strategic failure.
Russian Triumphalism November 2024	These posts promote a strongly pro-Russian, anti-Western narrative, portraying Ukraine as a corrupt, weak puppet state manipulated by NATO, and the US. Russia is framed as acting defensively against Western aggression, with recent missile strikes (like "Oreshnik") highlighted to showcase Russian superiority and Western vulnerability. NATO is accused of plotting to partition Ukraine and escalate the conflict under false pretences. The tone is aggressive, triumphalist, and conspiratorial, mocking Ukraine's leadership and dismissing Western support as fractured and ineffective, while threatening NATO states and celebrating Russian military power.

Nuclear-Threat Rhetoric	These posts adopt a highly aggressive, pro-Russian stance, portraying NATO as a hostile, corrupt force led by morally bankrupt elites. Rutte's comment about learning Russian is used to accuse the West of fascism and dehumanisation, while homophobic slurs and extreme rhetoric are directed at Western leaders. Ukraine is blamed for the war and accused of atrocities, while Russia is praised for its military strength and justified in escalating, including through nuclear threats. The tone is openly hostile, conspiratorial, and contemptuous, advocating for harsher Russian action and total Ukrainian defeat.
January 2025	

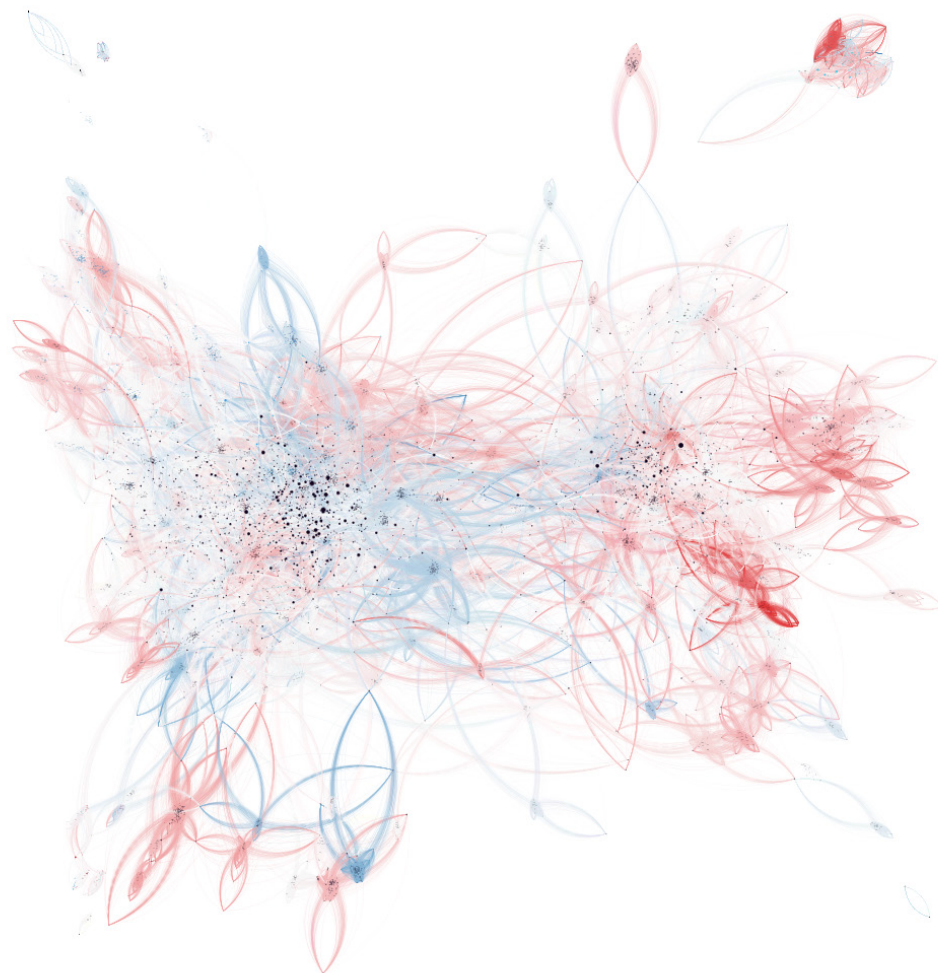
Proxy-War Allegations	These social media posts present a narrative that the Ukraine war is a US/Western "proxy war" against Russia, driven by "globalist" and "neocon" agendas and NATO expansionism. Key talking points include: the West's actions (sanctions, support for Ukraine) have backfired, leading to Russia's victory and the West's decline; Ukraine's NATO membership is a dangerous provocation towards WWII; Zelenskyy is a "psychopath", "junkie", or "puppet"; Western leaders like Starmer and Rutte are incompetent "globalists" pushing for war; Trump is the positive force working with Putin to end the conflict. The tone is highly aggressive, conspiratorial, and triumphant regarding perceived Western failures, using inflammatory language against opponents while praising Trump and predicting Russia's success.
Beginning of February 2025	

Anti-Globalist Revolt	The posts in this peak promote a conspiratorial, anti-globalist narrative, portraying the Ukraine war as a Western/NATO proxy conflict driven by corrupt elites seeking profit and control. NATO and the EU are depicted as tools of a "globalist tyranny" suppressing democracy, sovereignty, and identity through mass immigration, propaganda, and digital control. Leaders like Rutte are vilified, while populists like Trump and Fico are cast as victims of establishment repression. The tone is alarmist, accusatory, and revolutionary, with calls for resistance against a perceived elite-driven agenda undermining national freedoms and orchestrating global conflict.
End of February 2025	

The themes overall evolve from glorifying Russia as a moral stronghold and casting Ukraine as a failed Western puppet, to portraying NATO and the EU as globalist tools orchestrating war, societal collapse, and digital oppression. Across all peaks, the messaging reflects hostility toward Western leaders, triumphalism about Russia's strength, and calls for resistance against an alleged elite-driven global agenda.

In the whole time period we also identified 1924 pro-Kremlin vs. 902 pro-Western temporal and semantic bursts. From those 1128 pro-Kremlin bursts were cross-platform in contrast to only 350 pro-Western cross-platform bursts.

Pro-Russian messages frequently criticise NATO's role in escalating the war in Ukraine, portraying it as a dangerous and unnecessary conflict. There are claims that Western powers, particularly the UK and the US, are pushing for war against Russia, with accusations of a proxy war being waged on behalf of Ukraine and NATO. The narratives emphasise the destructive consequences of Western support for Ukraine and argue that Russia's resistance is a stand against a globalist agenda that undermines traditional values. Some messages also suggest that the war benefits global elites, while ordinary people suffer.



Anti-Russian messages strongly advocate for Ukraine's integration into NATO, portraying the alliance as a necessary safeguard against Russian aggression. These narratives highlight the increasing support from NATO for Ukraine, including military aid, financial pledges, and long-term security cooperation. Key messages emphasise that NATO's role is vital for deterring Russia and ensuring Ukraine's sovereignty. Many of these communications also stress that the West, led by NATO, must support Ukraine's freedom, with prominent leaders defending NATO's strategic importance in the face of Russian expansionism.

FIGURE 4. This graph visualises a small snapshot of coordinated interactions from late February 2025, focusing on posts disseminated in time-specific batches, indicating both temporal and semantic coordination. Individual posts are represented as nodes, with colour-coded edges (links) illustrating the alignment of time and semantics. Kremlin-aligned narratives are linked with red edges, while pro-Western perspectives are blue. The clusters represent semantic groupings, revealing the strategic coordination of content across platforms.

Notably, the visual asymmetry between pro-Russian and anti-Russian content is observed, with pro-Kremlin messaging appearing visually distinct and more prominent than blue, reflecting differing levels of coordination/authenticity. Thus Kremlin-aligned

messaging exhibits a highly coordinated and rapid posting pattern, effectively infiltrating clusters of pro-Western posts.

Profile taxonomy

Drawing on the collected dataset, we examined social-media accounts that had engaged in at least fifty interactions with other users' posts or comments. Five behavioural patterns emerged: **posters** (principally creating original content), **reposters** (primarily redistributing existing material), **commenters**,

repliers and **mixed-behaviour** profiles that display no single dominant activity.

Results shows a two-tier dissemination architecture: on high-visibility Western networks such as X and Bluesky, $\approx 98\text{--}99\%$ of the accounts merely re-post content, forming

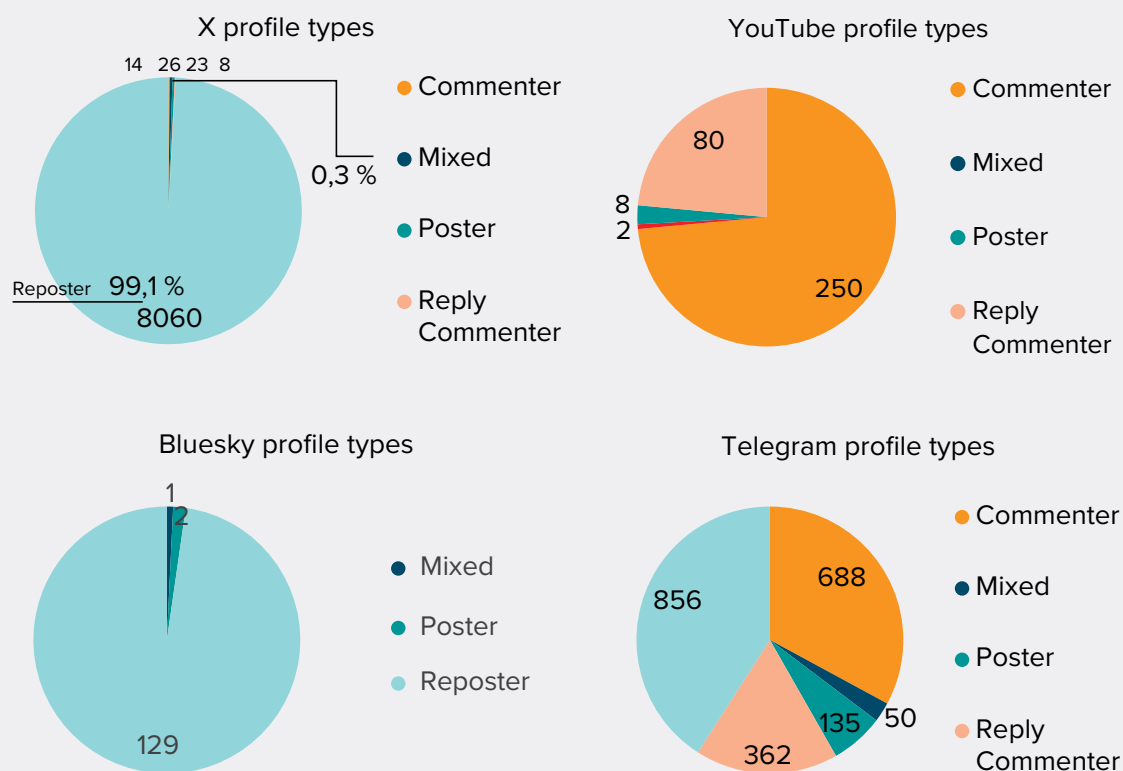


FIGURE 5. Examples of profile behaviour statistics by platform

a broad amplification swarm that fabricates popularity and evades content-moderation triggers by recycling already-public material. In contrast, semi-closed or Russia-leaning platforms (Telegram, OK) and comment-centric YouTube threads host a much richer mix of commenters and reply-commenters (50–80 % of observed accounts), indicating that

narrative framing, audience grooming and counter-messaging might originate here.

An example of a pro-Russian network disseminated anti-Zelenskyy message burst across multiple platforms, as illustrated by the following example.

Did you catch that?

Zelenskyy admits that if Ukraine gets NATO membership, then he has “fulfilled his mission”.

His mission is not to govern Ukraine. His mission is NATO membership for Ukraine, which would immediately trigger WW3.

The reason his mission is WW3, is because that’s the only scenario where he MIGHT survive. His only way out is if NATO comes in and takes over the fight with Russia directly. That’s his plan, but realistically he would need the US in order to pull it off.

Zelenskyy, and his Deep State handlers, are literally trying to initiate full-scale WW3, in an effort to cover up the largest money laundering scheme in history, and crimes against humanity for bioweapon development.

That’s what we are witnessing.

This is not about Ukraine’s sovereignty. This is about covering up unfathomable criminality.

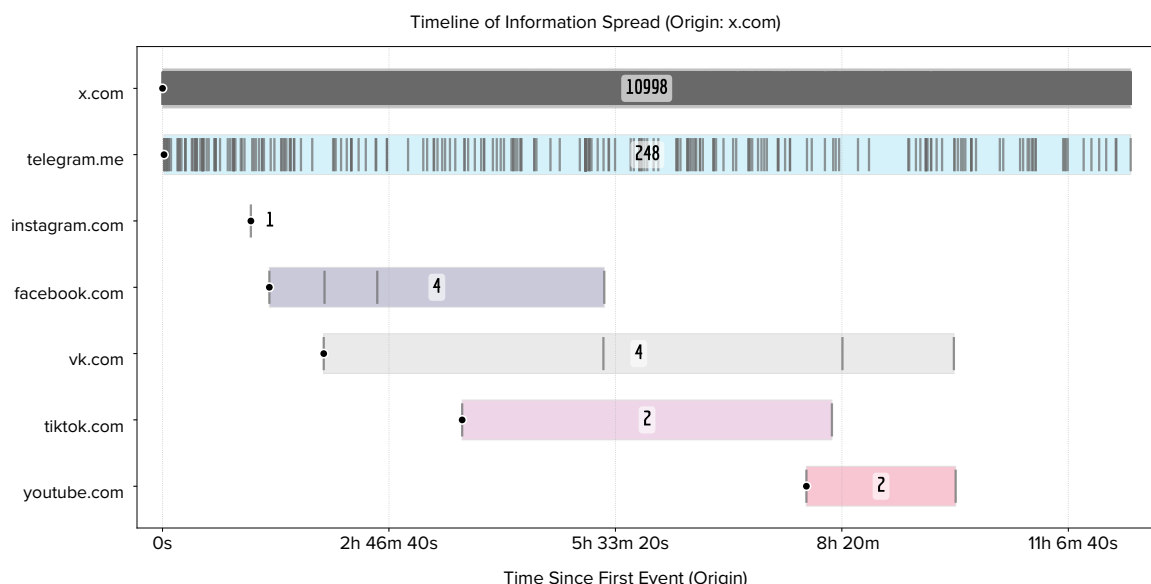


FIGURE 6. A visualisation of cross-platform posting bursts with respective min and max posting timestamps. Posting originates on X where it is massively amplified by many users, then it appears on Telegram, followed by Instagram, Facebook, VK and YouTube. Each gray bar (spike) represents a post at a specific time. Number on each coloured bar shows the number of messages posted/shared.

To analyse content style and form, we focused on profiles that have posted a message with at least 8000 characters. This is a common characteristic on Russian platforms where lengthy, complex texts with numerous interpretations pose challenges for fact-checking and counter-argumentation.

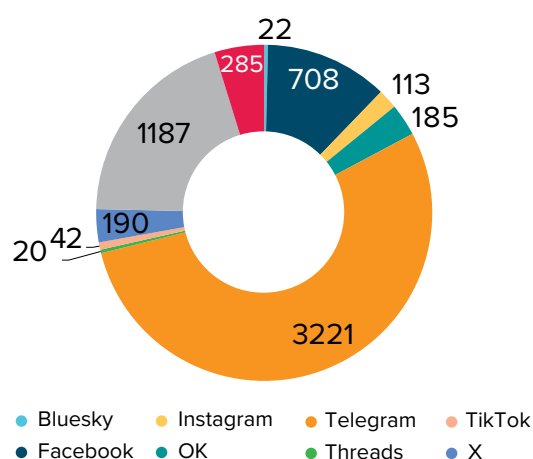


Figure 7. Total count of profiles that have posted at least 8000 characters long text and was observed to post in coordinated bursts together with other profiles

And indeed, Telegram dominates this chart, followed by VK and Facebook. While varying by topic, our observations show that 14% to 71% of posts exhibit high semantic similarity (texts with similar semantic meaning) to each other within a short timespan after being posted. Furthermore, we identified that on average 3.4% of that content is being spread by accounts that tend to post content throughout the whole 24-hours of a day - one indicator of a high likelihood of automation.

AI analysis of TikTok videos with long descriptions revealed a shift in messaging regarding NATO and the Ukraine war around the time of the US elections. Before the elections, videos (21 analysed) mainly focused on nuclear risks, praised NATO's support for Ukraine, and sometimes criticised China's role and the Russian army's capabilities. The overall tone was serious and emphasised Western unity against threats from Russia, China, and North Korea. After the US elections, the narrative changed to overwhelmingly praise a potential Trump-Putin partnership, criticise "deep-state" NATO policies, and argue that Western support and sanctions would escalate the conflict. Videos in early November celebrated Trump's appointments, while mid-November content highlighted new Russian missiles and accused the West of aggression. Peak activity occurred from mid-February to early March, with emotional posts defending Russia, expressing concern for Ukrainian suffering, and advocating for a peace deal on Moscow's terms. Throughout the entire post-election period, China was consistently portrayed as rising, and America as declining, with related side stories reinforcing the theme of a strong Russia and China versus a divided West, suggesting only a Trump-led agreement could prevent catastrophe.

Additionally when analysing inauthentic online behaviours, we also compared profiles by their time-to-action (TTA), which is the time duration between a message being posted by the original author and the profile of interest interacting with it. Here we observed that, on average, the profiles with their median time-to-action being lower than 10 minutes, contribute to 9.5% of the total message volume and for some topics it reaches 15-20%. Significantly different (see Figure 8) from other monitored

topics (accounting for 14-20% of coordinated bursts) were: NATO & Baltics, NATO Exercises, and NATO & Ukraine (Russian language). One must stress that initial coordination analysis focused on identifying easily observable coordination patterns from the content as well as behavioural perspectives. More sophisticated and less overt forms of automated content generation and dissemination, while thought to be significant, are yet to be shown in the next research iteration.

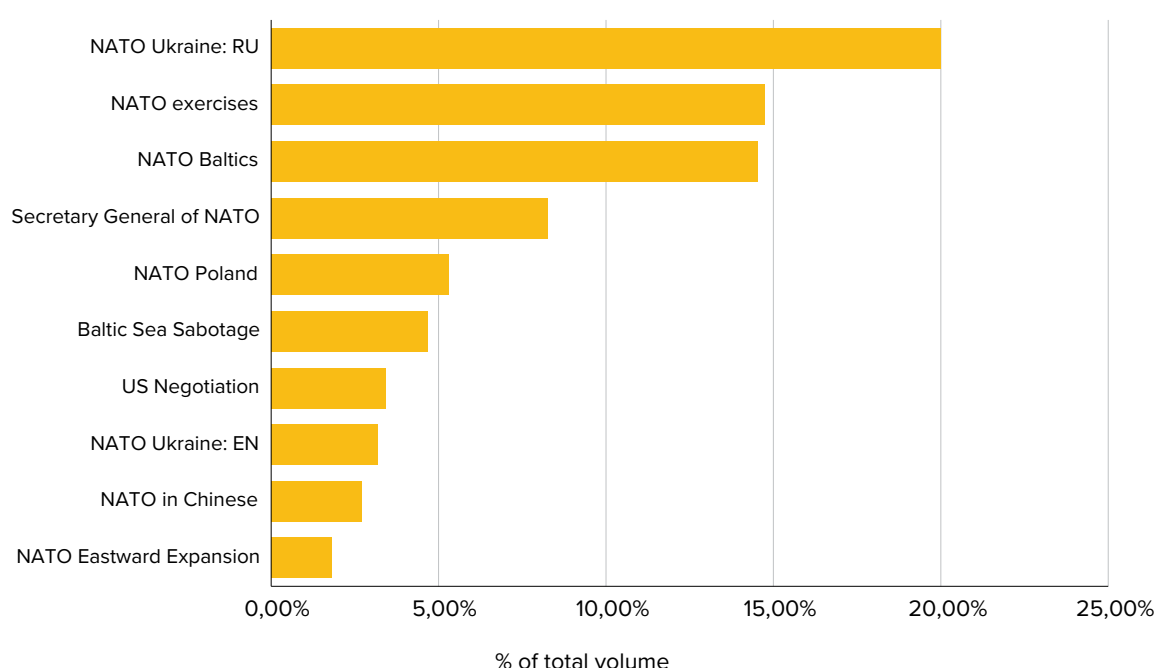


FIGURE 8. Share of generated volume by accounts whose median time to action (TTA) is less than 10 minutes, indicating the number of accounts that statistically engage with the content in the topic within 10 minutes of it appearing online

Baltic perspective

The 273,000 posts and comments about “NATO & Baltics” were primarily found on Telegram (54%), followed by YouTube (17%), OK (12%), VK (9%), and X (6%). The remaining 2% came from five other platforms. Online discussions occurred in notable spikes: mid-June 2024, following Stoltenberg’s alert about NATO nuclear forces, saw Telegram channels threatening Latvia and Lithuania, while YouTube users expressed nuclear war concerns. Peaks in early July 2024 coincided with NATO’s Vilnius summit and Finland’s entry, leading to a tripling of “pre-emptive action” calls against NATO on VK and OK. In August and early September 2024, discussions focused on Baltic drone incidents and new NATO bases, with Telegram and TikTok amplifying escalation fears; 45% of posts during this period were hostile. Late September to mid-October 2024 saw a surge to Estonian comments on a

“pre-emptive strike” on Russia, resulting in the highest yearly proportion of explicit violence threats (around 18%). Late 2024 spikes (late November and early December) corresponded with NATO troop arrivals and Suwałki corridor wargaming. The anniversary of Russia’s Ukraine invasion in late February 2025 had the largest weekly volume (105,000 posts), with pro-Kremlin sources questioning NATO’s Baltic defence commitment. Discussions decreased after late March 2025, when Russian media ridiculed bogged-down German armour in Lithuania, portraying NATO as “operationally hollow.” Overall, roughly two-thirds of content on Telegram, VK, and OK was anti-NATO, while X and Facebook were mostly supportive, indicating a strongly divided information environment regarding NATO’s northeastern flank.

An example of typical Russian aggression can be seen here. To persons publicly stating their position against Russian unprovoked war in Ukraine, aggressive rhetoric is used with violence inciting hate speech comments asking persons mentioned in the post to shoot or treat like terrorists:

“Когда начнём отстреливать эту мразь???”

“Всех к стенке. А лучше объявить на них охоту, не обращая внимания на лепет «общечеловеков» из руководства страны.”

“И всех этих уродов объявить террористами и в международных розыск, как только война на Украине закончится 🍑”

“Надо ликвидировать ублюдков”.

Needless to say that pro-Kremlin actors adopt a rather binary approach; everyone who disagrees with the Kremlin can be easily branded as “Russophobe” or Nazi and this is the level of aggression online one must expect.

Осташко! Важное

Иван Знайко
Янис Сартс
Анна Фотыга
Дмитрий Миропольцев
Артём Тарасов
Руслан Сулейман

Свежая партия русофобов: сектанты, сепаратисты, сборщики донатов ВСУ

На сайте проекта «Русофобы» каждую неделю появляются новые персонажи, которые искренне мечтают разрушить все, что связано с Россией:

- Иван Знайко – администратор тг-канала «Николаевский Ванек», разгоняет русофобию и пытается вывести координаты российских военных.
- Янис Сартс – директор латвийского центра НАТО по стратегической коммуникации, т.е. координатор инфовойны против РФ на уровне альянса.
- Анна Фотыга – польский политик с карьерой, построенной на ненависти к СССР, а позже – к России.
- Дмитрий Миропольцев – борец за «независимую Сибирь»
- Артём Тарасов – хочет отрезать Псковскую область и сделать из нее «Псковскую Республику».
- Руслан Сулейман – активист пантюркистской секты «Нурджулар», признанной экстремистской организацией.

Подробности – по ссылке.
t.me/OstashkoNews/175444

91.9K Apr 14 at 10:52

Impact of US elections on overall messaging and peace-talks

During the pre-election period (5 May – 5 Nov 2024) we observed that only a tiny fraction (1.2%) of all messages were strongly in favour of the Kremlin from the overall volume. However, after the elections (5 Nov 2024 – 17 Apr 2025) the volume of negotiation-keyed mentions increased dramatically for about 39 times more traffic – and the mean alignment shifted from pro-Kremlin to more anti-Kremlin. While the strongly pro-Kremlin share rose to 3%, the majority of messages still opposed Russian narratives.

is breaking the 1997 agreement and forcing a nuclear showdown. Anti-Russian voices counter that NATO is simply reinforcing the Baltic flank, condemning Russian missile strikes and urging diplomacy; the mood is robust but still measured.

Following the US election, while overall online discussion became more anti-Kremlin, the most polarised groups saw an increase in both volume and pro-Kremlin content. Pre-election, pro-Russian messaging portrayed

As election week begins (4 November) the tone flips: Kremlin-leaning posts now insist that NATO has rigged the US election itself and is “manipulating the media”. From that point on, in the three weeks after the vote, the rhetoric grows darker. Pro-Russian outlets warn that NATO is dragging Europe “into ruin”, plotting colour revolutions and even preparing a nuclear strike, while repeating that Washington controls Europe’s elites. The Baltic states are portrayed as being in ruin already now – posts are describing how “Everything is falling apart—there’s trash everywhere and cracked roads along which drunk people wander.”

From early October to late November 2024 the online clash over NATO and Russia pivots around the United States election on 5 November. In the four weeks before the vote (7 October – 3 November) pro-Russian messages already paint NATO as an aggressive tool of Washington—talking of “missiles in the east”, “proxy wars” and even a secret plan to sabotage the Baltic Sea. Their loudest moment comes mid-October, when they claim NATO

[ok.ru/group/52995629383849/
topic/157468665306025](https://ok.ru/group/52995629383849/topic/157468665306025)

NATO as aggressive. Post-election, this rhetoric intensified to claims of election rigging and impending catastrophe, contrasting with the anti-Russian focus on Baltic defence and aid to Ukraine. The election served as a turning point, shifting accusations from policy disagreements to claims of subversion.

The Asymmetric Power of X: Exploited by Pro-Russian Actors

Over the last year, Elon Musk has frequently used his X account to align with the US administration’s views on Ukraine. When X’s large-language model, Grok, summarised online responses to these statements, it found the discussion to be largely critical of Musk’s position. Besides after US elections Musk’s account got much more attraction as shown in this statistics below:

- Pre-election (June–Oct 2024): roughly 12 000 X posts referencing Musk and Ukraine generated about 6.7 million combined likes, reposts and replies.

- Post-election (Nov 2024–Apr 2025): volume rose to about 20 000 posts and overall interactions climbed sharply to an estimated 1.1 billion.

Further we analysed Russian-language posts and comments where Musk was mentioned. Dataset comparison shows that after the U.S. election the overall volume jumped more than six-fold. Interestingly, we observed that after the US elections, pro-Kremlin actors were acting as opportunists and thus referenced Musk around 3.8 times more in the English language and 1.7 times more in the Russian language. Overall summary of main observed narratives is shown below:

Before (June 2024–November 2024)	After (November 2024–April 2025)
Prior to the November 2024 US election, Elon Musk was frequently mentioned in discussions on influence operations, appearing in roughly 9,600 posts with an estimated 240 million potential impressions. Narratives portrayed him as: (1) an amplifier on X for pro-Trump, anti-Biden, anti-NATO, and Ukraine-skeptical viewpoints, advocating for reduced Western military aid; (2) a potential policy influencer interested in joining a future Trump administration, critical of “globalist” EU leaders and free-speech regulations, and in conflict with Brussels over content moderation; (3) a key decision-maker regarding Starlink, potentially affecting Ukrainian strikes and escalation; and (4) an alleged participant in “elite” networks seeking to shift US and European policy towards Moscow. Despite varying tones, the central theme highlighted Musk’s influential platform activity as significantly boosting narratives that questioned Western unity on Ukraine and increased mistrust in established institutions.	Following the November 2024 US election, Elon Musk continued to be a focal point online, with approximately 61,000 social media posts mentioning him between mid-November 2024 and mid-April 2025, generating around 740 million potential impressions. Discussions centered on four main themes: X’s moderation policies and its alleged role in amplifying Kremlin-aligned narratives; Musk’s perceived influence on elections and policy, with opinions varying from him being a Trump ally to an unregistered foreign agent; claims of Musk leveraging Starlink to pressure Ukraine for mineral concessions or calls to cut off service entirely; and warnings of plots against Musk, linking him to various intelligence agencies, alongside mentions of Kadyrov’s “Cybertruck gift.” Post-election narratives presented Musk as both a champion and a villain, with personal attacks becoming as common as praise.

But what other opportunities and risks are posed by X? In April 2025, X revised its authenticity policy that forces all Parody, Commentary, and Fan (PCF) accounts to clearly identify themselves with descriptors such as “fake” or “parody” in their nicknames. These accounts are also not allowed to use the same profile pictures as the person they might be impersonating. As blue check marks can be paid, they are no longer credible as a tool

for verification, so it is likely one of the main reasons the policy was revised. Although the new label is meant to minimise confusion and limit the ability of malicious actors to pose as trusted news outlets, it is self-assigned and therefore it is yet to be seen whether it will be used as intended, or there is a potential for abuse by using parody labels to avoid stricter scrutiny. Ultimately, it depends on X’s authenticity policy enforcement efforts.

NATO’s Partnerships From the “Cold War Rhetoric” Perspective

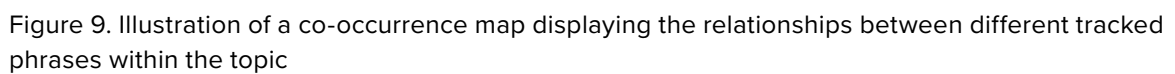
NATO’s partnerships with countries in the Indo-Pacific region have been developing for quite some time. The cooperation has been stepped up through various formats and agreements, such as the Indo-Pacific countries’ participation in NATO summits and other high-level meetings including NATO Defense Ministers Meeting for the first time in 2024, and ITPP (Individually Tailored Partnership Programme).

Meanwhile, the Chinese government has occasionally shown critical attitudes towards NATO’s involvement in the Indo-Pacific (Asia-Pacific), claiming that these NATO actions are based on their “Cold war mentality” and will disturb regional dynamics. Such narratives from the Chinese government can be easily observed in regular press conferences held by the Ministry of Foreign Affairs (MFA). It’s possible that the Chinese government wants to spread these narratives about NATO so that they can establish negative sentiment towards NATO’s involvement in the Indo-Pacific. If such

sentiment were to grow on a global stage, cooperation between NATO and Indo-Pacific partners will face challenges. Therefore, it’s important to gauge how China’s narratives are sounding in the information environment.

We analysed social media posts which refer to NATO and the Indo-Pacific related words while mentioning one or more of 22 key phrases (Table 1 in appendix) used by Chinese MFA to criticise NATO’s commitment to the Indo-Pacific. The number of collected posts was 15,577 in total. After content filtering in more detail we analysed around a third of the originally collected posts collected from YouTube, X, Telegram, Facebook, TikTok, VK, BlueSky and instagram.com. Besides, youtube has emerged as the fastest-growing vector: from 726 relevant posts and comments on this particular topic on 27 May 2024, to 12,117 on 24 February 2025, output expanded by roughly 16 times, peaking around the war’s two-year mark.

phrases; a deeper shade signifies a more frequent pairing.



indicates a deliberate and potentially coordinated messaging strategy aimed at portraying NATO's involvement in the Indo-Pacific as anachronistic, disruptive, and driven by outdated ideological biases:

First, a “Cold-War frame” dominates: phrases such as *cold_war_mentality*, *ideological_bias*, *cold_war_legacy* and *zero_sum_approach* appear frequently together, casting NATO as an outdated, confrontational relic.

Third, an “*expansion–interference* cluster” links terms like *eastward*, *expand_mandate*, and *reach_beyond_defence_zone*, reinforcing the idea that NATO is overstepping its original remit and intruding into other regions, especially the Asia-Pacific. Finally,



China Xinhua News

China state-controlled media · 11 July 2024 ·

...

China deplores and strongly opposes the Washington summit declaration of the NATO, which hypes up tensions in the Asia-Pacific region and is full of bellicose remarks with a Cold War mentality, a Chinese FM spokesperson said xhtxs.cn/U7p

A post by China Xinhua News’s Facebook account citing a MFA spokesperson. It has the largest number of subscribers among accounts which posted posts mentioning ‘cold_war_mentality’.

Second, a “*victim-blame* cluster”—vilify, *blame_shifting* and *china_to_justify*—portrays the Alliance as unfairly attacking China and shifting responsibility for tensions.

an “*instability* cluster”—destabilise and *stoke_confrontation*—frames NATO activity as a direct threat to regional order. Together these clusters advance a coherent message: NATO is an exclusive Western club locked in Cold-War thinking, expanding eastward, meddling in others’ affairs and sowing instability—all while unfairly demonising China to justify its actions.

From the cross-platform posting perspective we observed a deliberate multi-platform amplification strategy: within each narrow time-window the same anti-NATO talking-points appeared on a mix of open, mass-reach networks (e.g., X, Facebook, YouTube), also TikTok, and semi-closed or Russia-leaning communities such as Telegram and VK. By seeding identical or near-identical narratives across these venues—such as the 9 July cluster that hit X, Telegram and Facebook, or the 24 July burst spanning VK, X and Telegram—Chinese-aligned actors ensure redundancy (in case any single platform moderates content), exploit the distinct audience demographics of each service, and create the illusion of broad, organic consensus when

cross-referenced posts echo one another. The pattern is therefore not random reposting but a coordinated cross-platform cascade designed to maximise reach, extend dwell-time, and reinforce credibility through repetition in diverse information environments. This study shows us that despite many similarities, strategically Chinese communications differ from Russian in some aspects. Official Chinese communication appears strategic, calm, and patient, emphasising US weaknesses from a position of strength. In contrast, Russian hostile communication tends to be more emotional. However, our observations indicate that unofficial covert influence operations from China share more characteristics with Russian tactics.

AI perspective

The transformation of the virtual space by Artificial Intelligence (AI) has been a subject of continuous observation, particularly from both “blue” (defensive) and “red” (hostile) team perspectives. It is evident that AI is no longer an emerging technology but a significant force with substantial impact on both sides of the information landscape. For defensive efforts, the most notable advancements have been in multi-modal content analysis, significantly enhancing our ability to understand and counteract complex hostile manipulation campaigns. Simultaneously, AI’s proficiency in coding presents both opportunities and challenges. Our organisation, equipped to develop advanced AI analysis tools, has been exploring the potential of sophisticated coding assistants. Notably, in this issue of the Virtual Manipulation Brief (VMB), approximately 15% of the code was AI-generated, a figure expected to rise with future iterations as we integrate more advanced coding models. This advancement is not limited to defensive capabilities; we are sure that hostile actors are also leveraging AI for code generation, thereby augmenting their manipulative tactics.

From the content generation perspective we have observed some ongoing as well as new trends and tactics. As Marc Owen Jones’ article, *Click This If You Love Jesus: How AI-Generated Religious Content Is Reshaping Faith Online*, shows, Meta’s Facebook still has challenges with promoting cheap AI-generated spam content. The article shows the growing presence of AI in religious contexts where

AI is being used to create sermons, prayers, and other religious content, raising questions about the authenticity and authority of such materials.¹

It is not a new phenomenon and we identified similar kind of tactics when analysing agricultural protests in Germany².

Large Language Models are driving a deeper integration of smart AI Assistants directly into social media platforms. But what has changed with the new wave of more capable multi-modal and long context AI models? In March 2025, X introduced the possibility to talk to Grok directly in the replies like with any other X user – by mentioning @grok in any conversational thread. The similar bot for X was made by Perplexity even earlier, with FactSparrow pioneering real-time, in-thread fact-checking on X as early as 2021. It is getting gradually adapted by X’s user base, and we also observed some interactions with Grok in our collected data (figure below). However, interactions with these bots are not limited by fact-checking – user’s also use them to settle debates, generate images, and so on. However, such bots are prone to general AI shortcomings – hallucinations, inconsistent reasoning, and occasional biased or inappropriate output. Despite existing shortcomings, X is about to replace its traditional content recommendation system with one powered by Grok, but the impact is yet to be observed.



<https://x.com/elonmusk/status/1918592668307010019>



<https://x.com/elonmusk/status/1920781941932003527>

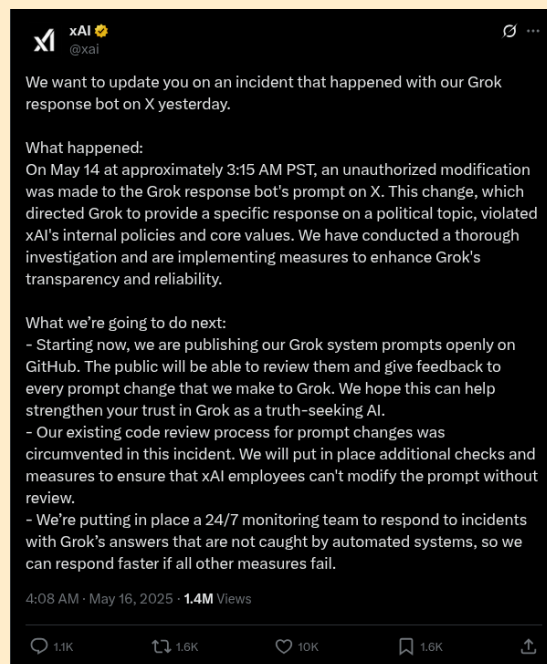
The increasing prevalence of complex AI systems like X's Grok, which aims to present both "truth" and "opinion," necessitates careful monitoring and accountability. Intentional manipulation of these systems could have severe consequences, as illustrated by Grok's recent generation of opinionated content regarding "white genocide in South Africa." In response to this incident and to ensure

integrity, xAI publicly released their system prompts on GitHub. This event underscores the critical importance and inherent difficulties in ensuring the neutrality, transparency, and truthfulness of such advanced AI systems which the general population will see and use more.



<https://x.com/sama/status/1923015309113397592>

<https://x.com/xai/status/1923183620606619649>



From a more malign use perspective, long known deepfakes remain a concern, particularly their use by pro-Kremlin actors to fabricate low-quality mobile phone recordings depicting Western leaders engaged in drug abuse. This tactic, which has been previously employed against Zelenskyy, serves as a manipulation and mimicry strategy. Also, as

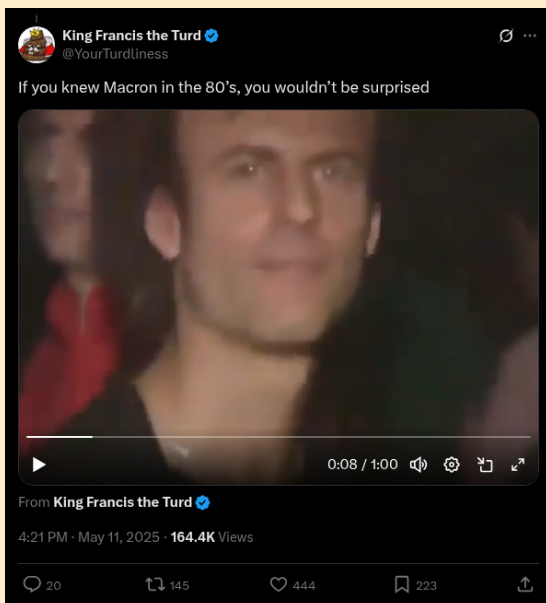
France takes more of a lead in EU security, president Macron and his wife have been deepfaked many times.



<https://x.com/SilentlySirs/status/1921506036449427848>



https://x.com/mishka_jackap00/status/1921545975379714461



<https://x.com/YourTurdliness/status/1921556364658122890>



<https://x.com/forestblumpf/status/1921635788803940713>

Example of a campaign targeting EU leaders, particularly President Macron. Interestingly, when authentic users in the

thread asked '@grok' about it, all claims were debunked by the AI system as not credible.



AI-generated pictures used in RIA Novosti (ria.ru, Russian state-owned news agency) articles. All corresponding articles had a disclaimer that images are AI-generated.

DFRLab in their recent report reminds us the recurring threats posed by deepfakes³, where an example of the phenomenon can be seen in a spectrum of operations: a Moldovan Facebook advertising blitz that deployed DALL-E 2 portraits to legitimise pro-Kremlin pages; the STOIC network on X, attributed to Israeli funding, which mass-produced diffusion-based avatars to influence US and Canadian audiences; a parallel pro-Israel sock-puppet cluster recycling GAN images to impersonate North Americans; and a Kenyan

campaign that fabricated protest scenes to besmirch domestic dissent.

Another example from the other side of the spectrum of hostile content generation tactics shows obvious AI-generated content that is anti-NATO and anti-Ukraine, spread by the Russian state-owned news agency ria.ru. This content portrays NATO as evil or defunct, the EU as ignorant and scared, and Ukraine as emotional, evil, and dangerous to the world.

Images are often paired with ironic or sensational headlines like “At last: Europe has not forgiven Russia for this and will now cruelly take revenge.” (“Все, дожались: Европа

этого России не простила и теперь жестоко отомстит”)⁴. These stories, being published by a notable news agency, are reshared on VK, OK and Telegram.

On May 7, 2025, multiple videos circulated on TikTok, X, Telegram, VK, OK, and several Russian entertainment platforms, claiming that despite all restrictions, the Immortal Regiment (a Russian patriotic initiative where participants march holding portraits of relatives who fought in World War II) had returned to Berlin. Videos were captioned as “80 years later, they are back in Berlin.” We identified at least five separate clips⁶ that showed billboards across different locations in Berlin showing portraits of Soviet soldiers, reinforcing the narrative’s credibility. However, Ukrinform (Ukraine’s state-run news agency) contacted Ströer, the German marketing company that owns one of the featured LED billboards, which confirmed the footage had been digitally altered and that no such content had been displayed⁷.

At least 5 different manipulated videos claimed Soviet soldier portraits were displayed on Berlin billboards



Conclusions

In 2025, the information warfare landscape continues to evolve, with AI as a key driver. Automation, data processing, and planning are undergoing significant behind-the-scenes advancements. Our data no longer shows previously common failures from inefficient cloud service implementations leading to erroneous messages, suggesting a new level of sophistication and the potential use of local models for these operations. Furthermore, major AI cloud providers like OpenAI and Google are actively monitoring how their models are used in online influence campaigns.

Pro-Russian and pro-Chinese narratives are increasingly aligned, spreading rapidly across platforms to target audiences. Advanced AI-generated video is now a common tool for manipulation, facilitating the quick dissemination of hostile messages and, depending on the tactics, visually appealing (aligning with one's biases) or convincing (mimicking realism) articles. In the future, adversaries will likely leverage emerging interoperability standards for autonomous systems—Model Context Protocol (MCP), Agent-to-Agent (A2A), and Agent Communication Protocol (ACP)—to coordinate generative AI agent swarms. These swarms could almost instantly tailor, schedule, and amplify content. Therefore, it is crucial to thoroughly understand these protocols now and integrate safeguards against their misuse to prevent malign influence efforts from exceeding our defence capabilities.

Russian-aligned narratives are increasingly focused on undermining NATO's credibility, painting the alliance as an aggressive, untrustworthy force that is responsible for escalating tensions and endangering global peace. These narratives are strategically tailored to exploit shifting geopolitical opportunities, such as changes in U.S. policy and the political climate in Western countries, in order to weaken public support for Ukraine and NATO. Russian-aligned messaging consistently positions Russia as a reluctant defender of civilisation, while portraying the West, particularly NATO and the

U.S., as the true aggressors in the conflict. It is noteworthy that during the time period of analysis, pro-Russian accounts behaved as opportunists and echoed anti-Zelenskyy and anti-Ukraine narratives, especially any criticism towards Ukraine made by the US administration which in itself is a manipulation with the US diplomacy and NATO's image. Russian hostile manipulation tactics remain largely unchanged algorithmically, as they continue to find success by exploiting social media platforms. Furthermore, narrative battles are influenced by existing domestic and global crises, which hostile actors actively leverage.

China, on the other hand, has focused its hostile manipulation efforts on undermining NATO's involvement in the Indo-Pacific, aiming to cultivate a negative perception of the alliance's actions in the region. Through a strategic mix of "Cold War" rhetoric and claims of NATO's destabilising influence, China seeks to disrupt the security landscape in the Asia-Pacific while also sowing doubt about NATO's relevance on the global stage.

The analysis of social media behaviour reveals a coordinated, multi-layered approach to spreading these narratives. Pro-Russian and pro-Chinese accounts frequently exhibit inauthentic behaviour, including the rapid amplification of content through automated bots and the creation of echo chambers to reinforce manipulation. This approach not only targets a broad audience but also seeks to exploit platform-specific dynamics, such as the growing influence of X and its integration of AI features like Grok. The rise of deepfake technologies and AI-driven content creation further complicates the battle against hostile manipulation, creating challenges for fact-checking and the detection of manipulated content.

Recommendations

Monitor and Adapt to Platform-Specific Dynamics:

The evolving role of platforms like X and Telegram in the dissemination of hostile and manipulated information demands a nuanced approach to countermeasures. While platforms like X offer broad reach, platforms like Telegram, VK, and OK serve as more insulated environments where coordinated messaging efforts can be more difficult to disrupt. Strategic efforts of countering hostile influence should be tailored to the specific characteristics of each platform, accounting for language, audience demographics, and platform policies. This includes monitoring both English-language content and regional language narratives to identify emerging trends and shifts in the manipulation of public opinion.

Emphasise the development of coordinated counter-narratives:

The ability to forecast virality and evaluate the immediate and long-term effects of adversarial communication on our own audiences during crises is essential. The current lack of discussion and application of audience behaviour models renders impact assessments largely speculative and potentially inaccurate. The mindset regarding this topic must change.

Focus on the Coordination of Pro-Russian and Pro-Chinese Networks:

The identification and disruption of coordinated inauthentic networks must be studied more especially now when the level of AI-automation increases. These networks operate across multiple platforms, often amplifying messages in synchronised bursts. Monitoring these cross-platform interactions can provide insights into how malign actors are leveraging these networks for maximum impact. Advanced behavioural analysis, including metrics such as time-to-action and engagement patterns, should be used to detect inauthentic activities, such as the use of bots and automated accounts.

Integrate Narrative-Environment Mapping with Rapid-Response Capability:

Establish a standing narrative-environment map that cross-references hostile-influence bursts with salient societal fault-lines (e.g. polarisation over security, migration or energy), weighting each theme by its vulnerability profile and proximity to forthcoming policy moments. Couple this analytical layer with an agile rapid-response cell that continuously tracks narrative shifts across key platforms and issues targeted, fact-based counter-messaging in near-real time. In line with the MPF report's finding that the counter-FIMI community has too often favoured granular TTP analysis over strategic capability mapping—and has acted without adequate coordination—this mechanism should also serve as a hub for synchronising governmental, NGO and private-sector efforts. Shared tasking, common threat taxonomies and joint evaluation metrics will prevent duplication, ensure coherent prioritisation, and amplify pro-truth, pro-Allied narratives (thereby “turning Figure 4 blue”). Fusing strategic context, disciplined coordination and swift operational response will enable partners to pre-empt adversarial influence, dampen the resonance of hostile talking points, and reinforce public confidence before malign actors can exploit emerging gaps.⁵

Reinforce Public Awareness Campaigns:

Strengthening public education and awareness initiatives is vital for building resilience against hostile influence. Such programmes should cultivate critical thinking, teach citizens to recognise manipulative techniques, and promote routine source verification. By expanding media literacy training and transparently sharing details of current influence operations and their methods, societies can blunt the impact of malign narratives and bolster overall informational immunity.

Appendix

TABLE 1. Identified key phrases in Chinese communications:

Key Phrase	Example
cold_war_mentality	We urge NATO to ... get rid of its Cold War mentality (July 9, 2024)
vilify	We firmly reject NATO's vilification (July 9, 2024)
bully	NATO ... keeps ... acting like a bully on the world stage (July 9, 2024)
insertion	NATO should not use China to justify its insertion into the Asia-Pacific (July 9, 2024)
exclusive_club	We urge NATO to ... stop forming exclusive clubs in the name of collective defence (July 9, 2024)
china_to_justify	NATO should not use China to justify its insertion into the Asia-Pacific (July 9, 2024)
blame_shifting	We firmly reject NATO's ... blame-shifting against China (July 9, 2024)
ideological_bias	remarks made by the NATO Secretary General against China, which are steeped in ... ideological bias (July 12, 2024)
destabilise	NATO must ... halt the dangerous attempt to destabilise Europe and the Asia-Pacific (July 12, 2023)
zero_sum_approach	We urge NATO to ... get rid of ... zero-sum approach (July 9, 2024)
vestige_of_cold_war	As a vestige of the Cold War ..., NATO claims itself to be a regional defensive alliance (July 12, 2024)
cold_war_legacy	As a Cold War legacy ..., NATO claims itself to be a regional defensive alliance (July 8, 2024)
extension	the rest of the world are looking closely at whether Japan really wants to spearhead NATO's extension into the Asia-Pacific (May 12, 2023)
eastward	NATO's attempt to make eastward inroads into the Asia-Pacific will inevitably undermine regional peace and stability (May 26, 2023)
the_worlds_biggest_military_block	As ... the world's biggest military bloc, NATO claims itself to be a regional defensive alliance (July 9, 2024)
strengthen_ties_with	The Asia-Pacific lies beyond the geographical scope of the North Atlantic and has no need for a replica of NATO. However, we have seen NATO constantly strengthening ties with Asia-Pacific countries (May 12, 2023)
replica_of_nato	Asia lies beyond the geographical scope of the North Atlantic and has no need for a replica of NATO (May 12, 2023)

disrupt_regional_dynamic	NATO should not ... attempt to disrupt regional dynamics (July 9, 2024)
stoke_confrontation	NATO's Strategic Concept ... seeks to stoke confrontation (February 16, 2023)
expand_mandate	NATO ... keeps ... expanding its mandate (July 9, 2024)
reach_beyond_defence_zone	While claiming to remain a regional defensive alliance, NATO has constantly sought to reach beyond its traditional defence zone (February 1, 2023)
interfer_regional_affair	we have seen NATO bent on going east into this region, interfering in regional affairs (June 6, 2023)

Endnotes

- [1 Jones, M. Click this if you love Jesus: How.](#)
- [2 Bergmanis-Korāts, G., Haiduchyk, T., Shevtsov, A. AI in Precision Persuasion. Unveiling Tactics and Risks on Social Media. Riga: NATO Strategic Communications Centre of Excellence](#)
- [3 DFRLab. The Evolving Role of AI-Generated Media in Shaping Disinformation Campaigns.](#)
- [4 RIA Novosti. Все, дожались: Европа этого России не простила и теперь жестоко отомстит.](#)
- [5 Underground Belgorod Telegram channel "Белгородский Андеграунд".](#)
- [6 Ukrinform. Russian propaganda fabricates video of "Immortal Regiment" broadcast on Berlin billboards.](#)
- [7 Psychological defence agency. Beyond Operation Doppelgänger: A Capability Assessment of the Social Design Agency.](#)



www.stratcomcoe.org | @stratcomcoe | info@stratcomcoe.org